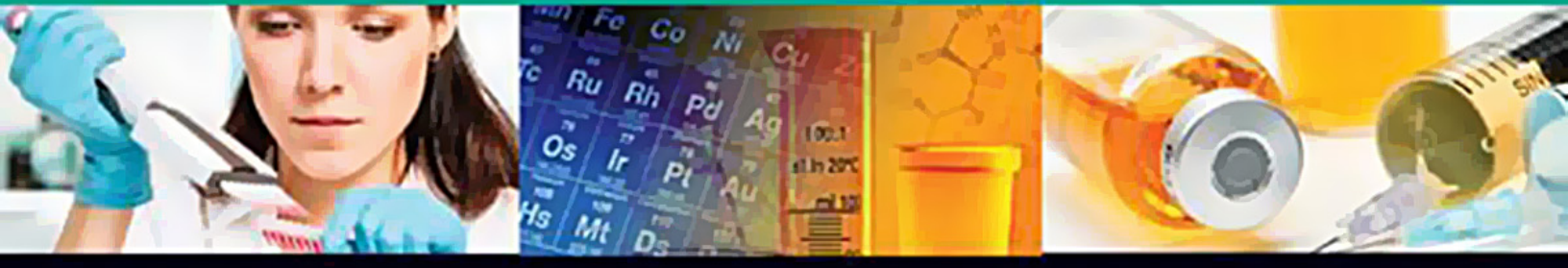
Edited by
Camille Georges Wermuth
David Aldous
Pierre Raboisson
Didier Rognan

THE PRACTICE OF MEDICINAL CHEMISTRY



FOURTH EDITION

AP

THE PRACTICE OF MEDICINAL CHEMISTRY

FOURTH EDITION

The Practice of Medicinal Chemistry

FOURTH EDITION

Edited by

CAMILLE GEORGES WERMUTH, PhD, PharmD

Founder of Prestwick Chemical, Strasbourg, France

DAVID ALDOUS, BSc, DPhil

Boston Head Lead Generation to Candidate Realisation, Boston, MA, USA

PIERRE RABOISSON, PhD, PharmD

Janssen Infectious Diseases and Diagnostics BVBA, Beerse, Belgium

DIDIER ROGNAN, PhD, PharmD

Laboratoire d'Innovation Thérapeutique, CNRS-Université de Strasbourg, Illkirch, France

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
125, London Wall, EC2Y 5AS
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

First edition 1996
Reprinted 2001
Second edition 2003
Third edition 2008

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-417205-0

Contents

7. Fragment-Based Drug Discovery

VENKATA VELVADAPU, BENNETT T. FARMER AND ALLEN B. REITZ

8. Molecular Variations Based on Isosteric Replacements

PAOLA CIAPETTI AND BRUNO GIETHLEN

9. Ring Transformations

CHRISTOPHE MORICE AND CAMILLE G. WERMUTH

10. Macrocycles: Under-Explored and Poorly Exploited Drug Class Despite the Proven Therapeutic Potential

PIERRE RABOISSON

III

PRIMARY EXPLORATION OF STRUCTURE-ACTIVITY RELATIONSHIPS

11. Conformational Restriction and Steric Hindrance in Medicinal Chemistry

PETER WIPF, ERIN M. SKODA AND ANDRÉ MANN

12. Application Strategies for the Primary Structure–Activity Relationship Exploration

CAMILLE G. WERMUTH, SERGE GRISONI, BRUNO VILLOUTREIX AND JEAN-PHILIPPE ROCHER

13. Substituent Groups

PATRICK BAZZINI AND CAMILLE G. WERMUTH

14. The Role of Functional Groups in Drug–Receptor Interactions

LAURENT SCHAEFFER

32. Chemical and Physicochemical Approaches to Solve Formulation Problems

33. Discover a Drug Substance, Formulate, and Develop It to a Product

VII

PHARMACEUTICAL AND CHEMICAL MEANS TO SOLUBILITY AND FORMULATION PROBLEMS

34. Drug Nomenclature

35. Web Alert: Using the Internet for Medicinal Chemistry

36. Protection of Inventions in Medicinal Chemistry

List of Contributors

Anne Badel UMR-S973, MTi, University Paris Diderot, Paris, France

Patrick Bazzini PC SAS, Bd Gonthier d'Andernach, Illkirch, France

Frans M. Belpaire Heymans Institute for Pharmacology, Gent, Belgium

Alexandre Borrel University of Helsinki, Helsinki, Finland; UMR-S973, MTi, University Paris Diderot, Paris, France

Koen Boussery Laboratory of Medical Biochemistry and Clinical Analysis, Faculty of Pharmaceutical Sciences, Gent University, Gent, Belgium

Christopher Brice Consultant, Luthi & Co., Intellectual Property Law Offices, of Tel Aviv, Israel

Anne-Claude Camproux UMR-S973, MTi, University Paris Diderot, Paris, France

David Cavalla Numedicus Ltd, Cambridge, England

Paola Ciapetti Head of Medicinal Chemistry, Novalyst Discovery, Illkirch, Cedex, France

Hongming Chen Chemistry Innovation Centre, Discovery Sciences, AstraZeneca R&D, Mölndal, Sweden

Yong Mi Choi-Sledeski Sanofi, Boston, MA, USA

Bernd Clement Pharmaceutical Institute, Christian Albrechts-University, Gutenbergstraße, Kiel, Germany

Gordon M. Cragg NIH Special Volunteer, Natural Products Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, United States National Cancer Institute, Frederick, Maryland, USA

Patrick M. Dansette Laboratoire de Chimie et Biochimie Pharmacologique et Toxicologique, Université Paris Descartes, Sorbonne Paris Cité, Paris

Pierre deMontigny Sanofi, Waltham, MA, U.S.A

Ji-Cui Dong INN Programme, World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms

Iain G. Dougall IGD Consultancy Ltd, Loughborough, Leics, UK

Ola Engkvist Chemistry Innovation Centre, Discovery Sciences, AstraZeneca R&D, Mölndal, Sweden

Bernard Faller Novartis Pharma AG, Werk Klybeck Klybeckstrasse, Basel, Switzerland

Bennett T. Farmer Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA

Delphine Flatters UMR-S973, MTi, University Paris Diderot, Paris, France

Bruno Galli Novartis Pharma AG, Lichtstrasse, Basel, Switzerland

Colette Geneix UMR-S973, MTi, University Paris Diderot, Paris, France

Bruno Giethlen Novartis Pharma AG, Basel, Switzerland

Serge Grisoni Pierre Fabre Research Institute, Toulouse, Cédex, France

Gerhard Gross Scientific Director at Janssen in Beerse, Belgium

David Harris Sanofi, Waltham, MA, U.S.A

Yuichi Hashimoto Institute of Molecular and Cellular Biosciences, The University of Tokyo, Japan

Chris Ho Sanofi, Waltham, MA, U.S.A

Andrew L. Hopkins College of Life Sciences, University of Dundee, Dundee, Scotland, UK

Hiba Abi Hussein UMR-S973, MTi, University Paris Diderot, Paris, France

Peter Imming Institut fuer Pharmazie, Martin-Luther-Universitaet Halle-Wittenberg, Halle, Germany

Minoru Ishikawa Institute of Molecular and Cellular Biosciences, The University of Tokyo, Japan

Tim Jonckers Janssen Research & Development Janssen Infectious Diseases – Diagnostics BVBA , Turnhoutseweg, Beerse, Belgium

Sabine Kopp World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms

David G.I. Kingston Department of Chemistry and the Virginia Tech Center for Drug Discovery, Virginia Tech, Blacksburg, Virginia, USA

Thierry Kogej Chemistry Innovation Centre, Discovery Sciences, AstraZeneca R&D, Mölndal, Sweden

Sophie Lasseur INN Programme, World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms

Dominique Lesuisse Lead Generation Board, Sanofi, AVENUE PIERRE BROSSOLETTE, CHILLY MAZARIN - France

Harvey Lieberman Sanofi ,Waltham, MA, USA

Richard Luthi Founding partner, Luthi & Co., Intellectual Property Law Offices, of Tel Aviv, Israel

Anne-Christine Macherey Prévention du Risque Chimique, Gif sur Yvette, France

André Mann LAAS-CNRS, Université Louis Pasteur, Strasbourg, France

Raffaella G. Balocco Mattavelli INN Programme, World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms

Christophe Morice Prestwick Chemical Inc., Boulevard Gonthier, Illkirch, France

Richard Morphy Eli Lilly, Erl Wood Manor, Windlesham, Surrey, UK

David J. Newman NIH Special Volunteer, Natural Products Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, United States National Cancer Institute, Frederick, Maryland, USA

Anne Olivier Sanofi-Aventis Recherche et Développement, Chilly-Mazarin, France

Michel Petitjean UMR-S973, MTi, University Paris Diderot, Paris, France

Pierre Raboisson Janssen Infectious Diseases and Diagnostics BVBA, Beerse, Belgium

Zoran Rankovic Eli Lilly, Indianapolis, IN, USA

Leslie Regad UMR-S973, MTi, University Paris Diderot, Paris, France

Allen B. Reitz Fox Chase Chemical Diversity Center, Inc., Doylestown, PA, USA

Bernd U. Riebesehl Technical Research & Development, Novartis Pharma AG, Basel, Switzerland

Jean-Philippe Rocher Pierre Fabre Research Institute, Castres, France

Tiago Rodrigues Swiss Federal Institute of Technology (ETH), Department of Chemistry and Applied Biosciences, Zürich, Switzerland

A. Romeo INN Programme, World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms

Jean-Michel Rondeau Novartis Institutes for BioMedical Research, Basel, Switzerland

Laurent Schaeffer Prestwick Chemical Inc., Boulevard Gonthier, Illkirch, France

Jean-Michel Scherrmann Department of Pharmaceutical Sciences, University Paris Descartes, Paris, France

Gisbert Schneider Swiss Federal Institute of Technology (ETH), Department of Chemistry and Applied Biosciences, Zürich, Switzerland

Herman Schreuder Sanofi, R&D LGCR/Struct., Design & Informatics FF, Frankfurt am Main, Germany

Wolfgang Sippl Medicinal Chemistry, Institute of Pharmacy, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany

Erin M. Skoda Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, USA

Olivier Taboureau UMR-S973, MTi, University Paris Diderot, Paris, France

Bernard Testa Department of Pharmacy, Lausanne University Hospital (CHUV), Lausanne, Switzerland

John Unitt Sygnature Discovery Limited, BioCity, Nottingham, UK

Johan Van de Voorde Vascular Research Unit, Ghent University, Gent, Belgium

Walter M.M. Van den Broeck Janssen Infectious Diseases BVBA, Beerse, Belgium

N. Murti Vemuri CMC Subject Matter Expert at BARDA/HHS, Conceptual MindWorks, Inc., Flemington, NJ, USA

Venkata Velvadapu Fox Chase Chemical Diversity Center, Inc., Doylestown, PA, USA

Franz Weiberth Sanofi, Waltham, MA, U.S.A

Bruno Villoutreix Université Paris Diderot, Sorbonne Paris Cité, Inserm, Paris, France

Camille G. Wermuth Founder of Prestwick Chemical, Strasbourg, France

Peter Wipf Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, USA

Gerhard Wolber Computer-Aided Molecular Design, Institute of Pharmacy, Freie Universität Berlin, Königin-Luisestr., Berlin, Germany

Foreword

The world's economy depends to a significant extent on our ability to deliver affordable and sustainable healthcare. As such, this new edition of *The Practice of Medicinal Chemistry* plays an important role in educating the next generation of scientists in the area as it goes beyond the simple delivery of new healing drugs to combat disease and illness. It enriches our knowledge and the very understanding of the lives of everyone on the planet. It is truly amazing how the different scientific disciplines can combine in this way to design and make such a wonderful array of functional molecules to serve some of our current needs. Nevertheless, the future presents enormous healthcare challenges that can only be met by appropriate investment and further fundamental scientific discovery. This new edition of *The Practice of Medicinal Chemistry* provides a unique scholarly compilation of the tools, techniques, and methods necessary to begin this journey of discovery, whether in industry or academia.

The book's practical overview differentiates this text from others. It provides a menu of topics that can be consulted individually, while also providing a holistic view covering the history of drug discovery through to the issues of today involving the consumption and production of pharmaceuticals. The process of drug discovery has become a highly complex operation requiring the medicinal chemist to acquire wide-ranging skills from areas such as biology, technology, modelling, delivery, physiochemistry, and synthesis. To pull this together in a single book is a heroic task that these authors have done magnificently.

As our science moves forward toward more biologicals, smaller volume products that are focused on patients, and more sustainable and flexible manufacturing in an ever more regulated environment, we will require new generations of creative individuals. They will need to be ever more innovative, using all the tools our modern society can offer. In particular, big-data mining, tissue sampling and genomic mapping, the "Internet of Things," and wearable health monitors will all be likely components in the armory of the next breed of medicinal chemist.

I regard this expanded textbook as essential reading for all those new to the field. It also provides a quality check for current practitioners in this rapidly evolving environment. The book also does not shy away from providing a future vision of the trends of the discipline. It is written by experts who elegantly convey their passion, experience, and insight for the benefit of all readers.

I therefore welcome this updated and expanded version of *The Practice of Medicinal Chemistry* and believe it provides—as did past editions—the bedrock of our discipline.

Steven V. Ley
Cambridge
http://www.leygroup.ch.cam.ac.uk/

Preface to the Fourth Edition

Bringing a new drug to patients is both a privilege and a challenge fraught with success and failure. A privilege because there can be no greater calling than to alleviate suffering to enable a healthier life.

> Without health, life is not life; it is only a state of languor and suffering—an image of death. ***Buddha***

A challenge fraught with success and failure as no drug makes it from idea to patients without experiencing success and failure.

> Success is not final; failure is not fatal. It is the courage to continue that counts. ***Winston Churchill***

Prof. Camille Wermuth recognized the need to capture in a single volume the essence of the disciplines needed by medicinal chemists, so as to enable those just entering the field or the seasoned professional to keep pace with the ever-changing nature of drug discovery and development. His vision became *The Practice of Medicinal Chemistry*, providing the medicinal chemistry community with access to experts from across the industry and academia who would share their knowledge to educate the community, thereby preparing the community to recognize and seize opportunities as they emerged.

> Fortune favors the prepared mind. ***Louis Pasteur***

The fourth edition has built off the previous editions. It is updated to reflect developments over the last seven years, including five new chapters on topics such as the evaluation of the biological activity of compounds and systems biology. More than seventy experts from ten countries have shared their insights and perspectives on the practice of medicinal chemistry.

The editorial work for the fourth edition has been shared by Camille Wermuth, Pierre Raboisson, Didier Rognan, and Dave Aldous. Odile Blin helped organize and shape how we initiated the fourth edition; we are indebted to her tireless professionalism. The editors wish to express their thanks to Molly M. McLaughlin and the Elsevier Academic Press, who have worked with us to keep this project moving forward.

I believe my final quote—from Jason Calacanis—captures the challenges medicinal chemists face every day.

> You have to have a big vision and take very small steps to get there. You have to be humble as you execute but visionary and gigantic in terms of your aspiration. In the Internet industry, it's not about grand innovation; it's about a lot of little innovations—every day, every week, every month—making something a little bit better. ***Jason Calacanis***

Medicinal chemistry is a highly collaborative and iterative process that has many paths. Being open, collaborative, and humble are qualities that will help you successfully navigate these paths from idea to patient.

Preface to the Third Edition

Like the preceding editions of this book, this third edition treats of the essential elements of medicinal chemistry in a unique volume. It provides a practical overview of the daily problems facing medicinal chemists, from the conception of new molecules through to the production of new drugs and their legal/economic implications. This edition has been updated, expanded and refocused to reflect developments in the past 5 years, including 11 new chapters on topics such as hit identification methodologies and cheminformatics. More than 50 experts in the field from eight different countries, who have benefited from years of practical experience, give personal accounts of both traditional methodologies and the newest discovery and development technologies, providing readers with an insight into medicinal chemistry.

A major change in comparison to the previous editions was the decision to alleviate my editorial burden in sharing it with seven section editors, each being responsible for one of the eight sections of the book. I highly appreciated their positive and efficacious collaboration and express them my warmest thanks (in the alphabetical order) to Michael Bowker, Hugo Kubinyi, John Proudfood, Bryan Reuben, Richard Silverman, David Triggle and Han van de Waterbeemd.

Another change was the decision taken by Elsevier/Academic Press to publish the book in full colors thus rendering it more pleasant and user-friendly. I take this occasion to thank Keri Witman, Pat Gonzales, Kirsten Funk and Renske van Dijk for having successively ensured the editorial development of the book. Taking into account that we had to work with a cohort of about 50 authors, each of them having his personality, his original approach and his main busy professional live, this was not an easy task. I am deeply indebted to my assistant Odile Blin for the way she had mastered, efficiently and with friendliness, all the secretarial work and particularly the contacts with the different authors and with the Elsevier development editors. As for the earlier editions, I also want to express my gratitude to my wife Renée and my daughters Delphine, Joëlle and Séverine for all their encouragements and for sacrificing many hours of family life in order to leave me enough free time to edit this new version of the "Medicinal Chemist's Bible."

My final thoughts go to the future readers of the book, and especially to the newcomers in Medicinal Chemistry having the curiosity to read the preface. I cannot resist giving them some advice for doing good science.

First of all, be open-minded and original. As Schopenhauer noted, the task of the creative mind is "not so much to see what no one has seen yet; but to think what nobody has thought yet, about what everyone sees." A wonderful illustration is found in Peter Hesse's cartoon below.

Second, always keep in mind that the object of Medicinal Chemistry is to synthetize new drugs useful for suffering patients. Like many scientists, medicinal chemists, have to navigate between two tempting reefs. On one side they should avoid doing "NAAR": non-applicable applied research, on the other side they may be attracted by "NFBR": non-fundamental basic search."

Third, convinced as they may be that the neighbors grass is always greener, they may be attracted to start their research in using as a hit a recently published competitor's product. In fact, the published compound may exhibit only a weak activity, therefore be very careful when starting a new program and never forget that the worst thing a medicinal chemist can do is to prepare a me-too of an inactive compound!

Camille G. Wermuth

Preface to the Second Edition

Like the first edition of *The Practice of Medicinal Chemistry* (nicknamed 'The Bible' by medicinal chemists) the second edition is intended primarily for organic chemists beginning a career in drug research. Furthermore, it is a valuable reference source for academic, as well as industrial, medicinal chemists. The general philosophy of the book is to complete the biological progress − Intellectualization at the level of function using the chemical progress Intellectualization at the level of structure (Professor Samuel J. Danishevsky, *Studies in the chemistry and biology of the epothilones and eleutherobins*, Conference given at the XXXIVémes Rencontres Internationales de Chimie Th6rapeutique, Facult6 de Pharmacie, Nantes, 8−10 July, 1998).

The recent results from genomic research have allowed for the identification of a great number of new targets, corresponding to hitherto unknown receptors or to new subtypes of already existing receptors. The massive use of combinatorial chemistry, associated with high throughput screening technologies, has identified thousands of hits for these targets. The present challenge is to develop these hits into usable and useful drug candidates. This book is, therefore, particularly timely as it covers abundantly the subject of drug optimization.

The new edition of the book has been updated, expanded and refocused to reflect developments over the nine years since the first edition was published. Experts in the field have provided personal accounts of both traditional methodologies, and the newest discovery and development technologies, giving us an insight into diverse aspects of medicinal chemistry, usually only gained from years of practical experience.

Like the previous edition, this edition includes a concise introduction covering the definition and history of medicinal chemistry, the measurement of biological activities and the three main phases of drug activity. This is followed by detailed discussions on the discovery of new lead compounds including automated, high throughput screening techniques, combinatorial chemistry and the use of the internet, all of which serve to reduce preclinical development times and, thus, the cost of drugs. Further chapters discuss the optimization of lead compounds in terms of potency, selectivity, and safety; the contribution of genomics; molecular biology and X-ray crystallization to drug discovery and development, including the design of peptidomimetic drugs; and the development of drug-delivery systems, including organ targeting and the preparation of pharmaceutically acceptable salts. The final section covers legal and economic aspects of drug discovery and production, including drug sources, good manufacturing practices, drug nomenclature, patent protection, social-economic implications and the future of the pharmaceutical industry.

I am deeply indebted to all co-authors for their cooperation, for the time they spent writing their respective contributions and for their patience during the editing process. I am very grateful to Didier Rognan, Paola Ciapetti, Bruno Giethlen, Annie Marcincal, Marie-Louise Jung, Jean-Marie Contreras and Patrick Bazzini for their helpful comments.

My thanks go also to the editorial staff of *Academic Press* in London, particularly to Margaret Macdonald and Jacqueline Read. Last but not least, I want to express my gratitude to my wife Renée for all her encouragements and for her comprehensiveness.

Camille G. Wermuth

Preface to the First Edition

The role of chemistry in the manufacture of new drugs, and also of cosmetics and agrochemicals, is essential. It is doubtful, however, whether chemists have been properly trained to design and synthesize new drugs or other bioactive compounds. The majority of medicinal chemists working in the pharmaceutical industry are organic synthetic chemists with little or no background in medicinal chemistry who have to acquire the specific aspects of medicinal chemistry during their early years in the pharmaceutical industry. This book is precisely aimed to be their 'bedside book' at the beginning of their career.

After a concise introduction covering background subject matter, such as the definition and history of medicinal chemistry, the measurement of biological activities and the three main phases of drug activity, the second part of the book discusses the most appropriate approach to *finding a new lead compound or an original working hypothesis*. This most uncertain stage in the development of a new drug is nowadays characterized by high-throughput screening methods, synthesis of combinatorial libraries, data base mining and a return to natural product screening. The core of the book (Parts III to V) considers the *optimization of the lead in terms of potency, selectivity, and safety*. In 'Primary Exploration of Structure-Activity Relationships', the most common operational stratagems are discussed, allowing identification of the portions of the molecule that are important for potency. 'Substituents and functions' deals with the rapid and systematic optimization of the lead compound. 'Spatial Organization, Receptor Mapping and Molecular Modelling' considers the three-dimensional aspects of drug-receptor interactions, giving particular emphasis to the design of peptidomimetic drugs and to the control of the agonist- antagonist transition. Parts VI and VII concentrate on the definition of satisfactory drug-delivery conditions, i.e. means to ensure that the molecule reaches its target organ. Pharmacokinetic properties are improved through adequate chemical modifications, notably prodrug design, obtaining suitable water solubility (of utmost importance in medical practice) and improving organoleptic properties (and thus rendering the drug administration acceptable to the patient). Part VIII, 'Development of New Drugs: Legal and Economic Aspects', constitutes an important area in which chemists are almost wholly self taught following their entry into industry.

This book fills a gap in the available bibliography of medicinal chemistry texts. There is not, to the author-editor's knowledge, any other current work in print which deals with the practical aspects of medicinal chemistry, from conception of molecules to their marketing. In this single volume, all the disparate bits of information which medicinal chemists gather over a career, and generally share by word-of-mouth with their colleagues, but which have never been organized and presented in coherent form in print, are brought together. Traditional approaches are not neglected and are illustrated by modern examples and, conversely, the most recent discovery and development technologies are presented and discussed by specialists. Therefore, *The Practice of Medicinal Chemistry* is exactly the type of book to be recommended as a text or as first reading to a synthetic chemist beginning a career in medicinal chemistry. And, even if primarily aimed at organic chemists entering into pharmaceutical research, all medicinal chemists will derive a great deal from reading the book.

The involvement of a large number of authors presents the risk of a certain lack of cohesiveness and of some overlaps, especially as each chapter is written as an autonomic piece of information. Such a situation was anticipated and accepted, especially for a first edition. It can be defended because each contributor is an expert in his/her field and many of them are 'heavyweights' in medicinal chemistry. In editing the book I have tried to ensure a balanced content and a more-or-less consistent style. However, the temptation to influence the personal views of the authors has been resisted. On the contrary, my objective was to combine a plurality of opinions, and to present and discuss a given topic from different angles. Such as it is, this first edition can still be improved and I am grateful in advance to all colleagues for comments and suggestions for future editions.

Special care has been taken to give complete references and, in general, each compound described has been identified by at least one reference. *For compounds for which no specific literature indication is given, the reader is referred to the Merck Index*.

The cover picture of the book is a reproduction of a copperplate engraving designed for me by the late Charles Gutknecht, who was my secondary school chemistry teacher in Mulhouse. It represents an extract of Brueghel's engraving *The alchemist ruining his family in pursuing his chimera*, surmounted by the aquarius symbol. Represented on the left-hand side is my lucky charm caster oil plant (*Ricinus communis L., Euphorbiaceae*), which was the starting point of the pyridazine chemistry in my laboratory. The historical cascade of events was as follows: cracking of caster oil produces n-heptanal and aldolization of n-heptanal – and, more generally, of any enolisable aldehyde or ketone – with pyruvic acid leads to a-hydroxy-γ-ketonic acids. Finally, the condensation of these keto acids with hydrazine yields pyrodazones. Thus, all our present research on pyridazine derivatives originates from my schoolboy chemistry, when I prepared in my home in Mulhouse n-heptanal and undecylenic acid by cracking caster oil!

Preparing this book was a collective adventure and I am most grateful to all authors for their cooperation and for the time and the effort they spent to write their respective contributions. I appreciate also their patience, especially as the editing process took much more time than initially expected.

I am very grateful to Brad Anderson (University of Utah, Salt Lake city), Jean-Jacques André (Marion Merrell Dow, Strasbourg), Richard Baker (Eli Lilly, Erl Wood, UK), Thomas C. Jones (Sandoz, Basle), Isabelle Morin (Servier, Paris), Bryan Reuben (London South Bank University) and John Topliss (University of Michigan, Ann Arbor) for their invaluable assistance, comments and contributions.

My thanks go also to the editorial staff of Academic Press in London, Particularly to Susan Lord, Nicola Linton and Fran Kingston, to the two copy editors Len Cegielka and Peter Cross, and finally, to the two secretaries of our laboratory, Franqois Herth and Marylse Wernert.

Last but not least, I want to thank my wife Renée for all her encouragement and for sacrificing evenings an Saturday family life over the past year and a half, to allow me to sit before my computer for about 2500 hours!

Camille G. Wermuth

SECTION ONE

General Aspects of Medicinal Chemistry

CHAPTER

1

Medicinal Chemistry: Definitions and Objectives, Drug Activity Phases, Drug Classification Systems

Peter Imming

Institut fuer Pharmazie, Martin-Luther-Universitaet Halle-Wittenberg, Halle, Germany

OUTLINE

Medicinal chemistry remains a challenging science which provides profound satisfaction to its practitioners. It intrigues those of us who like to solve problems posed by nature. It verges increasingly on biochemistry and on all the physical, genetic and chemical riddles in animal physiology which bear on medicine. Medicinal chemists have a chance to participate in the fundamentals of prevention, therapy and understanding of diseases and thereby to contribute to a healthier and happier life. **A Burger** [1]

I. DEFINITIONS AND OBJECTIVES

A. Medicinal Chemistry and Related Disciplines and Terms

A definition of medicinal chemistry was given by a IUPAC specialized commission: "*Medicinal chemistry* concerns the discovery, the development, the identification and the interpretation of the mode of action of biologically active compounds at the molecular level. Emphasis is put on drugs, but the interests of the medicinal chemist are not restricted to drugs but include bioactive compounds in general. Medicinal chemistry is also concerned with the study, identification, and synthesis of the metabolic products of these drugs and related compounds" [2].

Drugs—natural and synthetic alike—are chemicals used for medicinal purposes. They interact with complex chemical systems of humans or animals. Medicinal chemistry is concerned with this interaction, focusing on the organic and biochemical reactions of drug substances with their targets. This is one aspect of drug chemistry.

Other important aspects are the synthesis and the analysis of drug substances. The two latter aspects together are sometimes called *pharmaceutical chemistry*, but the synthesis of drugs is considered by some people—mainly chemists—to be part of medicinal chemistry, denoting analytical aspects as pharmaceutical chemistry. In German faculties of pharmacy, the literal translations of pharmaceutical and medicinal chemistry—Pharmazeutische and Medizinische Chemie—are used synonymously.

The general study of drugs is called *pharmacy* or *pharmacology*. A common narrower definition of pharmacology concentrates on the fate and effects of a drug in the body. *Clinical chemistry*, a different subject, is concerned with the determination of physiological and pathophysiological parameters in body fluids, such as enzyme activities and metabolites in blood and urine. The term *biopharmacy* has been reserved for the investigation and control of absorption, distribution, metabolism, excretion, and toxicology (ADMET) of drug substances.

Some further terms are more or less synonymous with medicinal chemistry: *(molecular) pharmacochemistry, drug design, selective toxicity*. The French equivalent to medicinal chemistry is chimie thérapeutique, and the German terms are Medizinische/Pharmazeutische Chemie and Arzneimittelforschung.

In academia, medicinal chemistry is a major subject in most pharmacy faculties—both for undergraduates and in research—and in many chemistry faculties. In the pharmaceutical industry, medicinal chemistry is at the heart of the search for new medicines.

The main activities of medicinal chemists are evident in the analysis of their most important scientific journals (e.g., *Journal of Medicinal Chemistry, European Journal of Medicinal Chemistry, Bioorganic and Medicinal Chemistry, ChemMedChem, Archiv der Pharmazie, Arzneimittelforschung, Chemical and Pharmaceutical Bulletin*).

The *objectives of medicinal chemistry* are as easily formulated as they are difficult to achieve: find, develop, and improve drug substances that cure or alleviate diseases (see below, Section I.C.) and understand the causative and accompanying chemical processes (see below, Section III.A).

Medicinal chemistry is an interdisciplinary science covering a particularly wide domain situated at the interface of organic chemistry with life sciences such as biochemistry, pharmacology, molecular biology, genetics, immunology, pharmacokinetics, and toxicology on one side, and chemistry-based disciplines such as physical chemistry, crystallography, spectroscopy, and computer-based techniques of simulation, data analysis, and data visualization on the other side.

B. Drugs and Drug Substances

Drugs are composed of *drug substances* (syn. *active pharmaceutical ingredients, APIs*) and *excipients* (syn. *ancillary substances*). The combination of both is the work of pharmaceutical technology (syn. *galenics*) and denoted a *formulation*.

In 2014, the World Drug Index contained over 80,000 marketed and development drug substances [3]. The United States *Orange Book* listed approx. 3,500 products in 2014, and the United States Pharmacopeia contains monographs of approx. 1,400 small-molecules Active Pharmaceutical Ingredients (APIs) and 160 biologic drug substances [4]. In 2013 in Germany, the "Rote Liste" contained approximately 6,000 drugs in 7,500 formulations representing approximately 2,000 APIs [5]. The WHO Essential Medicines List held approximately 350 drug substances in 2013 that WHO claims sufficient for the treatment of approx. 90 percent of all diseases where drugs are useful [6].

What makes a chemical "drug-like?" Because of the versatility of their molecular targets (see below), there can be no universal characteristic of drug substances. However, since the general structure of the target organisms is identical, generalizations as to drug substance structure are possible for biopharmacy [7,8]. For a chemical to be readily absorbed by the gut and distributed in the body, its size, hydrophilicity/lipophilicity ratio, stability toward acid media and hydrolytical enzymes, etc. have to meet defined physicochemical criteria. A careful analysis of reasons for drug attrition revealed that only 5 percent were caused by pharmacokinetic difficulties, whereas 46 percent were due to insufficient efficacy and 33 percent to adverse reactions in animals or humans [9]. Since both wanted and unwanted effects are due to the biological activity, 79 percent of drug candidates had unpredicted or wrongly predicted sum activities.

Predictions of toxicity from molecular features are still precarious [10–12]. Only rather general rules are for sure; such as avoidance of very reactive functional groups, for example, aldehyde because of oxidative instability and haptene nature; α,β-unsaturated carbonyl compounds and 2-halopyridines because of their unspecific reactivity as electrophiles. Torcetrapib is a typical example of toxicity—or adverse effects—challenges. It was an antiatherosclerotic drug candidate promising to become a blockbuster when in latter phase III of clinical trials, an increased risk of mortality led the company to discontinue its development. It was not clear whether the effects

were caused by the mechanism of action—inhibition of cholesteryl ester transfer protein—some other effect or an interaction with another drug. This is just one instance that "it isn't that simple [and] nothing's obvious and nothing's for certain" in rational drug development [13].

C. Stages of Drug Development

Most drugs were discovered rather than developed [14]. That is why a large number of drug substances are natural products or derivatives thereof. It is a matter of debate if ethnic medicines or nature still hold gems as yet undiscovered by pharmacy [15,16]. Synthetic substance collections ("libraries") have been created through (automated) organic chemistry. The very high number and diversity of natural and synthetic chemical entities is faced with an equally growing number of potential reaction partners (targets) from biochemical and pathophysiological research.

In virtual, biochemical and cell-based testing, compound selections are run against an isolated or physiologically embedded target that may be involved in the disease process [17]. Compounds that exceed a certain threshold value in binding to the target or modulation of some functional signal behind it, are called *hits*. If the identity and purity of the compound and the assay result are confirmed in a multipoint activity determination, the compound rises to the status of *validated hit*. From this one hopes to develop leads. A *lead* is a compound or series of compounds with proven activity and selectivity in a screen and fulfills some drug development criteria such as originality, patentability, and accessibility (by extraction or synthesis). Molecular variation hopefully tunes the physicochemical parameters so that it becomes suitable for ADME. An example of a small optimization algorithm is shown in Figure 1.1.

If the resulting *optimized lead* (*preclinical candidate*) displays no toxicity in cell and animal models, it becomes a *clinical candidate*. If this stands the tests of efficacy and safety in humans and overcomes marketing hurdles, a *new drug entity* will enter the treasure trove of pharmacy. Box 1.1 illustrates that activity is a necessary but not sufficient quality of medicines. There is, of course, no ideal drug in the real world, but one has to find a relative optimum. This often means developing a drug that has a different side-effect profile than drugs marketed for the same therapeutic indication so prescriptions can be tailored to the ways different patients react to a drug.

The role of medicinal chemistry is most prominent in steps one and two of drug development:

1. The discovery step, consisting of the choice of the therapeutic target (biochemical, cellular, or *in vivo* model; see below) and the identification or discovery and production of new active substances interacting with the selected target.
2. The optimization step that deals with the improvement of an active compound. The optimization process primarily takes into account the increase in potency, selectivity, and decrease in toxicity. Its characteristics are the establishment of structure–activity relationships, ideally based on an understanding of the molecular mode of action.

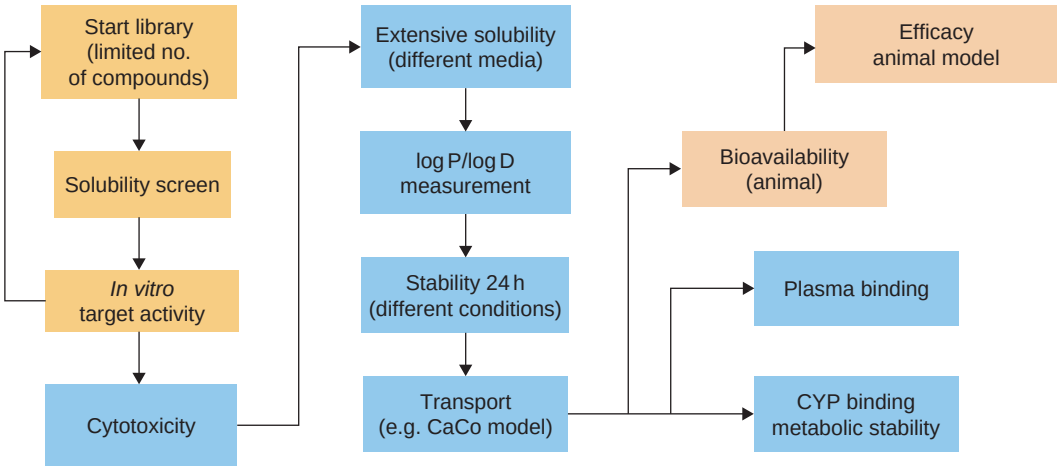


FIGURE 1.1 Example of an optimization algorithm. *Source: Adapted from a presentation by Dr. U. Heiser, Probiodrug AG, Halle, Germany, reproduced with permission.*

3. The formulation step, whose purpose is the continuation of the improvement of the pharmacokinetic properties and the fine-tuning of the pharmaceutic properties of active substances to render them suitable for clinical use. This can consist—to name just a few instances—of the preparation of better absorbed compounds, of sustained release formulations, and of water-soluble derivatives or in the elimination of properties related to the patient's compliance (irritation, painful injection, undesirable organoleptic properties). For an example, see Figure 1.2.

The main tasks of medicinal chemistry consist of the optimization of the following characteristics:

a. Higher affinity and target-intrinsic activation for better clinical activity so the dosage and nonspecific side effects will be as low as possible. There are no examples of drugs that are dosed below 10 mg/day that cause idiosyncratic adverse drug reactions. For drug substances that have to be given in higher doses—i.e., the majority—medicinal chemistry tries to find active derivatives that will be metabolized in a safe way [18]. This includes assaying for inhibition of or reaction with key enzymes of biotransformation, such as oxidases of the cytochrome type, some of which are highly demanded by food constituents and xenobiotics including drug substances [19]. Medicinal chemistry tries to prepare drugs that are not metabolized by bottleneck enzymic pathways [20].

b. Better selectivity, which may lead to a reduction of unwanted side effects. This sometimes entails the assaying of a very high number of other targets; for example, an antidepressive serotonin re-uptake inhibitor has to be tested against all subtypes of serotonin, adrenaline, and dopamine receptors, plus many other key receptors and enzymes.

BOX 1.1

THE IDEAL NEW DRUG SUBSTANCE

- New chemical entity for patentability and registration.
- Maximum four-step synthesis with, for example, no heavy metal catalysts and no environmentally problematic waste; no chromatographic purification steps; purity>99 percent.
- Stable up to 70°C even in humid air and light.
- Solid-state properties (crystalline, not polymorphous, not hygroscopic) that make it a perfect partner for (tablet) compaction.
- Solubility in water sufficient for the production of stable blood-isotonic solutions.
- Oral bioavailability .90 percent with no interindividual variation.
- Very high activity and pharmacokinetic profile enable once-a-day-dosage at 5–10 mg.

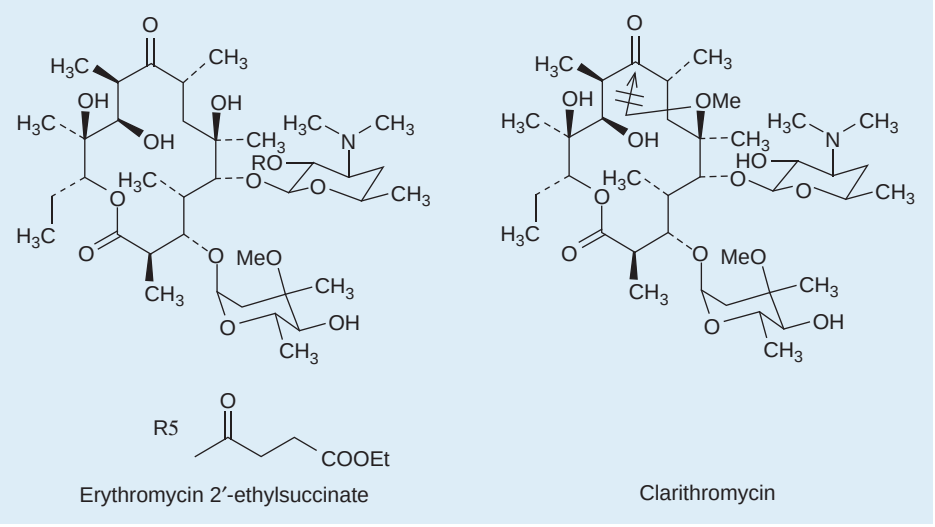


FIGURE 1.2 An example of fine-tuning of pharmacologically active chemicals: Erythromycin 29-ethylsuccinate and clarithromycin are semisynthetic derivatives of the macrolide antiinfective erythromycin. The small molecular change in the former leads to the elimination of bitterness which is important as this class of drugs is often used in pediatrics and administered as a syrup. In the latter, because hemiketal formation is no longer possible (arrow), clarithromycin is stable in the acidic milieu of the stomach (pH 2).

In spite of the high number of compounds, targets, and assays, the development pipeline of new chemical entities as drug substances has not got fuller in the past 20 years. For possible explanations, see the discussion of drug targets below and Ref. [9].

II. DRUG ACTIVITY PHASES

The progression of a drug into the body, to its target(s), and out again can be broken down into three mechanistically distinct phases, the second and third being partly simultaneous. During drug development, all three phases are investigated interdependently, because structural changes required for one phase must not abolish suitability in another phase.

A. The Pharmaceutical Phase

Drug substances are applied orally (preferred mode) or parenterally (e.g., by subcutaneous or intravenous injection, rectally, or through inhalation). A combination of the skills of medicinal chemists and pharmaceutical technologists has to provide the drug candidate in suitable formulations. For tablets, the drug substance needs to be crystalline and not have a low melting point. For injections, it should be water soluble (e.g., as a salt). The required structural features must be compatible with the pharmacological activity, of course.

B. The Pharmacokinetic Phase

For this phase, medicinal chemists and biopharmacists work together to design a compound that will have suitable ADME parameters. Sufficient solubility in an aqueous medium for absorption and blood transport has to be combined with sufficient lipophilicity for passage through cell membranes. If an active compound is too hydrophilic and at the same time contains a carboxylic acid group, for instance, conversion to a simple ester will facilitate absorption. Once in the blood, unspecific esterases will catalyze hydrolysis to the active carboxylic acid form. Such an ester is an instance of a *prodrug*.

Drug substances should remain active and in the body for a period of time that is neither too short nor too long. For many drugs, a metabolic and/or excretion rate that enables "once a day" dosage is sought. Sometimes this requires the identification of sites in the molecule that will be metabolized quickly with concomitant loss of activity. The vasodilator iloprost, for instance, was developed from the endogeneous mediator prostacyclin that has very short half-life both *in vivo* and on the shelf. Modification of several chemically and metabolically vulnerable positions yielded a stable and active derivative—a highly sophisticated product of synthetic medicinal chemistry (Figure 1.3) [21]. By contrast, sometimes functionality is introduced for the acceleration of biotransformation and excretion. Articaine is a local anesthetic of the anilide type. Systemically, it interferes with heart rate—an unwelcome side effect in dentistry. That is why articaine contains an additional ester group. Once in the blood stream, this will be hydrolyzed quickly to an inactive carboxylic acid (Figure 1.4) [22]. Medicinal chemistry here has come full circle, as anilide local anesthetics were developed from ester anesthetics like procain in order to prolong activity.

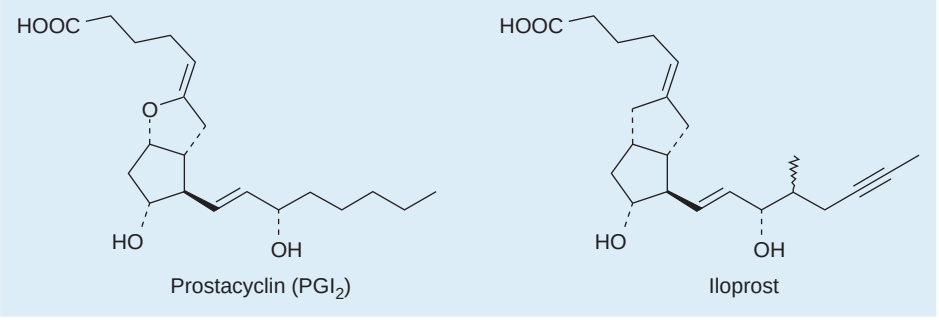


FIGURE 1.3 Prostacyclin and its synthetic analog, iloprost, that combines activity with sufficient *ex vivo* and *in vivo* stability.

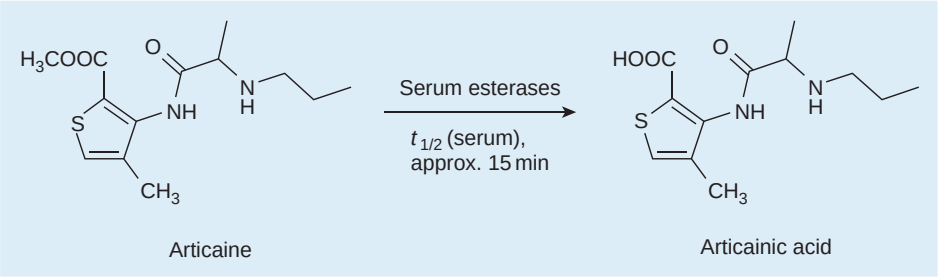


FIGURE 1.4 Articaine, a common local anesthetic dentists use, and its inactive metabolite that is formed off the scene of painful action. The value for $t_{1/2}$ is from the reference Oertel R, Ebert U, Rahn R, Kirch W. The effect of age on pharmacokinetics of the local anesthetic drug articaine. Reg Anesth Pain Med 1999;24:524–8.

C. The Pharmacodynamic Phase

While pharmacokinetics investigates what the body does to the drug, pharmacodynamics is concerned with what the drug does to the body. Most scientists who consider themselves medicinal chemists will be most comfortable with and interested in this phase. They will cooperate with biochemists and pharmacologists to elucidate mechanistic details of the interaction of the drug with its target(s), a topic we will treat in the Section III.

D. The Road to Successful Drug Development?

In the past years, many analyses have appeared that try to explain the dearth of new drug substances in the face of billions of dollars that have been spent, billions of assay data points that have been accumulated, and ten thousands of virtual and thousands of real hits that have been generated. By comparison, the Belgian medicinal chemist Paul Janssen and his relatively small group had tremendous success in the development of new drug entities and activities [23]. It was postulated that the individualization rather than integration of research guidelines into successive hypes (e.g., "as target subtype selective as possible"; "ADME rules have to be strictly adhered to"; "modeling programs automatically give a correct representation of molecules"; "the more combinatorial ligands, the more hits") is responsible for the disappointing state of drug discovery. What is needed is to keep what we already know about how successful drugs were actually discovered or invented [24], while providing an atmosphere of creativity in a team of scientists from various disciplines. Summarizing their long-lasting experiences in antibacterial research, an industrial team concluded that for this therapeutic area at least, synthesizing novel chemical structures that interact with and block *established* targets in new ways is a robust strategy [24,25].

So what does the increasing knowledge of targets mean for medicinal chemistry? This subject will be introduced in the following paragraphs and discussed in detail in later chapters.

III. DRUG CLASSIFICATION SYSTEMS

Classification systems help with understanding what a drug actually does at the molecular level (classification by target), and they are indispensable for categorizing the large number of drug substances (classification by clinical effect).

A. Classification by Target and Mechanism of Action

1. Targets

Targets are molecular structures, chemically definable by at least a molecular mass, that will undergo a specific interaction with chemicals that we call drugs because they are administered to treat or diagnose a disease [26]. To be meaningful, the interaction has to have a connection with the clinical effect(s). It is very challenging to prove that the interaction of a drug substance with a specific molecular target indeed triggers the clinical effect(s).

A clinically relevant target might consist not of a single biochemical entity but the simultaneous interference of a number of receptors. Only this multi-target interaction will give a net clinical effect that might be considered beneficial. It is only by chance that the current *in vitro* screening techniques will identify drugs that work through such targets.

The number of targets presently used is still open to discussion in medicinal chemistry, but various approaches concurred in finding several hundred. The number of potential targets, however, was estimated to be

several hundred thousand in view of the manifold protein complexes, splicing variants, and possible interventions with signaling pathways [26,27]. The problem with counting is two-fold: first, the identification of the reaction partners of drug substances in the body; and second, exactly what to define and count as the target. A target definition derived from the net effect rather than the direct chemical interaction will require input from systems biology, a research field that promises to affect the drug discovery process significantly [28]. At the other end of the scale of precision, we can define some targets very precisely on the molecular level. For example, we can say that dihydropyridines block the CaV1.2a splicing variant in heart muscle cells of L-type high-voltage activated calcium channels.

The actual depth of detail used to define the target is primarily dependent on the amount of knowledge available about the target and its interactions with a drug. Even if the target structure has already been determined, the molecular effect of the drug could still not be fully described by the interactions with one target protein alone. For example, antibacterial oxazolidinones interact with 23S-rRNA, tRNA, and two polypeptides, ultimately leading to inhibition of protein synthesis [29]. In this case, a description of the mechanism of action that only includes interactions with the 23S-rRNA target would be too narrowly defined. In particular, in situations in which the dynamic actions of the drug substance stimulate or inhibit a biological process, it is necessary to move away from the description of single proteins, receptors and other targets to view the entire signal chain as the target. and so on to view the entire signal chain as the target.

Lists that classify all marketed drug substances according to target, with references, were published. An excerpt is given in Table 1.1 [26].

2. *Mechanisms of Action*

An effective drug target comprises a biochemical system rather than a single molecule. Present target definitions are static. We know this to be insufficient, but techniques to observe the dynamics of drug–target interactions are just being created. Most importantly, we are not able to gauge the interaction of the biochemical "ripples" that follow the drug's initial molecular effect. The first molecular step of drug activity consists in mass-action governed drug–target recognition. For clinically observable activity, a series of biochemical steps need to follow that have to shift physiological equilibria in a transient way. Indeed, the gap between chain and circles of molecular events and clinical effects is still wide open, as reflected by the complementarity of target and phenotypic-oriented drug discovery approaches [30].

Although the term "mechanism of action" itself implies a classification according to the dynamics of drug substance effects at the molecular level, the dynamics of these interactions are only speculative models at present, and so mechanism of action can currently only be used to describe static targets, as discussed above.

All drugs somehow interfere with signal transduction, receptor signaling, and biochemical equilibria. For many drugs we know—and for most we suspect—that they interact with more than one target. So there will be simultaneous changes in several biochemical signals, and there will be feedback reactions of the pathways disturbed. In most cases, the net result will not be linearly deducible from single effects. For drug combinations, this is even more complicated. Awareness is also increasing of the nonlinear correlation of molecular interactions and clinical effects. For example, the importance of receptor–receptor interactions (receptor mosaics) was summarized for G-protein-coupled receptors (GPCRs), resulting in the hypothesis that cooperativity is important for the decoding of signals, including drug signals [31]. Table 1.2 lists examples of dynamic molecular mechanisms of drugs. Table 1.1 is the excerpt of an attempt at a complete list of drug targets. Notably, inhibitors and antagonists by far outnumber effectors, agonists, and substitutes. It appears that reconstitution of biochemical and pharmacological balances is more easily achieved by blocking excessive or complementary pathways rather than by substitution or repair of deficient or defective biochemical input.

Greater knowledge of how drugs interact with the body (e.g., mechanisms of action, drug–target interactions) has led to a reduction of established drug doses and inspired the development of newer, highly specific drug substances with a known mechanism of action. However, a preoccupation with the molecular details has resulted in a tendency to focus only on this one aspect of the drug effects. For example, cumulative evidence suggests that the proven influence of certain psychopharmaceuticals on neurotransmitter metabolism has little to do with the treatment of schizophrenia or the effectiveness of the drug for this indication [32]. With all our efforts to understand the molecular basis of drug action, we must not fall into the trap of reductionism. For antibacterial research, multitargeting is now considered to be essential [33]. More generally, in recent years the limits of the reductionist approach in drug discovery have become painfully clear. Nobel laureate Roald Hoffmann put it this way: "Chemistry reduced to its simplest terms, is not physics. Medicine is not chemistry … knowledge of the specific physiological and eventually molecular sequence of events does not help us understand what [a] poet has to say to us" [34]. The cartoon (Figure 1.5) illustrates this point. Although it is too early for systems biology

TABLE 1.1 The Main Drug Target Classes with Examples of Targets and Ligands. A Full List Can Be Found in Ref. [26]

Target class	Target subclass	Target example	Drug substance example (activity)
Enzymes	Oxidoreductases	Aldehyde dehydrogenase	Disulfiram (inhibitor)
	Transferases	Protein kinase C	Miltefosine (inhibitor)
	Hydrolases	Bacterial serine protease	β-Lactams (inhibitors)
	Lyases	DOPA decarboxylase	Carbidopa (inhibitor)
	Isomerases	Alanine racemase	D-cycloserine (inhibitor)
	Ligases (syn. synthases)	Dihydropteroate synthase	Sulphonamides (inhibitors)
Proteins	Growth factors	Vascular endothelial growth factor	Bevacizumab (antibody)
	Immunoglobulins	CD3	Muromonab-CD3 (antibody)
	Integrins	α4-Integrin subunit	Natalizumab (antibody)
	Tubulin	Human spindle	Vinca alkaloids (development inhibitors)
Substrates, metabolites	Substrate	Asparagine	Asparaginase (enhanced degradation)
	Metabolite	Urate	Rasburicase (enhanced degradation)
Receptors	Direct ligand-gated ion channel receptors	γ-Aminobutyric-acid (GABA)-	Barbiturates (allosteric agonists)
	G-protein-coupled receptors	Acetylcholine receptors	Pilocarpine (muscarinic receptor agonist)
		Opioid receptors	
		Prostanoid receptors	Buprenorphine (κ-opioid antagonist)
			Misoprostol (agonist)
	Cytokine receptors	TNFα receptors	Etanercept (receptor mimic)
	Integrin receptors	Glycoprotein IIb/IIIa receptor	Tirofiban (antagonist)
	Receptors associated with a tyrosine kinase	Insulin receptor	Insulin (agonist)
	Nuclear receptors, steroid hormone receptors	Mineralocorticoid receptor	Aldosterone (agonist)
	Nuclear receptors, other	Retinoic acid receptors	Isotretinoin (RARα agonist)
Ion channels	Voltage-gated Ca^{2+} channels	L-type channels	Dihydropyridines (inhibitors)
	K^+ channels	Epithelial K^+ channels	Diazoxide (opener)
	Na^+ channels	Voltage-gated Na^+ channels	Carbamazepine (inhibitor)
	Ryanodine-inositol 1,4,5-triphosphate receptor Ca^{2+} channel	Ryanodine receptors	Dantrolene (inhibitor)
	Transient receptor potential Ca^{2+} channel	TRPV1 receptors	Acetaminophen metabolite (inhibitor)
	Chloride channels	Mast cell chloride channels	Cromolyn sodium (inhibitor)
Transport proteins	Cation-chloride cotransporter family	Thiazide-sensitive NaCl symporter	Thiazide diuretics (inhibitors)
	Na^+/H^+ antiporters		Amiloride (inhibitor)
	Proton pumps	H^+/K^+ ATPase	Omeprazole (inhibitor)
	Eukaryotic (putative) sterol transporter (EST) family	Niemann-Pick C1 like 1 protein	Ezetimibe (inhibitor)
	Neurotransmitter/Na^+ symporter family	Serotonin/Na^+ symporter	Paroxetine (inhibitor)

(*Continued*)

TABLE 1.1 (Continued)

Target class	Target subclass	Target example	Drug substance example (activity)
DNA, RNA	Nucleic acids	Bacterial 16S-RNA	Aminoglycosides (protein synthesis inhibition)
	Ribosome	Bacterial 30S subunit	Tetracyclines (protein synthesis inhibition)
Physicochemical mechanism	Ion exchange	Hydroxide	Fluoride (enhanced acid stability of adamantine)
	Acid binding	In stomach	Hydrotalcite
	Adsorptive	In gut	Charcoal
	Surface-active	On oral mucosa	Chlorhexidine (disinfectant)
	Oxidative	On skin	Permanganate (disinfectant)
	Reductive	Disulphide bonds	*N*-acetylcysteine (mucolytic)
	Osmotically active	In gut	Lactulose (laxative)

TABLE 1.2 Examples of Dynamic (Process) Mechanisms of Drug Action

Dynamic mechanism	Example
Covalent modifications of the active center	Acylation of bacterial transpeptidases by β-lactam antibiotics
Drugs that require the receptor to adopt a certain conformation for binding and inhibition	Trapping of K^+ channels by methanesulphonanilide antiarrhythmic agents
Drugs that exert their effect indirectly and require a functional background	The catechol *O*-methyltransferase inhibitor entacapone, the effect of which is due to the accumulation of nonmetabolized dopamine
Anti-infectives that require the target organism to be in an active, growing state	β-Lactam antibacterials
Molecules requiring activation (prodrugs)	Enalaprilate, paracetamol
Modifications of a substrate or cofactor	Asparaginase, which depletes tumour cells of asparagine;[a] isoniazide, which is activated by mycobacteria leading to an inactive covalently modified NADH[b]
Simultaneous modulation of several signaling systems	GPCR receptor mosaics for the decoding of drug signals
Fluctuations of physiological signaling molecules	Dopamine fluctuations after administration of cocaine, followed by a gradual increase in steady state dopamine concentration[c]

[a]*Graham ML. Pegaspargase: a review of clinical studies. Adv Drug Deliv Rev 2003;55:1293–302.*
[b]*Larsen MH, Vilchèze C, Kremer L, Besra GS, Parsons L, Salfinger M, et al. Overexpression of inhA, but not kasA, confers resistance to isoniazid and ethionamide in Mycobacterium smegmatis, M. bovis BCG, and M. tuberculosis. Mol Microbiol 2002;46:453–66.*
[c]*Heien ML, Khan AS, Ariansen JL, Cheer JF, Phillips PE, Wassum KM, et al. Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats. Proc Natl Acad Sci USA 2005;102:10023–8.*

to provide clear-cut protocols for medicinal chemistry, translational medicine [35] and other integrative research efforts stress the functional as opposed to reductionist character of living systems, hopefully improving the success rate of drug research [36].

B. Other Classification Systems

From a pharmaceutical standpoint, many different criteria can be used to classify medications: type of formulation, the frequency with which it is prescribed or recommended, price, refundibility, prescription or nonprescription medication, etc. If a classification of the APIs is undertaken, numerous possibilities are revealed as well. At the end of the 19th century, drug substances were classified the same as other chemical entities: by nature of their primary elements, functional moieties, or organic substance class. Recently, the idea of classifying drug substances strictly according to their chemical constitution or structure has been revived. Databases attempt to gather

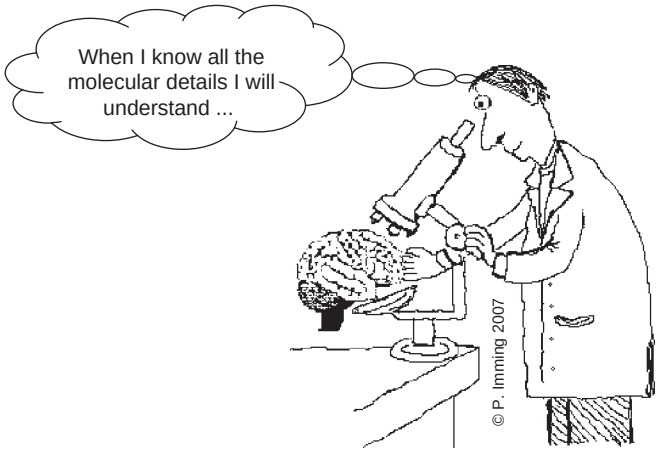


FIGURE 1.5 Searching for molecular mechanisms … "The meaning of the message will not be found in the chemistry of the ink." Sperry R. Brain circuits and functions of the mind. Cambridge: Cambridge University Press; 1990. *Source: Roger Sperry, neurophysiologist, Nobel Prize in Medicine, 1981.*

and organize information on existing or potential drug substances according to their chemical structure and diversity. The objective is to create—virtual or real—substance or fragment "libraries" that contain pertinent information about possible ligands for new targets (e.g., an enzyme or receptor) of clinical interest [37,38], and, more importantly, to understand the systematics of molecular recognition (ligand−receptor) [39,40].

The most commonly used classification system for drug substances is the *ATC system* [41]. It was introduced in 1976 by the Nordic Council on Medicines as a method for carrying out drug utilization studies throughout Scandinavia. In 1981, the World Health Organization recommended the use of the ATC classification for all global drug utilization studies, and in 1982 founded the WHO Collaborating Centre for Drugs Statistics Methodology in Oslo to establish and develop the method. The ATC system categorizes drug substances at five different levels according to (1) the organ or system on which they act (anatomy), (2) therapeutic and pharmacological properties, and (3) chemical properties. The first level comprises the main anatomical groups, while the second level contains the pharmacologically relevant therapeutic subgroup. The third level consists of the pharmacological subgroup, and the fourth the chemical subgroup. The fifth level represents the chemical substance (the actual drug entity). Drugs with multiple effects and different target organs can be found more than once within the system. The anti-inflammatory agent diclofenac, for instance, has three ATC numbers, one of them being M01AB05. This key breaks down to: M01 (musculo-skeletal system; anti-inflammatory and antirheumatic agents, nonsteroids); M01AB (acetic acid derivatives and related substances); and 05 (diclofenac in M01AB). The two other keys classify diclofenac as a topical agent and its use for inflammation of sensory organs.

While ATC is better suited if the emphasis is on therapeutic use, the TCAT system [26,42] puts the target chemistry first, particularly suiting the medicinal chemical approach.

References

[1] Burger A. Preface. In: Hansch C, Sammes PG, Taylor JB, editors. Comprehensive medicinal chemistry. Oxford: Pergamon Press; 1990. p. 1.
[2] Wermuth CG, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry. IUPAC Recommendations 1997, Part I. In: Adam J, editor. Annual reports in medicinal chemistry. San Diego, CA: Academic Press; 1998. p. 385−95. For Part II, see: <http://www.iupac.org/home/publications/provisional-recommendations/published/published-container/glossary-of-terms-used-in-medicinal-chemistry-part-ii.html>.
[3] World Drug Index. Available from Thomson Scientific at: <http://www.thomsonreuters.com/world-drug-index/>.
[4] US Pharmacopeial Convention. Available from: <http://www.usp.org/>.
[5] Rote Liste 2013. Frankfurt/Main: Rote Liste Service GmbH. <http://www.rote-liste.de/>; 2014.
[6] World Health Organisation. WHO model list of essential medicines. Available from: <http://www.who.int/medicines/publications/essentialmedicines/en/>; 2013.
[7] Alqahtani S, Mohamed LA, Kaddoumi A. Experimental models for predicting drug absorption and metabolism. Expert Opin Drug Metab Toxicol 2013;9:1241−54.
[8] Cheng F, Li W, Liu G, Tang Y. In silico ADMET prediction: recent advances, current challenges and future trends. Curr Top Med Chem 2013;13:1273−89.
[9] Schuster D, Laggner C, Langer T. Why drugs fail—a study on side effects in new chemical entities. Curr Pharmaceut Des 2005;11:3545−59.

[10] Guengerich FP, MacDonald JS. Applying mechanisms of chemical toxicity to predict drug safety. Chem Res Toxicol 2007;20:344–69.
[11] Bai JP, Abernethy DR. Systems pharmacology to predict drug toxicity: integration across levels of biological organization. Annu Rev Pharmacol Toxicol 2012;53:451–73.
[12] Segall MD, Barber C. Addressing toxicity risk when designing and selecting compounds in early drug discovery. Drug Discov Today 2014;19:688–93. Available from: http://dx.doi.org/10.1016/j.drudis.2014.01.006.
[13] Lowe D. In the pipeline: trocetrapib. Chemistry World 2007;4:16.
[14] Sneader W. Drug discovery. Chichester: Wiley-Interscience; 2008.
[15] Li JW, Vederas JC. Drug discovery and natural products: end of an era or an endless frontier? Science 2009;325:161–5.
[16] Koehn FE, Carter GT. Industrial natural product chemistry for drug discovery and development. Nat Prod Rep 2014;31:35–60.
[17] Potyrailo R, Rajan K, Stoewe K, Takeuchi I, Chisholm B, Lam H. Combinatorial and high-throughput screening of materials libraries: review of state of the art. ACS Comb Sci 2011;13:579–633.
[18] Kalgutkar AS, Soglia JR. Minimizing the potential for metabolic activation in drug discovery. Expert Opin Drug Metab Toxicol 2005;1:91–141.
[19] Zlokarnik G, Grootenhuis PDJ, Watson JB. High throughput P450 inhibition screen in early drug discovery. Drug Discov Today 2005;10:1443–50.
[20] Pelkonen O, Raunio H. In vitro screening of drug metabolism during drug development: can we trust the predictions? Exp Opin Drug Metab Toxicol 2005;1:49–59.
[21] Skuballa W, Schillinger E, Stuerzebecher CS, Vorbrueggen H. Prostaglandin analogs. Part 9. Synthesis of a new chemically and metabolically stable prostacyclin analog with high and long-lasting oral activity. J Med Chem 1986;29:313–15.
[22] Oertel R, Rahn R, Kirch W. Clinical pharmacokinetics of articaine. Clin Pharmacokin 1997;33:417–25.
[23] Black J. A personal perspective on Dr. Paul Janssen. J Med Chem 2005;48:1687–8.
[24] Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nat Rev Drug Discov 2007;6:29–40.
[25] Jordan AM, Roughley SD. Drug discovery chemistry: a primer for the non-specialist. Drug Discov Today 2009;14:731–44.
[26] Imming P, Sinning C, Meyer A. Drugs, their targets and the nature and number of drug targets. Nat Rev Drug Discov 2006;5:821–34.
[27] DrugBank. Available from: <http://www.drugbank.ca/>.
[28] Lowe JA, Jones P, Wilson DM. Network biology as a new approach to drug discovery. Curr Opin Drug Discov Devel 2010;13:524–6.
[29] Colca JR, McDonald WG, Waldon DJ, Thomasco LM, Gadwood RC, Lund ET, et al. Cross-linking in the living cell locates the site of action of oxazolidinone antibiotics. J Biol Chem 2003;278:21972–9.
[30] Swinney DC. Phenotypic vs. target-based drug discovery for first-in-class medicines. Clin Pharmacol Ther 2013;93:299–301.
[31] Fuxe K, Marcellino D, Borroto-Escuela DO, Frankowska M, Ferraro L, Guidolin D, et al. The changing world of G protein-coupled receptors: from monomers to dimers and receptor mosaics with allosteric receptor-receptor interactions. J Recept Signal Transduct Res 2010;30:272–83.
[32] Hyman SE, Fenton WS. What are the right targets for psychopharmacology? Science 2003;299:350–1.
[33] Silver LL. Multi-targeting by monotherapeutic antibacterials. Nat Rev Drug Discov 2007;6:41–55.
[34] Hoffmann R. Speech at the Nobel banquet. Available from: <http://nobelprize.org/nobel_prizes/chemistry/laureates/1981/hoffmann-speech.html>; 1981.
[35] Albani S, Prakken B. The advancement of translational medicine-from regional challenges to global solutions. Nat Med 2009;15:1006–9.
[36] McGonigle P, Ruggeri B. Animal models of human disease: challenges in enabling translation. Biochem Pharmacol 2014;87:162–71.
[37] Peng Z. Very large virtual compound spaces: construction, storage and utility in drug discovery. Drug Discov Today Technol 2013;10:e387–94.
[38] Wilde F, Link A. Advances in the design of a multipurpose fragment screening library. Expert Opin Drug Discov 2013;8:597–606.
[39] Hendlich M, Bergner A, Gunther J, Klebe G. Relibase: design and development of a database for comprehensive analysis of protein–ligand interactions. J Mol Biol 2003;326:607–20. Available from: <http://relibase.ccdc.cam.ac.uk/>.
[40] Parenti MD, Rastelli G. Advances and applications of binding affinity prediction methods in drug discovery. Biotechnol Adv 2012;30:244–50.
[41] WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index. Available from: <http://www.whocc.no/atc_ddd_index/>; 2015.
[42] Imming P, Buss T, Dailey LA, Meyer A, Morck H, Ramadan M, et al. A classification of drug substances according to their mechanism of action. Pharmazie 2004;59:579–89.

CHAPTER

2

Evaluation of the Biological Activity of Compounds: Techniques and Mechanism of Action Studies

Iain G. Dougall[1] *and John Unitt*[2]

[1]IGD Consultancy Ltd, Loughborough, Leics, UK; [2]Sygnature Discovery Limited, BioCity, Nottingham, UK

OUTLINE

I. INTRODUCTION

Drugs are molecules that are designed to perturb biological systems (cells, isolated tissues, whole animals, and ultimately patients). The responses observed are, in most cases, the result of the drug interacting with proteins, which have the capacity to convert chemical information into biological information. These proteins include plasma membrane bound receptors such as G protein coupled receptors (GPCRs) and tyrosine kinase receptors, ion channels (both ligand gated and voltage operated), enzymes, transporters, and transcription factors such as the nuclear hormone receptors (NHRs), which bind to specific consensus sequences of DNA and modulate gene transcription. Many of these target classes have been very successfully exploited to produce pharmacological agents designed to treat human (and animal) diseases (Figure 2.1). However, a significant number of potentially therapeutically useful drug targets have proven to be less tractable to small molecule approaches. In many cases this is because the interactions in question are protein–protein in nature and therefore difficult to modulate with conventional drugs. This realization has driven significant efforts in the field of biologicals. These large molecular weight agents, which include monoclonal antibodies, oligonucleotides, and small interfering RNAs (siRNA),

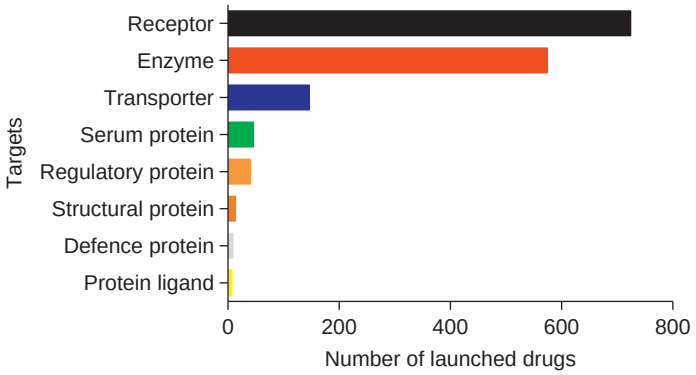


FIGURE 2.1 Number of Launched Small Molecule Drugs by Protein Target Class. As of May 2014, according to Thomson Reuters Cortellis.

have significantly increased the armamentarium of researchers allowing them to probe the role of previously intractable targets in human disease (see Chapter 3). However, to date, with the exception of monoclonal antibodies (such as the anti-TNFs) very few of these biological agents have made it to the market. In addition, the higher cost of large molecules means that small molecule drug programs remain an attractive proposition, even for chemically challenging targets.

The focus of this chapter is the description of how small molecules synthesized by medicinal chemists are assessed for biological activity (although many of the principles are equally applicable to large molecules). Such agents have been the mainstay of pharmacological treatment of human disease for decades and remain a very important class of drugs in the continuing search for new medicines to address unmet clinical needs. Historically, medicinal chemists used the naturally occurring ligands or substrates (for enzymes) of target proteins as starting points for small molecule-based research programs, although nowadays "hits" from high throughput screens (HTS), fragment screens, or *in silico* screens are more likely to act as initial "leads." Irrespective of the chemical basis of the program, the assays employed, the data generated, and their subsequent analysis form the basis of screening cascades that are designed ultimately to identify and progress molecules with appropriate properties for clinical testing. If the target is novel, and in the absence of definitive data linking it to human disease (e.g., genetic association data such as the link between the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) dysfunction and cystic fibrosis (CF)), such clinical testing provides the ultimate validation (or invalidation) of the target.

II. DRUG DISCOVERY APPROACHES AND SCREENING CASCADES

A. Target Based Screening

Modern day drug discovery programs largely center on target based screening, i.e., they aim to identify compounds that modulate the activity of a target that is potentially implicated in a human disease. To this end, compounds are typically tested in a range of *in vitro* biological assays designed to measure primary activities (potency, intrinsic activity, and/or efficacy), selectivity (activity versus related and unrelated targets), cellular toxicity, and physiologically relevant activity. The primary assay sits at the top of the screening cascade and data derived from it drive understanding of structure-activity relationships (SAR), allowing compound optimization. Criteria are set at each level of the cascade for compound progression to the next assay. Compounds with suitable robust properties progress to animal model testing with the aim of showing activity in a "disease relevant" setting as a prelude to picking a candidate drug (CD) for clinical trials. As drug discovery programs progress from early (Hit and Lead Identification) to late (Lead Optimisation, Candidate Selection) phases, the screening cascade evolves to become increasingly complex. Figure 2.2A illustrates a typical example of a screening cascade, although it is important to realize that the cascade used will be target dependent.

B. Phenotypic Screening

Although target based screening has proven to be very successful in the discovery of new medicines, a major disadvantage of the approach is that the evidence linking the target to the disease is often relatively weak, and the hypothesis is only proven (or disproven) after considerable investment of time, effort, and money.

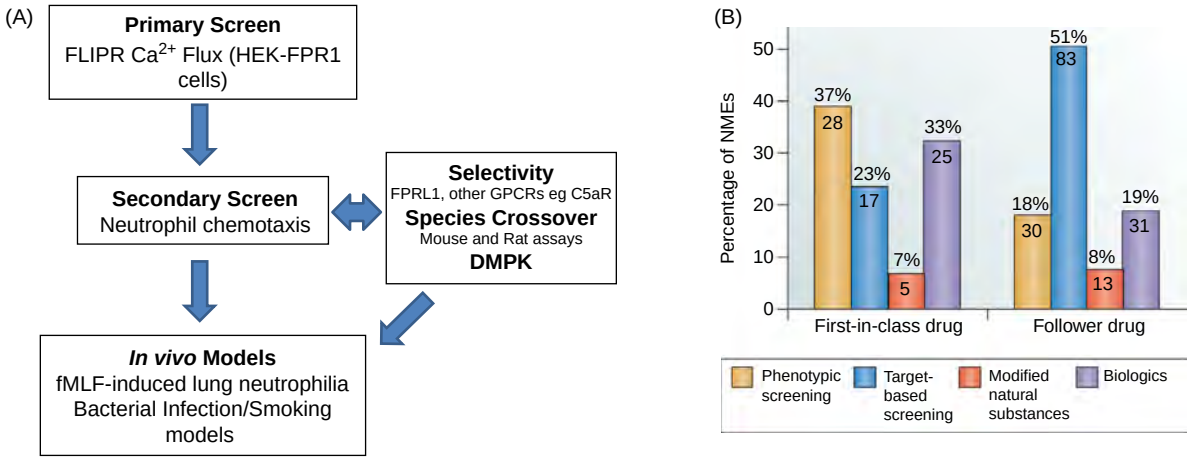


FIGURE 2.2 **Screening Cascades and Drug Discovery Approaches.** (A) A typical screening cascade for an early discovery program. In this instance the aim was to discover antagonists of the formyl peptide receptor, FPR1, as potential treatments for COPD. The receptor mediates a number of effects on a range of cell types but is perhaps best known as an activator of neutrophils. Hence the focus of the cascade is on neutrophil biology. As the project proceeded, a number of other assays were introduced to the cascade including neutrophil and macrophage superoxide generation and airway smooth muscle contraction. (B) The distribution of new drugs discovered between 1999 and 2008, according to the discovery strategy. *Reprinted by permission from Swinney & Anthony, 2011 [1].*

Phenotypic screening, on the other hand, starts from the premise that the assay readout has high disease relevance, and therefore active compounds are more likely to be clinically efficacious. The cell-based assays that are typically employed also have the advantage that "hits" by definition have cellular activity: in target based approaches using isolated proteins, this property often has to be built in later. The disadvantage of this approach is that the mechanism of action (MOA) of the compounds in the assay is usually unknown, so subsequent optimization of "hits" involves significant de-convolution activities. Nevertheless, phenotypic screens are becoming more popular in the drug industry, driven by the increased availability of novel higher throughput technologies and the success of this approach in identifying new "first in class" molecules [1] (Figure 2.2B). Examples of areas in which the approach has proven fruitful are neglected parasitic diseases such as human African trypanosomiasis and cystic fibrosis. In the former example, compounds can be screened for trypanocidal activity against the whole parasite [2]. In the case of CF, the well-understood genetic basis of the disease has encouraged screening for compounds that improve the functional activity of the defective protein (CFTR). This latter example is somewhat of a "halfway house" between a target-based screen and a phenotypic screen: the target is clearly known, but improved function of CFTR can result from compound interaction with CFTR itself or with proteins involved in its processing, trafficking and ion channel function. The recent approval of the CFTR potentiator, Kalydeco™ [3,4] is testament to the power of this approach.

III. *IN VITRO* ASSAYS

As outlined above, the initial phases of a target based screening cascade typically employ a range of *in vitro* assays. The exact system(s) used will be target and mechanism dependent. For example, a project targeting antagonists may use a binding assay as the primary screen, whereas one targeting agonists is more likely to use a functional assay. The following sections give a basic introduction to some of the more commonly used types of primary assays outlining their advantages and disadvantages. The measurements made in these assays that are typically reported to medicinal chemists, the properties of various different classes of drugs, and the principles underlying their analysis are also described.

A. Primary Assays

Glossary

A glossary of commonly used biochemical/pharmacological parameters is presented in Table 2.1 to assist the reader's understanding of the following sections.

TABLE 2.1 Glossary of Key Pharmacological/Biochemical Terms

Pharmacological/ biochemical term	Definition
K_A (pK_A)	Standard pharmacologic convention for the equilibrium dissociation constant of an agonist receptor complex with units of M. It is a measure of affinity. ($pK_A = -\log_{10}K_A$)
K_B (pK_B)	Convention for the equilibrium dissociation constant of an antagonist receptor complex determined in a functional assay. It has units of M and is a measure of affinity. ($pK_B = -\log_{10}K_B$)
K_D (pK_D)	Convention for the equilibrium dissociation constant of a ligand receptor complex measured in a binding assay. It has units of M. ($pK_D = -\log_{10}K_D$)
K_i (pK_i)	The K_B for an antagonist (or inhibitor) but measured in a binding study or enzyme assay. It has units of M. ($pK_i = -\log_{10}K_i$)
K_m	The Michaelis constant K_m is the substrate concentration at which the reaction rate is half of V_{max}. It has units of M and is measure of the substrate's affinity for the enzyme.
$[A]_{50}$ or EC_{50} (pA_{50} or pEC_{50})	The effective concentration of an agonist producing 50 percent maximal response to that particular drug (not necessarily 50 percent of the maximal response of the system). It has units of M and is a measure of agonist potency. (pA_{50} and $pEC_{50} = -\log_{10}A_{50}$ and $-\log_{10}EC_{50}$).
IC_{50} (pIC_{50})	The concentration (usually molar) of an inhibitor (receptor, enzyme antagonist) that blocks a given predefined stimulus by 50 percent. It is a measure of inhibitor potency but is an empirical value in that its magnitude can vary with the strength of the stimulus to be blocked. ($pIC_{50} = -\log_{10}IC_{50}$).
pA_2	The negative logarithm of the molar concentration of an antagonist that produces a 2-fold shift to the right of an agonist concentration-effect curve. It is a measure of antagonist potency.
Intrinsic Activity	A measure of agonist activity, it is the fractional response of an agonist (positive or inverse) relative to a standard full agonist. It is unit-less and ranges from 0 for antagonists to 1.0 for full agonists.
Efficacy (e or τ)	A measure of the capacity of an agonist to produce a physiological response. It is unit-less but can have both positive and negative values (for inverse agonists).
ED_{50}	The *in vivo* counterpart of EC_{50} referring to the dose (D) of agonist that produces 50 percent maximal effect

1. *Binding assays*

The aim of binding experiments is to determine the affinity (the strength with which a compound binds to the target site) of the compound for the biological target. They are the simplest and most robust assays. Today, binding assays are commonly run using recombinantly generated human protein or mammalian cell lines (such as human embryonic kidney 293 (HEK293) or Chinese hamster ovary (CHO) cells) engineered to express the human version of the target protein. Isolated protein, membrane preparations from cells, or whole cells can all be used to measure the affinity of test compounds. Isolated proteins are often employed for enzyme targets whereas membrane and cell preparations have been widely used in programs aimed at finding drugs that target GPCRs and ion channels. As it is impractical routinely to label test compounds, typically the measurements made are indirect, in that the ability of the test compound to inhibit binding of a standard labeled compound is assessed. Such assays depend of course on the availability of a suitably affine and selective labeled compound. Historically the label has been radioactive, but more recently fluorescently labeled compounds have also been employed. Increasingly, label free technologies such as that developed by Biacore [5], which use surface plasmon resonance to measure binding events as changes in molecular mass, are being used in drug discovery programs. This technique has the advantage of allowing real time measurements to be made. Thus, affinity, kinetics, and thermodynamics are easily studied (see Section IIIA (2)). It is also a very sensitive technique and therefore can detect the low affinity interactions that are typical of low molecular weight fragments (<250 Da). The disadvantages of this technique are that the development of successful protein target immobilization can take considerable time and effort, and its application is mainly with solubilised proteins like kinases rather than integral membrane receptors like GPCRs.

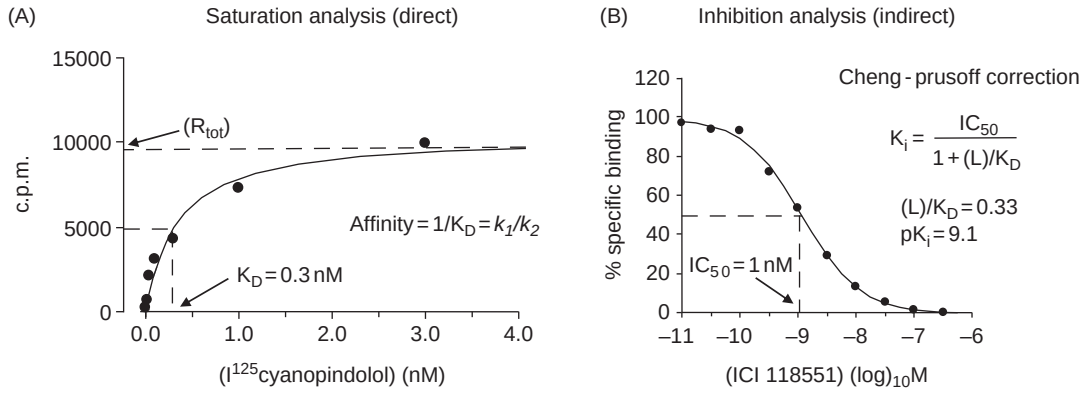


FIGURE 2.3 **Binding Assays: Direct and Indirect Measurements.** (A) A direct binding assay using I^{125} labeled cyanopindolol as a β_2-adrenoceptor ligand. The curve describes a rectangular hyperbola which saturates at high ligand concentration. The ligand dissociation constant (K_D) was estimated as 0.3 nM and is a measure of the ligand affinity. (B) A typical inhibition analysis using membranes expressing the human β_2-adrenoceptor and employing 0.1 nM I^{125} cyanopindolol as the labeled ligand. The displacing ligand, the selective β_2-adrenoceptor antagonist ICI 118551, produces complete inhibition of the specific binding yielding an IC_{50} of 1 nM. In this instance $[L]/K_D$ is <1.0 so the IC_{50} is a good estimate of the K_i as calculated by the Cheng-Prusoff equation. *Unpublished data.*

2. *Binding Studies: Principles and Analysis*

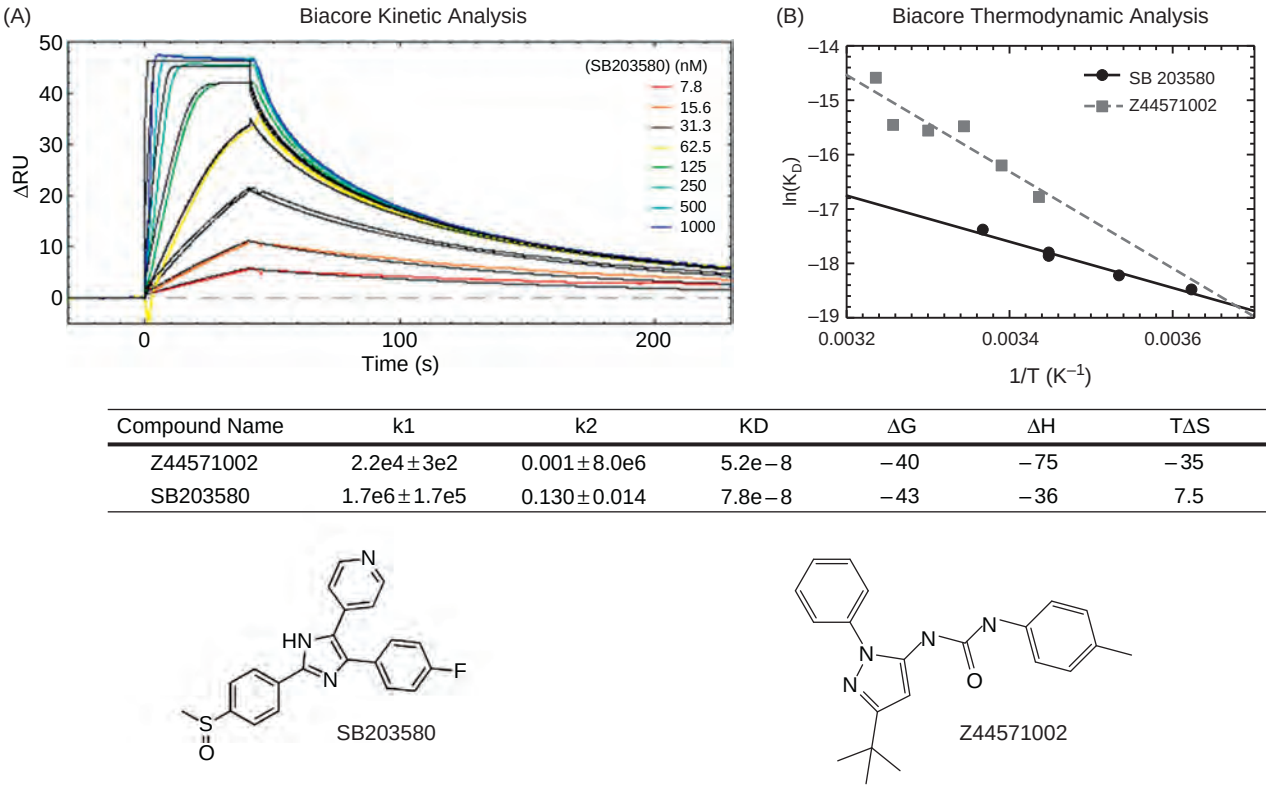
The simplest model of drug-receptor (or more generally protein) interaction is the Law of Mass Action in which the drug binds reversibly to the protein at a single site. Under such conditions, regardless of the assay employed, ligand (L) binding to its receptor (R) at equilibrium is described by the following equation:

$$[LR] = \frac{[R_{tot}][L]}{[L] + K_D} \tag{2.1}$$

where [LR] represents the concentration of ligand occupied receptors, $[R_{tot}]$ the total receptor pool and K_D is the ligand equilibrium dissociation constant (offset rate constant (k_2) divided by the onset rate constant (k_1)), a measure of the affinity of the ligand for its receptor. Eq. (2.1) describes a saturable curve with all the receptors being occupied at high ligand concentrations. K_D represents the concentration of ligand that occupies 50 percent of $[R_{tot}]$. Thus, in principle the direct binding of a labeled compound to the target can be simply measured and the K_D estimated from the midpoint of the saturation curve (Figure 2.3A). However, as it is clearly not practical to label all test compounds, the affinity of these is measured indirectly by assessing the displacement of a labeled ligand. In such experiments, the ability of the test compound to inhibit a single concentration (usually at approximately the K_D or below) of labeled ligand is measured. They typically yield a sigmoidal curve (when the drug concentration is expressed in log form) from which the IC_{50} (concentration of the test compound that produces 50 percent inhibition of the specific binding of the labeled compound) can be measured (see Figure 2.3B). Curve parameter estimates are usually derived from direct fitting of the experimental data to simple mathematical equations as described in Section IIIA (6). Assuming that the interaction between the labeled compound and the test compound is competitive, the dissociation constant (K_i) of the test compound can be calculated from the Cheng-Prusoff [6] equation:

$$K_i = \frac{IC_{50}}{1 + [L]/K_D} \tag{2.2}$$

Proof of the assumption that the interaction is competitive requires further experimentation, such as studying the inhibition by the test compound with different concentrations of the labeled ligand. As is evident from Eq. (2.2), at concentrations of L in excess of K_D the IC_{50} estimate will increase proportionately (i.e., higher concentrations of L will require higher concentrations of test compound to displace it). It is therefore imperative that when IC_{50} values are used to compare compound activities, that this is done under identical conditions (i.e., $[L]/K_D$ should be constant).



Compound Name	k1	k2	KD	ΔG	ΔH	TΔS
Z44571002	2.2e4±3e2	0.001±8.0e6	5.2e−8	−40	−75	−35
SB203580	1.7e6±1.7e5	0.130±0.014	7.8e−8	−43	−36	7.5

FIGURE 2.4 **Binding Assays: Kinetic and Thermodynamic Measurements.** Direct binding assay using the Biacore label free technique. The sensorgram (A) illustrates the time course of SB203580 binding to immobilized mitogen activated kinase p38α. The y-axis shows the mass change resulting from compound binding to p38α. At t = 0 a range of SB203580 concentrations were passed across the immobilized p38α to measure net association, and then at t = 50 s compound is replaced with buffer to initiate dissociation. The table shows the association (k_1 ($M^{-1}s^{-1}$)) and dissociation (k_2 (s^{-1})) rate constants as well as the equilibrium dissociation constants (K_D (M)) for two compounds. (B) Thermodynamic analysis of two p38α inhibitors using Biacore. Enthalpy and entropy components of binding derived from the Van't Hoff analysis are detailed in the attached table. ΔG, ΔH and TΔS values are in kJ/mol. *Unpublished data.*

Since IC_{50} values do not infer a particular MOA, they are routinely used to compare the activities of compounds in binding assays. Other modes of compound binding beyond simple competition are discussed in the sections below on enzymes and functional studies.

The kinetics of drug binding are also most easily studied in simple binding assays rather than functional assays where association and dissociation rates are more likely to be affected by diffusion barriers. Label free techniques such as Biacore are being increasingly utilized for this purpose as illustrated in Figure 2.4A for inhibitors of p38α. Interestingly the compounds in this example have similar affinities (52 and 78 nM), but markedly different kinetics (e.g., k_1 values 2.2×10^4 and 1.7×10^6 $M\,s^{-1}$). Slow kinetics and hence longer residence times are potentially advantageous compound characteristics [7,8], therefore such measurements are becoming increasingly important in drug discovery programs. An interesting example of a clinically used drug with unusual kinetics is the muscarinic antagonist, Tiotropium. It binds M_2 and M_3 receptor subtypes nonselectively (K_i values of 0.1 to 0.2 nM), but the compound has a much slower off rate (>10-fold) at the M_3 subtype, enough to make it a physiologically selective M_3 antagonist [9].

Finally, thermodynamic studies can be employed to determine the relative contributions of enthalpy and entropy to a compound's binding energy. The Gibbs free energy of binding (ΔG) is made up of enthalpic and entropic contributions and for reversible binding events can described as:

$$\Delta G = \Delta H - T\Delta S \tag{2.3}$$

where ΔH defines enthalpy and ΔS, entropy at temperature T.

The relationship between binding affinity (K_D) and temperature (T) is defined by the Van't Hoff equation:

$$\ln(K_D) = (\Delta H/R)(1/T) - \Delta S/R \tag{2.4}$$

TABLE 2.2 Enzyme Assay Techniques

Enzyme class	Product	Principle	Type	Detection examples
Protein kinases	Phosphoprotein	Labeled antibody	Homogeneous	AlphaLisa
	Phosphoprotein	Labeled antibody	Multi-step	ELISA
	ADP	Coupling enzyme	Multi-step	Luciferase (luminescent product)
Metabolic enzymes	Small molecule	Radioactivity	Multi-step	Radiometric
	• no chromophore	Radioactivity/Proximity	Homogeneous	Radiometric (SPA)
	• no antibody detection	Mass/charge	Homogeneous	LC-MS (RapidFire)
		Lipophilicity/charge	Multi-step	HPLC
		Coupling enzyme	Multi-step	Variety (secondary product)
Methyltransferases	Methylhistone	Labeled antibody	Homogeneous	AlphaLisa
	Methylhistone	Labeled antibody	Multi-step	ELISA
	H_2O_2	Coupling enzyme	Multi-step	Peroxidase (Light-based product)
	Formate	Coupling enzyme	Multi-step	Variety (secondary product)

where R is the Gas Constant. Thus by performing binding kinetics at different temperatures and then plotting ln (K_D) data as a function of 1/T, the relative enthalpy and entropy contributions to the compound's binding energy can be determined (Figure 2.4B). Such measurements allow chemistry to be steered towards optimization of either component or both. Generally in drug design, effort is focused on making optimal interactions (electrostatic, H-bonding etc.) with the target (enthalpic), whilst rigidifying the compound to reduce conformational (entropic) contributions (i.e., reduction in rotational freedom on binding).

3. *Enzyme Assays*

Enzymes are highly specific biological catalysts evolved to perform a broad range of biochemical transformations under physiological conditions. They operate in multiple locations (e.g., cytoplasmic, lysosomal, and extracellular) and under different spatial constraints (e.g., membrane bound, soluble, multimer). The reactions catalyzed by enzymes range in complexity from simple one-step chemical oxidations (e.g., alcohol dehydrogenase) to targeted, subtle protein modifications (e.g., kinases, methylases, etc.). The nature of an enzyme's activity is driven by precise substrate recognition at the active site (where the catalysis takes place) and through other regulatory sites.

Nearly all enzyme-targeted drugs are inhibitors, so most enzyme assays are designed to detect inhibitors by measuring the blockade of product production − substrate depletion is generally not used as the high starting background makes it technically more difficult. In order to develop an enzyme assay, one needs active and pure enzyme, substrates (e.g., protein, lipid, sugar, metabolite etc.), and a way to measure product formation and a good understanding of the optimal conditions for enzyme activity. Sufficient enzyme needs to be produced in a functionally active state at high purity (>95 percent) and in large amounts (>10 mg). To achieve this, enzymes are expressed recombinantly at high levels in various cell systems with tags (e.g., histidines) attached to aid purification. These tags are genetically encoded such that they are expressed at the C or N termini of the protein, distal from the active site and so less likely to affect the enzyme's activity. Once enzyme overexpression has been achieved, the enzyme is purified from the lysed cells by affinity chromatography using the attached tag (e.g., nickel column for his tag), followed by size exclusion (gel filtration) or pI (ion exchange). Enzyme production is not always straightforward and can require considerable optimization to reproducibly deliver a highly pure and active product. Substrates are usually commercially available, but if proteins, they may have to be made in the same way as the enzyme target. There are many different ways to measure product formation (Table 2.2), but the guiding principles are summarized below:

- Most enzyme assay readouts are now:
  - Light-based (e.g., fluorescence, luminescence, absorbance, fluorescence polarisation, HTRF, etc.)
  - Homogeneous (i.e., no separation steps, e.g., AlphaScreen)
  - Scalable (amenable to 384-well plates and HTS)
- Rarely:
  - ELISA (multiple wash steps)
  - Radiometric or HPLC (usually for metabolite, small molecule products)

The overall enzyme catalyzed reaction process is summarized below:

$$E + S \leftrightarrow ES \leftrightarrow EP \rightarrow E + P$$

where E (enzyme), S (substrate), ES (enzyme:substrate complexes), EP (enzyme:product complexes) and P (product). Enzyme reactions are generally studied under steady state conditions in which [S] is in excess of [E] and the reaction rate is linear—for most enzymes, [E] is nM and [S] is μM or mM. Initially upon addition of substrate to enzyme, only a very small percentage of total substrate is turned over and the rate is linear, but as more substrate is consumed it becomes rate-limiting and the enzyme velocity slows, usually when >10 percent substrate is used. The relationship between [S] and initial enzyme rate (v) was initially described by Briggs and Haldane [10]:

$$v = \frac{V_{max}[S]}{[S] + K_m} \tag{2.5}$$

Where V_{max} is the maximal rate (when $[S] \gg [E]$) and K_m is the Michaelis constant ([S] at which v is $V_{max/2}$).

In essence, enzyme assays are designed to optimally measure product formation and its inhibition by test compounds (I). It is important to show that the rate of product formation is proportional to [E] over the time course of the assay (i.e., steady state), so that a decrease in product rate by [I] relates directly to a reduction in active [E] due to inhibitor occupancy or indirect reduction in the number of substrate accessible active sites (i.e., formation of [EI]). Enzyme inhibitor mechanisms are discussed later on in Section IIIA (4). To minimize insolubility issues during dilution, compounds are usually dissolved in DMSO (anhydrous) to 10 mM and subsequently diluted in DMSO in half-logarithmic steps (usually 7) to produce a range of concentrations. These compound DMSO solutions are then diluted in assay media (large dilution, e.g., 25-fold) and then into the assay (small dilution, e.g., 4-fold) such that the final assay [DMSO] is tolerated (typically <1 percent (v/v)). Visual inspection for insolubility can be easily monitored during this process and fed back to the project team.

As with receptor functional and binding assays, compound potency is usually measured using an IC_{50} value (or pIC_{50} ($-\log_{10}IC_{50}$)) and typically determined from an eight point concentration inhibition curve using a four parameter logistic fit:

$$\%I = I_{min} + \frac{(I_{max} - I_{min})[I]^n}{[I]^n + [IC_{50}]^n} \tag{2.6}$$

where I_{max} is the maximal inhibition (usually ~100 percent), I_{min} the minimal inhibition (~0 percent) and n the slope of the curve.

Robert Copeland's book, *Evaluation of Enzyme Inhibitors in Drug Discovery* is highly recommended further reading [11].

4. *Types of Enzyme Inhibition and Their Analysis*

A. REVERSIBLE INHIBITORS

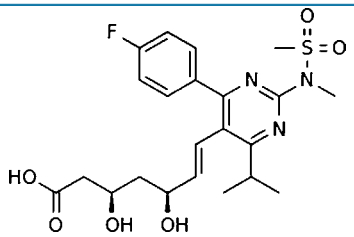
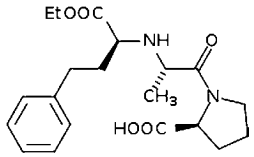
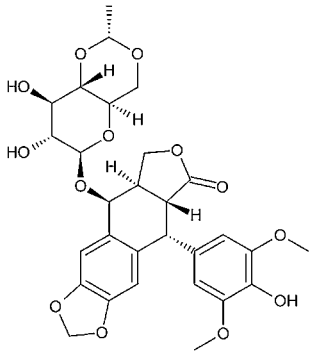
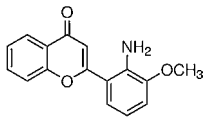
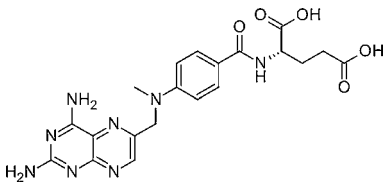
The majority of enzyme inhibitor drugs are reversible in that removal of the inhibitor (e.g., by dialysis) fully restores the enzymatic activity. Such inhibitors bind to their target enzyme through a combination of noncovalent interactions, such as hydrogen bonding or ionic, hydrophobic, and Van der Waals interactions, and don't generally undergo any chemical transformation while enzyme bound. Their behavior is described by the following equation:

$$E + I \rightleftharpoons EI$$

where E represents the active enzyme, I the reversible inhibitor, and EI the inactive inhibitor-bound enzyme. Examples of drugs that are reversible enzyme inhibitors and their mechanisms of action are shown in Table 2.3.

Testing for reversible inhibition relies on separation of the inhibitor from the inhibitor bound enzyme, which can be achieved using differences in enzyme and inhibitor mass (i.e., enzyme: >30 000 Da, inhibitor: ~400 Da) using a variety of techniques (e.g., dialysis, gel filtration, ultracentrifugation, etc.). By reducing free [I], EI complex dissociates leading to the recovery of enzyme activity. It is important during the pre-incubation of inhibitor with enzyme prior to reversibility that substrate is included to ensure the enzyme turns over and the inhibitor is exposed to all enzyme states during its catalytic cycle. Two common techniques to demonstrate reversible enzyme inhibition are jump dilution and immobilization. In the former, enzyme is incubated with inhibitor at $10 \times IC_{50}$ to give ~90 percent inhibition, and then, after sufficient time to allow EI formation, the mixture is rapidly ("jump") diluted 100-fold in assay buffer so that the final [I] is 10-fold below the IC_{50} such that if fully

TABLE 2.3 Examples of Reversible Enzyme Inhibitor Drugs

Example	Structure	Target
Crestor Competitive		HMG CoA Reductase
Enalapril Competitive		Angiotensin converting enzyme
Etoposide Noncompetitive		Topoisomerase II
PD 098059 Noncompetitive		MEK
Methotrexate Uncompetitive		Dihydrofolate reductase
Lithium Uncompetitive	Li^+	IMPase

reversible only ~9 percent inhibition would be expected. In contrast, if the inhibitor is irreversible, the jump dilution would have little effect on the enzyme-inhibitor complex and the enzyme's activity would still be inhibited by ~90 percent. Technical assay conditions may need to be fine-tuned to ensure slow dissociation is examined and that sufficient active enzyme is present post "jump" dilution. The immobilization technique depends on the ability to irreversibly immobilize the enzyme in a 96-well plate such that sufficient activity is retained for inhibitor studies. Immobilized enzyme is incubated with inhibitor and substrate to determine the pre-wash IC_{50}, followed by washing to remove substrate and inhibitor, then re-measurement of enzyme activity with substrate (post-wash). Figure 2.5 illustrates this, showing no change in IC_{50} with washing for a mechanism-based, 2-thioxanthine, irreversible myeloperoxidase inhibitor [12], but complete loss of inhibition on washing with a reversible inhibitor [13].

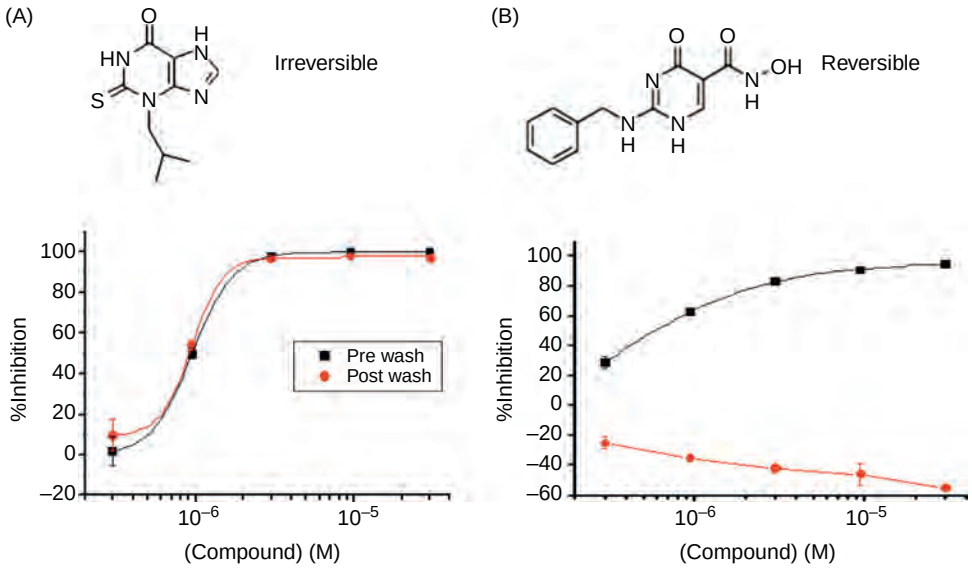


FIGURE 2.5 **Reversibility of Enzyme Inhibition.** MPO inhibition by an irreversible (A) and a reversible (B) inhibitor using plate-based immobilized enzyme and removal by washing. Inhibition curves are shown for both types of compound before (black circles) and after (red circles) inhibitor removal. *Unpublished data*.

B. IRREVERSIBLE INHIBITORS

In some cases, enzymes can be irreversibly inhibited through formation of a covalent bond between the enzyme and the inhibitor. Such inhibition due to the inherent chemical reactivity of an inhibitor is usually too nonspecific and promiscuous to be a useful drug mechanism. More commonly, irreversible drugs are specifically recognized by their target enzyme as "pseudo-substrates" and converted to reactive products that covalently inactivate the enzyme. This mechanism-based irreversible inhibition provides target selectivity by virtue of structural recognition of the inhibitor by the enzyme and the specific chemistry of the enzyme's active site.

$$E + I \rightleftharpoons EI \rightarrow E - I$$

Irreversible inhibition is eventually "reversed" over days *in vivo* after inhibitor administration has stopped, by the de novo synthesis of new enzyme to replace that inactivated by the inhibitor. Examples of drugs that are irreversible enzyme inhibitors are shown in Table 2.4.

The advantages of an irreversible drug are that with time it will inhibit all the enzyme such that high potencies and ligand efficiencies can easily be obtained, leading to long duration of action *in vivo*. Potential disadvantages are a higher risk of poor specificity, inability to quickly reverse *in vivo* effects if required, a need for a more complex set of *in vitro* assays to drive SAR (i.e., IC_{50} isn't sufficient on its own), and reactive inhibitor intermediates formed during enzyme inactivation have the potential to react with other proteins to form immunogenic adducts.

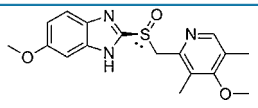
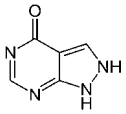
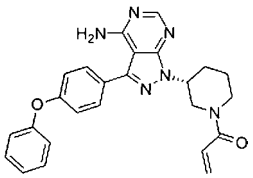
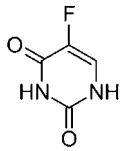
C. COMPETITIVE INHIBITORS

Inhibitors can reversibly bind to the free form of the enzyme, to substrate-bound enzyme catalytic forms, or to both, such that an inhibitor's potency can have different relationships to [S]. This is summarized by the enzyme turnover reaction scheme in the presence and absence of reversible inhibitors in Figure 2.6A. If inhibitor and substrate binding are mutually exclusive (i.e., inhibitor and substrate cannot bind to the enzyme at the same time), the inhibitor is competitive with respect to that substrate. Hallmark features of a competitive inhibitor are an increase in K_m, but no effect on the V_{max} as illustrated in Figure 2.6B. Competitive inhibitors generally bind at the enzyme's active site and compete with substrate for occupancy.

D. NONCOMPETITIVE INHIBITORS

A noncompetitive inhibitor exhibits affinity for both the free enzyme (E) and the enzyme-substrate complexes and thus is defined by two equilibrium constants, one for EI (K_i) and one for the ESI complex (αK_i). α describes the relative affinity of I for E and ES. For example, when $\alpha = 1$, then I has equal affinity for E and ES. Key

TABLE 2.4 Examples of Irreversible Enzyme Inhibitor Drugs

Drug	Structure	Target
Nexium		H^+/K^+ ATPase in gastric parietal cells
Allopurinol		Xanthine oxidase
Ibrutinib		Bruton's tyrosine kinase
5-Fluorouracil		Thymidylate synthase

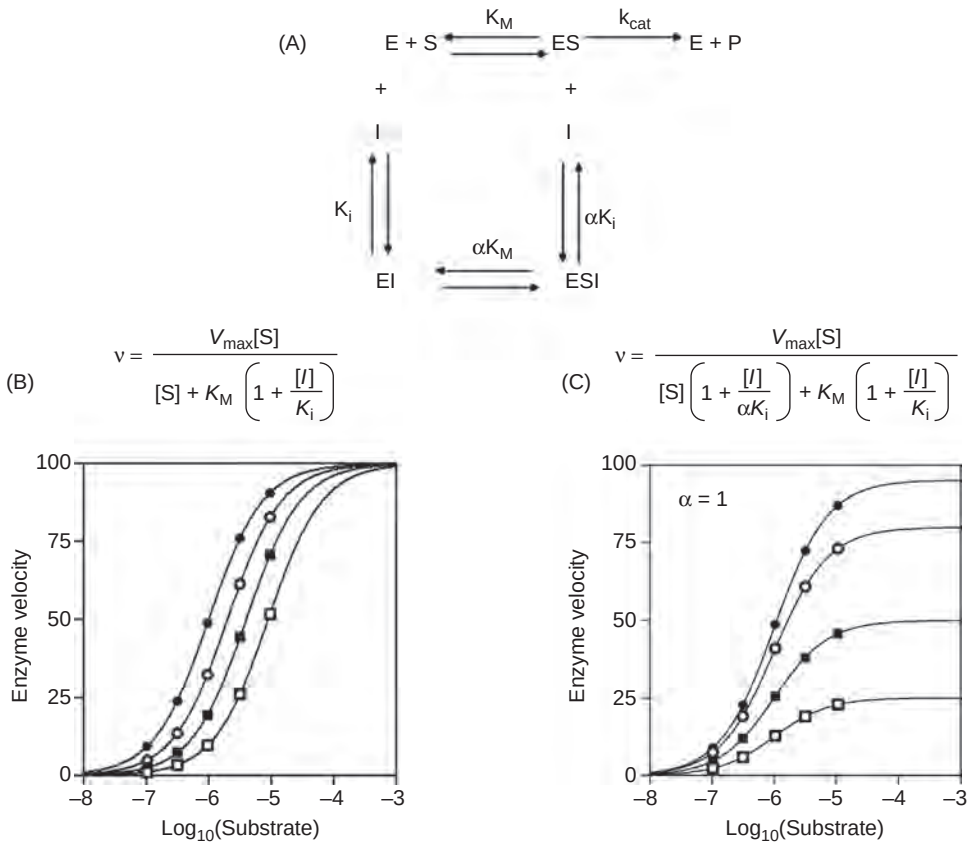


FIGURE 2.6 **Competitive and Noncompetitive Enzyme Inhibition.** Equations and graphs illustrating a general enzyme reaction scheme for reversible inhibitors (A), the substrate dependency of the steady state velocity for an enzyme in the presence of a range of competitive (B), and noncompetitive (C) inhibitor concentrations.

$$v = \frac{V_{max}[S]}{[S]\left(1 + \frac{[I]}{\alpha K_i}\right) + K_M}$$

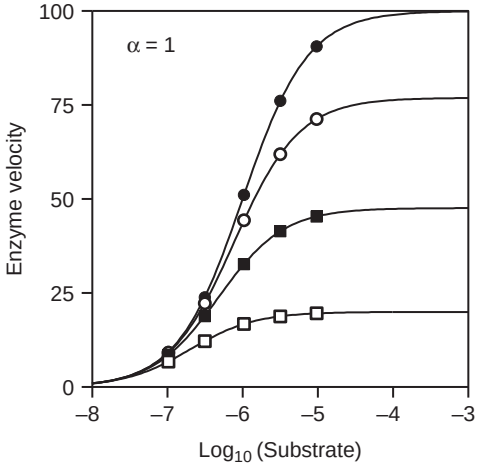


FIGURE 2.7 **Uncompetitive Enzyme Inhibition.** Equation and graph illustrating the substrate dependency of the steady state velocity for an enzyme in the presence of a range of uncompetitive inhibitor concentrations.

features of a noncompetitive inhibitor are no effect on substrate K_m and a decrease in V_{max} as illustrated in Figure 2.6C. Noncompetitive inhibitors tend to bind to the enzyme at sites distinct from the active site and exert their effects allosterically.

E. UNCOMPETITIVE INHIBITORS

Uncompetitive inhibitors only recognize and interact with ES and subsequent downstream catalytic species with no binding to free enzyme. Thus to exhibit enzyme binding, uncompetitive inhibitors require formation of ES and inhibition of enzyme activity is characterized by a decrease in both substrate K_m and V_{max} (see Figure 2.7). Since uncompetitive inhibitors only block processes beyond ES formation, one might expect only V_{max} to be suppressed with no effect on K_m, but as the inhibitor binds to and stabilizes the ES complex, it makes it more difficult for S to dissociate or be converted to product, increasing enzyme affinity for S and so reducing substrate K_m. This mode of action is attractive for drug design as the inhibitors bind to the enzyme target only when the target is active and substrate present. Uncompetitive inhibitors decrease substrate K_m and V_{max} as well as exhibiting higher inhibition with increasing [S] as illustrated in Figure 2.7. From the equations and graphs describing the three modes of enzyme inhibition (Figures 2.6 and 2.7), it can be seen that competitive (I only binds E with affinity K_i) and uncompetitive (I only binds ES with affinity αK_i) are special cases of noncompetitive inhibition (I binds both E and ES with affinities K_i and αK_i respectively).

5. Functional Assays

Binding assays as described above (Sections IIIA(1) and (2)) provide information on the affinity of compounds, but they do not generally indicate if they are agonists. That is, they do not provide information on their ability to elicit functional responses (efficacy). Such information is crucial for projects that are aiming to identify agonists as therapeutic agents, but functional assays are also important for inhibitor based projects as they allow a) compounds with unwanted agonism to be identified and b) confirmation that activity detected in binding assays translates into inhibition of functional readouts.

Traditionally pharmacologists used pieces of isolated tissue (typically smooth muscle preparations from laboratory animals) to generate functional data, but nowadays cellular systems (engineered or native cell lines) are routinely employed. The former assays have the advantage that the readout (e.g., contraction or relaxation) was often physiologically relevant, but the disadvantages of very low throughput and potential issues with species cross-over (see Section IIID (3)). Cellular assays have the advantage of increased throughput, allowing very large numbers of compounds to be screened. They also routinely employ the human ortholog of the protein (either endogenously expressed or engineered into a cell type that does not usually express it) ensuring that compounds chosen for progression have good activity at the clinical target. The disadvantages are that the readouts used in these assays are usually "upstream" of the physiologically relevant response and may, for various reasons, not always mirror the latter. Typical measurements include the generation of 2nd messengers such as cyclic

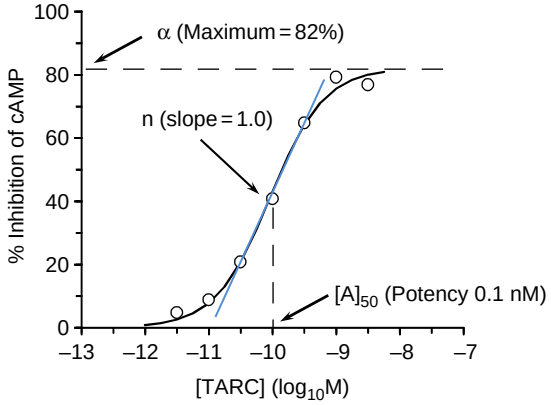


FIGURE 2.8 Agonist Concentration-Effect Curves. An agonist concentration-effect curve highlighting the parameters that define it: α (the maximum effect); $[A_{50}]$ (potency), the concentration of agonist that produces 50 percent of α; and n, a measure of the slope. $[A_{50}]$ values are often quoted as pA_{50} ($-\log_{10}[A_{50}]$). In this example, the $[A_{50}]$ is 10^{-10} M and the pA_{50} is 10.0. Data shows TARC (CCL17) induced inhibitions of cAMP levels in CHO cells stably transfected with human CCR4. *Unpublished data.*

adenosine monophosphate (cAMP), Ca^{2+}, or inositol phosphates, and beta arrestin movement for GPCRs, ion fluxes or membrane potential changes for ion channels, and gene transcription for NHRs. The technologies for measuring such readouts are constantly evolving but include FLIPR®, which uses a range of fluorescent dyes to measure changes in intracellular calcium concentration or membrane potential; antibody based technologies such as AlphaLISA, which can be used to detect a range of substances including cAMP; and reporter gene assays, which use fluorescent or luminescent proteins under the control of target gene promoters to assess drug-induced gene transcription. Again a major advantage of these technologies is their suitability for medium- to high-throughput screening. The nature of the readout can, however, generate difficulties in data analysis and interpretation. For example, changes in intracellular Ca^{2+} levels in response to agonists in FLIPR® assays are often transient in nature, which can result in failure to reach true equilibrium when potent competitive antagonists are studied. In such hemi-equilibrium cases, the antagonists appear to be insurmountable, and this has the potential to introduce errors into affinity estimations (see [14] and Section IIIA 6D).

6. *Functional Studies and Their Analysis*

By definition, functional studies involve analyzing agonist responses, either alone or in the presence of antagonists (or inhibitors). Central to these analyses is the generation of agonist concentration-effect curves. Their defining properties are described below. Subsequent sections describe the different classes of agonists and outline how their interaction with antagonists is analyzed to yield antagonist affinity estimates and MOA information.

A. AGONIST CONCENTRATION-EFFECT (E/[A]) CURVES

Agonist concentration-effect curves are typically sigmoidal (s-shaped) when plotted in semi-logarithmic form ($E/\log_{10}[A]$) and are described by four parameters: 1) a lower asymptote (β), which represents the basal state of the system; 2) an upper asymptote (α), which represents the maximum effect that the agonist produces in the system; 3) a location or potency ($[A_{50}]$ or EC_{50}), which represents the concentration of agonist that produces an effect equal to 50 percent of α-β and 4) a slope parameter (n), which is a measure of the gradient of the curve at the $[A_{50}]$ level. A number of computational programs are available that allow estimates of these parameters to be made by fitting experimental E/[A] curve data to the following form of the Hill equation (a saturable function that adequately describes curves of varying gradients):

$$E = \beta + \frac{(\alpha - \beta)[A]^n}{[A]^n + [A_{50}]^n} \tag{2.7}$$

In practice, $\beta = 0$ in the majority of cases. That is, the basal effect level is ascribed a value of zero, and therefore most E/[A] curve data can be adequately described by a 3-parameter Hill equation as illustrated in Figure 2.8. It is the analysis of how the three curve parameters (α), $[A]_{50}$, and n are affected by experimental manipulation that allows drug-receptor interactions to be quantitatively described in terms of affinity (binding) and efficacy (response-eliciting capacity).

B. FULL AGONISTS, PARTIAL AGONISTS, AND INVERSE AGONISTS

The first step in agonist action is the formation of a reversible agonist-receptor (AR) complex, a process that is generally assumed to be governed by the Law of Mass Action. Accordingly, the equilibrium concentration of

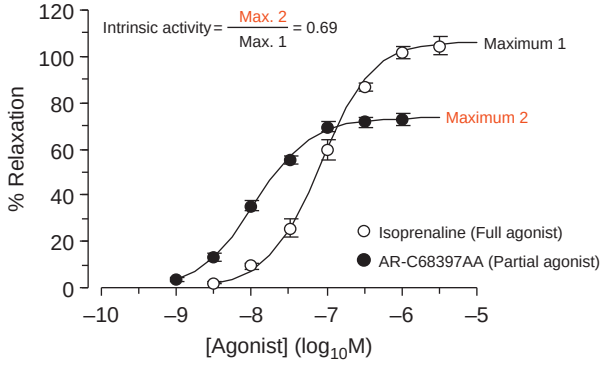


FIGURE 2.9 Full Agonists, Partial Agonists and Intrinsic Activity. Experimental data showing the β_2-adrenoceptor mediated smooth muscle relaxing activity of the reference full agonist isoprenaline and the partial agonist AR-C68397AA in guinea pig isolated tracheal rings. The intrinsic activity of AR-C68397AA was 0.69. *Unpublished data.*

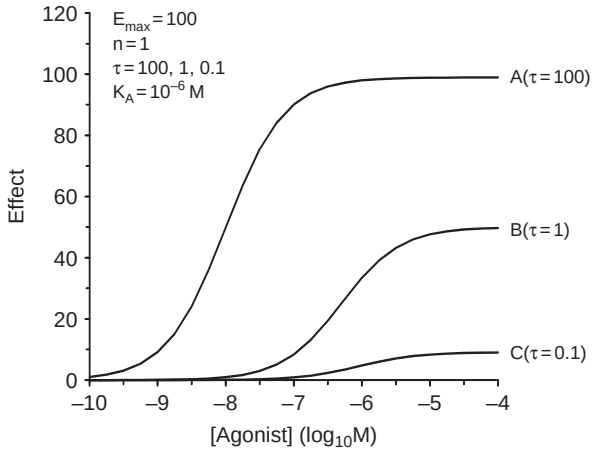


FIGURE 2.10 System Dependence of Drug Effects. Simulated curves showing how the curve parameters of an agonist change in systems with varying receptor expression. The receptor expression range is 1000-fold from curve A to curve C. In system A, the drug exhibits full agonism (IA = 1.0 and high efficacy ($\tau = 100$)); in system B, it shows partial agonism (IA = 0.5 and low efficacy ($\tau = 1$)); and in system C, it shows very weak partial agonism (IA = 0.09 and very low efficacy ($\tau = 0.1$)). In system C the drug effectively behaves as an antagonist as this level of IA is difficult to detect in most assay systems. The Operational Model of Agonism [15] was used to simulate the data. In this model, τ is a measure of the efficacy of the agonist and incorporates both drug (intrinsic efficacy) and system (receptor number ($[R_{tot}]$) and coupling efficiency) parameters.

agonist occupied receptors is a rectangular hyperbolic function (a special case of the Hill function where n = 1) of the agonist concentration (identical to Eq. (2.1)). This curve is defined by a maximal value of $[R_{tot}]$, the total receptor concentration, and a midpoint value of K_A, the agonist dissociation constant. K_A determines how well the agonist binds; that is, it is a measure of the affinity of the agonist for its receptors. Agonist occupancy is subsequently amplified into functional effect by the biochemical/biophysical machinery of the cell/tissue, and this is what is measured experimentally in the form of an E/[A] curve. The efficiency of this transduction process can vary between agonists and across systems (i.e., it is both drug and tissue dependent). Agonist efficacy is a measure of the efficiency of the transduction process. Full agonists have high efficacies and therefore can elicit the maximum effect (E_{max}) that the test system is capable of generating. Partial agonists by contrast have low efficacy and cannot elicit a maximum response (Figure 2.9).

Measuring the efficacy (and affinity) of full agonists is not straightforward because their occupancy is efficiently converted into effect and thus the $[A_{50}]$ is much lower than the K_A. An experimental manipulation that decreases the efficacy of the agonist to a level where it behaves as a partial agonist (where $[A_{50}]$ approximates K_A) is therefore required. Irreversible antagonists have been used for this purpose as they covalently modify receptors, thereby decreasing $[R_{tot}]$ (see Section IIIA 6D). An important consequence of efficacy being both a drug- and system-dependent parameter is that an agonist can demonstrate different behaviors in different systems. Thus, a drug that exhibits partial agonism in one system may be a full agonist in another (with higher $[R_{tot}]$ or more efficient transduction machinery) or effectively an antagonist in yet another (with lower $[R_{tot}]$ or less efficient transduction machinery) (Figure 2.10).

As alluded to above, it can be difficult to measure the affinity and efficacy of agonists, and typically the information reported to the medicinal chemist is the potency ($[A_{50}]$ or, more often, pA_{50} ($-\log_{10}[A_{50}]$)) and the intrinsic activity (IA) of the compound. The latter is a measure of the maximal activity of the test compound relative to a reference full agonist [16]. If the test agonist produces a maximum response less than the reference agonist, then the IA will be <1.0. For example, in Figure 2.9 AR-C68397AA produces a maximum effect that is 69 percent of the reference full agonist isoprenaline, and thus is a partial agonist with an IA of 0.69. Such compounds are

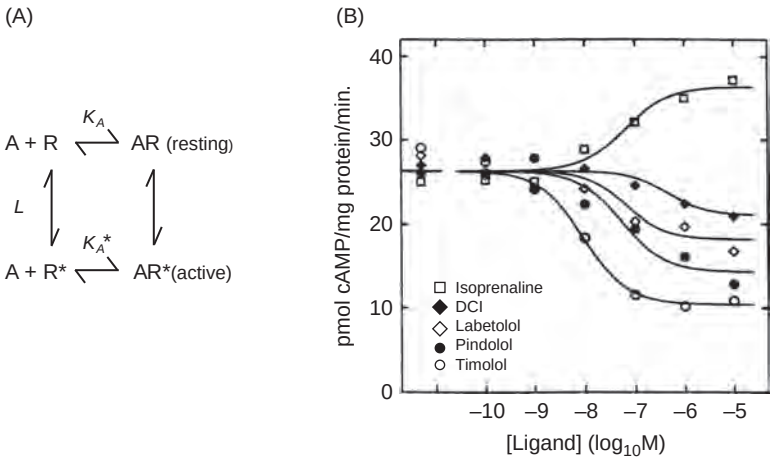


FIGURE 2.11 **Inverse Agonism.** (A) A simple two-state model in which the receptor exists in R (resting) or R* (active) states. In the absence of agonist (A), the distribution of the two states is governed by the equilibrium constant *L*. The agonist has affinities for the two states governed by the dissociation equilibrium constants, K_A and K_A*. (B) Production of cAMP in membranes of Sf9 cells expressing the human β_2-adrenoceptor. Isoprenaline exhibits positive agonism but all the other ligands tested show varying degrees of inverse agonism (i.e., they decrease the basal level of cAMP production). *Panel (B) reproduced with permission from Chidiac et al., 1994 [20].*

useful to the medicinal chemist as they help direct efforts to optimize the efficacy of compounds for therapeutic benefit. For example, identification of partial agonists were important staging posts in the development of the antagonists propranolol and cimetidine [17,18]. Finally, it is important to emphasize that the IA scale does not discriminate between full agonists (i.e., all full agonists will have an IA of 1.0 but they may have different efficacies).

Until relatively recently, agonist efficacy was considered only as a positive scalar associated with increased receptor activity. This dogma was challenged by the discovery of the phenomenon of constitutive receptor activation and compounds that showed inverse agonism [19,20] That is, they decreased the level of constitutive activation, demonstrating negative efficacy (Figure 2.11). The most likely mechanism for inverse agonism is that such compounds have a selectively higher affinity for the inactive state of the receptor and thereby uncouple spontaneously coupled (active) receptor species [21,22]. The existence of multiple receptor states also offers an explanation for the phenomenon of "signaling bias or agonist trafficking," whereby one agonist may direct signaling to a particular cascade while another agonist may not [23]. To date, inverse agonism has largely been a property detected in genetically engineered cells systems where receptors (or modified receptors) can be expressed at supra-physiological levels. Many of the compounds that exhibit inverse agonism in such systems behave as competitive (neutral) antagonists with zero efficacy in more physiologically relevant assays. As such, the therapeutic relevance of inverse agonism remains largely unknown, but this now well-documented phenomenon has changed the way pharmacologists view drug-receptor interactions as well as resulting in the re-classification of drugs that were formerly thought to be competitive antagonists (e.g., Ranitidine and Propranolol). Importantly, designing compounds with inverse agonist properties and/or signaling bias offers the medicinal chemist further opportunities in tailoring compounds to address unmet clinical needs. The advancement of a biased μ-opioid receptor agonist (TRV130) into clinical testing as an analgesic with low side-effect potential serves as an example of how such new concepts of receptor function are being exploited [24].

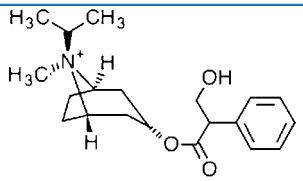
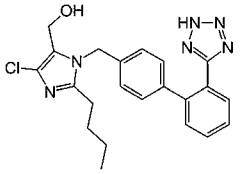
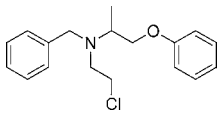
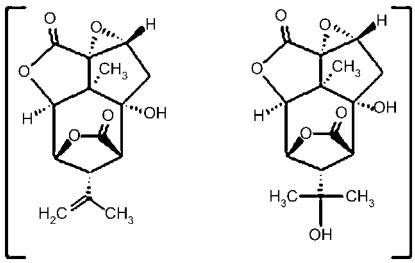
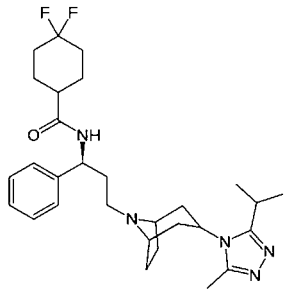
C. OPTIMIZING AGONISTS

As discussed above, agonists bind to and activate receptors. The optimization of agonist properties therefore relies on designing compounds with both good affinity and appropriate efficacy. Affinity can be measured in ligand binding assays, but functional assays are required to provide estimates of IA or efficacy. In most cases, the aim of agonist based projects is to identify high potency, high efficacy agonists so that the drug dose ultimately administered will be small and the effect large. In some instances however, partial agonists can have therapeutic advantages. Thus, if the desirable therapeutic effect is observed in a tissue with high receptor number/coupling but an undesirable side-effect is mediated in a tissue with low receptor number/coupling, a partial agonist of appropriate efficacy could produce agonism in the former but be "silent" in the latter (compare curves A and C in Figure 2.10). Finally, as described above, designing compounds that signal selectively through a particular pathway may provide further levels of therapeutic control.

D. ANALYSIS OF ANTAGONISTS

As with enzyme inhibitors, several different classes of antagonists with distinct mechanisms of action including irreversible competitive, reversible competitive, noncompetitive, and allosteric have been identified

TABLE 2.5 Examples of Various Classes of Receptor Antagonists

Example	Structure	Target
Ipratropium Competitive		M_3-receptor
Losartan Competitive		AT_1-receptor
Phenoxybenzamine Irreversible		α-adrenoceptors (nonselective)
Picrotoxin Noncompetitive		$GABA_A$-receptor
Maraviroc Allosteric		CCR5

(Table 2.5). Their blockade of agonist-induced effects can be surmountable (rightward displacement of the E/[A] curve with no depression of the maximum (α)) or insurmountable (depression of the maximal agonist response (α)). It is important to realize that the profile of antagonism observed can show system dependence; that is, an antagonist can exhibit surmountable activity in one assay system and insurmountable activity in another, despite having the same mechanism of action (see Figure 2.14A and B). A common example of this phenomenon is the behavior of high affinity competitive antagonists in FLIPR® assays in which the changes in intracellular calcium levels measured are typically transient in nature. In these circumstances, true equilibrium is not reached as the agonist does not have sufficient time to access antagonist bound receptors resulting in apparent nonsurmountable antagonism [25,26]. This contrasts with the behavior of such antagonists in systems where agonist responses are sustained (e.g., in many isolated tissue systems), true equilibrium is reached and the antagonism is surmountable.

The interaction of an antagonist with its receptors is described by a single parameter, affinity which equates to potency (unlike agonists where potency is dependent on both affinity and efficacy). By definition, antagonists have an IA = 0 in functional assays, in which their affinity is measured by studying their receptor interaction with an agonist. The affinity of antagonists can also be measured in binding assays (see Section IIIA (1)),

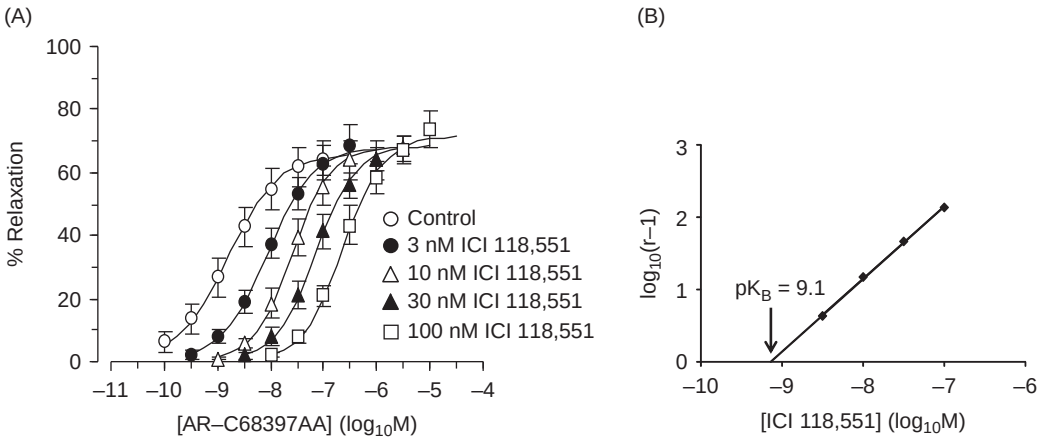


FIGURE 2.12 **Competitive Antagonism in Functional Assays.** Antagonism of the AR-C68397AA β_2-adrenoceptor mediated relaxation of rabbit isolated saphenous vein by the competitive antagonist ICI 118,551 (A). Note the concentration-dependent parallel rightward displacement of the control curves. (B) Illustrates the displacements (r values) in Schild plot form derived from one of the 5 experiments that make up in the data in (A). The plot has a slope of unity and the intercept on the x-axis yields an estimate of 9.1 for the pK_B ($-\log_{10}K_B$). *Unpublished data.*

although if such systems are used it is essential to confirm lack of efficacy by subsequent testing in functional assays. The following sections discuss the properties and analysis of the various classes of antagonists.

COMPETITIVE ANTAGONISTS Reversible competitive antagonists are probably the most important class of antagonists, and a large number of clinically used drugs fall into this class. As outlined above for agonists, the first step in the action of these drugs is the formation of a reversible, relatively short-lasting, drug-receptor complex governed by the Law of Mass Action. In this mode of antagonism, the binding of the agonist and antagonist is mutually exclusive. The presence of the antagonist therefore decreases the probability that an agonist-receptor interaction will occur. To achieve the same degree of agonist occupancy—and therefore the same effect—in the presence of the antagonist as in its absence, the agonist concentration must be increased. The factor (r) by which it must be increased depends on both the concentration of antagonist ([B]) used and on how well it binds (K_B). This relationship, which was first described by Schild [27], is shown below:

$$r - 1 = [B]^n / K_B \tag{2.8}$$

where $r = [A_{50}]/[A_{50}]^c$ (location parameter of the E/[A] curve in the presence of the antagonist/location parameter of the E/[A] curve in the absence (c = control) of the antagonist), K_B is the antagonist equilibrium dissociation constant, and n represents the stoichiometry of the interaction between the antagonist and the receptors (n = 1, when one molecule of antagonist binds to one receptor molecule).

Experimentally, a K_B is estimated by studying the interaction of an agonist and antagonist over a wide range of antagonist concentrations (the wider, the better). This is necessary because drugs which are not reversible competitive antagonists may appear to be so within a narrow range of concentrations. If the antagonist is truly competitive, it should produce parallel rightward displacement (i.e., no change in midpoint slope (n) occurs) of the E/log[A] curves with no change in the maximal response (α) (see Figure 2.12A). The analysis involves fitting experimentally derived values of r at different concentrations of antagonist to the following form of Eq. (2.8) [28]. See Figure 2.12B.

$$\log_{10}(r - 1) = n\log_{10}[B] - \log_{10}K_B \tag{2.9}$$

Consistency of the data with Eq. (2.9) is judged by the finding of a linear plot with a slope (n) of 1.0. Under these conditions, the intercept on the x-axis ($\log_{10}[B]$) gives an estimate of K_B. When n is significantly different from 1, the intercept gives an estimate of pA_2 ($-\log_{10}K_B/n$). The pA_2 is an empirical estimate of antagonist affinity and equates to the negative logarithm of the concentration of antagonist that produces a two-fold rightward shift (r = 2) of the control E/[A] curve. Nonlinearity and slopes other than unity can result from many causes. For example, a slope of greater than 1 may indicate incomplete antagonist equilibration or removal of the antagonist from the biophase (receptor compartment). A slope that is significantly less than 1 may indicate removal of the agonist by a saturable uptake process, or it may result from the interaction of the agonist with more than one receptor. In the latter case, the Schild plot may be nonlinear with a clear inflexion. All of these potential complicating factors have been described in detail previously by Kenakin [29].

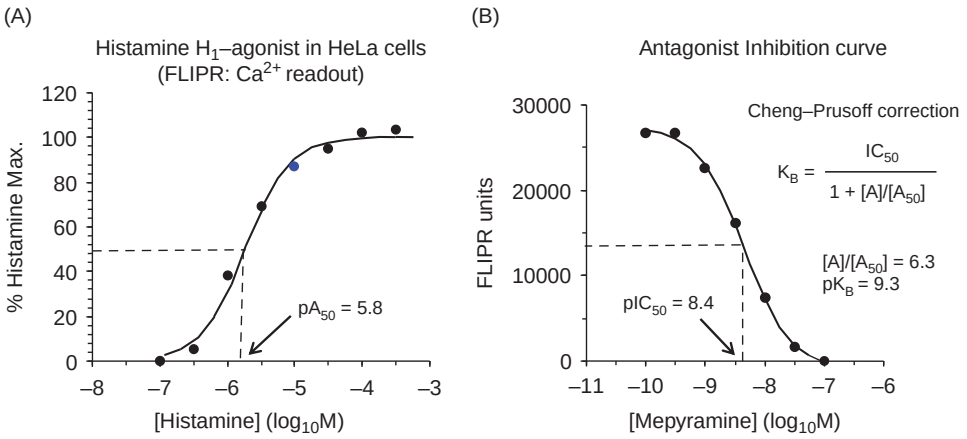


FIGURE 2.13 **Cheng-Prusoff Analysis of Antagonism in Functional Assays.** (A) Histamine concentration effect curve generated in HeLa cells which endogenously express the human H_1-receptor. 10 μM histamine (blue symbol) was chosen as the concentration of agonist to be used to assess the inhibitory effects of the H_1-receptor antagonist mepyramine. (B) Mepyramine completely inhibited the Ca^{2+} induced histamine response and yielded a pIC_{50} value of 8.4. As $[A]/[A]_{50}$ was significantly greater than 1, the affinity (pK_B) estimate (9.3) is considerably greater than the pIC_{50}. *Unpublished data.*

Although Schild type analysis is the most robust method of assessing antagonist behavior in functional assays, the needs of modern high-throughput drug discovery programs dictate that it is used sparingly to assess the mechanism of action for priority compounds. Routine screening of antagonist properties will more likely be assessed by doing a simpler functional Cheng-Prusoff type experiment (Figure 2.13) in which the effects of several concentrations of the test compound on the response to a single concentration of agonist are studied. The experimental data can then be fitted to the following equation [30]:

$$K_B = \frac{IC_{50}}{(2+([A]/[A_{50}])^n)^{1/n} - 1} \tag{2.10}$$

As was outlined above for binding studies, the estimated IC_{50} is dependent on the concentration of ligand employed. In this case, the concentration of agonist ([A]) relative to its $[A_{50}]$ dictates the IC_{50} (and hence the estimated K_B). Practically, the experimenter usually employs a concentration of agonist that is as close to the $[A_{50}]$ as possible so that the IC_{50} is a good estimate of the K_B. The shape of the agonist E/[A] curve is also important as evidenced by the inclusion of the slope parameter (n) in this form of the Cheng-Prusoff equation. When $n = 1$ the equation simplifies to a form equivalent to Eq. (2.2). Such analysis, although higher throughput, does not discriminate different modes of action of test compounds. For example, it will not differentiate competitive from noncompetitive compounds. Without additional proof that the interaction of agonist and antagonist is competitive, it is more appropriate to use the measured IC_{50} as a measurement of antagonist potency rather than calculating a K_B. As is the case with enzyme and binding assays the factor $[A]/[A]_{50}$ should be kept constant so that IC_{50} values of different compounds can be easily compared.

IRREVERSIBLE, NONCOMPETITIVE, AND ALLOSTERIC ANTAGONISTS Several other forms of antagonists have been identified and will be discussed briefly. Irreversible antagonists form covalent bonds with the receptor protein and thus prevent binding of agonists. They therefore effectively decrease the receptor pool and, by so doing, decrease the response eliciting capacity (efficacy) of the agonist, as this is dependent on $[R_{tot}]$. As equilibrium is not attained, the antagonist affinity cannot be measured, but such agents (e.g., Phenoxybenzamine) form the basis of the receptor inactivation method developed by Furchgott [31] for estimating agonist affinities and efficacies. Practically, however, the reliance of this method on the availability of suitable alkylating agents excludes its use in most receptor systems.

Noncompetitive antagonists bind to receptors and make them functionally inoperative either by preclusion of agonist binding (due to negligible dissociation of the antagonist during the response-gathering phase of the experiment) or through some other biochemical mechanism that obviates agonist effect (e.g., pore blockade of ion channels). Under these circumstances, increasing the agonist concentration cannot overcome the effect of the antagonist, and a distinctive feature of noncompetitive antagonists is the depressive effect they have on the

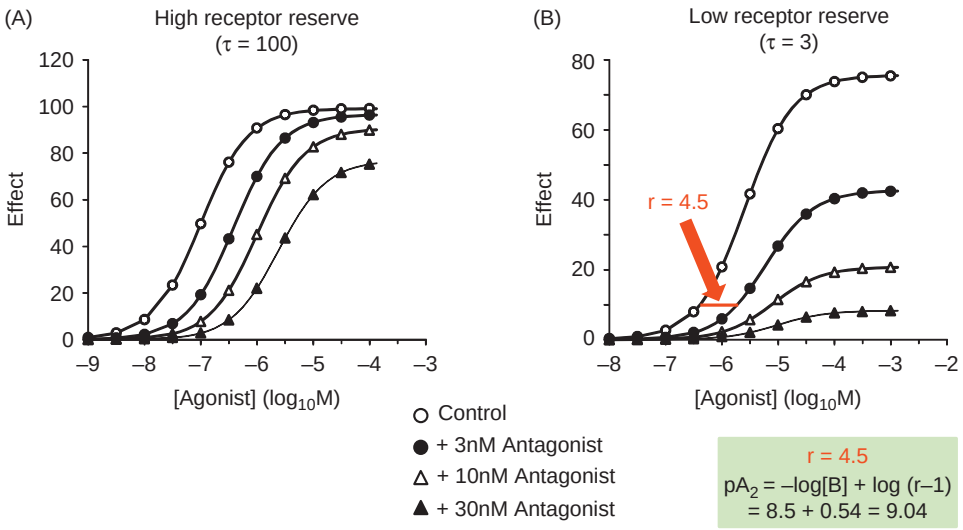


FIGURE 2.14 **Noncompetitive Antagonism in Functional Assays.** Simulations showing the effect of a noncompetitive antagonist on responses to the same agonist in a system with high receptor reserve (A) or low receptor reserve (B). Increasing concentrations of the antagonist (3, 10, 30 nM) cause more marked depression of the agonist maximum effect in the low reserve system. Data was simulated using a form of the Operational Model of agonism that assumes that antagonist binding precludes binding of the agonist [32]. The model parameters used were $E_m = 100$, $n = 1$, $\tau = 100$ (high reserve) or $\tau = 3$ (low reserve), $pK_A = 5.0$, $pK_B = 9.0$. Estimates of the antagonist affinity (pK_B) can be made by fitting data directly to this model or approximated as $pA_2 = -\log[B] + \log(r - 1)$ when a concentration-ratio (r) measured at low response levels is used (B).

maximal agonist response (α). The magnitude of the depression will however depend on the agonist under study and the system used. This relates to the concept of receptor reserve whereby maximum agonist effects can be achieved at low levels of receptor occupancy (binding)—for example, 10 percent occupancy may be enough to produce a maximum response and therefore there is a 90 percent receptor reserve. Receptor reserve depends on both the receptor number ($[R_{tot}]$) and the efficiency of stimulus-response coupling as well as the intrinsic efficacy of the agonist. Hence, noncompetitive antagonists will have differing capabilities to depress the maximal response to the same agonist in different systems. The same will be true for different agonists in the same system. The potency of noncompetitive antagonists can be estimated using various models but as a "rule of thumb" the pA_2 ($-\log[B] + \log(r - 1)$), as defined above for competitive antagonists, gives a reasonably accurate estimate of the antagonist affinity (pK_B) when measured at low levels of agonist response [14]. See Figure 2.14B.

All of the modes of antagonism described above are orthosteric; that is, the antagonist blocks access of the agonist to its binding site through steric hindrance. Allosteric antagonists in contrast bind to their own site on the receptor to induce a change in conformation of the receptor, which in turn alters the affinity or efficacy of the receptor for the agonist [33,34]. It is now clear that allosteric ligands can both increase and decrease the affinity and efficacy of other ligands, so allosteric modulators is a more appropriate term. Indeed, perhaps the best known therapeutically used allosteric modulators are the benzodiazepines, which increase the conductance of the $GABA_A$ receptor. One of the key properties of allosteric modulators is their saturability of effect, which can be evidenced in functional experiments such as Schild analysis where a curvilinear plot results (Figure 2.15). Similarly, in Cheng-Prusoff type analyses, such antagonists will produce less than 100 percent inhibition of the agonist response. This behavior results from the fact that while the allosterically modified receptor may have diminished affinity (and/or efficacy) for the agonist, the agonist can still produce receptor activation in the presence of the modulator. As is evident from Figure 2.15B, use of concentration-ratios (r) at low antagonist concentrations can yield reasonably accurate estimates of compound affinity.

Allosteric modulation offers a number of potential advantages over orthosteric antagonists. First, they can modify (i.e., reduce or increase by a small amount) endogenous agonist signals without completely blocking them, thus allowing fine-tuning of responses. Second, there is the potential to increase the duration of allosteric effect by loading the receptor compartment with large concentrations of modulator. Such large concentrations will have no further effect than to prolong the saturated allosteric effect (i.e., the saturability of the allosteric ligand can be used to limit effect but increase duration). Another potential advantage of allosterism is increased selectivity. Orthosteric antagonists often have limited selectivity across receptor subtypes. For example, most

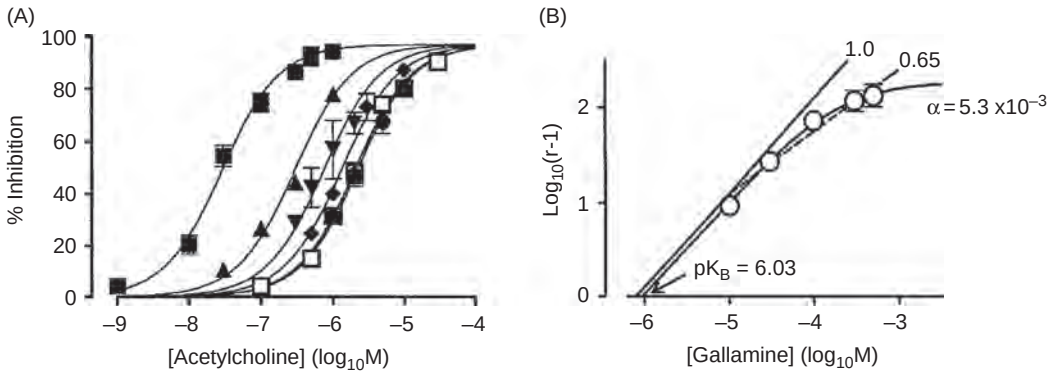


FIGURE 2.15 **Allosteric Antagonism.** (A) The effects of acetylcholine (Ach) on the electrically evoked contractions of the guinea pig left atrium in the absence (■) or presence of the allosteric modulator gallamine at the following concentrations: 10 μM (▲), 30 μM (▼), 100 μM (♦), 300 μM (□), and 500 μM (●). (B) The Schild plot of the data shown in (A). The solid line (slope = 1) denotes the behavior expected for a competitive antagonist, whereas the dashed line shows the best fit linear regression (and associated slope factor) through the points. The curve through the points and associated parameter estimates represent the fit to an allosteric model. The estimated pK_B was 6.03 and the α value of 5.3×10^{-3} equates to a gallamine-induced decrease in the affinity of ACh of 189-fold. *Reproduced with permission from Christopoulos and Kenakin, 2002 [35].*

muscarinic receptor antagonists exhibit poor selectivity between the five known subtypes (M_1–M_5), presumably because they are competing with acetylcholine for very similar recognition sites. However, the surrounding protein structure of the receptors are sufficiently different to offer the potential for selective stabilization of receptor conformations by allosteric modulators. These potential advantages of allosteric modulators remain largely theoretical as very few such agents have to date reached the market. Nevertheless, the approval of the CCR5 antagonist Maraviroc (Selzentry®) for the treatment of human immunodeficiency virus (HIV) infection demonstrated the feasibility of this approach. This compound inhibits HIV entry by binding to a receptor site distinct from where the viral gp120 envelope protein binds [36,37].

Finally, although the discussion above focuses on receptors, allosteric modulation of enzyme function is a well-known phenomenon. The availability of binding sites distinct from those for the substrate again offers the potential for increased selectivity. For example, compounds designed to bind to an allosteric site in a particular protein kinase are likely to have improved selectivity over compounds targeting the ATP binding site.

B. Compound Interference in Primary Assays: Artifacts and False Positives

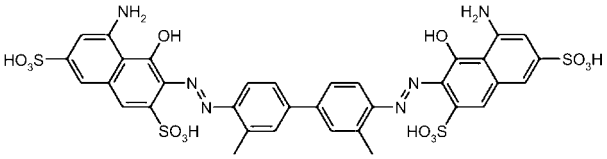
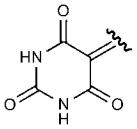
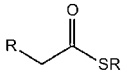
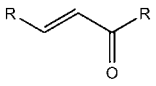
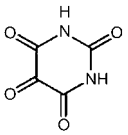
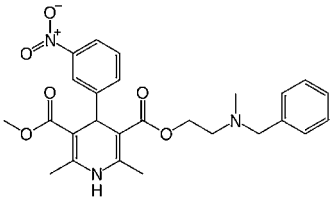
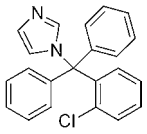
Over the last decade, high throughput screening of large compound collections has been used to successfully identify new chemical starting points for drug discovery programs, notably the CCR5 antagonist Maraviroc. However, the identification of true actives (drugs that interact specifically with the target of interest) from a HTS output has in many cases been hampered by co-detection of large numbers of "active" compounds with either undesirable and/or nonselective mechanisms (i.e., false positives). Compounds can work in an assay nonspecifically (i.e., not target related) through a variety of mechanisms, some of which are very easy to detect while others are more intractable. This is briefly summarized in Table 2.6 and reviewed elsewhere [38].

Careful design of the primary assay and selection of compound library is therefore important in minimizing the propensity to detect these undesirable promiscuous inhibitors. For example, the simple inclusion of detergent and/or protein in biochemical assays can have a profound effect on minimizing false positive detection [39].

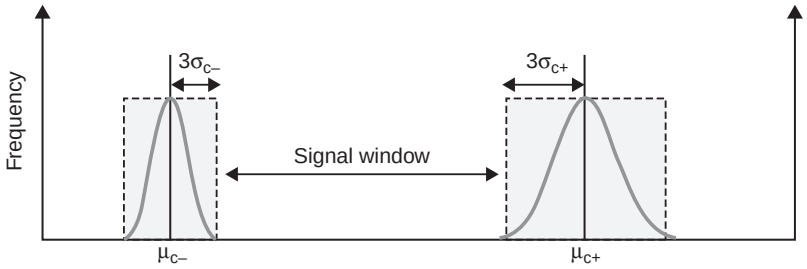
C. Assay Biostatistics

It is not only important to understand how the biological properties of compounds are measured in various *in vitro* assay systems but also to appreciate how consistently an assay performs. This is key to defining criteria for validating compound test data, identifying SAR, and directing medicinal chemistry effort. Assay consistency is usually measured by testing a standard compound (if one is available) and various controls (usually positive and negative conditions to define signal-to-background) in each experiment. This is then used to monitor inter-assay performance on an experiment to experiment basis to ensure the assay can consistently measure the standard compound and retain an acceptable assay window. Experiments are invalidated if the assay value for the

TABLE 2.6 Assay Interference Compounds

Property	Typical structures	Identification	Ease of detection
Interference with assay signal (e.g., fluorescent, singlet O_2 quenchers)	e.g., Trypan Blue 	• Signal only counter-screen in absence of target • Chemical structure and properties	Easy
Irreversible protein damage (e.g., oxidants)	e.g., Alkylidene barbiturates 	• Various redox assays • Chemical structure and properties	Easy
Reactivity (e.g., acyl halides)	 Thioesters  Michael acceptor	• Time-dependent inhibition • Chemical structure and properties	Medium
Activity due to impurity	Free metal ions	• No SAR • Re-synthesis to high purity and retest	Medium
Chemical instability[a]	e.g., alloxan 	• Variable assay results that track with time after synthesis/ dissolution • Activity can be linked to instability • Compound stability assay by LC/MS	Easy
Very low solubility (usually <10 μM in aqueous)	e.g., nicardipine 	• Partial maximal inhibition due to lower free [compound] at high total concentrations • Tendency for assay data at high [compound] to be more variable • Test for solubility (light scattering etc.) • Increase assay [DMSO] if tolerated to improve solubility • Assay signal interference due to insolubility especially at high [compound]	Easy
Compound aggregation[b,c]	e.g., Clotrimazole 	• Flat SAR, but not always • Nonselective within target class • Incomplete inhibition curves • No inhibition with detergent • Potency ranks with nanoparticle/ aggregate detection by DLS • Insensitive to [enzyme] • Noncompetitive • Can be time dependent	Difficult

[a] *Dorfmueller HC, Borodkin VS, Blair DE, Pathak S, Navratilova I, van Aalten DM. Substrate and product analogues as human O-GlcNAc transferase inhibitors. Amino Acids 2010;40:781–92.*
[b] *LaPlante SR, Carson R, Gillard J, Aubry N, Coulombe R, Bordeleau S, et al. Compound aggregation in drug discovery: Implementing a practical NMR assay for medicinal chemists J Med Chem 2013;56:5142 − 50.*
[c] *Sink R, Gobec S, Pečar S, Zega A. False positives in the early stages of drug discovery. Curr Med Chem 2010;17(34):4241.*



$$Z' = \frac{|\mu_{c+} - \mu_{c-}| - (3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|} = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|}$$

FIGURE 2.16 **Assay Biostatistics: Z′ Values.** A graphical representation of variation of positive (C+) and negative (C−) assay signals and how they are used to calculate Z′ values. The table shows the guiding principles in using Z′ values to measure assay performance. *Reproduced with permission from Kenakin [41].*

Z′ values and high-throughput screening assays.

Z′ Value	Description of Assay	Comments
Z′ = 1	No variation (σ = 0) or infinite band of separation	Ideal assay
1> Z′ ≥ 0.5	Large dynamic range	Excellent assay
0.5 > Z′ > 0	Small dynamic range	Adequate assay
0	No band of separation, σ_{c+} and σ_{c-} touch	Dubious quality
< 0	No band of separation, σ_{c+} and σ_{c-} overlap	Impossible for screening

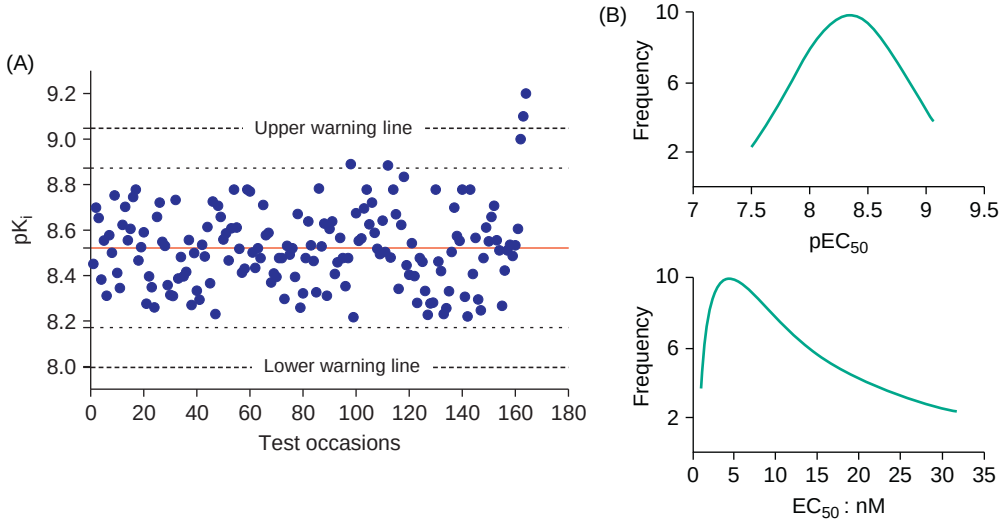


FIGURE 2.17 **Assay Biostatistics: Control Charts and Averaging Biological Data.** (A) Control Chart. pKi values for a compound (y-axis) run as an assay quality control over a number of occasions (x-axis). Solid line is the arithmetic mean. The dotted lines (2x SD (95 percent)) and dashed lines (3x SD (99.7 percent)) on either side of the mean are the warning lines. Data points that drift beyond these warning lines indicate concern over the quality of the assay data on these occasions. Unpublished data. (B) Log normal distribution of pEC_{50} values (upper panel) and skewed distribution of EC_{50} values (lower panel) converted from the same pEC_{50} values. *Reproduced with permission from Kenakin [42].*

standard falls outside an acceptable range and/or the assay signal significantly changes. Typically a Z′ value [40] is routinely calculated to measure both the variability of the assay window (signal−background), the signal, and the background values (Figure 2.16). For more detail see Assay Validation (http://www.ncbi.nlm.nih.gov/books/NBK83783/).

Control charts (Figure 2.17A) are used to monitor assay performance by tracking the standard compound pEC_{50}/pIC_{50} values over the duration of a drug discovery project. The 95 percent confidence interval indicates the range of pEC_{50} values within which the true population value can be expected to be found with 95 percent

certainty - this is generally a 3–4-fold range. Individual experiments in which the standard pEC_{50} falls outside this range should be investigated for any deviation in assay protocol, reagent batch, or process, and the experimental data rejected.

Overall an experiment is acceptable if the measured value (say IC_{50}) for the standard falls within an acceptable potency range (i.e., mean ± 95 percent C.I.) and $Z' > 0.5$ for the signal and background values.

It is important to note that biological measurements are not always normally distributed and may need to be transformed prior to applying statistical tests to understand assay variability [43]. For example, IC_{50}, EC_{50}, K_i, etc. values are not normally distributed, so the mean IC_{50} of 10 μM and 100 μM is not the arithmetic mean, 55 μM, but the geometric mean, 32 μM. In contrast, pIC_{50}, pEC_{50}, pK_i, etc. are usually normally distributed, so these values should be used to interpret SAR and assay variability (e.g., SEM and SD_{n-1}) (Figure 2.17B).

A common question asked by the project chemist is, "What is the significant difference in potency between compound A and compound B as I need to know if compound A is more potent so I can design/synthesize the next compounds based on this SAR?" The best answer to this is to determine the minimal significance ratio (MSR) ideally with a set of compounds with a broad range of potencies. Essentially, the MSR is the smallest potency ratio between two compounds that is statistically significant. For a good assay, the MSR is <3, meaning that a 3-fold minimum in compound potency difference is significant. See http://www.ncbi.nlm.nih.gov/books/NBK83783/ for further detail on the experimental design and statistics behind an MSR calculation. Of course, determining the MSR depends on already having inhibitors with a range of potency and is generally performed in the assay development phase or when SAR unexpectedly changes or assay performance drifts.

D. Selectivity, Cytotoxicity, and Species Cross-over.

Before compounds are progressed to more complex i*n vitro* assays and to *in vivo* testing, it is important to assess their selectivity, cytotoxicity liabilities, and activity at the target protein of the model species to be used.

1. Selectivity

Selectivity is typically initially tested at closely related family members of the target of interest. Thus if the target was, for example, the purinoceptor $P2Y_{12}$, then the activity of lead compounds at the other P2Y receptor subtypes would likely be assessed relatively early. As compounds are further optimized, selectivity screening widens and in this case might include other purinoceptors (P2X receptors and Adenosine receptors) and ultimately a large panel of receptors, enzymes, ion channels, and transporters. Such assays can either be established "in house" or outsourced to one of the many contract research organizations that offer such screening services. Practically, there is a technical and cost limitation to the number of selectivity assays that can be run, but it is not unusual for CDs to be tested in several hundred different assays. Such testing allows compounds with good selectivity profiles to be identified, which greatly facilitates interpretation of data generated in more complex i*n vitro* systems and *in vivo* models. Ultimately, the hope is that good selectivity delivers clinical candidates with excellent efficacy and minimal side-effects.

2. Cytotoxicity

A compound is cytotoxic when it damages the substructure or function of a cell, often leading to cell death, while toxicity generally refers to the damaging effects of compounds on whole organisms. Cytotoxicity is used to try to predict the *in vivo* toxicity of compounds and is often measured in parallel with drug target cell functions to gauge how specific a compound is in affecting the desired cell function; the bigger the difference in compound potency between modulation of a desirable cell function and unwanted gross cytotoxicity, the better the compound. Cytotoxicity is most often determined by simply measuring the ability of a compound to kill cells, and a variety of methods are available (Table 2.7), each with different strengths and weaknesses (see [44] for further detail). Assay selection is most often based on cost, throughput, technical capability, and target cell sensitivity.

3. Species Crossover

Usually—but not always—primary screening is carried out using the human version/ortholog of the target protein. This has the obvious advantage that compounds are screened against the ultimate target of clinical candidates. This strategy does, however, necessitate that lead compounds are subsequently tested against the other species that are typically used in drug discovery programs for both efficacy and toxicology testing. The most common species employed are mice, rats, and dogs, but other species (guinea pigs, rabbits, mini-pigs, macaques,

TABLE 2.7 Cytotoxicity Assay Guiding Principles

Cytotoxic principle	Biomarker	Detection	Pros	Cons
Cell metabolism unable to provide sufficient energy (in the form of ATP) for viable cell function	Intracellular [ATP]	CellTiter-Glo®	Sensitive Fast Stable signal HTS	Expensive
Oxidative metabolism (mitochondrial) unable to provide sufficient energy in the form of ATP for cell function	Dye reduction	Alamar Blue (Resazarin) MTT WST	Cheap Fast HTS	Compound interference
Loss of plasma membrane integrity such that intracellular enzymes leak out or extracellular dye leaks in	LDH/AK release Trypan Blue exclusion	ToxilLght™	Easy Cheap	Compound interference Cell counting, so low throughput

etc.) may also be used. In many cases, compounds have good activity at the target protein in these species and therefore appropriate dosing regimens and straightforward data interpretation is possible. In instances (usually receptor targets) where poor species cross-over is observed (low or no activity at the target in typically used species), it may be necessary to investigate less commonly used species or even to initiate a parallel chemistry effort to identify compounds that do have good species cross-over.

E. Cellular and Tissue Functional Responses

The initial *in vitro* characterization of test compounds is usually carried out in assay systems designed for sensitivity, speed, reproducibility, and cost rather than physiological or pathophysiological relevance (unless phenotypic screening is being employed). The data generated is an important first stage in identifying interesting compounds, but subsequent confirmation of activity in systems with better links to the targeted disease mechanism(s) is essential for compound progression. Differences in receptor number and/or coupling strength, compound penetration, and metabolism are among the factors that may vary between the primary and more disease relevant assays and can markedly change compound activity (see above). Positive data in the latter systems significantly increases confidence that lead compounds will modulate physiological processes implicated in human disease. The systems employed will be target specific but often include a range of cellular and tissue assays. Typical examples of such assays include assessing the effects of compounds on the movement (chemotaxis) of white blood cells for anti-inflammatory targets or on the tone of airway smooth muscle for bronchodilators (Figure 2.18). Such experiments can be complemented by *ex vivo* studies as described below and also aid in the interpretation of *in vivo* experiments.

IV. *EX VIVO* ASSAYS

Ex vivo assays allow the effects of compounds to be studied in biological samples (e.g., tissues, blood, or cells) removed from intact animals or humans. Such experiments are very useful in that they provide evidence that the target and mechanism of action under study is operational in the species studied, as well as demonstrating drug absorption and penetration to the desired site of action. Binding studies or functional studies can be used to measure compound activity. For example, binding assays are often employed to measure receptor occupancy in brain samples. A typical experiment might involve orally administering a test compound to rats at various doses (usually a minimum of four doses plus vehicle is needed) and then terminating the animals at an appropriate time post-dose. The whole brain (or relevant area) is then removed and homogenized and the ability of the test compound in the sample to inhibit the binding of a radio-ligand to the target of interest is measured. This allows the fractional occupancy of the test compound to be plotted as a function of dose administered. In the case of functional experiments, readouts such as up-regulation of adhesion molecules or shape change of inflammatory cells in blood samples are typically used. Such readouts have the advantage that they can often also be employed in early clinical trials as proof of mechanism biomarkers. This approach has been successfully applied to several potential anti-inflammatory targets including CXCR2, stimulation of which induces neutrophil activation. Lazaar

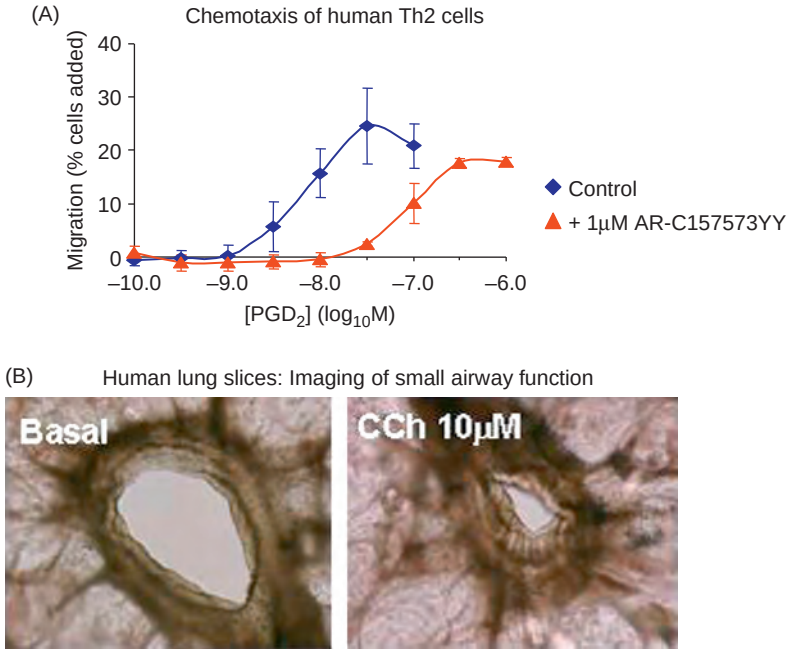


FIGURE 2.18 **Disease Relevant Cellular and Tissue Functional Responses.** (A) PGD_2-induced migration of human Th2 cells. These effects are mediated by the DP_2 (CRTh2) receptor as evidenced by the blockade of the responses with the selective DP_2 receptor antagonist, AR-C157573YY ($pA_2 = 7.0$). Unpublished data. (B) Carbachol (CCh)-induced constriction of a small airway (approximately 2mm in diameter) in a slice of human lung. This response is mediated by the M_3-muscarinic receptor and explains the utility of M_3-antagonists such as Tiotropium in diseases such as COPD which are characterized by restricted airflow. *Unpublished data.*

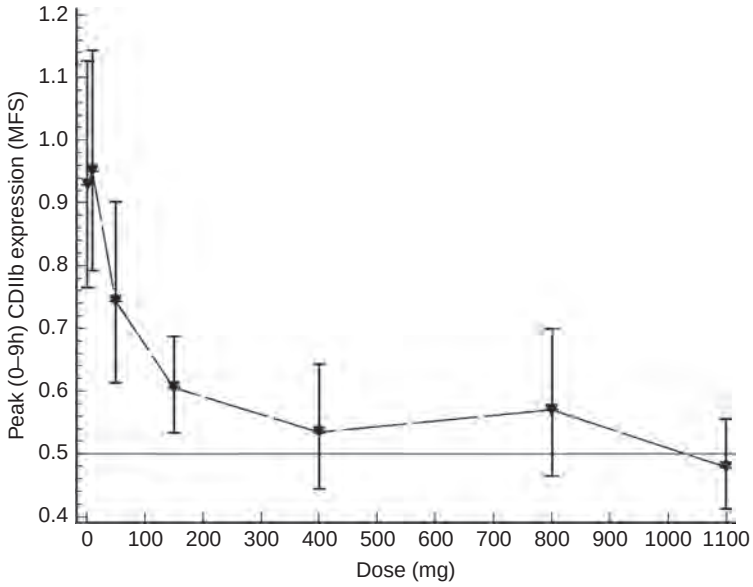


FIGURE 2.19 ***Ex vivo* Assays.** *Ex vivo* assessment of the effect of the CXCR2 antagonist, SB-656933 on CD11b expression on peripheral blood neutrophils following stimulation of whole blood samples with 30 nM CXCL1. A dose-dependent inhibition of CD11b was observed reaching maximum inhibition at concentrations of 400 mg SB-656933 and above. *Reproduced with permission from Lazaar et al., 2011 [45].*

and colleagues [45] used CXCL1 (GRO)-induced up-regulation of the adhesion molecule CD11b to study the *ex vivo* antineutrophilic effect of the oral CXCR2 antagonist, SB-656933, thereby allowing the effective dose of the antagonist to be estimated (Figure 2.19). By studying different time-points after dosing, *ex vivo* experiments can also provide information on the duration of action of compounds.

V. *IN VIVO* ASSAYS

Compounds with suitable potency, efficacy, and selectivity at the primary target and that have shown activity in disease relevant *in vitro* systems also need to have good pharmacokinetics (PK) if they are to deliver *in vivo* efficacy and duration of action. To this end, compounds are routinely tested in a range of *in vitro* assays that serve as models of drug absorption (e.g., Caco-2 cells), metabolism (e.g., rat microsomes, human hepatocytes, cytochrome P450 enzyme assays), and distribution (e.g., plasma protein and tissue binding) prior to *in vivo*

dosing. The physiochemical properties of molecules that are known to be "drug like" provide guidelines for medicinal chemists to optimize the PK properties of compounds, although this often proves more challenging than the optimization of the primary activities of potency and efficacy. The route of drug administration also dictates the properties that need to be optimized. For example, inhaled drugs ideally have low oral bioavailability, whereas oral drugs need good bioavailability. Once compounds with appropriate *in vitro* properties have been identified, they can then be assessed for *in vivo* activity as described below.

A. Pharmacokinetic Models

Pharmacokinetics is discussed in detail in Chapter 23 and mentioned here only in terms of the information such studies yield for compound optimization. The main purpose of PK models is to provide information on "dose to man" and dosing frequency. The driving force in PK is the speed with which the drug is cleared from the body. Clearance (Cl) is typically measured by following the drug concentration in plasma after intravenous administration of the drug in rats (or other species). Simplistically, the elimination of the drug from the body can be approximated by the exit of a substance from a single compartment via a first order elimination process. From this, the volume of distribution (V_d) of the drug can be estimated as the total amount of drug in the body/plasma concentration. Cl and V_d are the primary parameters required to describe PK and are related to the drug elimination rate constant (k) by:

$$k = \frac{Cl}{V_d} \tag{2.11}$$

and since drug half-life $t_{1/2} = 0.693/k$, then:

$$t_{1/2} = \frac{0.693\ V_d}{Cl} \tag{2.12}$$

Thus it can be seen that a reduction in Cl leads to a slower elimination and therefore a longer $t_{1/2}$. Similarly, an increase in V_d (increased tissue binding and sequestration of drugs away from the plasma) leads to a reduction of accessibility to elimination and a subsequent increase in $t_{1/2}$. The $t_{1/2}$ determines the duration of action of the drug after a single dose but will also determine the time to reach steady state on chronic dosing since this is the mirror image of disappearance. Thus drugs with long $t_{1/2}$ may take weeks to reach steady state. Ideally, drugs are administered once or twice a day, so Cl and V_d are key parameters for the medicinal chemist to optimize. As most drugs are designed to be orally active, bioavailability (F) is another important factor in drug design. F defines the fraction of the dose that that reaches the systemic circulation (i.e., is absorbed and survives first-pass metabolism in the liver). It is calculated as:

$$F = \frac{\text{Area under the curve after an oral dose}}{\text{Area under the curve after an equivalent } i.v. \text{ dose}} \tag{2.13}$$

A typical plot of plasma concentration as a function of time (i.e., PK profile) following oral dosing of a compound is illustrated in Figure 2.20A [46].

Ultimately, the dose and frequency of dose are aimed at keeping the drug concentration above the effective concentration in humans for as long as possible without eliciting adverse effects. Allometric scaling (which is based on body weight) can be used to predict the Cl of drugs in humans based on measurements made in preclinical species.

B. Efficacy Models

Whereas PK models measure drug concentration, efficacy models provide pharmacodynamic (PD) readouts (i.e., they measure drug effects). Such efficacy models range in complexity from simple acute readouts of mediators or cell numbers to more clinically relevant measurements such as tumor size after chronic dosing. Irrespective of the readout, they are designed to provide further confidence that the drug target being investigated is clinically relevant and its modulation will lead to a desirable clinical outcome. The species used include mice, rats, guinea pigs, rabbits, dogs, and macaques, and various routes of drug administration are employed (intraperitoneal, subcutaneous, intravenous, inhaled, oral, etc.). As described above, it is important to know that compounds under test have good activity at the target in the species under study and that the compound has a

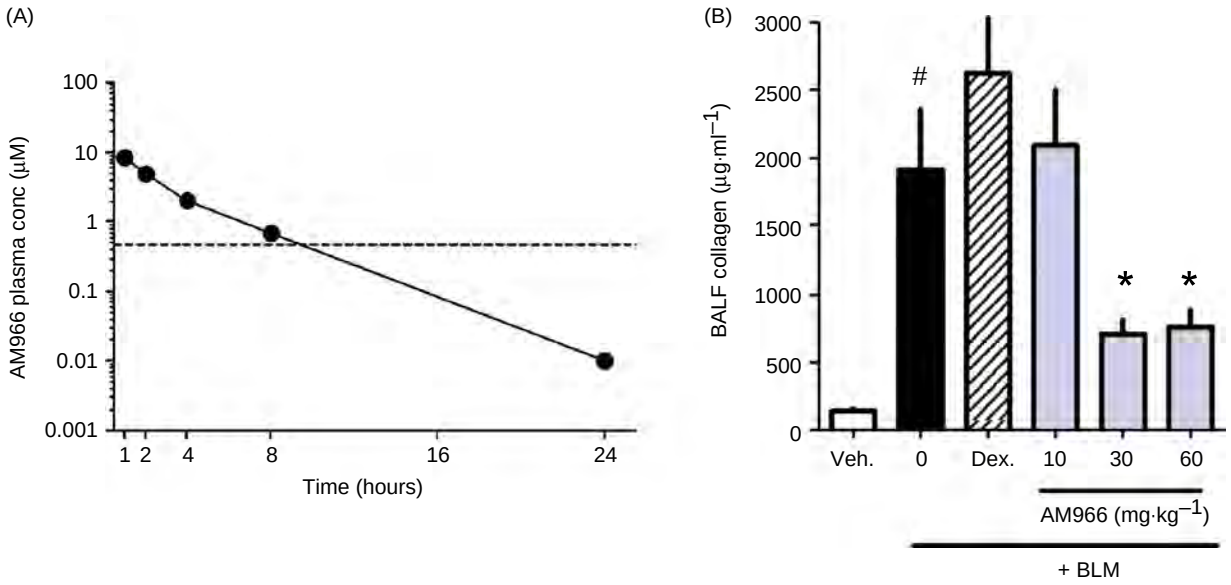


FIGURE 2.20 ***In vivo* Assays.** (A) PK profile of the LPA_1 receptor antagonist, AM699, after oral (10 mg/kg) administration to mice. The dashed line represents the IC_{50} value for AM699 mediated inhibition of LPA responses in CHO cells expressing the murine LPA_1 receptor. The profile indicates that at doses of 10 mg/kg and above, the compound provides IC_{50} coverage for ≥ 8 hours. (B) The efficacy of AM699 in a mouse 14-day bleomycin model. The compound was administered twice a day for 14 days and soluble collagen measured in the bronchoalveolar lavage fluid as a marker of fibrosis (at day 14). Dexamethasone was included as a comparative compound since steroids are used as treatments for IPF despite their dubious efficacy. *Reproduced with permission from Swaney et al., 2010 [46].*

PK profile that allows adequate target coverage for the duration of the PD model. In instances where the ortholog of the human target does not exist in mice or it has a different function, it may be possible to employ transgenic animals or xenografts. An example of the former approach is the transgenic expression of human ICAM-1 in mice allowing them to be infected with human rhinovirus (mouse ICAM-1 does not recognize rhinoviruses that infect humans) [47]. In the cancer field, xenografts of human tumor cells grown in mice are widely used models. Although a vast range of models are available, it should be recognized that many of the efficacy models routinely employed in drug discovery programs are somewhat poor predictors of clinical efficacy. Rather, they allow the investigators to study the effects of drugs on particular mechanisms of action that may (or may not) be relevant to the clinical disease being targeted. As an example, bleomycin-induced lung inflammation and fibrosis is widely employed as a model to study the effects of drugs aimed at treating idiopathic pulmonary fibrosis (IPF) (Figure 2.20B) [46]. Whereas the resultant lung pathology has some similarity with that observed in IPF, profound inflammation precedes the fibrosis, a sequence of events that does not appear to occur in the human disease. Nevertheless such models allow compound activity to be optimized and builds confidence that *in vitro* activity is indicative of *in vivo* activity. Ideally, compound dose-response information is generated over a wide dose range (just as it should be in *in vitro* studies), allowing estimations of ED_{50} or ID_{50} (effective doses of agonists or antagonists/inhibitors producing 50 percent of the maximum response) values and maximum effects to be generated and compared with *in vitro* estimates. PK measurements made during efficacy models allow PK-PD relationships to be further explored which assists in refining "dose to man" estimates made from prior PK studies.

Finally, for some diseases, useful animal models do not exist or are poorly characterized, necessitating progression of drugs straight to the clinic from *in vitro* testing. CF is a good example of this in that compounds that correct or potentiate CFTR function *in vitro* and that have good PK can readily be assessed for activity in humans using surrogates of CFTR function such as nasal potential difference or skin sweat chloride.

C. Safety Testing

At the end of drug discovery and prior to going into human clinical studies, regulatory authorities (e.g., FDA and EMA) have to be convinced that a potential new drug is safe as well as efficacious. To achieve this, data are

TABLE 2.8 Examples of Safety Testing Models

Assays	Test	Purpose
In vitro	Ames[a]	Bacterial assay to measure a compound's mutagenic potential—ability to change gene DNA sequence—and a predictor of carcinogenesis.
	hERG[b] Cardiac ion channel	Binding/functional assay to determine if a compound inhibits hERG activity, a predictor of cardiac arrhythmia
In vivo	Repeat-dose animal study	Assessing the effects of long term compound exposure on in life (e.g., weight, behavior) and terminal readings (e.g., organ histopathology)

[a]*Mortelmans K, Zeiger E. The Ames Salmonella microsome mutagenicity assay. Mutat Res 2000;455(1–2):29–60.*
[b]*Sanguinetti MC, Tristani-Firouzi M. hERG potassium channels and cardiac arrhythmia. Nature 2006;440(7083):463–9.*

submitted from nonclinical pharmacological, pharmacokinetic and toxicological animal studies. At this stage, an understanding of the nonclinical safety profile is essential and should aim to cover three areas:

- establish a safe initial dose level of the first human exposure
- identify parameters for clinical monitoring of potential adverse effects
- special toxicity (e.g., genotoxicity, carcinogenicity, reproduction toxicity)

Generally, a tailored combination of *in vitro* and *in vivo* animal tests is performed (see Table 2.8 for examples), which will include studies of the drug's toxicity on organs targeted by that compound, as well as determining if there are any long term carcinogenic effects or toxic effects on mammalian reproduction.

Importantly, it is assumed that animals and humans respond to administered chemicals in essentially the same way and that exposing animals to the maximal level of compound possible is a valid approach to predicting low incidence human toxic responses. These data are then used to determine a "No Observable Adverse Effect Level" (NOAEL) and then a safe starting dose is estimated for human clinical trials by allometry (scaled by animal/human, size/shape differences).

Acknowledgements

We thank Dr. Martin Redhead, senior bioscientist at Sygnature, for providing and analyzing the Biacore data, and Dr. Iain Walters, principal medicinal chemist at Sygnature, for his helpful comments in reviewing this chapter. We would also like to thank AstraZeneca colleagues past and present who generated much of the data presented here.

References

[1] Swinney DC, Anthony J. How were medicines discovered? Nat Rev Drug Discov 2011;10:507–19.
[2] Gilbert IH. Drug discovery for neglected diseases: molecular target-based and phenotypic approaches. J Med Chem 2013;56:7719–26.
[3] Van Goor F, Hadida S, Grootenhuis PD, Burton B, Cao D, Neuberger T, et al. Rescue of CF airway epithelial cell function *in vitro* by a CFTR potentiator, VX-770. PNAS 2009;106(44):18825–30.
[4] Davis PB, Yasothan U, Kirkpatrick P. Ivacaftor. Nat Rev Drug Discov 2012;11:349–50.
[5] Anderson K, Karlsson R, Löfås S, Franklin G, Hämäläinen M. Label-free kinetic binding data as a decisive element in drug discovery. Exp Op Drug Disc 2006;1:439–46.
[6] Cheng YC, Prusoff WH. Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochem Pharmacol 1973;22:3099–108.
[7] Vauquelin G, Charlton SJ. Long-lasting target binding and rebinding as mechanisms to prolong *in vivo* drug action. Br J Pharmacol 2010;161:488–508.
[8] Swinney DC. The role of binding kinetics in therapeutically useful drug action. Curr Opin Drug Discovery Dev 2009;12:31–9.
[9] Tautermann CS, Kiechle T, Seeliger D, Diehl S, Wex E, Banholzer R, et al. Molecular basis for the long duration of action and kinetic selectivity of Tiotropium for the Muscarinic M_3 Receptor. J Med Chem 2013;56(21):8746–56.
[10] Briggs GE, Haldane JBS. A note on the kinetics of enzyme action. Biochem J 1925;19:338–9.
[11] Copeland R. Evaluation of enzyme inhibitors in drug discovery: a guide for medicinal chemists and pharmacologists. 2nd ed. John Wiley & Sons; 2013.
[12] Tidén AK, Sjögren T, Svensson M, Bernlind A, Senthilmohan R, Auchère F, et al. 2-thioxanthines are mechanism-based inactivators of myeloperoxidase that block oxidative stress during inflammation. J Biol Chem 2011;286(43):37578–89.
[13] Forbes LV, Sjögren T, Auchère F, Jenkins DW, Thong B, Laughton D, et al. Potent reversible inhibition of myeloperoxidase by aromatic hydroxamates. J Biol Chem 2013;288(51):36636–47.
[14] Kenakin T, Jenkinson S, Watson C. Determining the potency and molecular mechanism of action of insurmountable antagonists. J Pharmacol Exp Ther 2006;319(2):710–23.

[15] Black JW, Leff P. Operational models of pharmacological agonism. Proc R Soc Lond [Biol] 1983;220:141–62.
[16] Ariëns EJ. Affinity and intrinsic activity in the theory of competitive inhibition. I. Problems and theory. Arch Int Pharmacodyn Ther 1954;99(1):32–49.
[17] Stapleton MP. Sir James Black and propranolol. The role of the basic sciences in the history of cardiovascular pharmacology. Texas Heart Inst J 1997;24:336–42.
[18] Parsons ME. Histamine H_2-receptor antagonists. In: Leff P, editor. Receptor based drug design, vol. 89. New York: Marcel Dekker, Inc; 1998. p. 195–206.
[19] Costa T, Herz A. Antagonists with negative intrinsic activity at delta opioid receptors coupled to GTP-binding proteins. Proc Natl Acad Sci USA 1989;86:7321–5.
[20] Chidiac P, Hebert TE, Valiquette M, Dennis M, Bouvier M. Inverse agonist activity of beta-adrenergic antagonists. Mol Pharmacol 1994;45:490–9.
[21] Leff P. The two-state model of receptor activation. Trends Pharmacol Sci 1995;16:89–97.
[22] Bennett KA, Tehan B, Lebon G, Tate CG, Weir M, Marshall FH, et al. Pharmacology and structure of isolated conformations of the adenosine A_{2A} receptor define ligand efficacy. Mol Pharmacol 2013;83:949–58.
[23] Kenakin T, Williams M. Defining and characterizing drug/compound function. Biochem Pharmacol 2014;87:40–63.
[24] Soergel DG, Subach RA, Burnham N, Lark MW, James IE, Sadler BM, et al. Biased agonism of the mu opioid receptor by TRV130 increases analgesia and reduces on-target adverse effects versus morphine: a randomized, double-blind placebo-controlled crossover study in healthy volunteers. Pain 2014;155(9):1829–35.
[25] Miller TR, Witte DG, Ireland LM, Kang CH, Roch JM, Masters JN, et al. Analysis of apparent non-competitive responses to competitive H(1)-histamine receptor antagonists in fluorescent imaging plate reader-based calcium assays. J Biomol Screen 1999;4:249–58.
[26] Charlton SJ, Vauquelin G. Elusive equilibrium: the challenge of interpreting receptor pharmacology using calcium assays. Br J Pharmacol 2010;161:1250–65.
[27] Schild HO. pAx and competitive drug antagonism. Br J Pharmacol 1949;4:277–80.
[28] Arunlakshana O, Schild HO. Some quantitative uses of drug antagonists. Br J Pharmacol 1959;14:48–58.
[29] Kenakin TP. The Schild regression in the process of receptor classification. Can J Physiol Pharmacol 1982;60:249–65.
[30] Leff P, Dougall IG. Further concerns over Cheng-Prusoff analysis. Trends Pharmacol Sci 1993;14:110–12.
[31] Furchgott RF. The use of β-haloalkylamines in the differentiation of receptors and in the determination of dissociation constants of receptor-agonist complexes. In: Harper NJ, Simmonds AB, editors. Advances in Drug Research, 3. New York: Academic Press; 1966. p. 21–55.
[32] Schmidt JA, Bell FM, Akam E, Marshall C, Dainty IA, Heinemann A, et al. Biochemical and pharmacological characterization of AZD1981, an orally available selective DP2 antagonist in clinical development for asthma. Br J Pharmacol 2013;168(7):1626–38.
[33] Urwyler S. Allosteric modulation of family C G-protein-coupled receptors: from molecular insights to therapeutic perspectives. Pharmacol Rev 2011;63:59–126.
[34] Müller CE, Schiedel AC, Baqi Y. Allosteric modulators of rhodopsin-like G protein-coupled receptors: opportunities in drug development. Pharmacol Ther 2012;135:292–315.
[35] Christopoulos A, Kenakin T. G protein-coupled receptor allosterism and complexing. Pharmacol Rev 2002;54:323–74.
[36] Muniz-Medina VM, Jones S, Maglich JM, Galardi C, Hollingsworth RE, Kazmierski WM, et al. The relative activity of "function sparing" HIV-1 entry inhibitors on viral entry and CCR5 internalization: Is allosteric functional selectivity a valuable therapeutic property? Mol Pharmacol 2009;75:490–501.
[37] Scholten DJ, Canals M, Maussang D, Roumen L, Smit MJ, Wijtmans M, et al. Pharmacological modulation of chemokine receptor function. Br J Pharmacol 2012;165:1617–43.
[38] Shoichet BK. Screening in a spirit haunted world. Drug Discov Today 2006;11(13–14):607–15.
[39] Ryan AJ1, Gray NM, Lowe PN, Chung CW. Effect of detergent on "promiscuous" inhibitors. J Med Chem 2003;46(16):3448–51.
[40] Zhang J-H, Chung TD, Oldenburg KR. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. J. Biomol Screening 1999;4:67–72.
[41] Kenakin TP. A pharmacology primer: Theory, applications, and methods. 3rd ed. Elsevier Academic Press; 2009. p 149–77.
[42] Kenakin TP. A pharmacology primer: Theory, applications, and methods. 3rd ed. Elsevier Academic Press; 2009. p 1–19.
[43] Hancock AA, Bush EN, Stanisic D, Kyncl JJ, Lin CT. Data normalization before statistical analysis: keeping the horse before the cart. Trends Pharmacol Sci 1988;9(1):29–32.
[44] Peternel L, Kotnik M, Prezelj A, Urleb U. Comparison of 3 cytotoxicity screening assays and their application to the selection of novel antibacterial hits. J Biomol Screen 2009;14(2):142–50.
[45] Lazaar AL, Sweeney LE, MacDonald AJ, Alexis NE, Chen C, Tal-Singer R. SB-656933, a novel CXCR2 selective antagonist, inhibits *ex vivo* neutrophil activation and ozone-induced airway inflammation in humans. Br J Clin Pharmacol 2011;72(2):282–93.
[46] Swaney JS, Chapman C, Correa LD, Stebbins KJ, Bundey RA, Prodanovich PC, et al. A novel, orally active LPA(1) receptor antagonist inhibits lung fibrosis in the mouse bleomycin model. Br J Pharmacol 2010;160:1699–713.
[47] Bartlett NW, Walton RP, Edwards MR, Aniscenko J, Caramori G, Zhu J, et al. Mouse models of rhinovirus-induced disease and exacerbation of allergic airway inflammation. Nat Med 2008;14(2):199–204.

CHAPTER

3

Drug Targets, Target Identification, Validation, and Screening

Walter M.M. Van den Broeck

Janssen Infectious Diseases BVBA, Beerse, Belgium

OUTLINE

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are or what your name is, if it doesn't agree with experiment, it's wrong. ***Richard P. Feynman (American theoretical physicist 1918–1988)***

I. INTRODUCTION

For ages, humans have been using medicinal substances without tools like DNA microarrays to identify them. Instead, they were guided by theories like the concept of the four humors in Greco-Roman medicine or by spiritual systems like animism. The chances are high that modern medicinal chemists would fully reject these rationales. Today we believe that the essential first step in the discovery of a new cure for a disease is the identification of the protein that is at the basis of that disease. The chances are high that medicinal chemists would fully agree with this rationale, but maybe they shouldn't. In this chapter, we will see why.

First, we examine why the definition of a drug target is already a bit misleading. Then we explore whether the mantra "first a target, then a drug" is a good guideline. We compare the three most used strategies for drug discovery today and assess the role of target identification in these strategies. The next question is what kind of targets we should try to identify. Is the search for the cause of a disease a fruitful road to find new cures? Can we find cures altogether? Finally, after having established the difference between the two meanings of target identification, we describe briefly the current and most frequently used methods to identify and validate drug targets.

II. WHAT IS A DRUG TARGET?

In 1891, Paul Ehrlich was experimenting with dyes to stain bacteria. He had already made outstanding contributions in treating infected patients with antitoxins. Together with vaccines, these account for the successful immunotherapy. Ehrlich saw this immunotherapy as chemical reactions between very complex structures. At that time, the concepts of cells and microorganisms were very new and nobody understood the composition of cells. Maybe a cell was one big molecule, (i.e., a cell-molecule). Ehrlich believed that cell-molecules had side-chains to receive nutrients from outside, and he called these side-chains receptors. He thought that bacteria also had receptors and that the staining of bacteria was a chemical reaction between the dye molecule and the receptors. What if this reaction could not only stain the bacteria but also kill them? What if this dye could do the same in an infected patient? Ehrlich showed that methylene blue was taken up by the malaria parasite and had modest effects in two patients. He was extremely excited by this and coined the term "chemotherapy." The difference with immunotherapy was that now the antitoxins—which were very complex and difficult to produce and standardize—could be replaced by well-identified chemicals (small molecules) that were easier to produce and handle.

We owe the concept that a drug acts by binding to a target molecule to Paul Ehrlich. In his own words:,"Corpora non agunt nisi fixata" or "substances don't act unless they are bound." Today this concept is still valid. The *Oxford Dictionary of Biochemistry* defines a drug target as "a biological entity (usually a protein or gene) that interacts with, and whose activity is modulated by, a particular compound." Peter Imming [1] uses the following working definition: a molecular structure (chemically definable by at least a molecular mass) that will undergo a specific interaction with chemicals we call drugs because they are administered to treat or diagnose a disease. The interaction has a connection with the clinical effect(s).

These definitions could give the misleading impression that a drug−target interaction is a one-to-one relation [2], as if every drug acted by binding to one and only one single specific target. This impression is further strengthened by the ambition of every medicinal chemist, starting with Paul Erhlich himself, to synthesize a "magic bullet," an ultra-specific compound that would bind only to the target and to nothing else. However, evidence is growing that many drugs are successful just because they act on multiple different—not co-located—targets, potentially even hitting several pathways together [3]. Of the 1366 drugs reported in DrugBank 2.0, about 960 have more than one therapeutic target [4], a phenomenon called polypharmacology. As a consequence, searching for a super-selective drug may not always lead to the most active compound. In this perspective, target-based drug screening is not well suited to discover these so called "dirty drugs."

The one-to-one relation also doesn't fit with drugs that act by binding to a complex of proteins or even a complex between proteins and nucleic acids. Many proteins form dimers, trimers, or even more complex constellations. In these cases, the drug binding pocket could contain parts of two or more proteins. But the target discovery tools are less well suited to find such targets.

Yet another—very obvious—violation of the one-to-one relation is that the same pocket can accommodate many different small molecules. A substantial part of all new drugs is based on this promiscuous behavior of many pockets. The production of close analogues—or, more pejoratively, "me too drugs" —is often seen as a risk averse and profit driven strategy. Nevertheless, these drugs often result in an important incremental progress in activity, side effect profile, or administration facility [5].

A less obvious violation of the one-to-one relation is the fact that a protein can contain multiple pockets. Usually these pockets are all different and could partially overlap, be indirectly connected by allosteric regulation, or be completely separated. The binding to these different pockets could result in different effects. For example, the binding with nucleoside drugs to the active site of a viral polymerase makes it more difficult for the virus to build resistance than with nonnucleoside drugs that have their binding site outside the active site of the enzyme.

These comments make the picture of a drug target more complex. We could define a drug target as the minimal constellation of molecules that elicit a medically desired effect when bound by a drug.

III. THE PURPOSE OF TARGET IDENTIFICATION

Before exploring the plethora of methods to identify drug targets, we should discuss the role and the value of target identification in the drug discovery process. We will describe the role of target identification in the following three drug discovery strategies for small molecules:

- Target-Based Screening Strategy
- Phenotypic Screening Strategy
- Fast Follower Strategy

A. Target-Based Screening

Target identification is the cornerstone of target-based screening. The concept underlying this strategy is that at the most fundamental level, most drugs work by binding to a specific target. Therefore, if you want to make a truly new drug, the first thing you have to do is to find a new target. The next step is to find small molecules that bind to this target, preferably as specific as possible. This procedure looks so overwhelmingly self-evident, innovative, and scientific that the complete pharmaceutical research community has been dreaming for decades about realizing this strategy. With tremendous efforts, some even succeeded in making drugs this way (e.g., mercaptopurine and cimetidine), but in general the tools were inadequate.

Beginning in the 1980s with the breakthrough in gene technologies, along with the invention of the extremely versatile polymerase chain reaction in 1983 and the publication of the human genome by HUGO and Craig Venter in 2001, pharmaceutical scientists finally received the tools they needed to turn the blind old-fashioned dull drug screening into a highly rational, hypothesis-driven, reductionistic and efficient drug discovery engine. Target-based screening was now possible, and the entire industry embraced it, largely replacing the phenotypic screens [6,7]. Even today, in many presentations on drug discovery for the general public and in many textbooks [8] and publications [9,10], the mantra "first a target, then a drug" is still presented as the main road for drug discovery. The technological advancements are indeed enormous. Today we can sequence the genome of entire organisms in days and measure gene activity in single cells. We can trace individual molecules as they move around in a cell. We can screen millions of compounds in miniaturized and robotized high-throughput assays. Crystallographers can observe protein targets at atomic scale. Faster than ever before, chemists can synthesize very complex molecules, and these can be quantified and identified in very small amounts. Bio-informatics can mine big databases and simulate biological pathways and systems. These technological advancements are certainly as profound and extensive as those in the electronics industry. Many people in the field, particularly molecular biologists and young managers, expected to see an explosion of new drugs against diseases formerly untreatable. But today, most diseases are still here, and the only thing that really exploded was the cost to discover and develop new drugs. In a recent article [11], the authors plotted the number of drugs that could be developed with 1 billion dollar over the years, beginning as early as 1950. The investments were corrected for inflation. It's remarkable that the exponential decrease in output is almost constant over the entire time-span. There is no such thing as a dramatic revolution in increased output. This constant exponential decrease in itself is not scientific proof that there is something wrong with the target-based screening strategy. There could be—and there certainly are—other reasons that could explain the steady increase in R&D cost per drug. But the least thing it proves is that target-based screening and all the new technologies have not brought the expected quantum leap in R&D efficacy. A more specific investigation [12] tracked down the research strategies for all 259 drugs that were approved by the FDA between 1999 and 2008. For the so called first-in-class small molecules (molecules with a new target, not the me-too ones) 38 percent came out of target-based screening. The other 62 percent came out of phenotypic screening. And 62 percent is even an underestimation of the success-rate of phenotypic screening, because this strategy was used less by industry. (Figure 3.1).

Target-based screening is now more and more brought into question [6,12–17]. Although this strategy has certainly led to many successes, it has failed more than expected. Often the targets thrown up by this reductionistic, bottom-up approach were wonderful *in vitro* but disastrous in the clinic due to lack of efficacy or unexpected toxicity [18].

B. Phenotypic Screening

The under-performance of the above described target-based screening leads us to the question whether it's possible to develop drugs without knowing the target in the first place. The answer is, of course, a big "yes." Aspirin

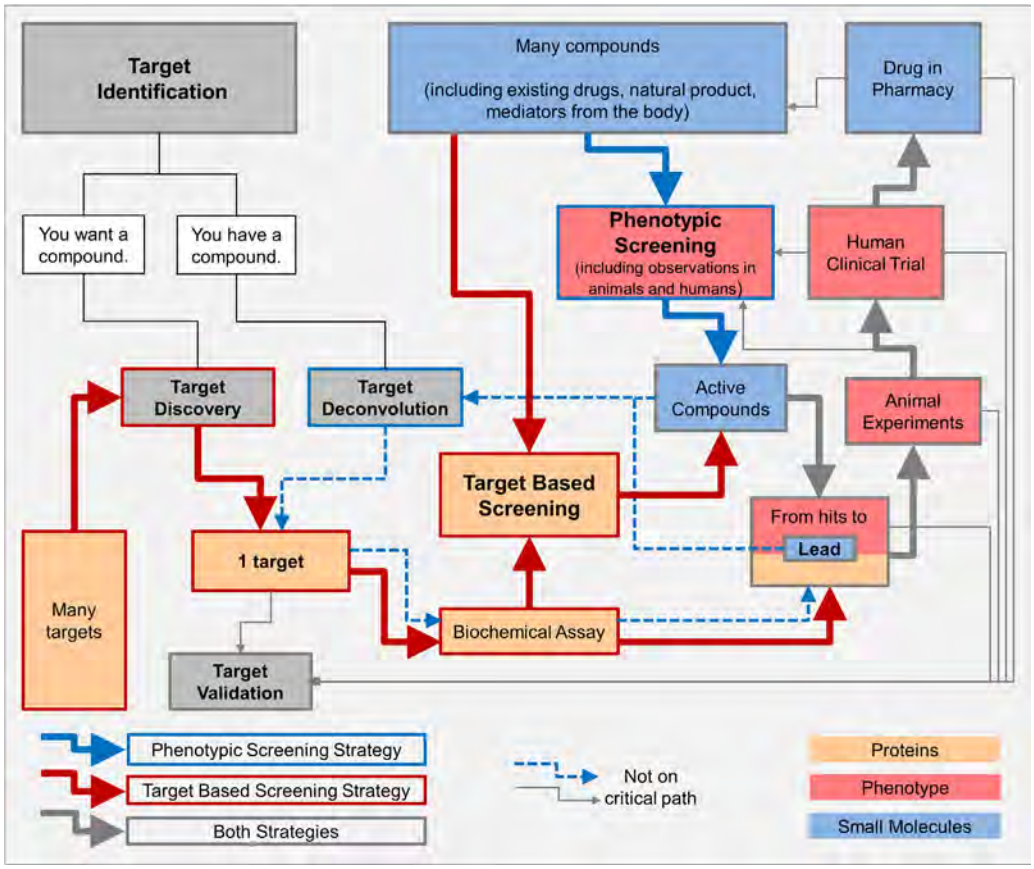


FIGURE 3.1 Target identification can be split into target deconvolution and target discovery. The former is used to identify the target of an active compound, usually obtained by phenotypic screening (blue arrows). The latter is used to discover a new target whose modulation would be of medical use. This target is the starting point for a target-based drug screening project (brown arrows) resulting in an active compound. Target-based screening is a longer process because you first have to screen for a target before you can screen for a compound. In phenotypic screening, the target identification can be done in parallel with the further lead optimisation or even be omitted, and therefore is not on the critical path (dotted arrows).

was synthesized in 1897, but its mechanism of action only discovered by Vane [19,20] in 1971 and its target in 1976. Morphine was used for ages, but its main target, the μ-opioid receptor, was identified by Pert and Snyder [21] in 1973, while other targets are still under investigation. The targets of the general anesthetics are only gradually emerging within the last decade, with the $GABA_A$ receptor as the most prominent one [22]. We could tell similar stories for the benzodiazepines, corticosteroids [9], cyclosporine and FK506 [12,23,24], sulfonylureas, antipsychotic drugs, fibrates, vinca alkaloids, and many antidepressants [8]. And although the targets for most antidepressants have been determined by now, their mechanism of action is still a mystery [25–27]. It may also be that the actual declared targets for some drugs are in fact only off-target effects and that the real targets will be discovered in the future. The reader will appreciate that we can't give an example of this last group today.

It is safe to say that most "first in class drugs" developed before the 1980s were discovered by phenotypic screening. First, you discover an active compound, then you try to determine the target. This is exactly the opposite of target-based screening. Today, phenotypic screening is often seen as cellular screening. But cells are only the smallest living organisms that can build up a phenotype out of their genome. All the experiments in which we examine the effect of at least one compound on the phenotypic level are phenotypic screens. A Neanderthal observing papaver somniferum reducing his tooth pain was performing a kind of phenotypic screen. There is an enormous variety in types of cells, cell combinations, tissues, and animals that can be studied [17]. One could rank these experiments on a scale going from the more reductionistic ones on the left to the more holistic ones on the right. The most holistic and closest to real-life situations are experiments in humans, better known as clinical trials. All other experiments more to the left—however sophisticated and ingenious they might be—are always only an approximation of the real thing. Figuring out a mutation in a gene for leptin in a family with extreme obesity is

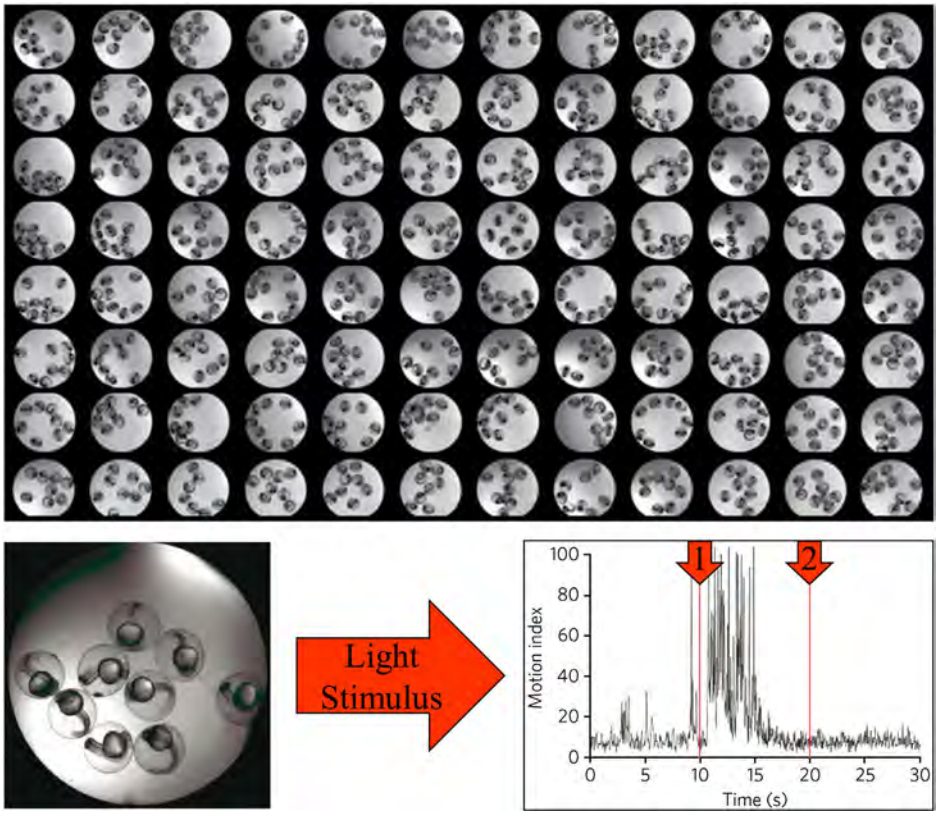


FIGURE 3.2 Eight to ten zebrafish embryos are put in every single well of a 96-well plate. The aggregate motor activity per well is recorded over a 30-second time span, during which two light stimuli are given at 10 and 20 seconds (red arrows). The reactions to the light flashes are translated to distinct behavioral patterns that can be used to evaluate potential neuroactive drugs. Using an automated platform, 5,000 embryos could be screened per microscope-day. *Source: Adapted from reference [28].*

only an approximation of the possible genes causing obesity in the whole population. Indeed, the developed leptin analog was extremely effective, but only in the very small group of people carrying the mutation [13].

Case studies and observations of side-effects in clinical trials (or via drug surveillance) represent an extremely valuable source of information. They form an unplanned phenotypic screen of the highest level. Although a clinical trial is a planned phenotypic experiment, the unexpected side-effect that could eventually result in a drug repositioning exercise was not planned. Drug repositioning and obtaining lead compounds based on unplanned observations are not extremely rare, but they can hardly be conceived as a planned research strategy.

Phenotypic screenings of limited but well-chosen sets of compounds in animals have been a very productive way to find new drugs in the period stretching from the end of World War II until the 1980s. Today, animals are not used anymore to screen compounds for having interesting (unexpected) effects. They are used to confirm expected effects (and to study the pharmacokinetic properties and safety profile of compounds). But the zebrafish is changing this again. For those not familiar with this fish—which is never served in restaurants—it is a very small (4 cm) fish with transparent embryos that for several reasons has become a favorite vertebrate animal model in research. One can easily put 10 embryos in one well of a 96-well microtiter plate. The behavior of these embryos upon treatment with a compound can be monitored automatically. For example, the team of Randall Peterson [28] measured the motor activity upon stimulation with intensive light and could link the so called photomotor response profiles to different classes of neuroactive drugs. In an automated platform, 5,000 embryos could be observed per microscope per day. In total they screened 14,000 different compounds, and the behavioral information of more than a quarter million embryos was collected. (Figure 3.2).

Most phenotypic screens today are still performed in cells. Again, we have to be aware that cells are only an approximation of tissues and organs. And cells in laboratories—often immortalized cancer cell-lines, possibly manipulated to mimic a disease state—could behave totally differently than normal cells. Nevertheless, these cellular phenotypic screens have many advantages over biochemical screens with purified targets. A major advantage is that in a cell, the target is in its normal biological context: it is present in the real compartment of the cell, in the real constellation with other proteins and membranes, at the real concentration, in the real phosphorylation-ubiquitination-glycosylation-acylation and whatever state, and embedded in its full metabolic and regulatory pathways.

A second advantage is that cells contain more than one target. Throw a potential bactericidal compound on a culture of *M. tuberculosis* and you are testing its activity on 614 vital proteins, about 3,400 other proteins, and an unknown set of nonprotein targets [29]. In contrast, in a biochemical screen you only test one known target at a time and it can take years to set up an assay for a second protein. Then there is a good chance that this protein is in an artificial, biologically irrelevant state. The only advantage is that the high-throughput capacity is usually higher, but critics would say that this only helps to produce an even bigger pile of irrelevant hits.

A third advantage is that phenotypic screens do not fail to identify prodrugs, which need to be converted to active drug by a host cell or a bacterial cell. Prodrugs are not active in biochemical assays.

For the fourth advantage of phenotypic screening, it is instructive to look at the search for antivirals against the hepatitis C virus (HCV). Compared with a human or even with a single cell, a virus is a simple organism. You sequence its little genome and try to figure out the function of the handful of viral proteins. You can bet that all of them must be vital and hence must represent good drug targets, especially the enzymes. Target-based screening makes much sense here. You synthesize the enzyme, make a biochemical screen, test thousands of compounds, and—with some luck—you obtain active inhibitors. If, after further optimization, these inhibitors have an acceptable oral bioavailability, if they have a good half-life, if they can penetrate the liver cells, and if they have no safety issues, you have a drug. This turned out to be the case for compounds targeting the HCV protease enzyme, while compounds targeting the HCV polymerase enzyme [30] are in late stage clinical development. However, the class of replicase complex inhibitors—formerly known as NS5A inhibitors—could only be found by phenotypic screening. These compounds have picomolar *in vitro* activities and about eight of them are now in phase 2–3 studies [30]. The exact role of the nonstructural viral NS5A protein and the way the compounds bind to it is still a matter of debate. This enigmatic and intractable protein would never have been selected as a target in the target-based screening strategy. This example shows that phenotypic screening can be used to access less obvious targets. Mark Fishman, president of the Novartis Institutes for BioMedical Research (NIBR), who initiated a phenotypic screening effort when he moved to Novartis in 2002, puts it this way: "For me it's a discovery tool. The single biggest impediment to drug discovery is the small number of new, validated targets that we have. Phenotypic screening is one way of moving beyond well-defined targets from the literature to discover new therapeutic targets and new disease biology." [6] Another example is the discovery of bedaquiline, the first anti-tuberculosis drug with a new mechanism of action in 40 years [31]. Bedaquiline [32], synthesized by J. Guillemont, was discovered by a phenotypic screen and its target, the bacterial ATP synthase, was identified afterwards. Again, this target would probably never have been chosen as starting point in a target-based screening campaign and can actually not even be tested in isolation because of its complexity.

What's the role of target identification in the phenotypic screening approach? There are two cases.

- In cases where you use phenotypic screening to find active compounds, it is still very useful to identify the target. First, building a SAR based on activities measured with a biochemical assay gives you a higher resolution because you exclude all the variations caused by cell penetration, off-target binding, cell compartmentalization, and so on. But one has to keep the cellular activity under close surveillance. Working years in the chemistry labs to build a compound that is super-active in the biochemical assay but doesn't penetrate the cell anymore is a waste of time. Second, when 3D-stuctures of the target are available, molecular modeling could assist the SAR. Third, knowing the target could also help in exploring the toxicity and the role of the pathways in human and animal models. Fourth, in general it is substantially more challenging to get a compound without a target through the consecutive approval committees within and outside the company. However, knowledge of the target is not always a regulatory requirement for drug approval. Between 1999 and 2008 the FDA has approved nine compounds with unknown targets [12,19].
- In cases where you use phenotypic screening as a way to find new targets, the role of target identification is self-evident. This approach to finding new targets becomes more and more a valid option for three reasons. First, there is the increased capability to read out ever more complex phenotypes, there is the increased capacity to handle more animals by using robots, and there is the renewed interest in phenotypic screening in general. Second, the methods to determine a target—as will be described further in this chapter—are becoming increasingly efficient. And third, targets found via phenotypic screening are more likely to become valid targets.

C. Fast Follower Strategy

The fast follower strategy consists in synthesizing analogs of an existing drug in the hope of obtaining a compound with at better profile than the starting drug. This strategy is often regarded as not innovative, while there

are many examples where the second or third so called me-too drug offered real clinical improvements [33]. This strategy has many advantages. There is no need to invest in the discovery of a target, and the target is already validated in the best possible way: in humans. Besides the known therapeutic efficacy and safety profile, even the financial risks and benefits can be estimated with more certainty. The probability that at least some of the new analogs will have similar or even better activities is usually high. The new compounds should, of course, fall outside the patent of the original drug. Given the virtually infinite chemical space and the creativity of medicinal chemists, this requirement is an easily attainable goal. It becomes trickier to stay outside the competitor's patents applications on the same target during the 18 months before they are published. Given these advantages, it should not come as a surprise that the number of fast follower drugs is rather high. Of the 259 drugs approved by the FDA between 1999 and 2008, 164 (63 percent) were follower drugs. In the subset of small molecules, the percentage was even higher at 74 percent [12].

The role of target identification in the fast follower strategy is mainly limited to the confirmation that the new compounds still work on the same target. Target knowledge can, of course, be used to drive the SAR.

Chemists have to be aware that the fast follower approach is less attractive for biologists than for chemists. Biologists can be very creative and use their full intellectual potential in proposing new hypotheses on working mechanisms, in the endeavor to find new targets and in the design of new assays. But in the fast follower approach, most of this early work is already done and described by the originators. Biologists merely have to reproduce parts of this work. For the medicinal chemists, however, the fast follower approach stretches their chemical knowledge and creativity to the limit. They don't have to reproduce but instead come with something better in a new way.

IV. TARGET OPTIONS AND TREATMENT OPTIONS

In this section, we deal with questions like "What are good targets?"; "Which targets should we try to discover?"; "What can we achieve with targets?"; and "What are the therapeutic targets and treatment options?"

Today, most drugs are discovered and developed by commercial organizations. Even if their shareholders would not be interested in making money, companies at least have to make sufficient profit to cover the cost of new research and development. Therefore, the selection of a target is driven by scientific as well as commercial considerations [5]. In an ideal world, the commercial value of a drug should be in parallel. (Figure 3.3).

with its therapeutic value. The therapeutic value can be determined by calculating the DALYs (Disability Adjusted Life Years) that a disease or condition causes. Severe but rare diseases result in fewer DALYs compared to less severe but very common chronic diseases. In the real world, one should add the difference in purchasing power of patients to the equation. Even then, predicting the return of a drug has proven to be very hard. For example, several companies were not interested in licensing atorvastatin (Lipitor) [5]. Here we will focus on the scientific considerations. Given a selected disease, what are the chances of finding a small molecule against it? What is a good target if we want to cure the disease? Or is this already a wrong starting point?

In some languages, the word "drug" or "medicine"—not to mention the word "cure" —is translated by a word that literally means "agent that cures." That sounds fair and familiar, but most of us don't realize how wrong this really is. The majority of drugs synthesised and sold today don't cure at all. A cure should result in the permanent end to the specific instance of the disease, without further need for medication. A cure also implies that a relapse should not be the consequence of an inadequate treatment. Proton pump inhibitors or H2 antagonists can cure a patient with a duodenal ulcer within several weeks. More precisely, the temporary suppression of the normal gastric acid production permits the body to repair the gastric mucosa. But they do not remove the major causative agent, the bacterium Helicobacter pylori, and within a year the patient can relapse. Therefore, antibiotics are added to complete the treatment. Cases where a patient relapses after antibiotic treatment could be due to inadequate cure (i.e., not all bacteria were killed) or to reinfection from outside.

A quick look at the top 200 pharmaceutical products by 2009 worldwide sales [34] gives us plenty of examples of drugs that don't cure: cholesterol lowering drugs, anticoagulants, anti-asthmatics, anti-psychotics, anti-rheumatics, insulin, anti-diabetics, anti-Alzheimer products, blood pressure lowering drugs, painkillers, erythropoietin, anti-epileptics, erectile dysfunction products, immunosuppressive agents, nasal vasoconstrictors, hypnotics, sedatives, vaccines, urinary incontinence products, anti-Parkinson products, and narcotic analgesics. The pharmacological classes that do cure—at least sometimes—form a much shorter list; anti-infective drugs and anti-cancer drugs are the most prominent ones. Anti-ulcerants, as explained above, are a bit of a special case. The acute use of anticoagulants to dissolve a blood clot in the coronary arteries just after a heart attack could also be regarded as curative.

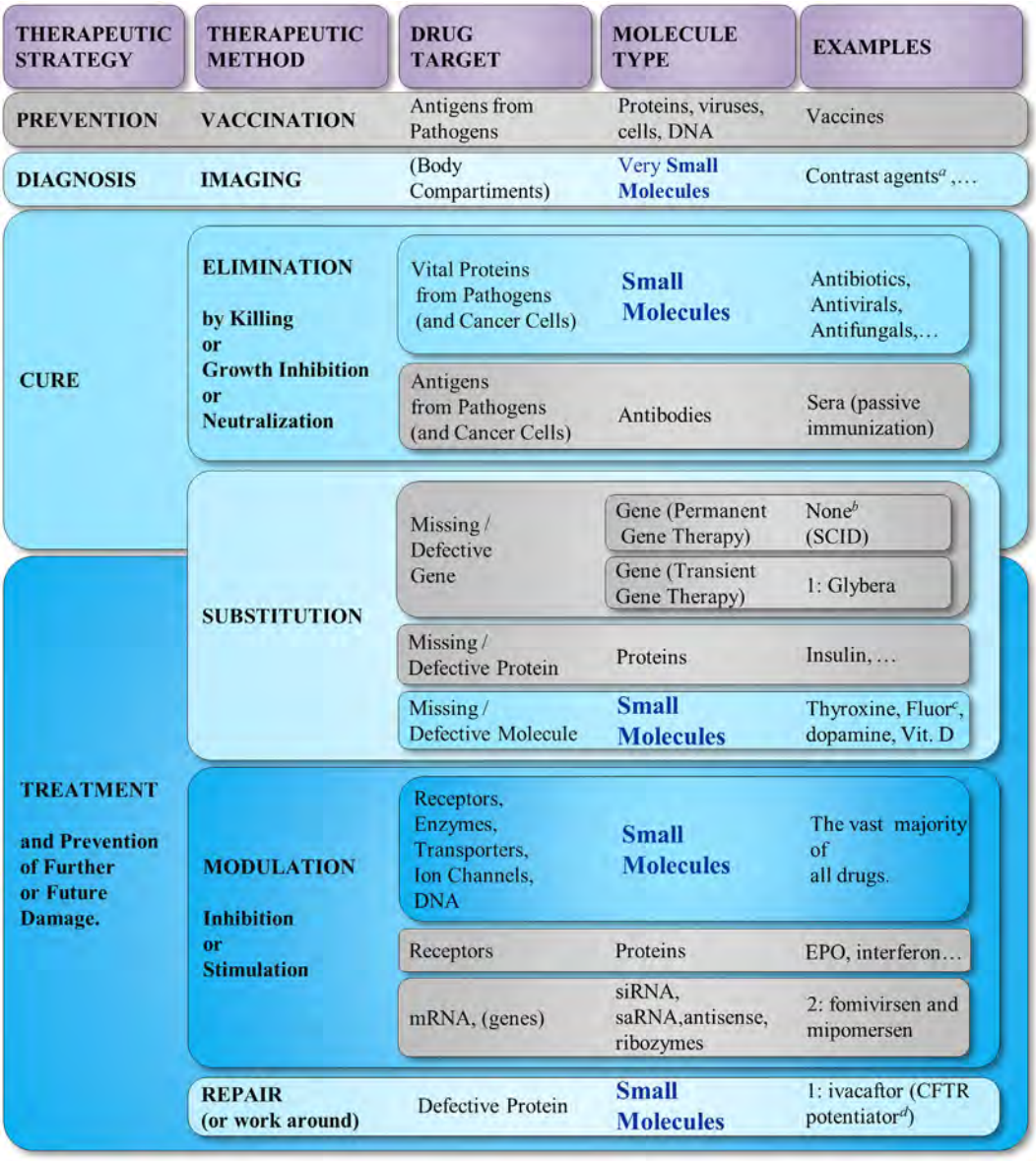


FIGURE 3.3 A simplified overview of the drug landscape. The areas in blue are covered by small molecules. The darker the blue, the more predominant the small molecules are. Biologicals are in grey. In general, it's very hard to cure a disease with small molecules, except for infections and cancer. Therefore, the targets for small molecules can best be found in the pathways that could modify the disease. The proteins that are on the causative path may be of indirect use but are rarely direct targets for small molecules, except again for infections and cancer. Some remarks to the figure: a. By definition, only molecules that are introduced into the body can form part of the diagnostic drugs and they are mainly used for imaging. b. Although there have been some curative gene therapy interventions in the past (like for SCID), at the moment the only approved gene therapy is Glybera, but that works only transiently and is therefore grouped under treatment. c. Nutrient supplements are grouped under the treatment of a shortage with preventive effects. d. Ivacaftor could also be seen as an allosteric agonist.

The observation that most drugs don't cure is not mind-blowing, but surprisingly this observation is often ignored by scientists looking for a new target. They persist in looking for the cause of a disease. They are looking for the gene mutations that cause cystic fibrosis, schizophrenia, Alzheimer disease, diabetes, obesity, and others. They are looking for the medical insults in early life that may cause epilepsy or autism or depression. But when they finally determine the deletion of three nucleotides that result in the loss of amino acid phenylalanine in position 508 of the CFTR protein, causing this protein to fold abnormally and finally being responsible for two-thirds of the cystic fibrosis cases, what then? Can they undo the deletion with a small molecule? No. Can they replace the CFTR gene with a small molecule? No. Can they replace the cystic fibrosis transmembrane conductance regulator by a small molecule? No. What about when scientists finally discover that 5 percent of schizophrenia is

caused by being born in winter or by hypoxia during fetal development? Can they change the birthday of the patient with a small molecule? Can they influence fetal conditions decades after birth with a small molecule? We could go on with other examples. The bottom line is that in general it is not possible to remove the cause of a disease with small molecules. Only when the cause is a cancer cell or a pathogen, things that don't have to be repaired or replaced but simply have to be killed, only then can small molecules cure.

There are of course exceptions. The small molecule ivacaftor is the first approved molecule that repairs a defective protein. The defective protein is the CFTR protein with the less frequent G551D mutation that accounts for 4–5 percent cases of cystic fibrosis. The word "repair" is perhaps not the best term, because ivacaftor most probably facilitates the impaired channel gating via an allosteric potentiating effect [35]—a kind of patch, so to speak. Ivacaftor has to stay in place to keep the protein channel open. It works only for a small subset of the cystic fibrosis patients. Again, ivacaftor is an exception, and although we can't exclude that science will change this tomorrow, the lesson today is that proteins or genes that cause disease could be very interesting but are rarely good drug targets for small molecules.

What are then better targets? Beta2-adrenoceptor agonists are effective in the acute treatment of asthma, although asthma is not caused by a defective sympathetic nervous system stimulation. The underlying cause of essential hypertension is unknown but can be treated—not cured—with alfa1-adrenoreceptor antagonists (among others). These examples show that it is not mandatory to explore the full pathogenesis of a disease to invent drugs that relieve symptoms, that counteract potential lethal situations, and that prevent further damage or future damage. It also points to where we should look for targets. In order to mitigate or counteract the effects of a disease, we should look for powerful master switches: proteins that are specialized in modulating cell functions. These proteins have a name: receptors. So it makes complete sense that 60 percent of all targets are located on the cell surface [36] and that 44 percent of the human targets are receptors [37]. Of these receptors, 82 (42 percent) are G protein-coupled receptors targeted by 357 unique drugs [37]. Of the 1,062 drugs that act on human targets, 563 (53 percent) of them bind to receptors. In fact, except for the 88 cancer targets, almost all other human targets are involved in disease modifying therapy. We should also not forget that more and more drugs have nothing to do with disease. Examples include women who want to become temporally infertile and the elderly who want to smooth out their wrinkles or extend their normal life-span.

In most articles [1,2,36–38] that estimate the number of actual and potential targets, the authors are rather prudent and conservative. Despite our 20,000 genes, the number of targets is often estimated to be only a few thousands. We are not going to refute this guesstimate but will make three remarks:

1. The famous visionary and science fiction writer Arthur C. Clarke once said that "If an elderly but distinguished scientist says that something is possible, he is almost certainly right; but if he says that it is impossible, he is very probably wrong." [39]
2. In the years just before the Wright brothers took off with their motorized airplane, elderly distinguished scientists declared that it was impossible to fly with machines heavier-than-air. These scientists apparently never observed a bird.
3. In the years just after the first antisense drugs were approved by the FDA, imagine how many more drug targets we would have if we were able to deliver easily [40] antisense and siRNA constructs into human cells. Imaging we would be able to activate genes with saRNA or be able to repair genes with rather small molecules [41].

V. TARGET DECONVOLUTION AND TARGET DISCOVERY

Figure 3.1 illustrates the two different meanings and applications of target identification. To avoid confusion, we will call the first target deconvolution and the second target discovery. In literature, target identification is often used for both, although target deconvolution is used more and more.

1. Target deconvolution.

   With target deconvolution, we mean that we have an active compound for which we want to find out, elucidate, assess, determine, and identify the target on which it is working. Target deconvolution [42] is the more recent appellation. Most often the active compound comes out of a phenotypic screening project. Therefore, target deconvolution is seen as an important and almost indispensable component of the phenotypic screening strategy. It is, however, not on the critical path in the sense that a project could advance to lead optimisation and in principle even to clinical trials and market approval without target knowledge.

The target identification can be executed in parralel with the further process. As stated before, the FDA approved 9 drugs [12,19] without known targets between 1999 and 2008.

2. Target discovery.

With target discovery, we mean that we want to find a potential target that can be used in a target-based screening project to obtain active compounds. The starting point is different, as we have no compound. We have a disease for which we want to find, screen, and explore possible targets. Target discovery is an absolute indispensable starting point for every target-based screening campaign. Target discovery can also be a part of academic research into the pathogenesis of a disease. One valid way to discover new targets is to start a phenotypic screening followed by target deconvolution of the lead compound. Most of the methods used for target discovery can also be used for target deconvolution.

VI. METHODS FOR TARGET IDENTIFICATION AND VALIDATION

There are dozens of methods to identify drug targets and the field is heavily based on extremely rapidly evolving technologies such as molecular biology, miniaturization, microscopy, automation, and informatics. Although we selected the most frequently used methods, the list is certainly not complete and risks becoming outdated very rapidly. No single method or guideline can be applied to every target, and the methods are complementary. There is also no common way to classify the different methods [9,23,42–50]. Therefore, we have gathered them in an overview (Table 3.1) so that the reader can classify them using multiple criteria.

A. Affinity Chromatography

Concept. An immobilized small molecule is used as a bait to fish the best binding protein(s) out of a mixture [42,44–46,51].

Input. An active small molecule and a mixture of possible targets (proteins).

Output. A fraction of the mixture that is enriched in targets that have a high binding affinity to the small molecule.

Group. Affinity based method. Direct method. Compound-centric chemical proteomics method. Target deconvolution. Bottom-up. One of the most widely used methods, especially for targets that are only present in mammalian cells or whole organisms [44].

Description. Affinity chromatography is a method to separate or purify mixtures based on different affinities toward a solid phase. Here the mixture is a cell or tissue lysate that contains the target(s) for the small molecule of interest. The small molecule is chemically attached to the solid phase. This solid phase (also referred to as stationary phase, column packing, matrix, beads, or resin) often consists of agarose or sepharose. The cell lysate acts as the mobile phase and is incubated with the solid phase. The components that have the highest affinity for the small molecule will bind preferentially to the solid phase. The unbound components of the cell lysate are then washed away or eluted from the column using an appropriate buffer. The next step is the elution of the bound target. This can be done with a buffer that disrupts the interaction between the target and the immobilized small molecule. Alternatively, the target can be eluted using an excess of free small molecules. The purified target now remains to be identified using advanced mass spectrometry or immunodetection on a SDS-PAGE gel.

Requirements. This method can only be applied when one has already obtained a (small) molecule by phenotypic screening or other sources. The second requirement is that it must be possible to chemically attach the small molecule to the solid phase in such a way (or under the assumption) that it doesn't interfere with the binding toward the target protein. It can be helpful to dispose of a SAR to identify parts of the molecule that can be used to attach the linker.

Advantages. This method has been very successful, and a large number of modifications are available to overcome several drawbacks. The most important advantage is that although the binding happens *in vitro*, the targets in the cell lysate are still in a very natural state [42,51]. The cells can be primary cells, even from human biopsies. These cells contain the entire proteome for that cell-type and disease-state. This is an unbiased probe compared to individually recombinant synthesized proteins in artificial prokaryotic cells, missing post-translational modifications. In the lysed cells, the proteins were embedded in their full metabolic and regulatory pathways, possibly in some relevant activation and labeling state.

TABLE 3.1 Some Methods for the Deconvolution, Discovery and Validation of Targets

Method	Objective	Group	Small molecule state	Binding occurs in	Target state	Target production	Detection of binding	Target identification
Affinity Chromatography (A)	T. Deconvolution	affinity methods	linked to solid phase	vitro	in lysate	normal + lysate	affinity separation	mass spectrometry
Fractionation (A)	T. Deconvolution	affinity methods	labeled but free	vitro	in lysate	normal + lysate	fraction is labeled	mass spectrometry
Phage Display (A)	T. Deconvolution	affinity methods	linked to solid phase	vitro	linked to phage	expression-cloning	affinity separation	sequencing
mRNA Display (A)	T. Deconvolution	affinity methods	linked to solid phase	vitro	linked to mRNA construct	expression-cloning	affinity separation	sequencing
Haploinsufficiency Profiling (HIP) (C)	T. Deconvolution	genetic methods	totally free	cells	normal	modulated expression	change in phenotype	sequencing
siRNA (C)	T. Deconvolution	genetic methods	totally free	cells	normal	modulated expression	change in phenotype	sequencing
Resistant mutants (D)	T. Deconvolution	genetic methods	totally free	cells	normal vs mutated	normal	change in phenotype	sequencing
siRNA (E)	T. Discovery T. Validation	genetic methods	na	na	normal	modulated expression	na	change in phenotype
Yeast three-hybrid (F)	T. Deconvolution	genetic methods	free but with linker	cells	linked to activating domain	expression-cloning	change in phenotype	sequencing
DNA microarrays (G)	T. Discovery	genetic methods	na	na	normal	normal vs altered	na	spot location + database
Protein microarrays (G)	T. Deconvolution	microarrays	labeled but free	vitro	linked to solid phase	not defined	affinity separation	spot location + database
Gene expression Profiling (H)	T. Deconvolution	Profiling	totally free	cells	normal	normal	change in transcriptome	profile comparison
Analysis of patho-physiology (I)	T. Discovery	top-down approach	useful	cells	normal	normal	not defined	not defined
Study of "old" drugs (J)	T. Discovery	top-down approach	not defined	not defined	not defined	not defined	not defined	not defined
Systems biology (K)	T. Discovery	in silico	in silico	in silico	in silico	in silico	in silico	in silico
Simulation of a human patient (L)	complete knowledge	in silico	in silico	in silico	in silico	in silico	in silico	in silico

The letters in brackets refer to the section in this chapter where the method is described. T.: Target. Na: not applicable. Normal means that the target is in its normal physiological state and location. The grouping of methods is rather arbitrary. Some authors group phage display under genetic methods, while there are reasons to group protein microarrays under affinity methods.

Disadvantages. First, the immobilization of the small molecules is not a routine operation and can abrogate the biological interaction with the protein. It will not be possible to find a good linker for every compound. Second, the assumption that the protein with the highest affinity to the small molecule is likely to be the biological target can be wrong. Also, when the small molecule has to bind to more than one protein in order to exert its biological activity, only the most abundant or strongest binding protein will be recovered. Third, the method works best for high-affinity bindings with dissociation constants ranging from 10^{-7} to 10^{-15}M. The lower the binding affinity, the more target proteins will be lost [42]. Fourth, this method is best suited to soluble proteins. Membrane bound proteins or nonproteins will be hard to identify with the standard set up [46]. Fifth, the conditions have to be adapted for every small molecule. This is not a high-throughput assay. Further disadvantages are nonspecific protein binding, the use of complex instrumentation, and the trial and error aspects in the fine-tuning [44].

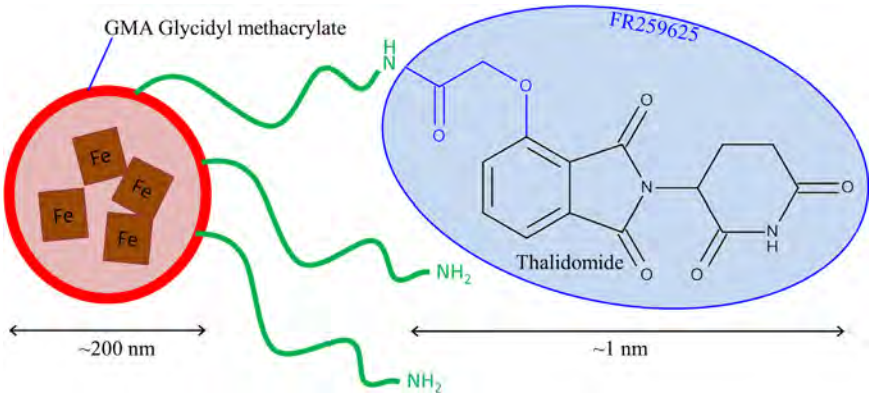


FIGURE 3.4 The FR529625 carboxy derivative (in blue ellipse) of thalidomide (in black) is bound to the amino-groups of the linkers (in green) of the FG beads (in red). The ferrite-glycidyl methacrylate (FG) beads allow for magnetic separation of target molecules that bind to thalidomide from the cell extract. The coating with GMA (Glycidyl mathecrylate) has a very low nonspecific protein adsorption. *Redrawn based on reference [54].*

Examples. Because affinity chromatography is one of the most used methods for target deconvolution, the number of examples is huge. Some examples are: tubulexin A binding to CSE1L and tubulin, Methyl gerfelin binding to glyoxalase 1, piperlongumine binding to GSTP1, Imatinib binding to kinases, and resveratrol binding to eIF4A

Case study. A dramatic example is the identification of a primary target of thalidomide teratogenicity in 2010 by the group of Ito [52]. Thalidomide was marketed during the late 1950s as a safe sedative because no lethal dose could be established in animals. Therefore, it was prescribed to thousands of pregnant women, especially for morning sickness. Soon thereafter, an epidemic of typical birth defects broke out. Thalidomide could be identified as the cause and was removed from the market, but for more than 10,000 children it was too late. Thalidomide was the biggest disaster in pharmaceutical history and was the start of a much more stringent FDA. But, amazingly, the drug is back. After the serendipitous discovery in 1965 by Dr. Sheskin of its miraculous effect on erythema nodosum leprosum [53], a very painful leprosy complication, the drug is now used under strict control for the treatment of leprosy and multiple myeloma. Because the effects of the drug are that unique, several hundreds of clinical trials are now undertaken, mainly in the field of cancer. Unfortunately, in South America children are again born with limb and other birth defects due to noncompliant use of thalidomide. If the mechanism of action of thalidomide's teratogenicity could be found, safe analogs could perhaps be conceived. Many mechanisms were proposed, but the direct target was not known. The group of Ito performed a classic affinity purification using magnetic FG beads. To bind the compound to the beads, a carboxylic derivate (FR259625) was made [54] and bound covalently to the amino-group of the Glycidyl methacrylate. The loaded beads were then incubated with HeLa cell extracts (a cancer cell once taken without permission from the cervix of Henrietta Lacks). The beads were washed very stringently and bound proteins finally eluted using free thalidomide [55]. The eluate was put on a SDS gel electrophoresis and silver stained. Only two protein bands could be found and identified with mass spectrometry: cereblon (CRBN) and DNA binding protein 1 (DDB1). The latter in fact did not bind directly to thalidomide but was piggy-backing on cereblon. After identification of the target, several target validation experiments were performed, as will be described at the end of this chapter. (Figure 3.4 and Figure 3.5)

Variations.

1. Reduction of sample complexity. When the drug binds many proteins in an unspecific way, one can try first to eliminate the proteins that are not of interest [56].
2. Compound analogs. A way to validate the target is to synthesize compounds with weaker and stronger affinities for the protein and then to examine the correlation between the phenotypic effect and the affinity [44]. The extreme is to compare with a nonactive analog [42].
3. Biotinylation. When biotin can be attached as linker, the biotin can then be bound with high affinity to streptavidin.
4. Fractionation. In this method, the small molecule is not immobilized but free. The molecule is radioactively or fluorescently labeled and incubated with the cell lysate. This mixture is then fractionated and the radioactivity or fluorescence is measured to determine which fraction contains the bound target protein. Radioactive labeling can be a solution when it's impossible to attach a linker to the molecule, but this is an expensive approach. The fractionation or separation can be done using affinity chromatography, 2D gel electrophoresis, or any other technique [44].

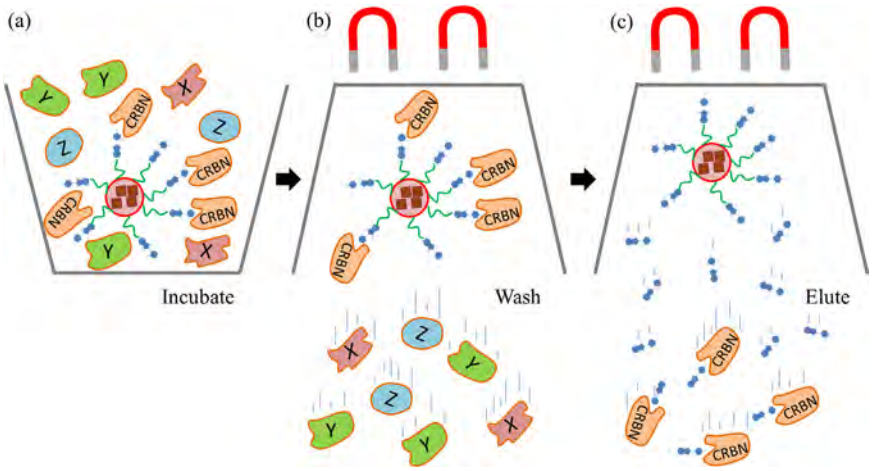


FIGURE 3.5 (a) The magnetic FG beads are incubated with a cell extract. The protein CRBN (cereblon) binds to thalidomide (blue molecule) while other proteins (X, Y, Z…) don't. (b) Using the magnets, it's easy to wash the beads and remove the nonbinding proteins. (c) In a last step the CRBN protein is eluted using free thalidomide. *Based on reference [55].*

5. Magnetic beads. Ferrite-glycidyl methacrylate (FG) beads can be used to allow magnetic separation of the solid phase from the cell extract [54].
6. Quantitative mass spectrometry. Different methods using isotopes (iTRAQ, ICAT, SILAC) [51] can help to make mass spectrometry more quantitative.
7. Phage display. The different target proteins are displayed by phages on their surface. The phage population is separated by affinity chromatography with the small molecule and the enriched elution is amplified in bacteria. With this target enriched phage collection, another affinity round is then performed. This way, low abundant binding proteins are amplified [42].
8. mRNA display. Similar to the principle above. The target proteins are linked to their encoding mRNA's instead of a phage. After affinity enrichment, the mRNA can be amplified with PCR and translated *in vitro* to obtain more of the target protein. With this target enriched collection, a second cycle is started [42].

B. Genetic Methods

Description. Today the molecular biology tools to play with genes are extremely versatile and powerful, yet at the same time often easy to use at a gradually decreasing cost. As a consequence, more and more methods for target identification and validation are directly or indirectly based on them.
A first approach is modulate gene expression for target deconvolution. The principle is simple. Imagine a (small) drug molecule that changes the phenotype of a cell by blocking an unknown target protein. What will happen when we change the concentration of this protein by modulating the expression of its gene? Decreasing the target concentration will make it easier to block the target with the same amount of drug. The cells become more sensitive to the drug. Decreasing the target concentration to zero by deleting the gene or suppressing the expression completely will (in an ideal situation) mimic the phenotype induced by the drug. An alternative to deleting the gene is to mutate the gene in such a way that the protein loses its function, also resulting in mimicking the drug phenotype. Increasing the target protein concentration by overexpressing its gene will make the cells less sensitive to the drug. How can we use all these modulations to identify the drug target when we don't know the gene to modulate? Well, we "simply" modulate all possible genes one-by-one, and each time we look for a change in drug induced phenotype. When there is a change, we have found the target gene and hence the target protein, or to put it more accurately, we found at least one target.

Down-regulating expression can be done in many ways:

- deletion of the gene (knock down [57], used in the Haploinsufficiency profiling in yeast method);
- mutating the gene (randomly chemical induced mutations);
- binding to the DNA with so called zinc finger proteins [58];
- binding to the mRNA with antisense RNA; or
- silencing the gene with siRNA.

Up-regulating expression can be done, for example, by transfecting the cells with cDNA or plasmids or viral vectors. A new tool to activate expression is the use of saRNA, but this is still under debate [59,60].
A further extension to the above approach is not to modulate existing genes but to introduce new ones. For example, transfect yeast with all human genes one-by-one, like in the yeast three hybrid method, or make viruses that express human proteins on their surface like in the phage display method.
The modulation methods for target deconvolution are known under the name "chemical genomic methods."
A second approach is to modulate gene expression for target discovery. Again we modulate the genes one-by-one, and we look for changes in phenotype induced by this modulation, usually without giving any drug. This way we can infer the function of a gene. We can also look for specific phenotypic changes. When, for example, cancer cells don't grow anymore after a certain gene is down-regulated, than the gene-product is possibly an interesting drug target.
A third approach is to modulate a specific gene to validate that a given protein is indeed a good target. Modulation, usually down-regulation, mimics the effect of a drug without yet having a drug for the target. This down-regulation can even be done in animals to further add confidence to the target. When one has a drug, an extra confirmation consists in mutating the target gene in such a way that the protein keeps its function but can't be blocked anymore by the drug, rendering the cell insensitive to the drug. This would prove that the drug doesn't work via other mechanisms.
A fourth approach is to identify existing modulations in gene expression (including mutations) associated with diseases or drug resistance, using DNA micro-arrays or sequencing. (Figure 3.6).

C. Haploinsufficiency Profiling in Yeast

Concept. Lowering the expression of a target by deleting one of its two genes in the diploid yeast makes the strain more sensitive to a drug that acts on that target and often results in decreased growth [33,44,46,58,61].
Input. At least one compound of interest and a collection of strains, each with a deleted gene.
Output. Genes that code for proteins that directly or indirectly interact with the compound.
Group. Direct Methods. Genetic deconvolution methods. Chemical genomics.
Description. Imagine the picture of a drug that inhibits the growth of yeast by inhibiting a certain enzyme. Increasing the abundance of this enzyme will render the compound less effective. Decreasing the abundance of this enzyme will make the yeast more sensitive to the drug. One way to decrease the abundance of the enzyme is to shut down one copy of every two genes in the diploid yeast that codes for the enzyme. This method is based on the natural phenomenon of haploinsufficiency: in diploid organisms, deletions or mutations in one copy of the diploid set of a gene may result in an abnormal phenotype, especially when it is a vital gene. A deletion of a gene that doesn't code for the enzyme will not change the sensitivity to the drug, unless the gene

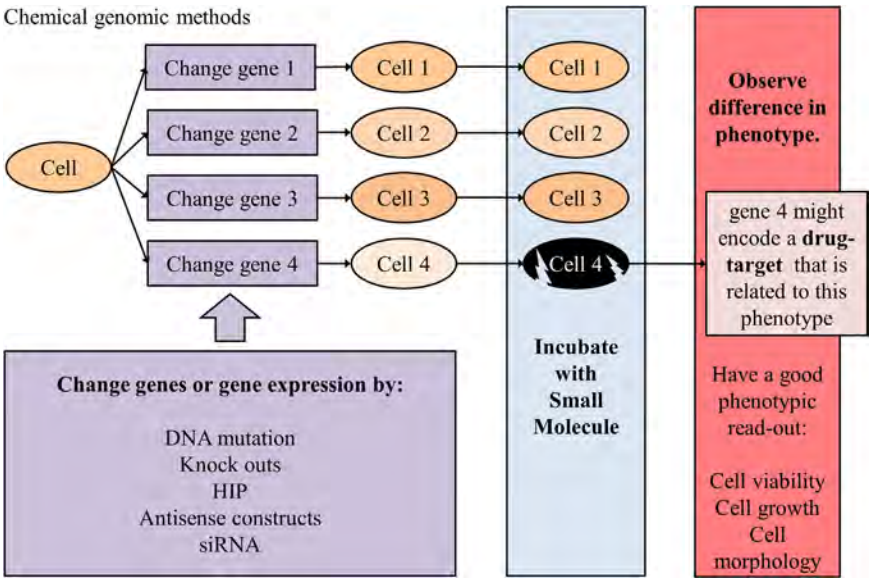


FIGURE 3.6 A collection of cells is made in which all possible genes are modulated one-by-one, and each time we look for a difference in drug induced phenotype. When a target is, for example, less expressed, then the cell could become more sensitive to the drug. Here crippled cell 4 indicates that gene 4 could code for a protein interacting with the drug.

has something to do with the pathway that comprises the enzyme or—more indirectly—results in the inhibition of an efflux pump, thereby increasing the concentration of the small molecule. Hence, observing which deletions affect the growth in the presence or absence of the compound can pinpoint possible targets of that compound or at least components that are part of the working or not working of the drug or even its toxicity. The Yeast KnockOut (YKO), a pooled collection of S. cerevisiae strains where each one of the more than 6,000 genes are completely deleted and replaced by a unique 20 base pair sequence—a "barcode" —is now available [33]. When the pooled collection grows in contact with the small molecule, some strains will grow less. This can be monitored by extracting the genomic DNA and quantifying the "barcodes" with PCR. This way the strain that contains the known deletion that interacts with the small molecule is identified.

Requirements. This method can only be applied when one already has a (small) molecule obtained by phenotypic screening or other sources.

Advantages.

1. The small molecule can interact with the proteins in their normal *in vivo* constitution (except for an artificially decreased abundance in one specific protein) [46].
2. The small molecule does not have to be linked to a bead or a hybrid protein. This not only makes life easier but also guarantees that the binding capacity is not hindered or influenced.
3. Only 0.1−1.0 mg compound per assay is required.
4. Because all genes are tested in the pooled collection, the identification of multi-target interactions is possible. Also, proteins that do not bind at all with the compound but have an effect because they are part of the pathway could be identified. Given the fact that good drugs are often "dirty drugs" that hit multiple targets, this method is well placed to study not only the target but also the whole mechanism of action of a drug.
5. There is a good chance that the cellular localization of the yeast protein mimics that of the human homolog. This further increases the real-life path that the drug has to follow to reach its targets.

Disadvantages.

1. In the yeast three hybrid, one can test (in principle) any human protein. But the protein is not in its normal state and habitat. In the haploinsufficiency model, the proteins are in their normal state and habitat, but this only approximates the human proteins and conditions. Human proteins that lack a yeast homolog will not be picked up.
2. A second limitation is that only interactions between small molecules and target proteins that impair cell growth (a specific phenotypic condition) will be picked up [33].

Examples.

Brefeldin A binding on Sec7p, doxorubicin binding on SIZ1, cisplatin binding on FCY2, NMD2, NOT3, SKY1, methotrexate binding to DFR1, FOL1, FOL2., 5-fluorouracil binding to CDC21, RRP6, RRP41, RRP44, RRP46, NOP4, MAK21, SSF1, YPR143W and tunicamycin binding to ALG7p, HAC1, GFA1.

Variations.

The key of the HIP method is that the target of interest is decreased in concentration. Deleting one copy of the diploid set is one option to achieve this, but one could also use RNA interference to reduce the target protein concentration. Other variations just do the opposite: increase the target concentration using plasmid overexpression. A nice experiment is to combine the under- and overexpression. When a reduced protein makes the cells more sensitive to the drug and at the same time makes the cell more resistant to the drug when overexpressed, then the probability that this protein interacts with the small molecule is high [46].

D. Analysis of Resistant Mutants

Concept. When a bacterium is treated with an antibiotic, resistant mutants can be selected and their genome sequenced. The mutated gene often reveals the target of the antibiotic [62].

Input. A drug and a resistant mutant.

Output. The mutated gene.

Group. Genetic methods. Target deconvolution.

Description. This method is very useful in determining the target of molecules obtained from a phenotypic screen with bacteria or viruses. As viruses and bacteria replicate quickly with a high frequency of spontaneous mutations, they easily acquire mutations resulting in resistance to a given drug. Selecting resistant mutants and sequencing their complete genomes to identify the mutation resulting in resistance is nowadays a very feasible exercise.

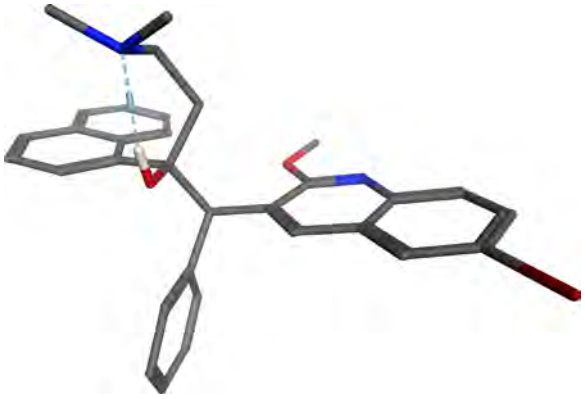


FIGURE 3.7 Structure of bedaquiline rendered in MOETM based on crystallographic data in reference [32]. The target of bedaquiline, the ATP synthase of *M. tuberculosis*, was identified by sequencing resistant mutants of the bacterium.

Case study. Using a phenotypic screen, the diarylquinoline bedaquiline (R207910) was found to be active against *M. tuberculosis* (TB) with a minimum inhibitory concentration (MIC) of 0.030 μg/mL. The activity was specific for mycobacteria and resistant strains were still sensitive to several other TB drugs, suggesting a new mechanism of action. The group of Koen Andries compared sequences of resistant and sensitive strains and identified two point mutations in the atpE gene, coding for a protein of the mycobacterial ATP synthase [31,63]. The ATP synthase enzyme is essential for the generation of energy in mycobacteria, and had previously not been described as a target for a TB drug or in fact any antibiotic. In December 2012, the FDA granted bedaquiline accelerated approval, based on phase II clinical trial data, as part of combination therapy to treat adults with multi-drug resistant pulmonary tuberculosis when other alternatives are not available [39]. (Figure 3.7)

E. siRNA for Target Validation

Concept. Temporary suppression of a gene-product with siRNA mimics the effect of an antagonistic drug. The value of the protein as a potential drug-target can be tested without having a drug. More recently discovered saRNA seems to up-regulate genes and could perhaps mimic some agonists [64].
Input. A potential valuable drug target.
Output. The phenotypic effect of the decreased production of the target protein.
Group. Target validation. Genetic methods. This method can, however, also be used for target discovery and target deconvolution.
Description. To add confidence that a protein would be a valuable drug target to treat a disease or at least to confirm that the protein is implicated in the disease, one could remove the protein by simply deleting the gene for it. But this is rather drastic and in reality far from simple, especially in animals. The gene could be vital during embryonic development and the production of knock-out animals is time consuming and demanding. It would be much more convenient if we could turn off the gene during a limited period (e.g., during the experiment). By preference, this would be done just with an injection, almost like a drug.
This is possible with siRNA. siRNA stands for small interference RNA and consists of 21−25 nucleotides long, double-stranded RNA molecules, complementary to a sequence in the gene one would like to knockdown. The detailed mechanism of the knockdown is rather complicated, involving several enzymes, and is still not fully understood. The bottom line, however, is that the mRNA is cleaved and hence the translation is blocked. The nice thing is that in principle any mRNA can be deleted this way. One only needs to have a unique sequence of 21 nucleotides belonging to the mRNA. In practice, not every sequence works equally well, and other sequences turn out not to be selective enough. Delivery of the siRNAs into the cell is another challenge. High pressure injection of naked siRNAs into the tail vein of mice resulted in effective uptake into the liver, kidney, lung, and muscle. Delivery of siRNA via expression out of a viral vector is a useful option. Even more sophisticated is the creation of a cell (or transgenic animal) that expresses the siRNA under control of a promotor that can be activated with an existing small molecule.

Advantages.

1. The biggest advantage is that you can study the effect of inhibiting a target without having a drug that does the job.
2. Down-regulating gene expression with siRNA can mimic drug effects in a much better way than knocking out a gene by deletion or mutation. With siRNA, the cell or animal can develop and behave normally as long as no siRNA is given. The siRNA also has only a temporary effect, almost like a drug.
3. In principle one only needs to have a unique "antisense" sequence of 21−25 nucleotides to block a protein. No knowledge of the structure of the protein is required. In theory, all antagonistic drugs could be replaced by siRNA molecules that could be designed in silico in—let's exaggerate—a few hours. The fact that the last ten years only two such drugs made it to the market is due to the famous delivery problem, and to some other issues like off-target effects and lack of efficacy.
4. siRNA is a relatively fast and inexpensive method.

Disadvantages.

1. Down-regulating a gene is not the same as blocking a specific part of the gene-product. A drug could, for example, completely block one function of a receptor, but the receptor is still present and can have other functions that are not affected. Knockingdown the gene will decrease all functions of that protein.
2. Down-regulating a gene can have more effect than just a decrease in concentration of the one gene-product.
3. The delivery of siRNA remains a challenge.
4. The activity and selectivity of the siRNA sequences are hard to predict.
5. The down-regulation is not 100 percent; there might still be some protein synthesis ongoing.

Variations.

Here we described the use of siRNA to block one specific gene in order to validate its gene-product as a drug target. siRNA can, however, also be used for target discovery by blocking genes one-by-one in order to correlate genes with phenotypic effects. And in combination with a (small) molecule, siRNA can be used like HIP to determine the target of the drug

Additional comments.

A small disadvantage is that only antagonists and enzyme inhibitors can be mimicked with siRNA. But this could change since the recent discovery [59] that short double stranded RNA complementary with the promotors of a gene can strongly activate the transcription. These so called saRNA (for small activating RNA) can mimic agonistic drug effects. Whether they do this via direct activation of the promotor or by silencing an upstream repressor is still under debate [60].

F. Yeast Three-Hybrid System

Concept. A cell (often yeast) is genetically engineered to give a signal when a small molecule (the bait) binds with protein X (the prey). Every cell is transfected to express another protein X. Transfecting plasmids are made from a library of cDNA [42,44,65].

Input. A small molecule and a set of cells, each expressing one member out of a set of possible target proteins.

Output. A read-out that permits the identification and selection of the cell (or colony) in which the small molecule binds to the expressed protein. This protein is localized in the cell nucleus.

Group. Expression cloning approaches.

Description. This system also is based on the affinity between the compound and the target protein, but it is used to circumvent the problem of low protein abundance by expressing the protein in genetically engineered cells. To understand this method, one has to dive into some molecular biology. Because that is not the purpose of this book, we will not dive very deep. The method is derived from the well-known yeast two-hybrid system. The yeast two-hybrid system or Y2H was developed by Stanley Fields and Ok-Kyu Song [66] in 1989 in order to detect protein−protein interactions. The key underlying the Y2H assay is that most eukaryotic transcription factors can be split into two parts and still activate transcription on condition that the two parts are sufficiently close together. A direct physical contact between the two parts is not necessary. The two parts are the DNA binding domain (DBD) and the activation domain (AD). To study the interaction between known protein 1 and unknown protein X, Fields and Song made two hybrids: they fused protein 1 with the DNA binding domain (hybrid 1, called the bait) and protein X with the activation domain (hybrid 2, called the prey). These hybrids were expressed in the yeast cells by transfecting them with vectors (plasmids) encoding the two hybrid proteins. In fact, for every protein X a

different vector is made. When the bait (protein 1) "captures" the prey (protein X), they bring their attached domains to each other, thereby bringing the activation domain close enough to the transcription start site to activate the transcription of the reporter gene. In the original assay, the reporter gene was the lacZ gene of E. coli. This gene encodes for β-galactosidase, an enzyme that cleaves lactose into glucose and galactose. Lactose can now be replaced by artificial derivatives that produce a useful read-out when cleaved. For example, cleaving the derivative 5-bromo-4-chloro-3-indolyl-β-D-galactoside turns the cells blue. This indicates that protein 1 binds with protein X. In the yeast three-hybrid system, a third hybrid is made: a fusion between a small molecule 1 of interest and a ligand molecule like methotrexate. The protein 1 is replaced by a ligand binding protein like dihydrofolate reductase that will bind strongly to methotrexate. The net result is an assay in which the colonies will turn blue when the bait—now ending with small molecule 1—captures the prey, protein X. The blue cells are then selected, and protein X is identified based on its DNA (Figure 3.8).

Requirements. This method can only be applied when one already has a (small) molecule obtained by phenotypic screening or other sources. The second requirement is that it must be possible to chemically attach the small molecule to a ligand like methotrexate using a linker like polyethylene glycol and in such a way (or under the assumption) that it doesn't interfere with the binding toward the target protein. It can be helpful to dispose of a SAR to identify parts of the molecule that can be used to attach the linker. A third requirement is that one needs to make use of a cDNA library containing the proteins to study. The compilation of this library (genomic DNA, normalized cDNA, full-length cDNA, etc.) can affect the protein expression. Finally, the cell type should be appropriate for your proteins.

Advantages.

1. As the target proteins are expressed artificially, they can be in higher concentrations than in a cell lysate.
2. The whole mammalian proteome can be screened in intact cells.
3. Once you have the complete set of transfected yeast cells, the technology is rather easy and can be automated.

Disadvantages.

1. Yeast is not a mammalian cell. However, mammalian versions (like MASPIT [42]) are now available
2. When molecules can activate the reporter gene via other ways, this will result in false positives.
3. Many false negatives due to steric hindrance of the protein in its fusion complex.
4. The small molecule has to be linked to methotrexate.
5. Proteins have to be fused with the activation domain and are possibly not in their natural state.
6. The binding takes place in the nucleus of the cell, hence membrane penetration is part of the equation.

Examples.

Atorvastatin binding to PDE6D, kinase inhibitors like purvalanol B to bind with known targets (like CDK1, CDK5, CDK6) and unknown targets (CLK3, PCTK1, PCTK2, PAK4, RSK3, FYN, YES, EPHB2, and FLT4) [42], Sulfosalazine binding to SPR.

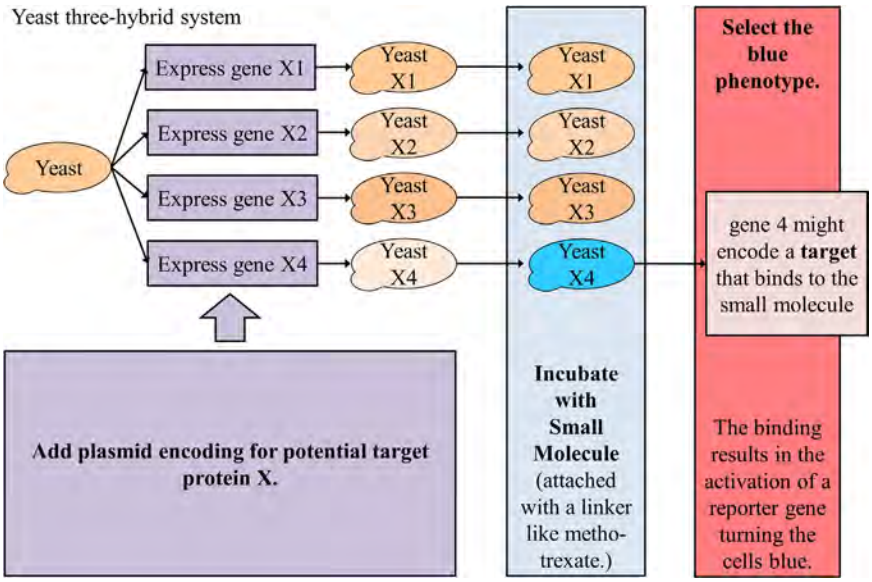


FIGURE 3.8 A collection of yeast cells is built in which all target proteins of interest are introduced one-by-one. This is done by introducing a gene construct that codes for the protein linked to a reporter gene activating domain. When a drug binds to the target protein, the reporter gene is activated. In case the reporter gene codes for β-galactosidase the yeast can be made to turn blue.

Variations.
Since its creation in 1989 the Y2H technique has seen a bewildering number of variations and improvements, very well described by Stynen et al. [65] We only mention the MASPIT variant. MASPIT stands for mammalian small molecule protein interaction trap. The protein hybrids are replaced by a JAK-STAT hybrid system inducing a reporter gene. The advantage, besides being in a mammalian cell, is that the binding occurs in the cytoplasm rather than in the nucleus.

G. DNA Microarrays

Concept. A DNA microarray is a surface with an array of microscopic spots of DNA. Each spot contains short single strand DNA molecules, all with the same unique sequence. Each gene expressed in a sample will hybridize with its corresponding spot, which can be identified using fluorescence readout [67–70].
Input. cDNA converted from mRNA out of a cell lysate.
Output. A list of all of up- and down-regulated genes.
Group. Target discovery. No small molecule needed. Target validation. Target deconvolution: see comparative profiling.
Description. Most disease phenotypes must be ultimately based on changes in gene expression in some cells. Hence, a list of all genes whose expression is up- or down-regulated in a disease can lead to potential drug targets. A gene that is mutated and causes a disease is probably less useful than a healthy gene whose altered expression or gene products are involved in the disease phenotype.
Today, microarrays are the technology of choice to discover these genes. The principle is simple, but the practice is a combination of art and science [69]. A DNA microarray is a surface of a few square centimeters that contains thousands or even millions of spots of DNA. Every single spot contains a unique probe consisting of single-strand DNA molecules with a unique sequence representing a specific gene. It is of the utmost importance that one knows which spot belongs to which gene. With a spot size of 16 μm, one can put 40,000 spots on a square centimeter, sufficient to probe the whole human transcriptome. Every spot contains 10 million identic single strand DNA molecules with a length of 25 to over 100 nucleotides long. These microarrays (or "gene chips") are manufactured industrially using several different techniques like robotic contact spotting, inkjet deposition, or on-chip synthesis using photolithography (Affymetrix) [69].
To identify which genes are expressed in a specific tissue (e.g., a liver cancer), one proceeds as follows. The mRNA is extracted and converted to cDNA because RNA is less stable. The cDNA can be concentrated or amplified when needed. The cDNA is labeled with a fluorochrome and incubated on the microarray chip for hybridization during several hours. Nonbound material is washed away, the chip is laser-scanned by a robot, and the data are processed, analyzed, and visualized [69–72]. To interpret the results, one should compare the gene expression from the liver cancer cells with the corresponding expression in normal liver cells. The method of choice to do this is to repeat the exercise with the normal cells on a second chip. Until very recently, a convenient alternative was to process the normal cells in parallel with the cancer cells from the beginning and use a different dye, usually green for control cells and red for the cells of interest. The red and green labeled cDNAs are then put together in equal amounts onto the chip. After hybridization and washing, the red spots indicate genes that are more expressed in the cancer cells than in the normal cells, and green spots indicate genes that are less expressed in the cancer cells than in normal cells. Equally expressed spots will be yellow. The problem with this pairwise analysis is that it is less suited for later comparisons due to standardization issues. For this reason, the pairwise analysis is increasingly abandoned.
Advantages.
1. Microarrays allow us to test the whole human transcriptome in one experiment.
2. The technology is now a standard laboratory tool and many services are available.
3. The technology is very versatile and can be used in a broader way than just for comparing the transcriptome. Other applications are chromosome aberrations studies, methylation analysis, single-nucleotide polymorphism detection, toxicogenomics, and diagnostic use. Today 45 million sequences can be probed with one chip.
4. The technology is becoming cheaper, comparable to PCR (Polymerase Chain Reaction).
5. With the "maskless" photolithographic production, probes can be made on-demand without having to interrupt the production process.
6. Limited sample need. For the transcriptome analysis, you need about 20,000 cells. But with a preparation step, one could start from a single cell [73].

Disadvantages.

1. When millions and millions of molecules interact and hybridize with each other you may expect some variation and unexpected things to happen, resulting in false positives and false negatives. For example, the cDNA can form tertiary structures preventing some parts from hybridizing with the probes.
2. Not all mRNAs are equally well extracted, and some mRNA are very unstable.
3. Important pharmacological targets such as GPCRs, ion channels, and transporters have mRNA in low concentrations and need enhanced sensitivity of the microarray to be picked up.
4. The transcriptome doesn't equal the proteome. mRNA can be regulated before translation, and many post-translational regulations and modifications to proteins can make them inactive and hence virtually absent.
5. Some closely related genes can easily hybridize with the probes of their close family members (called cross-hybridization).
6. The need for skilled technical personnel.

Examples.

Potential targets were identified related to asthma, SARS, arthritis, and systemic lupus erythematosus. In the cancer field, clinicians are using DNA microarrays to distinguish between cancers and to predict which treatments will have the most effect [67]. Using protein microarrays, SMIR4, a small-molecule inhibitors of rapamycin, has been found [74].

Variations.

1. The list of variations based on probes, array size, production methods, and applications is endless. Smaller spots allow the production of microarrays with millions of spots to interrogate the single-nucleotide polymorphisms (SNPs). Microarrays for yeast, viruses and bacteria, lab animals, and per chromosome are all available or can be produced.
2. Microarrays loaded with proteins (directly or via antibodies) permit the study of protein−protein interaction, protein−DNA/RNA interactions, and protein−small molecule interactions.
3. Microarrays with living cells attached to the surface by antibodies or proteins are used mainly in cancer research.

H. Comparative Profiling

Concept. A compound is added to a cell. This alters the gene-expression profile, called the "signature." This signature is then compared with a database of signatures obtained from treating cells with known compounds. The underlying idea is that compounds that produce about the same signature have about the same target(s) [45]. Gene-expression is a convenient profile, but the concept of comparative profiling is also applicable to protein expression, toxicity pattern, metabolome expression, morphological aspects, and other aspects.

Input. The gene-expression signature of an uncharacterized compound.

Output. Matching drug−target pairs based on profile similarity, to be further validated.

Group. Indirect. Deconvolution. Profiling. In silico. Genetic.

Description. As an example, we describe gene expression profiling using the Connectivity Map (also known as cmap) developed at the Broad Institute [75−77]. The concept is based on the pioneering work of Hughes in yeast [78]. The first step is to incubate relevant cells or tissues with the compound to be characterized. Cells from treated patients or animals could also be used. Then the complete gene-expression pattern is obtained using micro-arrays that cover the whole genome. The second step is to feed this gene-expression pattern into the cmap website [79] and click query. A pattern-matching algorithm based on the Kolmogorov-Smirnov statistic now compares the profile with all the profiles in the database. At the moment of writing, the so-called build02 database contains over 7,000 transcriptional expression profiles from the treatment of a very limited set of cultured human cell lines with 1,309 small molecules at a limited set of concentrations. As a comparison, build01 contained only 164 compounds. Pattern-matching algorithms can be fine-tuned in myriads of ways, but the cmap website keeps that very limited so that scientists at the bench don't need to be experts in bio-informatics or statistics. The main output of the algorithm is—to put it simply—a list of compounds ranked by how well their induced up- and down-regulated gene-expression mimics that of the test compound. It is important to note that the cells used to obtain the pattern to be searched (the "query signature") should not be from the same cell line and not even from the same species as the ones used to build the database. Of course, only genes that are present and expressed in both cell types can be compared. This is less a shortcoming than one would suspect, because many genes are up- or down-regulated in clusters, and missing one or two genes still permits finding the overall matching patterns. On the other hand, a compound like estrogen will not change the expression profile when the estrogen receptor is absent in the cell (Figure 3.9).

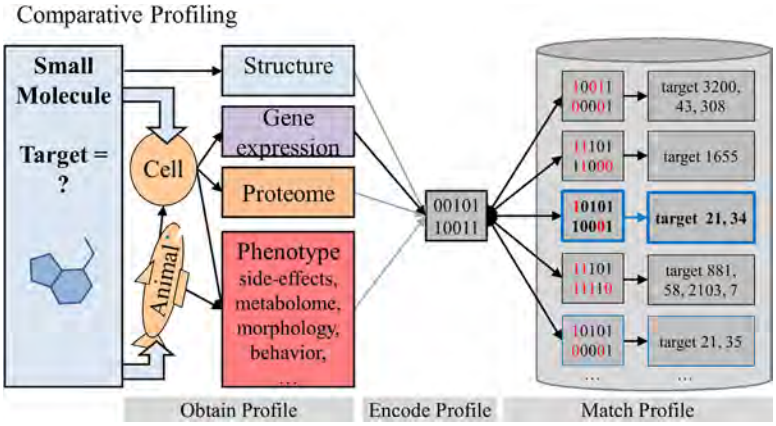


FIGURE 3.9 A method to infer the target of a compound is to first get a profile of a compound. A profile can be anything, such as the structure of the compound, the behavioral impact in an animal experiment, or the gene expression change in a cell. It is important that one uses a database containing the profiles of many other compounds together with their target(s). The profiles are encoded in such a way that matching algorithms can find the profiles in the database that most resembles the profile of the drug. The underlying hypothesis is that compounds with the same profile would also have the same target.

Requirements. Comparative profiling is only possible when there is something with which to compare. One has to use of a database that contains the signatures of many known compounds and a pattern-matching algorithm.
Advantages.
1. The availability of freely accessible profiling databases growing over time.
2. Web-based, easy-to-use interface.
3. The test compound is used in its normal state in a normal cell.
4. Multiple effects are measured, not just the binding with one single protein.

Disadvantages.
1. The output is not a guaranteed target but rather a proposal of possible targets or pathways. And when an expression match is found with a compound with unknown targets, you still have no target clue.
2. The cells used to build the database are special human cell lines, living on plastic, instead of primary cells living in the body.
3. Targets or pathways that were never hit before will not result in a match.

Examples.
Tools: the COMPARE NCI60 analysis [80], the JFCR39 database [81], the Connectivity Map database developed at the Broad Institute [75,76]. Drug-targets identified [45]: gedunin-HSP90, droxinostat-HDAC, bisebromoamide-actin, glucopiericidin A-glucotransporter, BNS22-topoisomearse II, Theonellamide F-3b-Hydroxysterols.
Case study.
The group of Wei and Armstrong [82] was confronted with the resistance to glucocorticoid treatment of childhood acute lymphoblastic leukemia (ALL), resulting in a poor prognosis. Comparing gene-expression profiles obtained from sensitive and resistant patients was not really helpful, but querying the Connectivity Map database pointed to the connection between glucocorticoid sensitivity and rapamycin. Rapamycin, a mTOR inhibitor, is an FDA-approved immunosuppressant, and it would be very promising if this drug could restore the sensitivity of the cancer to dexamethasone. Wei could prove this concept to work in cells. Another group showed that mTOR inhibitors are synergistic with methotrexate in mice [83]. Promising *in vitro* results with temsirolimus, another mTOR inhibitor, in prostate cancer didn't translate into a positive clinical outcome [84].
Variations.
1. Other databases, other algorithms, limited gene sets, other techniques to measure the gene expression. Example: the Luminex 1000 Profiling Approach [85].
2. Comparative profiling of protein-expression patterns.
3. Comparative profiling of side effects.
4. Comparative profiling of the metabolome.
5. Comparative profiling of the morphology of the cell.
6. Comparative profiling of the structure of the compound.

I. Analysis of the Pathophysiology

Concept. The detailed analysis of the pathophysiology could reveal new interesting targets.
Input. A disease of interest.
Output. Potential targets, pathways, or active compounds. Insight.

Group. Target discovery. Conventional. Top down.
Description. It sounds logical and scientific that the way to find a drug-target is the detailed study of a disease, unraveling and exploring the consecutive layers, starting from the macroscopic manifestations, drilling down to changes in tissues and cells, and ending with characterizing the molecular culprits. However, most of the time there is no need to travel down the entire way, nor to understand every element on the path to find a drug. When George Hitching and Gertrude Elion first started to synthesize analogs of purine to block DNA synthesis, the role of DNA was still uncertain. They knew that the purine metabolite was needed by bacteria to produce DNA, and so false purines could perhaps block the enzymes and function as "antimetabolites." In general, studying the functions and interactions of endogenous small molecules—metabolites, hormones, neurotransmitters, cytokines—is very rewarding. Formulating hypotheses, even when they turn out decades later to be wrong, is another useful approach. Finding targets that could be used to modify or counteract the consequences of a disease—instead of finding the cause of a disease—is often more fruitful, because removing the cause remains even today an almost unattainable goal. Infectious diseases are the biggest exceptions to this rule. Cancer is already less an exception. The reason is that in the case of bacteria, the complete organism can be put in a test tube. When a compound is able to kill the bacterium in a test tube, there is some chance that it will maintain that capacity in our body. Of course, the drug has to reach the bacterium in our body. For cancer, the situation is more complex. A single cancer cell is not the same thing as a cancer tissue with all its different cell types. Blood circulation is different in cancer tissues. Oxygen levels and acidity could be different. Analysis of the pathophysiology is a top-down, classic strategy that still works well today and certainly has a bright future with all the new molecular and imaging research tools available.
Advantages.
1. New targets and pathway can be found.
2. No need to start from an active molecule, although starting from endogenous molecules is often key.
3. Proven track record.
Disadvantages.
Takes a long time without guaranteed success. The war on cancer started decades ago.
Examples. Some targets identified by analysis of the pathophysiology are many viral enzymes like reverse transcriptase, dihydrofolate reductase, 5HT transporter, COX-2, cysteinyl leukotriene receptor, angiotensin-converting enzyme, estrogen receptor, and dopamine receptor.

J. The Study of Existing Drugs

Concept. The thorough study of "old" and even more recent drugs with the newest technologies and insights can reveal new potential targets [44].
Input. An existing drug.
Output. The drug-target or additional knowledge on the mechanism of action and affected pathways, including new subtypes of receptors or potential new targets.
Group. Target discovery. Target deconvolution. Conventional. Top down.
Description. Because in the past most drugs were discovered via phenotypic screening, a lot of drugs were on the to-do list for target deconvolution. Consequently many new targets were discovered this way. Since the decline in phenotypic screening in favor of target-based screening, the to-do list has shrunk. The recent revival of phenotypic screening and the screening of exotic (natural) molecules will further sustain the need for target deconvolution. The study of toxic effects of drugs can also lead to new targets or to deeper knowledge of existing target effects. Recently [86] a group revealed that COX-2 inhibition in the cardiomyocytes in mice resulted in enhanced susceptibility to induced arrhythmogenesis. The study of COX-2 also led to the discovery of the COX-3 enzyme, a splice variant which was subsequently found to be nonactive in humans.
Examples. $GABA_A$ receptor, COX enzymes, tubulin, L-type calcium channels, KATP channels, dopamine D2 receptor, PPARa, monoamine transporters.

K. Systems Biology

Concept. The study of the dynamic interactions between all components of a biological system.
Input. All kinds of so-called -omics data sets [10]. Modelling software.
Output. A better understanding of the role and value of specific targets in their biological context.

Group. In silico methods.
Description. To understand and predict the behavior of an organism, one should study not only all its individual parts in a reductionistic way but also their dynamic interactions in an integrated holistic perspective. The concept of systems biology is certainly not new. Simulating the dynamic interaction between two components of an axon (a sodium and a potassium channel) to explain how a nerve signal emerged was published in 1952 [87]. Today the number of components in the actual datasets is huge: 29,000 transcriptomic elements, 20–30 million epigenetic elements, 22,000–39,000 proteins (each present in a multitude of different states, like phosphorylation states) and around 40,000 small cellular biomolecules [10]. The next step is to collect the interactions between these components. The BioGRID database contains over 139,000 nonredundant interactions between over 18,000 human proteins. Then comes the software and the mathematical models to make predictions. As the models are fitted with the actual data, they are usually good in predicting what we know already. Most of the targets proposed so far using systems biology were not easily druggable and provoked a renewed interest in phenotypic screening. Systems biology is, however, an ideal working instrument to increase our understanding of the mechanism of actions of drugs in the body.
Example. Pyrvinium [10].

L. In Silico Simulation of the Human Patient

Concept. The ultimate goal of systems biology is the understanding of functioning of a whole organism [10]. Understanding in this field is achieved by modeling or simulation. We are, however, so far from the complete simulation of a patient that I think it is appropriate to address this topic with a sense of humor.
Input. A demand for a safe, effective, inexpensive oral drug against disease x.
Output. After pushing the button, the chemical formula of a perfect drug rolls out of the computer.
Group. Futuristic methods.
Requirements. An enormous dose of luck and a pile of quantumcomputers.
Additional comments. Trying to simulate a human patient in silico would redirect funding for drug research to the informatics industry. But one should be realistic. A human body contains about 10 trillion interacting cells and every cell contains about 23 trillion molecules (of which about 8 billion are proteins) that interact with one another. Today it's very hard to predict the behavior of one single small molecule in a very confined part of a protein. We can't predict consistently the structure of a protein out of the gene sequence. We can't consistently predict the function of a protein. Today we can't even predict consistently the interaction between one small molecule and a bunch of water molecules—called dissolution. One could say that we don't have to simulate all the billions of molecules to predict the overall behavior. Yet, a change in one molecule in one single cell can lead to the death of a person within a year. (Figure 3.10).

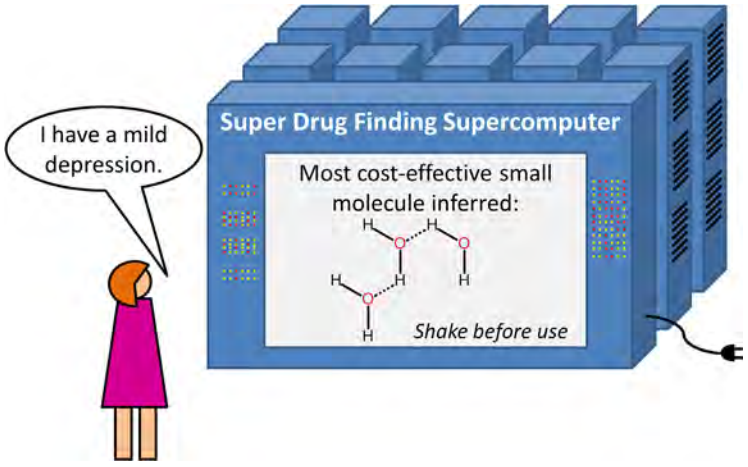


FIGURE 3.10 The complete simulation of a human patient at work.

VII. TARGET VALIDATION

It can't be stressed enough that the only validation that really matters is the proof in the clinic. It adds, of course, confidence that a target seems to work in dozens of different cell-types and in several animal models, but at the end this is all ersatz compared to the ultimate test: does a compound that effectively acts on the target really work in humans?

Further, target validation is not different from the usual scientific process. When an interesting target is described in a publication, the first step is to reproduce the experiments. Although not well studied, this reproducibility turns out to be a bit problematic. Florian Prinz examined 67 projects—mostly in oncology—and found that in only 20–25 percent of them the in-house findings were completely in line with the relevant published data [15,88].

The next step is to make variations in one of the three elements in the ligand–target–environment equation. Modulating the affinity of the small molecule for the target protein should correlate with a modulated activity of the molecule. Mutations in the binding domain of the protein should result in modulation or loss of activity of the ligand. Changing the cell-type should or should not change the effect of the molecule. As already stated, the use of siRNA is an invaluable tool for target validation. It can be used in cell experiments and animal models to study the effect of inhibiting or blocking a target without the need to have a small molecule.

As an example we continue with the thalidomide story. In 2010, Ito [52] identified cereblon as a target protein for thalidomide's teratogenicity using affinity chromatography. How did he validate the target? One way was to go *in vivo*. Thalidomide is not teratogenic in mice and rats. But what about in zebrafish? Thalidomide given to zebrafish embryos resulted in disturbed pectoral fins. Zebrafish have a cereblon orthologous gene, zcrbn, and its protein could also be affinity-extracted with thalidomide. Then Ito's team blocked the gene with antisense constructs, resulting in the same teratogenic effects as thalidomide. Injection of correct mRNA for the gene rescued the defects. They then made a mutant zCrbn protein that could not bind to thalidomide but was still functioning. Overexpression of this mutant made the embryos insensitive to thalidomide.

VIII. CONCLUSION

In the last decades, many successful new small molecules and biologicals have enriched our therapeutic capabilities. Yet despite all the new fantastic technologies, we have not seen an explosion in R&D efficacy. The small molecules that came out of the target-based screening strategy often missed efficacy or showed unexpected toxicity in clinical trials. Today, the follower strategy, based on targets validated in humans, is by far the most successful strategy measured by the number of drugs it produces. The second most productive strategy is phenotypic screening, which is probably even more successful than has been acknowledged. The renewed interest in phenotypic screening and the increased capabilities to handle thousands of animals like zebrafish in an automated way will result in an increased demand for target deconvolution. The methods for target deconvolution will remain a rapidly evolving and highly entangled landscape. The ultimate target validation method, however, will still be the human clinical trial. Chemists should think critically about targets that were identified based solely on highly reductionistic methods. Otherwise, they might blindly chase for years the perfect compound for the wrong target.

References

[1] Imming P, Sinning C, Meyer A. Drugs, their targets, and the nature and number of drug targets. Nat Rev Drug Discov 2006;5(10):821–34.
[2] Chen XP, Du GH. Target validation: a door to drug discovery. Drug Discoveries Ther 2007;1(1):23–9.
[3] Kubinyi H. Drug research: myths, hype, and reality. Nat Rev Drug Discov 2003;2(8):665–8.
[4] Boran AD, Iyengar R. Systems approaches to polypharmacology and drug discovery. Curr Opin Drug Discov Dev 2010;13(3):297–309.
[5] Knowles J, Gromo G. A guide to drug discovery: target selection in drug discovery. Nat Rev Drug Discov 2003;2(1):63–9.
[6] Kotz J. Phenotypic screening, take two. Sci-Bus Exchange 2012;5(15):1–3.
[7] Butcher EC. Can cell systems biology rescue drug discovery? Nat Rev Drug Discov 2005;4(6):461–7.
[8] Hill R, Rang HP. Drug discovery and development. 2nd ed. Edinburgh: Elsevier; 2012.
[9] Lindsay MA. Finding new drug targets in the 21st century. Drug Discov Today 2005;10(23–24):1683–7.
[10] Berg EL. Systems biology in drug discovery and development. Drug Discov Today 2014;19(2):113–25.

[11] Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. Nat Rev Drug Discov 2012;11(3):191–200.
[12] Swinney DC, Anthony J. How were new medicines discovered? Nat Rev Drug Discov 2011;10(7):507–19.
[13] Sams-Dodd F. Target-based drug discovery: is something wrong? Drug Discov Today 2005;10(2):139–47.
[14] Gabor Miklos GL. The human cancer genome project—one more misstep in the war on cancer. Nat Biotechnol 2005;23(5):535–7.
[15] Prinz F, Schlange T, Asadullah K. Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov 2011;10(9):712.
[16] Arrowsmith J. Trial watch: Phase II failures: 2008–2010. Nat Rev Drug Discov 2011;10(5):328–9.
[17] Zheng W, Thorne N, McKew JC. Phenotypic screens as a renewed approach for drug discovery. Drug Discov Today 2013;18 (21–22):1067–73.
[18] Arrowsmith J, Miller P. Trial watch: Phase II and phase III attrition rates: 2011–2012. Nat Rev Drug Discov 2013;12(8):569.
[19] Eggert US. The why and how of phenotypic small-molecule screens. Nat Chem Biol 2013;9(4):206–9.
[20] Vane JR, Botting RM. The mechanism of action of aspirin. Thromb Res 2003;110(5–6):255–8.
[21] Pert CB, Snyder SH. Opiate receptor: demonstration in nervous tissue. Science 1973;179(4077):1011–14.
[22] Garcia PS, Kolesky SE, Jenkins A. General anesthetic actions on GABA(A) receptors. Curr Neuropharmacol 2010;8(1):2–9.
[23] Stockwell BR. Chemical genetics: ligand-based discovery of gene function. Nat Rev Genet 2000;1(2):116–25.
[24] Schreiber SL, Crabtree GR. The mechanism of action of cyclosporin A and FK506. Immunol Today 1992;13(4):136–42.
[25] Svenningsson P, Kim Y, Warner-Schmidt J, Oh YS, Greengard P. p11 and its role in depression and therapeutic responses to antidepressants. Nat Rev Neurosci 2013;14(10):673–80.
[26] Flight MH. Mood disorders: room for improvement. Nat Rev Drug Discov 2013;12(8):578–9.
[27] Gulbins E, Palmada M, Reichel M, et al. Acid sphingomyelinase-ceramide system mediates effects of antidepressant drugs. Nat Med 2013;19(7):934–8.
[28] Kokel D, Bryan J, Laggner C, et al. Rapid behavior-based identification of neuroactive small molecules in the zebrafish. Nat Chem Biol 2010;6(3):231–7.
[29] Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol 2003;48 (1):77–84.
[30] Liang TJ, Ghany MG. Current and future therapies for hepatitis C virus infection. N Engl J Med 2013;368(20):1907–17.
[31] Andries K, Verhasselt P, Guillemont J, et al. A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis. Science 2005;307(5707):223–7.
[32] Petit S, Coquerel G, Meyer C, Guillemont J. Absolute configuration and structural features of R207910, a novel anti-tuberculosis agent. J Mol Struct 2007;837(1–3):252–6.
[33] Smith AM, Ammar R, Nislow C, Giaever G. A survey of yeast genomic assays for drug and target discovery. Pharmacol Ther 2010;127 (2):156–64.
[34] McGrath NA, Brichacek M, Njardarson JT. A graphical journey of innovative organic architectures that have improved our lives. J Chem Educ 2010;87(12):1348–9.
[35] Eckford PD, Li C, Ramjeesingh M, Bear CE. Cystic fibrosis transmembrane conductance regulator (CFTR) potentiator VX-770 (ivacaftor) opens the defective channel gate of mutant CFTR in a phosphorylation-dependent but ATP-independent manner. J Biol Chem 2012;287 (44):36639–49.
[36] Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006;5(12):993–6.
[37] Rask-Andersen M, Almen MS, Schioth HB. Trends in the exploitation of novel drug targets. Nat Rev Drug Discov 2011;10(8):579–90.
[38] Agarwal P, Sanseau P, Cardon LR. Novelty in the target landscape of the pharmaceutical industry. Nat Rev Drug Discov 2013;12 (8):575–6.
[39] <http://www.brainyquote.com/quotes/quotes/a/arthurccl100793.html> [accessed 01.11.13].
[40] Crunkhorn S. Trial watch: success in amyloidosis trials supports potential of systemic RNAi. Nat Rev Drug Discov 2013;12(11):818.
[41] Mullard A. Make or break for first splice-modulating agents. Nat Rev Drug Discov 2013;12(11):813–15.
[42] Terstappen GC, Schlupen C, Raggiaschi R, Gaviraghi G. Target deconvolution strategies in drug discovery. Nat Rev Drug Discov 2007;6 (11):891–903.
[43] Schenone M, Dancik V, Wagner BK, Clemons PA. Target identification and mechanism of action in chemical biology and drug discovery. Nat Chem Biol 2013;9(4):232–40.
[44] Titov DV, Liu JO. Identification and validation of protein targets of bioactive small molecules. Bioorg med chem 2012;20(6):1902–9.
[45] Tashiro E, Imoto M. Target identification of bioactive compounds. Bioorg Med Chem 2012;20(6):1910–21.
[46] Chan JN, Nislow C, Emili A. Recent advances and method development for drug target identification. Trends Pharmacol Sci 2010;31 (2):82–8.
[47] Cong F, Cheung AK, Huang SM. Chemical genetics-based target identification in drug discovery. Annu Rev Pharmacol Toxicol 2012;52:57–78.
[48] Stockwell BR. Exploring biology with small organic molecules. Nature 2004;432(7019):846–54.
[49] Ziegler S, Pries V, Hedberg C, Waldmann H. Target identification for small bioactive molecules: Finding the needle in the haystack. Angew Chem 2013;52(10):2744–92.
[50] Lindsay MA. Target discovery. Nat Rev Drug Discov 2003;2(10):831–8.
[51] Rix U, Superti-Furga G. Target profiling of small molecules by chemical proteomics. Nat Chem Biol 2009;5(9):616–24.
[52] Ito T, Ando H, Suzuki T, et al. Identification of a primary target of thalidomide teratogenicity. Science 2010;327(5971):1345–50.
[53] Sheskin J. Thalidomide in the treatment of lepra reactions. Clin Pharmacol Ther 1965;6:303–6.
[54] Ito T, Ando H, Handa H. Teratogenic effects of thalidomide: molecular mechanisms. CMLS, Cell Mol Life Sci 2011;68(9):1569–79.
[55] Ito T, Handa H. Deciphering the mystery of thalidomide teratogenicity. Congenit Anom (Kyoto) 2012;52(1):1–7.

[56] Graves PR, Kwiek JJ, Fadden P, et al. Discovery of novel targets of quinoline drugs in the human purine binding proteome. Mol Pharmacol 2002;62(6):1364–72.
[57] Walke DW, Han C, Shaw J, Wann E, Zambrowicz B, Sands A. In vivo drug target discovery: identifying the best targets from the genome. Curr Opin Biotechnol 2001;12(6):626–31.
[58] Wang S, Sim TB, Kim Y-S, Chang Y-T. Tools for target identification and validation. Curr Opin Chem Biol 2004;8(4):371–7.
[59] Li LC, Okino ST, Zhao H, et al. Small dsRNAs induce transcriptional activation in human cells. Proc Natl Acad Sci U S A 2006;103 (46):17337–42.
[60] Check E. RNA interference: hitting the on switch. Nature 2007;448(7156):855–8.
[61] Giaever G, Flaherty P, Kumm J, et al. Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. Proc Natl Acad Sci U S A 2004;101(3):793–8.
[62] Ho CH, Piotrowski J, Dixon SJ, Baryshnikova A, Costanzo M, Boone C. Combining functional genomics and chemical biology to identify targets of bioactive compounds. Curr Opin Chem Biol 2011;15(1):66–78.
[63] de Jonge MR, Koymans LH, Guillemont JE, Koul A, Andries K. A computational model of the inhibition of Mycobacterium tuberculosis ATPase by a new drug candidate R207910. Proteins 2007;67(4):971–80.
[64] Jones SW, Souza PM, Lindsay MA. siRNA for gene silencing: a route to drug target discovery. Curr Opin Pharmacol 2004;4(5):522–7.
[65] Stynen B, Tournu H, Tavernier J, Van Dijck P. Diversity in genetic in vivo methods for protein–protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. Microbiol Mol Biol Rev 2012;76(2):331–82.
[66] Fields S, Song O. A novel genetic system to detect protein–protein interactions. Nature 1989;340(6230):245–6.
[67] Jayapal M, Melendez AJ. DNA microarray technology for target identification and validation. Clin Exp Pharmacol Physiol 2006;33 (5–6):496–503.
[68] Leroy Q, Raoult D. Review of microarray studies for host-intracellular pathogen interactions. J Microbiol Methods 2010;81(2):81–95.
[69] Dufva M. DNA microarrays for biomedical research: methods and protocols. New York: Humana Press; 2009.
[70] Göhlmann H, Talloen W. Gene expression studies using affymetrix microarrays. Boca Raton: Taylor & Francis; 2009.
[71] Wouters L, Gohlmann HW, Bijnens L, Kass SU, Molenberghs G, Lewi PJ. Graphical exploration of gene expression data: a comparative study of three multivariate methods. Biometrics 2003;59(4):1131–9.
[72] Lewi PJ. Spectral mapping, a personal and historical account of an adventure in multivariate data analysis. Chemom Intell Lab Syst 2005;77(1–2):215–23.
[73] Kurimoto K, Saitou M. Single-cell cDNA microarray profiling of complex biological processes of differentiation. Curr Opin Genet Dev 2010;20(5):470–7.
[74] Huang J, Zhu H, Haggarty SJ, et al. Finding new components of the target of rapamycin (TOR) signaling network through chemical genetics and proteome chips. Proc Natl Acad Sci U S A 2004;101(47):16594–9.
[75] Lamb J. The Connectivity Map: a new tool for biomedical research. Nat Rev Cancer 2007;7(1):54–60.
[76] Lamb J, Crawford ED, Peck D, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science 2006;313(5795):1929–35.
[77] Lamb JG., T. Science in action: the connectivity map. <http://www.youtube.com/watch?v = W_bZOGL3ZbE>; 2013 [accessed 01.11.13].
[78] Hughes TR, Marton MJ, Jones AR, et al. Functional discovery via a compendium of expression profiles. Cell 2000;102(1):109–26.
[79] Connectivity Map 0.2., <http://www.broadinstitute.org/cmap/>; 2006 [accessed 01.11.13].
[80] Stefely JA, Palchaudhuri R, Miller PA, et al. N-((1-benzyl-1H-1,2,3-triazol-4-yl)methyl)arylamide as a new scaffold that provides rapid access to antimicrotubule agents: Synthesis and evaluation of antiproliferative activity against select cancer cell lines. J Med Chem 2010;53(8):3389–95.
[81] Yaguchi S, Fukui Y, Koshimizu I, et al. Antitumor activity of ZSTK474, a new phosphatidylinositol 3-kinase inhibitor. J Natl Cancer Inst 2006;98(8):545–56.
[82] Wei G, Twomey D, Lamb J, et al. Gene expression-based chemical genomics identifies rapamycin as a modulator of MCL1 and glucocorticoid resistance. Cancer cell 2006;10(4):331–42.
[83] Teachey DT, Sheen C, Hall J, et al. mTOR inhibitors are synergistic with methotrexate: an effective combination to treat acute lymphoblastic leukemia. Blood 2008;112(5):2020–3.
[84] Armstrong AJ, Shen T, Halabi S, et al. A phase II trial of temsirolimus in men with castration-resistant metastatic prostate cancer. Clin Genitourin Cancer 2013;11(4):397–406.
[85] Peck D, Crawford ED, Ross KN, Stegmaier K, Golub TR, Lamb J. A method for high-throughput gene expression signature analysis. Genome Biol 2006;7(7):R61.
[86] Wang D, Patel VV, Ricciotti E, et al. Cardiomyocyte cyclooxygenase-2 influences cardiac rhythm and function. Proc Natl Acad Sci U S A 2009;106(18):7548–52.
[87] Hodgkin AL, Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. J Physiol 1952;117(4):500–44.
[88] Haibe-Kains B, El-Hachem N, Birkbak NJ, et al. Inconsistency in large pharmacogenomic studies. Nature 2013.

SECTION TWO

Lead Compound Discovery Strategies

CHAPTER

4

Strategies in the Search for New Lead Compounds or Original Working Hypotheses

Camille G. Wermuth[1], *Bruno Villoutreix*[2], *Serge Grisoni*[3], *Anne Olivier*[4] *and Jean-Philippe Rocher*[5]

[1]Founder of Prestwick Chemical, Strasbourg, France; [2]Université Paris Diderot Sorbonne Paris Cité, Inserm, Paris, France; [3]Pierre Fabre Research Institute, Toulouse, Cédex, France; [4]Sanofi-Aventis Recherche et Développement, Chilly-Mazarin, France; [5]Pierre Fabre Research Institute, Castres, France

OUTLINE

So ist denn in der Strategie alles sehr einfach, aber darum nicht auch alles sehr leicht. (Thus in the strategy everything is very simple, but not necessarily very easy). ***Carl von Clausewitz*** [1]

I. INTRODUCTION

This chapter deals with the various strategies leading to active compounds and active compounds collections. The introduction of modern biology methods and technologies has driven the discovery process to a target-based

approach. However, the discovery strategies leading to new drugs may also still address systems-based approaches such as the phenotypic screening methods [2]. The primary objective is to identify original and attractive starting points for therapeutic discovery programs. Such programs typically begin with the search for "hits."

A. Hits and Leads

A hit is an active substance having a preferential activity for the target and which satisfies all of the following criteria [3]: (1) reproducible activity in a relevant bioassay, (2) confirmed structure and high purity, (3) specificity for the target under study, (4) confirmed potential for novelty, and (5) a chemically tractable structure, that is, molecules presenting a certain affinity for a target.

Identifying hits for a new target usually involves screening of a wide range of structurally diverse small molecules in an *in vitro* bioassay. Alternatively, small molecules can be screened for their potential to modulate a biological process thought to be critical in disease or in which the target is thought to play a major role. Miniaturization and robotics means that the number of compounds that can be screened has greatly increased, and several thousand compounds can be screened in one day.

Once a hit is discovered, its activity must be confirmed and validated. Typical hit validation criteria are as follows: (1) the hit must be active *in vitro* and be amenable to *in vivo activity in target or disease models*; (2) the hit should not display human ether-a-go-go-related (hERG) toxicity; (3) the analogs of the hit must display clear structure−activity relationships (SAR); (4) basic physicochemical and ADME properties of the hit series must be evaluated in order to identify potential unwanted properties in the series and to assess structure−properties relationships (SPR); (5) the hit should not contain chemically reactive functions; and (6) the hit must provide patent opportunities. Only then does it becomes a lead substance, commonly named "lead."

If a lead molecule emerges from these additional studies on SAR, absorption, distribution, metabolism, excretion (ADME), and toxicity, it acquires the "clinical drug candidate" status. After a short toxicological study, it fulfills the criteria required for administration to humans for initial clinical studies.

B. The Main Hit or Lead Finding Strategies

A retrospective analysis of the ways leading to discovery of new drugs suggests that five successful strategies can yield new hits and/or lead compounds [4,5]. The first strategy is based on the modification and improvement of already existing active molecules. The second one consists of the systematic screening of sets of arbitrarily chosen compounds on selected biological assays. The third approach resides in the retroactive exploitation of various pieces of biological information that sometimes result from new discoveries made in biology and medicine, and sometimes are just the fruits of more or less fortuitous observations. The fourth route to new active compounds is a rational design based on the knowledge of the molecular cause of the pathological dysfunction. The fifth strategy is based on the structural knowledge of the target combined with computational methods or biophysical technologies of ligand−protein interaction.

II. FIRST STRATEGY: ANALOG DESIGN

The most popular strategy in drug design is the synthesis of analogs of existing active molecules. The objective is to start with known active principles or "first-in-class" drugs and—by various chemical transformations—prepare new molecules (sometimes referred to as "follow-on" or "me-too compounds") for which an increase in potency, a better specific activity profile, improved safety, or a better formulation that is easier-to-handle by physicians and nurses or more acceptable to the patient are claimed.

A. Typical Examples

A typical illustration of this approach is found in the series of losartan analogs (Figure 4.1) or in the conazole series (Figure 4.2). All the compounds show similar structures and similar affinity for the angiotensin II receptor. As such they can be considered "full" analogs.

In the pharmaceutical industry, motivations for analog design are often driven by competitive and economic factors. Indeed, if the sales of a given medicine are high and the company is found in a monopolistic situation, protected by patents and trademarks, other companies will want to produce similar medicines, with some

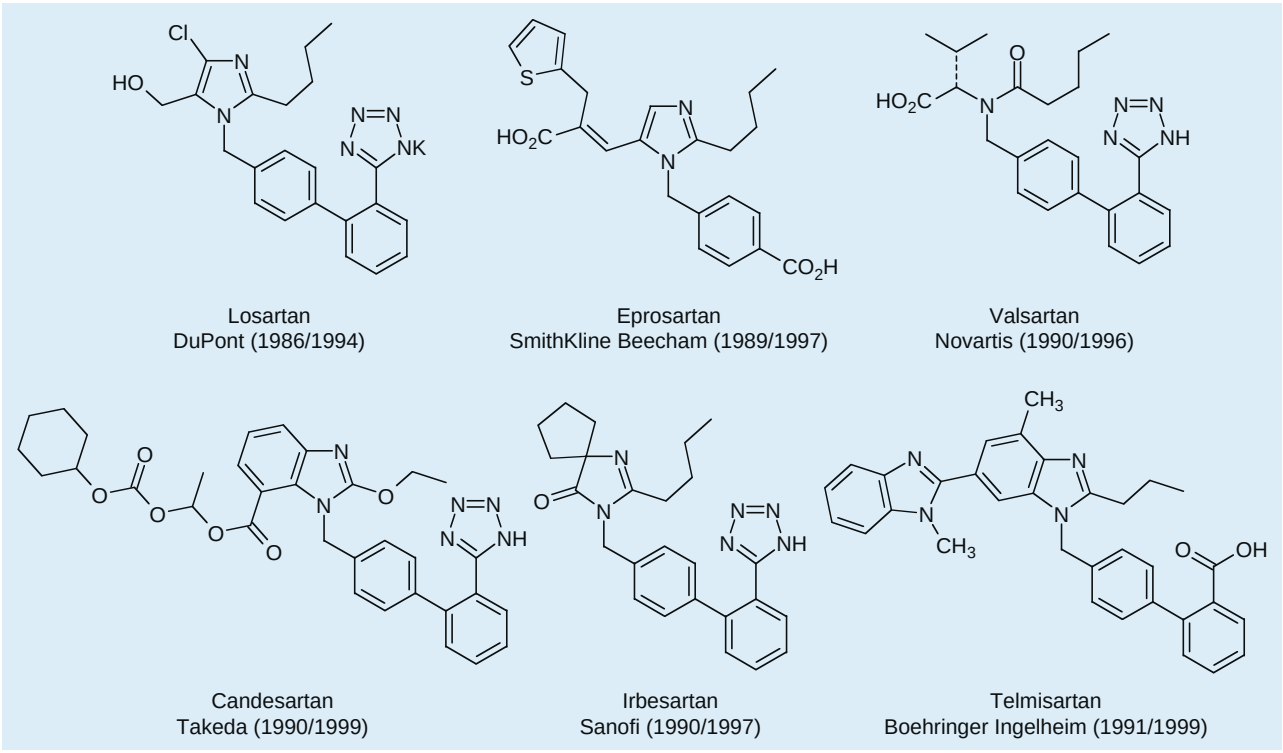


FIGURE 4.1 Angiotensin AT1 receptor antagonists derived from losartan. Despite their structural similarity, it can be assumed that the corresponding discoveries were made independently. The first year under parentheses is the basic patent year; the second one is the year of the first launch.

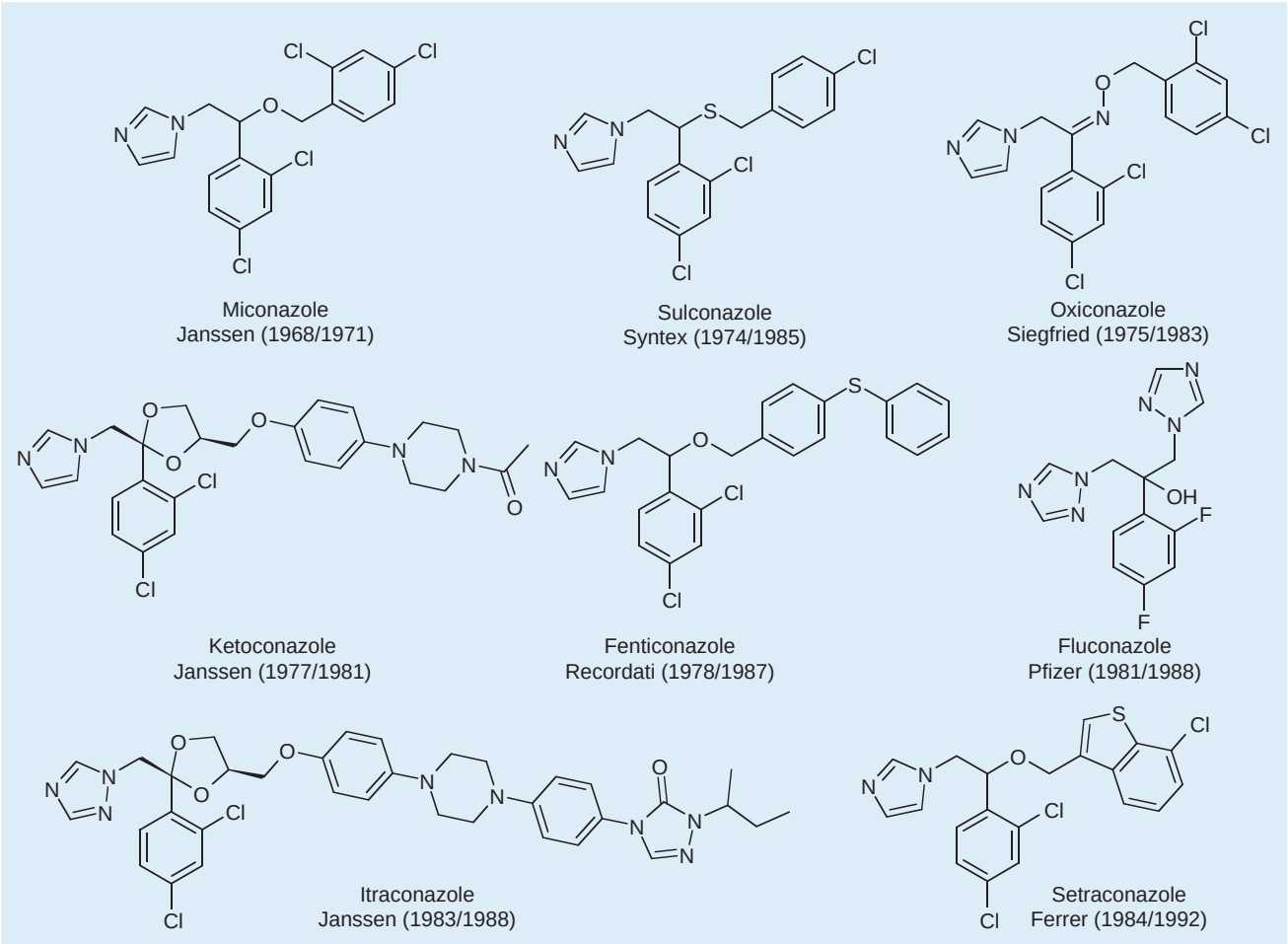


FIGURE 4.2 An example of me-too compounds (full analogs) is given by micoconazole-derived fungistatics, which act by inhibition of the ergosterol biosynthesis. The first year under parentheses is the basic patent year; the second one is the year of the first launch.

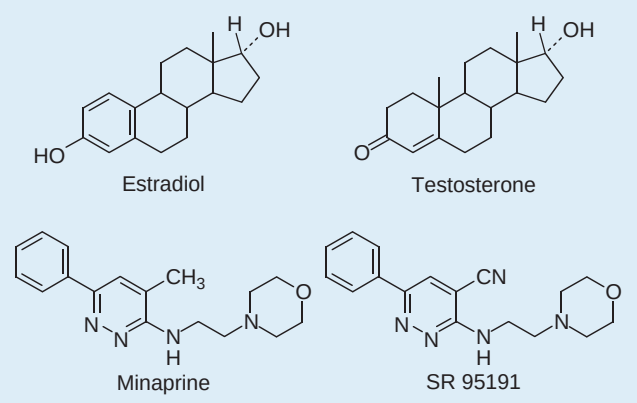


FIGURE 4.3 Some examples of structural analogs. Despite their structural analogy, these compounds present different pharmacological activities.

therapeutic improvements if possible. They will therefore use the already commercialized drug as a lead compound and search for ways to modify its structure and some of its physical and chemical properties while retaining or improving its therapeutic properties.

B. The Different Categories of Analogs

The term analogy, derived from the Latin and Greek *analogia*, has been used in natural sciences since 1791 to describe structural *and* functional similarity [6]. Extended to drugs, this definition implies that the analog of an existing drug molecule shares chemical and therapeutical similarities with the original compound. This definition suggests three categories of drug analogs: (1) analogs presenting chemical and pharmacological similarity; (2) analogs presenting only chemical similarity; and (3) analogs displaying similar pharmacological properties but presenting totally different chemical structures.

Analogs of the first category, presenting at the same time chemical and pharmacological similarities, can be considered as "full" or "true" analogs (Figures 4.1 and 4.2). These analogs correspond to the class of drugs often referred to as "me-too compounds." Usually, they are improved versions of a pioneer drug over which they present a pharmacological, pharmacodynamic, or biopharmaceutical advantage. Other examples are the angiotensin-converting enzyme (ACE) inhibitors derived from captopril, the histamine H_2 antagonists derived from cimetidine, and the hydroxymethyl-glutaryl-CoA reductase (HMG-CoA reductase) inhibitors derived from mevinolin. Such analogs are designed for industrial and marketing reasons with the same justifications as those which are valid for any other industrial products such as laptop computers or automobiles.

The second class of analogs, which are made of chemically similar molecules and for which we propose the term "structural analogs," contain compounds originally prepared as close and patentable analogs of a novel lead, but for which the biological assays revealed totally unexpected pharmacological properties. A historical example of the emergence of a new activity is provided by the discovery of the antidepressant properties of imipramine, which was originally designed as an analog of the potent neuroleptic drug chlorpromazine. Observation of an "emergent" activity can be purely fortuitous or can result from a voluntary and systematic investigation.

Another example, illustrating that chemical similarity does not necessarily mean biological similarity, is found for steroid hormones. Testosterone and progesterone, although being chemically very close, have very different biological functions (Figure 4.3). Similarly, minaprine is a dopaminergic drug, whereas its cyano analog SR 95191 is a potent MAO-A inhibitor [7].

For the third class of analogous compounds, chemical similarity is not observed, but they share common biological properties. We propose the term "functional analogs" for such compounds. Examples are the neuroleptics chlorpromazine and haloperidol or the tranquillizers diazepam and zopiclone (Figure 4.4). Despite different chemical structures, they show similar affinities for the dopamine and the benzodiazepine receptors, respectively. The design of such drugs is presently facilitated, thanks to virtual screening of large libraries of diverse structures.

C. Pros and Cons of Analog Design

Analog design lacks originality and has often been a source of criticism of the pharmaceutical industry [10]. Each laboratory wants to have its own antiulcer drug, its own antihypertensive, etc. These drug copies are called "me-too

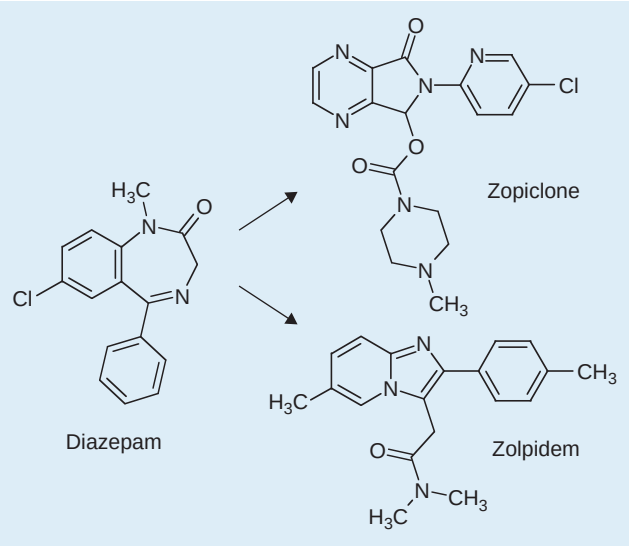


FIGURE 4.4 Zopiclone and zolpidem are selective benzodiazepine receptor agonists not related chemically to benzodiazepines [8,9].

products." Generally, the owner firm of the original drug continues to prepare new analogs, to ensure both a maximum perimeter of protection of its patents and to remain the leader in a given area. For these reasons, the chemical transformation of known active molecules constitutes the most widespread practice in the pharmaceutical research.

A key driver is the certainty to deliver an active drug with the desired therapeutic value. It is indeed extremely rare, and practically improbable, that a given biological activity is unique to a single molecule. Molecular modifications allow the preparation of additional products for which one can expect, if the investigation has been sufficiently prolonged, a comparable activity to that of the copied model—perhaps even a better one. This factor is comforting for the copiers as well as for the financiers that subsidize them. It is necessary, however, to keep in mind that the original inventor of a new medicine possesses a technological and scientific advantage over the copier and that he too was able to design a certain number of copies of his own compound before he selected the molecule insuring the best compromises between activity, secondary effects, toxicity, and invested money.

A second element favoring the copy derives from the information already gained, which then facilitates subsequent pharmacological and clinical studies. As soon as the pharmacological models that served to identify the activity profile of a new prototype are known, it suffices to apply them to the therapeutic copies. In other terms, the pharmacologist will know in advance what kind of activity they desire and which tests they will have to apply to select the wanted activity profile. In addition, during clinical studies, the original research undertaken with the lead compound will serve as a reference and can be transposed as unchanged to the evaluation of the copy. Criticism of this approach is a result of the obvious fact that, in selecting a new active molecule by means of the same pharmacological models as were used for the original compound, one will inevitably end with a compound presenting a nearly identical activity profile and thus the innovative character of such research is greatly diminished.

Finally, financial arguments may play in favor of the therapeutic copy. Thus, it may be important—and even vital—for a pharmaceutical company or for a national industry to have its own drugs rather than to subcontract a license. Indeed, in paying dues of license, an industry deprives its own research. Moreover, the financial profitability of a research based on me-too drugs can appear to be higher, because no investment in fundamental research is required. The counterpart is that the placement on the market of the copy will naturally occur later than that of the original drug, and thus it will make it more difficult to achieve a high sales ranking, all the more so because the me-too drug will be in competition with other copies targeting a similar market.

In reality, the situation is more subtle because very often the synthesis of me-too drugs is justified by a desire to improve the existing drug. Thus, for penicillins, the chemical structure that surrounds the β-lactamic cycle is still being modified. Current antibiotics that have been derived from this research (e.g., the cephalosporins) are more selective, more active on resistant strains, and can be administered by the oral route. They are as different from the parent molecule as a recent car compared to a forty-year-old model. A systematic comparison from a large set of diverse classes of first-in-class and follow-on drugs has shown that minimal chemical changes may lead to strong and advantageous therapeutic properties [11]. In other terms, innovation can result from the sum of a great number of stepwise improvements as well as from a major breakthrough.

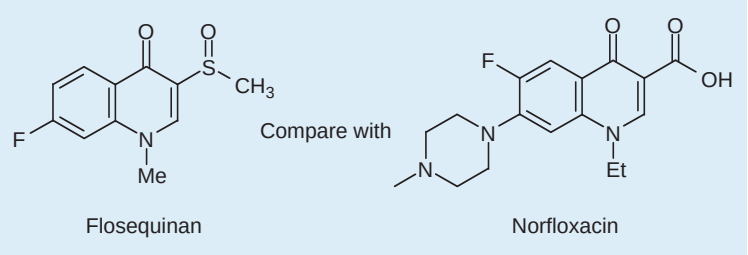


FIGURE 4.5 The striking analogy between the vasodilator drug flosequinan and the quinolone antibiotic norfloxacin.

It can also happen that during the pharmacological or clinical studies of a me-too compound, a totally new property not present in the original molecule appears unexpectedly. Due to the emergence of such a new activity, the therapeutic copy becomes in turn a new lead structure. This was the case for imipramine, initially synthesized as an analog of chlorpromazine and presented to the clinical investigators for the study of its antipsychotic profile [12]. During imipramine's clinical evaluation, it demonstrated much more activity against depressive states than against psychosis. Since 1954, imipramine has truly opened a therapeutic avenue for the pharmacological treatment of depression.

On its way to becoming Viagra, the compound UK-92,480, prepared in 1989 by the Pfizer scientists in Sandwich, England, went first from a drug for hypertension to a drug for angina. Then it changed again when a 10-day toleration study in Wales turned up its unusual side effect: penile erections [13]. It seems probable that a similar emergence of a new activity occurred with flosequinan, which is a sulfoxide bioisostere of the quinolone antibiotics (Figure 4.5). This compound turned out to be a vasodilator and cardiotonic drug, having totally lost any antibiotic activity [14].

III. SECOND STRATEGY: SYSTEMATIC SCREENING

This method consists of screening new molecules, whether they are synthetic or of natural origin, on an animal model or on any biological test without having in mind hypotheses about its pharmacological or therapeutic potential. It rests on the systematic use of selective batteries of experimental models destined to mimic closely the pathological events. The trend is to undertake *in vitro* rather than *in vivo* tests: binding assays, enzyme inhibition measurements, activity on isolated organs or cell cultures, etc. In practice, systematic screening can be achieved in two different manners. The first one applies a very exhaustive pharmacological investigation to a small number of chemically sophisticated and original molecules. This is known as "extensive screening." In contrast, the second one strives to find, among a great number of molecules (several hundreds or thousands), one that could be active in a given indication. This is "random screening."

A. Extensive Screening

Extensive screening is generally applied to totally new chemical entities coming from an original effort of chemical research or from a laborious extraction from a natural source. For such molecules, the high investment in synthetic or extractive chemistry justifies an extensive pharmacological study (central nervous, cardiovascular, pulmonary, and digestive systems, antiviral, antibacterial, or chemotherapeutic properties, etc.) to detect if there exists an interesting potential linked to these new structures. In summary, a limited number of molecules is studied in a thorough manner (vertical screening). It is by such an approach that the antihistaminic and later the neuroleptic properties of the amines derived from phenothiazine were identified. Initially, these compounds had been submitted with negative results to a limited screening study only directed toward possible chemotherapeutic, antimalarial, trypanocidal, and anthelmintic activities.

Original chemical research is also at the origin of the discovery of the benzodiazepines by Sternbach [15]. By the way, this author specifies that the class of compounds he was seeking had to fulfill the following criteria: (1) the chemical series had to be relatively unexplored; (2) it had to be easily accessible; (3) it had to allow a great number of variations and transformations; (4) it had to offer some challenging chemical problems; and (5) it had to "look" as if it could lead to biologically active products.

The extensive screening approach has often led to original molecules. It is, however, highly dependent on the skill and the intuition of the medicinal chemist and even more on the talents of the pharmacologist, who has to be able to adapt and to orient tests as soon as findings evolve to reveal the real therapeutic potential of the molecule under study.

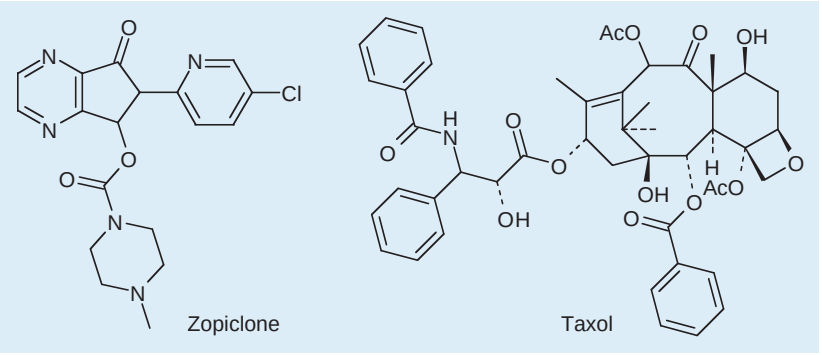


FIGURE 4.6 Drugs discovered by random screening.

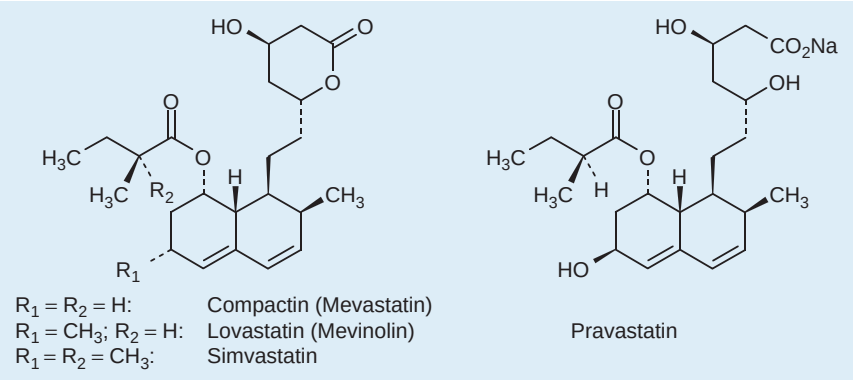


FIGURE 4.7 The natural compounds compactin (mevastatin) and lovastatin block the cholesterol biosynthesis in inhibiting the enzyme HMG-CoA reductase. The later developed compounds simvastatin and pravastatin are semi-synthetic analogs. The open-ring derivative pravastatin is less lipophilic and therefore presents less central side effects. For all these compounds the ring-opened form is the actual active form *in vivo*.

More recent examples of successful systematic screening campaigns are seen with the discovery of cyclopyrrolones (e.g., zopiclone; Figure 4.6) as ligands for the central benzodiazepine receptor [8,16], or of taxol as an original and potent anticancer drug (for a review see Suffness [17]).

Besides the target-based drug discovery strategy followed by the pharmaceutical industry during the last decades, phenotypic screening has been revisited with the emergence of biochemical and molecular biology techniques allowing medium throughput screening on systems-based assays [2,18].

B. Random Screening

In this case, the therapeutic objective is fixed in advance, and—contrary to the preceding case—a great number (several thousands) of molecules are tested but only on a limited number of experimental models. This method has been used for the discovery of new antibiotics. By submitting samples of earth collected in countries from all over the world to a selective antibacterial and antifungal screening, the rich arsenal of anti-infectious drugs that are presently at the disposal of the clinicians was developed. During World War II, an avian model in chickens infected with *Plasmodium gallinaceum* was used for the massive screening of thousands of potential antimalarials. The objective was to solve the problem of the shortage of quinine by finding a synthetic antimalarial. Unfortunately, no satisfying drug was found. Massive screening was implemented in Europe and the United States to discover new anticancer [19] and anti-epileptic drugs. Here again the problem was to select some predictive but inexpensive cellular or animal models. A common criticism of these methods is that they constitute—by the absence of a rational lead—a sort of fishing. Besides, the results are highly variable: nil for the discovery of new antimalarials, rather weak for the anticancer drugs, but excellent in their time for the discovery of antibiotics.

Among more recent successes of this approach, one can mention the discovery of lovastatin, also called mevinolin (Figure 4.7) [20,21], which was the basis of a new generation of hypocholesterolemic agents, acting by inhibition of HMG-CoA reductase.

Sometimes unexpected findings result from systematic screening applied in an unprejudiced manner. An example is found in the tetracyclic compound BMS-192548 extracted from Aspergillus niger WB2346 (Figure 4.8). No one would *a priori* forecast that BMS-192548 exhibits central nervous system (CNS) activities. Actually, the compound turns out to be a ligand for the neuropeptide Y receptor preparations [22].

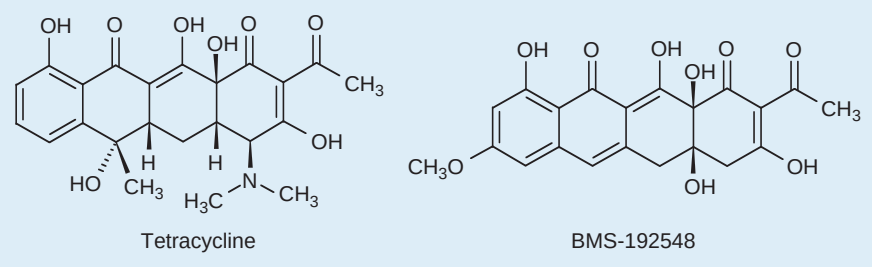


FIGURE 4.8 Unexpected CNS activity of the tetracycline analog BMS-192548.

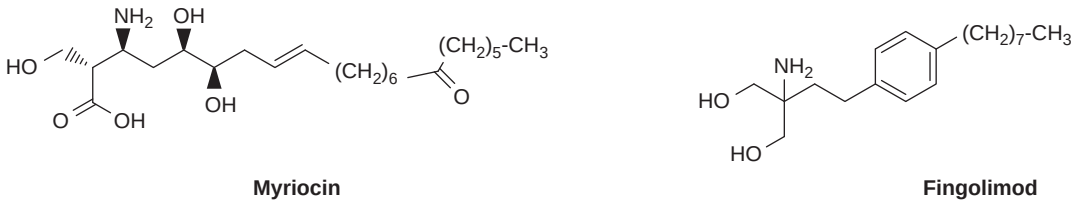


FIGURE 4.9 Fingolimod, a multiple sclerosis drug based on the immunosuppressant fungal metabolite myriocin.

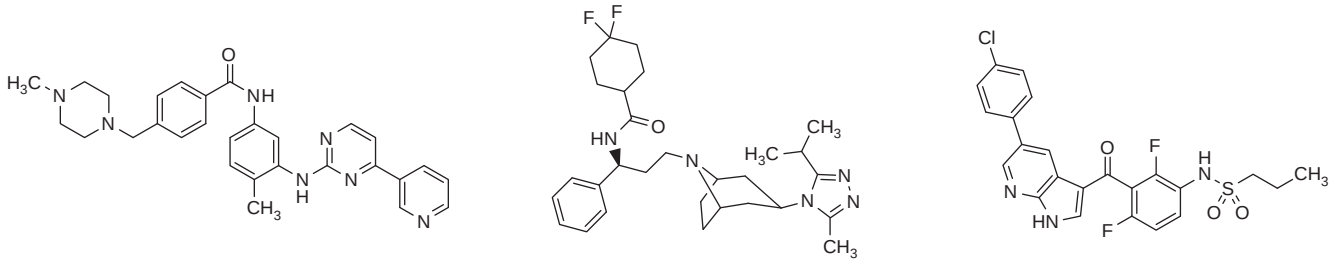


FIGURE 4.10 Examples of drugs originating from HTS (imitanib, maraviroc) and FBDD (vemurafenib).

Another example is the multiple sclerosis drug fingolimod (Gilenya), which is a molecule resulting from the optimization of the fungal natural product myriocin used in traditional Chinese medicine (Figure 4.9) [23].

C. High-Throughput Screening

The arrival of robotics in the 1980s along with the miniaturization of the *in vitro* testing methods meant it became possible to combine the two preceding approaches. In other words, screen millions of compounds on a large number of biological targets. This strategy has delivered numerous drugs (Figure 4.10) [24]. High-throughput screening (HTS) is usually applied to the displacement of radioligands and to the inhibition of enzymes. The present trend is to replace radioligand-based assays with fluorescence-based measurements. Primary sources can also be crudely purified vegetal extracts or fermentation fluids. In this latter case, one proceeds to the isolation and to the identification of the responsible active principle [25,26] only when an interesting activity is observed. HTS is treated in this chapter, where we will describe the general principles of screening set design.

Because it is now possible for a pharmaceutical company to screen several million molecules simultaneously on 30–50 different biochemical tests, the problem becomes feeding the robots interesting molecules and consequently designing "smart" compound collections. To this end, the role of the medicinal chemist is crucial in the preparation of the compound collection to the hit validation triage step and beyond; these steps involve multiple tasks where the interaction with chemoinformaticians plays a major role [27]. Several methods leading to high-quality hit discovery have been described [28]. HTS like methods have also been used to de-orphanize selected targets [29].

Despite the success of this approach, several limitations appear:

- the coverage of the molecular diversity, even when increasing the size of the libraries;
- the quality and the management of the collections;
- the management of the results; and
- the economic issue.

For this last point, the moderate level of success in converting hits from the screening step to drug candidates has tempered its use. To assist the process, several chemo-informatic or modeling tools can be used. For instance, documented examples of *in silico* screening performed prior to the biochemical screening have provided either a focused compound subset dedicated to a target or a diversified or rationally design set if poor data are available on the target [30]. In many cases, one of the goals of using *in silico* approaches is to build a set of compounds with a size corresponding to the capacities of the users. We have to keep in mind several numbers: a usual hit rate for a screening is around 0.1 percent, but in some cases screening does not highlight any hits or only poorly usable ones. However, there are examples showing that *in silico* screenings—when they can be applied—can lead to a hit rate of 10 percent or even more [31].

Compound collection and in silico filters

The preparation of a compound collection for screening is as much art as science, and it benefits from the use of chemoinformatic approaches and from the skills of medicinal chemists [32–34]. This is a critical step because the quality of the screening set of compounds will have a major impact on the next phases of the project.

The size of the compound collections may vary from about 100 thousand to about 5 million compounds to be handled with the current technologies (not mentioning economic concerns), and the chemical space involving small chemical compounds below MW 600 is almost infinite. Therefore, lessons learned from the past must guide the selection of the pertinent screening set. At least three key concepts can be applied to generate the best collection: diversity, the removal or flagging of unwanted compounds, and the design of target-focused compound collections. In addition, with regard to filtering, there are two different philosophies. One school of thought suggests that it might be premature to perform any compound selection during the preparation of the screening collection because the compounds will undergo significant changes during the optimization process. The other philosophy considers that in many situations the clinical candidates are not so dramatically different from the initial hits [35], so the screening sets should contain high-quality molecules free of toxicophores and in the appropriate physico-chemical property range. A possible solution could be to be to find a reasonable balance between the two concepts.

Diversity is an important component that is often used to assess the quality of a library (or to design a high quality collection). Chemical diversity is best ensured by having several representatives of each compound scaffold. Diversity or similarity can be generated using various *in silico* methods (search for the presence of chemical substructures, machine learning approaches, etc.). The removal or the flagging of unwanted compounds is often performed by using various *in silico* filters.

A well-known filter has been proposed by Lipinski and co-workers (the so-called Rule of Five), essentially focusing on orally bioavailability [36], but since 1997 many others have been reported: the Rule of Three for fragment-based drug discovery [37]; the GSK 4/400 rule that states that a molecule with a log P above 4 and a MW above 400 will damage many ADME-Tox properties [38]; and the 3/75 rule which state that compounds with a computed logP <3 and TPSA (topological polar surface area) >75 gave a 6-fold reduction in prevalence of *in vivo* toxicity versus compounds with ClogP >3 and TPSA <75 [39]. The identification of toxicophores, the structural alerts [40–42], and the pan assay interference compounds or PAINS [43] are crucial for the design of high-quality screening sets and for the analysis of the hits coming from HTS.

It is possible to find several valuable commercial packages to prepare a collection and to perform ADMET predictions, such as Schrodinger's Ligprep, Openeye's Filter, or tools from MOE, Tripos, Accelrys, ICM, and Molecular Discovery, among many others. Fortunately, free packages with source codes, open databases, and online services are also available [44].

Several techniques described above provide a means of identifying compounds that should not be included in a screening collection. Besides these "generic" filters, one way of improving the hit rates is to design screening libraries that are targeted to a particular gene family. There are a number of approaches to designing such targeted or focused screening libraries. When protein crystal structures or homology models are available, a screening library can be designed from a structure-based perspective using docking and scoring engines. An alternative and often complementary approach consists in the design of a screening library based on the structural information coming from known active compounds. For gene families such as kinases and G-protein

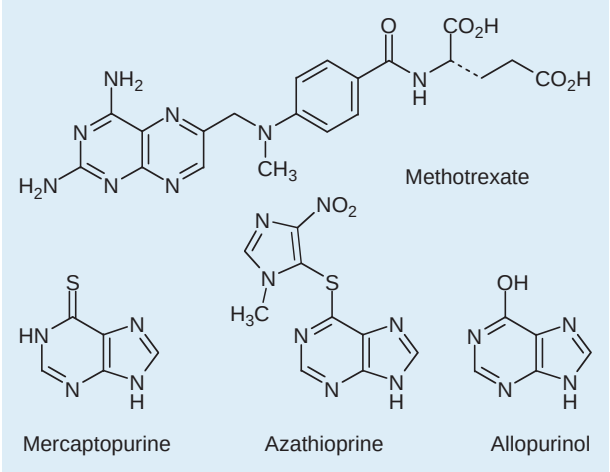


FIGURE 4.11 Departing from methotrexate, simple intermediates led to new drugs. Mercaptopurine and azathioprine are immunosuppressants, and allopurinol is used in the treatment of gout.

coupled receptors, hundreds of known active compounds can be used as starting points for a ligand-based design search and corresponding generation of specialized collections [45]. For some new types of challenging targets, including protein–protein interaction, conventional compound libraries are not fully suitable and are associated with extremely low hit rates and poorly diverse starting points [46]. The topology of the small ligand binding pockets located at the interface between two interacting protein partners is such that more aromatic and more tridimensional molecules (among other molecular descriptors) are needed. In such cases—and especially in the absence of any structural information—machine learning approaches can be used to prepare specific collections presumably enriched in inhibitors of protein–protein interactions [47].

D. Screening of Synthesis Intermediates

As synthesis intermediates are chemically connected to final products and as they often present some common groupings with them, it is unlikely that they equally share some pharmacological properties. For this reason, it is always prudent to submit these compounds to a pharmacological evaluation. Among drugs discovered this way, one finds the tuberculostatic semicarbazones. They were initially used in the synthesis of antibacterial sulfathiazoles. Subsequent testing of isonicotinic acid hydrazide, destined for the synthesis of a particular thiosemicarbazone, revealed the powerful tuberculostatic activity of the precursor that has become since then a major antitubercular drug (isoniazide).

Inhibitors of the enzyme *dihydrofolate-reductase* such as methotrexate (Figure 4.11) are used in the treatment of leukemia. During the search for methotrexate analogs, a very simple intermediate, mercaptopurine, was also submitted to testing. It turned out to be active but relatively toxic. Subsequent optimization led to azathioprine, a prodrug releasing mercaptopurine *in vivo*.

Azathioprine was found to be more potent as an immunosuppressive agent than the previously used corticoids and was systematically used in all organ transplantations until the advent of cyclosporine. Another intermediate in this series, *allopurinol*, inhibits xanthine-oxidase and therefore is used in the treatment of gout [48].

E. New Leads from Old Drugs: The SOSA Approach

The SOSA approach (SOSA 5 = **S**elective **O**ptimization of **S**ide **A**ctivities) represents an original alternative to high-throughput screening (HTS) [4,49–53]. It consists of two steps:

1. Screen newly identified pharmacological targets of a limited set (approximately 1000 compounds) of well-known drug molecules for which bioavailability and toxicity studies have already been performed and which have proven usefulness in human therapy. By definition, in using such a library, all hits that are found are drug-like.

2. Optimize hits (by means of traditional, parallel, or combinatorial chemistry) in order to increase the affinity for the new target and decrease the affinity for the other targets. The objective is to prepare analogs of the hit molecule in order to transform the observed "side activity" into the main effect and to reduce strongly or abolish the initial pharmacological activity.

The rationale behind the SOSA approach lies in the fact that in addition to their main activity, almost all drugs used in human therapy show one or several side effects. In other words, if they are able to exert a strong interaction with the main target, they also exert weaker interaction with some other biological targets. Most of these targets are unrelated to the primary therapeutic activity of the compound. The objective of the medicinal chemist is then to proceed to a reversal of the affinities, the identified side effect becoming the main effect and vice versa. Many cases of activity profile reversals by means of the SOSA approach have been published.

A typical illustration of the SOSA approach is given by the development of *selective ligands for the endothelin ET_A receptors* by scientists from Bristol-Myers-Squibb [54,55]. Starting from an in-house library, the antibacterial compound sulfathiazole (Figure 4.12) was an initial, but weak, hit ($IC_{50} = 69\ \mu M$). Testing of related sulfonamides identified the more potent sulfisoxazole ($IC_{50} = 0.78\ \mu M$). Systematic variations led finally to the potent and selective ligand BMS-182874. *In vivo*, this compound was orally active and produced a long-lasting hypotensive effect.

Further optimization guided by pharmacokinetic considerations led the BMS scientists to replace the naphtalene ring by a diphenyl system [55]. Among the prepared compounds, BMS-193884 ($ET_AK_i = 1.4$ nM; $ET_BK_i = 18\,700$ nM) showed promising hemodynamic effects in a phase II clinical trial for congestive heart failure. More recent studies led to the extremely potent antagonist BMS-207940 (edonentan; $ET_AK_i = 10$ pM) presenting an 80 000-fold selectivity for ET_A versus ET_B. The bioavailability of this compound is 100 percent in rats, and it exhibits oral activity already at a 3 μM/kg dosing [55].

Another example is the antidepressant minaprine (Figure 4.13). In addition to reinforcing serotoninergic and dopaminergic transmission, this amino-pyridazine possesses weak affinity for muscarinic M_1 receptors ($K_i = 17\ \mu M$). Simple chemical variations diminished the dopaminergic and serotoninergic activities and boosted the cholinergic activity up to nanomolar concentrations [56–58].

Similarly, chemical variations of the D_2/D_3 non-selective benzamide sulpiride (Figure 4.14) led to compound Do 897, a selective and potent D_3 receptor partial agonist [59].

As mentioned above, a differentiating peculiarity of this type of library is that it is constituted by compounds that have already been safely given to humans. Thus, if a compound were to "hit" with sufficient potency on an

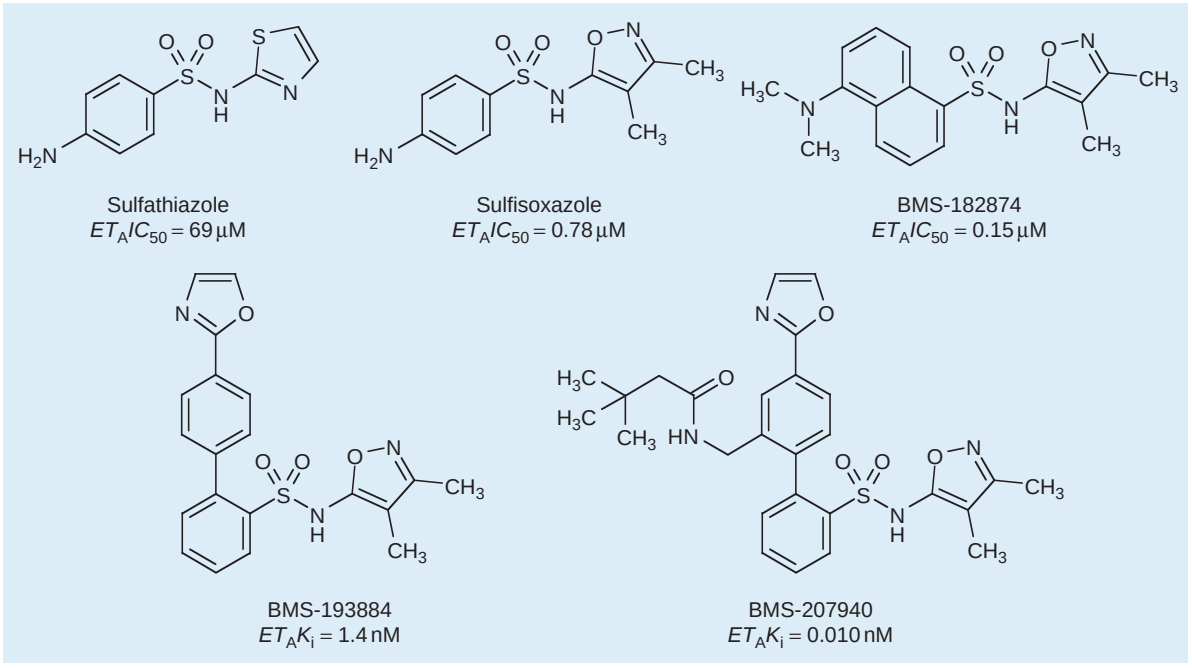


FIGURE 4.12 A successful SOSA approach allowed the identification of the antibacterial sulfonamide sulfathiazole as a ligand of the endothelin ET_A receptor and its optimization to the selective and potent compounds BMS-182874, BMS-193884 and BMS-207940 [54,55].

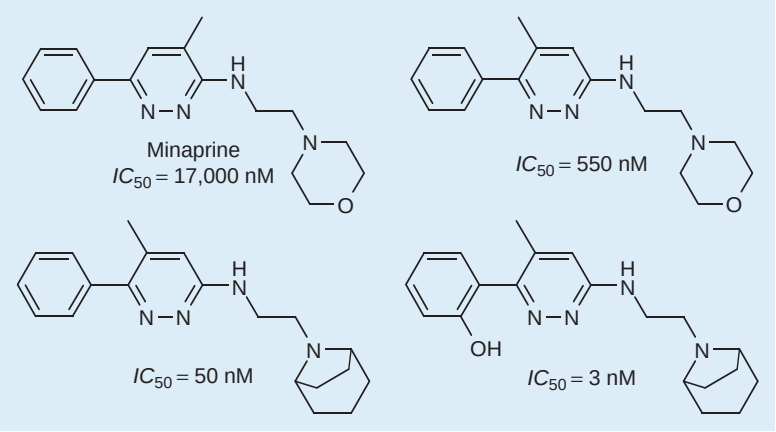


FIGURE 4.13 Progressive passage from minaprine to a potent and selective partial muscarinic M1 agonist [54,56,57].

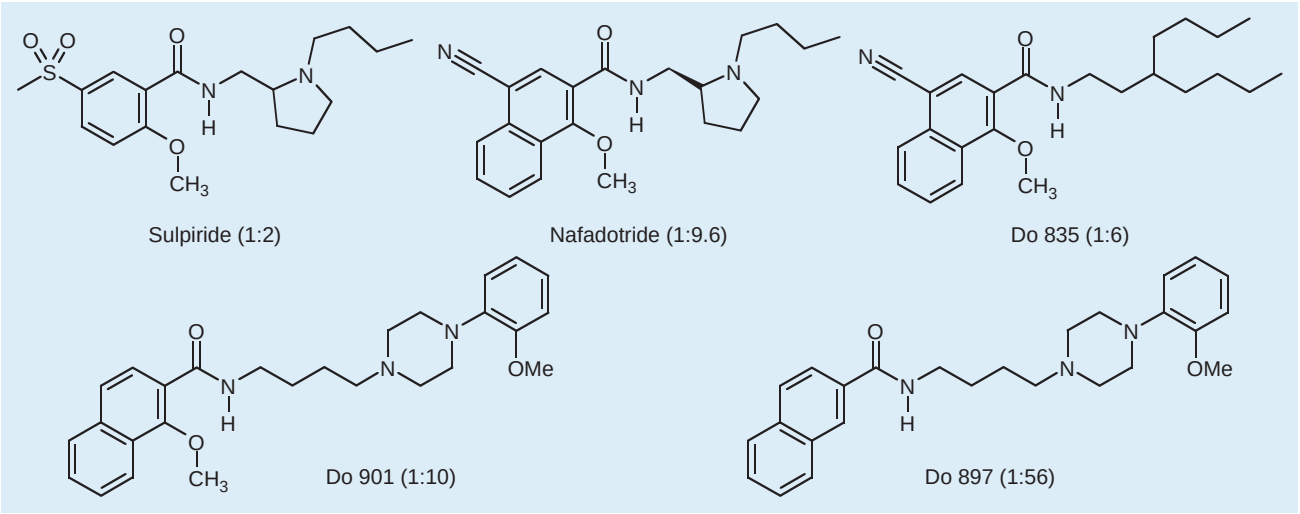


FIGURE 4.14 The progressive change from the D_2/D_3 receptor nonselective antagonists to the highly D_3-selective compound Do 897 [58]. The numbers between parentheses indicate the D_2/D_3 affinity ratio.

orphan target, there is a high chance that it could rapidly be tested in patients for proof of principle. Alternatively, if one or more compounds hit but with insufficient potency, optimized analogs can be synthesized, and the chances that these analogs will be good candidate drugs for further development are much higher than if the initial lead is toxic or not bioavailable. One of these new-type of chemical libraries, the Prestwick Chemical Library, is available [60]. It contains 1120 biologically active compounds with high chemical and pharmacological diversity as well as known bioavailability and safety in humans. Over 90 percent of the compounds are well-established drugs, and 10 percent are bioactive alkaloids. For scientists interested in drug-likeness such a library comply adequately with the quest for "drug-like" leads.

IV. THIRD STRATEGY: EXPLOITATION OF BIOLOGICAL INFORMATION

A major contribution to the discovery of new active principles comes from the exploitation of biological information. This means information which relates to a given biological effect (fortuitous or voluntary) provoked by some substances in humans, in animals, or even in plants or bacteria. When such information becomes accessible to the medicinal chemist, it can serve to initiate a specific line of therapeutic research. Originally, the observed biological effect can simply be noticed without any knowledge of how it works.

A. Exploitation of Observations Made in Humans

The activity of exogenous chemical substances on the human organism can be observed under various circumstances: ethnopharmacology, popular medicines, clinical observation of secondary effects or adverse events, fortuitous observation of activities of industrial chemical products, etc. Because the harvested information is observed directly in humans, this approach presents a notable advantage.

1. Study of Indigenous Medicines (Ethnopharmacology)

Natural substances were for a long time the unique source of medicines. At present, they constitute 30 percent of the used active principles—and probably more (approximately 50 percent) if one considers the number of prescriptions that utilize them, particularly since use of antibiotics plays a major role [61]. Behind most of these substances one finds indigenous medicines. As a consequence, *ethnopharmacology* represents a useful source of lead compounds. Historically, we are indebted to this approach for the identification of the cardiotonic digitalis glucosides, the opiates, and the cinchona alkaloids. Curare was obtained from a South American plant used for a long time by natives to make arrow poisons. The cardiotonic glucosides of the *Strophantus* seeds and the alkaloid eserine from the Calabar beans are other examples of active drugs originally used by natives as poisons. The *Rauwolfia serpentina* has been used for centuries in India before western medicine became interested in its tranquillizing properties and extracted reserpine from it. Atropine, pilocarpine, nicotine, ephedrine, cocaine, theophylline, and innumerable other medicines have thus been extracted from plants to which popular medicine attributed therapeutic virtues.

Despite its extremely useful contributions to the modern pharmacopoeia—such as artemisin and huperzine—folk medicine is a rather unreliable guide in the search for new medicines. This is illustrated by the example of antifertility agents. According to natives of some islands of the Pacific, approximately 200 plants would be efficient in reducing male or female fertility. Extracts have been prepared from 80 of these plants and have been administered at high dosings to rats during periods of four weeks and more without observing the slightest effect upon pregnancies or litter sizes [62]. When ethnopharmacology and the natural substance chemistry end in the discovery of a new active substance, this latter is first reproduced by total synthesis. It is then the object of systematic modifications and simplifications that aim to recognize by trial and error the minimal requirements that are responsible for the biological activity.

2. Clinical Observation of Side Effects of Medicines

The clinical observations of entirely unexpected side effects constitute a quasi-inexhaustible source of clues in the search for lead compounds. Indeed, beside the wanted therapeutic action, most drugs possess side effects. These are accepted either from the beginning as a necessary evil or recognized only after some years of use. When side effects present a medical interest by themselves, a planned objective can be the dissociation of the primary from the side effect activities (i.e., enhance the activity originally considered as secondary and diminish or cancel the activity that initially was dominant). Promethazine, for example, an antihistaminic derivative of phenothiazine, is burdened with significant sedative effects. The merit of a clinician such as Laborit [63] has been to promote the utilization of this side effect and to direct research toward better profiled analogs. This impulse was the origin of the birth of chlorpromazine, the prototype of a new therapeutic series, the neuroleptics, whose existence was unsuspected until then and has revolutionized the practice of psychiatry [12,64]. Innumerable other examples can be found in the literature, such as the hypoglycemic effect of some antibacterial sulfamides, the uricosuric effect of the coronaro-dilating drug benziodarone, the antidepressant effect of isoniazide, an antitubercular drug, and the hypotensive effect of β-blocking agents.

This last example is beautifully illustrated by the discovery of the potassium channel activator cromakalim [65]. Cromakalim is the first antihypertensive agent to be shown to act exclusively through potassium channel activation [66]. This novel mechanism of action involves an increase in the outward movement of potassium ions through channels in the membranes of vascular smooth muscle cells, leading to relaxation of the smooth muscle. The discovery of this compound can be summarized as follows: β-adrenergic receptor blocking drugs were not thought to have antihypertensive effects when they were first investigated. However, pronethalol, a drug that was never marketed, was found to reduce arterial blood pressure in hypertensive patients with *angina pectoris*. This antihypertensive effect was subsequently demonstrated for propranolol and all other β-adrenergic antagonists [67]. Later on, there were some doubts that blockade of the β-adrenergic receptors was responsible for the hypotensive activity, and attempts were made to dissociate, in the classical β-blocking molecules, the β-blockade from the antihypertensive activity. Among the various conceivable molecular variations that are possible for the

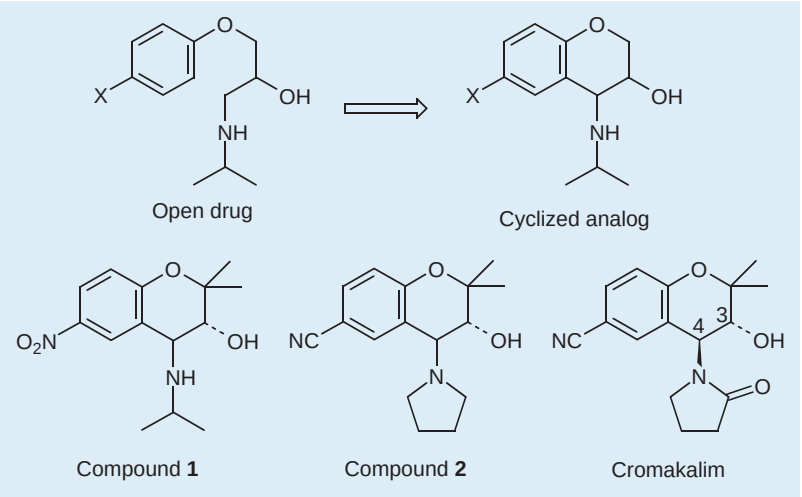


FIGURE 4.15 The clinical observation of the hypotensive activity of the "open" (and therefore flexible) β-blocking agents was the initial lead to cyclized analogs devoid of β-blocking activity but retaining the antihypertensive activity (Stemp and Evans [65]).

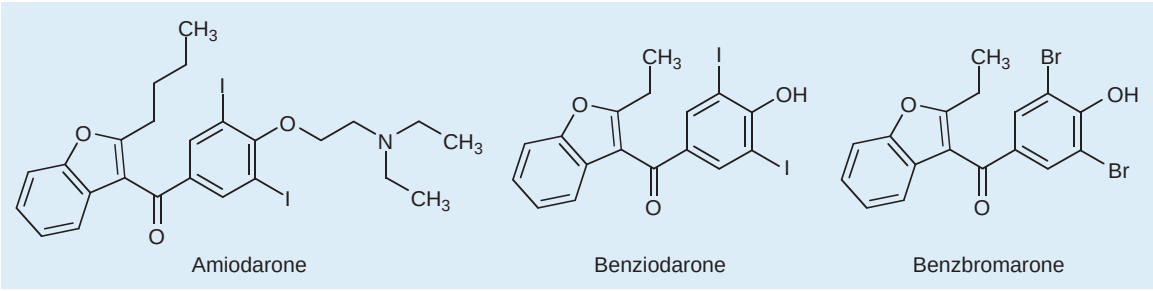


FIGURE 4.16 Structures of the arones.

flexible β-blockers, it was found that conformational restriction obtained in cyclizing the carbon atom bearing the terminal amino group onto the aromatic ring yielded derivatives devoid of β-blocking activity but retaining the antihypertensive activity (Figure 4.15).

One of the first compounds prepared (compound **1**, Figure 4.15) was indeed found to lower blood pressure in hypertensive rats by a direct peripheral vasodilator mechanism. No β-blocking activity was observed. Optimization of the activity led to the 6-cyano-4-pyrrolidinylbenzopyran (compound **2**), which was more than a 100-fold more potent than the nitro derivative. The replacement of the pyrrolidine by a pyrrolidinone (which is the active metabolite) produced a 3-fold increase in activity and the optical resolution led to the (−)-3*R*, 4*S* enantiomer of cromakalim (BRL 38227) that concentrates almost exclusively the hypotensive activity [65,68,69].

3. *New Uses for Old Drugs*

In some cases, a new clinical activity observed for an old drug is sufficiently potent and interesting to justify the immediate use of the drug in the new indication. This is illustrated hereafter.

Amiodarone, for example (Figure 4.16), was introduced as a coronary dilator for angina. Concern about corneal deposits, discoloration of skin exposed to sunlight, and thyroid disorders led to the withdrawal of the drug in 1967. However, in 1974 it was discovered that amiodarone was highly effective in the treatment of a rare type of arrhythmia known as the Wolff-Parkinson-White syndrome. Accordingly, amiodarone was reintroduced specifically for that purpose [70].

Benziodarone, initially used in Europe as a coronary dilator, proved later on to be a useful uricosuric agent. Presently, it is withdrawn from the market due to several cases of jaundice associated with its use. The corresponding brominated analog, benzbromarone, was specifically marketed for its uricosuric properties.

Thalidomide, was initially launched as a sedative/hypnotic drug (Figure 4.17), but withdrawn because of its extreme teratogenicity. Under restricted conditions (no administration during pregnancy or to any woman of childbearing age), it found a new use as immunomodulator. Particularly it seems efficacious for the treatment of *erythema nodosum leprosum*, a possible complication of the chemotherapy of leprosy [71].

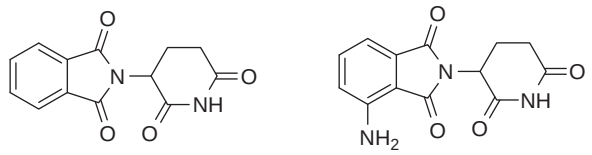


FIGURE 4.17 Structure of thalidomide and panalidomide. The marketed compounds are the racemates.

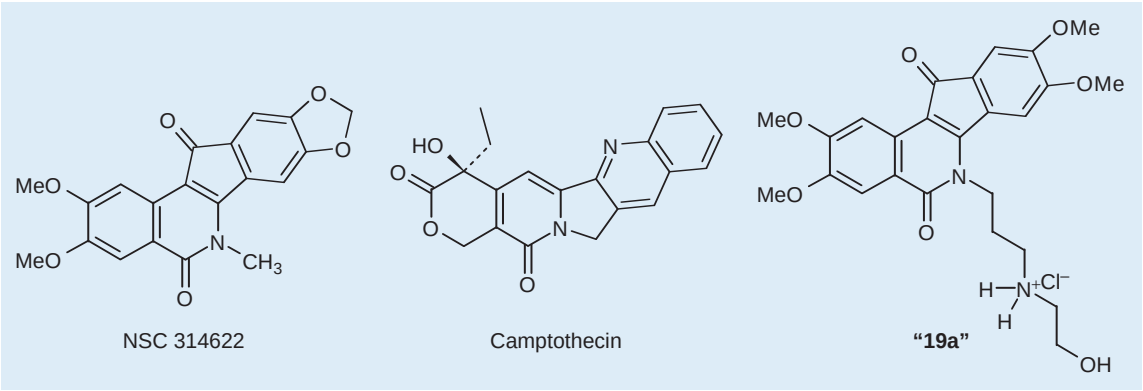


FIGURE 4.18 The indenoisoquinoline NSC 314622 resurfaced 20 years after its first testing as a top 1 inhibitor [72].

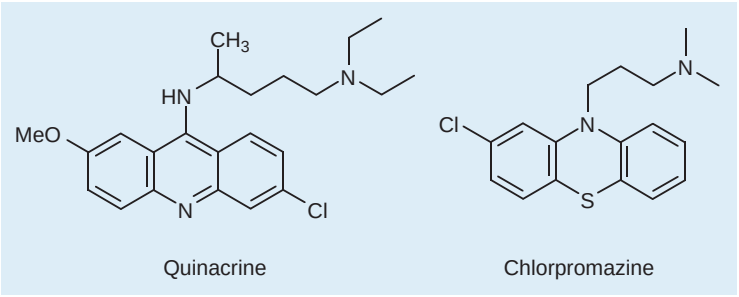
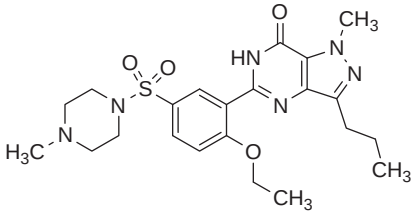


FIGURE 4.19 Old drugs, new use. The antimalarial drug quinacrine and the antipsychotic drug chlorpromazine are able to inhibit prion infection [73].

The antiangiogenic properties of thalidomide have been a topic of intense investigations leading to the discovery of the amino-thalidomide pomalidomide (Pomalyst), which has been approved in both the US and Europe for the treament of refractory multiple myeloma. In 1978, the synthesis of the *indenoisoquinoline* NSC 314622 (Figure 4.18) was reported as the result of an unexpected transformation during a synthesis of nitidine chloride. Given its weak antitumor activity, it was not investigated further. Twenty years later, NSC 314622 resurfaced as a potential topoisomerase I (top I) inhibitor and served as lead structure for the design of cytotoxic non-camptothecin top I inhibitors such as the compound "19a" [72].

In 2001, the antimalarial drug *quinacrine* and the antipsychotic drug *chlorpromazine* (Figure 4.19) were shown to inhibit prion infection in cells. Prusiner et al. [73]. identified the drugs independently and found that they inhibit conversion of normal prion protein into infectious prions and clear prions from infected cells. Both drugs can cross over from the bloodstream to the brain, where the prion diseases are localized.

A more recent example is provided by the discovery of the use of sildenafil (Viagra®, Figure 4.20), a phosphodiesterase type 5 (PDE5) inhibitor, as an efficacious, orally active agent for the treatment of male erectile dysfunction [74,75]. Initially, this compound was brought to the clinic as an hypotensive and cardiotonic substance and its usefulness in male erectile dysfunction resulted clearly from the clinical observations. Then sildenafil (Viagra) has been itself the topic of chemical transformations, the "nitrogen walk" refinement of its core leading to the discovery of the follow-on drug valdenafil (Figure 4.5).



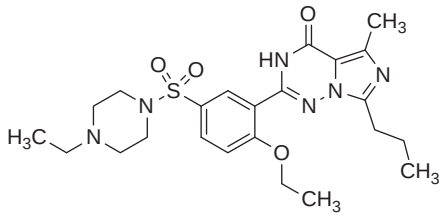


FIGURE 4.20 Structure of the PDE5 inhibitors sildenafil [74,75] and valdenafil [2].

In many therapeutic families, each generation of compounds induces the birth of the following one. This happened in the past for the sulfamides, penicillins, steroids, prostaglandins, and tricyclic psychotropics families, and one can draw real genealogical trees representing the progeny of the discoveries. More recent examples are found in the domain of ACE inhibitors and in the family of histaminergic H_2 antagonists.

Research programs based on the exploitation of side effects are of great interest in the discovery of new starting points as far as they depend on information about activities *observed directly in humans* instead of animals. On the other hand, they do allow the detection of new therapeutic activities *even when no pharmacological models in animals exist*.

4. *The Fortuitous Discovery of Activities of Industrial Chemical Products*

During the industrial manufacture of nitroglycerin, toxic manifestations due to this compound—particularly strong vasodilating properties—were observed in workers. From there came the utilization of this substance—and later on of other nitric esters of aliphatic alcohols—in *angina pectoris* and as cerebral vasodilators. In an analogous manner it was observed during the manufacture of the sulfa drug sulfathiazole that 2-amino-thiazole, one of the starting materials, was endowed with antithyroidal properties. This observation fostered the use of this compound, and of amino-thiazoles in general, for the treatment of thyroid gland hyperactivity. Tetraethylthiurame disulfide was originally used as antioxidant in the rubber industry. After having manipulated it, workers felt an intolerance to alcohol. Therefore, this product was proposed for alcohol withdrawal cures (disulfiram). On the molecular level, the mode of action of disulfiram rests on the inhibition of the enzyme aldehyde-dehydrogenase that normally insures the oxidation of acetaldehyde into acetic acid. The intake of alcohol under disulfiram provokes an accumulation of acetaldehyde that achieves a real intoxication of the patient. Another example of a fortuitous discovery is given by the example of probucol. This antihyperlipoproteinemic compound was originally synthesized as an antioxidant for plastics and rubber [76,77].

B. Exploitation of Observations Made in Animals

We find here all the research done by physiologists that has been the basis of the discovery of vitamins, hormones, and neurotransmitters, and the fall-outs of various pharmacological studies when they were performed *in vivo*. Other observations made on animals, often in a more or less fortuitous manner, have led to useful discoveries. An example is provided by the dicoumarol-derived anticoagulants.

The discovery of the anticancer properties of the alkaloids of *Vinca rosea* constitutes a particularly beautiful example of pharmacological feed-back. Preparations from this plant had the reputation in some popular medicines of possessing antidiabetic virtues. During a controlled pharmacological test, these extracts were proven to be devoid of hypoglycemic activity. On the other hand, it was frequently observed that the treated rats died from acute septicemia. A study of this phenomenon showed that it was due to massive leukopenia. In taking the leukocytes count as the activity end-point criterion, it became possible to isolate the main alkaloid, vinblastine [78]. At the same time in another laboratory, routine anticancer screening had revealed the activity of the crude extract on the murine leukemia [79]. Subsequently, the antileukemic activity became a screening tool. Out of thirty alkaloids isolated from various periwinkles, four (vinblastine, vinleurosine, vincristine, and vinrosidine) were found active in human leukemias [80].

Analogs of L-arginine with modifications at the terminal guanidino nitrogen and/or the carboxyl terminus of the molecule have been widely used for their ability to inhibit the production of nitric oxide (NO) and are thought to be competitive antagonists of nitric synthase. In studies designed to elucidate the role of NO in the gastrointestinal tract, an inhibitory effect of N^G-nitro-L-arginine methyl ester (L-NAME) on cholinergic neural responses was sometimes observed. This inhibitory effect was shown to be consistent with a blockade of the muscarinic receptors [81].

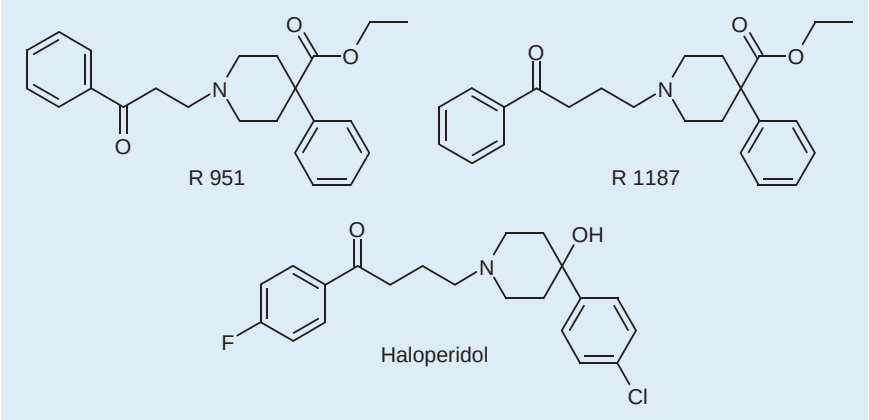


FIGURE 4.21 The passage from pethidine-related opiate analgesics to the dopaminergic antagonist haloperidol.

Remember too that it was the research of insecticides that led to the discovery of the organophosphorus acetylcholinesterase inhibitors by Schrader at the Bayer laboratory [82]. The study of their mechanism of action has shown that they act by acylation of a serine hydroxyl in the catalytic site of the enzyme. This was one of the first examples describing a molecular mechanism for an enzymatic inhibition.

Replacement by Janssen et al. [83]. of the *N*-methyl group of pethidine by various propio- and butyrophenones led to potent analgesics such as R951 and R1187 (Figure 4.21). During their pharmacological study it was noted that mice that had been injected with these drugs became progressively calm and sedated. The resemblance of the sedation to that produced by chlorpromazine encouraged Janssen to synthesize analogs of R1187 in the hope that one might be devoid of analgesic activity while retaining tranquilizing activity.

From this effort, haloperidol emerged in 1958 as the most potent tranquillizer yet to have been discovered. It is 50−100 times as potent as chlorpromazine, with fewer side effects [83,84].

C. Exploitation of Observations Made in the Plant Kingdom and in Microbiology

Among the numerous discoveries that we owe to botanists and pharmacognosts, the interest in tryptophan metabolites has to be invoked, especially the interest for indolylacetic acid [85]. This compound acts as growth hormone in plants. *Para*-chlorinated phenoxyacetic acids (Methoxone (4-Chloro-2-methylphenoxy) acetic acid) (MCPA); 2,4-Dichlorophenoxyacetic acid (2,4-D)); are mimics (bioisosteres) of indolylacetic acid and show similar phytohormonal properties: at high doses they serve as weed control agents. Ring-chlorinated phenoxyacetic acids have been later introduced in molecules as varied as meclofenoxate (cerebral metabolism), clofibrate (lipid metabolism), and ethacrynic acid (diuretic).

The 5-hydroxylated analog of indolacetic acid is the principal urinary metabolite of serotonin. On the basis of two biochemical observations, the possible role of serotonin in inflammatory processes and the increase of urinary metabolites of tryptophan in rheumatic patients, Shen, from the Merck Laboratories, designed anti-inflammatory compounds derived from indolacetic acid. Among them he found indomethacin in 1963, one of the most powerful nonsteroidal anti-inflammatory drugs currently known [86].

A particularly rich contribution of this approach in the therapeutic area has been the discovery and the development of penicillin (see Chapter 1 in third edition of this book). It initiated the discovery of many other major antibiotics such as chloramphenicol, streptomycin, tetracyclines, and rifampicine.

In conclusion, whatever its origin may have been, the use of biological information constitutes a rich source of new lead molecules for research. It presents the advantage of offering creative approaches not dependent on routine pharmacological models. Once the lead molecule is identified, it will immediately be the object of thorough studies to elucidate its molecular mechanism of action. Simultaneously, one will proceed to the synthesis of structural analogs, as well as to the establishment of structure−activity relationships, and to the optimization of all indispensable parameters for its development: potency, selectivity, metabolism, bioavailability, toxicity, and cost price. In other terms, even if the initial discovery was purely fortuitous, subsequent research must be marked by a very important effort of rationalization.

V. FOURTH STRATEGY: PLANNED RESEARCH AND RATIONAL APPROACHES

The approaches that we have described up to now owe a great deal to chance (screening, fortuitous discoveries), or they lack originality (therapeutic copies). A more scientific approach is based on the knowledge of the incriminated molecular target: enzyme, receptor, ion channel, signaling protein, transport protein, or DNA. The progresses in molecular and structural biology allowed the identification and characterization of several hundred new molecular targets and made it possible to envisage the design of drugs at a more scientific level.

A. L-DOPA and Parkinsonism

A historical example in which the key information that rendered possible a rational approach to drug design is the discovery of the usefulness of L-3,4-dihydroxy-phenylalanine (L-DOPA) in the treatment of Parkinson's disease. Thus, since it was observed in patients suffering from parkinsonism that the dopamine levels in the basal ganglions were much lower than those found in the brains of healthy persons [87], a symptomatic, but rational, therapy became possible. This therapy consists of administering to patients the L-DOPA. This amino acid is able to cross the blood-brain barrier and is then decarboxylated into dopamine by brain DOPA-decarboxylase. Initial clinical studies were undertaken by Cotzias, Van Woert, and Schiffer [88]. Several hundred thousand patients have benefited from this treatment. However, 95 percent of the DOPA administered by the oral route is decarboxylated in the periphery before having crossed the blood-brain barrier. To preserve the peripheral DOPA from this unwanted precocious degradation, a peripheral inhibitor of DOPA-decarboxylase is usually added to the treatment. An additional improvement of the treatment is the simultaneous addition of an inhibitor of catechol *O*-methyltransferase such as tolcapone or entacapone.

Other examples of the rational approach in pharmacology are the discovery of inhibitors of the ACE or that of antagonists of histaminergic H_2-receptors.

B. Inhibitors of the ACE

ACE catalyses two reactions that are supposed to play an important role in the regulation of the arterial pressure: (1) conversion of angiotensin I, which is an inactive decapeptide, into angiotensin II, an octapeptide with a very potent vasoconstrictor activity; and (2) inactivation of the nonapeptide bradykinin, which is a potent vasodilator (Figure 4.22).

An inhibitor of the converting enzyme would therefore constitute a good candidate for the treatment of patients suffering from hypertension. The first substance developed in this sense was teprotide, a nonapeptide presenting an identical sequence to that of some peptides isolated in 1965 by Ferreira from the venom of *Bothrops jararaca*, a Brazilian viper (Figure 4.23).

Teprotide inhibits in a competitive manner the degradation of angiotensin I by the converting enzyme. The presence of four prolines and a pyroglutamate renders this peptide relatively resistant to hydrolysis, but not to a sufficient degree to allow its oral administration. In the search for a molecule offering better bioavailability, the

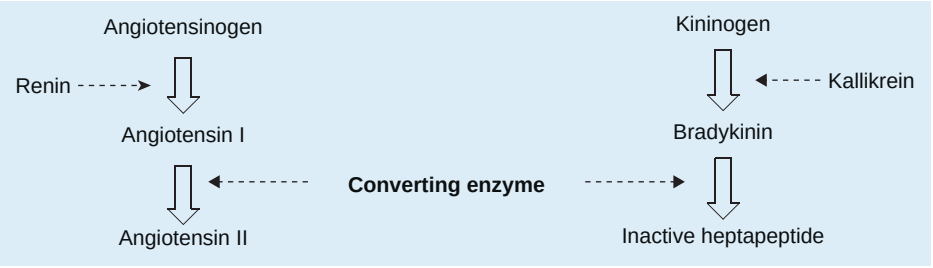


FIGURE 4.22 Scheme of the renin-angiotensin and of the kallikrein–kinin systems. The converting enzyme (a carboxy-dipeptidyl-hydrolase) is common to the two systems.

pyro-Glu-Trp-Pro-Arg-Pro-Glu-Ile-Pro-Pro-OH

FIGURE 4.23 The structure of the nonapeptide teprotide.

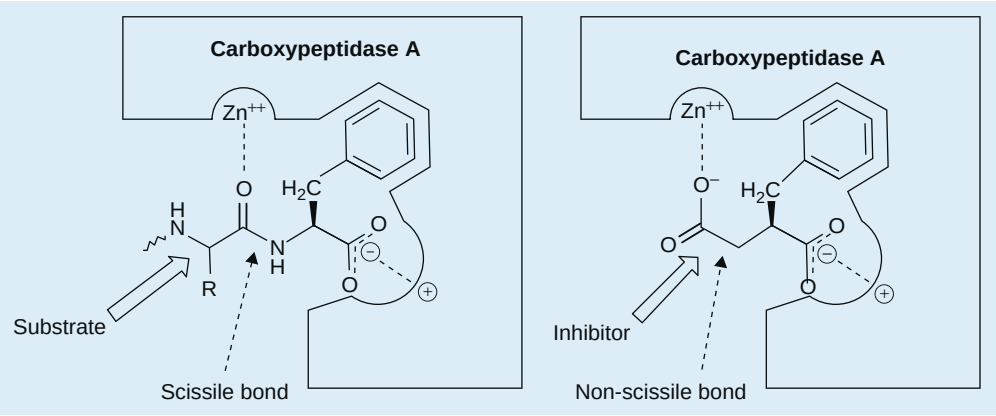


FIGURE 4.24 Interactions between carboxypeptidase A and a substrate (left) or an inhibitor (right). *Adapted after Cushman et al. [89].*

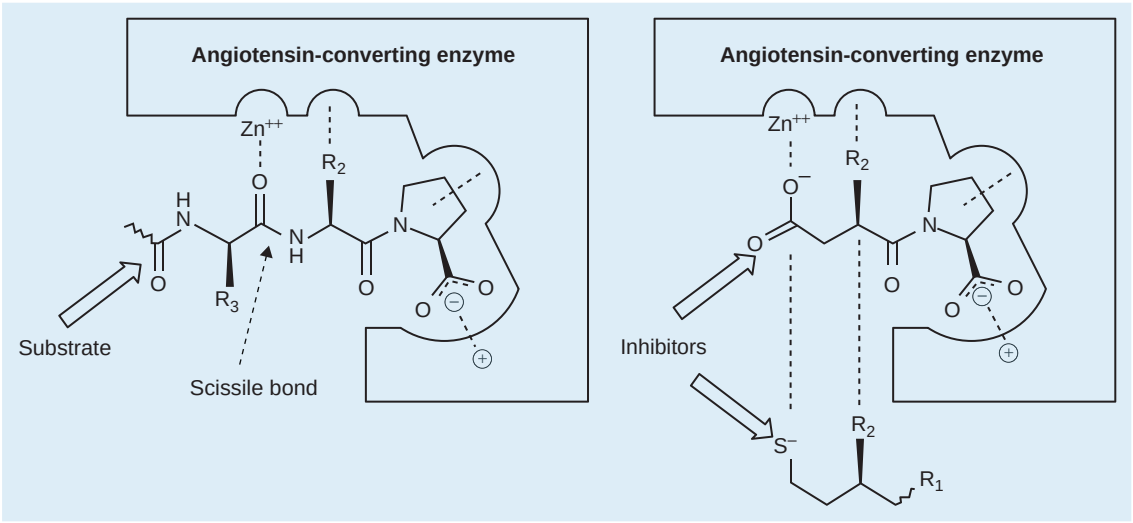


FIGURE 4.25 Interactions between ACE (a dipeptidyl carboxypeptidase) and a substrate (left) or inhibitors (right). *Adapted after Cushman et al. [89].*

reasoning of the Squibb scientists rested on the analogy of ACE with the bovine carboxypeptidase A [89]. In fact, both enzymes are carboxypeptidases; carboxypeptidase A detaches only one C-terminal amino acid while the converting enzyme detaches two. Furthermore, it was known that the active site of carboxypeptidase A comprises three important elements for the interaction with the substrate (Figure 4.24): an electrophilic center, establishing an ionic bond with a carboxylic function; a site capable of establishing a hydrogen bond with a peptidic C-terminal function; and an atom of zinc, solidly fixed on the enzyme and serving to form a coordinating bond with the carbonyl group of the penultimate (the scissile) peptidic function.

By identifying that the conversion enzyme had a similar function, however altered by one amino acid unit (cleavage of the second peptidic bond instead of the first, departing from the terminal carboxyl group), scientists of the Squibb company have conceived the model drawn on Figure 4.25.

According to this model, *N*-succinyl amino acids such as the succinyl prolines shown in Figure 4.25 (right) should be able to interact with each of the above-mentioned sites based on, first their proline carboxyl (ionic bond), their amide function (hydrogen bond), and then on the carboxyl of the succinyl moiety (coordination with the zinc atom). These compounds should then be able to act as competitive and specific inhibitors of the converting enzyme. Therefore, a series of *N*-succinyl amino acids were prepared, and the *N*-L-proline derivative **1** (Figure 4.26) showed some activity ($IC_{50} = 330\ \mu M$). Amino acids other than L-proline lead to less active succinyl derivatives; this result is in agreement with the fact that several peptidic inhibitors (notably teprotide) also possess a proline in the C-terminal position. In the present example, *N*-succinyl-L-proline was selected as lead compound. The next task

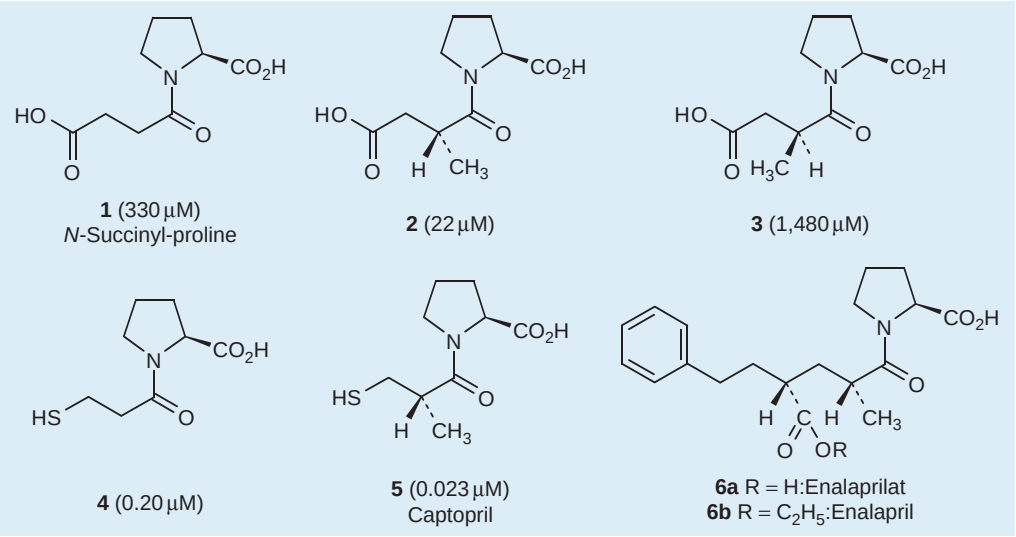


FIGURE 4.26 Structures of some key compounds in the development of captopril and enalapril [89].

was to optimize its activity, and this was done by researching the best interaction with the active site of the enzyme. Two steps were decisive in this quest: fishing for hydrophobic pockets and the research for a better coordinant for the zinc atom (Figure 4.26). The exploration of hydrophobic pockets was achieved by substituting the succinyl moiety with methyl groups (four possibilities taking into account the regio- and the stereoisomers). Structure **2**, methylated at position β to the amide, appeared clearly more active than **1** ($IC_{50} = 22$ instead of 330 μM). In this process, one observes an important stereoselectivity, since the IC_{50} value of epimer **3** of the compound **2** drops to 1480 μM. The best coordination with the zinc was achieved by replacing the carboxyl function by a mercapto group. The gain resulting from this modification was extremely important as shown by the comparison of compounds **1** and **4** or also **2** and **5**. Compound **5** (SQ 14225) with an IC_{50} of 0.023 μM, and a K_i of 0.0017 μM is active by the oral route and has been introduced in therapy under the designation of captopril.

It is interesting to observe that the gain in affinity obtained by the introduction of the mercapto function can be obtained by adding an additional hydrophobic interaction. Thus, scientists from Merck developed enalaprilat **6a**, a compound of comparable effectiveness for which the additional hydrophobic interaction is provided by a phenethyl residue. But enalaprilat is poorly absorbed orally; therefore, the commercial compound is enalapril **6b**, the corresponding ethyl ester.

C. Discovery of the H_2-Receptor Antagonists

Research to develop specific antagonists of the H_2 histamine receptor for the treatment of gastric ulcers has also proceeded through a rational process [90,91]. Starting from the observation that the antihistaminic compounds known at that time (antagonists of H_1-receptors) were incapable of antagonizing the gastric secretion provoked by histamine, Black and his collaborators envisaged the existence of an unknown subclass of the histamine receptor (the future H_2-receptor). From 1964 on, they initiated a program of systematic search for specific antagonists for this receptor.

The starting point was the guanyl-histamine (Figure 4.27) that possesses weak antagonistic properties against the gastric secretion induced by histamine. The lengthening of the side chain of this compound clearly increased the H_2-antagonistic activity, but a residual agonist effect remained. In replacing the strongly basic guanidino function by a neutral thiourea, *burimamide* was obtained. Although very active, this compound was rejected for its low oral bioavailability.

The addition of a methyl group in position 4 of the imidazolic ring, followed by the introduction of an electron-withdrawing sulfur atom in the side chain, finally led to a compound that was both very active and less ionized, properties which improved its absorption by the oral route. The derivative thus obtained, *metiamide*, was excellent and—moreover—ten times more potent than burimamide. However, because of its thiourea grouping, metiamide was tainted with side effects (agranulocytosis, nephrotoxicity) that would limit its clinical use.

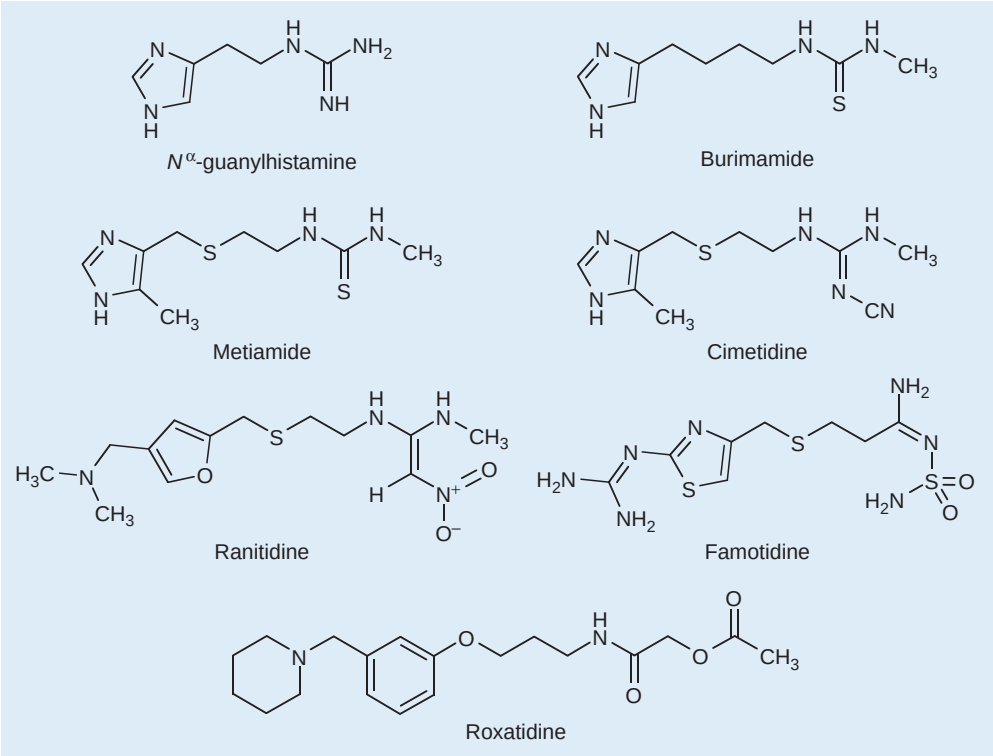


FIGURE 4.27 Structures of some key compounds in the development of H_2-receptor antagonists.

The replacement of the thiourea by an isosteric grouping having the same pK_a (*N*-cyanoguanidine) led finally to *cimetidine*, which became a medicine of choice in the treatment of gastric ulcers. Later on, it appeared that the imidazolic ring present in histamine and in all H_2-antagonists discussed hitherto was not indispensable to the H_2-antagonistic activity. Thus, *ranitidine*, which possesses a furan ring, has appeared to be even more active than cimetidine. The same proved to be true for *famotidine* and *roxatidine*.

VI. FIFTH STRATEGY: APPLYING BIOPHYSICAL TECHNOLOGIES AND COMPUTATIONAL METHODS

This last strategy combines the increasing knowledge that we have on various targets and the new technologies that we can apply to assist the discovery of small molecules interacting both physically and functionally with the target protein. Progress in molecular and structural biology technologies allowed the identification and experimental characterization of several hundred molecular targets and facilitated the design of drugs at a more rational level.

A. Biophysical Technologies

Knowledge of the three-dimensional structures of protein targets has the potential to greatly accelerate drug discovery, but technical challenges and time constraints have traditionally limited its use to lead optimization. Its application is now being extended beyond structure determination into new approaches for lead discovery. Large panels of biophysical methods are currently used, such as:

- protein crystallization [92];
- ligand-protein co-crystallization and soaking [93];
- ligand-protein H^3 and N^{15}-NMR [94];
- Surface Plasmon Resonance [95];
- Differential Scanning Fluorimetry [96]; and
- Mass spectrometry [97]

Structure–activity relationships by nuclear magnetic resonance have been widely used to detect ligand binding and to give some indication of the location of the binding site. X-ray crystallography has the advantage of defining ligand-binding sites with greater certainty. High-throughput approaches make this method applicable to screening to identify molecular fragments that bind protein targets and to defining precisely their binding sites. X-ray crystallography can then be used as a rapid technique to guide the elaboration of the fragments into larger molecular weight compounds that might be useful leads for drug discovery. The Fragment-Based Drug Discovery (FBDD) approach is detailed in Chapter 7, and we will only stress some key features.

In addition, the integration of these techniques with computational chemistry technologies support the rational design and optimization of hits.

1. *Hits from Fragments*

This approach starts from the screening of a fragment set by technologies allowing the measure of low-affinity interaction with the target protein. The protein characteristics determine the most appropriate technology to be used. For example, in general it is difficult to perform NMR with proteins larger than 40 kDa, however, crystallography could provide valuable structural information. As the size of the compound set does not usually exceed a few thousand, selection criteria tailored to the target should be established. Following hit identification, the challenge is in the conversion of a fragment having a binding affinity in the millimolar range 0.1–1 mM (Kd) into a compound showing a biochemical effect (Kd below 10 μM). The crucial step is the growing of the ligand toward higher binding and biochemical activity on the target [98] using the structural information gained from NMR or crystallography.

2. *Finding Extra Interaction: the Example of a Covalent Drug*

Covalent drugs are not new, and chemists did not wait for these technologies to design such products., Today, however, structural biology information helps to make the design of such molecules more efficient [99]. This method can be applied to the generation of covalent inhibitors, taking advantage of the identification of a nucleophilic group (i.e., Ser OH, Cys SH). The acceptor functionality can be modulated according to the expected profile of the molecule, which is carefully evaluated by specific biochemical assays. The anticancer drug Ibrutinib derives from this structure-based approach [100].

B. Computational Methods

1. *Virtual Screening*

Virtual screening can be defined as a set of computational methods that analyzes large databases or collections of compounds in order to identify potential hit candidates. This search can be performed on corporate libraries and/or on virtual libraries [30].

These *in silico* experiments can complement HTS (and are indeed often combined with the screening campaigns); they can also be performed prior to experimental screening or after HTS to rescue some compounds potentially missed by the *in vitro* readouts (so-called latent hits).

Virtual screening approaches have been traditionally subdivided into two main approaches: ligand-based screening and receptor-based screening. In the ligand-based screening, 2D or 3D chemical structures or molecular descriptors of known actives (and sometimes inactive molecules) are used to retrieve other compounds of interest from a database by applying similarity sourcing techniques or by seeking a common substructure, pharmacophore, or shape parameters within the active set. In structure-based (or receptor-based) screening, the library database compounds are docked into a binding site (or over the entire surface), then ranked using one or several scoring functions. The process can then be repeated if deemed appropriate using different types of post-processing approaches. Flexibilities of the receptor and of the ligand can be treated by different strategies and at different stages of the process depending on the amount of computer resources available and on the type of target. The ligand- and structure-based methods can be combined if the necessary information is available.

Two examples of molecules (among many) that have been developed with the help of *in silico* approaches include tirofiban and cevoglitazar (Figure 4.28). Tirofibran is a reversible antagonist of fibrinogen binding to glycoprotein IIb/IIIa that prevents blood clotting. This compound is one of the first FDA approved drug whose origins can be traced back to a pharmacophore-based virtual screening design [101]. Cevoglitazar, a potent PPAR-α/γ dual agonist was developed by combining the virtual screening of 3D databases and structure-based drug design.

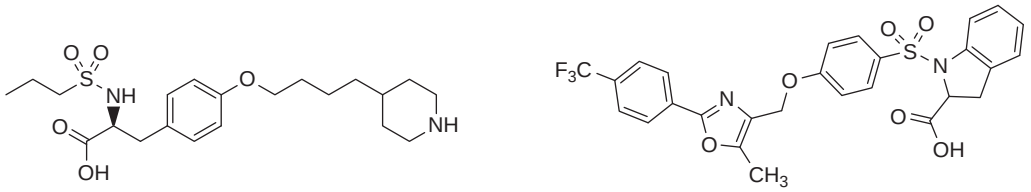


FIGURE 4.28 Examples of molecules originating from virtual screening approaches.

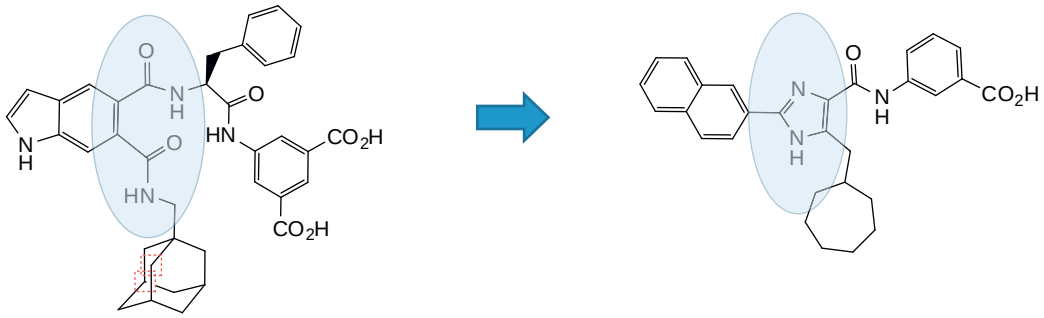


FIGURE 4.29 A scaffold hopping strategy towards novel CCK_2 antagonists.

2. *Computational Drug Repurposing*

Several *in silico* approaches have been developed to facilitate compound repositioning. Different concepts are used to try to connect a small molecule with a disease and/or a target. Some approaches involve a compound similarity search, sometimes combined with machine learning strategies; others involve reverse docking or reverse pharmacophore profiling or binding pocket comparisons. Other strategies—known as signature-based drug-repositioning methods—make use of gene signatures derived from disease "omic" data obtained with or without treatments to discover unknown off-targets or unknown disease mechanisms (transcriptomic methods). Side-effect methods highlight drugs with new disease indications by representing a disease on the basis of a set of side effects associated with the treatments. Depending on the new indication, the repurposed drug may need further medicinal chemistry optimization [102–104].

3. *Scaffold Hopping*

Scaffold hopping aims to discover structurally novel compounds by replacing the central core (also called scaffold, chemotype, or framework) structure of a known active compound. The principle of scaffold hopping is to retain enough similarity in the template compound to keep the activity while altering the central core and changing the properties of the molecule or moving it into novel chemical space. Key substituents that are involved in the binding are often retained to maintain biological activity. Usually, the core replacement leads to a new ADMET and IP profile, solving potential problems with the original compound. These computational approaches have been integrated more and more systematically into drug discovery efforts during the last 15 years, and published work in the field has increased exponentially [105].

Numerous methods that enable the generation of scaffold hopping exist and have been compared, including heterocyclic replacement, ring opening and ring closure, peptidomimetics, topology based hopping, 2D fingerprint, and 3D pharmacophore [106–108].

This technology is illustrated by the example in Figure 4.29. The scientists working in the computational chemistry company Cresset® have identified a new CCK2 prototype antagonist by using the new force field, which includes condensed representation of electrostatic, hydrophobic, and shape properties. Such pharmacophoric fingerprints are used to find biologically equivalent replacements for key moieties, enabling the finding of new structures in a new chemical space [109,110]. The compounds identified maintained activity and selectivity for

this receptor over CCK1. In addition, the reduction in molecular weights coupled with lower polarities greatly reduced levels of biliary elimination, which was an issue. It turned them into good lead compounds for the development of drug candidates whose structures were not obviously related to those of the parents.

In the scaffold hopping approach, commercial scaffold libraries can be screened. In addition, new chemical space can be explored by generating virtual libraries of synthesizable compounds like those proposed by Pitt in 2009 [111].

VII. CONCLUSION

The discovery of new lead compounds—and possibly new drugs—can be schematically classified into five approaches. These consist of the improvement of already existing drugs, of systematic screening, of retroactive exploitation of biological information, of attempts toward rational design, and of the use of the target protein structural information. It would be imprudent to compare hastily the merit of each of these approaches. Indeed, "poor" research can end with a universally recognized medicine, and—conversely—a brilliant rational demonstration can remain sterile. It is therefore of highest importance, given the random nature of discovery and the virtual impossibility of planned invention of new active principles, that decision-makers in the pharmaceutical industry employ all five strategies described, and that they realize that these strategies are not mutually exclusive. On the other hand, once a lead compound is discovered and characterized, it would be inappropriate not to study its molecular mechanism of action. Every possible effort should be made in this direction. In conclusion, all strategies resulting in identification of lead compounds are *a priori* equally good and advisable, provided that the research they induce afterwards is done in a rational manner.

References

[1] von Clausewitz C. Vom kriege [On War]. Drittes buch, 1. Kapitel strategie 18. 1973 ed. Bonn: Ferdinand Dümmlers Verlag; 1832 [in German].
[2] Eder J, Sedrani R, Wiesmann C. The discovery of first-in-class drugs: origins and evolution. Nat Rev Drug Discov 2014;13:577–87.
[3] Ramesha CS. How many leads from HTS? Comment. Drug Discov Today 2000;5:43–4.
[4] Wermuth CG. Selective optimization of side activities: another way for drug discovery. J Med Chem 2004;47:1303–14.
[5] Hopkins A, Lanfear J, Lipinski CA, Beeley L. Chemical tools for indications discovery. In: Doherty A, editor. Annual reports in medicinal chemistry, vol. 40. San Diego, CA: Academic Press; 2005. p. 339–48.
[6] Rey A. Dictionnaire historique de la langue française [Historical dictionary of the French language], vol. 1. Paris: Dictionnaires Le Robert; 1992. [in French].
[7] Worms P, Kan J-P, Wermuth CG, Roncucci R, Bizière K. SR 95191 a selective inhibitor of type A monoamine oxidase with dopaminergic properties. I. Psychopharmacological profile in rodents. J Pharmacol Exp Ther 1987;240:241–50.
[8] Jeanmart C, Cotrel C. Synthèse de (chloro-5 pyridyl-2)-6 méthyl-4 pipérazinyl-1) carbonyloxy-5 oxo-7 dihydro-6,7 5H-pyrrolo [3,4-b] pyrazine. C R Acad Sci Paris Ser C 1978;287:377–8. [in French].
[9] Doble A, Canto T, Piot O, Zundel J-L, Stutzmann J-M, Cotrel C, et al. The pharmacology of cyclopyrrolone derivatives acting at the GABAA/benzodiazepine receptor. Adv Biochem Psychopharmacol 1992;47:407–18.
[10] Angell M. The truth about the drug companies: how they deceive us and what to do about it. New York: Random House; 2004.
[11] Giordanetto F, Boström J, Tyrchan C. Follow-on drugs: how far should chemists look? Drug Discov Today 2011;16:722–32.
[12] Thuilier J. Les dix ans qui ont changé la folie [The decade that changed madness]. Paris: Robert Laffont; 1981. p. 253–257. [in French].
[13] Kling J. From hypertension to angina to Viagra. Mod Drug Discov 1988;1:31–8.
[14] Mannhold R. BTS-49465. Drug Fut 1986;11:177–8.
[15] Sternbach LH. The benzodiazepine story. J Med Chem 1979;22:1–7.
[16] Blanchard JC, Boireau A, Garret C, Julou L. In vitro and in vivo inhibition by zopiclone of benzodiazepine binding to rodent brain receptors. Life Sci 1979;24:2417–20.
[17] Suffness M. Taxol: from discovery to therapeutic use. In: Bristol J, editor. Annual reports in medicinal chemistry, vol. 21. San Diego: Academic Press; 1993. p. 305–14.
[18] Lee JA, Uhlik MT, Moxham CM, Tomandl D, Sall DJ. Modern phenotypic drug discovery is a viable, neoclassic pharma strategy. J Med Chem 2012;55:4527–38.
[19] Boyd MR, Paull KD. Some practical considerations and applications of the national cancer institute in vitro anticancer drug discovery screen. Drug Dev Res 1995;34:91–109.
[20] Endo A. Compactin (ML-236B) and related compounds as potential cholesterol-lowering agents that inhibit HMG-CoA reductase. J Med Chem 1985;28:401–5.
[21] Lee TJ. Synthesis, SARs and therapeutic potential of HMG-CoA reductase inhibitors. Trends Pharmacol Sci 1987;8:206–8.
[22] Shu YZ, Cutrone JQ, Klohr SE, Huang S. BMS-192548, a tetracyclic binding inhibitor of neuropeptide Y receptors, from *Aspergillus niger* WB2346. II. Physico-chemical properties and structural characterization. J Antibiot 1995;48:1060–5.

[23] Strader CR, Pearce CJ, Oberlies NH. Fingolimod (FTY720): a recently approved multiple sclerosis drug based on a fungal secondary metabolite. J Nat Prod 2011;74:900–7.
[24] Swinney DC, Anthony J. How where new medicines discovered? Nat Rev Drug Discov 2011;10:507–19.
[25] Nisbet LJ, Westley JW. Developments in microbial products screening. In: Bristol J, editor. Annual reports in medicinal chemistry, vol. 21. San Diego: Academic Press; 1986. p. 149–57.
[26] Hylands PJ, Nisbet LJ. The search for molecular diversity (I). In: Bristol J, editor. Annual reports in medicinal chemistry, vol. 26. San Diego: Academic Press; 1991. p. 259–69.
[27] Dhalin JL, Walters MA. The essential roles of chemistry in high-throughput screening triage. Future Med Chem 2014;6:1265–90.
[28] Keseru GM, Makara GM. Hit discovery and hit-to-lead approaches. Drug Discov Today 2006;11:741–8.
[29] Wigglesworth MJ, Wolfe LA, Wise A. Orphan seven transmembrane receptor screening. Ernst Schering Found Symp Proc 2006;2:105–43.
[30] Clark DE. What has virtual screening ever done for drug discovery. Expert Opin Drug Discov 2008;3:841–51.
[31] Sun D, Chuaqui C, Deng Z, Bowes S, Chin D, Singh J, et al. A kinase-focused compound collection: compilation and screening strategy. Chem Biol Drug Des 2006;67:385–94.
[32] Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 2007;6:881–90.
[33] Hann MM, Keserü GM. Finding the sweet spot: the role of nature and nurture in medicinal chemistry. Nat Rev Drug Discov 2012;11:355–65.
[34] Leeson PD, St-Gallay SA. The influence of the organizational factor on compound quality in drug discovery. Nat Rev Drug Discov 2011;10:749–65.
[35] Perola E. An analysis of the binding efficiencies of drugs and their leads in successful drug discovery programs. J Med Chem 2010;53: 2986–97.
[36] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23:3–25.
[37] Jhoti H, Williams G, Rees DC, Murray CW. The "rule of three" for fragment-based drug discovery: where are we now? Nat Rev Drug Discov 2013;12:644–5.
[38] Gleeson P. Generation of a set of simple, interpretable ADMET rules of thumb. J Med Chem 2008;51:817–34.
[39] Hughes JD, Blagg J, Price DA, Bailey S, Decrescenzo GA, Devraj RV, et al. Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorg Med Chem Lett 2008;18:4872–5.
[40] Ellison CM, Enoch SJ, Cronin MT. A review of the use of in silico methods to predict the chemistry of molecular initiating events related to drug toxicity. Expert Opin Drug Metab Toxicol 2011;7:1481–95.
[41] Benigni R, Bossa C. Mechanisms of chemical carcinogenicity and mutagenicity: a review with implications for predictive toxicology. Chem Rev 2011;111:2507–36.
[42] Rishton GM. Non-leadlikeness and leadlikeness in biochemical screening. Drug Discov Today 2003;8:86–96.
[43] Baell JB, Holloway GA. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 2010;53:2719–40.
[44] Villoutreix BO, Lagorce D, Labbé CM, Sperandio O, Miteva MA. One hundred thousand mouse clicks down the road: selected online resources supporting drug discovery collected over a decade. Drug Discov Today 2013;18:1081–9.
[45] Sheppard DW, Lipkin MJ, Harris CJ, Catana C, Stouten PF. Strategies for small molecule library design. Curr Pharm Des 2014;20:3314–22.
[46] Macarron R, Banks MN, Bojanic D, Burns DJ, Cirovic DA, Garyantes T, et al. Impact of high-throughput screening in biomedical research. Nat Rev Drug Discov 2011;10:188–95.
[47] Sperandio O, Reynès CH, Camproux AC, Villoutreix BO. Rationalizing the chemical space of protein–protein interaction inhibitors. Drug Discov Today 2010;15:220–9.
[48] Böhm H-J, Klebe G, Kubinyi H. Wirkstoffdesign [Drug design]. Heidelberg: Spektrum Akademischer Verlag; 1996. p. 599 [in German].
[49] Wermuth CG. Search for new lead compounds: the example of the chemical and pharmacological dissection of aminopyridazines. J Heterocycl Chem 1998;35:1091–100.
[50] Wermuth CG, Clarence-Smith K. "Drug-like" leads: bigger is not always better. Pharm News 2000;7:53–7.
[51] Poroikov V, Akimov D, Shabelnikova E, Filimonov D. Top 200 medicines: can new actions be discovered through computer-aided prediction? SAR QSAR Environ Res 2001;12:327–44.
[52] Wermuth CG. The "SOSA" approach: an alternative to high-throughput screening. Med Chem Res 2001;10:431–9.
[53] Wermuth CG. Possible alternatives to high-throughput screening. In: Chorgade MS, editor. Drug discovery and development, vol. 1. Hoboken, NJ: John Wiley & Sons; 2006. p. 213–32.
[54] Stein PD, Hunt JT, Floyd DM, Moreland S, Dickinson KEJ, Mitchell C, et al. The discovery of sulfonamide endothelin antagonists and the development of the orally active ETA antagonist 5-(dimethylamino)-N-(3,4-dimethyl-5-isoxazolyl)-l-naphthalenesulfonamide. J Med Chem 1994;37:329–31.
[55] Murugesan N, Gu Z, Spergel S, Young M, Chen P, Mathur A, et al. Biphenylsulfonamide endothelin receptor antagonists. 4. Discovery of N-[[(4,5-dimethyl-3-isoxazolyl)amino]sulfonyl]-4-(2-oxazolyl)[1,1′-biphenyl]-2-yl]methyl]-N,3,3-trimethylbutanamide (BMS-207940), a highly potent and orally active ETA selective antagonist. J Med Chem 2003;46:125–37.
[56] Wermuth CG, Schlewer G, Bourguignon J-J, Maghioros G, Bouchet M-J, Moire C, et al. 3-Aminopyridazine derivatives with atypical antidepressant, serotonergic and dopaminergic activities. J Med Chem 1989;32:528–37.
[57] Wermuth CG. Aminopyridazines: an alternative route to potent muscarinic agonists with no cholinergic syndrome. Il Farmaco 1993; 48:253–74.
[58] Wermuth CG, Bourguignon J-J, Hoffmann R, Boigegrain R, Brodin R, Kan J-P, et al. SR 46559A and related aminopyridazines are potent muscarinic agonist with no cholinergic syndrome. Bioorg Med Chem Lett 1992;2:833–6.

[59] Pilla M, Perachon S, Sautel F, Garrido F, Mann A, Wermuth CG, et al. Selective inhibition of cocaine-seeking behaviour by a partial dopamine D_3 agonist. Nature 1999;400:371–5.
[60] Prestwick Chemical, Inc. KSN Suite 1475, Washington, DC, 20006-1202. <http://www.prestwickchemical.com/> In.
[61] Kleemann A, Engel J. Pharmazeutische wirkstoffe: synthese, patente, anwendungen [Pharmaceutical drugs: synthesis, patents, applications] (Preface). Stuttgart: Georg Thieme Verlag; 1982. [in German].
[62] Price JR. Antifertility agents of plant origin. In: Austin CR, Perry JS, editors. A symposium on agents affecting fertility. Boston: Little, Brown, and Co; 1965. p. 3–17.
[63] Laborit H, Huguenard P, Alluaume R. Un nouveau stabilisateur végétatif, le 4560 RP [A new vegetative stabilizer, 4560 RP]. Presse Méd 1952;60:206–8. [in French].
[64] Maxwell RA, Eckhardt SB. Drug discovery: a casebook and analysis. Clifton, NJ: Humana Press; 1990.
[65] Stemp G, Evans JM. Discovery and development of cromakalim and related potassium channel activators. In: Ganellin CR, Roberts SM, editors. Medicinal chemistry, the role of organic chemistry in drug research. 2nd ed. London: Academic Press; 1993. p. 141–62.
[66] Hamilton TC, Weir SW, Weston AH. Comparison of the effects of BRL 34915 and verapamil on electrical and mechanical activity in rat portal vein. Br J Pharmacol 1986;88:103–11.
[67] Gerber JG, Nies AS. Antihypertensive agents and the drug therapy of hypertension. In: Goodman-Gilman A, Rall TW, Nies AS, Taylor P, editors. The pharmacological basis of therapeutics. 8th ed. New York: Pergamon Press; 1990. p. 784–813.
[68] Evans JM, Fake CS, Hamilton TC, Poyser RH, Watts EA. Synthesis and antihypertensive activity of substituted trans-4-amino-3,4-dihydro-2,2-dimethyl-2H-1-benzopyran-3-ols. J Med Chem 1983;26:1582–9.
[69] Evans JM, Fake CS, Hamilton TC, Poyser RH, Showell GA. Synthesis and antihypertensive activity of 6,7-disubstituted trans-4-amino-3,4-dihydro-2,2-dimethyl-2H-1-benzopyran-3-ols. J Med Chem 1984;27:1127–31.
[70] Sneader W. Drug prototypes and their exploitation. Chichester: John Wiley & Sons; 1996. p. 242.
[71] Iyer CGS, Languillon J, Ramanujam K. WHO coordinated short-term double-blind trial with thalidomide in the treatment of acute lepra reactions in male lepromatus patients. Bull WHO 1971;45:719–32.
[72] Cushman M, Jayaraman M, Vroman JA, Fukunaga AK, Fox BM, Kohlhagen G, et al. Synthesis of new indeno[1,2-c]isoquinolines: cytotoxic non-camptothecin topoisomerase I inhibitors. J Med Chem 2000;43:3688–98.
[73] Korth C, May BCH, Cohen FE, Prusiner SB. Acridine and phenothiazine derivatives as pharmacotherapeutics for prion disease. Proc Natl Acad Sci USA 2001;98:9836–41.
[74] Terret NK, Bell AS, Brown D, Ellis P. Sildenafil (Viagra), a potent and selective inhibitor of type 5 cGMP phosphodiesterase with utility for the treatment of male erectile dysfunction. Bioorg Med Chem Lett 1996;6:1819–24.
[75] Boolell M, Allen MJ, Ballard SA, Gepi-Attee S, Muirhead GJ, Naylor AM, et al. An orally activetype 5 cyclic GMP-specific phosphodiesterase inhibitor for the treatment of penile erectile dysfunction. Int J Urol Res 1996;8:47–52.
[76] Foye WO, Lemke TL, Williams DA. Principles of medicinal chemistry. Baltimore: Lea & Febiger; 1995. p. 995.
[77] Neuworth MB, Laufer RJ, Barnhart JW, Sefranka JA, McIntosh DD. Synthesis and hypocholesterolemic activity of alkylidenedithio bisphenols. J Med Chem 1970;13:722–5.
[78] Noble RL, Beer CT, Cutts JH. Role of chance-observations in chemotherapy: *Vinca rosea*. Ann NY Acad Sci 1958;76:882–94.
[79] Johnson IS, Wright HF, Svoboda GH, Vlantis J. Antitumor principles derived from *Vinca rosea* Linn. I. Vincaleukoblastine and leurosine. Cancer Res 1960;20:1016–22.
[80] Johnson IS, Armstrong JG, Gorman M, Burnett Jr JP. The vinca alkaloids: a new class of oncolytic agents. Cancer Res 1963;23:1390–427.
[81] Buxton ILO, Cheek DJ, Eckman D, Westfall DP, Sanders KM, Keef KD. N^G-Nitro-l-arginine methyl ester and other alkyl esters of arginine are muscarinic antagonists. Circ Res 1993;72:387–95.
[82] Schrader G. Die Entwicklung neuer Insektizide auf Grundlage von Organischen Fluor und Phosphorverbindungen [The development of new insecticides on the basis of organic fluorine and phosphorus compounds]. Monographie 62. Weinheim: Verlag Chemie; 1952. [in German].
[83] Janssen PAJ, Van de Westeringh C, Jagneau AWM, Demoen PJA, Hermans BKF, Van Daele GHP, et al. Chemistry and pharmacology of CNS depressants related to 4-(4-hydroxy-4-phenylpiperidino)butyrophenone. Part I: synthesis and screening data in mice. J Med Pharm Chem 1959;1:281–97.
[84] Janssen PAJ. The evolution of the butyrophenones, haloperidol and trifluperidol, from meperidine-like 4-phenylpiperidines. Int Rev Neurobiol 1965;8:221–63.
[85] Albert A. Selective toxicity. 6th ed. London: Chapman and Hall; 1979. p. 221.
[86] Shen TY. Perspectives in non-steroidal antiinflammatory agents. Angew Chem Int Ed 1972;11:460–72.
[87] Ehringer H, Hornykiewicz O. Verteilung von Noradrenalin und Dopamin (3-hydroxytyramin) im Gehirn des Menschen und ihr Verhalten bei Erkrankungen des extrapyramidalen systems [Distribution of norepinephrine and dopamine (3-hydroxytyramine) in the human brain and their behavior in diseases of the extrapyramidal system]. Klin Wochenschr 1960;38:1236–9. [in German].
[88] Cotzias GE, Van Woert MH, Schiffer LM. Aromatic amino acids and modification of Parkinsonism. N Engl J Med 1967;276:374–9.
[89] Cushman DW, Cheung HS, Sabo EF, Ondetti MA. Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids. Biochemistry 1977;16:5485–91.
[90] Black JW, Duncan WAM, Durant JC, Ganellin CR, Parsons ME. Definition and antagonism of histamine H2-receptors. Nature 1972;236:385–90.
[91] Ganellin CR. Cimetidine. In: Bindra JS, Lednicer D, editors. Chronicles of drug discovery, vol. 1. New York: John Wiley & Sons; 1982. p. 1–38.
[92] Blundell TL, Jhoti H, Abell C. High-throughput crystallography for lead discovery in drug design. Nat Rev Drug Discov 2002;1:45–54.
[93] Davies TG, Tickle IJ. Fragment screening using X-ray crystallography. Top Curr Chem 2012;317:33–59.
[94] Vajpai N, Strauss A, Fendrich G, Cowan-Jacob SW, Manley PW, Grzesiek S, et al. Solution conformations and dynamics of Abl kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib. J Biol Chem 2008;283:18292–302.

[95] Zeder-Lutz G, Choulier L, Besse M, Cousido-Siah A, Figueras FXR, Didier B, et al. Validation of surface plasmon resonance screening of a diverse chemical library for the discovery of protein tyrosine phosphatase1b binders. Anal Biochem 2012;421:417–27.
[96] Cummings MD, Farnum MA, Nelen MI. Universal screening methods and applications of Thermofluor. J Biomol Screen 2006;11:854–63.
[97] Maple JM, Garlish RA, Rigau-Roca L, Porter J, Whitcombe I, Prosser CE, et al. Automated protein-ligand interactionscreening by mass spectrometry. J Med Chem 2012;55:837–51.
[98] Bottegoni G, Favia AD, Recanatini M, Cavalli A. The role of fragment-based and computational methods in polypharmacology. Drug Discov Today 2012;17:23–34.
[99] Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. Nat Rev Drug Discov 2011;10:307–17.
[100] Pan Z, Scheerens H, Li SJ, Schultz BE, Prengeler PE, Burrill AC, et al. Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. Chem Med Chem 2007;2:58–61.
[101] Hartman GD, Egbertson MS, Halczenko W, Laswell WL, Duggan ME, Smith RL, et al. Non-peptide fibrinogen receptor antagonists. 1. Discovery and design of exosite inhibitors. J Med Chem 1992;35:4640–2.
[102] Ekins S, Williams AJ, Krasowski MD, Freundlich JS. In silico repositioning of approved drugs for rare and neglected diseases. Drug Discov Today 2011;16:298–310.
[103] Liu Z, Fang H, Reagan K, Xu X, Mendrick DL, Slikker Jr W, et al. In silico drug repositioning: what we need to know. Drug Discov Today 2012;18:110–15.
[104] Jin G, Wong STC. Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines. Drug Discov Today 2014;19:637–44.
[105] Schneider G, Neidhart W, Giller T, Schmid G. "Scaffold-hopping" by topological pharmacophore search: a contribution to virtual screening. Angew Chem Int Ed 1999;38:2894–6.
[106] Sun H, Tawa G, Wallqvist A. Classification of scaffold hoping approaches. Drug Discov Today 2012;17:310–24.
[107] Böhm HJ, Flohr A, Stahl M. Scaffold hopping. Drug Discov Today Technol 2004;1:217–24.
[108] Langdon SR, Ertl P, Brown N. Biosisosteric replacement and scaffold hopping in lead generation and optimization. Mol Inform 2010;29:366–85.
[109] Cheeseright T, Mackey MD, Rose S, Vinter A. Molecular field extrema as descriptors of biological activity: definition and validation. J Chem Inf Model 2006;46:665–76.
[110] Low CMR, Buck IM, Cooke T, Cushnir JR, Kalindjian SB, Kotecha A, et al. Scaffold hopping with molecular field points: identification of a cholecystokinin-2 (CCK_2) receptor pharmacophore and its use in the design of a prototypical series of pyrrole- and imidazole-based CCK_2 antagonists. J Med Chem 2005;48:6790–802.
[111] Pitt WR, Parry DM, Perry BG, Groom CR. Heteroaromatic rings of the future. J Med Chem 2009;52:2952–63.

CHAPTER

5

Natural Products as Pharmaceuticals and Sources for Lead Structures*

David J. Newman[1], *Gordon M. Cragg*[1] *and David G.I. Kingston*[2]

[1]NIH Special Volunteer, Natural Products Branch, Developmental Therapeutics Program, Division of Cancer Treatment and Diagnosis, United States National Cancer Institute, Frederick, Maryland, USA; [2]Department of Chemistry and the Virginia Tech Center for Drug Discovery, Virginia Tech, Blacksburg, Virginia, USA

OUTLINE

Accuse not Nature, she hath done her part; do thou but thine. **Milton, Paradise Lost**

*Note: This chapter reflects the opinions of the authors, not necessarily those of the US Government

I. INTRODUCTION

Throughout the ages, humans have relied on nature to cater for their basic needs, not the least of which are medicines for the treatment of a myriad of diseases. Plants, in particular, have formed the basis of sophisticated traditional medicine systems, with the earliest records documenting the uses of approximately 1000 plant-derived substances in Mesopotamia, and the Ebers Papyrus dating from 1500 BCE in Egypt, documenting over 700 drugs, mostly of plant origin [1]. The first record of the Chinese *Materia Medica* documenting 52 prescriptions dates from about 1100 BCE, and was followed by works such as the Shennong Herbal (c.100 BCE; 365 drugs) and the Tang Herbal (659 CE; 850 drugs) [2]. Documentation of the Indian Ayurvedic system also dates from before 1000 BCE (Charaka; Sushruta and Samhitas with 341 and 516 drugs respectively) [3,4].

The Greeks and Romans contributed substantially to the rational development of the use of herbal drugs in the ancient Western world. Dioscorides, a Greek physician (100 CE), accurately recorded the collection, storage, and use of medicinal herbs during his travels with Roman armies throughout the then "known world," while Galen (130–200 CE), a practitioner and teacher of pharmacy and medicine in Rome, is well known for his complex prescriptions and formulae used in compounding drugs. However, it was the Arabs who preserved much of the Greco-Roman expertise during the Early Middle Ages (5th–12th centuries), and who expanded it to include the use of their own resources together with Chinese and Indian herbs unknown to the Greco-Roman world. A comprehensive review of the history of medicine may be found on the website of the National Library of Medicine (NLM), United States National Institutes of Health (NIH), at http://www.nlm.nih.gov/hmd/medieval/arabic.html.

Plant-based systems continue to play an essential role in healthcare, and their use by different cultures has been extensively documented [5,6]. It has been estimated by the World Health Organization that approximately 80 percent of the world's inhabitants rely mainly on traditional medicines for their primary health care, while plant products also play an important role in the health care systems of the remaining 20 percent of the population who mainly reside in developed countries [7].

II. THE IMPORTANCE OF NATURAL PRODUCTS IN DRUG DISCOVERY AND DEVELOPMENT

The continuing value of natural products as sources of potential chemotherapeutic agents has been reviewed by Newman and Cragg [8]. An analysis of the sources of new drugs from 1981 to 2010 classified these compounds as N (an unmodified natural product), NB (defined botanical), ND (a modified natural product), S (a synthetic compound with no natural product conception), S*, S*/NM (a synthetic compound with a natural product pharmacophore; /NM indicating competitive inhibition), and S/NM (a synthetic compound showing competitive inhibition of the natural product substrate). This analysis indicated that 66 percent of the 1073 small molecule, new chemical entities (NCEs) are formally synthetic, but 19 percent correspond to synthetic molecules containing pharmacophores derived directly from natural products classified as S* and S*/NM. Furthermore, 11 percent are actually modeled on a natural product inhibitor of the molecular target of interest, or mimic (i.e., competitively inhibit) the endogenous substrate of the active site, such as ATP (S/NM). Thus, only 36 percent of the 1073 NCEs can be classified as truly synthetic (i.e., devoid of natural inspiration) in origin (See Figure 5.1). A recent, unpublished analysis of 1135 small molecule NCEs reported from 1981 to 2012 indicates a similar division between the S category (35 percent) and the naturally-derived and inspired categories (65 percent). In the area of anti-infectives (antibacterial, -fungal, -parasitic, and -viral), close to 70 percent are naturally derived or inspired (N; ND; S*; S*/NM; S/NM), while in the cancer treatment area 77.8 percent are in this category, with the figure being 63 percent if the S/NM category is excluded [8].

In recent years, a steady decline in the output of the R&D programs of the pharmaceutical industry was reported, with the number of new small molecule active substances—also known as new chemical entities (NCEs)—hitting a twenty year low of eighteen in 2004. Though rising a little, only twenty were approved by all approving agencies world-wide in 2010 [9]. Recently, in its summary of "Novel New Drugs 2013", the FDA reported that twenty-seven new drugs (NMEs) were approved by them in 2013, while from 2004 through 2012, the average number approved by the FDA was twenty-six) [10]. This reflects a continuing relatively low output of NCEs (equivalent to NMEs). While various factors have been held to blame for this downturn, it is significant that the past twenty years have seen a decline in interest in natural products on the part of major pharmaceutical companies in favor of reliance on new chemical techniques—such as combinatorial chemistry—for generating molecular libraries. The realization that the

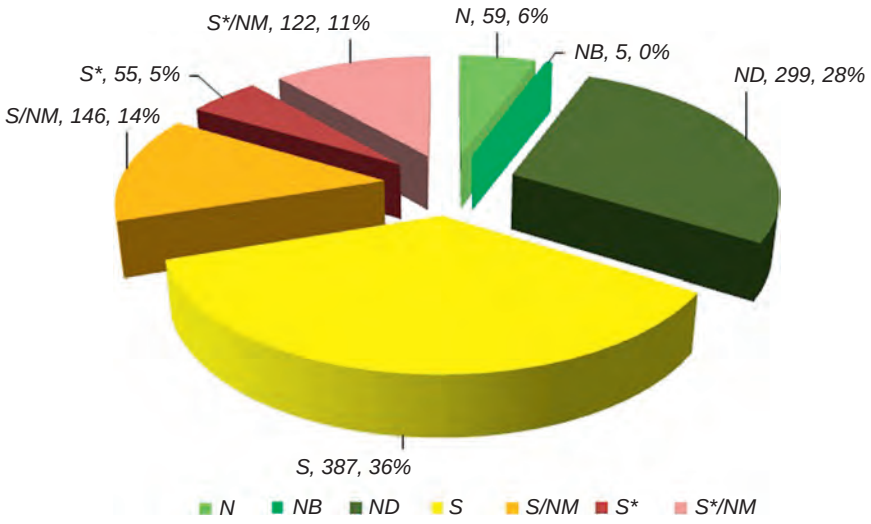


FIGURE 5.1 Sources of drugs.

number of NCEs in drug development pipelines is declining may have led to the rekindling of interest in "rediscovering natural products" [11], as well as the heightened appreciation of the value of natural product-like models in "improving efficiency" in so-called diversity-oriented synthesis [12]. The urgent need for the discovery and development of new pharmaceuticals for the treatment of cancer, AIDS, and infectious diseases, as well as a host of other diseases, demands that all approaches to drug discovery be exploited aggressively. It is clear that nature has played—and will continue to play—a vital role in the drug discovery process [13]. Very recently Nicolaou, in his essay on "Advancing the Drug Discovery and Development Process," stated "The three-dimensional structures of natural products and their wealth of chiral centers should serve as an inspiration and motivation for drug designers. The dimensionality and chirality of biological receptors and the fact that natural products have evolved along and against such biomolecules explains their diverse, potent, and often selective biological properties. Employing them and molecules like them as leads and introducing some of their structural features in drug designs makes good sense and should be a complementary approach to the currently employed drug design practices." He later added, "It might be true that natural products chemistry requires longer-term plans and higher initial investments, but in the long run, the endeavor pays off as demonstrated by its rich and glorious history" [14].

A. The Origin of Natural Products

While the contributions of natural secondary metabolites (all nonproteinaceous natural products would fall under this term) to modern medicine are abundantly clear, the question of their origins has long intrigued chemists and biochemists. Six major hypotheses have been proposed, and these have been well summarized by Haslam [15]: {1} simple waste products; {2} substances that had a functional metabolic role that has now been lost; {3} products of random mutations; {4} examples of "evolution in progress"—pool of compounds from which new biochemical processes can emerge; {5} "The *processes* of secondary metabolism, rather than the *products* (secondary metabolites) are important;" and {6} products that play a key role in the organism's survival, providing defensive substances or other physiologically important compounds.

Although each of the above has (or has had) its supporters, Williams et al. [16] and Harborne [17], among others, argue convincingly that the weight of the evidence is behind the sixth hypothesis. Indeed, it seems reasonable to assume that in many instances the production of these complex and often toxic chemicals has evolved over eons as a means of chemical defense by essentially stationary organisms, such as plants and many marine invertebrates. Thus, pupae of the coccinellid beetle, *Epilachna borealis*, appear to exert a chemical defensive mechanism against predators through the secretion of droplets from their glandular hairs containing a library of hundreds of large-ring (up to ninety-eight members) macrocyclic polyamines [18]. These libraries are built up from three simple (2-hydroxyethylamino)-alkanoic acid precursors, and are clear evidence that combinatorial chemistry has been pioneered and widely used in nature for the synthesis of biologically active compound libraries. A further example is the venom composed of combinatorial libraries of several hundred peptides injected by

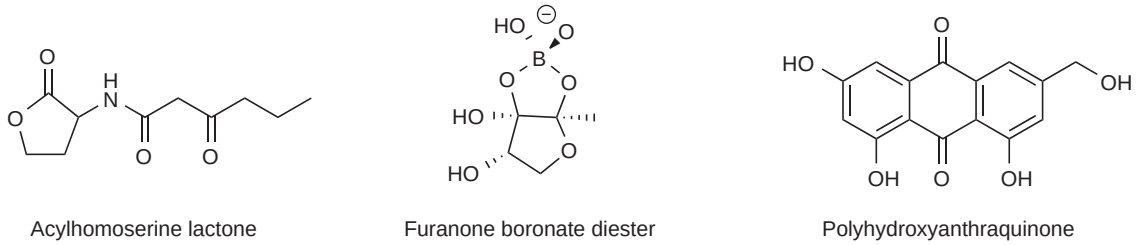


FIGURE 5.2 Quorum-sensing compounds and potential QS-inhibitors.

species of the cone snail genus, *Conus*, to stun their prey prior to capture [19,20]. One component of this mixture has been developed as Ziconotide, a nonnarcotic analgesic that is currently marketed as Prialt® [21].

Microorganisms are reported to kill sensitive strains of the same or related microbial species through excretion of antimicrobial toxins [22], resembling the process of allelopathy whereby plants release toxic compounds in order to suppress the growth of neighboring plants [23,24]. Bacteria also use a cell-to-cell "chemical language" as a signaling mechanism known as quorum-sensing to control the density of population growth and so-called biofilm formation. The best studied of these compounds are the acyl homoserine lactones (AHLs) exemplified by compounds—such as N-3-oxohexanoyl-l-homoserine lactone (Figure 5.2) from *Vibrio fisheri*, and a furanone boronate diester (Autoinducer-2; Al-2)—that appears to be a universal signal (Figure 5.2) promoting the activation of genes promoting virulence, spore formation, biofilm formation, and other phenomena [25,26].

Research into the inhibition of quorum sensing as a strategy for the discovery of novel antimicrobial agents to control the virulence of deadly and increasingly drug-resistant bacteria such as *Pseudomonas aeruginosa*, *Escherichia coli*, and *Staphylococcus aureus* is progressing as a means to combat infections caused by these and other pathogens [27–31]. A recent paper in addition to those referenced above shows the potential of the often observed but rarely investigated fungal "guttates" (also known as "exudates") as sources of quorum sensing inhibitors against *S. aureus*. This work demonstrated that polyhydroxyanthraquinones (Figure 5.2) are structures with potential [32]. Two very recent papers showing the potential of natural products or modified natural products as inhibitors of biofilms were published in the journal *Tetrahedron* in mid-2014. The first by Melander et al. dealt with the potential for indole-containing compounds as potential control agents [33], and the second by Fletcher et al. covered a variety of natural product structures with inhibitory potential [34]. Both of these papers show a variety of structural types that can be modified by classical medicinal chemistry to improve their activities.

B. The Uniqueness of the Natural Products Approach

Natural products are generally complex chemical structures, whether they are cyclic peptides like cyclosporin A or complex diterpenes like paclitaxel. Inspection of the structures that are discussed in Section IV.E is usually enough to convince any skeptic that few of them would have been discovered without application of natural products chemistry. It has already been noted that "the three-dimensional structures of natural products and their wealth of chiral centers should serve as an inspiration and motivation for drug designers" [14], In addition, recognition and appreciation of the value of natural product-like models in "improving efficiency" in so-called diversity-oriented synthesis has also been mentioned [12].

Structural diversity is not the only reason why natural products are of interest to drug development, since they often provide highly selective and specific biological activities based on mechanisms of action. Two very good examples of this are the HMG-CoA reductase inhibition exhibited by lovastatin, and the tubulin-assembly promotion activity of paclitaxel. These activities would not have been discovered without the natural product lead and investigation of their mechanisms of action. A striking illustration of the influence of natural products on many of the enzymatic processes operative in cell cycle progression may be found in a paper published by Meijer in 2003 on natural products and the cell cycle [35], with a modified version shown below in Figure 5.3. The bioactivity of natural products stems from the hypothesis that essentially all natural products have some receptor-binding activity. The problem is to find to which receptor a given natural product is binding. Viewed another way, a given organism provides the investigator with a complex library of unique bioactive constituents, analogous to the library of crude synthetic products initially produced by combinatorial chemistry techniques. The natural products approach can thus be seen as complementary to the synthetic approach, each providing access to (initially) different lead structures. In addition, development of an active natural product structure by

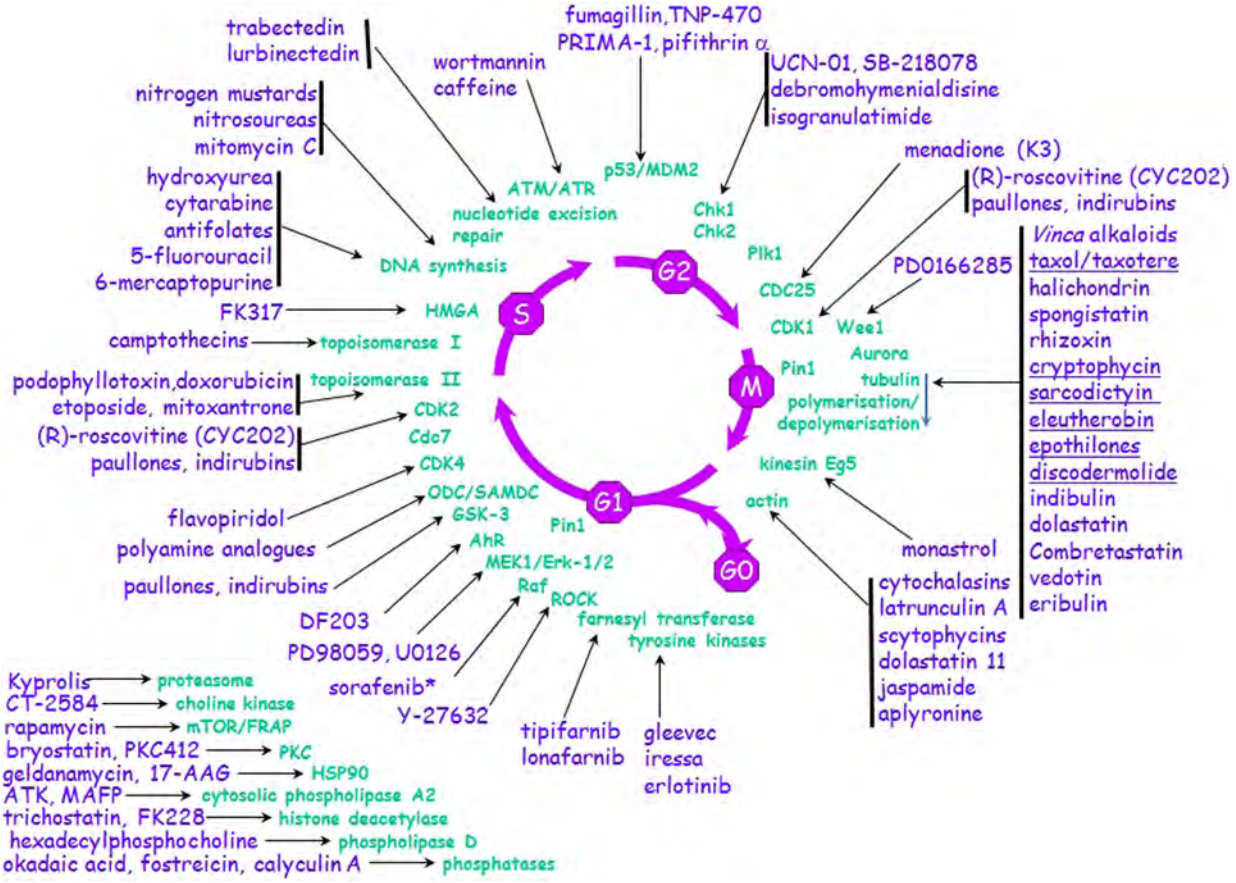


FIGURE 5.3 Natural products and the cell cycle.

combinatorially directed synthesis is an extremely powerful tool. The task of the natural products researcher is thus to select those compounds of pharmacological interest from the "natural combinatorial libraries" produced by extraction of organisms. Fortunately, the means to do this efficiently are now at hand.

C. The Impact of New Screening Methods

In the early days of natural products research, new compounds were simply isolated at random, or at best by the use of simple broad-based bioactivity screens based on antimicrobial or cytotoxic activities. Although these screens did result in the isolation of many bioactive compounds [9], they were considered to be too nonspecific for the next generation of drugs. Fortunately, a large number of robust and specific biochemical and genetics-based screens using transformed cells, a key regulatory intermediate in a biochemical or genetic pathway, or a receptor–ligand interaction (often derived from the explosion in genomic information since the middle 1990s) are now in routine use. These screens will permit the detection of bioactive compounds in the complex matrices that are natural product extracts with greater precision. High-throughput assays, where large numbers of samples can be screened in a short period of time, are becoming less expensive, and such assays are moving from the industrial or industrial-academic consortium-based groups to academia in general, with specific expression systems being employed as targets for natural product lead discovery [36]. The application of new techniques, including new fluorescent assays, NMR, affinity chromatography, and DNA microarrays, led to significant advances in the effectiveness of high throughput screening [37,38].

The screens themselves are all highly automated and are high-throughput screens (HTS; upwards of 50,000 assay points per day in a number of cases). The successful application of such high throughput screening was illustrated by the discovery of a new antibiotic, platensimycin, by a team of scientists from Merck Research Laboratories. It has *in vitro* activity against several drug-resistant bacteria, is a selective FabF inhibitor, and was discovered through the testing of a library of 250,000 natural product extracts in a custom-designed assay

involving an engineered strain of *Staphylococcus aureus* incorporating the fatty acid synthase pathway enzyme, FabF [39]. In early 2014, Ghosh and Xi published an up-to-date article on the syntheses surrounding both platensimycin and its close structural relative platencin. This article should be consulted in order to appreciate the manifold ways in which these complex structures can now be synthesized under conditions that allow the production of novel analogues [40].

III. THE DESIGN OF AN EFFECTIVE NATURAL-PRODUCTS-BASED APPROACH TO DRUG DISCOVERY

There are four major elements in the design of any successful natural products-based drug discovery program: acquisition of biomass, effective screening, bioactivity-driven fractionation, and rapid and effective structure elucidation (which includes dereplication). Although some of these have been mentioned earlier, it is instructive to bring them together here.

A. Acquisition of Biomass

The acquisition of biomass has undergone a very significant transition from the days when drug companies and others routinely collected organisms without any thought of ownership by—or reimbursement to—the country of origin. Today, thanks to the Convention on Biodiversity (CBD) and similar documents and agreements, such as the US National Cancer Institute's Letter of Collection (NCI's LOC: http://dtp.nci.nih.gov/branches/npb/loc.pdf), all ethical biomass acquisitions now include provisions for the country of origin to be recompensed in some way for the use of its biomass. It should be noted that the LOC predated the CBD by three years. Its tenets, as a minimum, must be adhered to by any investigator who has his or her collections funded by the NCI/NIH. Such recompense to the country of origin is best provided through formal agreements with government organizations and collectors in the host country, with such agreements providing not only for reimbursement of collecting expenses but also for further benefits (often in the form of milestone and/or royalty payments) in the event that a drug is developed from a collected sample. The impact of the CBD on natural products research as related to the experiences of the NCI and the US government-sponsored International Cooperative Biodiversity Groups (ICBG) program has been discussed [41], with discussions of the methods used by various groups published [42–44].

It is axiomatic that all samples collected, irrespective of type of source, must if at all possible be fully identified to genus and species. Such identification is usually possible for all plant species, but it is not always possible for microbes and marine organisms. Voucher specimens should be provided to an appropriate depository in the host country as well as to a similar operation in the home country of the collector. The selection of plant samples often raises the question of the ethnobotanical/ethnopharmacological approach versus a random approach. The former method, which usually involves the selection of plants that have a documented (written or oral) use by native healers, is attractive in that it can tap into the empirical knowledge developed over centuries of use by large numbers of people. In addition, the bioactive constituents may be considered as having had a form of continuing clinical trial in humans. The benefits of this approach have been extolled in several articles in 1990 that predate the CBD [7,45,46], and one author provides personal experience of the effectiveness of some jungle medicines [47]. The weakness of the ethnobotanical approach has always been that it is slow, requiring careful interviewing of native healers by skilled scientists, including ethnobotanists, anthropologists, trained physicians, and pharmacologists. In addition, the quoted folkloric activity in the collected plant(s) may not be detectable by the particular screens in use by the screening laboratory. Where ethnobotanical approaches have the highest possibility of success is in studies related to overt diseases/conditions such as parasitic infections, fungal sores, and contraception/conception, to name a few. In such cases, there are adequate controls, even on the same patient. Where there does not yet appear to be any successful relationship is in diseases such as cancer and HIV-related conditions, where extensive laboratory-based testing of the patient is required for an accurate diagnosis.

1. Classical Natural Sources: Untapped Potential

Despite the intensive investigation of terrestrial flora, it was estimated that only 5–15 percent of the approximately 300,000 species of higher plants has been systematically investigated, chemically and pharmacologically [48,49], and this figure has not materially altered since these publications. Exposure of the roots of

hydroponically grown plants to chemical elicitors has been reported to selectively and reproducibly induce the production of bioactive compounds [50], while feeding of seedlings with derivatives of selected biosynthetic precursors can lead to the production of nonnatural analogues of the natural metabolites. This has been demonstrated in the production of nonnatural terpene indole alkaloids related to the vinca alkaloids through the feeding of seedlings of *Catharanthus roseus* with various tryptamine analogues [51]. Further work on aeroponically-mediated production of plant secondary metabolites is discussed later in the chapter (cf. Section III.H).

The potential of the marine environment as a source of novel drugs remains comparatively unexplored [52,53]. Until recently, the investigation of the marine environment has largely been restricted to tropical and subtropical regions, but colder climes are now being explored as discussed later in Section III.B.2 on extremophiles.

B. The Unexplored Potential of Microbial Diversity

For more than fifty years, large and small pharmaceutical companies fermented millions of soil isolates in order to find new compounds from microbial sources, and then over time realized that they were isolating the same metabolites. For a multiplicity of reasons, most companies decided to jettison their fermentation-based discovery programs from the middle of the 1980s to the early 2000s and distributed their microbial collections to smaller groups or simply archived them.

1. Next Generation Sequencing/Epigenetics

Today, due to the development of new sequencing technologies, access to microbial genomes coupled to chemoinformatic tools permits the elucidation and activation of known and previously unknown biosynthetic pathways in both prokaryotes and eukaryotes. Due to the fast evolution of next-generation sequence technologies (also known as NGS or post-Sanger sequencing), biologists and chemists now have volumes of sequence data for the genes involved in natural product biosynthesis. To give an idea of the drop in cost, Sanger sequencing provides data at a cost of around $500.00/Mbp, but some NGS technologies can rapidly produce data for well under $3.00/Mbp [54]. To give some perspective as to the use of low-cost sequencing, a paper by Shendure [54] has been cited over 1500 times at the time of writing (September 2014), and a very recent paper, though not on microbes, demonstrates the power of the current techniques by deriving the complete genomic sequence of a Neanderthal and comparing the genome structure with other hominoids [55].

As a result of this very significant drop in cost, a number of large-scale sequencing initiatives were initiated, such as the Genomic Encyclopedia of Bacteria and Archaea [56] and the Human Microbiome Project [57]. Thus, the number of complete published whole microbial genome sequences has skyrocketed, providing more opportunities for genome mining to discover novel natural products, though one should be cautious and bear in mind the caveats expressed in the recent review by Smith al., where they point out that function also has to be identified [58].

Investigation of microbial genome sequences over the last fifteen or so years has revealed that a large number of the potential biosynthetic gene clusters in microorganisms are “cryptic” or “silent.” For example, sequencing *Streptomyces* genomes has revealed that each strain has the potential to produce at least twenty or more secondary metabolites, but only a fraction of these have been identified by use of conventional fermentation methods [59]. Due to the then current wisdom that fungal secondary metabolites were spread across their chromosomes, it was not until relatively recently that these organisms were accorded their just due. From conventional genomic studies rather than by genetic manipulation of promoters, von Döhren reported that the model soil fungus *Aspergillus nidulans* had twenty-eight putative PKS and twenty-four NRPS gene clusters. Thus, at least fifty-two secondary metabolites might be produced if these could be activated [60].

This report was followed by a 2012 review from the Keller group at Wisconsin that extended the analyses to eight other species of *Aspergillus* [61]. In this report, they identified a lower limit of thirty-three and an upper limit of seventy-nine putative clusters, but they excluded any terpene synthases on the eight chromosomes of this genus, hence the potential just in this genus is immense. If one adds to this basic genomic work the effects of using the genetic probe *LaeA*, then another area of control of such clusters in *Aspergillus* species opens up [62]. Further examples from the same research group also demonstrates the power of these techniques in identifying bioactive metabolites from this genus, thus adding to the basic structures that medicinal chemists can work with [63].

The use of epigenetic manipulation was shown to produce novel previously undescribed structures covering a variety of disease states. In particular, work by the Cichewicz group at the University of Oklahoma on the remodeling of fungal metabolism used external histone deacetylase and DNA methyltransferase treatments to induce cryptic clusters to produce [64,65]. In addition to exogenous applications of epigenetic agents, one can also use the simpler technique of co-culturing different microbes (crossing kingdoms or simply genera) and eliciting what are probably protective agents produced under the competitive stresses of one microbe upon another [66].

Microbes from other environments also received their fair share of attention as potential sources of novel agents using genomic methods. The use of metagenomic gene-mining has been well covered since the original report of the Handelsman group in 1998 demonstrating this new frontier [67]. Since then, many groups worldwide have utilized the technique, with a recent publication from Nováková and Farkašovský commenting on the mining of natural sources for new secondary metabolites [68].

2. *Extremophiles*

Extremophilic microbes (extremophiles) abound in extreme habitats. These include acidophiles (acidic sulfurous hot springs), alkalophiles (alkaline lakes), halophiles (salt lakes), piezo (baro)- and (hyper)thermophiles (deep-sea vents) [69–73], and psychrophiles (arctic and antarctic waters, alpine lakes) [74].

An unusual group of acidophiles that thrive in acidic, metal-rich waters has been found in an abandoned mine-waste disposal site, a polluted environment which is generally toxic to most prokaryotic and eukaryotic organisms. In this work the novel sesquiterpenoid and polyketide-terpenoid metabolites berkeleydione and berkeleytrione (Figure 5.4), showing activity against metalloproteinase-3 and caspase-1, activities relevant to cancer, Huntington's disease, and other diseases, were isolated from *Penicillium* species found in the surface waters of Berkeley Pit Lake in Montana [75–77]. Further information on the biosynthetic potential of this source was reported by the original authors in 2013 when they reported on other unusual molecules from this source, some with biological activity [78]. Quite recently, Arto's group gave an excellent account of the scalable synthesis of (-)-berkelic acid on a gram scale up to the final step, and then the prototypical production on a milligram scale at 55 percent yield due to not wishing to lose material in side reactions on the final coupling step [79].

Further examples of the unusual molecules produced by extremophilic microbes can be seen from the work reported by the Lin group on spiromastixones A-O from a deep sea fungus [80], the discussion on bioactive natural products from cold seas by Liu et al [81], the review by Mosey and Floreancig on the pederin/mycalamide family [82],—which should be read in conjunction with the work reported later by the Piel group—and the very recent and excellent review covering reports from 2009 to 2013 by Skropeta and Wei on recent advances in deep-sea natural products [83].

3. *Plant Metabolites and Endophytes/Rhizosphere Microbes*

Within the past two decades, there has been a dramatic increase in the number of reports of endophytes, which are best defined as microbes that reside inside the living tissues of plants without having deleterious effects and—from the perspective of this chapter—are involved in producing valuable therapeutic plant-derived secondary metabolites. We will also comment upon microbes that are not endophytic as in the definition above, but are part of the plant's rhizosphere (the area encompassed by the roots), since in one very interesting case the precursor for a well-known "plant product" is produced by such a microbe.

From extrapolation studies, it is reasonably certain that each of the approximately 350,000 plant species on earth may serve as a host to one or more endophytes [84,85]. This is a trend that was initially identified in the marine area when natural products that were thought to be produced by the invertebrate from which they were isolated were later found to be produced by symbiotic/commensal microbes. A significant number of reports from India and the People's Republic of China, in addition to the initial report from Stierle and Strobel on the production of Taxol® [86], demonstrated that endophytic fungi produce at least the following four major classes of "plant-derived" natural products: taxanes [87], podophyllotoxins [88,89], camptothecins [90,91], and the vinca alkaloids [92–94]. Summaries of the then known endophytic fungi that produce nominally plant secondary metabolites were published in 2012 [95,96].

Recently, the very well-known "plant metabolite" swainsonine (Figure 5.4), which is the causative agent of "locoism," has now been shown to be produced by endogenous fungal endophytes in a number of plant genera [97]. A number of species of the fungal genus *Undfilium* have been identified in these plants [97]. There are reports of other free-living fungi producing the same material, and they are predators of other organisms, including red clover [98]. The endophyte appears to be vertically transmitted in the seed or seed coat, and thus is

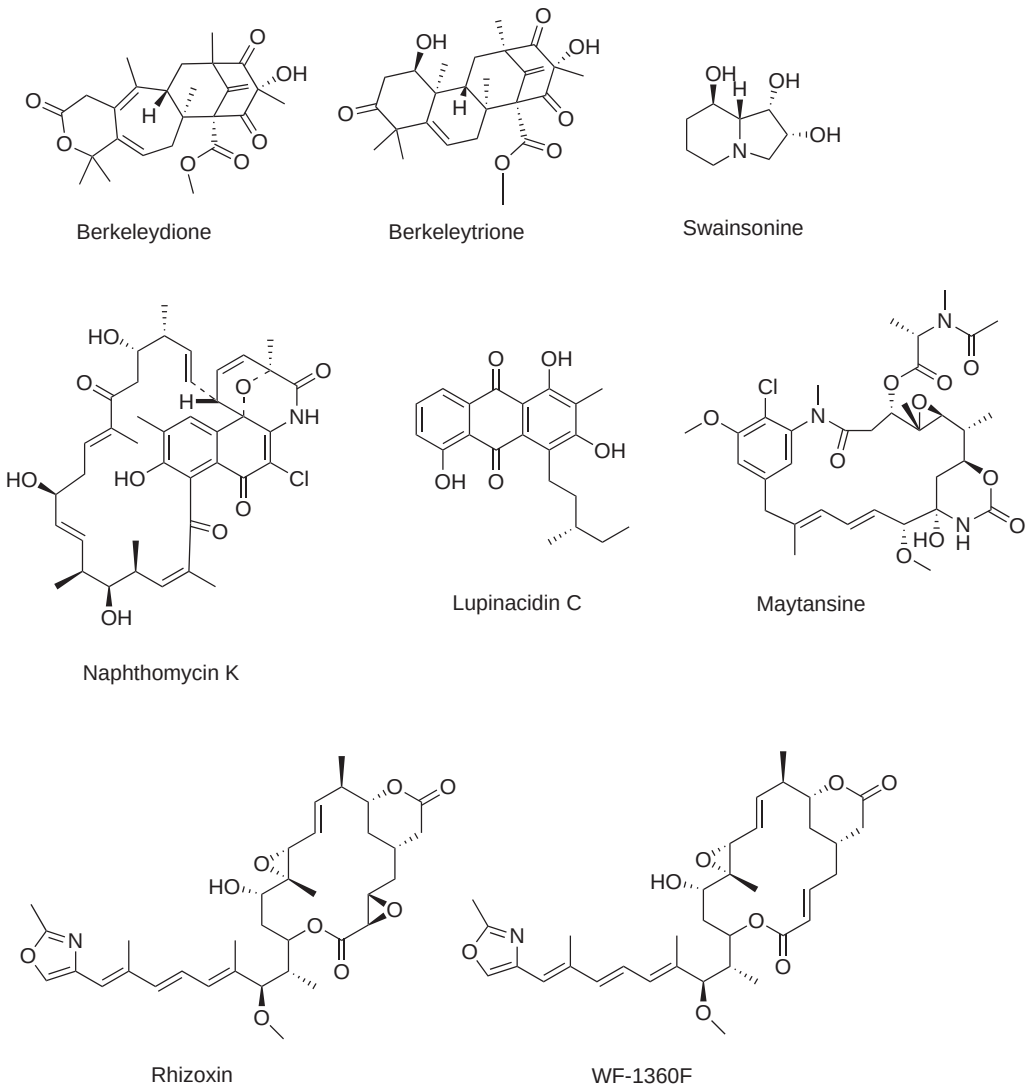


FIGURE 5.4 Compounds from extremophiles and endophytes.

reminiscent of the methods of transmission of the putative producer of the marine antitumor agent, bryostatin, as shown by Haygood and her collaborators [99].

In addition to endophytic fungi, there have also been limited reports of endophytic actinobacteria reported to enhance and/or produce new secondary metabolites [85,100,101]. Thus, Lu and Shen reported in 2007 the isolation and identification of a new cytotoxic ansamycin, naphthomycin K (Figure 5.4), produced by an endophytic *Streptomyces sp. CS* isolated from the medicinal plant *Maytenus hookeri* [102]. Then in 2011, Igarashi and coworkers identified the new anthraquinone, lupinacidin C, isolated from the endophytic actinomycete, *Micromonospora lupini*, that coexisted in the root nodules of the legume *Lupinus angustifolius* [103]. This compound exhibited anti-invasive activity against murine colon cancer cells.

A. THE MAYTANSINE SAGA

An excellent example of organisms working together to produce a bioactive substance is the story of the antitumor agent maytansine (Figure 5.4). For many years, maytansine and congeners were thought to be exclusively plant-derived secondary metabolites isolated from the *Maytenus ovatus* plant collected in Ethiopia and first reported by Kupchan et al. in 1972 [104]. This metabolite was also isolated from *M. buchananii* and *Putterlickia verrucosa* plants. Because the structure, a 19-membered halogenated ansamycin, was similar to those reported from bacteria, was an unusual structure for a plant secondary metabolite, and was present in some but not all

individual *P. verrucosa* plants, a search commenced for microorganisms (fungal or bacterial endophytes) that could produce its core structure. This was successful, and in 1977 investigators at Takeda Industries in Japan reported the discovery of the family of ansamitocins [105]. These structures were maytansine derivatives with either an ester or a hydroxyl moiety at C3, from two subspecies of *Nocardia* (subsequently renamed as *Actinosynemma pretiosum*) isolated from the *Carex* species of grassy plants. From these and other studies, researchers considered that the P-3 precursor was produced by an endophyte or symbiont in the rhizosphere that is subsequently taken up by the plant and converted into maytansine. In 2013, Wings et al reported that axenic cultures of a producing plant species, *P. verrucosa*, did not produce maytansine, and reported as a result of their studies using modern genomic techniques that the *A. pretiosum* ssp. *auranticum* bacterium present in the rhizosphere of the plant is involved in maytansine biosynthesis [106].

Only a handful of mostly higher plants and their corresponding endophytes have been investigated, leaving the vast majority of plants to be studied. Not only can endophyte–plant interactions induce the production of new compounds, but endophyte–endophyte interactions within plants also have the potential to produce new secondary metabolites as plants are unlikely to be colonized by just a single microbe. Biosynthetic genes can be up- or down-regulated in endophytes as a result of interacting with other microorganisms within their environment [107,108]. Signaling molecules analogous to bacterial quorum sensors and other elicitors are thought to be involved in activating cryptic biosynthetic gene clusters.

B. RHIZOXIN

An interesting example of endo-symbiosis between a fungus and a bacterium has been discovered in the case of rice seedling blight where the toxic metabolite, rhizoxin (Figure 5.4), originally isolated from the contaminating *Rhizopus* fungus, was initially thought to be produced by a symbiotic *Burkholderia* bacterial species [109,110]. This unexpected finding revealed a complex symbiotic-pathogenic relationship, extending the fungal–plant interaction to a third, key bacterial player, thereby offering potentially new avenues for pest control. This observation was consistent with the discovery of four *Rhizopus* species that would produce rhizoxin on fermentation and two others in the same genus that did not. Later work using combinations of *Rhizopus* strains that either produced rhizoxin or stopped at the monoepoxy compound WF-1360F (Figure 5.4), which was originally thought to be a by-product of biosynthesis, demonstrated that, by mixing and matching fungal hosts (strains that either produced rhizoxin or only WF-1360F) and the endosymbiotic bacterium from each, the *Burkholderia* bacterium produced the monoepoxy precursor, and then the fungus produced rhizoxin via a second fungal-specific interaction. These results led to a revison of the bipartite system to a novel tripartite system, and aptly demonstrates what can be done at this moment in time to "interrogate" biochemical processes [111].

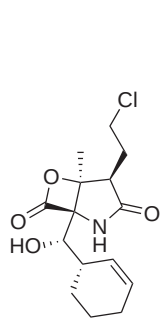
4. Further Involvement of Rhizosphere and Related Microbes

Two recent papers in Nature described new generation sequencing in the identification of bacterial microbiota colonizing the root rhizosphere, soil, and endophytic compartments (within the roots) of *Arabidopsis thaliana* [112,113]. Both studies identified similar phyla of bacteria inhabiting the endophytic compartments of *A. thaliana* and demonstrated that they are significantly dissimilar compared to those found in plant-free soil and the root rhizosphere. Notably, the microbiota of the endophytic compartment is influenced by soil type, and some variation was observed among plants of different genotypes and developmental stages. All of these observations suggest that there is a large possibility of finding more unusual endophytic microbes and symbiotic interactions that have the capability of producing new secondary metabolites. As more reports on the metagenomic sequencing of plant microbiomes are published, these will facilitate the dissection of endophyte–endophyte and endophyte–plant interactions.

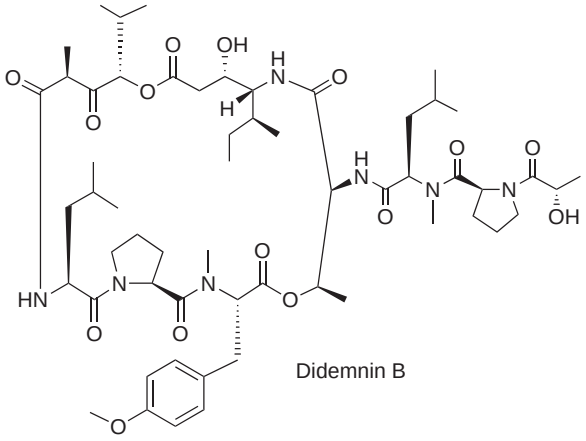
5. Culturable Marine Microbes (Free-living and Endosymbionts)

A. SALINOSPORAMIDE A; FROM A FREE-LIVING MARINE MICROBE

Although this compound could be listed under marine antitumor agents, we think that the story of salinosporamide is best covered in this section due to the interactions that it uncovered. The compound was first reported from the marine-sourced bacterium *S. tropica* by the Fenical and Jensen groups at Scripps Institute of Oceanography (SIO) in 2003 [114] and identified as a novel proteasome inhibitor (Figure 5.5). In 2007, Udwary et al. reported the complete genome sequence of *S. tropica*, identifying seventeen potential biosynthetic gene clusters, including those comprising the salinosporamide locus [115]. This led to extensive studies by the groups of



Salinosporamide



Didemnin B

FIGURE 5.5 Compounds from free-living marine microbes.

Moore, Fenical, and Jensen at the Scripps Institute of Oceanography, all of whom collaborated in methods of eliciting the products of some of these clusters, with extension into other microbial genera. This work through the middle of 2010 was covered in an excellent article by Lane and Moore, which gave a plethora of examples where these techniques had been used in unlocking the genomic information in marine-derived microbes [116].

Because of these advances in identifying biosynthetic gene clusters, the initial promise of combinatorial biosynthesis is now being realized, in that one may either utilize unnatural starting substrates in order to alter the final product, or—by mixing and matching specific gene clusters—one may generate novel structures. An excellent example of the latter technique is the biosynthetic engineering in order to produce fluorosalinisporamide from the salinosporamide producer *S. tropica*. By replacing the salL chlorinase gene in *S. tropica* with the fluorinase gene flA from *Streptomyces cattleya* and growing the $salL^{-}$ flA^{+} mutant strain in the presence of inorganic fluoride [117], the fluorine analogue of the natural product was obtained. Although natural products containing fluorine are rare, a fair percentage of all marketed drugs contain this atom. Thus, a genetic-based system that would incorporate fluorine into a molecule by biosynthetic techniques may well be of interest in the future.

B. DIDEMNIN B; FROM A FREE-LIVING MARINE MICROBE

Another example of the potential of marine microbes is the current story of didemnin B (Figure 5.5); again this could be listed under marine antitumor agents, but due to recent publications, it is probably best discussed here. In 2011, a Japanese group reported the identification of didemnin B from the fermentation of a marine-derived *Tristella mobilis* strain [118]. The genus *Tristella* was first identified from waste water in Thailand and reported by Shi et al in 2002 [119], with a second species, *T. bauzanensis*, identified in 2011 by Zhang et al. from soil in a heavy metal contaminated site in Italy [120]. Thus, until the report by Tsukimoto et al., the only record of this genus came from terrestrial sources, albeit aqueous.

In 2012, a Chinese group from Hong Kong working closely with the SIO and researchers in the Department of Chemistry at the University of California, San Diego, published the complete biosynthetic pathway to didemnin B from a free-living *Tistrella mobilis* strain isolated in the Red Sea. In addition to the pathway, they also demonstrated in real-time the production of didemnin B as the microbe was growing by use of scanning laser mass-spectrometry. Thus, there was absolutely no question that this free-living microbe produced didemnin B [121]. Work is continuing on other didemnins and related compounds.

6. "As Yet" Uncultured Marine Microbial Symbionts

Before discussing the papers mentioned in the next paragraph, an interesting discussion of the metabolic requirements for marine microbes under extreme energy limitations was published in 2013, pointing out that work with marine microbes has to take into account the vastly different flux of metabolites required for even basic growth patterns, compared with microbes from "normal" environments [122]. It should be also pointed out that uncultured microbes may well have entirely different requirements if they are not "free-living" but are part of a symbiotic arrangement.

There have been some seminal papers in the fields of biosynthesis of secondary metabolites in marine-derived organisms in the last few years describing techniques to "unlock" the biosynthetic processes in what are called

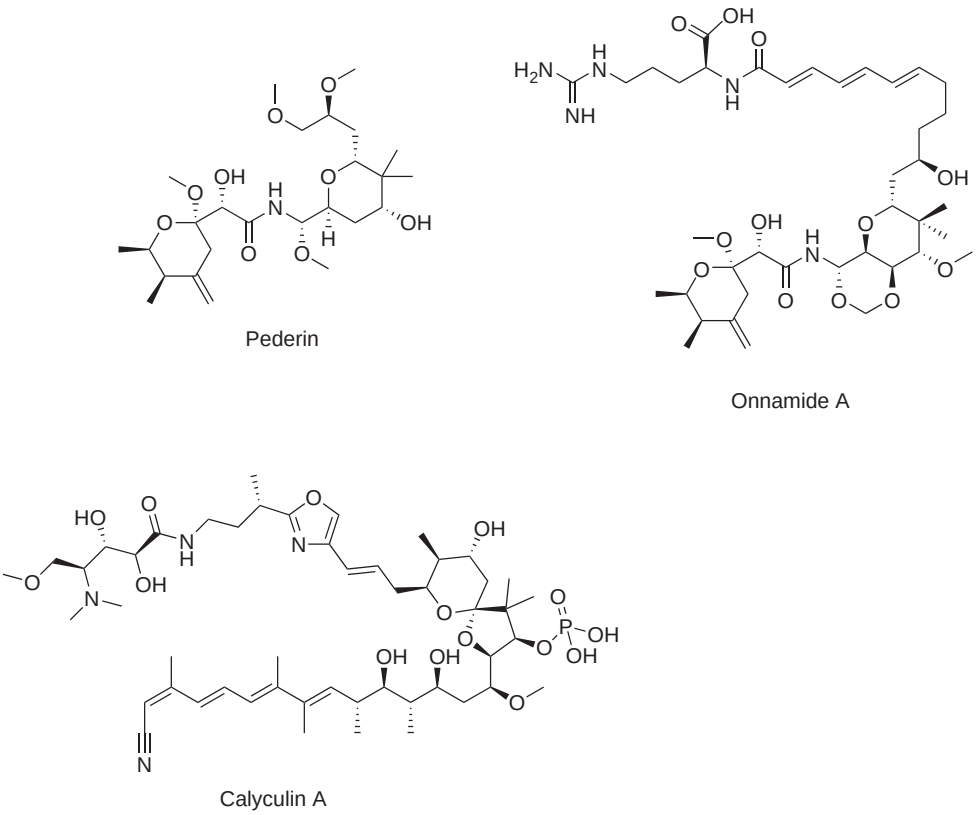


FIGURE 5.6 Examples of microbial natural products from uncultured marine microbes.

"as yet" uncultured microbial symbionts. Definitive evidence now shows that many bioactive compounds isolated from various macro-organisms are actually metabolites synthesized by symbiotic bacteria [123]. These include the pederin class of antitumor compounds (Figure 5.6) isolated from beetles of genera *Paederus* and *Paederidus*, which have also been isolated from several marine sponges, an example being onnamide A (Figure 5.6) [123–125]. Recently, Mosey and Floreancig published an excellent review on the isolation, chemical synthesis, and medicinal chemistry of this class of compounds [82].

We should also point out that a significant number of antitumor agents isolated from marine invertebrates closely resemble products from terrestrial microbes. Although the actual microbes have not been cultivated as yet, their genomes have been interrogated to identify the putative producing clusters [52,126]. These include the ecteinascidins and bryostatins as initial examples, but the 2011 review from the Crews' group at the University of California, Santa Cruz, gives many more examples of this type of relationship [127]. A more recent paper from the Crews' group covers compounds probably produced by Gram-negative organisms [128]. It, too, should be consulted for more examples.

However, pride of place at this moment in time should go to the work reported by the Piel group in 2014 in an article in Nature [129]. In this seminal paper, they showed conclusively that the vast majority of the materials isolated from a Pacific sponge, *Theonella swinhoei* Y (the yellow variant), are produced by the as yet uncultured "*Entotheonella* spp," which was isolated by single–cell techniques from the sponge homogenate and then totally sequenced by next generation genomic methods. Analyses of the genomic results demonstrated that, contrary to accepted dogma, the peptidic portions of the compounds were ribosomally produced and then "tailored" by a series of enzymes from the same organism in an assembly-line fashion, to give over forty identified compounds, including the pederin series mentioned earlier. Then very recently, the same nominal microbe was shown to be responsible for the production of another class of compounds from the sponge *Discodermia calyx*, the protein phosphatases 1 and 2A inhibitor calyculin (Figure 5.6) [130].

Finally, the paper by Rocha-Martin et al. that was recently published gives an excellent overview of the manifold ways in which "systems biology" can be used to aid in the discovery of bioactive compounds from marine microbes [131].

7. Combinatorial Biosynthesis

Great advances were made in the late 1990s and early 2000s in understanding the role of multifunctional polyketide synthase enzymes (PKSs) in bacterial aromatic polyketide biosynthesis, and many such enzymes have been identified, together with their encoding genes. Through the rapidly increasing analysis of microbial genomes, a multitude of gene clusters encoding for polyketides, NRPs, and hybrid polyketide-NRP metabolites have been identified, thereby providing the tools for engineering the biosynthesis of novel "nonnatural" natural products through gene shuffling, domain deletions, and mutations. Examples of novel analogues of anthracyclines, ansamitocins, epothilones, enediynes, and aminocoumarins produced by combinatorial biosynthesis of the relevant biosynthetic pathways were reviewed by Shen et al. in 2005 [132] and then updated in 2012 by van Lanen and Shen [133].

Although the commercialization of processes appeared to wane following the takeover of Kosan, Inc., by Bristol Myers Squibb in the mid-2000s, recent papers have shown that the process is still continuing, in addition to the reviews mentioned in the previous paragraph. Thus, Bachmann et al. suggest that accelerated genome mining is materially aiding in discovering novel natural products [134], and the Goss group have just published a paper that combines chemical synthesis with microbial biosynthetic processes [135]. With the advent of information from genomic studies (see Sections III.B.1 and V.D), it is probable that this method will be applied more generally in the next few years.

C. Extraction

In the case of microbes and marine organisms, extraction is normally carried out on the whole organism (though now some groups are isolating the commensal/associated microbes from marine invertebrates before a formal extraction). With plants however, which may be large and have well differentiated parts, it is common to take multiple samples from one organism and to extract them separately. The methods used by the NCI have been reviewed [136] and are summarized at http://npsg.ncifcrf.gov/. The procedures used for extraction vary with the nature of the sample, and in some cases are dependent upon the nature of the ultimate assay. Thus, a number of screens are sensitive to the tannins and complex carbohydrates that are extractable from a variety of organisms, and systems have been developed that permit easy removal of such "nuisance" compounds before assay [137,138]. In the simplest cases, however, extraction with a lower alcohol (methanol to isopropanol) will bring out compounds of interest, though in most cases a sequential extraction system is utilized, with compounds being extracted with solvents of ever-increasing polarity.

D. Screening Methods

As mentioned earlier, the advent of new and robust high-throughput screens has had—and continues to have—a major impact on natural products research in the pharmaceutical industry. One screen that has been described in detail is the National Cancer Institute's sixty cell-line cytotoxicity screen for anticancer agents [139]. Although this is not a true receptor-based screen, it has now been developed into a system whereby a large number of molecular-targets within the cell lines may be identified by informatics techniques, and refinements are continuing [140]. An assay based on differential susceptibility to genetically modified yeast strains was described in 1986 [141], and led to many screens based upon genetically modified yeasts, but in a number of screens the low permeability of the unmodified yeast cell wall to chemical compounds was overlooked. Thus, data from such screens—particularly those designed with gene deletions—should be carefully scrutinized, since a large number were based upon hosts without a "modified" cell wall.

At the time of writing the previous edition in 2007, no reports had been published about the use of crude extracts tested in 1536 well plates, let alone against protein−protein interactions in 1536 well plates using fluorescent techniques. In a recent paper, a multidisciplinary group from academia and the NCI demonstrated that such a screening program was successful in identifying natural product extracts and subsequently purified natural products from the initial hits that interfered with some of the apoptotic Bcl-2 proteins [142].

These assays were performed in volumes of less than 10 μl and required acoustic dispensing of the crude organic extracts and pin-tool dispensing of the concomitant aqueous extracts, both under conditions that did not denature the proteins used and would be reproducible. In addition, secondary assays in a dose response format, including counter screens, both using *in cellulo* systems in 384 well plates and volumes of about 30 μl, were also necessary in order to confirm that results observed in the *in vitro* conditions could occur within the cell (transfer

across membranes, etc.). Thus, by use of sophisticated robotic systems using *in vitro* and *in cellulo* assays, Hassig et al. were able to identify novel agents from an initial screen of close to 150,000 crude extracts from plants, microbes, and marine invertebrates [142].

Very recently two review articles on screening asked the question as to whether phenotypic screening is still relevant at the moment, since target-related screening had become the *sine qua non* of high-throughput screening in the last fifteen years. The review articles came out back-to-back in the *Nature Reviews* series, one covering cancer screening [143], and the other an analysis of the screening paradigms used to discover the 113 "first in class" drugs approved between 1999 and 2013 by the US Food and Drug Administration (FDA) [144]. This latter review was an extension of a previous analysis by Swinney and Anthony in 2011 in the same journal [145]. Basically, both of the analyses showed that in cancer drugs, target-directed drug discovery was predominant, though this is not surprising as a majority of drugs approved in the time frames used were kinase inhibitors. However, a significant number of cancer drugs were from phenotypic or combined screens. Of the thirty-three "systems based drugs" (meaning other than target-based screening) in the Eder et al. review [144], 12 or 35 percent were based on natural products, whereas only two of the seventy-eight target-based drugs were from natural product structures. We should point out that in the analyses by two of the authors of this chapter, these figures (2/78) would be higher due to differences in definition and would be part of the natural product inspired series in our analyses in the 2012 paper [8].

However, what both of these reviews demonstrate is that phenotypic screening is still a valid process, particularly when linked to other methods of demonstrating the particular interactions that are occurring in the whole cells used, as discussed below.

We mentioned above the use of yeast strains that had been genetically modified so that they could demonstrate that a specific gene or gene set were affected by either an extract or a pure compound, thus giving an idea as to the particular locus with which a compound/extract was interacting. An assumption in all of these assays based upon a specific gene/gene construct was that the background of the yeast cell was not involved. Work published by the Nislow and Giaever team very recently in *Science* shows that this is not always the case, and that there may at times be an effect from the background genetic make-up. They demonstrated by screening ~3000 compounds chosen from over 50,000 drug-like compounds using haploinsufficiency profiling (a strain that lacks one of the copies of the ~1100 essential genes) and a homozygous profiling (where ~4800 homozygous deletion strains identify the nonessential genes that "buffer" the targeted pathways) that the HIPHOP profile from this chemogenomics platform demonstrated the genome-wide response to a given compound [146].

Another method that has come of age since the last edition is the use of interfering RNA or RNAi based operations where the interfering RNA strands cause knockout of specific gene sequences. These small pieces of ribonucleic acid enable screeners to investigate responses in live cells and also, in some cases, *in vivo*, thus aiding in the identification of potential targets of drugs from both target and phenotypic results [147].

Thus, screening of natural products today can involve a variety of approaches that can utilize materials from crude extracts through to purified compounds, whole cells that may or may not contain modified targets, isolated targets, modified yeast cells, or—as is often the case—selected combinations of the methods in order to both discover an activity of interest and then to identify its locus (often loci) of activity.

E. Isolation of Active Compounds

The isolation of the bioactive constituent(s) from a given biomass can be a challenging task, particularly if the active constituent of interest is present in very low amounts. The actual procedure will depend to a large extent on the nature of the sample extract. A marine sample, for example, may well require a somewhat different extraction and purification process than that derived from a plant sample. Nevertheless, the essential feature in all of these methods is the use of an appropriate and reproducible bioassay to guide the isolation of the active compound. It is also extremely important that compounds that are known to inhibit a particular assay, or those that are nuisance compounds, be dereplicated (identified and eliminated) as early in the process as possible. Procedures for doing this have been discussed [148–155], and various new approaches to isolation and structure elucidation have been reviewed [156,157].

F. Structure Elucidation

Structure elucidation of the bioactive constituent depends almost exclusively on the application of modern instrumental methods, particularly high-field NMR with cryoprobe systems where the level of compound is in

the sub-milligram range, and MS^n. These powerful techniques, coupled in some cases with selective chemical manipulations, are usually adequate to solve the structures of most secondary metabolites up to 2 kD molecular weight. X-ray crystallography is also a valuable tool if crystallization of the material can be induced, and in some cases it is the only method to assign absolute configurations unambiguously. Nowadays, the determination of the amino acid sequences of polypeptides or peptide-containing natural products up to 10–12 kD is a relatively straightforward task, requiring less than five milligrams of a polypeptide. In addition, MS^n techniques have developed to the stage where polypeptides containing unusual amino acids not recognized by conventional sequence techniques can be sequenced entirely by MS. Recent examples of the power of coupled MS techniques can be seen by inspection of the 2013 article by the Oberlies group [158], and the very recent reviews by the Dorrenstein group on current and future MS techniques [159,160].

G. Further Biological Assessment

Once the bioactive component has been obtained in pure form and shown to be either novel in structure or to exhibit a previously unknown function (if it is a compound that is in the literature), it must be assessed in a series of biological assays to determine its efficacy, potency, toxicity, and pharmacokinetics. These will help to position the new compound's spectrum of activity within the portfolio of compounds that a group may be judging for their utility as either drug candidates or leads thereto. If an idea can be gained as to its putative mechanism of action (MOA) (assuming that the screening techniques used in its discovery were not MOA-driven) at this stage, then it too can help as a discriminator in the prioritizing process.

H. Procurement of Large-Scale Supplies

Once a compound successfully completes evaluation in the initial biological assays, then larger amounts of material will be required for the studies necessary if activity and utility are maintained as the compound proceeds along the path from "hit" to a "drug lead" and then to a "clinical candidate." The supplies could be made available by cultivation of the plant or marine starting material, or by large-scale fermentation in the case of a microbial product. Chemical synthesis or partial synthesis may also be possible if the structure of the active compound is amenable to large-scale synthesis. The example of paclitaxel is instructive here. After initial large-scale production by direct extraction from *Taxus brevifolia* bark, it is now generally produced by a semisynthetic procedure starting from the more readily available precursor 10-desacetylbaccatin III [161].

Another method of obtaining adequate supplies of a natural plant product is by utilizing plant tissue culture methods. Although there are a few examples of the commercial production of secondary metabolites by plant cell culture (shikonin being perhaps the best known [162]), the application of this technique to commercial production of pharmaceuticals had not found general acceptance, primarily for economic reasons [163]. However, the development of viable methods for the large-scale production of paclitaxel (Taxol®) has illustrated that this technology can now be successfully applied to the production of a major drug for commercial purposes [164]. Plants grown hydroponically or aeroponically under controlled conditions have been shown to produce higher biomass yields than those grown in traditional soil culture [165,166]. In addition, growth of *Withania somnifera* under aeroponic conditions has been reported to produce two major withanolides, withaferin A and 2,3-dihydrowithaferin A-3-O-sulfate, with a combined un-optimized yield of 0.5–0.9 percent, exceeding the yield (0.2–0.3 percent) of withaferin A reported for soil-grown plants [167]. These results indicate that aeroponic and hydroponic techniques may be of value in maximizing the production of plant-derived drugs for advanced preclinical and clinical development.

The discovery that several major anticancer drugs originally isolated from plants may be produced by associated endophytic fungi (Section III.B.3) opens up further avenues for exploring the large-scale production of plant-derived pharmaceuticals. Likewise, the probable role of microbial symbionts in the production of bioactive agents from marine macro-organisms (Section III.B.5) offers similar opportunities for scaling up the production of marine-derived pharmaceuticals. It is interesting to note that the anticancer agent Yondelis® (ecteinascidin 743), originally isolated from the tunicate *Ecteinascidia turbinata*, is now produced on a large scale by semisynthesis from the antibiotic cyanosaframycin B, produced through fermentation of the bacterium *Pseudomonas fluorescens* (Section IV.F.3).

In a relatively few cases, total synthesis has provided a viable route to large-scale production of important bioactive agents. A good example was the marine-derived anticancer agent discodermolide [168], which entered

Phase I clinical trials but did not progress to later phases due to toxicity. In contrast, the modification of the halichondrin B skeleton to produce E7389 (eribulin; Section IV.H.1) by total synthesis is an excellent example of modification of a very complex molecule to a slightly less complex agent that is now approved for the second-line treatment of resistant breast cancer [169].

I. Determination of Structure-Activity Relationships/Further Structural Modifications

The initial hit isolated from the biomass, irrespective of source, is not necessarily the lead required for further development into a drug. It may be too insoluble, not be potent enough, or be broadly rather than specifically active. Once the structure has been determined, synthetic chemistry—involving both conventional and combinatorial methods—may be used in order to generate derivatives/analogues that have the more desirable characteristics of a potential drug. The role of these methods in the development of anticancer agents has been reviewed [170], and several examples of these types of processes are shown in Section IV. In all cases, the modified compounds are screened in suitable direct- and counterscreens in order to identify the necessary substituents for a given activity or to identify a partial structure as the "pharmacophore" for a given biological activity that can then be further developed. Excellent discussions of the overall process are given in the recent reviews on kinases by Liu et al. [171] and on anti-TB drugs by Coxon et al. [172]. These should be consulted for specific examples of the techniques involved.

The method using yeast-based systems (Section III.D) as described by Lee et al. [146] certainly has the potential to suggest where particular agents may interact with cellular metabolism and thus point to areas to be further investigated for specific interactions with cellular processes. It also has the advantage of eliminating potential mechanisms rather rapidly.

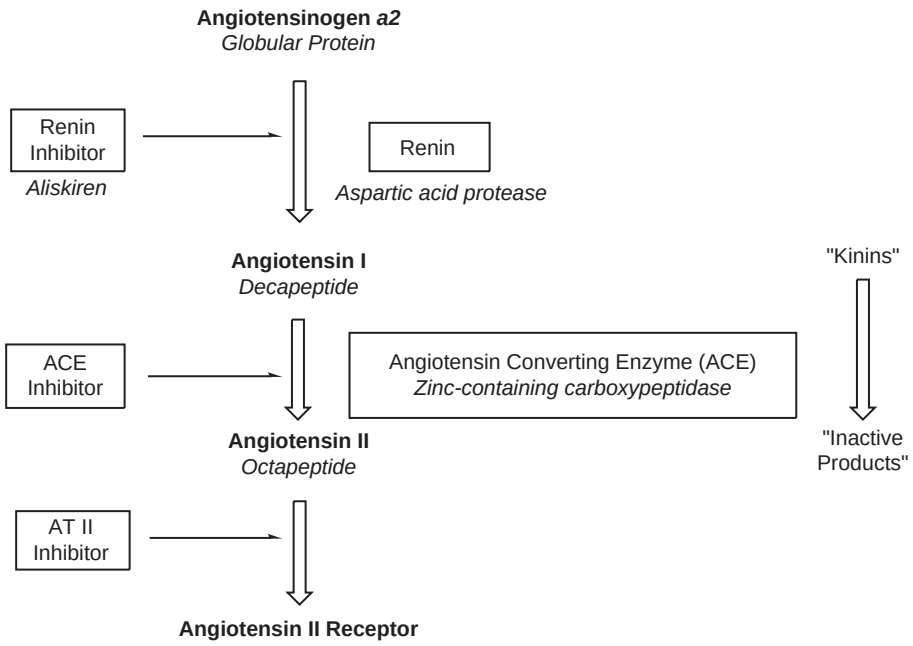
IV. EXAMPLES OF NATURAL PRODUCTS OR ANALOGUES AS DRUGS

A. Antihypertensives

1. *Direct Renin Inhibitors*

In the previous editions, we covered the angiotensin converting enzyme (ACE) inhibitors and the compounds loosely known as the "sartans," synthetic compounds modeled upon the angiotensin II substrate and its interaction with the angiotensin I receptor. What has been known since the initial discovery of the angiotensin system (Scheme 5.1) is that the first enzyme in the cascade is renin, an aspartic proteinase, that cleaves four residues off

SCHEME 5.1



of angiotensinogen, and then ACE converts the decapeptide angiotensin I to the octapeptide angiotensin II by cleavage between the eighth and ninth residue at the carboxy-terminus. It was knowledge that this enzyme was a zinc-containing carboxy-peptidase that enabled the Squibb group back in the 1970s to synthesize the pseudo-dipeptide Captopril® as the first ACE inhibitor to be approved by the FDA.

However, the "prime target" in the system is inhibition of renin, since that is the enzyme that starts the cascade and—unlike ACE—it does not hydrolyze the "kinin" peptides (e.g., bradykinin). Renin was known to be an aspartic proteinase, and it could be inhibited by the bacterial peptide pepstatin (Figure 5.7) [173]. Another now well-known class of aspartic proteinases are the HIV proteases, and work by Sigal's group at Merck demonstrated that pepstatin would also inhibit that class of proteases with a K_i of 1.1 μM, but the inhibition was not totally competitive [174].

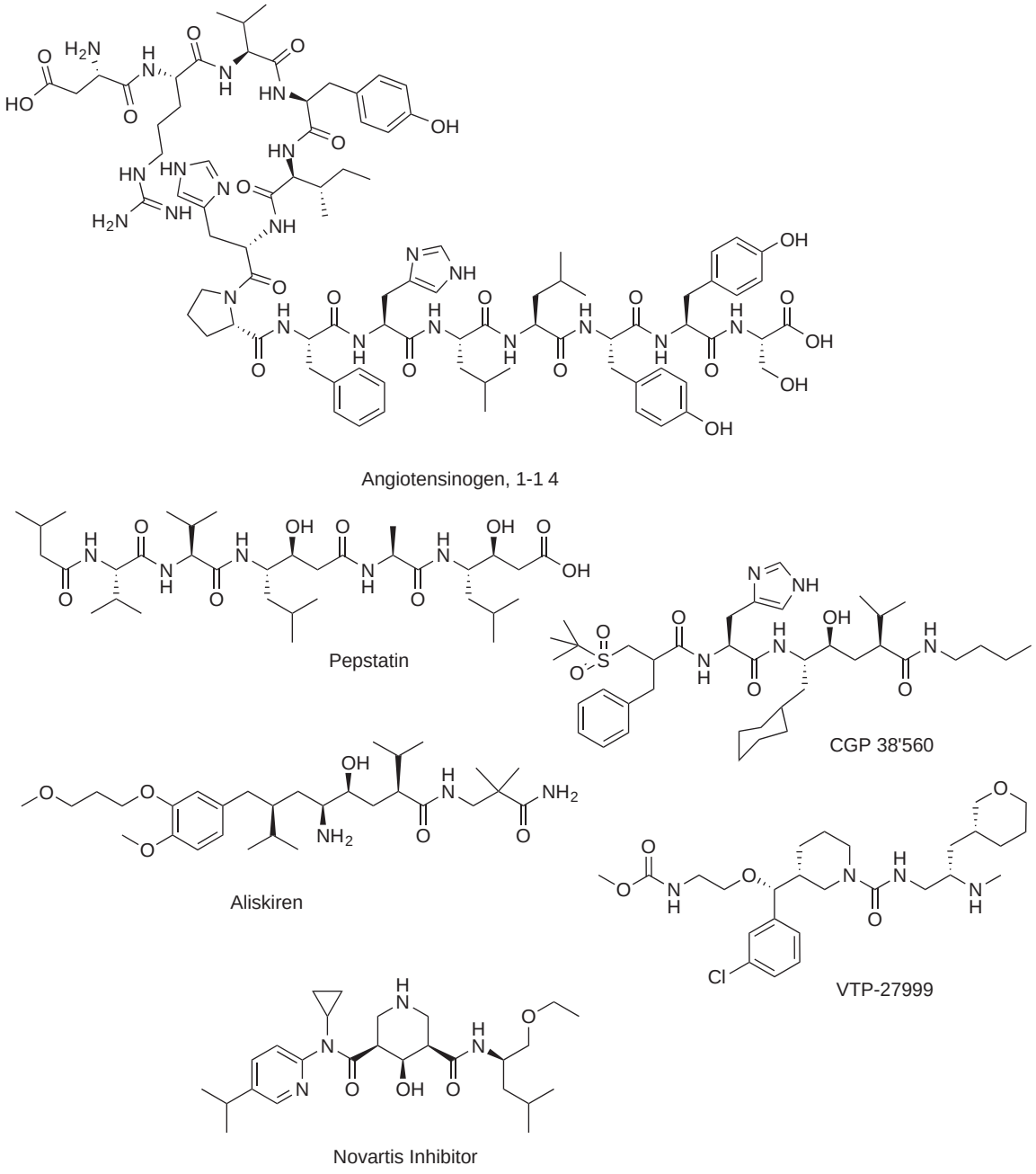


FIGURE 5.7 Natural product-based renin inhibitors.

Pepstatin (Figure 5.7) contains the unusual amino acid statine, which contains as a dipeptide mimic a hydroxyethylene isostere. It was the basis of a long-term project at Merck to synthesize renin inhibitors and later HIV-protease inhibitors based on this substituent mimicking the transition state of the aspartic proteinase/substrate pair [175,176]. Although all of their peptide structures did not provide a renin inhibitor that was approved as a drug, their work demonstrated the potential for such substitutions to be effective drug leads, albeit from Ciba-Geigy (now Novartis), en route to an orally active renin inhibitor. The first of what were known as type-I inhibitors [177] contained the dipeptide isostere (2*S*, 4S, 5S)-5-amino-4-hydroxy-2-isopropyl-6-cyclohexyl-hexanoic acid at the P1-P1′ position, and also mimicked angiotensinogen (Figure 5.7) from residue P3 to P1′ (nomenclature as Schetchter and Berger [178]).

The story of the search for orally-active renin inhibitors, which though formally nonpeptidic still contained the hydroxyethylene transition state dipeptide isostere identified from the bacterial inhibitor, is given in detail by the Novartis scientists in two papers. The search involved significant computerized structure activity relationships using the crystal structure of human renin to optimize the chemistry before finally leading to the drug candidate SPP-100 that became the drug aliskiren (Figure 5.7), which gained FDA approval in March 2007 and EMA approval in August 2007. The first paper, published in 2000, gives the chemical basis for the initial discoveries of pseudo-peptidic agents, and the use of structure-based drug design with modifications around the initial type-I inhibitor (CGP 38′560; Figure 5.7) [179]. The second paper, published in 2003, gives the next chapter in the story—the work leading up to aliskiren [180]. In 2006, a review of the literature on renin inhibitors was published in the Lancet, and the authors made the classic mistake of denoting the first renin inhibitor, pepstatin, to be a "synthetic" agent [181], even though the abstract of the reference that they quoted stated that the material was obtained from a culture filtrate of an actinomycete and demonstrated short action *in vivo* activity in a dog [173]. Thus, one has to go back to the original literature report in a fair number of cases to find out the actual source.

A thorough analysis of the various molecules and routes that led to aliskiren was published by Novartis scientists in 2010 [182] and should be consulted for the full story. However, other pharmaceutical companies and academic groups are also very active in this field, now that this first-in-class drug has shown efficacy and the problems that can be seen if different combinations of antihypertensive drugs are used. Thus, a fixed dose combination of aliskiren and valsartan, though approved in the USA in 2009, was later withdrawn in 2012, and the corresponding application in the EU was withdrawn before approval.

The academic side of drug discovery has published some recent reviews on direct renin inhibitors. The first is a study by Politi et al. in 2011 on the conformation of aliskiren in solution and when bound to its receptor [183]. The second, by members of the same group, was a study showing the calculated binding of aliskiren to a model of the HIV protease, and that the SGLT-2 inhibitor canagliflozin, which was approved by the FDA and the EMA in 2013 as a treatment for type II diabetes, and the HIV protease inhibitor darunavir approved in 2006 may have cross-activities in renin inhibition as well as their regular approved pharmacological targets, thus repurposing these compounds [184]. In addition to these papers, Nam and Ko reported in 2012 a total synthesis of aliskiren starting from a pseudo-symmetrical bis-lactone, in contrast to the other published syntheses that utilized convergent syntheses from two different chiral fragments, thus demonstrating a novel approach to the drug [185].

In 2011, groups from the small pharmaceutical company Vitae Pharmaceuticals and the large company GlaxoSmithKline published on the discovery of VTP-27999 (Figure 5.7), where the protonated secondary amine in this case falls between the two catalytic aspartates, corresponding to the hydroxyethylene dipeptide isostere in aliskiren and other renin inhibitors of the same basic structure. Currently, this compound is in phase I clinical trials. Very recently, Novartis scientists published the structure of a highly selective inhibitor where the transition-state isostere is now a substituted piperidine (Figure 5.7) [186]. However, without the original identification of the transition state isostere from pepstatin, none of these compounds would have been made. Finally, a review covering the preparation of biologically active peptides has just been published and makes interesting reading when the methods are compared with those covering the synthesis of pseudo-peptides that inhibit aspartic proteinases [187].

B. Anticholesterolemics

In previous editions, we covered the story of the discovery and development of the "statins," anticholesterolemic compounds based upon the fungal metabolite compactin, and how all of the agents that were approved—including the best-selling drug of all time, atorvastatin—had the same "warhead" as compactin, and inhibited HMG-CoA reductase, the essential step in cholesterol biosynthesis in mammals.

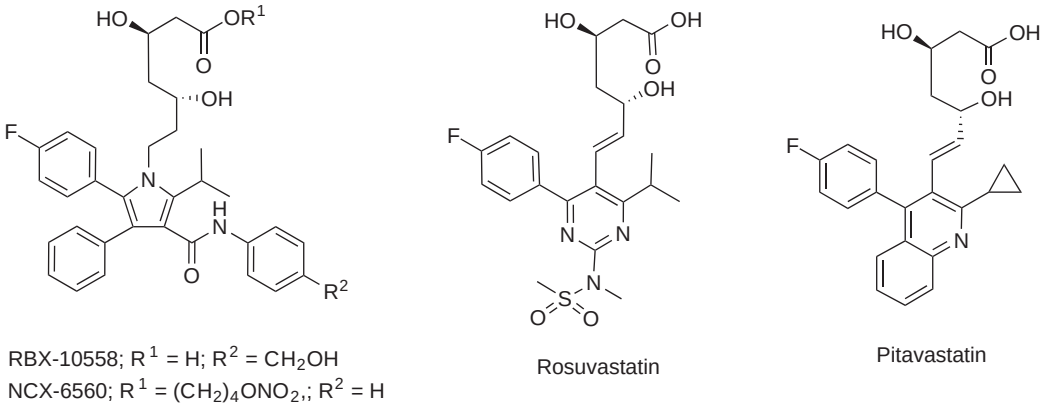


FIGURE 5.8 Modified statin-based anticholesterolemics.

1. *Modified Statins*

Since those editions, some variations on the same theme have entered clinical trials, with examples being RBx-10558 (Figure 5.8), a hydroxymethyl derivative of atorvastatin in Phase II, and NCX-6560 (Figure 5.8), a nitro-derivative of atorvastatin in Phase I. This latter compound is an interesting example, as it is a nitric oxide donating compound that is being directed towards both atherosclerosis and dyslipidemia [188,189].

However, what is interesting from a medicinal chemistry aspect is the utilization of microbial enzymes to produce essential chiral components for the synthesis of rosuvastatin (Figure 5.8) and pitavastatin (Figure 5.8). Thus, scientists at Lek Pharmaceuticals were able to produce enantiomerically pure acetyloxymethylene-lactol, which was then oxidized to the lactone. A volumetric yield of 50 g $L^{-1}.h^{-1}$ using fed-batch whole *E. coli* fermentations was possible with titres of 100 g. L^{-1} [190]. A later publication from the same group showed how, by further use of *E. coli*, the chemical oxidation was performed biologically using PQQ-dependent glucose dehydrogenase, overexpressed in the same organism [191].

C. Immunosuppressives

1. *Cyclosporin Derivatives*

In the previous edition, we covered the history of cyclosporin, tacrolimus, and rapamycin. Since that discussion, a number of variations on the rapamycin molecule have been utilized as antitumor agents rather than being further developed as immunosuppressive agents, though one human metabolite, novolimus, where the methoxyl group at C7 is demethylated, is now in use as a component of a stent recently approved in the EU.

The modified cyclosporin derivative voclosporin (Figure 5.9) was recently assigned orphan drug status in the EU for the treatment of noninfectious uveitis after a rather checkered past in being licensed to many companies and in many countries in the last fifteen or so years. It is currently in Phase II trials for nephritis and lupus, and syntheses from cyclosporin were published in the review by McIntyre and Castaner in 2004 [192] and in greater detail by Maeng et al. in 2012 [193].

2. *Fingolimod and KRP203*

In 2010, the FDA approved the immunosuppressant fungal metabolite derived agent fingolimod (FTY720; Figure 5.9) as the first novel treatment for multiple sclerosis for close to forty years. This compound is an activator of sphingosine-1-phosphate receptors and was derived from the fungal metabolite ISP-1 or myriocin (Figure 5.9). It was first synthesized by the Fujita group at Kyoto University in 1992 with a report in 1995 [194]. It first received approval in Russia in 2010, which was closely followed by FDA approval for treatment of multiple sclerosis. A discussion of the route to the drug from the original fungal metabolite was given by Strader et al. in 2011 [195]. This is not the end of the story with this molecule, however, as similarly to the rapamycins it is now a possible antitumor drug. It does not appear to function via the same mechanism observed when acting as an immunosuppressant [196], since the phosphorylated derivative is not necessary for its cytotoxicity. Recently, another close relative to FTY720, the compound known as KRP203 (Figure 5.9) that functions via the same

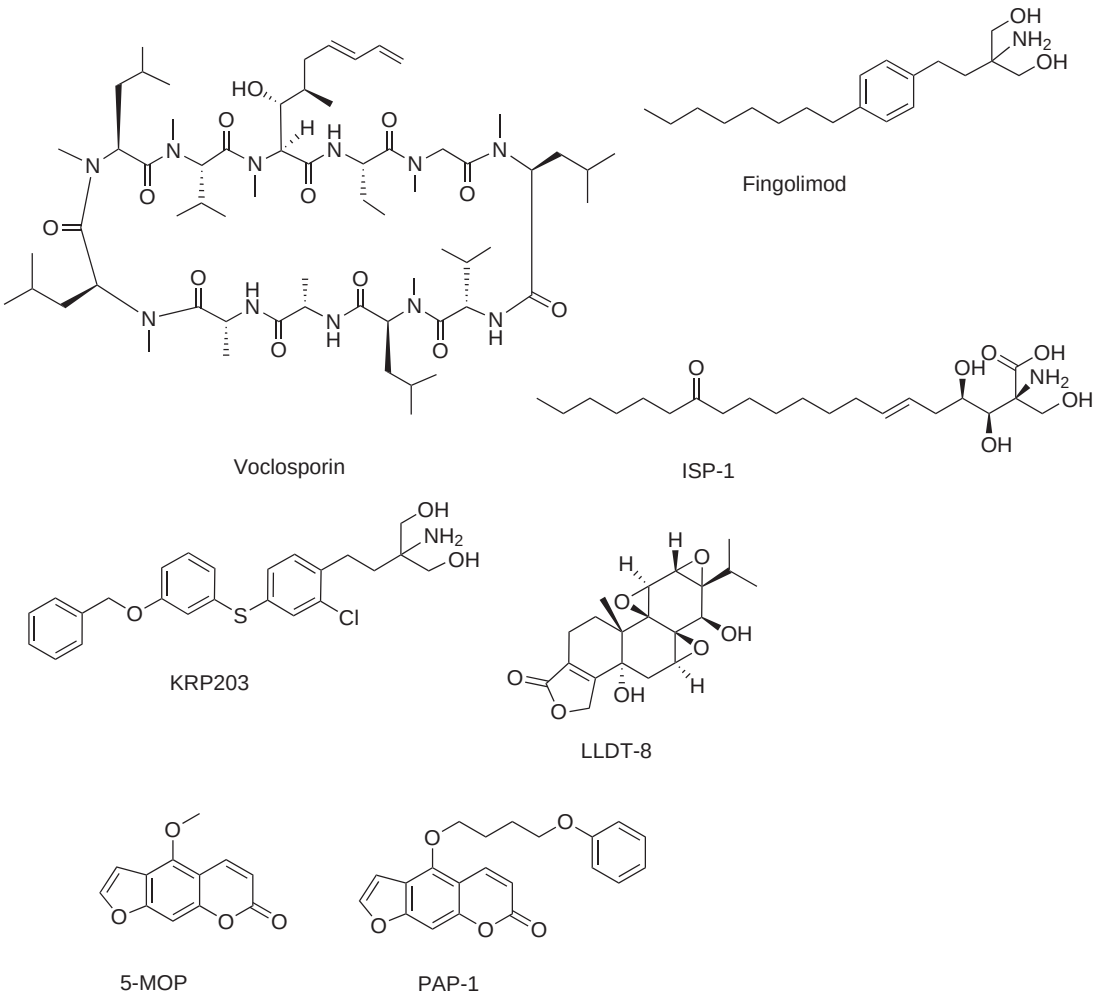


FIGURE 5.9 Immunosuppressives.

mechanism, was shown to be—like FTY720—active against amyloid-beta production in neurons and may also have other uses neurologically [197,198]. It is currently in Phase II trials for autoimmune diseases. A Phase I trial for treatment of patients undergoing stem cell transplantation is underway in Switzerland, sponsored by Novartis.

3. *Triptolide Derivative LLDT-8*

An old structure from traditional Chinese medicine (TCM) has been subtly modified to give (5*R*)-5-hydroxy-triptolide (LLDT-8) (Figure 5.9). It was reported in 2005 to show *in vitro* and *in vivo* immunosuppressive effects [199], with a recent review from the same group covering further work on the molecule [200]. The compound is currently in Phase I trials in the People's Republic of China for treatment of arthritis and rheumatoid arthritis.

4. *Psoralen Derivatives*

Another old natural product structure, 5-methoxypsoralen (5-MOP; Figure 5.9), had been used by Bohuslavizki et al. in 1994 as a potential Kv1.3 channel blocker [201,202]. However, this molecule was not specific enough for Kv1.3, so by using classical methods they modified 5-MOP via Psora-4 to PAP-1 (Figure 5.9) [203]. Molecules with this type of mechanism of action are now known to be potential drugs against T-cell mediated autoimmune diseases [204], and this compound under a variety of code names (AS-77, PAP-1, or SPS-4251) is in Phase I trials in Germany as a topical treatment for psoriasis.

D. Antibiotics

1. General Comments

In the last ten or so years, almost all of the major pharmaceutical companies have ceased work on antibiotics, due mainly to a perceived lack of return on investment (ROI) often linked to the large number of patients that have to be treated in Phase III clinical trials. In the last few years, although companies had used massive screening campaigns to search for antibiotics using target-based approaches, none of them produced significant results and none appeared to have utilized natural products as their source of agents. The 2007 review by then GSK scientists should be read in order to see the amount of work involved. Over seventy targets were screened against the GSK chemical libraries over a seven year period without a successful ending [205].

2. Current Antibiotics in Clinical Use and Development

In the seven years since the publication of the last edition, natural products and derivatives of natural products have continued to be the sources of most of the antibacterial agents approved in the period from 2007 to mid-2014. In 2011, the macrolidic antibiotic fidaxomicin (Figure 5.10) was approved for the treatment of *Clostridium difficile*-associated diarrhea. This compound, known originally as Tiacumicin B, was isolated from the fermentation of *Dactylosporangium auranticum* subsp. *hamdenensis* (NRRL 18085). Another compound that was isolated from a culture of *Actinoplanes deccanensis* (ATCC 21983) was named lipiarmycin A3 and is identical to tiacumicin B [206]. Currently, it has been approved almost worldwide for this use under a variety of different licensing agreements.

Modified natural products are still valid sources of antibacterials in all current cases involving the modification of very old antibiotic structures by medicinal chemists. Thus, retapamulin (Figure 5.10) from GSK was approved in 2007 and is a modern derivative of an antibiotic approved many years ago for veterinary usage. Two cephalosporins have been approved since 2008. The first, ceftobiprole medocaril (Figure 5.10), was originally approved in 2008, withdrawn in 2010, then resubmitted and approved in the EU in 2013 for hospital-acquired pneumonia. This relatively rapid approval, withdrawal, and reapproval is most unusual for a pharmaceutical compound. The second, ceftaroline fosamil acetate (Figure 5.10), was approved in the USA in late 2010 for treatment of methicillin resistant *S. aureus* (MRSA) skin infections and launched in early 2011.

Finally, just to demonstrate that fifty-year-old structures still have validity when medicinal chemists apply their craft to what may be considered "privileged structures," in 2009 the substituted vancomycin derivative telavancin HCl (Figure 5.10) was approved by the FDA for the treatment of complicated skin and skin structure infections (cSSSI) caused by Gram-positive bacteria [207]. Very recently, the FDA approved two further modified glycopeptides of the vancomycin class for the treatment of acute Gram-positive bacterial skin and soft tissue infections under the names dalbavancin and oritavancin, respectively. The synthetic route to dalvabancin that started with the naturally occurring glycopeptide A-40,926 was described by Malabarba et al. in 1995 [208], and in 2012 the crystal structure of this agent bound to its bacterial cell wall target was reported. This may well open up modification methods for improving the binding of such agents to their target in the presence of serum proteins [209]. The synthesis of oritavancin starting from the glycopeptide antibiotic A82846B was described by Cooper et al in 1996 [210], and an analysis of this drugs televancin and dalbavancin was recently published by Liapikou and Torres [211].

E. Anticancer Agents from Terrestrial Microbes

1. Epothilones

In the previous edition, we covered the epothilones in detail. Since that time, many have not survived clinical trials and as of early September 2014 the following are the relative numbers. The Thomson-Reuters Integrity™ database contains 199 epothilone structures. Filtering for "lead compounds" lowers the count to seventy-three, and with the filter "under active development" the number is reduced to four. Ixabepilone was launched in 2007 and is in clinical trials for other cancers, having been approved initially for resistant breast carcinomas. Sagopilone (ZK-EPO) and epothilone B are in Phase II (structures are in the 3rd edition), leaving only the Danishevfsky group's modified molecule isoxazolefludelone (Figure 5.11) [212,213] in Phase I and actively recruiting patients. A recent review by Altmann and Schinzer has excellent coverage of the epothilones, particularly from a chemistry aspect [214].

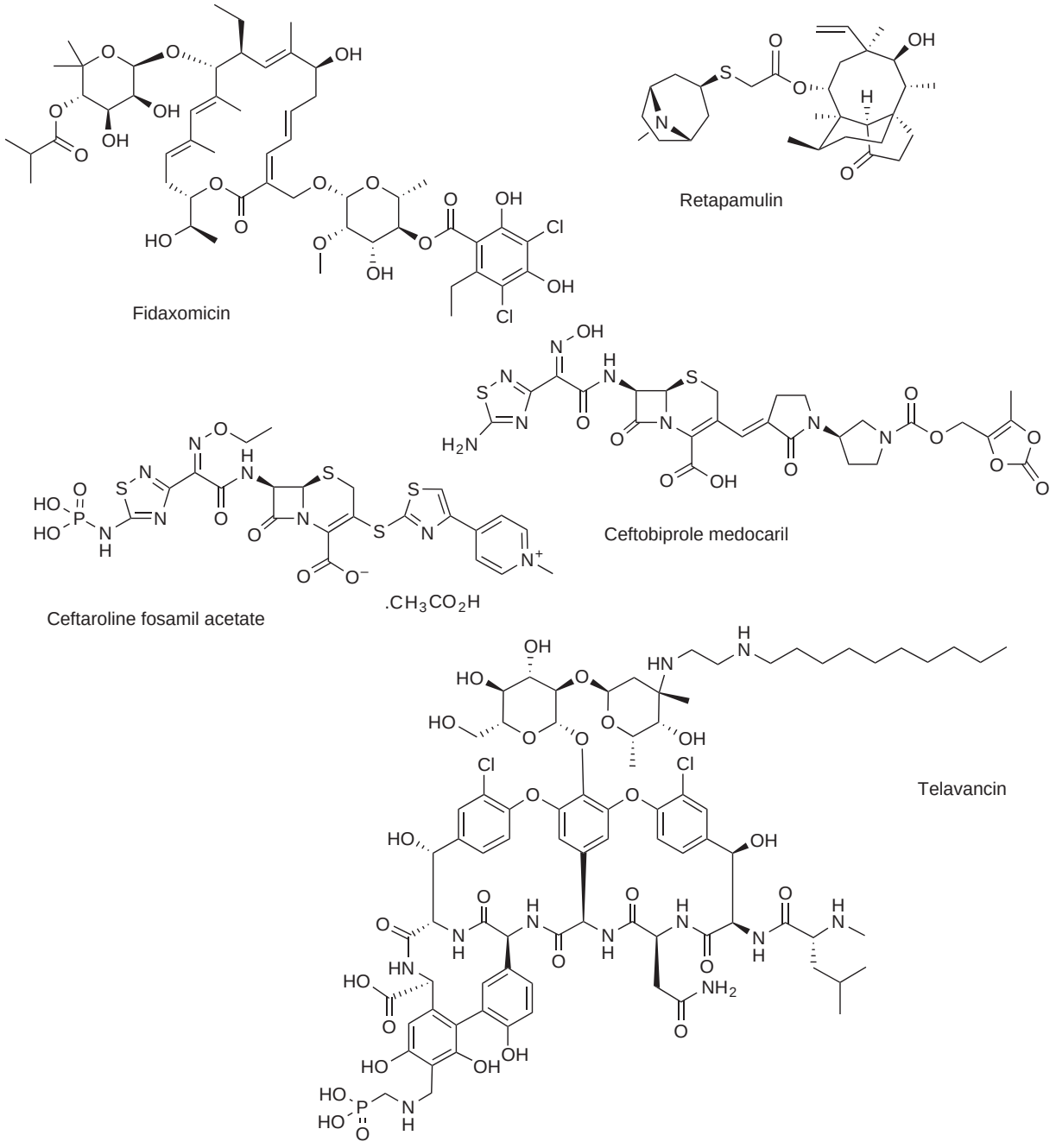


FIGURE 5.10 Natural-product based antibacterial agents.

2. *Other Examples*

There are a significant number of antitumor agents derived from microbial sources in use and in clinical trials that have been covered in a 2013 review by one of the authors of this chapter. This review can be consulted for the other major classes of antitumor agents from terrestrial microbial sources [126].

F. Anticancer Agents from Marine-Sourced Microbes

1. *General Comment*

Although not often realized, some anticancer agents from marine sources are not usually listed as having a microbial source. As mentioned previously (in Section III.B.5.a), these would include salinosporamide A.

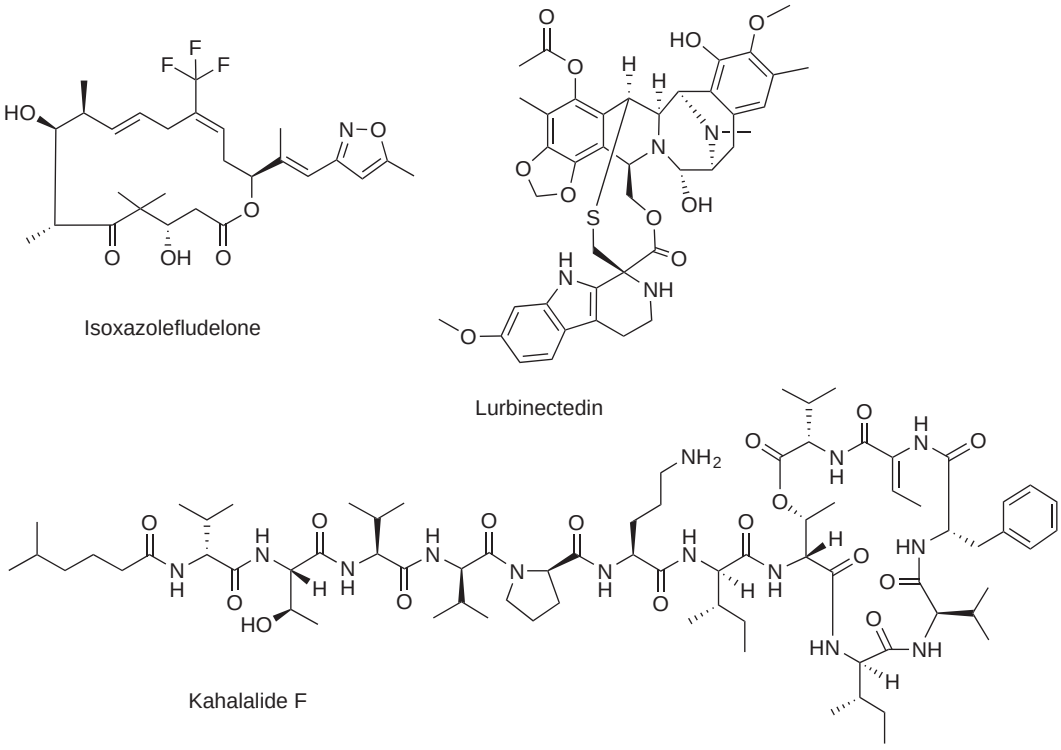


FIGURE 5.11 Terrestrial and marine microbial anticancer agents.

2. *Kahalalides*

The initial compound in this series, kahalalide F (KF; Figure 5.11), was isolated from the mollusk *Elysia rufescens* and was found to be a component of the green alga on which it was feeding. Since the amount of material isolated from the mollusk was significantly higher than that found in the alga, the probability is that the material was being absorbed from the alga and concentrated in the mollusk. A patent on a vibrio species that might be the actual producer was filed. Much fuller details are available in the 2011 review by Gao and Hamann [215]. PharmaMar had a derivative of KF in clinical trials but removed it in late 2012, though kahalalide F is still listed on the EU clinical trials web site under EudraCT number 2004-001254-29 in Phase II in Spain at the time of writing (late September 2014).

3. *Et743 and Derivatives*

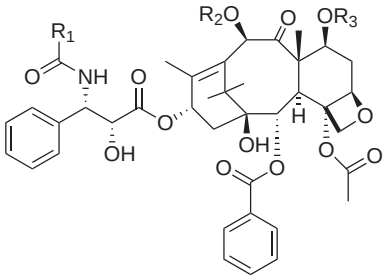
Currently, a derivative of trabectedin known as Lurbinectedin (Figure 5.11) is in Phase II trials, and it can be considered to be derived from a microbial product, even though the microbe that is considered the producing organism for ecteinascidin 743 has not been cultured as yet. The full biosynthetic pathway was elaborated by Sherman's group and reported in 2011 [216].

G. Anticancer Agents from Plants

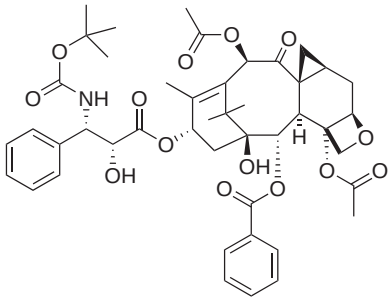
The initial studies of plant natural products as potential anticancer agents were made in the 1950s, and the vinca alkaloids and the podophyllotoxin analogues etoposide and teniposide were the first fruits of these investigations. Later work led to the very important taxane drugs and the camptothecin analogues.

1. *Paclitaxel*

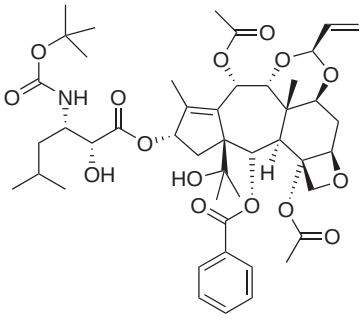
Paclitaxel (Taxol® (Figure 5.12) is the most exciting plant-derived anticancer drug discovered in the last fifty years. It occurs along with its key precursor 10-deacetylbaccatin III in the bark and leaves of various *Taxus* species [217]. It was found to act by promoting the assembly of tubulin into microtubules, and the discovery of this activity in 1979 by Schiff and Horwitz [218] was an important milestone in its development as a drug. After an



Paclitaxel (TaxolTM) R_1 = Ph, R_2 = $COCH_3$, R_3 = H
Docetaxel R_1 = $OC(CH_3)_3$, R_2 = R_3 = H
Cabazitaxel R_1 = $OC(CH_3)_3$, R_2 = R_3 = CH_3



Larotaxel



TPI-287, NBT-287, ARC-100

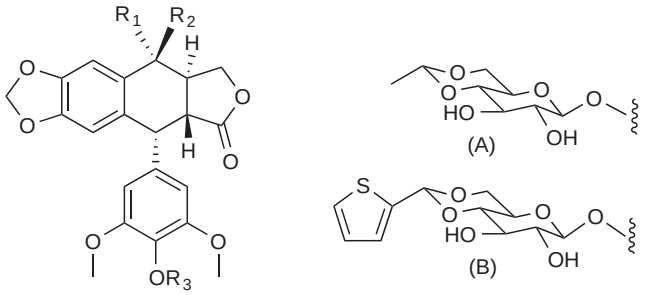
FIGURE 5.12 Taxane anticancer agents.

extended period of development [219], it was finally approved for clinical use against refractory ovarian cancer in 1992 and against refractory breast cancer in 1994. It then became a blockbuster drug, with annual sales peaking at $1.6 billion in 2000 [220] when it lost its protected status and became a generic drug. Annual sales are now approximately $200 million.

The success of paclitaxel spurred an enormous amount of work on the synthesis of analogues, and this work has been summarized in several reviews [164,221–225]. The first analogue to be developed was its close chemical relative docetaxel (Taxotere®; Figure 5.12) [223,226]. The albumin-bound formulation of paclitaxel known as Abraxane® was launched in 2005. This formulation offers some important clinical advantages compared with the original Cremophor formulation [227]. The dimethyl ether paclitaxel analogue cabazitaxel (Figure 5.12) was approved in 2010 [228,229]. One of its advantages is that it is less susceptible to drug resistance than paclitaxel and docetaxel. In other developments, larotaxel (Figure 5.12) was in Phase III clinical trials in combination with cisplatin, but the sponsor of the trial elected to discontinue development since the combination did not improve outcomes in comparison with gemcitabine/cisplatin [230]. Clinical trials are continuing on ortataxel [231] and NBT-287/TPI-287, where it should be noted that the structure of this compound has been revised (Figure 5.12) from the one originally published. Two new formulations of paclitaxel are also in Phase III trials: NK105, a paclitaxel-incorporating micellar nanoparticle [232], and paclitaxel poliglumex [233]. The reviews cited above should be consulted for information on other new analogues in development.

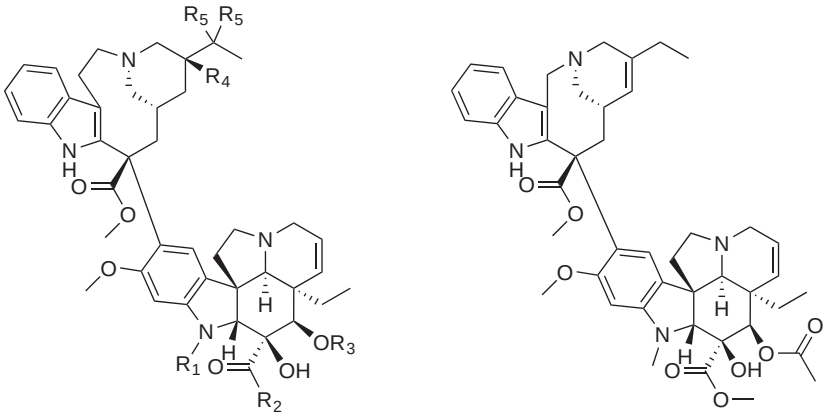
2. *Podophyllotoxin Derivatives*

Podophyllotoxin (Figure 5.13) was originally isolated as the major cytotoxic component of the May Apple, *Podophyllum peltatum*, and was placed into clinical trials, but it was shelved due to intractable toxicities [234]. Elegant work then led to the semisynthetic compounds etoposide (Figure 5.13) and teniposide (Figure 5.13), which are currently in clinical use worldwide [235]. The water-soluble derivative etopophos (Figure 5.13) was approved for intravenous use [236], and the newer drug candidates NK 611, GL-331, TOP-53, and tafluposide had been developed through various stages [237] but have not become drugs. None are in still in clinical trials as of September 2014. Although podophyllotoxin is a tubulin polymerization inhibitor, etoposide and teniposide function as DNA topoisomerase II inhibitors. This and other aspects of these drugs are discussed in more detail in a recent review [237].



Podophyllotoxin $R_1 = OH, R_2 = H, R_3 = CH_3$
Etoposide $R_1 = H, R_2 = A, R_3 = H$
Teniposide $R_1 = H, R_2 = B, R_3 = H$
Etopophos $R_1 = H, R_2 = A, R_3 = OP(O)(OH)_2$

FIGURE 5.13 Podophyllotoxin derivatives.



Vinblastine $R_1 = CH_3, R_2 = OCH_3, R_3 = COCH_3, R_4 = OH, R_5 = H$
Vincristine $R_1 = CHO, R_2 = OCH_3, R_3 = COCH_3, R_4 = OH, R_5 = H$
Vindesine $R_1 = CH_3, R_2 = NH_2, R_3 = H, R_4 = OH, R_5 = H$
Vinflunine $R_1 = CH_3, R_2 = OCH_3, R_3 = COCH_3, R_4 = H, R_5 = F$

Vinorelbine

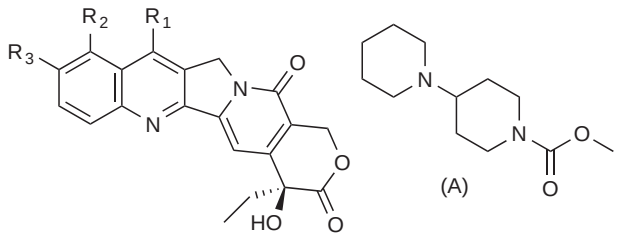
FIGURE 5.14 Natural and semi-synthetic vinca alkaloids.

3. *Vinca Alkaloids*

The bisindole alkaloids vinblastine and vincristine (Figure 5.14) were isolated from the Madagascan periwinkle as a result of a serendipitous observation made while studying extracts of the plant as a source of hypoglycaemic agents [238]. Both alkaloids are tubulin polymerization inhibitors, and they act at a different site on the tubulin molecule to that of the taxanes. Vinblastine is used in the treatment of bladder and breast cancers and Hodgkin's disease, while vincristine is used in the treatment of leukemias and lymphomas. Many variations on the basic structure of the *Vinca* alkaloids have been synthesized, and currently vinorelbine, vindesine, and vinflunine (Figure 5.14) have been approved for clinical use. Vinorelbine is used for the treatment of nonsmall-cell lung cancer, while vindesine has an important role in the treatment of acute lymphoblastic leukemia [239]. Vinflunine was approved in 2010 and is also in Phase III studies for the treatment of solid malignancies. One of these studies is for treatment of platinum resistant transitional carcinoma of the urothelium [240]. Further details of the chemistry and clinical applications of the vinca alkaloids are provided in a recent review [241].

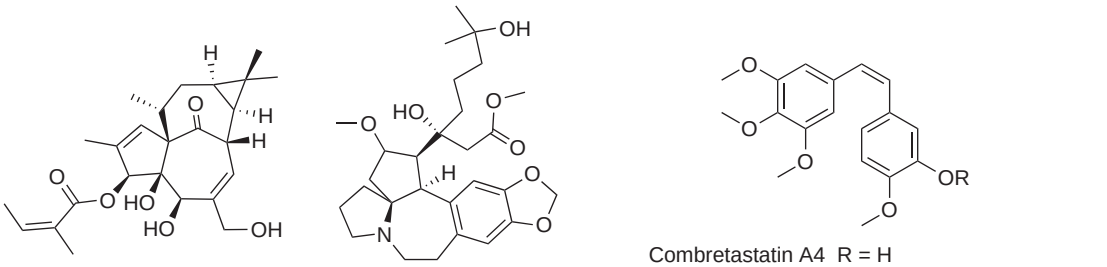
4. *Camptothecin and Derivatives*

The alkaloid camptothecin (Figure 5.15) was first isolated from the Chinese tree *Camptotheca acuminata* by Wani and Wall contemporaneously with their discovery of Taxol® [242]. It was first evaluated clinically as the sodium salt of the ring-opened lactone, but it demonstrated severe bladder toxicity and was dropped. It was resurrected as a result of its very specific biochemical activity as an inhibitor of topoisomerase I, and the



Camptothecin	$R_1 = R_2 = R_3 = H$
Topotecan	$R_1 = H$, $R_2 = CH_2N(CH_3)_2$, $R_3 = OH$
Irinotecan	$R_1 = CH_2CH_3$, $R_2 = H$, $R_3 = (A)$
Belotecan	$R_1 = CH_2CH_2NHCH(CH_3)_2$, $R_2 = R_3 = H$

FIGURE 5.15 Camptothecins and analogues.



Ingenol mebutate

Omacetaxine mepesuccinate

Combretastatin A4 R = H

Combretastatin A4 phosphate R = OPO_3^{2-} $2Na^+$

FIGURE 5.16 Other plant-derived anticancer agents.

semisynthetic derivative topotecan (Figure 5.15; Hycamptin®) has been approved as a second-line therapy for advanced ovarian cancer and for recurrent small-cell lung cancer. The analogue irinotecan (Figure 5.15; Camptosar®) has also been approved for treatment of advanced colorectal cancer, and belotecan (Figure 5.15) has been approved in South Korea for ovarian and small-cell lung cancers. All these camptothecin analogues function as topoisomerase I inhibitors, thus differentiating them from the podophyllotoxin derivatives and the anthracyclines such as Adriamycin, which function as topoisomerase II inhibitors. Further details of the chemistry and clinical applications of the camptothecins are provided in two recent reviews [243,244].

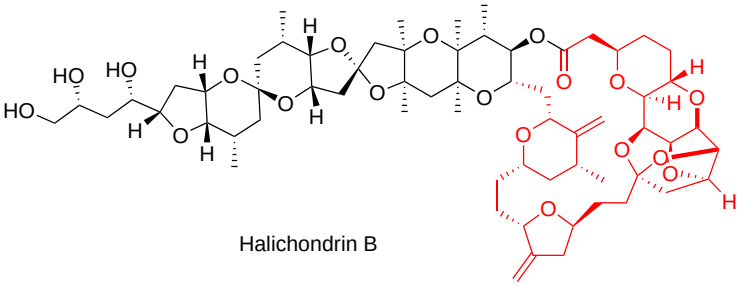
5. Other Plant-Derived Anticancer Agents

Ingenol mebutate (Picato™) is an ester of the diterpene ingenol (Figure 5.16) that was approved for treatment of actinic keratosis (a premalignant skin condition) in 2012 [245]. Homoharringtonine (Figure 5.16; Omacetaxine mepesuccinate, Synribo™) is a naturally-occurring ester of the alkaloid cephalotaxine isolated from various trees of the *Cephalotaxus* genus [246]. It is a protein synthesis inhibitor, and was approved in 2012 for treatment of chronic or accelerated phase chronic myeloid leukemia [247]. Combretastatin A4 phosphate (Figure 5.16) is a vascular disrupting agent; its progenitor combretastatin A4 was obtained from the bark of the African tree *Combretum caffrum* [248]. Its major potential application is as an adjunct to conventional chemotherapy [249].

H. Marine Agents (nonmicrobial to date)

1. Eribulin

The complex marine polyether halichondrin B (Figure 5.17) was first isolated from a Japanese sponge by Uemura and his colleagues [250], and was subsequently isolated again by Pettit from an *Axinella* sp. collected in Palau [251]. The compound had excellent bioactivity and showed a pattern of activity in the NCI 60-cell line screen comparable to the vinca alkaloids and paclitaxel [251]. The compound was only isolated in miniscule yield, however, and its complex structure appeared to make total synthesis impractical as a source for drug development. Fortunately, in the course of synthetic studies on the synthesis of halichondrin B [252], a group at



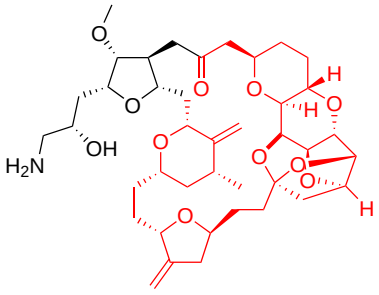


FIGURE 5.17 Halichondrin B and eribulin.

Eisai Research Institute in the USA, working closely with Kishi's synthetic group at Harvard, discovered that certain macrocyclic ketone analogues of the right hand half of halichondrin B retained all or most of the potency of the parent compound [253]. This key observation was then used as the impetus for the heroic large-scale synthesis of the analogue E7389 (Figure 5.17; Eribulin), which was approved in the USA for metastatic breast cancer in 2010. The development of E7389 (Eribulin) has been described in a recent series of papers, including discussion of the large-scale production of the most complex synthetic drug made to date [254–257].

I. Antimalarial Agents

1. *Quinine and its Analogues*

Malaria is a major scourge of humanity, and the discovery of new antimalarial drugs is a worldwide health imperative. The alkaloid quinine (Figure 5.18), isolated from the bark of the South American tree *Cinchona officinalis* and other trees of the same genus, was the first effective antimalarial agent to be discovered, and it has been called "the drug to have relieved more human suffering than any other in history" [258]. It served humanity well for about 300 years, although resistance to the drug was first noted in 1910, and it is no longer recommended by the WHO as a first line treatment for malaria. It was largely replaced in the mid-20th century by the synthetic analogue chloroquine (Figure 5.18), but resistance to this drug emerged in 1957, and it is no longer of value in many areas of the world [259]. Several other synthetic antimalarial agents have been based on the quinine pharmacophore, including mefloquine (Figure 5.18), primaquine, mepacrine, and amodiaquine, but each of these has its own disadvantages, including drug resistance and—in the case of mefloquine—psychotic side effects in some individuals.

2. *Artemisinin and its Analogues*

The Chinese plant *Artemisia annua* was used for the treatment of fevers in Chinese medicine as long ago as 340 CE, but its active constituent was not identified by Chinese scientists as qinghaosu (Figure 5.18; artemisinin) until 1971 [260]. Artemisin proved to be an excellent antimalarial agent, and artemisinin-based combination therapies (ACTs) are currently the preferred first-line antimalarial treatment for *Plasmodium falciparum*, although resistance has been observed in the Cambodia–Thailand border area [261]. Artemisinin's unusual endoperoxide bridge is a key to its mechanism of action, although the details of this mechanism are still being sorted out. A recent computational study suggested that artemisinin gets activated by iron, which in turn inhibits the calcium pump PfATP6 [262]. The importance of artemisinin has led to several synthetic and semisynthetic approaches to its production

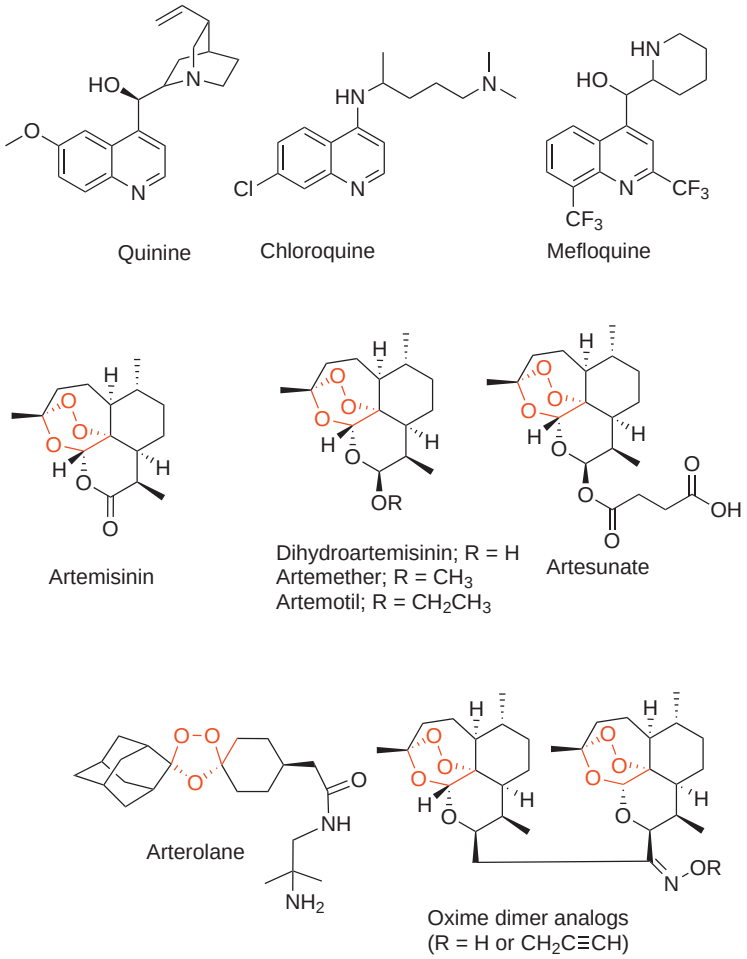


FIGURE 5.18 Natural antimalarial agents and analogues.

to complement its isolation from *A. annua*. The most fruitful current approach is the chemical conversion of artemisinic acid, produced by engineered strains of *Saccharomyces cerevisiae* [263] to artemisinin [264,265].

Many analogues and derivatives of artemisinin have been prepared in attempts to improve its activity and utility. The analogues dihydroartemisinin, artemether, artemotil, and artesunate have all found clinical use [266], while the fully synthetic arterolane and various oxime and other dimers [267] have promising activities (Figure 5.18). Other potential antimalarial natural products are described in two recent reviews [258,268].

J. Other Natural Products

The examples given above are simply a selection of the natural product and natural product analogues that have entered clinical use. Several recent reviews cover natural products as drugs and sources of structures [8,225,269–272]. These references should be consulted for further examples of how natural products have led to novel drugs for a multiplicity of diseases, and for insights into the future potential of natural products in drug discovery.

V. FUTURE DIRECTIONS IN NATURAL PRODUCTS AS DRUGS AND DRUG DESIGN TEMPLATES

A. Introduction

The probability that a directly isolated natural product (e.g., adriamycin or paclitaxel) will become a drug for a given disease is relatively low, except perhaps in the realm of antibiotics. However, the strength of the natural products approach is that these "base molecules" can serve both as leads to new active structures and as probes

for new mechanisms of action. In a similar fashion, combinatorial biosynthesis can be utilized to produce what are now being called "unnatural natural products," where the biosynthetic machinery of a microbial cell is dissected and the relevant genes are "mixed and matched," followed by expression in a suitable heterologous host. Such compounds may be used in their own right or could be the starting materials for further synthetic modifications. In addition, novel methods of chemical syntheses that have the potential to produce base "natural product" molecules that can be optimized for specific medicinal chemistry purposes are now being reported. That these ideas are not just pipe dreams can be seen in the following examples.

B. Combinatorial Chemistry

Detailed analyses of active natural product skeletons have led to the identification of relatively simple key precursor molecules that form the building blocks for use in combinatorial synthetic schemes that have produced numbers of potent molecules, thereby enabling structure activity relationships to be probed. Thus, in the study of the structure-activity relationships of the epothilones, solid-phase synthesis of combinatorial libraries was used to probe regions of the molecule important to retention or improvement of activity [273].

The approach above was further developed by the use of an active natural product as the central scaffold in the combinatorial approach in order to generate large numbers of analogues for structure-activity studies, the so-called "parallel synthetic approach." This embodies the concept of "privileged structures," originally proposed by Evans et al. [274] and advanced further by Nicolaou et al. using benzopyrans as scaffolds to produce "libraries from libraries" that had very significantly different biological activities than the starting base structure [275–278].

The Waldmann group developed the so-called "Biology-Oriented Synthesis of Natural Product-Inspired Libraries (BIOS)" approach. In the BIOS approach, a natural product inhibiting a protein target through interaction with a specific protein topology (fold) represents a biologically validated starting point for the development of closely related structures. These products may inhibit proteins with similar folds, since there are approximately a 1000 different topological folds in proteins [279–281]. This approach is similar to the privileged structure concept, but with the added dimension of using protein topology patterns as the basis for subsequent screens [281,282].

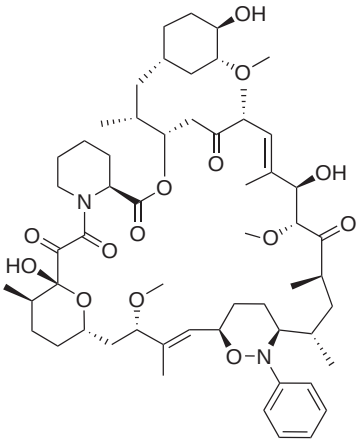
A somewhat similar approach to protein–ligand interactions has been developed by the Quinn group, who hypothesized that the interior Protein Fold Topology (PFT) characteristics of key enzymes involved in the biosynthetic pathways of a particular class of natural products (i.e., the active site topology of the final enzyme in the biosynthetic pathway) should mimic the PFT characteristics of the active site of proteins targeted by members of that natural product class. This was found to be the case with flavonoids, where the PFT characteristics of the relevant biosynthetic enzymes shared PFT characteristics of the phosphoinositol-3-kinase (PI3K) active site that is targeted by flavonoids [283]. Thus, a comparison of the biosynthetic PFT characteristics of natural products with the PFT characteristics of various targets should permit the identification of potential targets for the natural product classes, and vice versa [284]. Quinn's process can be thought of as "inside to target" whereas Waldmann's can be from "outside to target."

In the early days of combinatorial chemistry, the emphasis was on production of massive libraries of what turned out to be "flat" compounds. These did not generate many leads, and later enthusiasm waned for this approach. The use of natural product-like compounds as scaffolds, referred to as "Diversity Oriented Synthesis," led to the generation of smaller, more meaningful combinatorial libraries. This move was exemplified by the work of the Schreiber group, who combined the simultaneous reaction of maximal combinations of sets of natural-product-like core structures ("latent intermediates") with peripheral groups ("skeletal information elements") in the synthesis of libraries of over 1000 compounds bearing significant structural and chiral diversity [285–287].

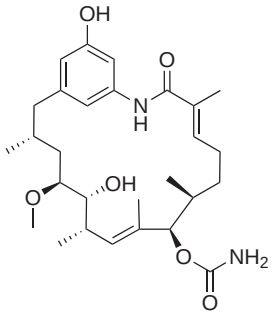
C. Natural Products as Design Templates

Natural products can also be used as design templates for more traditional chemical syntheses. This approach has been highly successful, as illustrated by some of the examples discussed in the previous sections where a modified natural product became the drug rather than the natural product itself. This section will thus simply give some additional examples to illustrate the power of the approach.

A recent book edited by Osbourn et al. [288] covers a wide area of natural-product chemistry and biology and contains short chapters covering case studies on how molecules were developed from the basic structure. Two



ILS-920



BC265

FIGURE 5.19 New compounds from natural product structures.

examples will suffice, one a modification of the basic rapamycin skeleton and the other a nonnatural ansamycin derivative with Heat Shock Protein 90 (HSP90) inhibitory activity. In the first case, Graziani described the route from rapamycin to ILS-920 (Figure 5.19) and the biological assays that demonstrated blood brain barrier penetratability, demonstrating the interplay of microbiology, chemistry, and bioassay [289]. The second case study described the development using biosynthetic medicinal chemistry of BC265 (Figure 5.19) via modification of the macbecin biosynthetic cluster, in order to generate compounds that would lead to a nonquinonoid HSP90 inhibitor [290]. Early publications on aspects of the work with the macbecin cluster were published in 2008 and should be consulted as well [291,292].

D. Metabolomics; Interaction of Microbes, Genomics and Chemistry

Recently, metabolomic profiling of natural product extracts, coupled with pharmacological profiling using appropriate bioassays, has been applied as a means to expedite the drug discovery process. The metabolomic approach enables the early detection and characterization of bioactive compounds of potential interest without the need for isolating active principles. Targeted isolation can then be tailored to the nature of compounds of interest. The application of metabolomic based techniques to drug discovery and development has been reviewed [293], and several reviews related to the discovery of bioactive compounds from natural sources have been published [294–297].

Very recently, three papers have been published that show the interplay between genomics, metabolomics, and natural product structures. Two were published in the same issue of *Current Opinion in Microbiology* in 2014. One, from Brady's group at Rockefeller University [298], discusses methods and techniques used in discovering new antibiotics from diverse environmental biomes. The second, from Rhee's group at Cornell Medical College [299], discusses the advances in instrumentation that allow dissection of the metabolic processes in cells. In the third, from Fischbach's group at the University of California, San Francisco, they demonstrated that by "genome mining" the massive amount of information in the human biome project they could identify genes from *lactobacilli* isolated from human vaginal specimens that produce a member of the thiopeptide antibiotic family, which they have named lactocillin. This agent is close to the clinical candidate LFF571 that is currently in Phase II clinical trials against *C. difficile* [300]. These three papers show the vast possibilities for the discovery of novel agents from the massive amounts of data now available for interrogation as to biosynthetic clusters, then coupling to medicinal chemistry. This is illustrated by the case of LFF571, where a 4-aminothiazoyl analogue of the naturally occurring GE2270 A became the lead candidate.

VI. SUMMARY

We have attempted in this chapter to show how varied are the natural product structures with bioactivity that can be isolated from nominal sources such as plants and marine invertebrates. Nowadays, it is highly probable that the actual producer of the majority of compounds found may well be microbes (cultivatable or as yet not

cultured) or a nominal source linked to a microbe within the macroorganism. The product found may often be a simple modification by the host or a commensal microbe, or an almost direct transfer from the epiphytic or endophytic microbe to the host from which it was identified.

Finally, the new frontier in natural product discovery may well be the interrogation of the microbial genomes that have been sequenced and are either published or are in proprietary databases at the moment. An example of the latter is the work being reported by the company Warp Drive Bio on massive sequencing of actinomycete genomes in order to search for gene sequences that have the potential to produce novel structures [301].

> Mother Nature has the compounds. It is our job to find and develop them for the good of all.

References

[1] Borchardt JK. The beginnings of drug therapy: ancient mesopotamian medicine. Drug News Perspect 2002;15:187–92.
[2] Huang KC. The pharmacology of chinese herbs. 2nd ed. Boca Raton, FL: CRC Press; 1999.
[3] Kapoor LD. CRC handbook of ayurvedic medicinal plants. Boca Raton, FL: CRC Press; 1990.
[4] Dev S. Ancient-modern concordance in ayurvedic plants: some examples. Environ Health Perspect 1999;107:783–9.
[5] Moerman DE. Medicinal plants of Native America. Ann Arbor, MI: University of Michigan Museum of Anthropology; 1986.
[6] Johnson T. CRC Ethnobotany desk reference. Boca Raton, FL: CRC Press; 1999.
[7] Farnsworth NR. The role of ethnopharmacology in drug development. In: Chadwick DJ, Marsh J, editors. Bioactive compounds from plants. Chichester: Wiley; 1990. p. 2–21.
[8] Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. J Nat Prod 2012;75:311–35.
[9] Newman DJ, Cragg GM, Snader KM. The influence of natural products upon drug discovery. Nat Prod Rep 2000;17:215–34.
[10] US Food and Drug Administration. Summary of novel new drugs. Available from: <http://www.fda.gov/downloads/drugs/developmentapprovalprocess/druginnovation/ucm381803.pdf>; 2013.
[11] Rouhi M. Rediscovering natural products. Chem Eng News 2003;81:77–91.
[12] Borman S. Improving efficiency. Chem Eng News 2006;84:56–78.
[13] Li JW, Vederas JC. Drug discovery and natural products: end of an era or an endless frontier? Science 2009;325:161–5.
[14] Nicolaou KC. Advancing the drug discovery and development process. Angew Chem Int Ed 2014;53:9128–40.
[15] Haslam E. Secondary metabolism: fact and fiction. Nat Prod Rep 1986;3:217–49.
[16] Williams DH, Stone MJ, Hauck PR, Rahman SK. Why are secondary metabolites (natural products) biosynthesized? J Nat Prod 1989;52:1189–208.
[17] Harborne JB. Role of secondary metabolites in chemical defence mechanisms in plants. In: Chadwick DJ, Marsh J, editors. Bioactive compounds from plants. Chichester: Wiley; 1990. p. 126–39.
[18] Schröeder FC, Farmer JJ, Attygalle AB, Smedley SR, Eisner T, Meinwald J. Combinatorial chemistry in insects: a library of defensive macrocyclic polyamines. Science 1998;281:428–31.
[19] Han TS, Teichert RW, Olivera BM, Bulaj G. Conus venoms: a rich source of peptide-based therapeutics. Curr Pharm Des 2008;14:2462–79.
[20] Lewis RJ, Dutertre S, Vetter I, Christie MJ. Conus venom peptide pharmacology. Pharmacol Rev 2012;64:259–98.
[21] Wallace MS. Ziconotide: a new nonopioid intrathecal analgesic for the treatment of chronic pain. Expert Rev Neurother 2006;6:1423–8.
[22] Czaran TL, Hoekstra RF, Pagie L. Chemical warfare between microbes promotes biodiversity. Proc Natl Acad Sci U S A 2002;99:786–90.
[23] Bais HP, Vepachedu R, Gilroy S, Callaway RM, Vivanco JM. Allelopathy and exotic plant invasion: from molecules and genes to species interactions. Science 2003;301:1377–80.
[24] Li ZH, Wang Q, Ruan X, Pan CD, Jiang AD. Phenolics and plant allelopathy. Molecules 2010;15:8933–52.
[25] Everts S. Bacterial conversations. Chem Eng News 2006;84:17–26.
[26] Rice SA, McDougald D, Kumar N, Kjelleberg S. The use of quorum-sensing blockers as therapeutic agents for the control of biofilm-associated infections. Curr Opin Investig Drugs 2005;6:178–84.
[27] Njoroge J, Sperandio V. Jamming bacterial communication: new approaches for the treatment of infectious diseases. EMBO Mol Med 2009;1:201–10.
[28] Sintim HO, Smith JA, Wang J, Nakayama S, Yan L. Paradigm shift in discovering next-generation anti-infective agents: targeting quorum sensing, c-di-GMP signaling and biofilm formation in bacteria with small molecules. Future Med Chem 2010;6:1005–35.
[29] Roy V, Adams BL, Bentley WE. Developing next generation antimicrobials by intercepting AI-2 mediated quorum sensing. Enzyme Microb Technol 2011;49:113–23.
[30] Bhardwaj AK, Vinothkumar K, Rajpara N. Bacterial quorum sensing inhibitors: attractive alternatives for control of infectious pathogens showing multiple drug resistance. Recent Pat Antiinfect Drug Discov 2013;8:68–83.
[31] Scutera S, Zucca M, Savoia D. Novel approaches for the design and discovery of quorum-sensing inhibitors. Expert Opin Drug Discov 2014;9:353–66.
[32] Figueroa M, Jarmusch AK, Raja HA, El-Elimat T, Kavanaugh JS, Horswill AR, et al. Polyhydroxyanthraquinones as quorum sensing inhiitors from the guttates of Penicillium restrictum and their analysis by desorption electrospray ionization mass spectrometry. J Nat Prod 2014;77:1351–8.
[33] Melander RJ, Minvielle MJ, Melander C. Controlling bacterial behavoir with indole-containing natural products and derivatives. Tetrahedron 2014;70:6363–72.

[34] Fletcher MH, Jennings MC, Wuest WM. Draining the moat: disrupting bacterial biofilms with natural products. Tetrahedron 2014;70:6373–83.
[35] Meijer L. Le cycle de division cellulaire et sa regulation [The cycle of cellular division and its regulation]. Oncologie 2003;5:311–26 [in French]
[36] Klekota J, Brauner E, Roth FP, Schreiber SL. Using high-throughput screening data to discriminate compounds with single-target effects from those with side effects. J Chem Inf Model 2006;46:1549–62.
[37] Liu B, Li S, Hu J. Technological advances in high-throughput screening. Am J Pharmacogenomics 2004;4:263–76.
[38] Mullin R. Drug discovery. Chem Eng News 2004;82:23–32.
[39] Wang J, Soisson SM, Young K, Shoop W, Kodali S, Galgoci A, et al. Platensimycin is a selective FabF inhibitor with potent antibiotic properties. Nature 2006;441:358–61.
[40] Ghosh AK, Xi K. Platensimycin and platencin. In: Hanessian S, editor. Natural products in medicinal chemistry. Switzerland: Wiley-VCH Verlag; 2014. p. 271–300.
[41] Cragg GM, Katz F, Newman DJ, Rosenthal J. The impact of the United Nations convention on biological diversity on natural products research. Nat Prod Rep 2012;29:1407–23.
[42] ten Kate K, Wells A. Benefit-sharing case study. The access and benefit-sharing policies of the United States National Cancer Institute: a comparative account of the discovery and development of the drugs calanolide and topotecan. Submission to the executive secretary of the convention on biological diversity. Kew, UK: The Royal Botanic Gardens; 1998.
[43] Rosenthal JP. Drug discovery, economic development, and conservation: the international cooperative biodiversity groups. Pharm Biol 1999;37:5.
[44] ten Kate K, Laird SA. The commercial use of biodiversity: access to genetic resources and benefit-sharing. London, UK: Earthscan Publications, Ltd; 1999.
[45] Balick MJ. Ethnobotany and the identification of therapeutic agents from the rainforest. In: Chadwick DJ, Marsh J, editors. Bioactive compounds from plants. Chichester: Wiley; 1990. p. 22–39.
[46] Cox PA. Ethnopharmacology and the search for new drugs. In: Chadwick DJ, Marsh J, editors. Bioactive compounds from plants. Chichester: Wiley; 1990. p. 40–55.
[47] Plotkin MJ. Conservation, ethnobotany, and the search for new jungle medicines: pharmacognosy comes of age…again. Pharmacotherapy 1988;8:257–62.
[48] Balandrin MF, Kinghorn AD, Farnsworth NR. Plant-derived natural products in drug discovery and development: an overview. In: Kinghorn AD, Balandrin MF, editors. Human medicinal agents from plants. Washington, DC: Amer Chem Soc; 1993. p. 2–12.
[49] Raskin I, Ribnicky DM, Komarnytsky S, Ilic N, Poulev A, Borisjuk N, et al. Plants and human health in the twenty-first century. Trends Biotechnol 2002;20:522–31.
[50] Poulev A, O'Neal JM, Logendra S, Pouleva RB, Timeva V, Garvey AS, et al. Elicitation, a new window into plant chemodiversity and phytochemical drug discovery. J Med Chem 2003;2542–7.
[51] McCoy E, O'Connor SE. Directed biosynthesis of alkaloid analogs in the medicinal plant Catharanthus roseus. J Am Chem Soc 2006;128:14276–7.
[52] Newman DJ, Cragg GM. Marine natural products and related compounds in clinical and advanced preclinical trials. J Nat Prod 2004;67:1216–38.
[53] Newman DJ, Hill RT. New drugs from marine microbes: the tide is turning. J Ind Microbiol Biotechnol 2006;33:539–44.
[54] Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotech 2008;26:1135–45.
[55] Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, et al. The complete genome sequence of a Neanderthal from the Altai Mountains. Nature 2014;505:43–9.
[56] Wu D, Hugenholtz P, Mavromatis K, Pukall R, Dalin E, Ivanova NN, et al. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature 2009;462:1056–60.
[57] Birren BW, Clifton SW, Highlander SK, Nelson KE, Petrosino JF, Sodergren E, et al. A catalog of reference genomes from the human microbiome. Science 2010;328:994–9.
[58] Smith MI, Turpin W, Tyler AD, Silverberg MS, Croitoru K. Microbiome analysis: from technical advances to biological relevance. F1000Prime 2014;6:51.
[59] Ohnishi Y, Ishikawa J, Hara H, Suzuki H, Ikenoya M, Ikeda H, et al. Genome sequence of the streptomycin-producing microorganism Streptomyces griseus IFO 13350. J Bact 2008;190:4050–60.
[60] von Döhren H. A survey of nonribosomal peptide synthetase (NRPS) genes in Aspergillus nidulans. Fungal Genet Biol 2009;46: S45–52.
[61] Sanchez JF, Somoza AD, Keller NP, Wang CCC. Advances in Aspergillus secondary metabolite research in the post-genomic era. Nat Prod Rep 2012;29:351–71.
[62] Jain S, Keller N. Insights to fungal biology through LaeA sleuthing. Fungal Biol Rev 2013;27:51–9.
[63] Lim FY, Keller NP. Spatial and temporal control of fungal natural product synthesis. Nat Prod Rep 2014;31:1277–86.
[64] Williams RB, Henrikson JC, Hoover AR, Lee AE, Cichewicz RH. Epigenetic remodeling of the fungal secondary metabolome. Org Biomol Chem 2008;7:435–8.
[65] Cichewicz RH. Epigenome manipulation as a pathway to new natural product scaffolds and their congeners. Nat Prod Rep 2010;27:11–22.
[66] Zuck KM, Shipley S, Newman DJ. Induced production of N-formyl alkaloids from Aspergillus fumigatus by co-culture with Streptomyces peucetius. J Nat Prod 2011;74:1653–7.
[67] Handelsman J, Rondon MR, Brady SF, Clardy J, Goodman RM. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 1998;5:R245–9.
[68] Nováková J, Farkašovský M. Bioprospecting microbial metagenome for natural products. Biologia 2013;68:1079–86.
[69] Persidis A. Extremophiles. Nat Biotechnol 1998;16:593–4.

[70] Abe F, Horikoshi K. The biotechnological potential of piezophiles. Trends Biotechnol 2001;19:102–8.
[71] Rossi M, Ciaramella M, Cannio R, Pisani FM, Moracci M, Bartolucci S. Extremophiles 2002. J Bacteriol 2003;185:3683–9.
[72] Amato A. Microbes live near undersea CO_2 lake. Chem Eng News 2006;84:14.
[73] Short PL. New Zealand plays to its strengths. Chem Eng News 2007;85:20–1.
[74] Cavicchioli R, Siddiqui KS, Andrews D, Sowers KR. Low-temperature extremophiles and their applications. Curr Opin Biotechnol 2002;13:253–61.
[75] Johnson DB, Hallberg KB. The microbiology of acidic mine waters. Res Microbiol 2003;154:466–73.
[76] Stierle AA, Stierle DB, Kemp K. Novel sesquiterpenoid matrix metalloproteinase-3 inhibitors from an acid mine waste extremophile. J Nat Prod 2004;67:1392–5.
[77] Stierle DB, Stierle AA, Hobbs D, Stokken J, Clardy J. Berkeleydione and berkeleytrione, new bioactive metabolites from an acid mine organism. Org Lett 2004;6:1049–52.
[78] Stierle AA, Stierle DB. Bioprospecting in the Berkeley pit: the use of signal transduction enzyme inhibition assays to isolate bioactive secondary metabolites from the extremophilic fungi of an acid mine waste lake. Stud Nat Prod Chem 2013;39:1–45.
[79] Arto T, Mendoza A, Fañanás FJ, Rodríguez F. (-)-Berkelic acid: lessons learned from our investigations on a scalable total synthesis. Strategies Tactics Org Synth 2014;10:33–50.
[80] Niu S, Liu D, Hu X, Proksch P, Shao H, Lin W. Spiromastixones A-O, antibacterial chlorodepsidones from a deep-sea-derived Spiromastix sp. fungus. J Nat Prod 2014;77:1021–30.
[81] Liu J-T, Lu X-L, Liu X-Y, Gao Y, Jiao B-H, Zheng H. Bioactive natural products from the Antarctic and Arctic organisms. Mini-Revs Med Chem 2013;13:617–26.
[82] Mosey RA, Floreancig PE. Isolation, biological activity, synthesis, and medicinal chemistry of the pederin/mycalamide family of natural products. Nat Prod Rep 2012;29:980–95.
[83] Skropeta D, Wei L. Recent advances in deep-sea natural products. Nat Prod Rep 2014;31:999–1025.
[84] Strobel G, Daisy B. Bioprospecting for microbial endophytes and their natural products. Microbiol Mol Biol Rev 2003;67:491–502.
[85] Strobel G, Daisy B, Castillo U, Harper J. Natural products from endophytic microorganisms. J Nat Prod 2004;67:257–68.
[86] Stierle A, Strobel G, Stierle D. Taxol and taxane production by Taxomyces andreanae, an endophytic fungus of Pacific yew. Science 1993;260:214–16.
[87] Li J-Y, Sidhu RS, Bollon A, Strobel GA. Stimulation of taxol production in liquid cultures of Pestalotiopsis microspora. Mycolog Res 1998;102:461–4.
[88] Eyberger AL, Dondapati R, Porter JR. Endophyte fungal Isolates from Podophyllum peltatum produce podophyllotoxin. J Nat Prod 2006;69:1121–4.
[89] Puri SC, Nazir A, Chawla R, Arora R, Riyaz-ul-Hasan S, Amna T, et al. The endophytic fungus Trametes hirsuta as a novel alternative source of podophyllotoxin and related tetralin lignans. J Biotech 2006;122:494–510.
[90] Amna T, Puri SC, Verma V, Sharma JP, Khajuria RK, Musarrat J, et al. Bioreactor studies on the endophytic fungus Entrophosphora infrequens for the production of an anticancer alkaloid camptothecin. Can J Microbiol 2006;52:189–96.
[91] Puri SC, Verma V, Amna T, Qazi GN, Spiteller M. An endophytic fungus from Nothapodytes foetida that produces camptothecin. J Nat Prod 2005;68:1717–19.
[92] Guo B, Li H, Zhang L. Isolation of a fungus producing vinblastine. J Yunnan Univ 1998;20:214–15.
[93] Zhang LQ, Guo B, Li H, Zeng S, Shao H, Gu S, et al. Preliminary study on the isolation of endophytic fungus of Catharanthus roseus and its fermentation to produce products of therapeutic value. Zhong Cao Yao (Chinese Tradit Herb Drugs) 2000;31:805–7.
[94] Yang X, Zhang L, Guo B, Guo S. Preliminary study of a vincristine-producing endophytic fungus isolated from leaves of Catharanthus roseus. Zhong Cao Yao (Chinese Tradit Herb Drugs) 2004;35:79–81.
[95] Chandra S. Endophytic fungi: novel sources of anticancer lead molecules. App Microbiol Biotechnol 2012;95:47–59.
[96] Kusari S, Hertweck C, Spiteller M. Chemical ecology of endophytic fungi: origins of secondary metabolites. Chem Biol 2012;19:792–8.
[97] Grum DS, Cook D, Baucom D, Mott IW, Gardner DR, Creamer R, et al. Production of the alkaloid swainsonine by a fungal endophyte in the host Swainsona canescens. J Nat Prod 2013;76:1984–8.
[98] Cook D, Gardner DR, Pfister JA. Swainsonine-containing plants and their relationship to endophytic fungi. J Agric Food Chem 2014;62:7326–34.
[99] Lopanik NB, Shields JA, Buchholz TJ, Rath CM, Hothersall J, Haygood MG, et al. *In vivo* and *in vitro* trans-acylation by bryP, the putative bryostatin pathway acyltransferase derived from an uncultured marine symbiont. Chem Biol 2008;15:1175–86.
[100] Zheng L, Li G, Wang X, Pan W, Li L, Hua L, et al. Nematicidal endophytic bacteria obtained from plants. Ann Microbiol 2008;58:569–72.
[101] Tiwari R, Kalra A, Darokar MP, Chandra M, Aggarwal N, Singh AK, et al. Endophytic bacteria from Ocimum sanctum and their yield enhancing capabilities. Curr Microbiol 2010;60:167–71.
[102] Lu C, Shen Y. A novel ansamycin, naphthomycin K from Streptomyces sp. J Antibiot 2007;60:649–53.
[103] Igarashi Y, Yanase S, Sugimoto K, Enomoto M, Miyanaga S, Trujillo ME, et al. Lupinacidin C, an inhibitor of tumor cell invasion from Micromonospora lupini. J Nat Prod 2011;74:862–5.
[104] Kupchan SM, Komoda Y, Court WA, Thomas GJ, Smith RM, Karim A, et al. Tumor inhibitors. LXXIII. Maytansine, a novel antileukemic ansa macrolide from Maytenus ovatus. J Am Chem Soc 1972;94:1354–6.
[105] Higashide E, Asai M, Ootsu K, Tanida S, Kozai Y, Hasegawa T, et al. Ansamitocin, a group of novel maytansinoid antibiotics with antitumour properties from Nocardia. Nature 1977;270:721–2.
[106] Wings S, Müller H, Berg G, Lamshöft M, Leistner E. A study of the bacterial community in the root system of the maytansine containing plant Putterlickia verrucosa. Phytochemistry 2013;91:158–64.
[107] Bandara WMMS, Seneviratne G, Kulasooriya SA. Interactions among endophytic bacteria and fungi; effects and potential. J Biosci 2006;31:645–50.

[108] Nützmann H-W, Reyes-Dominguez Y, Scherlach K, Schroeckh V, Horn F, et al. Bacteria-induced natural product formation in the fungus Aspergillus nidulans requires Saga/Ada-mediated histone acetylation. Proc Nat Acad Sci U S A 2011;108:14282–7.
[109] Partida-Martinez LP, Hertweck C. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. Nature 2005;437:884–8.
[110] Partida-Martinez LP, Monajembashi S, Greulich K-O, Hertweck C. Endosymbiont-dependent host reproduction maintains bacterial-fungal mutualism. Curr Biol 2007;17:773–7.
[111] Scherlach K, Busch B, Lackner G, Paszkowski U, Hertweck C. Symbiotic cooperation in the biosynthesis of a phytotoxin. Angew Chem Int Ed 2012;124:9753–6.
[112] Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core Arabidopsis thaliana root microbiome. Nature 2012;488:86–90.
[113] Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, et al. Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 2012;488:91–5.
[114] Feling RH, Buchanan GO, Mincer TJ, Kauffman CA, Jensen PR, Fenical W. Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus Salinospora. Angew Chem Int Ed 2003;42:355–7.
[115] Udwary DW, Zeigler L, Asolkar RN, Singan V, Lapidus A, Fenical W, et al. Genome sequencing reveals complex secondary metabolome in the marine actinomycete Salinispora tropica. Proc Natl Acad Sci U S A 2007;104:10376–81.
[116] Lane AL, Moore BS. A sea of biosynthesis: marine natural products meet the molecular age. Nat Prod Rep 2011;28:411–28.
[117] Eustáquio AS, O'Hagan D, Moore BS. Engineering fluorometabolite production: fluorinase expression in Salinispora tropica yields fluorosalinosporamide. J Nat Prod 2010;73:378–82.
[118] Tsukimoto M, Nagaoka M, Shishido Y, Fujimoto J, Nishisaka F, Matsumoto S, et al. Bacterial production of the tunicate-derived antitumor cyclic depsipeptide didemnin B. J Nat Prod 2011;74:2329–31.
[119] Shi BH, Arunpairojana V, Palakawong S, Yokota A. Tistrella mobilis gen. nov., sp. nov., a novel polyhydroxyalkanoate-producing bacterium belonging to alpha-Proteobacteria. J Gen Appl Microbiol 2002;48:335–43.
[120] Zhang D-C, Liu H-C, Zhou Y-G, Shinner F, Margesin R. Tistrella bauzanensis sp. nov., isolated from soil. Int J Sys Evol Microbiol 2011;61:2227–30.
[121] Xu Y, Kersten RD, Nam S-J, Lu L, Al-Suwailem AM, Zheng H, et al. Bacterial biosynthesis and maturation of the didemnin anti-cancer agents. J Am Chem Soc 2012;134:8625–32.
[122] Hoehler TM, Jorgensen BB. Microbial life under extreme energy limitation. Nat Rev Microbiol 2013;11:83–93.
[123] Piel J. Metabolites from symbiotic bacteria. Nat Prod Rep 2004;21:519–38.
[124] Piel J, Hofer I, Hui D. Evidence for a symbiosis island involved in horizontal acquisition of pederin biosynthetic capabilities by the bacterial symbiont of Paederus fuscipes beetles. J Bacteriol 2004;186:1280–6.
[125] Piel J, Butzke D, Fusetani N, Hui D, Platzer M, Wen G, et al. Exploring the chemistry of uncultivated bacterial symbionts: antitumor polyketides of the pederin family. J Nat Prod 2005;68:472–9.
[126] Giddings L-A, Newman DJ. Microbial natural products: molecular blueprints for antitumor drugs. J Ind Microbiol Biotechnol 2013;40:1181–210.
[127] Radjasa OK, Vaske YM, Navarro G, Vervoort HC, Tenney K, Linington RG, et al. Highlights of marine invertebrate-derived biosynthetic products: their biomedical potential and possible production by microbial associants. Bioorg Med Chem 2011;19:6658–74.
[128] Still PC, Johnson TA, Theodore CM, Loveridge ST, Crews P. Scrutinizing the scaffolds of marine biosynthetics from different source organisms: gram-negative cultured bacterial products enter center stage. J Nat Prod 2014;77:690–702.
[129] Wilson MC, Mori T, Rückert C, Uria AR, Helf MJ, Takada K, et al. An environmental bacterial taxon with a large and distinct metabolic repertoire. Nature 2014;506:58–62.
[130] Wakimoto T, Egami Y, Nakashima Y, Mori T, Awakawa T, Ito T, et al. Calyculin biogenesis from a pyrophosphate protoxin produced by a sponge symbiont. Nat Chem Biol 2014;10:648–55.
[131] Rocha-Martin J, Harrington C, Dobson ADW, O'Gara F. Emerging strategies and integrated systems microbiology technologies for biodiscovery of marine bioactive compounds. Mar Drugs 2014;12:3516–59.
[132] Thomas MG, Bixby KA, Shen B. Combinatorial biosynthesis of anticancer natural molecules. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. Boca Raton, FL: Taylor & Francis; 2005. p. 519–52.
[133] Van Lanen SG, Shen B. Combinatorial biosynthesis of anticancer natural products. In: 2nd ed. Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products, 2012. Boca Raton, FL: Taylor & Francis; 2012. p. 671–98.
[134] Bachmann BO, Van Lanen SG, Baltz RH. Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? J Ind Microbiol Biotechnol 2014;41:175–84.
[135] Mahoney KPP, Smith DRM, Bogosyan EJA, Goss RJM. Access to high value natural and unnatural products through hyphenating chemical synthesis and biosynthesis. Synthesis 2014;46:2122–32.
[136] McCloud TG. High throughput extraction of plant, marine and fungal specimens for preservation of biologically active molecules. Molecules 2010;15:4526–63.
[137] Beutler JA, McKee TC, Fuller RW, Tischler M, Cardellina II JH, McCloud TG, et al. Frequent occurrence of HIV-inhibitory sulfated polysaccharides in marine invertebrates. Antiviral Chem Chemother 1993;3:167–72.
[138] Cardellina II JH, Munro MHG, Fuller RW, Manfredi KP, McKee TC, Tischler M, et al. A chemical-screening strategy for the dereplication and prioritization of HIV-inhibitory aqueous natural-products extracts. J Nat Prod 1993;56:1123–9.
[139] Boyd MR, Paull KD. Some practical considerations and applications of the National Cancer Institute in vitro anticancer drug discovery screen. Drug Dev Res 1995;34:91–109.
[140] Developmental Theraputics Program. <http://dtp.nci.nih.gov/>; 2014.
[141] Johnson RK, Bartus HF, Hofmann GA, Bartus JO, Mong S-M, Faucette LF, et al. Discovery of new DNA-reactive drugs. In: Hanka LJ, Kondo T, White RJ, editors. In vitro and in vivo models for detection of new antitumor drugs, 1986. Tokyo: University of Tokyo Press; 1986. p. 15–26.

[142] Hassig CA, Zeng F-Y, Kung P, Kiankarimi M, Kim S, Diaz PW, et al. Ultra high throughput screening of natural product extracts to identify pro-apoptotic inhibitors of Bcl-2 family proteins. J Biomol Screening 2014;19:1201–11.
[143] Moffat JG, Rudolph J, Bailey D. Phenotypic screening in cancer drug discovery - past, present and future. Nat Rev Drug Discov 2014;13:588–602.
[144] Eder J, Sedrani R, Wiesmann C. The discovery of first-in-class drugs: origins and evolution. Nat Rev Drug Discov 2014;13:577–87.
[145] Swinney DC, Anthony J. How were new medicines discovered? Nat Rev Drug Discov 2011;10:507–19.
[146] Lee AY, St Onge RP, Proctor MJ, Wallace IM, Nile AH, Spagnuolo PA, et al. Mapping the cellular response to small molecules using chemogenomic fitness signatures. Science 2014;344:208–11.
[147] Mohr SE, Smith JA, Shamu CE, Neumuller RA, Perrimon N. RNAi screening comes of age: improved techniques and complementary approaches. Nat Rev Mol Cell Biol 2014;15:591–600.
[148] Marston A, Hostettmann K. Modern separation methods. Nat Prod Rep 1991;8:391–413.
[149] Hostettmann K, Wolfender J-L, Rodriguez S. Rapid detection and subsequent isolation of bioactive constituents of crude plant extracts. Planta Med 1997;63:2–10.
[150] Jaki B, Franzblau S, Pauli GF. An N.M.R. method towards the routine chiral determination of natural products. Phytochem Anal 2004;15:213–19.
[151] Gronquist M, Meinwald J, Eisner T, Schroeder FC. Exploring uncharted terrain in Nature's structure space using capillary NMR spectroscopy: 13 steroids from 50 fireflies. J Am Chem Soc 2005;127:10810–11.
[152] Wolfender JL, Queiroz EF, Hostettmann K. Phytochemistry in the microgram domain - a LC-NMR perspective. Mag Res Chem 2005;43:697–709.
[153] Konishi Y, Kiyota T, Draghici C, Gao JM, Yeboah F, Acoca S, et al. Molecular formula analysis by an MS/MS/MS technique to expedite dereplication of natural products. Anal Chem 2007;79:1187–97.
[154] Lambert M, Wolfender JL, Staerk D, Christensen B, Hostettmann K, Jaroszewski JW. Identification of natural products using HPLC-SPE combined with CapNMR. Anal Chem 2007;79:727–35.
[155] Queiroz EF, Wolfender JL, Hostettmann K. Modern approaches in the search for new lead antiparasitic compounds from higher plants. Curr Drug Targ 2009;10:202–11.
[156] Friesen JB, Pauli GF. G.U.E.S.S. - A generally useful estimate of solvent systems for CCC. J Liq Chromatog RelTechnol 2005;28:2777–806.
[157] Simmler C, Napolitano JG, McAlpine JB, Chen SN, Pauli GF. Universal quantitative NMR analysis of complex natural samples. Curr Opin Biotechnol 2014;25:51–9.
[158] El-Elimat T, Figueroa M, Ehrmann BM, Cech NB, Pearce CJ, Oberlies NH, et al. High-resolution MS, MS/MS, and UV database of fungal secondary metabolitees as a dereplicatoin protocol for bioactive natural products. J Nat Prod 2013;76:1709–16.
[159] Bouslimani A, Sanchez LM, Garg N, Dorrenstein PC. Mass spectrometry of natural products: current, emerging and future technologies. Nat Prod Rep 2014;31:718–29.
[160] Mohimani H, Liu W-T, Kersten RD, Moore BS, Dorrestein PC, Pevzner PA. NRPquest: coupling mass spectrometry and genome mining for nonribosomal peptide discovery. J Nat Prod 2014;77:1902–9.
[161] Cragg GM, Schepartz SA, Suffness M, Grever MR. The taxol supply crisis. New NCI policies for handling the large-scale production of novel natural product anticancer and anti-HIV agents. J Nat Prod 1993;56:1657–68.
[162] Fujita Y. Industrial production of shikonin and berberine. In: Bock G, Marsh J, editors. Applications of plant cell and tissue culture. Chichester: Wiley; 1988. p. 157–74.
[163] Fowler MW, Cresswell RC, Stafford AM. An economic and technical assessment of the use of plant cell cultures for natural product synthesis on an industrial scale. In: Chadwick DJ, Marsh J, editors. Bioactive compounds from plants. Chichester: Wiley; 1990. p. 228–38.
[164] Kingston DGI. Taxol and its analogs. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed. Boca Raton, FL: CRC Press/Taylor and Francis; 2012. p. 123–75.
[165] Hayden AL, Giacomelli GA, Hoffmann JJ, Yokelson TN. Aeroponics: an alternative production system for high-value root crops. Acta Horticulturae 2004;629:207–13.
[166] Hayden AL. Aeroponic and hydroponic systems for medicinal herb, rhizome, and root crops. HortScience 2006;41:536–8.
[167] Xu YM, Gao S, Bunting DP, Gunatilaka AAL. Unusual withanolides from aeroponically grown Withania somnifera. Phytochemistry 2011;72:518–22.
[168] Mickel SJ, Niederer D, Daeffler R, Osmani A, Kuesters E, Schmid E, et al. Large-scale synthesis of the anti-cancer marine natural product (+)-discodermolide. Part 5: linkage of fragments C1-6 and C7-24 and finale. Org Process Res Dev 2004;8:122–30.
[169] Yu MJ, Kishi Y, Littlefield BA. Discovery of E7389: a fully synthetic macrocyclic ketone analog of halichondrin B. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed Boca Raton, FL: CRC Press/Taylor and Francis; 2012. p. 317–45.
[170] Cragg GM, Grothaus PG, Newman DJ. Impact of natural products on developing new anti-cancer agents. Chem Rev 2009;109:3012–43.
[171] Liu J, Hu Y, Waller DL, Wang J, Liu Q. Natural products as kinase inhibitors. Nat Prod Rep 2012;29:392–403.
[172] Coxon GD, Cooper CB, Gillespie SH, McHugh TD. Strategies and challenges involved in the discovery of new chemical entities during early-stage tuberculosis drug discovery. J Infect Dis 2012;205:S258–64.
[173] Gross F, Lazar J, Orth H. Inhibition of the renin-angiotensinogen reaction by pepstatin. Science 1972;175:656.
[174] Darke PL, Leu C-T, Davis LJ, Heimbach JC, Diehl RE, Hill WS, et al. Human immunodeficiency virus protease. Bacterial expression and characterization of the purified aspartic protease. J Biol Chem 1989;264:2307–12.
[175] Evans BE, Rittle KE, Homnick CF, Springer JP, Hirshfield J, Veber DF. A stereocontrolled synthesis of hydroxyethylene dipeptide isosteres using novel, chiral aminoalkyl epoxides and .gamma.-(aminoalkyl)-.gamma.-lactones. J Org Chem 1985;50:4615–25.
[176] DeSolms SJ, Giuliani EA, Guare JP, Vacca JP, Sanders WM, Graham SL, et al. Design and synthesis of HIV protease inhibitors. Variations of the carboxyterminus of the HIV protease inhibitor L-682,679. J Med Chem 1991;34:2852–7.
[177] Ripka AS, Rich DH. Peptidomimetic design. Curr Opin Chem Biol 1998;4:439–52.

[178] Schechter I, Berger A. On the size of the active site of proteases. I. Papain. Biochem Biophys Res Commun 1967;27:157–62.
[179] Rahuel J, Rasetti V, Maibaum J, Rueger H, Goschke R, Cohen N-C, et al. Structure-based drug design: the discovery of novel non-peptide orally active inhibitors of human renin. Chem Biol 2000;7:493–504.
[180] Wood JM, Maibaum J, Rahuel J, Grutter MG, Cohen N-C, Rasetti V, et al. Structure-based design of aliskiren, a novel orally effective renin inhibitor. Biochem Biophys Res Commun 2003;308:698–705.
[181] Stassen JA, Li Y, Richart T. Oral renin inhibitors. Lancet 2006;368:1449–56.
[182] Webb RL, Schiering N, Sedrani R, Maibaum J. Direct renin inhibiotrs as a new therapy for hypertension. J Med Chem 2010;53:7490–520.
[183] Politi A, Leonis G, Tzoupis H, Ntountaniotis D, Papadopoulos MG, Grdadolnik SG, et al. Conformational properties and energetic analysis of aliskiren. Mol Inf 2011;30:973–85.
[184] Tzoupis H, Leonis G, Megariotis G, Supuran CT, Mavromoustakos T, Papadopoulos MG. Dual inhibitors for aspartic proteases HIV-1 PR and Renin: advancements in AIDS-hypertension-diabetes via molecular dynamics, inhibition assays, and binding free energy calculations. J Med Chem 2012;55:5784–96.
[185] Nam G, Ko SY. A total synthesis of aliskiren. Helv Chim Acta 2012;95:1937–45.
[186] Ehara T, Irie O, Kosaka T, Kanazawa T, Breitenstein W, Grosche P, et al. Structure-based design of substituted piperidines as a new class of highly efficacious oral direct renin inhibitors. ACS Med Chem Lett 2014;5:787–92.
[187] Palomo JM. Solid-phase peptide synthesis: an overview focused on the preparation of biologically relevant peptides. RSC Adv 2014;4:32658–72.
[188] Momi S, Impagnatiello F, Guzzetta M, Caracchini R, Guglielmini G, Olivieri R, et al. NCX 6560: a nitric oxide-releasing derivative of atorvastatin, inhibits cholesterol biosynthesis and shows anti-inflammatory and anti-thrombotic properties. Eur J Pharmacol 2007;570:115–24.
[189] Baetta R, Granata A, Miglietta D, Oliva F, Arnaboldi L, Bonomo A, et al. Nitric oxide-donating atorvastatin attenuates neutrophil recruitment during vascular inflammation independent of changes in plasma cholesterol. Cardiovasc Drugs Ther 2013;27:211–19.
[190] Oslaj M, Cluzeau J, Orkic D, Kopitar G, Mrak P, Casar Z. A highly productive, whole cell DERA chemoenzymatic process for production of key lactonized side-chain intermediates in statin synthesis. PLoS one 2013;8:e62250.
[191] Vajdic T, Oslaj M, Kopitar G, Mrak P. Engineered, highly productive biosynthesis of artificial, lactonized statin side-chain building blocks: the hidden potential of Escherichia coli unleashed. Metab Eng 2014;24:160–72.
[192] McIntyre JA, Castaner J. ISA-247 - Calcineurin inhibitor treatment of renal transplant rejection treatment of psoriasis. Drugs Future 2004;29:680–6.
[193] Maeng JH, Yang ZC, Manning DD, Masih L, Cao YY, Pattamana KG, et al. Organozirconium chemistry on cyclosporin: a novel process for the highly stereoselective synthesis of (E)-ISA247 (Voclosporin) and close analogues. Synthesis 2012;44:63–8.
[194] Adachi K, Kohara T, Nakao N, Arita M, Chiba K, Mishina T, et al. Design, synthesis, and structure-activity relationships of 2-substituted-2-amino-1,3-propanediols - discovery of a novel immunosuppressant, FTY720. Bioorg Med Chem Lett 1995;5:853–6.
[195] Strader CR, Pearce CJ, Oberlies NH. Fingolimod (FTY720): a recently approved multiple sclerosis drug based on a fungal secondary metabolite. J Nat Prod 2011;74:900–7.
[196] Zhang L, Wang HD, Ji XJ, Cong ZX, Zhu JH, Zhou Y. FTY720 for cancer therapy (Review). Oncol Rep 2013;30:2571–8.
[197] Takasugi N, Sasaki T, Ebinuma I, Osawa S, Isshiki H, Takeo K, et al. FTY720/Fingolimod, a sphingosine analogue, reduces amyloid-beta production in neurons.
[198] Martin R, Mireia S. Sphingosine-1 phosphate and central nervous system. In: Oldstone MBA, Hugh R, editors. Sphingosine-1-phosphate signaling in immunology and infectious diseases, 2014. Switzerland: Springer International Publishing; 2014. p. 149–70.
[199] Zhou R, Zhang F, He PL, Zhou WL, Wu QL, Xu JY, et al. (5R)-5-hydroxytriptolide (LLDT-8), a novel triptolide analog mediates immunosuppressive effects in vitro and in vivo. Int Immunopharmacol 2005;5:1895–903.
[200] Tang W, Zuo JP. Immunosuppressant discovery from Tripterygium wilfordii Hook f: the novel triptolide analog (5R)-5-hydroxytriptolide (LLDT-8). Acta Pharmacol Sin 2012;33:1112–18.
[201] Bohuslavizki KH, Hinck-Kneip C, Kniep A, Koppenhofer E, Reimers A. Reduction of MS-related scotoma by new class of potassium channel blockers from Ruta graveolens. Neuroopthalmol 1993;13:191–8.
[202] Bohuslavizki KH, Hansel W, Kniep A, Koppenhofer E, Niemoller K, Sanmann K. Mode of action of psoralens, benzofurans, acridones and coumarins on the ionic currents in intact myelinated nerve firbres and its significance in demyelating diseases. Gen Physiol Biophys 1994;13:309–28.
[203] Schmitz A, Sankaranarayanan A, Azam P, Schmidt-Lassen K, Homerick D, Hänsel W, et al. Design of PAP-1, a selective small molecule Kv1.3 blocker, for the suppression of effector memory T cells in autoimmune diseases. Mol Pharmacol 2005;68:1254–70.
[204] Beeton C, Wulff H, Standifer NE, Azam P, Mullen KM, Pennington MW, et al. Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. Proc Natl Acad Sci U S A 2006;103:17414–19.
[205] Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nat Rev Drug Discov 2007;6:29–40.
[206] Revill P, Serradell N, Bolos J, Tiacumicin B. Macrolide antibiotic, treatment of C difficile-associated diarrhea. Drugs Future 2006;31:494–7.
[207] Leadbetter MR, Adams SM, Bazzini B, Fatheree PR, Karr DE, Krause KM, et al. Hydrophobic vancomycin derivatives with improved ADME properties: discovery of telavancin (TD-6424). J Antibiot 2004;57:326–36.
[208] Malabarba A, Ciabatti R, Scotti R, Goldstein BP, Ferrari P, Kurz M, et al. New semisynthetic glycopeptides MDL 63,246 and MDL 63,042, and other amide derivatives of antibiotic A-40,926 active against highly glycopeptide-resistant VanA enterococci. J Antibiot 1995;48:869–83.
[209] Economou NJ, Nahoum V, Weeks SD, Grasty KC, Zentner IJ, Townsend TM, et al. A carrier protein strategy yields the structure of dalbavancin. J Am Chem Soc 2012;134:4637–45.

[210] Cooper RD, Snyder NJ, Zweifel MJ, Staszak MA, Wilkie SC, Nicas TI, et al. Reductive alkylation of glycopeptide antibiotics: synthesis and antibacterial activity. J Antibiot 1996;49:575–81.
[211] Liapikou A, Torres A. Emerging drugs on methicillin-resistant Staphylococcus aureus. Exp Opin Emerg Drugs 2013;18:291–305.
[212] Chou T-C, Zhang X, Zhong Z-Y, Li Y, Feng L, Eng S, et al. Therapeutic effect against human xenograft tumors in nude mice by the third generation microtubule stabilizing epothilones. Proc Natl Acad Sci U S A 2008;105:13157–62.
[213] Wilson RM, Danishefsky SJ. On the reach of chemical synthesis: creation of a mini-pipeline from an academic laboratory. Angew Chem Int Ed 2010;49:6032–56.
[214] Altmann K-H, Schinzer D. Chemistry and biology of epothilones. In: Hanessian S, editor. Natural products in medicinal chemistry. Weinheim, Germany: Wiley-VCH Verlag; 2014. p. 83–125.
[215] Gao J, Hamann MT. Chemistry and biology of kahalalides. Chem Rev 2011;111:3208–35.
[216] Rath CM, Janto B, Earl J, Ahmed A, Hu FZ, Hiller L, et al. Meta-omic characterization of the marine Invertebrate microbial consortium that produces the chemotherapeutic natural product ET-743. ACS Chem Biol 2011;6:1244–56.
[217] Wani MC, Taylor HL, Wall ME, Coggon P, McPhail AT. Plant antitumor agents. VI. The Isolation and structure of Taxol, a novel antileukemic and antitumor agent from Taxus brevifolia. J Am Chem Soc 1971;93:2325–7.
[218] Schiff PB, Fant J, Horwitz SB. Promotion of microtubule assembly in vitro by taxol. Nature 1979;277:665–7.
[219] Suffness M. Taxol: from discovery to therapeutic use Academic Press. In: Bristol JA, editor. Ann Rep Med Chem. San Diego: Academic Press; 1993. p. 305–14.
[220] Jacoby M. The top pharmaceuticals that changed the world. C&E News 2005;83:(25), 120.
[221] Wang X, Itokawa H, Lee K-H. Structure-activity relationships of taxoids. In: Itokawa H, Lee K-H, editors. Taxus: the genus taxus. London: Taylor & Francis; 2003. p. 298–86.
[222] Cragg GM, Newman DJ. A tale of two tumor targets: topoisomerase I and tubulin. The Wall and Wani contribution to cancer chemotherapy. J Nat Prod 2004;67:232–44.
[223] Dubois J. Recent progress in the development of docetaxel and paclitaxel analogues. Expert Opin Ther Patents 2006;16:1481–96.
[224] Kingston DGI, Newman DJ. Taxoids: cancer-fighting compounds from nature. Curr Opin Drug Disc Devel 2007;10:130–44.
[225] Ojima I, Kamath A, Seitz JD. Taxol, taxoids, and related taxanes. In: Hanessian S, editor. Natural products in medicinal chemistry. Weinheim, Germany: Wiley-VCH Verlag; 2014. p. 127–80.
[226] Guenard D, Gueritte-Voegelein F, Potier P. Taxol and taxotere: discovery, chemistry, and structure-activity relationships. Acc Chem Res 1993;26:160–7.
[227] Green MR, Manikhas GM, Orlov S, Afanasyev B, Makhson AM, Bhar P, et al. Abraxane, a novel Cremophor-free, albumin-bound particle form of paclitaxel for the treatment of advanced non-small-cell lung cancer. Ann Oncol 2006;17:1263–8.
[228] Galsky MD, Dritselis A, Kirkpatrick P, Oh WK. Cabazitaxel. Nat Rev Drug Disc 2010;9:677–8.
[229] Bouchard H, Semiond D, Risse M-L, Vrignaud P. Novel taxanes: cabazitaxel case study. In: Fischer J, Ganellin CR, Rotella DP, editors. Analogue-based drug discovery III. Weinheim, Germany: Wiley-VCH Verlag; 2013. p. 319–41.
[230] Sternberg CN, Skoneczna IA, Castellano D, Theodore C, Blais N, Voog E, et al. Larotaxel with Cisplatin in the first-line treatment of locally advanced/metastatic urothelial tract or bladder cancer: a randomized, active-controlled, phase III trial (CILAB). Oncology 2013;85:208–15.
[231] Beer M, Lenaz L, Amadori D. Phase II study of ortataxel in taxane-resistant breast cancer. J Clin Oncol 2008;26 Abstract 1066.
[232] Kato K, Chin K, Yoshikawa T, Yamaguchi K, Tsuji Y, Esaki T, et al. Phase II study of NK105, a paclitaxel-incorporating micellar nanoparticle, for previously treated advanced or recurrent gastric cancer. Invest New Drugs 2012;30:1621–7.
[233] Paz-Ares L, Ross H, O'Brien M, Riviere A, Gatzemeier U, Von Pawel J, et al. Phase III trial comparing paclitaxel poliglumex vs docetaxel in the second-line treatment of non-small-cell lung cancer. Br J Cancer 2008;98:1608–13.
[234] Hartwell JL, Schrecker AW. Chemistry of podophyllum. In: Zechmeister L, editor. Fortschritte chemie der organische naturstoffe [Progress in the chemistry of organic natural products], 1958. Vienna: Springer-Verlag; 1958. p. 83–166. [in German].
[235] Stahelin HF, von Wartburg A. The chemical and biological route from podophyllotoxin glucoside to etoposide: ninth cain memorial award lecture. Cancer Res 1991;51:5–15.
[236] Hande KR. Etoposide: four decades of development of a topoisomerase II inhibitor. Eur J Cancer 1998;34:1514–21.
[237] Lee K-H, Xiao Z. Podophyllotoxin and analogs. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed. Boca Raton, FL: CRC Press; 2012. p. 95–122.
[238] Noble RL, Beer CT, Cutts JH. Role of chance observation in chemotherapy: vinca rosea. Ann NY Acad Sci 1958;76:882–94.
[239] Dancey J, Steward WP. The role of vindesine in oncology: recommendations after 10 years' experience. Anti-Cancer Drugs 1995;6:625–36.
[240] Bellmunt J, Fougeray R, Rosenberg JE, von der Maase H, Schutz FA, Salhi Y, et al. Long-term survival results of a randomized phase III trial of vinflunine plus best supportive care versus best supportive care alone in advanced urothelial carcinoma patients after failure of platinum-based chemotherapy. Ann Oncol 2013;24:1466–72.
[241] Gueritte F, Fahy J. The vinca alkaloids. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed. Boca Raton, FL: CRC Press; 2012. p. 177–98.
[242] Wall ME. Camptothecin and taxol: discovery to clinic. Med Res Rev 1998;18:299–314.
[243] Rahier NJ, Thomas CJ, Hecht SM. Camptothecin and its analogs. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed. Boca Raton, FL: CRC Press; 2012. p. 5–25.
[244] Giannini G. Camptothecin and Analogs. In: Hanessian S, editor. Natural products in medicinal chemistry. Weinheim, Germany: Wiley-VCH Verlag; 2014. p. 181–224.
[245] Gras J. Ingenol mebutate: a new option for actinic keratosis treatment. Drugs Today 2013;49:15–22.
[246] Itokawa H, Hitotsuyanagi Y, Lee K-H. Homoharringtonine and related compounds. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed. Boca Raton, FL: CRC Press; 2012. p. 65–93.

[247] Gandhi V, Plunkett W, Cortes JE. Omacetaxine: a protein translation inhibitor for treatment of chronic myelogenous leukemia. Clin Cancer Res 2014;20:1735–40.
[248] Pinney KG, Pettit GR, Trawick ML, Jelinek C, Chaplin DJ. The discovery and development of the combretastatins. In: Cragg GM, Kingston DGI, Newman DJ, editors. Anticancer agents from natural products. 2nd ed. Boca Raton, FL: CRC Press; 2012. p. 27–63.
[249] Nagaiah G, Remick SC. Combretastatin A4 phosphate: a novel vascular disrupting agent. Future Oncol 2010;6:1219–28.
[250] Hirata Y, Uemura D. Halichondrins - antitumor polyether macrolides from a marine sponge. Pure Appl Chem 1986;58:701–10.
[251] Pettit GR, Herald CL, Boyd MR, Leet JE, Dufresne C, Doubek DL, et al. Antineoplastic agents. 219. Isolation and structure of the cell growth inhibitory constituents from the western Pacific marine sponge Axinella sp. J Med Chem 1991;34:3339–40.
[252] Choi H-W, Demeke D, Kang F-A, Kishi Y, Nakajima K, Nowak P, et al. Synthetic studies on the marine natural product halichondrins. Pure Appl Chem 2003;75:1–17.
[253] Zheng W, Seletsky BM, Palme MH, Lydon PJ, Singer LA, Chase CE, et al. Macrocyclic ketone analogues of halichondrin B. Bioorg Med Chem Lett 2004;14:5551–4.
[254] Yu MJ, Zheng W, Seletsky BM. From micrograms to grams: scale-up synthesis of eribulin mesylate. Nat Prod Rep 2013;30:1158–64.
[255] Chase CE, Fang FG, Lewis BM, Wilke GD, Schnaderbeck MJ, Zhu X. Process development of halaven®: synthesis of the C1-C13 fragment from d -(-)-gulono-1,4-lactone. Synlett 2013;24:323–6.
[256] Austad BC, Benayoud F, Calkins TL, Campagnaa S, Chase CE, Choi H-W, et al. Process development of halaven®: synthesis of the C14-C35 fragment via iterative Nozaki-Hiyama-Kishi reaction-Williamson ether cyclization. Synlett 2013;24:327–32.
[257] Austad BC, Calkins TL, Chase CE, Fang FG, Horstmann TE, Hu Y, et al. Commercial manufacture of halaven®: chemoselective transformations en route to structurally complex macrocyclic ketones. Synlett 2013;24:333–7.
[258] Negi AS, Gupta A, Hamid AA. Combating malaria with plant molecules: a brief update. Curr Med Chem 2014;21:458–500.
[259] Wongsrichanalai C, Pickard AL, Wernsdorfer WH, Meshnick SR. Epidemiology of drug-resistant malaria. Lancet Infect Dis 2002;2:209–18.
[260] Klayman DL. Qinghaosu (artemisinin): an antimalarial drug from China. Science 1985;228:1049–55.
[261] Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, et al. Artemisinin resistance in Plasmodium falciparum malaria. N Engl J Med 2009;361:455–67.
[262] Shandilya A, Chacko S, Jayaram B, Ghosh I. A plausible mechanism for the antimalarial activity of artemisinin: a computational approach. Sci Rep 2013;3:2513.
[263] Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, et al. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 2013;496:528–32.
[264] Chen H-J, Han W-B, Hao H-D, Wu Y. A facile and scalable synthesis of qinghaosu (artemisinin). Tetrahedron 2013;69:1112–14.
[265] Turconi J, Griolet F, Guevel R, Oddon G, Villa R, Geatti A, et al. Semisynthetic artemisinin, the chemical path to industrial production. Org Proc Res Dev 2014;18:417–22.
[266] Krishna S, Bustamante L, Haynes RK, Staines HM. Artemisinins: their growing importance in medicine. Trends Pharmacol Sci 2008;29:520–7.
[267] Mott BT, Tripathi A, Siegler MA, Moore CD, Sullivan DJ, Posner GH. Synthesis and antimalarial efficacy of two-carbon-linked, artemisinin-derived trioxane dimers in combination with known antimalarial drugs. J Med Chem 2013;56:2630–41.
[268] Wright CW. Recent developments in research on terrestrial plants used for the treatment of malaria. Nat Prod Rep 2010;27:961–8.
[269] Nogueira CR, Lopes LMX. Antiplasmodial natural products. Molecules 2011;16:2146–90.
[270] Kingston DGI, Newman DJ. Natural products as anticancer agents. In: Civjan N, editor. Natural products in chemical biology. New York: Wiley; 2012. p. 325–49.
[271] Rohena CC, Mooberry SL. Recent progress with microtubule stabilizers: new compounds, binding modes, and cellular activities. Nat Prod Rep 2014;31:335–55.
[272] Basmadjian C, Zhao Q, Bentouhami E, Djehal A, Nebigil CG, Johnson RA, et al. Cancer wars: natural products strike back. Front Chem 2014;2:1–18.
[273] Nicolaou KC, Roschangar F, Vourloumis D. Chemical biology of epothilones. Angew Chem Int Ed 1998;37:2014–45.
[274] Evans BE, Rittle KE, Bock MG, DiPardo RM, Fredinger RM, Whitter WL, et al. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. J Med Chem 1988;2235–46.
[275] Nicolaou KC, Pfefferkorn JA, Roecker AJ, Cao GQ, Barluenga S, Mitchell HJ. Natural product-like combinatorial libraries based on privileged structures. 1. General principles and solid-phase synthesis of benzopyrans. J Am Chem Soc 2000;122:9939–53.
[276] Nicolaou KC, Pfefferkorn JA, Mitchell HJ, Roecker AJ, Barluenga S, Cao GQ, et al. Natural product-like combinatorial libraries based on privileged structures. 2. Construction of a 10,000-membered benzopyran library by directed split-and-pool chemistry using NanoKans and optical encoding. J Am Chem Soc 2000;122:9954–67.
[277] Nicolaou KC, Pfefferkorn JA, Barluenga S, Mitchell HJ, Roecker AJ, Cao G-Q. Natural product-like combinatorial libraries based on privileged structures. 3. The "libraries from libraries" principle for diversity enhancement of benzopyran libraries. J Am Chem Soc 2000;122:9968–76.
[278] Nicolaou KC, Cao GQ, Pfefferkorn JA. Selenium-based solid-phase synthesis of benzopyrans II: applications to combinatorial synthesis of medicinally relevant small organic molecules. Angew Chem Int Ed 2000;39:739–43.
[279] Breinbauer R, Vetter IR, Waldmann H. From protein domains to drug candidates: natural products as guiding principles in the design and synthesis of compound libraries. Angew Chem Int Ed 2002;41:2878–90.
[280] Koch MA, Waldmann H. Protein structure similarity clustering and natural product structure as guiding principles in drug discovery. Drug Discov Today 2005;10:471–83.
[281] Van Hattum H, Waldmann H. Biology-oriented synthesis: harnessing the power of evolution. J Am Chem Soc 2014;136:11853–9.
[282] Wetzel S, Bon RS, Kumar K, Waldmann H. Biology-oriented synthesis. Angew Chem Int Ed 2011;50:10800–26.
[283] McArdle BM, Campitelli MR, Quinn RJ. A common protein fold topology shared by flavonoid biosynthetic enzymes and therapeutic targets. J Nat Prod 2006;69:14–17.

[284] Newman DJ. Natural products as leads to potential drugs: an old process or the new hope for drug discovery? J Med Chem 2008;51:2589–99.
[285] Burke MD, Berger EM, Schreiber SL. Generating diverse skeletons of small molecules combinatorially. Science 2003;302:613–18.
[286] Burke MD, Schreiber SL. A planning strategy for diversity-oriented synthesis. Angew Chem Int Ed 2004;43:46–58.
[287] Schreiber SL. Molecular diversity by design. Nature 2009;457:153–4.
[288] Osbourn A, Goss RJM, Carter GT, editors. Natural products: discourse, diversity, and design. Oxford, UK: Wiley Blackwell; 2014.
[289] Graziani EI. ILS-920: a rapamycin analog for ischemic stroke. In: Osbourn A, Goss RJM, Carter GT, editors. Natural products: discourse, diversity, and design. Oxford, UK: Wiley Blackwell; 2014. p. 469–81.
[290] Gregory MA, Moss SJ, Wilkinson B. BC265: a nonquinone ansamycin HSP90 inhibitor developed using biosynthetic medicinal chemistry. In: Osbourn A, Goss RJM, Carter GT, editors. Natural products: discourse, diversity, and design, 2014. Oxford, UK: Wiley Blackwell; 2014. p. 483–96.
[291] Martin CJ, Gaisser S, Challis IR, Carletti I, Wilkinson B, Gregory M, et al. Molecular characterization of macbecin as an Hsp90 inhibitor. J Med Chem 2008;51:2853–7.
[292] Zhang M-Q, Gaisser S, Nur-E-Alam M, Sheehan LS, Vousden WA, Gaitatzis N, et al. Optimizing natural products by biosynthetic engineering: discovery of nonquinone Hsp90 inhibitors. J Med Chem 2008;51:5494–7.
[293] Wishart DS. Applications of metabolomics in drug discovery and development. Drugs R D 2008;9:307–32.
[294] Hou Y, Braun DR, Michel CR, Klassen JL, Adnani N, Wyche TP, et al. Microbial strain prioritization using metabolomics tools for the discovery of natural products. Anal Chem 2012;84:4277–83.
[295] Tawfike AF, Viegelmann C, Edrada-Ebel R. Metabolomics and dereplication strategies in natural products. Methods Mol Biol 2013;1055:227–44.
[296] Yuliana ND, Janhangir M, Verpoorte R, Choi YH. Metabolomics for the rapid dereplication of bioactive compounds from natural sources. Phytochem Rev 2013;12:293–304.
[297] Krug D, Müller R. Secondary metabolomics: the impact of mass spectrometry-based approaches on the discovery and characterization of microbial natural products. Nat Prod Rep 2014;31:768–83.
[298] Charlop-Powers Z, Milshteyn A, Brady SF. Metagenomic small molecule discovery methods. Curr Opin Microbiol 2014;19:70–5.
[299] Aldridge BA, Rhee KY. Microbial metabolomics: innovation, application, insight. Curr Opin Microbiol 2014;19:90–6.
[300] Donia MS, Cimermancic P, Schulze CJ, Wieland Brown LC, Martin J, Mitreva M, et al. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. Cell 2014;158:1402–14.
[301] Vence T. Mining microbes for promising therapies: warp Drive Bio has industrialized its genomics platform in its hunt for natural products. Gen Eng Biotech News 2013 2013;33:12–14.

CHAPTER

6

In Silico Screening: Hit Finding from Database Mining

Tiago Rodrigues and Gisbert Schneider

Swiss Federal Institute of Technology (ETH), Department of Chemistry and Applied Biosciences, Zürich, Switzerland

OUTLINE

Simplicity is the ultimate sophistication. ***Leonardo da Vinci***

I. INTRODUCTION

The computer-assisted design of new chemical entities (NCEs) is one of the most promising fields in the development of new therapeutics. The co-evolution of theoretical and analytical chemistry, molecular biology, and computer technology has paved the way for the understanding of molecular mechanisms of drug action. Drug design has traveled a long journey since Brown and Frazer coined the "Schrödinger equation of drug design" in 1868: $\Delta\phi = f(\Delta C)$, which describes the basic relationship between a biological response ϕ and a chemical structure C. Finding suitable functional relationships probably is the most important aspect of computer-assisted medicinal chemistry. As simple as this task may appear, it is far from trivial to find suitable representations of chemical structures, to name just one of the crucial steps in structure-activity relationship (SAR) modeling. Due to the highly complex nature of living systems, reductionist approaches are unavoidable and necessary. Hence, the "traditional" procedure in drug design, or better *ligand design*, has been the virtual reconstruction of a 1:1 ligand–receptor complex and the derivation and quantification of the underlying forces of molecular interaction. This concept has been coming up with a number of success stories, but never to an extent that has considerably accelerated drug development or led to profoundly reduced or even

replaced animal studies or clinical trials. On the contrary, this simplistic view of drug action has often resulted in setbacks in the development of new drugs. For example, one of the most common anticancer drugs, Sunitinib, had been designed as a perfectly matching ligand, but the free base turned out to be insoluble in water. Derivatives were biologically very active but of course did not fit the original design anymore. Eventually, pharmacologists found that the derivatives were active at more than a hundred different subtypes of the target receptor tyrosine kinases. One might say that the ultimate great success of this drug development project was a lucky punch. In fact, serendipity has long played an important role in drug discovery. Fleming's *Eureka* moment while studying *Staphylococci* that lead to the discovery of penicillin is the perfect example of the early pharmaceutical research paradigm. Decades later, the effort rationally to conduct drug discovery programs has increased manifold, mostly driven by Hansch and Fujita's seminal work on structure-property relationships, as a consequent extension of the basic function introduced by Brown and Frazier [1]. Guided by a plethora of disciplines, where innovative chemistry and pharmacology play a paramount role, the pharmaceutical industry is constantly seeking new technologies that may heavily impact how drug discovery is carried out [2]. In this chapter, we present a succinct overview on selected molecular informatics and virtual screening approaches that are currently under mainstream use by both the pharmaceutical industry and academia, and assorted applications in computer-assisted hit finding, hit-to-lead optimization, *de novo* design, and computational target prediction.

A. High-Throughput Screening as Workhorse in Drug Discovery

With the swift development of robotics and biochemical assaying techniques, it has become possible to screen *in vitro* thousands to millions of compounds against several targets of interest, in a process commonly known as *high-throughput screening* (HTS). Obviously, the availability of a fast screening technology coupled to soaring numbers of test models, either biochemical of phenotypic, requires the development of new chemistry that meets the high demand for novel structures. Chemistry has thus become the bottleneck of an increasingly fast-paced drug discovery process. With the advent of combinatorial chemistry and in particular multicomponent reactions in the 1990s, the compound libraries in pharmaceutical companies were rapidly enriched in thousands of molecules available for testing [3]. However, at the time the increasing number of available screening compounds did not translate into a greater number of primary hits being successfully developed into leads and marketed drugs. In fact, HTS success rates were astonishingly low [2,4]. Several factors may influence this outcome and ultimately result in high attrition rates in clinical trials: Many drug targets are not properly validated and therefore not clinically relevant, and the hits and leads identified do not show appropriate pharmacokinetic (absorption, distribution, metabolism, and excretion; ADMET) and physicochemical properties. While HTS clearly has benefits and may in fact be considered as the main technology fuelling drug discovery pipelines today, it certainly is not accessible to all companies and academic groups, prompting these to embrace other options, as means for running their own programs [5,6].

B. Molecular Informatics in Drug Discovery

In a simplistic manner, molecular informatics—specifically the discipline of cheminformatics—deals with "the application of informatics methods to solve chemical problems" [7]. It was first brought to light with the emergent need to mine the ever increasing chemical compound and property databases. However, molecular informatics is not solely confined to the manipulation and processing of chemistry-related information. As an alternative to HTS, molecular informatics offers a multitude of *in silico* tools that are ubiquitous to practically all phases of drug discovery. Virtual screening methods provide a cost-effective and convenient solution for accessing chemical space and discovering hit and lead compounds in potentially shorter processing timeframes. Computational tools can also guide the hit-to-lead process and the optimization of lead compounds by tuning the binding kinetics, thereby improving affinity to the drug target and eventually drug potency; model physicochemical properties to improve the likelihood of suitable pharmacokinetics; and predict metabolic liabilities, toxicity, and *off*-target engagement, among others. In this regard, molecular informatics currently plays a transversal and crucial role specifically in the early phases of drug discovery and attempts to tackle the current loss of innovation expressed by low numbers of NCEs, while also improving the chance of advancing compounds to clinical trials and decreasing attrition rates in the late development stages.

II. *IN SILICO* SCREENING

Virtual library screening offers a serious complementary approach to HTS. As an example, in a computer-based screen of commercially available compounds, Shoichet and co-workers were able to identify inhibitors of dihydrofolate reductase that had not been retrieved by traditional HTS [8]. In virtual screening, the test compounds are selected by the use of software that explicitly or implicitly predicts compound binding to a drug target. In some cases, the hit rates of virtual screening, i.e., the number of compounds with the desired activity (e.g., receptor binding, enzyme inhibition) among the tested compounds, are significantly higher than in HTS [9]. Virtual screening provides rapid access to millions of compounds that can inclusively be purchased and tested. However, given the magnitude of drug-like chemical space (c. 10^{60} compounds; estimations vary), it is of the utmost importance to filter out undesirable molecules in a pre-screening process. Problematic synthetic accessibility, the presence of reactive groups, and poor predicted oral bioavailability via noncompliance to the Lipinski's "rule-of-5" and other drug-likeness estimations are some of the most commonly used pre-processing filters [9,10]. By applying these filters, the overall quality of the query database is improved, and the false-positive retrieval rate can be significantly reduced (a "false positive" is a compound that is erroneously predicted to have a certain desired activity or property). Also, a successful and meaningful prediction is largely dependent on the prior quality and amount of starting data, namely a model of the receptor structure (ligand-binding pocket) or a known ligand. Therefore, a sufficiently accurate three-dimensional structural model of the target (by x-ray crystallography, NMR, or sometimes even computational means) for the former or the binding pose for the latter case are ideally in hand for virtual compound database screening [11]. When both target and ligand structures are known, combinations of this information can be effectively employed for computer-assisted compound screening.

Virtual screening methods can be broadly categorized as ligand- or receptor-based. Both strategies are not only valid and comparable for a range of targets but also complementary, as different active ligands may be found using different software packages and scoring schemes [12].

A. Ligand-Based Virtual Screening

Ligand-based virtual screening is grounded on the hypothesis of similar ligands exerting similar biological activities. Hence, to conduct virtual screenings one needs at least one reference ligand (query) in a computer-readable molecular representation ("descriptor" or set of descriptors) and a similarity index by which the degree of pair-wise compound similarity can be measured (Figure 6.1). Commonly employed similarity indices are the Tanimoto-Jaccard similarity coefficient T (Eq. 6.1; often simply referred to as "Tanimoto index"), and the Minkowski metric D (Eq. 6.2). The Tanimoto index computes a value between zero and one indicating the

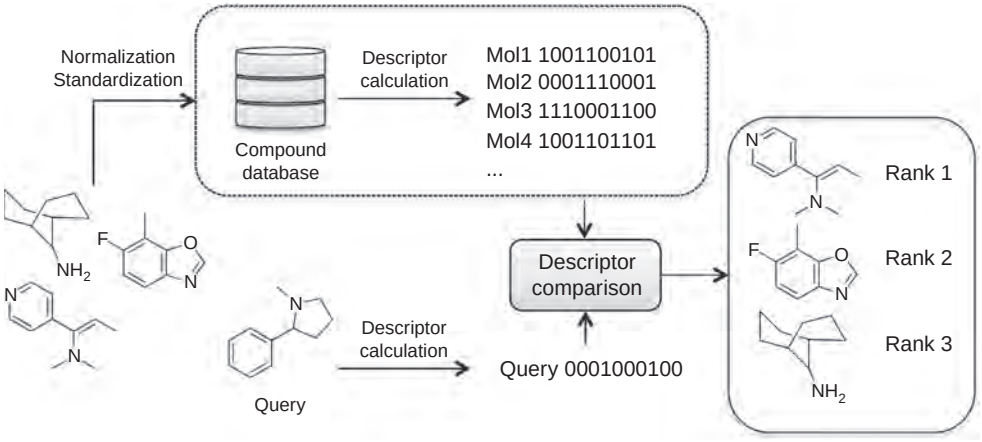


FIGURE 6.1 Schematic process of ligand-based similarity searching in a compound database. In the example, molecular structures are represented in terms of fingerprints ("bitstrings"). Various other descriptor types are possible and frequently used for similarity searching.

similarity of the fingerprint descriptors of molecules A and B. A value of one indicates fingerprint identity. In contrast, the Minkowski metric computes a value between zero (identity) and infinity.

$$T = \frac{c}{a+b-c}\text{, where} \tag{6.1}$$

a is the number of bits set in fingerprint A, b the number of bits set in fingerprint B, and c the number of common bits set in both fingerprints. In a substructure-based fingerprint representation, a "bit" corresponds to a certain substructure element.

$$D = \sqrt[p]{\sum_{i=1}^{n} |\mathbf{A}_i - \mathbf{B}_i|^p}\text{, where} \tag{6.2}$$

n is the number of descriptor elements. The most commonly employed forms of the Minkowski metric are the Manhattan distance ($p = 1$) and the Euclidian distance ($p = 2$).

Although a high-affinity query molecule is preferred and generally accepted as a good starting point for similarity-based virtual screening, low-affinity queries may also retrieve hits from a compound database. A wealth of screening techniques and software is available for that purpose, and its proper selection depends on the goal, suitability, and availability of descriptors and software, among other factors [13–15]. It has been shown that the actual performance of similarity searching largely depends on the choice of the molecular descriptors [16]. In general, these comparisons enable the calculation of a measure of similarity between the query compound and each of the molecules stored in a compound database. The latter are sorted in the order of decreasing similarity to the query, and the output from the search is provided as a ranked list to the user of a virtual screening software tool. In this result list, the chemical structures that are deemed to be most similar to the query structure—the "nearest neighbors"—are located on the top ranks.

Numerous descriptors have been devised for virtual screening and may be classified according to the underlying molecular dimensionality. Molecular weight is an example of a one-dimensional (1D) descriptor because only the molecular formula of the molecule is needed to calculate the value. Such descriptors are useful as filters for pre-processing the database search as they are intimately linked to the rule-of-5 [17]. A growing number of topological (2D) and 3D descriptors are of mainstream use in drug discovery, some of which are presented here. In parallel, a steady increase of small molecules' bioactivity data affords the perfect scenario for devising new virtual screening methods. Consequently, machine learning and "big data" mining methods are increasingly applied to identify compounds with the desired pharmacological activities and have become a viable alternative to conventional similarity search methods [18].

1. *Topological (2D) Descriptors*

Topological descriptors are solely based on the compound constitution and configuration, and are therefore independent from the molecular conformation. They have found broad applicability in drug discovery. Many straightforward similarity searches are based on 2D representations of molecular substructures [19,20]. Substructure-based similarity searching is most often used for retrieving a series of analogs of a given query compound. However, one should keep in mind that searching by substructure similarity bears the risk of compromising chemotype novelty among the reported hits. The number of occurrences of a given structural feature in a molecule and the computed octanol-water partition coefficient (clog$P$) are examples of 2D descriptors that can also be used as pre-processing filters in order to improve the quality of the screening outcome.

The Morgan algorithm [21] was developed as a connectivity code for describing chemical structures in the Chemical Abstracts Service (CAS). It applies an iterative process for allotting numeric identifiers to each nonhydrogen atom, using the identifier from the previous iteration. Once all atoms are maximally disambiguated, the final identifiers are used for a canonical atom numbering scheme. The Morgan algorithm and its derivatives set the ground for a wide variety of topological descriptors. For example, Extended-Connectivity Fingerprints (ECFPs) capture the intermediate atom identifiers as features rather than discarding them. There are several variations of this general scheme. For example, Hu et al customized ECFPs for improving the search efficiency and retrieval of active molecules [22]. ECFPs are "circular fingerprints" designed for chemical similarity searching. As such, the atom neighborhood diameter is an important parameter to the descriptor and is usually appended to the name, e.g., ECFP_4, ECFP_6, denoting the bond distance under consideration (four and six bonds distance in ECFP_4 and ECFP_6, respectively) [23,24]. Circular fingerprints are effective tools that have been successfully applied to a wide range of applications in drug discovery, despite being used initially for the evaluation of

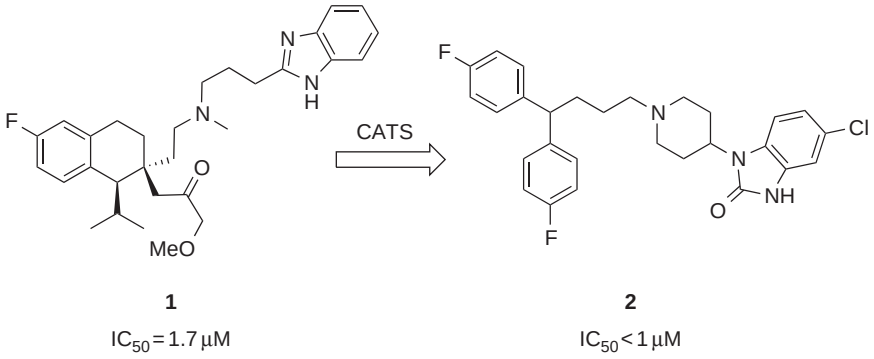
false-positive and false-negative hits in HTS results. In a similar fashion, ECFPs are frequently employed in ligand-based virtual screening campaigns to distinguish between active ligands from inactives, or simply as a means for compound clustering [25].

For example, ECFP_4 and ECFP_6 were used to analyze the structural similarity of virtual screening hits in discovery programs aiming at the identification of novel histamine H_1 and peroxisome proliferator-activated receptor (PPAR) ligands [26,27]. In another application, ECFPs were effectively applied as descriptors for the purpose of target prediction under the assumption that similar compounds exert comparable biological activities. In this case, one can predict a compound's biological function given only its 2D chemical structure and searching for structure similarity in libraries of drugs or lead candidates with known activities [28]. Similarity searches with ECFPs as descriptors can also be employed for the prediction of ADMET properties [29].

In analogy to substructure fingerprints, topological cross-correlation of generalized atom types is a simple molecular descriptor that leads to a compact, molecular size-independent description of potential pharmacophores [30]. The "*c*hemically *a*dvanced *t*emplate *s*earch" (CATS) [31] descriptor is an example of such a topological descriptor that incorporates the general idea of counting atom-pair distances. Distances are expressed as numbers of bonds along the shortest path connecting two nodes of nonhydrogen atoms in the molecular graph. Each atom is also assigned a specific type between hydrogen-bond donor (D), hydrogen-bond acceptor (A), positively charged (P), negatively charged (N), or lipophilic (L) features, for accurate node representation. Therefore, a total of 15 different atom-pairs (DD, DA, DP, DN, DL, AA, AP, AN, AL, PP, PN, PL, NN, NL, LL) suffice for the topological representation of any drug-like molecule. Distances of up to ten bonds are generally enough for exact molecular representation, leading to a 150-dimensional (15×10) vector, which can be used for screening large virtual compound libraries on the basis of this query correlation vector. In a prospective application of CATS, novel cardiac T-type Ca^{2+} channel blocking agents, e.g., molecule **2** (Figure 6.2) were discovered using Mibefradil (**1**, Figure 6.2) as seed structure. Notably, the "fuzziness" (permissiveness) of the CATS descriptor is ideal for a process termed "scaffold-hopping," i.e., the discovery of new and seemingly unrelated chemical scaffolds that preserve the essential function-determining pharmacophore features [31].

Among many other applications, CATS was also used in an artificial neural network method for the automatic identification of "frequent-hitters." Evidence suggests that topological descriptors intimately combined with machine-learning techniques are valuable tools for the prioritization of compounds from large databases [32]. In another application of CATS, ring systems of natural products were extracted and clustered on the basis of chemical similarity to trade drugs. Analysis of self-organizing maps (SOMs) allowed the identification of neurons representing clusters of molecules having certain topological pharmacophoric patterns that are common in natural products and synthetic drugs. Thus, natural compounds may provide interesting novel scaffold architectures that can be used in drug design [33]. As a further example, CATS and SOMs together were employed to obtain seventeen combinatorial products as adenosine A_{2A} receptor antagonists. Compound **3** showed a 121-fold selectivity for A_{2A}, with K_i (A_{2A}) = 2.4 nM and K_i (A_1) = 292 nM (Figure 6.3) [34].

On a similar note to *on*-target prediction and prioritization of virtual libraries, the CATS topological descriptor and a SOM-based clustering method were successfully applied to discover the *off*-target engagement for a small set of compounds, and predict targets for orphan molecules. In a library of known mGluR antagonists, e.g., **4**, several chemical entities exhibited binding constants in the micromolar range for dopamine receptors, offering proof-of-concept for using 2D descriptors in lead discovery for early recognition of potential side-effects [35]. Finally, compound **5** with the imidazopyridine privileged scaffold was successfully de-orphaned as a muscarinic M_1 receptor antagonist using a similar method [36].



1
$IC_{50} = 1.7\ \mu M$

2
$IC_{50} < 1\ \mu M$

FIGURE 6.2 Structures and activities of **1** (Mibefradil) and **2** (virtual screening hit) against the cardiac T-type Ca^{2+} channel.

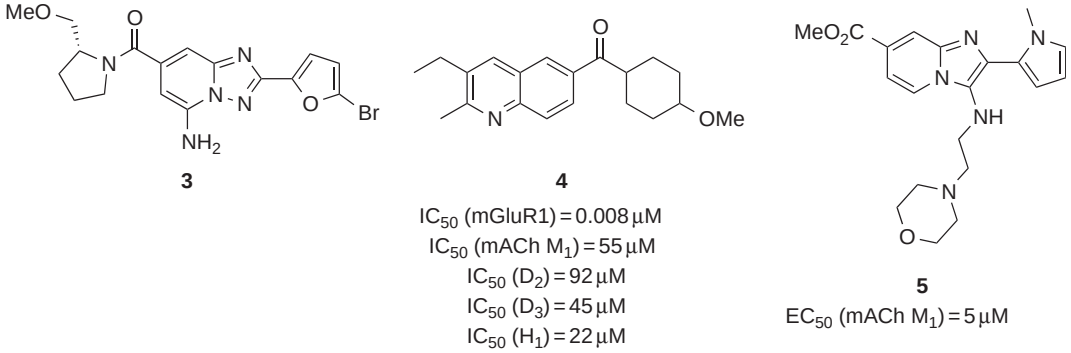


FIGURE 6.3 Structure of selective A_{2A} antagonist **3**, off-target inhibition of mGluR1 antagonist **4** and mACh M_1 modulator **5** discovered through application of self-organizing maps with CATS as descriptor. mGluR1—metabotropic glutamate receptor; mACh M_1—muscarinic acetylcholine receptor; $D_{2/3}$—dopamine receptor $D_{2/3}$; H_1—histamine receptor H_1.

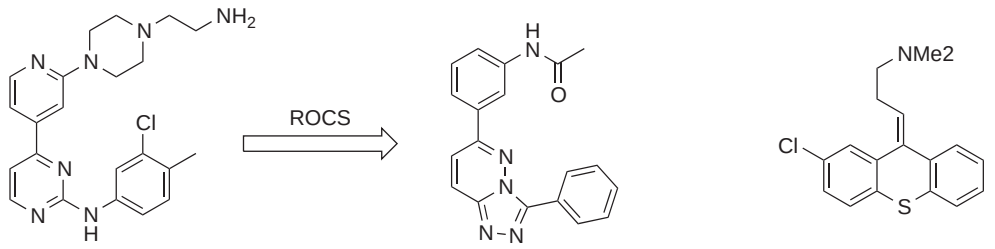
Taken together, these selected examples highlight the strength of ligand-based approaches for virtual screening, which only rely on the topological structure of molecules.

2. 3D Descriptors

Molecular fingerprints can be extended to 3D space by retaining the atom positions and the spatial arrangement of pharmacophoric features within the fingerprint by storing information on pairs, triplets, or quartets of features for multiple 3D conformations. Such features typically include hydrogen-bond donors and acceptors, negatively and positively ionizable groups, and hydrophobic and aromatic moieties. 3D pharmacophore features describe a molecule in an abstract manner, so different chemical structures can have very similar pharmacophore fingerprints. Therefore, similarity searches using pharmacophore fingerprints may result in hits with chemical scaffolds that differ from the query structure but similar pharmacological activity—a concept called "scaffold hopping." Scaffold hops can be achieved by both 2D and 3D descriptors. 3D descriptors depend on the molecular conformation and are most appropriate for representing geometry-dependent bioactivity. Hence, proper conformational sampling of the query database is of extreme importance. Since searches are conducted in respect to a query structure, knowledge of its bioactive conformation is important. These methods require the generation of a library of reasonable low energy conformers, on the assumption that some of those may be representative of bioactive conformations [37,38]. Several programs are available for generating 3D structures and conformers (e.g., Corina, [37,39] Catalyst [40], MOE [41]) employing approaches as varied as genetic algorithms [42] and other stochastic methods [43,44] and molecular dynamics [45] simulations. The prediction of the bioactive conformation of a ligand without experimental data is difficult, because in the majority of cases this particular conformation does not correspond to the energy minimum in the unbound state [46,47]. Nevertheless, irrespective of the importance of having accurate 3D data to start a virtual screening campaign, it has been suggested that the impact of the bioactive conformation on the overall database enrichment is limited [48–50].

A. SHAPE DESCRIPTORS

A popular technique for ligand-based virtual screening is to compare shapes of molecules. This has proven to be suitable and fast for screening millions of compounds with the aim of scaffold hopping. The Rapid Overlay of Chemical Structures (ROCS) [50] approach uses continuous functions derived from atom-centered Gaussians for calculating the volume overlap between molecules. Mismatched volumes are perceived as dissimilarities [51]. The technique was first used to identify new scaffolds for inhibitors of ZipA-FTsZ protein–protein interaction, starting from the bioactive conformation of a HTS hit, **6**. While the newly identified lead structures only bound weakly, e.g., **7**, they presented ample opportunity for further development, including improved synthetic accessibility than the query (Figure 6.4) [52]. In another application, ROCS was used to identify alternative chemical probes for the messenger nicotinic acid adenine dinucleotide phosphate (NAADP) [53] and low micromolar and ligand efficient scaffolds from four distinct template structures that served as enoyl-acyl carrier protein (ACP) reductase inhibitors. In the latter example, one of the most promising small molecules in this study presented activity in whole-cell Gram-positive and Gram-negative antibacterial assays [54]. Furthermore, ROCS has also been successfully employed for the fast reprofiling of FDA-approved drugs. Using known H_1 antagonists as query structures, a high retrieval rate was obtained by shape comparisons. Several approved drugs were found



6
$K_i = 12\,\mu M$

7
$K_D = 74\,\mu M$

8
$IC_{50} = 1\,nM$

FIGURE 6.4 Examples of active molecules found through application of shape similarity searching.

TABLE 6.1 Pharmacophore Feature Assignment in Four Software Packages

Feature	LigandScout	Catalyst	Phase	MOE
Hydrogen-bridge	Acceptor and donor features located at heavy atom	Acceptor and donor features located at heavy atom	Donors located at H atom and acceptors at heavy atom	Acceptor and donor features located at heavy atom
Lipophilic	Represented with tolerance spheres	Represented with tolerance spheres	Represented with tolerance spheres. Aromatic rings are not recognized as hydrophobic areas	Represented with tolerance spheres. Aromatic rings are not recognized as hydrophobic areas
Aromatic	Represented with tolerance and ring plane orientation	Represented with tolerance and ring plane orientation	Represented with tolerance and ring plane orientation	Depends on the pharmacophore scheme
Charge transfer	Tolerance spheres without explicit charges	Tolerance spheres without explicit charges	Tolerance spheres without explicit charges	Tolerance spheres with explicit charges

Adapted from Ref. [58].

to inhibit the H_1 receptor in the micromolar range, with chlorprothixene, **8**, being the most active ($IC_{50} = 1$ nM) [55]. ROCS-based alignment was also shown to be superior to explicit ligand-receptor docking in the prediction of correct active site orientation and metabolic sites of small molecules by cytochrome P450 (CYP2C9) [56], further evidencing the broad applicability of shape descriptors. Numerous variations of this similarity searching principle have been conceived and are productively used in hit and lead finding projects.

B. 3D PHARMACOPHORES

According to Wermuth's definition, [57] "a pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response. A pharmacophore does not represent a real molecule or a real association of functional groups, but a purely abstract concept that accounts for the common molecular interaction capacities of a group of compounds toward their target structure. The pharmacophore can be considered the largest common denominator shared by a set of active molecules. This definition discards a misuse often found in the medicinal chemistry literature, which consists of naming as pharmacophores simple chemical functionalities such as guanidines, sulfonamides or dihydroimidazoles (formerly imidazolines), or typical structural skeletons such as flavones, phenothiazines, prostaglandins or steroids."

In other words, regions encoding several types of potential receptor-interactions of the molecule in 3D space and inherently reflecting complementarity to the drug target can be used for comparing ligands, rather than using the explicit chemical structure of the molecules. Different schemes for assigning hydrogen-bond acceptors and donors, aromatic, hydrophobic, and other pharmacophoric points are employed depending on the software used (Table 6.1, Figure 6.5) [58]. Each feature is accompanied by a tolerance zone that can be occupied by atoms or groups of atoms capable of exerting such type of interaction. This kind of molecular representation is highly effective for database mining and scaffold hopping, because molecules with different structural motifs can

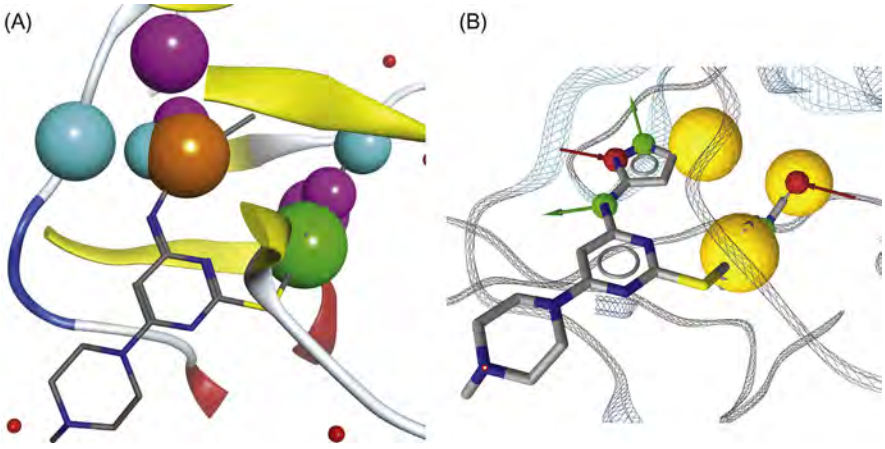


FIGURE 6.5 Pharmacophoric features of VX-680 (PDB-ID: 3E5A) [59] in the ATP-binding site of Aurora A kinase, using the software suites MOE (A) and LigandScout (B). Aromatic (orange), hydrogen-bond acceptor (cyan), hydrophobic (green) and hydrogen -bond donor (purple) features are highlighted in the MOE-derived model, while in (B) hydrophobic features (yellow), hydrogen-bond donor and acceptors (arrows) and exclusion volumes (grey) are highlighted.

present a similar chemical behavior and consequently the same biological effect (i.e., being bioisosteres). Catalyst, [60] LigandScout [61], and Phase [62] are representatives of commonly used computer programs in 3D pharmacophore-based virtual screening (Table 6.1).

Occasionally it is difficult to generate a unique model reflecting the pharmacophoric features of a structurally diverse set of ligands, which may be at least partly due to different binding sites and ligand binding modes. A multiple ligand alignment assumes identical binding modes, which is often not realistic. Making an error in this step of model building increases the likelihood of failure in the screening process. Thus, it is commendable to build as many models as there are chemotype classes within the reference ligand set. A typical workflow begins with the selection of suitable query structures for model generation and superimposition, ideally, of bioactive conformations. After generation of a conformer library for the set of query ligands, the software identifies conformers from the database that match with the previously assembled pharmacophore model, making this technology ideal for the rapid retrieval of hits [63].

3D pharmacophore models are among the most commonly used virtual library screening techniques. For example, Catalyst was used to identify novel high-affinity ligands for ERG2, emopamil binding protein (EBP), and the sigma-1 receptor (σ1) [64], while LigandScout was used to generate both ligand- and structure-based pharmacophore models for the identification of noncovalent and covalent cysteine protease cathepsin S (CatS) inhibitors [65]. Other examples of successful employment of 3D pharmacophore models are related to the development of a series of compounds selectively targeting MDR-1-expressing cells [66] and the design of a 5-point pharmacophore model for distinction between kinase frequent hitters and more selective inhibitors [67].

C. MISCELLANEOUS

A broad wealth of other 3D descriptors is used in computational hit and lead discovery. The comparative molecular field analysis (CoMFA) [68] samples steric and electrostatic fields around ligands to describe bioactive conformations that can then be used for 3D quantitative structure-activity relationship (QSAR) studies. Hence, CoMFA represents electrostatically and sterically favorable or unfavorable areas around the molecules for chemical interpretation of the models and understanding of the crucial regions for ligand binding. BCUT [69] and EVA [70] are other descriptor examples commonly used.

B. Receptor-Based Virtual Screening

1. Pharmacophore-Based Similarity Searching

A receptor-derived pharmacophore model approach is in all aspects similar to a purely ligand-based method, except that it uses the information on protein−ligand interactions from an experimentally determined 3D structure as key data for feature assignment. Despite the caveats of fitting mathematical models to electron densities for determining protein and ligand conformations, this technology has been increasingly used to screen virtual compound libraries. It is not only a complementary approach to molecular docking but potentially less demanding in computational resources and, consequently, of higher throughput.

2. *Molecular Docking*

Molecular docking is a popular technique among receptor-based methods available for screening virtual libraries of compounds. While predicting the ligand conformation within the user-defined binding site—the so-called "docking pose"—it attempts to provide an accurate structural model capable of rationalizing bioactivity or lack of. Therefore, automated ligand docking aims at the identification of chemical features responsible for molecular recognition and suggests how a medicinal chemist should modify a molecule to improve binding affinity. As simple and attractive as this technique may sound, one should always bear in mind that suggested poses are binding models on top of 3D structural models (x-ray structures). Consequently, the careful selection of starting data is fundamental to avoiding exponential propagation of inaccurate information that may lead irrevocably to failed pose predictions [71]. A docking job employs an algorithm that samples small molecules' conformations and searches for the best binding pose to be adopted in the binding site (i.e., the pose showing best complementarity to the pocket). Thus, a scoring function is necessary to rank the ligand-target fitting. It should be noted that scoring poses is not a trivial task, and a great deal of effort has been put together to improve current schemes. Enthalpic and entropic effects govern ligand binding according to the associated change in free energy (Eq. 6.3):

$$\Delta G = \Delta H - T\Delta S, \text{where} \tag{6.3}$$

G is the free energy of the ligand-receptor interaction, H is enthalpy, T is the absolute temperature, and S is entropy. While enthalpic contributions to ligand binding are currently modeled with some accuracy, the same is not true for the entropic contributions. Moreover, other challenges that may preclude the correct identification of binding poses are:

i. limited resolution of crystallographic targets;
ii. protein/enzyme flexibility;
iii. induced fit that occurs upon binding; and
iv. contribution of water and other solvent molecules in protein-ligand interactions.

A. PROTEIN FLEXIBILITY AND INDUCED FIT

State-of-the-art docking algorithms predict an incorrect binding pose for up to 50–70 percent of all ligands when only a single fixed receptor conformation is considered. However, accounting for protein flexibility as a whole is a complex and challenging issue due to the large conformational space that needs to be sampled [72]. Hence, in flexible docking the search algorithms usually model solely the residues in the active site either by molecular dynamics [73,74] or by computing rotamer libraries [75,76]. In the latter case, prior knowledge of common conformations (e.g., through crystallography) is needed. Alternatively, an ensemble of 3D protein structure conformations on which the user performs rigid docking may be used as a shortcut with similar success rates to flexible docking and significantly lower run times [77].

B. SCORING FUNCTION

Scoring functions probably are the most relevant ingredient to a successful docking program and represent a very active field of research. Search algorithms can efficiently identify the binding pose of small molecules, but the result only becomes significant if a scoring function is able to distinguish between correct and incorrect poses, or—in a broader sense—active and inactive ligands. It has been suggested that most scoring functions fail to show significant correlation with binding constants when confronted with novel ligand-receptor systems, despite being tuned to predict binding constants for a training set. Therefore, the focus of applied docking studies should be more on finding actives at all rather than correctly ranking all molecules from a compound database according to their predicted ΔG values [77,78].

Generally, scoring functions can be categorized as force field-based, empirical, or knowledge-based scoring functions. Force field-based functions quantify the sum of the receptor-ligand interaction and the steric strain induced by ligand binding. Van der Waals interactions, which play an important role in molecular recognition, are often described by Lennard-Jones potentials, while the electrostatic interactions are accounted for by a distance-dependent Coulombic term. Thus, the force field scoring functions are limited in the sense that they not include an entropic term [79,80]. On the other hand, empirical scoring functions are designed to estimate the binding free energy as a weighed sum of interaction terms, using measured binding constants as source data for coefficient calculation [81]. Therefore, a clear disadvantage of this method is its dependence on the training data, rendering it as possibly unsuitable for scoring chemically unrelated ligands [80]. Finally, knowledge-based functions use statistics for describing interatomic distances on the assumption that they implicitly reflect favorable or

unfavorable interactions between functional groups [78]. As no scoring function performs in a satisfactory manner, combining multiple scoring functions (i.e., consensus scoring) may prove a fruitful solution for diminishing the number of false-positives [82,83].

C. WATER

Correct modeling of water molecules in docking runs is of paramount importance, as they can be involved in ligand recognition via mediation of hydrogen-bridges or be displaced, thereby contributing to the entropic factor of the ligand-receptor interaction [84]. For example, many HIV protease inhibitors act as transition state mimetics by displacement of a catalytic water molecule [85]. Thus, accounting for water molecules in the binding pocket may potentially improve modeling outcomes, even though the exact position and orientation of these interactions is difficult to predict [86,87].

D. EXAMPLES

Several software packages are available for molecular docking [88], but only a few of them have become widely used. AutoDock [89], DOCK [90], FlexX [91], Glide [92], GOLD [93], and ICM [94] are a selected examples of benchmark packages. An extensive listing can be found elsewhere [88]. Several comparative studies have been published over time, highlighting the strengths and weaknesses of each software tool. While binding poses are generally well predicted, the accuracy in identifying the crystallized binding mode is where most programs vary [88]. For example, AutoDock 4.0 was recently used to discover novel antiplasmodial compounds as potential cytochrome bc_1 inhibitor starting points for medicinal chemistry optimization [95]. Several compounds were identified as micromolar fungal CYP53 inhibitors by relying on FlexX [96]. Glide was efficiently used for screening a library of compounds and retrieve 18 molecules with IC_{50} values ranging from 0.1 to 19 μM as hDHODH inhibitors while showing species selectivity [97]. Finally, ICM was applied to identify a novel class of pyrimidine-based compounds as low-micromolar inhibitors of Filamenting temperature-sensitive mutant Z (FtsZ) with activity against *Staphylococcus aureus* and *Escherichia coli* [98].

Automated ligand docking can deliver hits from compound databases, provided an accurate receptor model is available and a sensitive scoring function is used.

III. *DE NOVO* DESIGN

Computer-assisted *de novo* drug design was conceived with the aim of producing NCEs from scratch, while exploring a substantial fraction of uncharted chemical space [99]. The concept was introduced roughly twenty years ago and has already impacted the discovery of tractable leads for medicinal chemistry [100]. As an extension to *in silico* screening of physically available compounds stored in databases, *de novo* design generates virtual compounds for computational analysis (e.g., properties, similarity, docking, ranking). It is important to stress that, like conventional compound database screening, successful *de novo* design studies rarely focus on affording highly potent lead compounds in first instance but rather aim at finding innovative scaffolds for further development [101,102]. Therefore, *de novo* design of small molecules is currently perceived as a prime tool for advanced chemical space mining and scaffold-hopping [102,103]. Again, information about known ligands or the receptor of interest is key for the idea-generating step of *de novo* design, and secondary constraints for compound selection (e.g., log*P*, molecular weight) may be applied [99]. At the inception of *de novo* design most tools were largely atom-based for fine-grain chemical space probing and usually yielded highly tailored chemotypes. Therefore, despite the possibility of searching all chemical space synthetic, accessibility was rarely feasible. Currently, fragment-based approaches are largely employed in an attempt to mitigate the access issues to suggested structures and maintain some degree of "drug-likeness." In this section we provide a brief overview of current *de novo* design tools and success stories.

A. Receptor-Based *De Novo* Design

Receptor-based *de novo* design is founded on the principle of ligand-receptor complementarity. As such, accurate data on the receptor of interest is required for legitimate success expectations. Several software tools addressing this approach have been reported in the literature [99,102]. For example, Honma et al used LEGEND coupled to SEEDS for designing a CDK4 inhibitor [104]. Similarly, NovoFlap was used to propose a novel 5-HT_{1B}

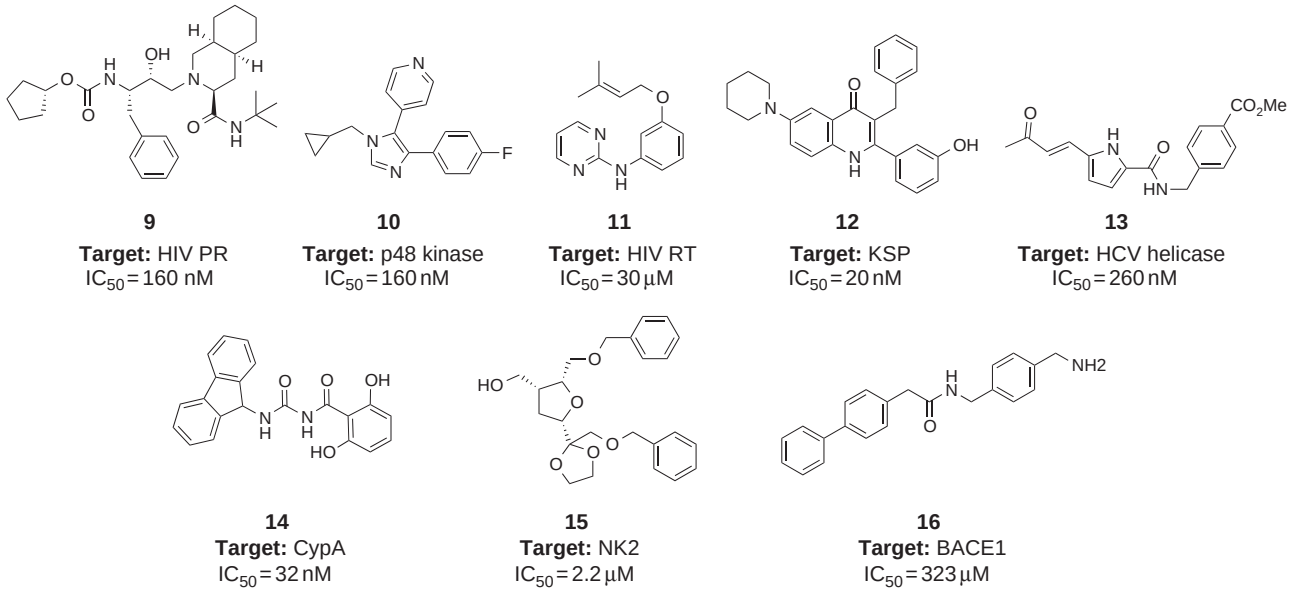


FIGURE 6.6 Examples of *de novo* designed compounds that were obtained by receptor-based design.

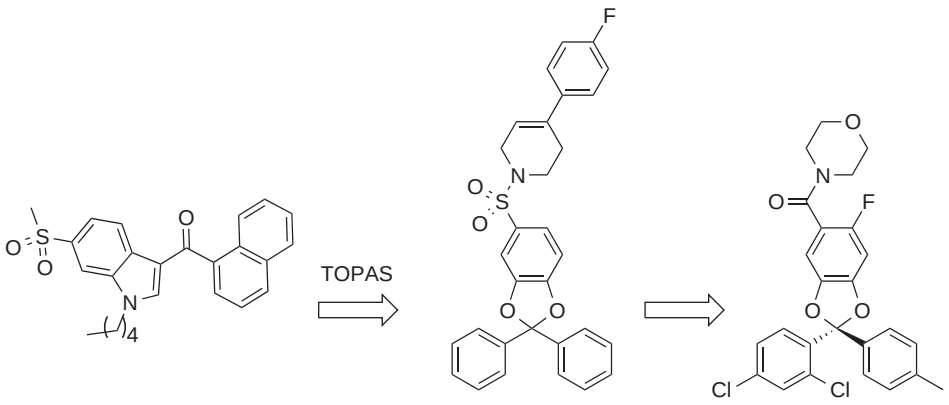
receptor chemotype [105]. BREED is a computer program that automates the common practice in medicinal chemistry of fragment swapping-joining between two distinct bioactive ligands. By employing this design strategy Pierce et al obtained new inhibitors of HIV protease inhibitors, including **9** (Figure 6.6) and p38 kinase, **10** [106]. On the other hand, BOMB applies a growth algorithm that was successfully used for designing a HIV reverse transcriptase (HIV RT) lead compound, **11** [107], relying on a free energy perturbation approach, whereas GeometryFit was used for designing a HIV gp41 inhibitor [108]. These selected examples show that there are numerous ways new chemical structures with the desired activity (receptor-binding) can be successfully generated *in silico*.

Among the most commonly used *de novo* design programs are LUDI, LigBuilder, and SPROUT. LUDI is a receptor-based software that screens fragment libraries. Fragments that have favorable predicted interactions within the binding pocket are then connected to new molecules. Using this fragment linking and growing strategy, a series of 1,4-dihydroquinolin-4-ones was designed as kinesin spindle protein (KSP) inhibitors, with **12** presenting an IC_{50} value of 20 nM [109], and quinazolines were discovered as FLT3 inhibitors [110]. Similarly, LigBuilder employs a genetic algorithm where the user is able to choose the design strategy (e.g., fragment linking or growing) [111]. For example, twenty-four high-affinity BRAF kinase inhibitors [112], thirty-two Cdc25 phosphatase A/B inhibitors amenable for further optimization, and a novel HCV helicase inhibitor **13** were identified using LigBuilder [113]. With the aim of improving the general quality of the designed molecules, LigBuilder v2.0—including ADMET filters and evaluation of synthetic feasibility for designed ligands—was recently reported to mitigate poor synthetic accessibility of LigBuilder 1.2 designs [114]. As a first proof-of-concept, Ni et al reported **14** as a nanomolar inhibitor of cyclophilin A (CypA) [115]. Finally, SPROUT maps the receptor and generates a skeleton that is then followed by atom substitution to afford novel molecules [116–118]. This software was successfully applied in the design of a new class of NK2 inhibitors (e.g., **15**) [119]. Moreover, in a recent application of SPROUT, Fishwick and co-workers discovered a series of non-peptidic BACE1 inhibitors based on the biphenylacetamide scaffold and identified **16** as a hit [120].

B. Ligand-Based *De Novo* Design

While receptor-based approaches have been of relevance for the rational design of hits and leads in drug discovery, they heavily rely on in-house or public availability of structural data for the receptor of interest. Hence, they are of limited use on receptors where current state of the art lacks 3D receptor models.

G-protein coupled receptors (GPCRs) represent a family of signal transduction gatekeepers with more than 800 members [121]. Although divided into classes and sub-families, they all share a seven-transmembrane helices' architecture. GPCRs trigger cascades of physiological responses through signal transduction as consequence



17
Target: CB1
K_i = 110 nM

18
Target: CB1
K_i = 13 nM

19
Target: CB1
K_i = 13 nM

FIGURE 6.7 *De novo* design and optimization of a CB1 inverse agonist.

of recognition and relay of extracellular stimuli to different intracellular proteins [122]. Thus, an imbalance of GPCR-mediated signaling is linked to numerous diseases [123,124]. Despite the prominent role in (patho)physiology, in-depth understanding of GPCR mechanisms on the molecular level is widely lacking. The conformational flexibility of these receptors has made crystallization challenging, and much of our current knowledge has been inferred from homology models to receptors within the same family. As of today only 110 GPCR complexes with different GPCR-ligand complexes have been reported [121]. Therefore, conventional structure-based *de novo* design of innovative chemotypes has been seldom applied to GPCRs.

TOPAS (TOPology Assigning System) was disclosed over a decade ago as an evolutionary algorithm for fragment-based *de novo* design that aims at generating NCEs by mimicking a template structure. A set of 25,000 fragment structures derived from 36,000 drugs function as building blocks for *de novo* design following eleven distinct reaction schemes that implement RECAP [125] fragmentations (e.g., amide, ester, amine, urea, ether, olefin synthesis). As such, re-assembly of small fragments using this set of chemical reactions partly overcomes the long-standing issue of synthetic feasibility of atom-based *de novo* designs. The CATS [31] topological pharmacophore descriptor was also implemented in TOPAS as a work-horse for measuring compound fitness [126], and Euclidean distances (Eq. 6.2) are calculated to express molecular similarity. As described before, the fuzzy nature of CATS is ideal for scaffold-hopping, allowing exploration of new activity islands in chemical space [31]. It was successfully applied to discover novel, patentable, and low micromolar cannabinoid (CB1) inhibitors (Figure 6.7) [127]. Four months of multi-dimensional optimization of **17** led to the identification of a CB1-selective low nanomolar inverse agonist, **18**, which was further developed to **19** [128]. Compound **19** presented high efficacy *in vitro* in a hypothermia assay and acceptable pharmacokinetic properties. A subsequent evolution of TOPAS by application of *pseudo*-retrosynthetic rules in a stock of chemically meaningful building blocks derived from drug molecules in conjunction with a similarity scoring led to the development of the software tool Flux (*F*ragment-based *L*igand B*u*ilder rea*x*ions) [129,130]. It uses an algorithm with mutation and crossover operators for structure breeding, thus using an evolution strategy, and the Manhattan distance (Eq. 6.2) between the seed structure and designed compounds as fitness function. In particular, Flux was successfully employed for *de novo* designing RNA ligands [131].

Based on the assumption that shape similarity is an important feature for bioisosteric replacement, the shape-based similarity method SQUIRREL (Sophisticated Quantification of Interaction Relationships) was also used for *de novo* design. It decomposes the molecular surface of a given template molecule to "shapelets" that represent local curvatures, and suggests replacements on the basis of fragment shape alignment. The surface is sectioned according to chemical rules, and the alignment performed by means of a graph-matching algorithm. In a prospective application, a pyrrole-based compound was designed as a nanomolar PPARα and PPARγ inhibitor [132].

The *de novo* design software DOGS (Design of Genuine Structures) was recently reported for a broad range of applications [133–135]. It applies a set of eighty-three established reactions to a database of more than 25,000 commercially available and curated building blocks for stepwise construction of innovative molecules.

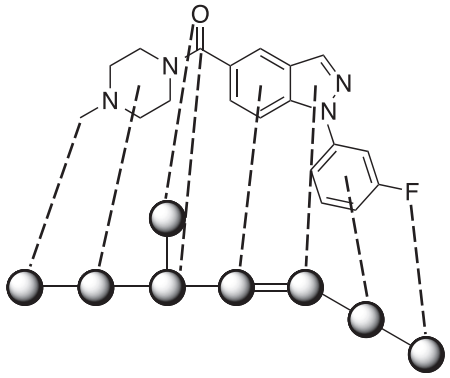
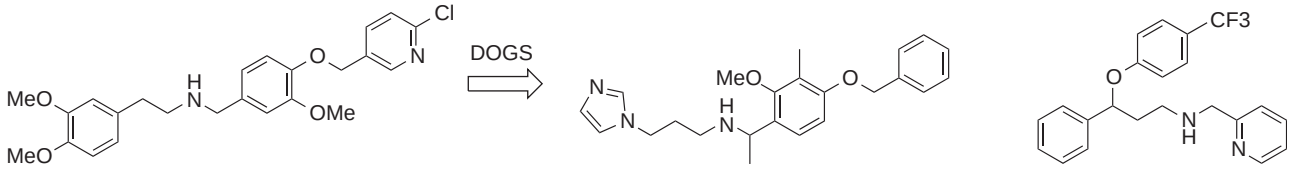


FIGURE 6.8 Molecular graph representation of a molecule with dashed lines denoting feature correspondence to the graph.



20
Target: hPlk-1
$IC_{50} = 0.2$ nM

21
Target: hPlk-1
$IC_{50} \sim 100$ nM

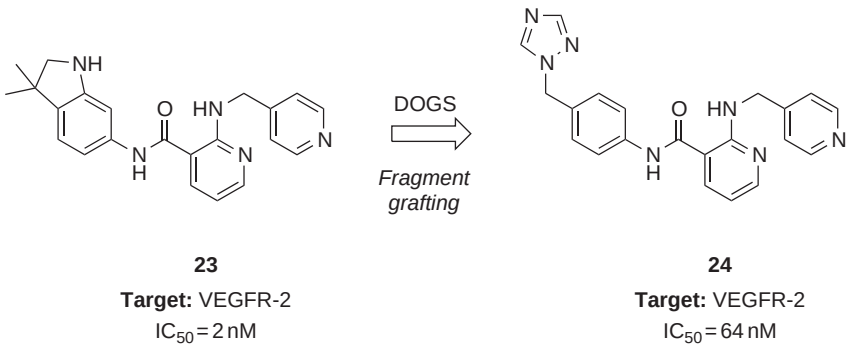
22
Target: hPlk-1
$IC_{50} \sim 50$ nM

FIGURE 6.9 *De novo* design of highly selective hPlk-1 inhibitors with DOGS.

From the chemistry toolbox implemented in DOGS, almost 80 percent of the generated scaffolds from virtual one-step synthesis products are not present in a large set of known bioactive molecules for human targets, indicating potential for new discoveries [136,137]. A design cycle in DOGS starts with random selection of the user-defined number of fragments. In the first instance, the algorithm selects the reaction to be applied. A building block or intermediate product typically contains functional groups that can serve as attachment points for subsequent reactions. After detection of a reaction subset that can be applied, the software evaluates which reaction type is more favorable through scoring of dummy products. Next, in a combinatorial-like approach, DOGS evaluates which intermediates obtain higher scores by applying the previously defined chemistry. The growth process continues until a designed molecule reaches 100 ± 30 percent of the template's molecular weight [136]. The quality of the designed molecules, including the intermediates, is evaluated by a ligand-based scoring scheme. DOGS uses a two-dimensional graph kernel method based on 2D topological structures as a scoring function and accounts for synthesizability by explicitly suggesting synthetic pathways [138]. Two different molecular graph representations were implemented to allow the user different levels of abstraction. While the molecular graph setting yields a finer-grained structural model, in the reduced graph mode only the general topological arrangement of the molecule is represented. In short, nonhydrogen atoms or moieties are denoted as vertices that are "covalently" bound by edges. Depending on the pharmacophoric feature (hydrogen-bond donor/acceptor, positive/negative charge, aromatic, lipophilic, or no type) a different "node color" is assigned for graph comparison (Figure 6.8).

DOGS was applied with great success to designing diverse, innovative, highly potent, selective and patent-free kinase inhibitors (Figure 6.9). For example, starting from **20**, a picomolar human Polo-like kinase (hPlk-1) inhibitor ($IC_{50} = 0.2$ nM), **21**, was generated as a type II-selective hPlk-1 inhibitor ($IC_{50} \sim 100$ nM) [139]. Further exploration of the *de novo* designed molecules led to the identification of **22** as a candidate for profiling. In biochemical assays, compound **30** showed a highly type II-selective effect against hPlk-1 and extraordinary ligand efficiency ($LE = 0.66$) [135].

Moreover, a highly selective vascular endothelium growth factor receptor-2 (VEGFR-2) inhibitor was designed from AMG-706 (**23**; Figure 6.10). Upon *de novo* design with DOGS and fragment grafting of a



23
Target: VEGFR-2
IC_{50} = 2 nM

24
Target: VEGFR-2
IC_{50} = 64 nM

FIGURE 6.10 Fragment-based *de novo* design of a VEGFR-2 subtype-selective **24** inhibitor.

hinge-binding motif, compound **24** exhibited the most selective kinase profile among VEGFR-2 inhibitors [140]. Importantly, the key feature for selectivity was generated *in silico* and the biochemical effects were translated in a cellular level [133].

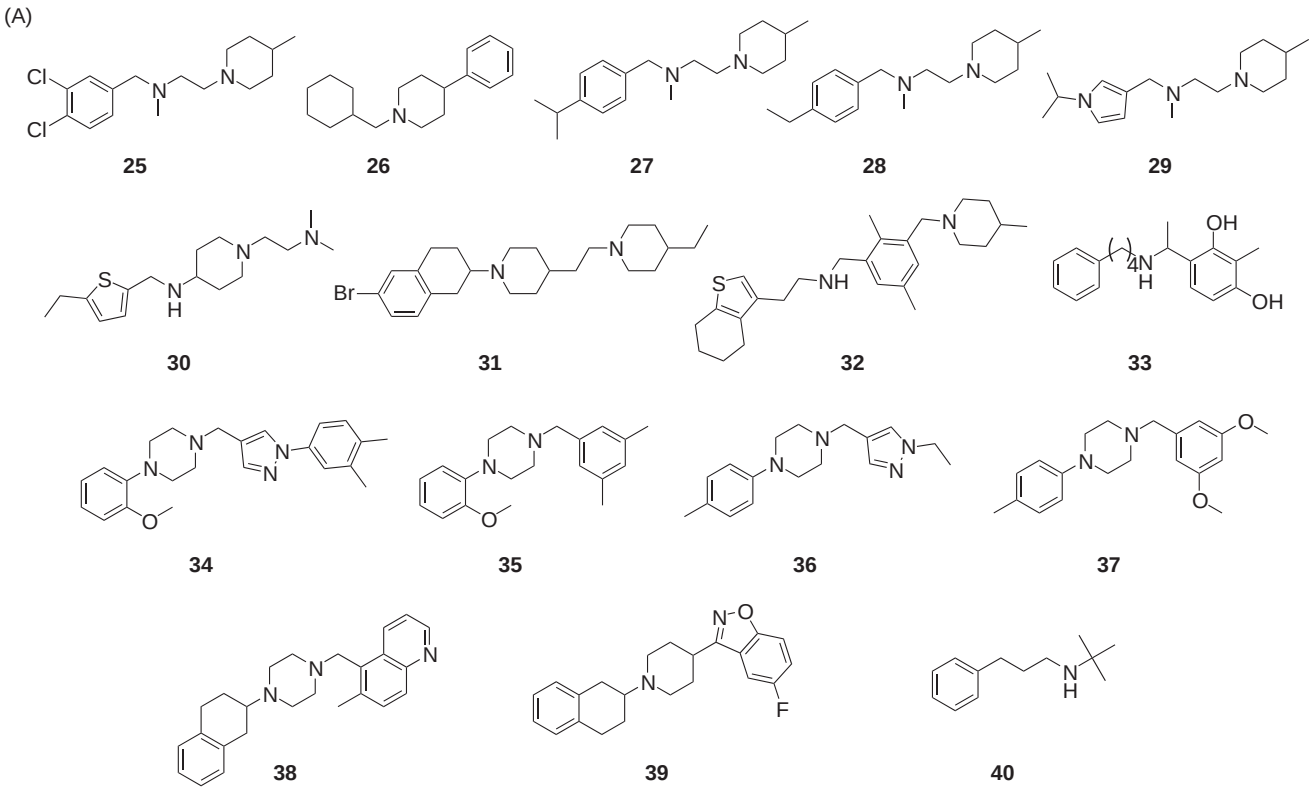
Finally, the MAntA software uses the well-established concept of combinatorial chemistry to generate NCE libraries, by applying a nature-inspired algorithm (i.e., ant-colony optimization) for building block prioritization [141,142]. Relying on publicly available structure-activity data (ChEMBL v14) [143], Reutlinger et al developed a predictive quantitative polypharmacology model encompassing 640 human drug targets. Taking the reductive amination as an example of a privileged reaction, novel subtype-selective and multi-target modulating dopamine D_4 antagonists as well as sigma-1, receptor-selective ligands were obtained with accurately predicted affinities (Figure 6.11). Nanomolar potency, high ligand efficiency, and an overall success rate of 90 percent demonstrate that MAntA may guide target-focused combinatorial chemistry [141].

This automated molecular design method is broadly applicable to other drug target classes and chemistry, including multistep routes, provided that reliable structure-activity data are available for constructing affinity prediction models. The actual computational design process is fast (within minutes on a desktop computer), so that both focused *de novo* library design and chemical synthesis can be executed within a day's work [142]. MAntA has been equally applied with success to short peptide sequences [144].

IV. CONCLUSIONS AND FUTURE DIRECTIONS

Computational medicinal chemistry is contributing to productivity within the pharmaceutical industry and early drug discovery in general by providing alternative and complementary technologies to HTS and facilitating big data handling. The prioritization of chemical series has always been a challenging task. For instance, the computation of ADMET-related parameters (e.g., log*P*/log*S*, metabolic liabilities and reactive sites), has allowed to focus medicinal chemistry efforts on restricted numbers of lead structures. Specific added value of computational tools in drug discovery lies in offering relatively affordable and swift means for the identification of hits and leads through rational experimental engagement. Independently from the availability of software packages for *in silico* database screening, the wealth of available biological data permits making an informed choice about whether ligand- or receptor-based approaches might be more appropriate for a specific research problem. While receptor-based approaches seem more appealing, as of today, the lack of understanding of entropic contributions to molecular recognition spearheads a considerable list of shortcomings. On the other hand, ligand-based approaches require a smaler amount of data to be safely deployed. The combination of several approaches not only provides a multifaceted view on research problems but also delivers a greater chance of succeeding in hit and lead finding. In this regard, the massive growth of biological and chemical data and the capacity of extracting and analyzing their information may significantly contribute to future drug discovery. Integrated systems comprising machine-learning methods for adaptive *de novo* design, *on-*/*off*-target affinity or bioactivity prediction, ADMET and patentability filters are being actively pursued both in the pharmaceutical industry and academia. Therefore, computational mining of chemical space has the potential to further contribute tailored NCEs and mitigate the perceived dearth of innovation in drug discovery.

(A)



(B)

Compound	Predictions (pK_i)					Experimental				
	σ_1	μ	δ	κ	H_3	σ_1	μ	δ	κ	H_3
25	9.6	3.5	2.5	3.4	6.3	100	17	0	47	0
26	9.3	4.3	3.5	4.4	6.0	99	20	0	54	6
27	9.1	3.4	2.8	3.1	7.1	100	5	4	4	33
28	8.8	3.1	2.6	2.7	7.1	97	3	0	0	21
29	8.1	2.7	2.3	2.5	6.0	94	6	0	19	23
30	7.8	3.9	3.4	3.6	6.2	82	7	0	3	54
31	8.1	5.0	4.2	5.2	6.3	98	11	0	10	48
32	7.9	5.3	4.6	4.5	6.9	84	72	0	35	18
33	5.1	2.7	2.3	2.9	5.2	39	14	3	12	0

	Predictions (pK_i)						Experimental					
	D_4	D_1	D_2	D_3	D_5	5-HT_{1a}	D_4	D_1	D_2	D_3	D_5	5-HT_{1a}
34	9.1	4.8	7.0	6.4	4.7	6.6	99	71	99	97	58	100
35	8.6	4.8	6.9	6.1	5.0	6.9	99	20	97	71	16	89
36	7.6	3.4	5.0	4.8	4.6	4.4	100	26	23	31	0	54
37	7.5	3.7	4.7	4.9	4.8	4.8	98	34	54	3	20	44
38	7.6	4.6	6.1	6.4	4.8	6.5	70	70	44	58	50	85
39	7.2	5.6	6.1	6.6	4.9	6.6	96	68	51	74	47	95
40	2.5	2.4	2.9	3.7	4.1	4.9	0	0	1	0	0	0

FIGURE 6.11 (A) MAntA designed small molecules. (B) Predicted activities and experimental values as percentage of inhibition at 2.5 μM in a radioligand-binding assay.

Glossary

Bioisostere Groups or substituents with similar physicochemical properties that produce roughly identical biological outcomes.
Chemotype Chemically distinct entity with a well-defined framework.
Data mining Process involving the discovery and filter of common patterns in a given descriptor space from a large dataset.
***De novo* design** Design from scratch of bioactive compounds by incremental construction of a ligand model.
Euclidean distance Segment length connecting two points.
Focused compound library Compound library designed on the basis of prior knowledge of critical features for bioactivity.
Frequent-hitter Compounds that hit many different targets in a wide range of assays, either by nonspecific binding to proteins or interference with screening assays.
Machine learning Field of study in computer science, applying artificial intelligence to learn patterns from datasets.
Manhattan distance Distance between two points based on vertical and horizontal trajectories.
Multi-dimensional optimization Simultaneous optimization of several constraints to optimal bioactivity.
Neural-network Machine learning model modeled on actual brain neural networks for pattern recognition.
***Off*-target** Receptor or enzyme to which a small molecule binds, wielding a secondary or undesired biological effect.
Pharmacophore Ensemble of steric and electronic features in a molecule that are necessary to ensure binding to a specific target and exerting a biological response.
Privileged structure Molecular scaffolds with versatile binding properties, such that a single scaffold is able to provide potent and selective ligands for a range of different biological targets through modification of functional groups [145].
Scaffold-hopping Process involving the discovery of structurally novel chemotypes by modification of the core scaffold of a known active compound displaying similar biological activity.
Self-organizing map Type of artificial neural-network for unsupervised data learning that produces a low-dimensional representation of the input data, through clustering of related compounds according to the molecular representation (descriptor) used.
Virtual Screening Ensemble of computer-based methods for the purpose of retrieving small molecules binding to a specific target from large compound databases.

References

[1] Hansch C, Fujita T. p-σ-π Analysis: a method for the correlation of biological activity and chemical structure. J Am Chem Soc 1964;86:1616–27.
[2] Drews J. Drug discovery: a historical perspective. Science 2000;287:1960–4.
[3] Macarron R. Critical review of the role of HTS in drug discovery. Drug Discov Today 2006;11:277–9.
[4] Lahana R. How many leads from HTS? Drug Discov Today 1999;4:447–8.
[5] Schneider G. Virtual screening: an endless staircase? Nat Rev Drug Discov 2010;9:273–6.
[6] Bleicher KH, Böhm H-J, Müller K, Alanine AI. Hit and lead generation: beyond high-throughput screening. Nat Rev Drug Discov 2003;2:369–78.
[7] Gasteiger J, Engel T, editors. Chemoinformatics: a textbook. Weinheim. Wiley-VCH Verlag; 2003..
[8] Brenk R, Irwin JJ, Shoichet BK. Here be dragons: docking and screening in an uncharted region of chemical space. J Biomol Screen 2005;10:667–74.
[9] Shoichet BK. Virtual screening of chemical libraries. Nature 2004;432:862–5.
[10] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 2001;46:3–26.
[11] Klebe G. Virtual ligand screening: strategies, perspectives, and limitations. Drug Discov Today 2006;11:580–94.
[12] Krüger DM, Evers A. Comparison of structure- and ligand-based virtual screening protocols considering hit list complementarity and enrichment factors. ChemMedChem 2010;5:148–58.
[13] Xue L, Bajorath J. Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. Comb Chem High Throughput Screen 2000;3:363–72.
[14] Bajorath J. Selected concepts and investigations in compound classification, molecular descriptor analysis, and virtual screening. J Chem Inf Comput Sci 2001;41:233–45.
[15] Livingstone DJ. The characterization of chemical structures using molecular properties. A survey. J Chem Inf Comput Sci 2000;40:195–209.
[16] Bender A, Jenkins JL, Scheiber J, Sukuru SCK, Glick M, Davies JW. How similar are similarity searching methods? A principal component analysis of molecular descriptor space. J Chem Inf Model 2009;49:108–19.
[17] Bajorath J. Integration of virtual and high-throughput screening. Nat Rev Drug Discov 2002;1:882–94.
[18] Geppert H, Vogt M, Bajorath J. Current trends in ligand-based virtual screening: molecular representations, data mining methods, new application areas, and performance evaluation. J Chem Inf Model 2010;50:205–16.
[19] Cramer III RD, Redl G, Berkoff CE. Substructure analysis: novel approach to the problem of drug design. J Med Chem 1974;17:533–5.
[20] Barnard JM. Substructure searching methods: old and new. J Chem Inf Comput Sci 1993;33:532–8.
[21] Morgan HL. The Generation of a unique machine description for chemical structures: a technique developed at Chemical Abstracts Service. J Chem Doc 1965;5:107–13.
[22] Hu Y, Lounkine E, Bajorath J. Improving the search performance of extended connectivity fingerprints through activity-oriented feature filtering and application of a bit-density-dependent similarity function. ChemMedChem 2009;4:540–8.
[23] Rogers D, Brown RD, Hahn M. Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up. J Biomol Screen 2005;10:682–6.
[24] Rogers D, Hahn M. Extended-connectivity fingerprints. J Chem Inf Model 2010;50:742–54.

[25] Hert J, Willett P, Wilton DJ, Acklin P, Azzaoui K, Jacoby E, et al. Comparison of topological descriptors for similarity-based virtual screening using multiple bioactive reference structures. Org Biomol Chem 2004;2:3256–66.
[26] Markt P, Petersen RK, Flindt EN, Kristiansen K, Kirchmair J, Spitzer G, et al. Discovery of novel PPAR ligands by a virtual screening approach based on pharmacophore modeling, 3D shape, and electrostatic similarity screening. J Med Chem 2008;51:6303–17.
[27] de Graaf C, Kooistra AJ, Vischer HF, Katritch V, Kuijer M, Shiroishi M, et al. Crystal structure-based virtual screening for fragment-like ligands of the uuman histamine H1 receptor. J Med Chem 2011;54:8195–206.
[28] Nettles JH, Jenkins JL, Bender A, Deng Z, Davies JW, Glick M. Bridging chemical and biological space: "Target Fishing" using 2D and 3D molecular descriptors. J Med Chem 2006;49:6802–10.
[29] Glen RC, Bender A, Arnby CH, Carlson L, Boyer S, Smith J. Circular fingerprints: flexible molecular descriptors with applications from physical chemistry to ADME. IDrugs 2006;9:199–204.
[30] Schneider G, Wrede P. Artificial neural networks for computer-based molecular design. Prog Biophys Mol Biol 1998;70:175–222.
[31] Schneider G, Neidhart W, Giller T, Schmid G. "Scaffold-hopping" by topological pharmacophore search: a contribution to virtual screening. Angew Chem Int Ed 1999;38:2894–6.
[32] Roche O, Schneider P, Zuegge J, Guba W, Kansy M, Alanine A, et al. Development of a virtual screening method for the identification of "frequent hitters" in compound libraries. J Med Chem 2002;45:137–42.
[33] Lee M-L, Schneider G. Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. J Comb Chem 2001;3:284–9.
[34] Schneider G, Nettekoven M. Ligand-based combinatorial design of selective purinergic receptor (A2A) antagonists using self-organizing maps. J Comb Chem 2003;5:233–7.
[35] Noeske T, Sasse BC, Stark H, Parsons CG, Weil T, Schneider G. Predicting compound selectivity by self-organizing maps: cross-activities of metabotropic glutamate receptor antagonists. ChemMedChem 2006;1:1066–8.
[36] Reutlinger M, Koch CP, Reker D, Todoroff N, Schneider P, Rodrigues T, et al. Chemically advanced template search (CATS) for scaffold-hopping and prospective target prediction for "orphan" molecules. Mol Inf 2013;32:133–8.
[37] Gasteiger J, Rudolph C, Sadowski J. Automatic generation of 3D-atomic coordinates for organic molecules. Tetrahedron Comput Method 1990;3:537–47.
[38] Chen I-J, Foloppe N. Conformational sampling of druglike molecules with MOE and Catalyst: implications for pharmacophore modeling and virtual screening. J Chem Inf Model 2008;48:1773–91.
[39] Sadowski J, Gasteiger J, Klebe G. Comparison of automatic threedimensional model builders using 639 X-ray structures. J Chem Inf Comput Sci 1994;34:1000–8.
[40] Güner O, Clement O, Kurogi Y. Pharmacophore modeling and three dimensional database searching for drug design using catalyst: recent advances. Curr Med Chem 2004;11:2991–3005.
[41] Chemical Computing Group, Inc. Molecular Operating Environment (MOE) 2012.10. 2012.
[42] McGarrah DB, Judson RS. Analysis of the genetic algorithm method of molecular conformation determination. J Comput Chem 1993;14:1385–95.
[43] Saunders M. Stochastic exploration of molecular mechanics energy surfaces: hunting for the global minimum. J Am Chem Soc 1987;109:3150–2.
[44] Saunders M. Stochastic search for the conformations of bicyclic hydrocarbons. J Comput Chem 1989;10:203–8.
[45] Chang G, Guida WC, Still WC. An internal coordinate Monte Carlo method for searching conformational space. J Am Chem Soc 1989;111:4379–86.
[46] Perola E, Charifson PS. Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding. J Med Chem 2004;47:2499–510.
[47] Takagi T, Amano M, Tomimoto M. Novel method for the evaluation of 3D conformation generators. J Chem Inf Model 2009; 49:1377–88.
[48] Renner S, Schwab CH, Gasteiger J, Schneider G. Impact of conformational flexibility on threedimensional similarity searching using correlation vectors. J Chem Inf Model 2006;46:2324–32.
[49] Zhang Q, Muegge I. Scaffold hopping through virtual screening using 2D and 3D similarity descriptors: ranking, voting, and consensus scoring. J Med Chem 2006;49:1536–48.
[50] Hawkins PCD, Skillman AG, Nicholls A. Comparison of shapematching and docking as virtual screening tools. J Med Chem 2007;50:74–82.
[51] Grant JA, Gallardo MA, Pickup BT. A fast method of molecular shape comparison: a simple application of a Gaussian description of molecular shape. J Comput Chem 1998;17:1653–66.
[52] Rush III TS, Grant JA, Mosyak L, Nicholls A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein–protein interaction. J Med Chem 2005;48:1489–95.
[53] Naylor E, Arredouani A, Vasudevan SR, Lewis AM, Parkesh R, Mizote A, et al. Identification of a chemical probe for NAADP by virtual screening. Nat Chem Biol 2009;5:220–6.
[54] Hevener KE, Mehboob S, Su PC, Truong K, Boci T, Deng J, et al. Discovery of a novel and potent class of F. tularensis enoyl-reductase (FabI) inhibitors by molecular shape and electrostatic matching. J Med Chem 2012;55:268–79.
[55] Vasudevan SR, Moore JB, Schymura Y, Churchill GC. Shape-based reprofiling of FDA-approved drug for the H1 histamine receptor. J Med Chem 2012;55:7054–60.
[56] Sykes MJ, McKinnon RA, Miners JO. Prediction of metabolism by cytochrome P450 2C9: alignment and docking studies of a validated database of substrates. J Med Chem 2008;51:780–91.
[57] Wermuth C-G, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1997). Annu Rep Med Chem 1998;33:385–95.
[58] Wolber G, Seidel T, Bendix F, Langer T. Molecule-pharmacophore superpositioning and pattern matching in computational drug design. Drug Discov Today 2008;13: n.p.

[59] Zhao B, Smallwood A, Yang J, Koretke K, Nurse K, Calamari A, et al. Modulation of kinase-inhibitor interactions by auxiliary protein binding: crystallography studies on Aurora a interactions with VX-680 and with TPX2. Protein Sci 2008;17:1791–7.
[60] Accelrys, Inc. Catalyst. <http://www.accelrys.com/>.
[61] Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening fi lters. J Chem Inf Model 2005;45:160–9.
[62] Dixon SL, Smondyrev AM, Rao SR. PHASE: a novel approach to pharmacophore modeling and 3D database searching. Chem Biol Drug Des 2006;67:370–2.
[63] Rodrigues T, Moreira R, Gut J, Rosenthal PJ, O Neill PM, Biagini GA, et al. Identification of new antimalarial leads by use of virtual screening against cytochrome bc1. Bioorg Med Chem 2011;19:6302–8.
[64] Laggner C, Schieferer C, Flechtner B, Poles G, Hoffmann RD, Glossmann H, et al. Discovery of high-affinity ligands of σ1 receptor, ERG2, and emopamil binding protein by pharmacophore modeling and virtual screening. J Med Chem 2005;48:4754–64.
[65] Markt P, McGoohan C, Walker B, Kirchmair J, Feldmann C, De Martino G, et al. Discovery of novel cathepsin S inhibitors by pharmacophore-based virtual high-throughput screening. J Med Chem 2008;48:1693–705.
[66] Hall MD, Salam NK, Hellawell JL, Fales HM, Kensler CB, Ludwig JA, et al. Synthesis, activity, and pharmacophore development for isatin-β-thiosemicarbazones with selective activity toward multidrug-resistant cells. J Med Chem 2009;52:3191–204.
[67] Aronov AM, Murcko MA. Toward a pharmacophore for kinase frequent hitters. J Med Chem 2004;47:5616–19.
[68] Cramer RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 1988;110:5959–67.
[69] Pearlman RS, Smith KM. Metric validation and the receptor-relevant subspace concept. J Chem Inf Comput Sci 1999;39:28–35.
[70] Ferguson AM, Heritage T, Jonathon P, Pack SE, Phillips L, Rogan J, et al. EVA: a new theoretically based molecular descriptor for use in QSAR/QSPR analysis. J Comput Aided Mol Des 1997;11:143–52.
[71] Brooijmans N, Kuntz ID. Molecular recognition and docking algorithms. Annu Rev Biophys Biomol Struct 2003;32:335–73.
[72] Totrov M, Abagyan R. Flexible ligand docking to multiple receptor conformations: a practical alternative. Curr Opin Struct Biol 2008;18:178–84.
[73] Hart TN, Read RJ. A multiple-start Monte Carlo docking method. Proteins 1992;13:206–22.
[74] Oshiro CM, Kuntz ID, Dixon JS. Flexible ligand docking using a genetic algorithm. J Comput Aided Mol Des 1995;9:113–30.
[75] Leach AR. Ligand docking to proteins with discrete side chain flexibility. J Mol Biol 1994;235:345–56.
[76] Bursulaya BD, Totrov M, Abagyan R, Brooks III CL. Comparative study of several algorithms for flexible ligand docking. J Comput Aided Mol Des 2003;17:755–63.
[77] Abagyan R, Totrov M. High-throughput docking for lead generation. Curr Opin Chem Biol 2001;5:375–82.
[78] Kitchen DB, Decomez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 2004;3:935–49.
[79] Wang R, Lu Y, Wang S. Comparative evaluation of 11 scoring functions for molecular docking. J Med Chem 2003;46:2287–303.
[80] Ferrara P, Gohlke H, Price DJ, Klebe G, Brooks III CL. Assessing scoring functions for protein-ligand interactions. J Med Chem 2004;47:3032–47.
[81] Böhm HJ. LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. J Comput Aided Mol Des 1992;6:593–606.
[82] Charifson PS, Corkery JJ, Murcko MA, Walters WP. Consensus scoring: a method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins. J Med Chem 1999;42:5100–9.
[83] Bissantz C, Folkers G, Rognan D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. J Med Chem 2000;43:4759–67.
[84] Ladbury JE. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. Chem Biol 1996;3:973–80.
[85] Brik A, Wong CH. HIV-1 protease: mechanism and drug discovery. Org Biomol Chem 2003;1:5–14.
[86] Verdonk KL, Chessari G, Cole JC, Hartshorn MJ, Murray CW, Nissink JW, et al. Modeling water molecules in protein-ligand docking using GOLD. J Med Chem 2005;48:6504–15.
[87] Scior T, Bender A, Tresadern G, Medina-Franco JL, Martínez-Mayorga K, Langer T, et al. Recognizing pitfalls in virtual screening: a critical review. J Chem Inf Model 2012;52:867–81.
[88] Moitessier N, Englebienne P, Lee D, Lawandi J, Corbell CR. Towards the development of universal, fast, and highly accurate docking/scoring methods: a long way to go. Br J Pharmacol 2008;153:S7–26.
[89] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, et al. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 1998;19:1639–62.
[90] Knegtel RM, Kuntz ID, Oshiro CM. Molecular docking to ensembles of protein structures. J Mol Biol 1997;266:424–40.
[91] Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–89.
[92] Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, et al. Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. J Med Chem 2006;49:6177–96.
[93] Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein–ligand docking using GOLD. Proteins 2003;52:609–23.
[94] Abagyan R, Totrov M, Kuznetsov D. ICM—A new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem 1994;15:488–506.
[95] Carrasco MP, Gut J, Rodrigues T, Ribeiro MHL, Lopes F, Rosenthal PJ, et al. Exploring the molecular basis of Qo bc1 complex inhibitors activity to find novel antimalarials hits. Mol Inf 2013;32:659–70.
[96] Berne S, Podobnik B, Zupanec N, Novak M, Kraševec N, Turk S, et al. Virtual screening yields inhibitors of novel antifungal drug target, benzoate 4-monooxygenase. J Chem Inf Model 2012;52:3053–63.
[97] Diao Y, Lu W, Jin H, Zhu J, Han L, Xu M, et al. Discovery of diverse human dihydroorotate dehydrogenase inhibitors as immunosuppressive agents by structure-based virtual screening. J Med Chem 2012;55:8341–9.

[98] Chan FY, Sun N, Neves MA, Lam PC, Chung WH, Wong LK, et al. Identification of a new class of FtsZ inhibitors by structure-based design and in vitro screening. J Chem Inf Model 2013;53:2131–40.
[99] Hartenfeller M, Schneider G. De novo drug design. Methods Mol Biol 2011;672:299–323.
[100] Schneider G, editor. De novo molecular design. Wiley-VCH; 2013..
[101] Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today 2004;9:430–1.
[102] Schneider G, Fechner U. Computer-based de novo design of drug-like molecules. Nat Rev Drug Discov 2005;4:549–663.
[103] Rodrigues T, Schneider G. Flashback forward: reaction-driven de novo design of bioactive compounds. Synlett 2014;25:170–8.
[104] Honma T, Hayashi K, Aoyama T, Hashimoto N, Machida T, Fukasawa K, et al. Structure-based generation of a new class of potent Cdk4 inhibitors: new de novo design strategy and library design. J Med Chem 2001;44:4615–27.
[105] Nugiel DA, Krumrine JR, Hill DC, Damewood Jr JR, Bernstein PR, Sobotka-Briner CD, et al. De novo design of a picomolar nonbasic 5-HT1B receptor antagonist. J Med Chem 2010;53:1876–80.
[106] Pierce AC, Rao G, Bemis GW. BREED: generating novel inhibitors through hybridization of known ligands. application to CDK2, p38, and HIV protease. J Med Chem 2004;47:2768–75.
[107] Jorgensen WL, Ruiz-Caro J, Tirado-Rives J, Basavapathruni A, Anderson KS, Hamilton AD. Computer-aided design of non-nucleoside inhibitors of HIV-1 reverse transcriptase. Bioorg Med Chem Lett 2006;16:663–7.
[108] Wang Y, Lu H, Zhu Q, Jiang S, Liao Y. Structure-based design, synthesis, and biological evaluation of new N-carboxyphenylpyrrole derivatives as HIV fusion inhibitors targeting gp41. Bioorg Med Chem Lett 2010;20:189–92.
[109] Jiang C, Yang L, Wu W-T, Guo Q-L, You Q-D. De novo design, synthesis, and biological evaluation of 1,4-dihydroquinolin-4-ones and 1,2,3,4-tetrahydroquinazolin-4-ones as potent kinesin spindle protein (KSP) inhibitors. Bioorg Med Chem Lett 2011;19:5612–27.
[110] Li WW, Chen JJ, Zheng RL, Zhang WQ, Cao ZX, Yang LL, et al. Taking quinazoline as a general support-nog to design potent and selective kinase inhibitors: application to FMS-like tyrosine kinase 3. ChemMedChem 2010;5:513–16.
[111] Wang R, Gao Y, Lai L. LigBuilder: a multi-purpose program for structure-based drug design. J Mol Model 2000;6:498–516.
[112] Park H, Jeong Y, Hong S. Structure-based de novo design and biochemical evaluation of novel BRAF kinase inhibitors. Bioorg Med Chem Lett 2012;22:1027–30.
[113] Park H, Bahn YJ, Ryu SE. Structure-based de novo design and biochemical evaluation of novel Cdc25 phosphatase inhibitors. Bioorg Med Chem Lett 2009;19:4330–4.
[114] Yuan Y, Pei J, Lai L. LigBuilder 2: a practical de novo drug design approach. J Chem Inf Model 2011;51:1083–91.
[115] Ni S, Yuan Y, Huang J, Mao X, Lv M, Zhu J, et al. Discovering potent small molecule inhibitors of cyclophilin A using de novo drug design approach. J Med Chem 2009;52:5295–8.
[116] Gillet VJ, Newell W, Mata P, Myatt G, Sike S, Zsoldos Z, et al. SPROUT: recent developments in the de novo design of molecules. J Chem Inf Comput Sci 1994;34:207–17.
[117] Gillet V, Johnson AP, Mata P, Sike S, Williams P. SPROUT: a program for structure generation. J Comput Aided Mol Des 1993;7:127–53.
[118] Gillet VJ, Myatt G, Zsoldos Z, Johnson AP. SPROUT, HIPPO, and CAESA: Tools for de novo structure generation and estimation of synthetic accessibility. Perspect Drug Discov 1995;3:34–50.
[119] Ali MA, Bhogal N, Findlay JBC, Fishwick CWG. The first de novo-designed antagonists of the human NK2 receptor. J Med Chem 2005;48:5655–8.
[120] Mok NY, Chadwick J, Kellett KA, Casas-Arce E, Hooper NM, Johnson AP, et al. Discovery of biphenylacetamide-derived inhibitors of BACE1 using de novo structure-based molecular design. J Med Chem 2013;56:1843–52.
[121] Stevens RC, Cherezon V, Katritch V, Abagyan R, Kuhn P, Rosen H, et al. The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. Nat Rev Drug Discov 2013;12:25–34.
[122] Rajagopal S, Rajagopal K, Lefkowitz RJ. Teaching old receptors new tricks: biasing seven transmembrane receptors. Nat Rev Drug Discov 2010;9:373–86.
[123] Wise A, Gearing K, Rees S. Target validation of Gprotein coupled receptors. Drug Discov Today 2002;7:235–46.
[124] Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov 2002;1:727–30.
[125] Lewell XQ, Judd DB, Watson SP, Hann MM. RECAP—Retrosynthetic combinatorial analysis procedure: a powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. J Chem Inf Comput Sci 1998;38:511–22.
[126] Schneider G, Lee M-L, Stahl M, Schneider P. De novo design of molecular architectures by evolutionary assembly of drug-derived building blocks. J Comput Aided Mol Des 2000;14:487–94.
[127] Rogers-Evans M, Alanine AI, Bleicher KH, Kube D, Schneider G. Identification of novel cannabinoid receptor ligands via evolutionary de novo design and rapid parallel synthesis. QSAR Comb Sci 2004;23:426–30.
[128] Alig L, Alsenz J, Andjelkovic M, Bendels S, Bénardeau A, Bleicher K, et al. Benzodioxoles: novel cannabinoid-1 receptor inverse agonists for the treatment of obesity. J Med Chem 2008;51:2115–27.
[129] Fechner U, Schneider G. Flux (1): a virtual synthesis scheme for fragment-based de novo design. J Chem Inf Comput Sci 2006;46:699–707.
[130] Fechner U, Schneider G. Flux (2): comparison of molecular mutation and crossover operators for ligand-based de novo design. J Chem Inf Model 2007;47:656–67.
[131] Schüller A, Suhartono M, Fechner U, Tanrikulu Y, Breitung S, Scheffer E, et al. The concept of template-based de novo design from drug-derived molecular fragments and its application to TAR RNA. J Comput Aided Mol Des 2008;22:59–68.
[132] Proschak E, Sander K, Zettl H, Tanrikulu Y, Rau O, Schneider P, et al. From molecular shape to potent bioactive agents II: fragment-based de novo design. ChemMedChem 2009;4:45–8.
[133] Rodrigues T, Kudoh T, Roudnicky F, Lim YF, Lin YC, Koch CP, et al. Steering target selectivity and potency by fragment-based de novo drug design. Angew Chem Int Ed 2013;52:10006–9.

[134] Rodrigues T, Roudnicky F, Koch CP, Kudoh T, Reker D, Detmar M, et al. De novo design and optimization of Aurora A kinase inhibitors. Chem Sci 2013;4:1229–33.
[135] Spänkuch B, Keppner S, Lange L, Rodrigues T, Zettl H, Koch CP, et al. Drugs by numbers: reaction-driven de novo design of potent and selective anticancer leads. Angew Chem Int Ed 2013;52:4676–81.
[136] Hartenfeller M, Eberle M, Meier P, Nieto-Oberhuber C, Altmann KH, Schneider G, et al. A collection of robust organic synthesis reactions for in silico molecule design. J Chem Inf Model 2011;51:3093–8.
[137] Hartenfeller M, Eberle M, Meier P, Nieto-Oberhuber C, Altmann KH, Schneider G, et al. Probing the bioactivity-relevant chemical space of robust reactions and common molecular building blocks. J Chem Inf Model 2012;52:1167–78.
[138] Hartenfeller M, Zettl H, Walter M, Rupp M, Reisen F, Proschak E, et al. DOGS: reaction-driven de novo design of bioactive compounds. PLoS Comput Biol 2012;8:e1002380.
[139] Schneider G, Geppert T, Hartenfeller M, Reisen F, Klenner A, Reutlinger M, et al. Reaction-driven de novo design, synthesis, and testing of potential type II kinase inhibitors. Future Med Chem 2011;3:415–24.
[140] Graczyk PP. Gini coefficient: a new way to express selectivity of kinase inhibitors against a family of kinases. J Med Chem 2007;50:5773–9.
[141] Reutlinger M, Rodrigues T, Schneider P, Schneider G. Combining on-chip synthesis of a focused combinatorial library with computational target prediction reveals imidazopyridine GPCR ligands. Angew Chem Int Ed 2014;53:582–5.
[142] Reutlinger M, Rodrigues T, Schneider P, Schneider G. Multi-objective molecular de novo design by adaptive fragment prioritization. Angew Chem Int Ed 2014;53(16):4244–8.
[143] Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 2012;40(database issue):D1100–7.
[144] Hiss JA, Reutlinger M, Koch CP, Perna AM, Schneider P, Rodrigues T, et al. Combinatorial chemistry by ant colony optimization. Future Med Chem 2014;6:267–80.
[145] DeSimone RW, Currie KS, Mitchell SA, Darrow JW, Pippin DA. Privileged structures: applications in drug discovery. Comb Chem High Throughput Screen 2004;7:473–94.

CHAPTER

7

Fragment-Based Drug Discovery

Venkata Velvadapu[1], *Bennett T. Farmer*[2] *and Allen B. Reitz*[1]

[1]**Fox Chase Chemical Diversity Center, Inc., Doylestown, PA, USA;**
[2]**Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA**

OUTLINE

Learn from yesterday; live for today; hope for tomorrow. The important thing is not to stop questioning. ***Albert Einstein***

Fragment based drug discovery (FBDD) has played an important role in the past decade, helping in many cases to streamline the complex process of drug discovery from target validation to new drug application. A total of eighteen drug candidates discovered by FBDD have been advanced to clinical trials so far, and Zelboraf® (vemurafenib, PLX4032) is the first FDA-approved drug discovered using fragment-based approaches. X-ray crystallography or NMR spectroscopy was used to identify thirteen of the eighteen FBDD drug candidates.

I. LIGAND–PROTEIN INTERACTIONS: FIRST PRINCIPLES

When Paul Ehrlich conceived of a drug as a magic bullet interacting as a key in a lock, he was not far from understanding the reality of ligand–protein interactions. Although target proteins are flexible and can adopt one or more of a manifold of induced conformations, binding sites on proteins have evolved to recognize a limited number of endogenous modulators and substrates and to exclude others.

A. Binding Energy as the Sum of the Parts

The Gibbs free energy (ΔG) of multivalent ligand–protein binding is the sum of the energies involving each of the substructures or fragments that comprise the ligand [1,2]. To the extent that we understand ligand–protein interactions on as small a scale as possible, we can design inhibitors and modulators from first principles more easily than the hit-or-miss approach implicit in high-throughput screening (HTS, see Figure 7.1). Fragments are typically clustered in diverse, low molecular weight libraries. Non-FBDD HTS screening libraries can be considered as combinations of fragments in various ways, but only sampling a small fraction of the total possible diversity space because of their larger molecular weight range. As ligands become larger and more complex, the probability of finding useful information for any randomly selected compound is exponentially smaller [3]. As hits are identified from HTS, typically with 1–10 μM potency, the individual fragments encompassed within those hits are typically not optimized for optimal interaction with the biological target. In cases where it is possible to screen the fragments separately in advance, one can design more efficient ligands *a priori* by further synthetic manipulation such as by linking different fragments together. Therefore, less complex and smaller molecules are better starting points early on in the drug discovery process. This is the essence of FBDD – determining which smaller molecular substructures or fragments interact optimally with targets of interest and how they bind, and then using that information to obtain better hits and leads for further consideration [4–17]. FBDD represented a paradigm shift in understanding the lead generation process in drug discovery, and is an attempt to get more information rapidly while doing the same amount of work overall.

B. Historical Development

Automation and miniaturization of biological assays led to the implementation of HTS in drug discovery in the early 1990s at most of the major pharmaceutical companies using their large internal compound libraries. With the advent of combinatorial chemistry, even larger numbers of compounds were synthesized and screened to determine activity against a variety of biological targets. Diversity in corporate screening libraries varies from

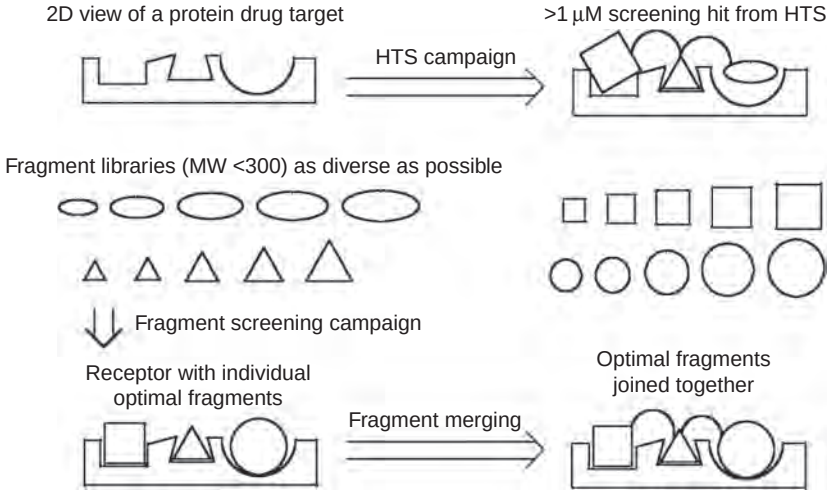


FIGURE 7.1 FBDD general principles.

small molecule fragments to the combination of many such small fragments linked together to form large and complex structures, prepared in the context of different programs. Libraries for HTS vary in the range of 10^4–10^6 compounds, and hit activity is often in the potency range of 1–10 μM, which is then followed by hit to lead (HTL) medicinal chemistry for further optimization [18,19].

However, many compound libraries prepared using combinatorial chemistry incorporate common chemical features suitable for automated synthesis, which limited overall diversity. Hits generated through HTS are often overly lipophilic with poor thermodynamic solubility. Many HTS libraries represent a relatively small fraction of possible chemical space, thus limiting the chance to identify reasonable starting points for lead optimization [20]. In addition, significant cost and time has to be dedicated to creating and enriching the diversity and quality of these libraries. Also, HTS demands a validated assay system, operational as an industrial process with a reasonable throughput and sufficient miniaturization to lower the cost per assay is often challenging and require huge amounts of resources that are usually not available for smaller companies.

During the 1990s, researchers at Abbott demonstrated structure–activity relationship development by nuclear magnetic resonance (SAR by NMR), which enabled the discovery of active fragments to complement those obtained by HTS [21]. High- or medium-throughput X-ray crystallography of compound libraries against protein targets were also initiated in 2000 in seminal work at Abbott [22]. Both of these techniques have provided valuable insight in early drug discovery research, with FBDD evolving to complement HTS in an attempt to discover more chemically-tractable smaller hits prior to starting chemical synthesis of analog libraries [6–20].

C. First Principles: Ligand Efficiency

Ligand efficiency or LE is a powerful tool to rank order hits from any screening campaign and is defined as:

LE = (pKi or pIC_{50})/(N_{hev} or number of heavy atoms) [23]
BEI = (p*Ki* or p*Kd*)/MW
SILE = $pIC_{50}/(N_{hev})^{0.3}$
LLE = p*Ki* − Log *P* (or Log D)
LELP = log *P*/LE

A retrospect analysis performed over a wide range of protein targets and ligands by Reynolds et al has shown that protein–ligand binding affinities are strongly influenced and vary by molecular size [23]. Binding Efficiency (BE), Fit Quality (FQ), and Size Independent Ligand Efficiency (SILE) can be used as alternate efficiency indices, while Ligand Lipophilic Efficiency (LLE) and Ligand-Efficiency-Dependent Lipophilicity (LELP) can be used as lipophilicity guiding indices. These metrics can be used to assess the relative merits of hits after screening and guide further optimization [24].

Changes in LE for analogs in lead optimization as measured by effects on MW as potency is increased for eighteen FBDD programs was investigated by Ferenczy and co-workers [24]. Biological data from the optimized compound and for one or several analogs in the optimization process were considered (Figure 7.2). The effectiveness of optimization was gauged by the slope of the line, as a smaller slope represented a more efficient process of optimization where there is greater improvement of affinity by a smaller increase in the molecular weight. The slope varied from values of 10.3 to 138.5, with 61.4 being the average with a standard deviation of 34.3. The broad extent of the variation of the slope suggests that optimization will in some cases be more dramatic than for others when using FBDD approaches.

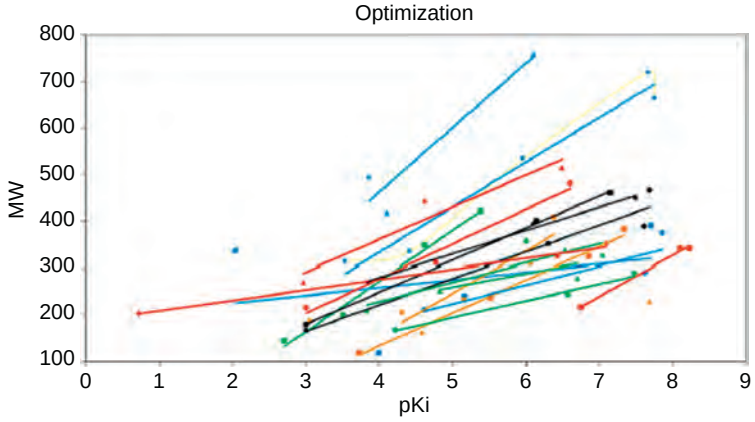


FIGURE 7.2 Ligand affinity (pKi) vs molecular weight for 18 FBDD programs. *Reprinted with permission from Ferenczy GG, Keseru GM. How are fragments optimized? A retrospective analysis of 145 fragments optimizations. J Med Chem 2013;56:2478–86. Copyright (2013) American Chemical Society.*

II. WHAT IS FRAGMENT-BASED DRUG DISCOVERY?

FBDD evolved as a different way to start the lead generation process in drug discovery, with the idea to collect more information while doing the same amount of work overall. FBDD provides the opportunity to design more efficient ligands *a priori*, as smaller molecules with less complex structures are chosen as starting points. These smaller fragments are often chosen in a way so that they provide synthetic handles to support further chemical manipulation and to grow the molecules either by linking, merging, or by substitution, with the aim of increasing binding potency and selectivity and optimizing ADME properties. Smaller molecules have the potential to access those binding sites that are inaccessible to the larger ones. Also, smaller compounds could provide more operational freedom for structural manipulations during multiple optimizations.

A. Overview of FBDD

Drug discovery research since 1990 has been heavily influenced by HTS technology and the combinatorial chemistry preparation of large compound libraries [19]. Large screening libraries and advances in screening automation have made HTS easier to perform. Projects that are initiated via HTS are of necessity focused on limited functional group manipulations based on the nature of the core scaffold and the substitution available or allowed [20]. Though many large, hydrophobic compounds—whose overall characteristics were not drug-like—emerged as hits, they did not serve as good lead generation starting points for lead optimization, and HTS in general has not yielded a high rate of return in terms of NMEs approved for therapy [25,26]. Because of this, consideration of drug-likeness of compounds in libraries was stressed to a much greater extent, and it was recognized that many inappropriate compounds are found in current screening collections that often result in false positives that waste time and effort during follow up. Hann et al related the complexity of a compound to the probability that it will experience a binding event [3]. In short, because the compounds in the current screening collections were both highly complex and overly hydrophobic, being derived mainly from previous LO campaigns, HTS provided only limited success in identifying novel hits that could be progressed into lead candidates, and then into marketed drugs.

One of the keys in FBDD is to use alternate biophysical read-outs of the binding of ligands with target proteins suitable for detecting low-affinity interactions to complement biochemical and functional assays. A seminal article by Fesik et al in 1996 coined the phrase "SAR by NMR," using NMR spectroscopy, as such a method is not dependent upon the detection of biological activity [21]. The SAR by NMR approach is to screen small, highly polar fragments, as opposed to larger members of screening libraries, and then to create lead candidates through the linking of multiple independently optimized fragments. An early example of the FBDD approach was the development of potent inhibitor **1** against stromelysin, where NMR screening had identified that **2** and **3** bound to stromelysin in two distinct, neighboring locations, namely the catalytic site and the S1′ pocket, respectively (Figure 7.3) [27].

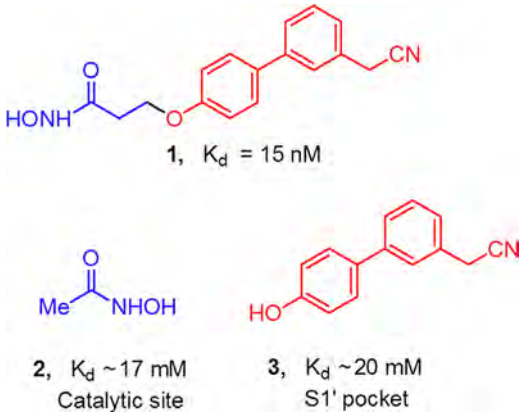


FIGURE 7.3 Fragment merging with dramatic improvement in potency.

B. The Concept of FBDD

Many HTS hits that are sufficiently potent to be considered further and the majority of lead candidates during lead optimization represent complex molecular structures comprised of multiple, interconnected ring systems, onto which any number of substituents are grafted. These structures are complex because of their size, extent of conformational freedom, and array of diverse chemical substitution. A careful analysis of such structures reveals that they can often be deconstructed into a set of chemical building blocks, such as heterocyclic and phenyl rings, and linkers comprised of amide, urea, ketone, and methylene functionalities (Figure 7.4). From these basic building blocks and linkers, one can envision constructing molecular structures that present only a limited number of optimized pharmacophores and degrees of conformational freedom.

These basic starting structures are called fragments, and Figure 7.4 shows hits generated from HTS compared to those from FBDD [5]. Smaller fragments with high ligand efficiency are more relatively useful as starting points in SAR development, although with fewer points of interaction than HTS hits with lower ligand efficiency and more points of interaction.

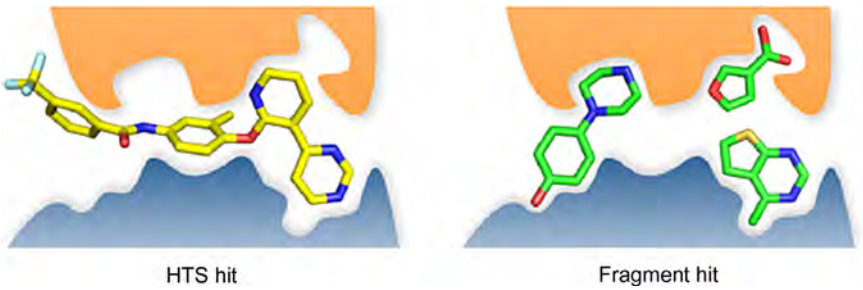


FIGURE 7.4 Interactions made by HTS hits vs. fragment hits where smaller fragments have better overall LE. *Reprinted with permission from Scott DE, Coyne AG, Hydson SA, Abell C. Fragment-based approaches in drug discovery and chemical biology. Biochemistry 2012;51:4990–5003. Copyright (2013) American Chemical Society.*

C. Fragments in FBDD

A fragment is a small molecule characterized by researchers at Astex as falling within the Rule of 3™ [28,29].

- Molecular weight < 300 Da
- Number of hydrogen bond donors ≤ 3
- Number of hydrogen bond acceptors ≤ 3
- cLogP (predicted) ≤ 3
- Number of rotatable bonds ≤ 3
- Polar surface area (PSA) ≤ 60 Å^2

Although FBDD is simple in principle, many factors play a significant role in converting fragments to leads, such as synthetic tractability, location of the binding site relative to the active site or the potential for allosteric modulation, and whether the particular binding site is unique to the original target protein or present in many different proteins [30].

The process of FBDD starts with the selection of a library of fragments. Common techniques used to detect a fragment binding event are NMR [31,32], surface plasmon resonance (SPR) [33,34], mass spectrometry (MS) [35,36], and X-ray crystallography [22,37]. X-ray crystallography is the only general technique that can reliably demonstrate the binding location and orientation of the fragment hit. While NMR can achieve these two outcomes in a step-wise fashion, X-ray crystallography provides both at once. Further stages in the FBDD process are shown in Figure 7.5, which shows it may take two to three years for any particular FBDD project to achieve validated or advanced lead status or entry into preclinical development.

D. Fragments Hits vs. HTS Hits

FBDD and HTS aim to discover lead candidates that have the desired potency with favorable selectivity and ADME properties. HTS seeks to identify one or more mature chemical starting points for hit-to-lead medicinal chemistry, whereas FBDD seeks to incrementally construct leads, often starting with less initial potency but equivalent or better LE. ADME (Adsorption, Distribution, Metabolism, and Excretion) properties play a very important role in determining the fate of the compounds *in vivo*. Metabolically liable and chemically reactive

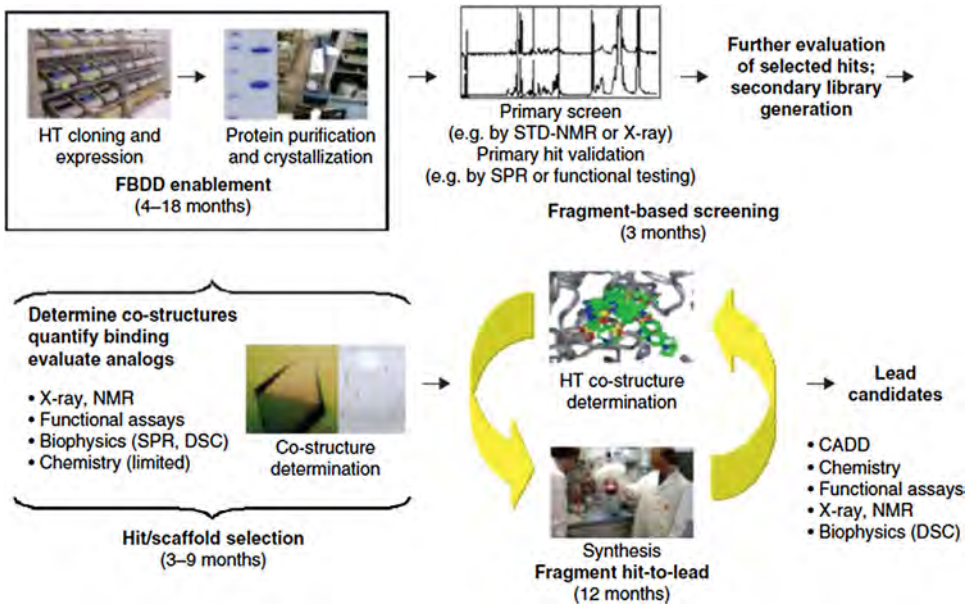


FIGURE 7.5 FBDD workflow with predicted timelines.

groups can be avoided during fragment elaboration in the hit-to-lead stage, unlike in complex HTS hits where they are already incorporated in structures from screening libraries.

III. CREATION AND ANALYSIS OF FBDD LIBRARIES

A. Fragment Library Design

Typical fragment libraries consist of 100−1,000 members and are generally much smaller than libraries used in HTS campaigns. Even with this lower number, greater diversity of structure and topology for the fragment MW range can often be achieved relative to the larger libraries. The potency of fragments upon screening is often weak (in the mM range), which can be dramatically improved during SAR development since a linear relationship frequently exists between MW and binding affinity [38].

Fragment libraries are typically selected with specific attributes in mind. For example, certain techniques require covalent attachment to either protein targets or solid support, which necessitates the incorporation of suitable functionality for that purpose, such as the sulfhydryl group. Some libraries for X-ray crystallography are designed to incorporate heavy atom substitution, such as with bromine, in order to solve the X-ray structures more rapidly [39]. Computational methods have been developed by both pharmaceutical companies and academic groups to prescreen fragment libraries as a way to select members for further study and consideration [40]. Hubbard et al utilized an *in silico* screening method to analyze fragment screens against twelve diverse targets to direct fragment library design [41]. A significant number of fragments overlapped as active compounds irrespective of target. For example, hits for protein−protein interactions tended to be more hydrophobic and heavier than fragment hits from other programs. Of the fragments hits against three kinases, 11 percent were common and at least 52 percent were unique to each kinase screened. An *in silico* screen using molecular graph theory to design fragment libraries was used by Nunez et al at Abbott to identify fragments against human trypsin [42]. Similarly, Hung et al have included non-aromatic sp^3-rich fragments in their fragment library, to generate fragment hits with high topological diversity [43]. It is a common strategy to create fragment libraries that bear representative substructures known to have favorable drug-like properties or are found to a high degree in currently marketed drugs [39]. Chemical novelty is not as important a criterion for inclusion in a primary fragment library, because unique composition of matter suitable for patent protection would be added later during elaboration and merging hit to lead medicinal chemistry.

B. Analysis of Fragments

Once a fragment has been determined to bind to a particular target protein, the magnitude of its binding must be established. Functional inhibition is routinely determined on initial fragment hits, although it is generally understood that interpretation of weak inhibition ($>10\,\mu M$) can often be confounded due to artifacts such as aggregation. The quality of binding is assessed based on stoichiometry, dose dependence, and location and orientation of the fragment hit. However, a fragment hit may exhibit a component of binding that involves one or more desired pockets or regions on a target protein. Competitive displacement experiments, using a potent inhibitor known to occlude the desired pockets or regions, can be used to determine sites of interaction [44].

X-ray crystallography and—to a lesser extent—NMR spectroscopy are the only general techniques that can reliably determine the binding location and orientation of fragment. Insight into the orientation of binding of a fragment hit helps to identify synthetic opportunities for accessing a desired neighboring pocket or region, or for linking with a different fragment hit that already occupies that neighboring pocket or region.

The crystals to be used for X-ray crystallography in an FBDD campaign must present a higher diffraction quality than is typically required by conventional structure-based drug design. A minimum resolution of <2.5 Å, typically on the order of 2 Å, is required to reliably orient a fragment that is bound to the target protein, because fragments are deliberately chosen to be small in size and therefore—in contrast to most HTS hits and LO compounds—typically present a rather compact and minimally asymmetric 3D structure. In contrast, reliable orientation of an HTS hit can often be achieved at a resolution of 3.2 Å.

Previously, prosecuting a greater number of less validated targets has been promulgated as a way to minimize the downside risk to the HTS/HTL approach. In contrast, the downside risk to the FBDD approach can be minimized by prosecuting a smaller number of targets that have achieved at least strong preclinical therapeutic validation. Such targets virtually guarantee that any quality LO compound will proceed into development, justifying the significant up-front resources required by the FBDD approach. Such targets are also typically pursued at many pharmaceutical companies and have many associated patents describing a sundry of inhibitor chemotypes and scaffolds, typically derived either directly from HTS/HTL campaigns or from patent-busting operations on previous HTS/HTL-derived compounds. Under these circumstances, the downside risk inherent to the FBDD approach is further reduced because the very strength of FBDD lies in its proven ability to span effectively a wider chemical space, which plays to chemical novelty. Also more polar chemical starting points can be selected, thereby avoiding at least the initial lipophilicity of many HTS hits that may later on become associated with specific ADME liabilities.

C. The Role of the Medicinal Chemist in FBDD

The fundamental differences between FBDD and HTS also mean that medicinal chemists play a different role in each approach. In the HTS follow up, the medicinal chemist is tasked to determine—often with limited chemistry resources—whether any HTS chemotypes represents chemical matter that can be progressed in a timely fashion into LO. Much data must be analyzed and synthetically challenging analogs are generally avoided. The HTS chemist is part informatician, part medicinal and synthetic chemist. In FBDD, the medicinal chemist plays the role of both a synthetic and structural chemist. The predictive power of the medicinal chemist for rational design is much enhanced by the use of NMR or X-ray structure studies, complemented by any functional assay data that is available on the target of interest. Since many starting fragments are commercially available and possess similar physicochemical properties, it is important to understand the nature of target. Knowledge of toxicity, stability, and bioisosterism can be used in the selection of appropriate fragment libraries [45,46].

IV. FRAGMENT SCREENING METHODS

After the original discovery of weakly biologically active or inactive hits from FBDD, the challenge for the medicinal chemist is to take this information and generate novel leads suitable for development into clinical candidates. Detecting a fragment inhibitory event is accomplished by increasing the concentration of the fragment in the functional assay, typically to 0.2−1.0 mM depending on the maximal solubility allowed. Once a fragment has been detected to bind to the target protein, the magnitude of its binding should be established, if possible.

A. Nuclear Magnetic Resonance

Ligand-based NMR techniques are based on detecting the ligand in a way that can be changed by the binding of that ligand to the protein target. One-dimensional ^{1}H NMR experiments such as CPMG (Carr-Purcell-Meiboom-Gill), water-ligand observed via gradient spectroscopy (WaterLOGSY), and saturation transfer difference (STD) are commonly used to detect ligand binding [5]. In addition to Abbott Laboratories, these techniques have been used in many other pharmaceutical companies such as at Astex Therapeutics, Evotec, Schering-Plough, and Vernalis [47–55].

Saturation transfer difference (STD) NMR is probably the most common NMR technique used to screen fragments for binding to protein targets. STD-NMR takes advantage of the enhanced cross-relaxation rate in a large protein system, and the fact that protons on the bound-state ligand become part of a larger proton–proton relay network [56,57]. Saturation of the protein methyl resonances is therefore quickly propagated through this relay network to the bound-state ligand, leading to a saturation of the proton resonances on the ligand. To this end, protein methyl resonances (~1–0.5 ppm) are chosen because many ligands of pharmaceutical interest have few protons that resonate in that region. In addition, methyl groups serve as relaxation sinks in proteins and are efficient propagators of proton saturation throughout the protein and hence the bound-state ligand. Protein methyl groups are at the terminus of amino acid side-chains protruding into the interior of most binding pockets, and are more likely to be proximal to the bound ligand [58]. STD-NMR studies are carried out such that ligand saturation, effected when the ligand is bound to the protein, is measured on the free ligand, not the bound-state ligand. Many factors such as ligand-to-protein ratio, absolute protein concentration, amino acid composition of the target binding site, and flexibility of the bound-state ligand contribute to the sensitivity of the STD-NMR experiment [58]. Often, competitive binding of the protein to ligand is required to eliminate false positives that can arise from precipitation or denaturation of the protein due to high ligand concentration. Information regarding fragment binding sites, occupancies, and the affinity of binding can be identified by using competitive displacement of a ligand with the known binding mode. STD-NMR methods were used in identifying fragment hits from a small library of ~1,400 fragments, where ten to twelve fragments were employed in the presence and absence of the previously known ATP-competitive ligand PU3 [59]. A total of eighty-two displaceable hits were identified as fragments binding to the ATP pocket of Hsp90 by Brough et al [60].

Researchers at Abbott also utilized NMR methods to identify fragment hits on the protein stromelysin, which is implicated in arthritis and tumor metastasis, whereas HTS screening of 115,000 compounds failed to generate lead compounds with potencies better than 10 μM [27]. Acetylhydroxamic acid (**2**) binds the catalytic zinc site, preventing metalloproteinase-3 (MMP-3) from autolytic degradation, with a dissociation constant of 17 mM (Figure 7.6). Upon selective screening of fragment libraries, a series of biphenyl compounds such as **4** were identified, which was further confirmed as the best compound after the synthesis of biaryl fragments that bound stromelysin. NMR was utilized to solve the structure of these fragments bound to stromelysin, which not only showed the chelation of **2** to the active site zinc but also confirmed a binding pocket that accommodated the biaryl functionality. Thus, by linking the two fragments, **5** was designed and prepared, which showed a good increase in potency and ligand efficiency. NMR-based structure determination provided evidence to support a particular binding orientation for compound **5**. These studies coupled with further SAR development led to the discovery of **6** (ABT-518) as a lead candidate, which was tested in Phase I trials for the treatment of cancer [61,62].

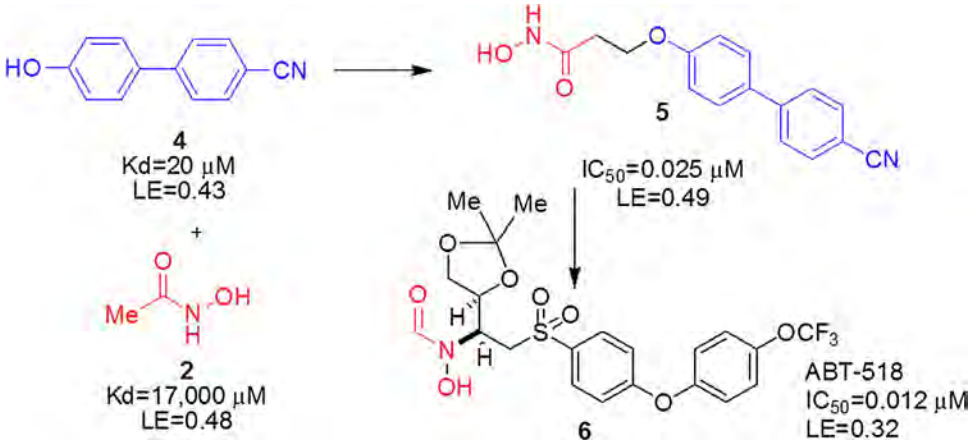


FIGURE 7.6 SAR by NMR in the discovery of ABT-518(**6**).

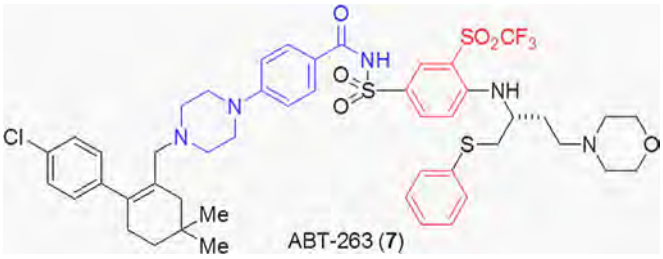


FIGURE 7.7 Fragment linkers were identified using NMR in the discovery of ABT-263(**7**).

NMR methods have also been used in identifying fragments that bind at the protein−protein interface of Bcl-xL. Fragment hits of potency 2 mM were identified from a library of 10,000 fragments. Subsequent synthetic elaboration led to the discovery of a compound whose NMR based structure bound to Bcl-xL helped in identifying an acyl sulfonamide linker. Further LO methods using this linker led to the discovery of ABT-263(**7**). (Figure 7.7) [63].

B. 2D (protein-based) Screening

Protein-based NMR techniques are more powerful than ligand-based techniques in characterizing ligand binding, but are much less widely applicable because the protein must be either ^{15}N or ^{13}C isotopically labeled. In cases where both these requirements are met, the binding of a ligand can be monitored based on the bound-state perturbation of the chemical shifts of specific proton and heteronuclear protein resonances. The specific protein resonances whose chemical shifts are thus perturbed serve to localize the ligand binding site, thereby eliminating the need to use competitive displacement experiments for this purpose, which helps to minimize false positives generated due to nonstochiometric, nonspecific binding caused in ligand-based NMR methods [64,65]. $^{1}H-^{15}N$ HSQC was used by Fresh et al to identify fragment hits that bind to a surface pocket of the Y220C p53 mutant [66].

C. X-ray Crystallography

The use of X-ray crystallography in FBDD has become more prevalent over the past fifteen years. Having a ligand−protein crystal structure provides both validated fragment hit and structural binding information in a single step [67]. X-ray crystallography is performed in FBDD by soaking and co-crystallization of the ligands with proteins. The hit rate found in typical examples when using X-ray crystallography in FBDD is c. 0.5−10 percent, and the most valuable information is obtained when multiple and unanticipated binding sites are identified. When proteins do not form robust crystals with added substrate or inhibitor, then fragments can be evaluated for their ability to displace a known ligand that co-crystallizes with the protein of interest upon soaking [68].

Though a powerful technique used to generate lead candidates, X-ray crystallography has several practical drawbacks. Only 20−30 percent of soluble proteins are amenable to crystallization, and membrane-bound targets such as G-protein-coupled receptors (GPCRs) and ion channels are generally excluded from screening in this way. It is important to have high resolution crystal structures (<2.5 Å^2) to properly interpret the position or confirmation of ligand binding. In certain cases, ligands can cause conformational changes that can take place during the interaction of the crystalline state of the protein [69,70]. Astex Pharmaceuticals has generated fragment hits using X-ray crystallography and has identified multiple efficient, synthetically tractable small molecules for lead optimization (Figure 7.8) [71]. For example, a screen of 500 fragments was conducted against cyclin-dependant kinase 2 (CDK2) by soaking the apo CDK2 crystals against combinations of four fragments. The CDKs are a family of serine-threonine protein kinases that regulate elements in cell cycle progression. A common ATP kinase binding site was observed in ~30 co-crystal structures generated against the screening hits. Four low molecular weight fragments were identified as weak potency hits with good ligand efficiencies and a suitable synthetic handle for elaboration. Key hydrogen bond interactions to the backbone residues at the hinge region of CDK2 (Glu_{81} and Leu_{83}) were conserved in all the bound fragments.

Having co-structures of fragment hits provides a greater opportunity to select the best such hits to be considered for lead optimization. Suitable vectors to access key regions, synthetic flexibility, and novelty are a few considerations in selecting fragments for further consideration. Compound **8** had an amine group that was suitably functionalized for picking up interactions with other residues on the protein. Hydrophobic substitution via an aryl group resulted in **9** with a 150-fold increase in activity. Further introduction of a sulfonamide at the 4-position yielded **10**, which showed a hydrogen bond interaction with Asp_{86} and a further increase of activity (1.9 μM IC_{50}, LE = 0.43) [71].

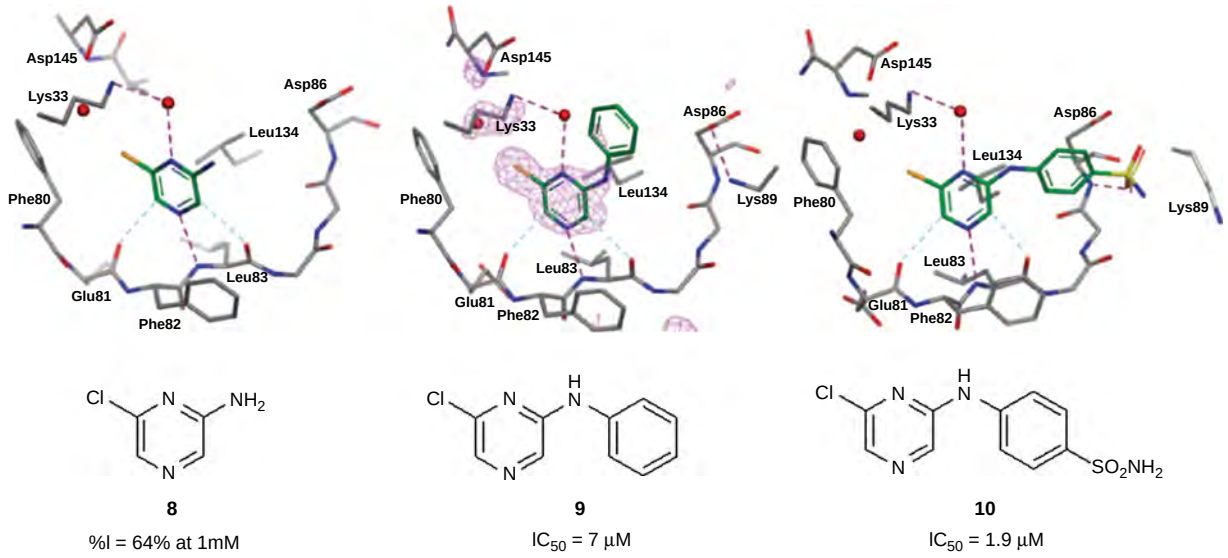


FIGURE 7.8 Lead optimization using fragment **8** guided by X-ray crystallography. *Reprinted with permission from Wyatt PG, et al. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (ABT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. J Med Chem 2008;51:4986. Copyright (2013) American Chemical Society.*

Similarly, substitution of **11** at the 7-postion with a hydrogen bond donor picked up an interaction with the protein back bone at the hinge region through Leu83 (Figure 7.9), increasing the potency 700-fold as in **12**. Further introduction of basic functionality off the 5 position led to **13** with a further increase in potency (0.03 μM IC_{50}, LE = 0.45). An X-ray crystal structure revealed that the 4-amino group formed hydrogen bonds with the carboxylates of Asp_{145} and Asn_{132}.

Structure-based lead optimization improving biological activity and drug-suitability ADME properties led to the discovery of ABT-7519 (**14**) (Figure 7.10), which is currently in Phase I/II clinical trials [4]. X-ray crystallography in FBDD has been extensively employed in additional drug discovery programs [72–75].

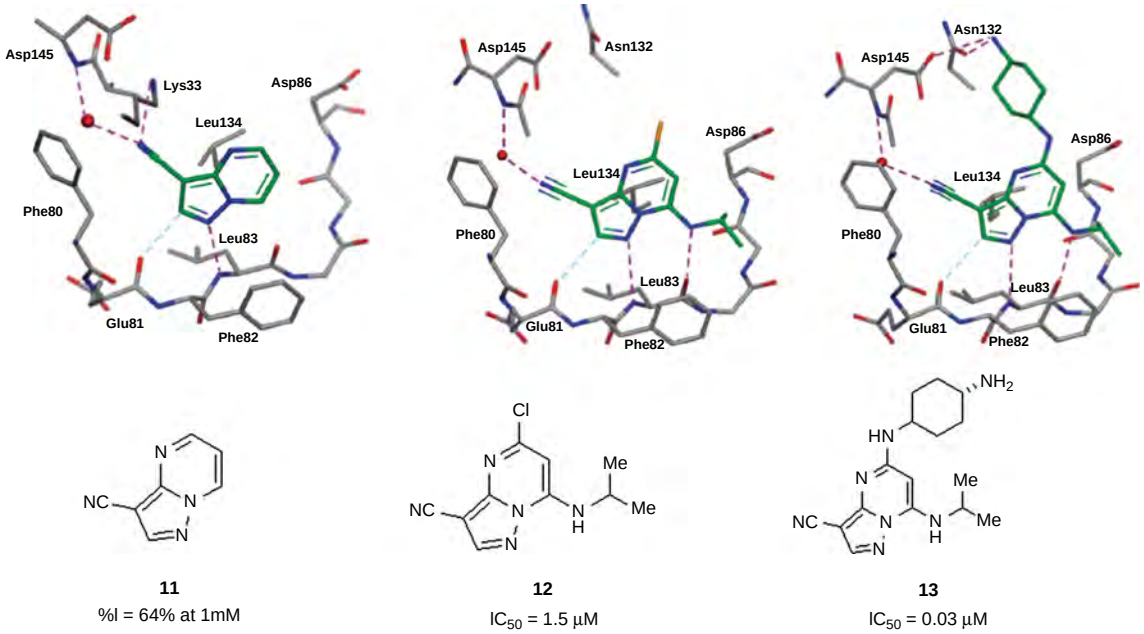


FIGURE 7.9 Lead optimization using fragment **11** guided by X-ray crystallography. *Reprinted with permission from Wyatt PG, et al. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (ABT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. J Med Chem 2008;51:4986. Copyright (2013) American Chemical Society.*

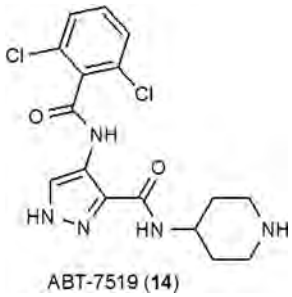


FIGURE 7.10 Structure of ABT-7519 (**14**).

V. OTHER BIOCHEMICAL AND BIOPHYSICAL METHODS

The value of screening fragments to obtain hits of low molecular weight and high LE is somewhat intuitive. However, the key has been to identify and exploit screening technologies that can detect the weak activity or binding that is inherent in the fragment-based approach. In addition to NMR and X-ray crystallography, several other approaches are now adapted for FBDD.

A. Substrate Activity Screening

Substrate activity screening is a fragment screening strategy in which substrates for a particular target protein are identified and then optimized rapidly [76,77]. For example, substrates for the cysteinyl protease cathepsin S bearing a fluorogenic group were optimized for cleavage [78]. Enhancement of substrate activity determines what structural features improve the binding interaction on the C-terminal side. After substrate activity is optimized, the scissile bond is modified to obtain an inhibitor, such as by the use of bioisosteric replacement of the scissile bond. For cathepsin S, a 15 nM Ki inhibitor was identified that was selective ($>$1,000-fold) relative to other cysteinyl proteases such as cathepsins B, L, and K.

B. SPR Spectroscopy

In most SPR spectroscopy cases, a protein is immobilized onto a metal-coated chip and ligands are allowed to flow past on it. Once the binding occurs, the change in the ligand-protein mass causes a change in the reflective property of the metal. An SPR biosensor chip is usually made of gold and has a target biomolecule covalently bound. The solutions of single fragments are then passed over it sequentially and as they bind to the biomolecule an increase in the surface mass is detected in real time [79–83]. Binding affinity and kinetics of fragments with molecular weight $<$100 Da can be calculated from the time-dependent fragment association-dissociation. A known ligand is tethered on the surface and the protein is allowed to complex, then a competing fragment is identified when decomplexation of the protein from the surface is detected. This technique can screen fragment libraries with as little as 25–50 μg of protein [79]. An alternate technology has been reported where the ligands are immobilized and assessed for the binding of the protein [84,85].

C. SAR by Mass Spectroscopy

Mass spectrometry techniques have been used as a fragment screening method where covalent or relatively-strong noncovalent ligand–protein complexation is detected using soft ionization electrospray mass spectrometry. In one variant, libraries of sulfhydryl-group-containing (-SH) fragments are incubated with protein targets mutated to incorporate Cys residues near the active site [86,87]. Those that demonstrate affinity for the protein may orient in such a way as to form a covalent disulfide bond with neighboring cysteine residues, which are then detected by MS analysis. This approach has a limitation where a sulfhydryl group may itself perturb the binding.

D. Interferometry and Isothermal Titration Calorimetry

Isothermal titration calorimetry is similar to SPR, involving changes in refractive index after binding to fragment leads [88,89]. ITC allows determination of the enthalpic and entropic contributions of a ligand when it

binds to the protein and significant amounts of energy is released. Greater ligand efficiency is observed for fragments that bind largely via enthalpic interactions. This technique requires larger quantities of protein compared to other techniques, and has a low throughput [90].

E. Virtual Screening

Computational methods can be used to predict structural information and understand the molecular interactions of ligands and proteins. Several companies (BioLeap, BiosolveIT, and MEDIT) extensively use computational methods for FBDD [91]. A range of scoring functions and docking methods examine the binding of fragments and drug-like compounds with multiple targets. This docking technology was used to predict a pool of compounds that consisted of fragments whose binding with X-ray crystal structures were already known, and ~50 percent of fragments and drug-like compounds were correctly docked compared with their X-ray crystal structures. Factors such as the free energy change of the ligand, conformational entropy, water interactions, desolvation, and protein confirmation changes are to be considered for producing the correct binding mode of the ligand that enables this technology to be used for screening [92].

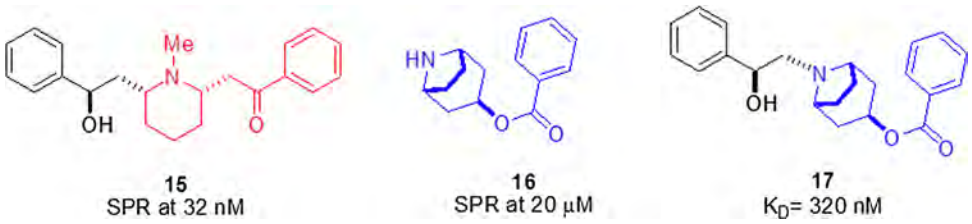
VI. FRAGMENT MERGING/LINKING/GROWING

Once fragment hits are validated, they need to be elaborated to improve potency, selectivity, LE, and novelty. Often X-ray crystal structure data provide the ideal place to start the elaboration campaign. When no crystal structure data are available, a balanced approach of *in silico* docking and validation using biochemical experiments has to take place. Three main approaches that are followed in the elaboration cycle are:

- fragment merging;
- fragment linking; and
- fragment growing.

A. Fragment Merging

Fragment merging involves combining the common structural features of fragments, substructures, or known ligands that complex with the protein [16,93]. A hybrid series can be generated based on different compounds in a series that help to identify important binding patterns and interactions. One such approach is seen in the discovery of fragments targeting acetylcholine binding protein (Figure 7.11).



15
SPR at 32 nM

16
SPR at 20 μM

17
K_D= 320 nM

FIGURE 7.11 Fragment elaboration by fragment merging of **15** and **16** to give **17**.

From the crystal structure, merged fragment **17** was identified with a K_D of 320 nM, identified by overlaying the crystal structure of the natural alkaloid lobeline (**15**, 32 nM) and a benzoate substituted nortropine fragment (**16**, 20 μM) [93]. Other FBDD programs have utilized this approach to identify fragments against Hsp90, Jun N-terminal kinase JNK3, and P13γ kinases [60,32,94].

B. Fragment Linking

Nonoverlapping fragments that bind at different sites (preferentially adjacent sites) can often be linked together to generate a new chemical series. This strategy, known as fragment linking, is the most common

FIGURE 7.12 Fragment elaboration by fragment linking of **18** and **19** to give **20**.

method of fragment elaboration. A compound derived by connecting the fragments should have a Gibbs free energy of binding better than the sum of the individual fragment binding energies. Fragments bind to the target with a particular confirmation, and an ideal linker will allow the fragment in the combined derivative to still adopt the same confirmation. Rigid linkers can reduce degrees of freedom and the entropic penalty upon binding event. On the other hand, flexible linkers can impart a greater number of rotatable bonds, causing an increase in the overall conformational space accessible both in the bound and solution phase, which can lower the selectivity for the particular target [95]. Often this strategy can help to identify hybrid series that are more potent than the fragments from which they were created. Howard et al identified **20**, having a 1.4 nM IC_{50} against thrombin (Figure 7.12), designed by combining *p*-chlorophenyltetrazole **19** (IC_{50} = 330 μM) bound at the S1 site of thrombin to that of amino alcohol **18** (IC_{50} = 100 μM) bound at an adjacent pocket [96].

C. Fragment Growing

It is a common practice to use traditional medicinal chemistry methods of elaboration to build on the insight provided by the crystal structure coordinates of particular fragments bound into protein targets. Hydrogen bond donors or acceptors are used for growing at a particular site on the fragment to exploit possible electrostatic interactions with the protein, and the value of substitution is corroborated by the potency of this elaborated fragment. The choice of fragments and substitution is influenced and guided by the local environment around the binding pocket. Factors such as potency, LE, synthetic tractability, and other physicochemical properties are generally taken into consideration [97]. Finding an optimal balance of substitution during the elaboration step is always a challenge. Though addition of hydrophobic functionality is often preferred for increasing potency, factors such as solubility, lack of aggregation, nonspecific binding, and poor bioavailability are to be considered. Nevertheless, any change in a particular series of compounds should be validated with biological experiments and preliminary ADME studies.

For example, Potter et al disclosed that improved potency (2 μM IC_{50}) was achieved for **22** via fragment growing starting from **21** (Figure 7.13). The initial phenyl imidazole fragment (180 μM IC_{50}) was elaborated by adding chlorine and a phenyl ethylamine group to give **22** with an IC_{50} of 2 μM [98].

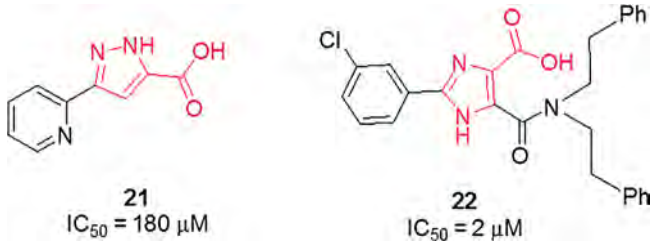


FIGURE 7.13 Fragment elaboration by fragment growing of **21** to give **22**.

VII. FRAGMENT HIT FOLLOW-UP, AND PITFALLS TO AVOID

A. How to Best Reduce False Positives (NMR, MS) and False Negatives (X-ray)

Ligand-based NMR suffers from nonspecific (i.e., nonstoichiometric) fragment binding. Although protein-based NMR is much less sensitive to such false positives in fragment-based screening, this technique is limited to the particular proteins to which it can be practically applied. Functional assays suffer from the effects of both high fragment concentration (0.2−1 mM) in the presence of nM protein concentrations, and from an inability to determine protein integrity after the assay is complete.

Past experience would suggest that the intersection of fragment hits from ligand-based NMR and those from a functional assay screen can produce a set of fragments with a rate of false positives suitable for immediate characterization by X-ray crystallography. However, fragment hits identified by only one screening method have been observed to present a significantly higher rate of false positives, and may therefore benefit from further biophysical characterization prior to crystallography, which may take the form of assessing dose-dependency, estimating the stoichiometry of binding, and estimating the potency of binding. Protein-based NMR, SPR, and differential scanning calorimetry (DSC) are potentially capable of providing information on one or more aspects of the further characterization.

B. Isothermal and Isothermal Titration Calorimetry and Further Secondary Analysis

SPR uses only unlabeled protein but requires that the protein retains binding upon being tethered to the SPR surface. ITC uses only unlabeled protein but may suffer in reliably quantifying weakly binding interactions. Protein-based NMR requires labeled protein and a high-resolution 2D $^{1}H-^{15}N$ or $^{1}H-^{13}C$ correlation spectrum. Both ITC and protein-based NMR methods require substantially more protein than SPR. All three methods can measure the dose-dependency of fragment binding and could therefore in principle yield an estimation of binding potency. For protein-based NMR, such measurements are extremely time consuming and have an implicit sensitivity that is maximally constrained by the extent of a net protein chemical-shift perturbation induced by 100 percent fragment occupancy. For ITC, the overall weak sensitivity of the measurement is often such that even establishing binding at a single fragment concentration is difficult. In general, SPR appears to be the most universally applicable technique for establishing dose-dependent binding or a dose−response binding curve. SPR sensitivity is governed by changes in molecular weight upon fragment binding, and is therefore expected to be more predictable from fragment to fragment for a given protein. Only small amounts of unlabeled protein are required, typically <1 mg for 100 fragments.

In terms of the stoichiometry of fragment binding, the most important requirement is that a sufficient level of fragment binding occurs at a single protein site; how many additional protein sites experience some level of occupancy is largely irrelevant. Therefore, only approaches that can discriminate between different populations of either bound fragments or fragment binding sites can provide such information. Protein-based NMR is one such approach, and competitive displacement ligand-based NMR is another. However, neither SPR nor ITC can provide such discrimination. While it is true that the total signal intensity measured in an SPR experiment can provide some insight as to whether fragment binding is a single or multiple event, this information provides no insight into either the total number of binding events or the relative proportion of each such event.

While the above discussion is focused mainly on approaches to triage a set of primary fragment hits down to select fragments most likely to yield an X-ray co-crystal structure, there is the added complication that the further optimization of any such fragment, whether it yields an X-ray co-structure or not, is dependent on readily being able to measure some parameter that is highly correlated with the binding potency of that fragment. An interesting point to realize is that a comparison of two different fragment co-structures with a given protein does not allow one to infer the relative binding potency of these two fragments. Therefore, developing fragment SAR based solely on co-crystal structures is not possible.

Once binding potencies of $<10\ \mu M$ have been achieved, the ability to develop fragment SAR is more likely, as this is the potency range in which most HTS hits fall already. The challenge lies in developing SAR for those select fragments whose binding potency remains $>10\ \mu M$. As fragments bind to proteins with a dissociation constant of 1 mM, many of them are not soluble at these concentrations. It is important to confirm that the fragments are soluble at the concentrations in the buffer solutions used for screening.

C. Pitfalls to Avoid

Lack of structural integrity of the fragment library can generate a false positive signal, and even 1 percent of a reactive impurity can contribute up to 10 μM in a fragment screen performed at 1 mM. False positives can be obtained with certain reactive functionalities and metal impurities that are known to interfere with the biochemical and biophysical methods used in the screening process. Certain compounds in screening libraries act as oxidizers generating hydrogen peroxide under the assay conditions, inactivating the protein. For example, **23**, **24**, and **25** are known to inhibit PTP1B by generating hydrogen peroxide under certain buffer conditions, which keeps the protein in a reduced state (Figure 7.14) [99,100].

These compounds can be reoxidized by ambient oxygen to generate hydrogen peroxide. Structurally similar compound **26** was not checked for hydrogen peroxide generation and was claimed to be a novel protein−protein inhibitor [101]. Later it was discovered that the activity was due to the generation of hydrogen peroxide [102−104].

Many small molecules tend to aggregate at higher concentrations and can affect the inhibition of proteins. Often, screens are conducted at higher concentrations to identify low affinity binders, and aggregation can lead to a false positive signal. It can be difficult to predict aggregation *a priori* based on the structure, and in fact approved drugs benzyl benzoate (**27**) and mefenamic acid (**28**) form aggregates at higher concentrations (Figure 7.15) [105].

It may be possible to prevent aggregate formation by adding small amounts of non-ionic detergent to the buffers used in compound-library screens [106].

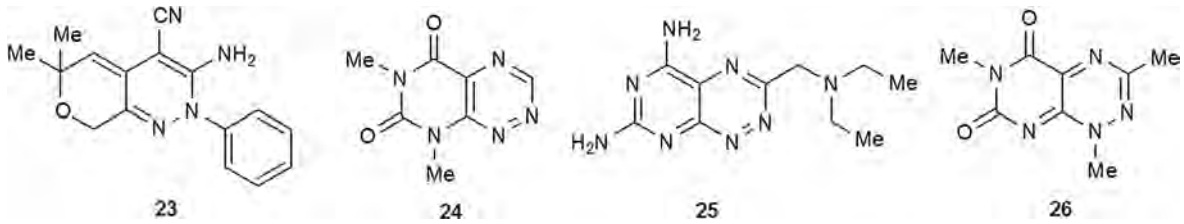


FIGURE 7.14 Agents known to generate hydrogen peroxide under biochemical assay conitions.

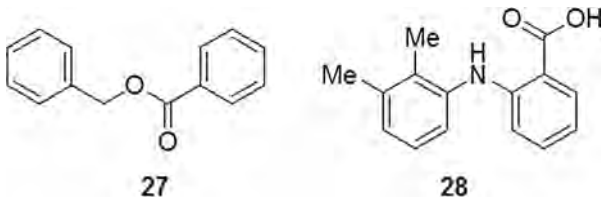


FIGURE 7.15 Drugs known to form aggregates at high concentrations.

VIII. ZELBORAF®, FIRST APPROVED DRUG FROM FBDD

FBDD has achieved a major milestone with the approval of the first drug approved using this method, Zelboraf (PLX4032, **32**, Figure 7.16) [107]. It is a selective kinase B-Raf enzyme inhibitor that is used for the treatment of metastatic melanoma. The goal of the PLX4032 project was to develop inhibitors that were selective for Raf kinases, particularly the oncogenic V600E mutation. Fragment screening initially showed activity against other kinases like Pim-1, but hit to lead medicinal chemistry led to the desired selectivity for the Raf kinases. A selected library of 20,000 scaffolds with MW 150-350 Da were screened at a concentration of 200 μM in functional assays against well-characterized kinases, and the initial hits were used for co-crystallography. 7-Azaindole **29** was elaborated using the bound co-crystal structure that appeared to have multiple binding orientations. Compound **30** was identified as the next fragment with increased potency and single binding mode. Subsequent medicinal chemistry led to the discovery of PLX 4720 and PLX 4032 with good selectivity against B-Raf and the V600E mutant, and were selective compared to other kinases [108]. Zelboraf (PLX4032) is a good example for fragment elaboration where selectivity was achieved during the lead optimization process, which may be the result of the appropriate selection of fragments during FBDD.

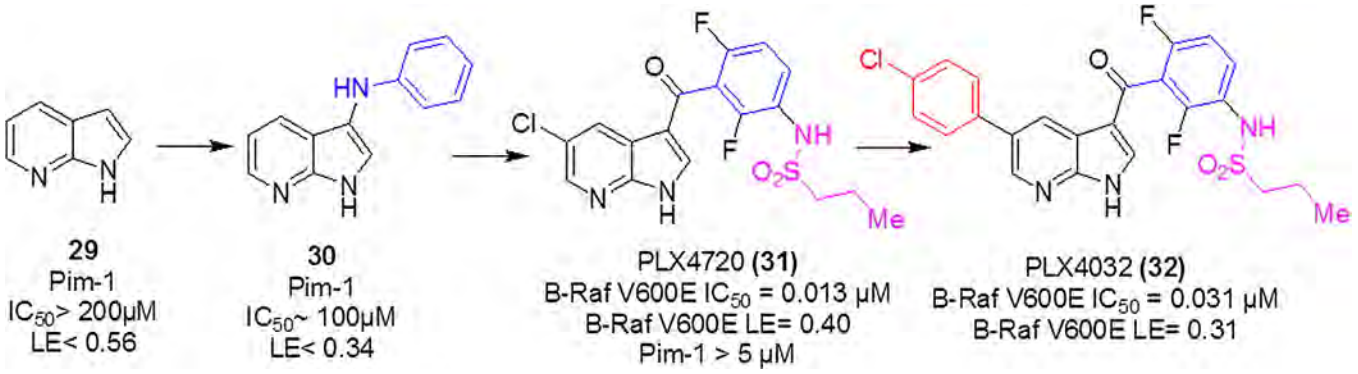


FIGURE 7.16 Zelboraf (PLX4032) as the first FDA approved drug discovered by FBDD.

IX. LIMITATIONS OF FBDD

An important caveat to FBDD is that fragments will not always orient in the same way individually as when combined together in an optimized structure. For example, β-lactamase inhibitor **33** was conceptually deconstructed into fragments **34**–**37**, which were prepared and individually evaluated for functional activity and binding mode by ligand–protein X-ray crystallography (Figure 7.17). Fragments **34**–**37** were found to bind to the protein in very different orientations when compared to the binding mode of **33**. In fact, **34** and **35** induced a conformational change in the protein itself and mapped into a previously unidentified tunnel carboxylate site. Fragment **34** was also seen to adopt a second binding mode that, along with **36**, bound into a different new distal carboxylate recognition site [109].

Low potency fragments are unsuitable for both whole cell screening and kinetic assays as an mM potency would require high fragment concentrations, which can result in false positives. Though widely applicable, FBBD failed to generate validated hits against an antibacterial target shikimate kinase obtained from *M. tuberculosis* [5].

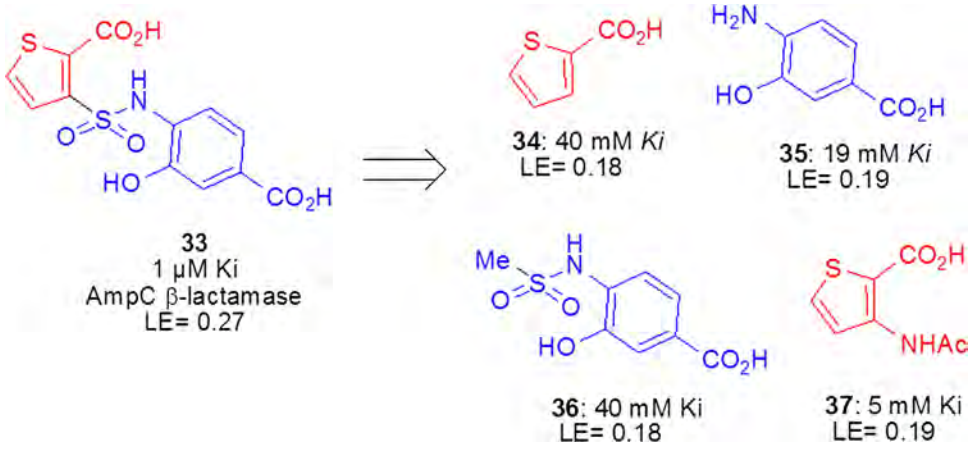


FIGURE 7.17 Elaborated fragment **33** lacked expected orientation as seen by individual fragment **34**–**37**.

X. TRENDS FOR THE FUTURE

FBDD has become an accepted approach at the early stage of many drug discovery programs. The past decade has seen major advances in fragment screening technologies, such as X-ray crystallography, NMR spectroscopy, and SPR. All of these seek to achieve a sustainable delivery of high-quality lead candidates that can be reliably progressed into the lead development and clinical pipeline. While new technologies facilitated fragment screening, the concepts of FBDD are still being applied on complex biological targets. In the future, FBDD will be adapted to more challenging targets, such as oligonucleotides, allosteric inhibitors, and protein–protein interactions.

Fragment libraries from both in-house and commercial sources have to be evaluated carefully. Most of these libraries lack extensive structural diversity and can contain flat heterocyclic molecules with modest overall complexity.

Limitations of NMR and X-ray screening methods are to be kept in mind when evaluating shallow surface exposed and highly hydrophobic pockets. A single-pocket binding site with one large, diffuse pocket is less preferred than a multi-pocket binding site with several smaller, more constrained subpockets. Better understanding of the target binding sites either through NMR or X-ray is needed to evaluate the quality of hits in guiding the lead optimization process. Extensive use of certain fragments, substructures, and substitution patterns has assisted in identifying novel fragment core structures for intellectual property protection.

Though structural data plays a crucial role, the techniques involved have some limitations. Characterization of targets like membrane proteins is difficult, as they are tough to crystallize and subject to functionally-relevant conformational changes not likely to be picked up by FBDD [110]. While identification of novel drugable targets and the application of FBDD to challenging targets is highly innovative, it is often challenged by an unpredictable and low success rate, limited resources, funding, and luck. X-ray crystallography remains the best tool for obtaining structural information of protein–ligand interactions and is an essential tool for FBDD.

While 2012 saw thirty-five new drugs entering the market, these would have entered clinical development three to seven years ago. Many pharmaceutical companies have downsized research operations and have outsourced jobs to low-cost providers as the result of many discovery programs not yielding the expected FDA approvals. This has resulted in greater opportunity for smaller biotech companies, academic labs, and research facilities to contribute and perform with limited resources. The concept of FBDD to start smaller and simpler has been validated, allowing smaller companies to address more versatile and challenging targets in the field of drug discovery.

References

[1] Jencks WP. On the attribution and additivity of binding energies. Proc Natl Acad USA 1981;78:4046–50.
[2] Nakamura CE, Abeles RH. Mode of interaction of β-hydroxy- β-methylglutaryl coenzyme A reductase with strong binding inhibitors: compactin and related compounds. Biochemistry 1985;24:1364–76.
[3] Hann MM, Leach AR, Harper G. Molecular complexity and its impact on the probability of finding leads for drug discovery. J Chem Inf Comput Sci 2001;41:856–64.
[4] Erlanson DA. Introduction to fragment-based drug discovery. Top Curr Chem 2012;317:1–32.
[5] Scott DE, Coyne AG, Hydson SA, Abell C. Fragment-based approaches in drug discovery and chemical biology. Biochemistry 2012;51:4990–5003.
[6] Baker M. Fragment-based lead discovery grows up. Nat Rev Drug Disc 2013;12:5–7.
[7] Murray CW, Verdonk ML, Rees DC. Experiences in fragment-based drug discovery. Trends Pharm Sci 2012;33:224–32.
[8] Murray CW, Rees DC. The rise of fragment-based drug discovery. Nat Chem 2009;1:187–92.
[9] Hubbard RE, Murray JB. Experience in fragment-based lead discovery. Methods Enzymol 2011;493:509–31.
[10] Ciulli A, Abell C. Fragment-based approaches to enzyme inhibition. Curr Opin Biotechnol 2007;18:489–96.
[11] Jahnke W, Flörsheimer A, Blommers MJJ, Paris CG, Heim J, Nalin CM, et al. Second-site NMR screening and linker design. Curr Top Med Chem 2003;3:69–80.
[12] Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat Rev Drug Disc 2007;6:211–19.
[13] Bartoli S, Fincham CI, Fattori D. The fragment-approach: an update. Drug Discov Today Technol 2006;3:425–31.
[14] Zartler ER, Shapiro MJ. Fragonomics: fragment-based drug discovery. Curr Opin Chem Biol 2005;9:366–70.
[15] Rees DC, Cosgreve M, Murray CW, Carr R. Fragment-based lead discovery. Nat Rev Drug Disc 2004;3:660–72.
[16] Erlanson DA, McDowell RS, O'Brien T. Fragment-based drug discovery. J Med Chem 2004;47:3463–82.
[17] Jahnke W, Erlanson DA. Fragment-based approaches in drug discovery. Weinheim: Wiley-VCH; 2006. pp. 1–369.
[18] Davis AM, Keeling DJ, Steele J, Tomkinson NP, Tinker AC. Components of successful lead generation. Curr Top Med Chem 2005;5:421–39.
[19] Michne W. Hit-to-lead chemistry: a key element in new lead generation. Pharmaceutical News 1996;3:19–21.
[20] Honma T. Recent advances in de novo design strategy for practical lead identification. Med Res Rev 2003;23:606–32.
[21] Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. Science 1996;274:1531–4.
[22] Nienaber VL, Richardson PL, Koighofer V, Bouska JJ, Giranda VL, Greer J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. Nat Biotechnol 2000;18:1105–8.
[23] Reynolds CH, Bembenek SD, Tounge BA. The role of molecular size in ligand efficiency. Bioorg Med Chem Lett 2007;17:4258–61.
[24] Ferenczy GG, Keseru GM. How are fragments optimized? A retrospective analysis of 145 fragments optimizations. J Med Chem 2013;56:2478–86.
[25] Brown D, Superti-Furga G. Rediscovering the sweet spot in drug discovery. Drug Disc Today 2003;8:1067–77.
[26] Hird N. Isn't combinatorial chemistry just chemistry? Drug Disc Today 2000;5:307–8.
[27] Hajduk PJ, Sheppard G, Nettesheim DG, Olejniczak ET, Shuker SB, Meadows RP, et al. Discovery of potent non-peptide inhibitors of stromelysin using SAR by NMR. J Am Chem Soc 1997;119:5818–27.

[28] Congreve M, Carr R, Murray C, Jhoti H. A "rule of three" for fragment-based lead discovery?. Drug Disc Today 2003;8:876–7.
[29] Schuffenhauer A, Ruedisser S, Marzinzik AL, Jahnke W, Blommers M, Selzer P, et al. Library design for fragment-based screening. Curr Top Med Chem 2005;5:751–62.
[30] Oblak M, Grdadolnik SG, Kotnik M, Jerala R, Filipic M, Solmajer T. In silico fragment-based discovery of indolin-2-one analogues as potent DNA gyrase inhibitors. Bioorg Med Chem Lett 2005;15:5207–10.
[31] Roberts GCK. Applications of NMR in drug discovery. Drug Disc Today 2000;5:230–40.
[32] Fejzo J, Lepre C, Xie X. Application of NMR screening in drug discovery. Curr Top Med Chem 2003;3:81–97.
[33] Dickopf S, Frank M, Junker HD, Maier S, Metz G, Ottleben G, et al. Custom chemical microarray production and affinity fingerprinting for the S1 pocket of factor VIIa. Anal Biochem 2004;335:50–7.
[34] Gestwicki JE, Hsieh HV, Pitner JB. Using receptor conformational change to detect low molecular weight analytes by surface plasmon resonance. Anal Chem 2001;73:5732–7.
[35] Moy FJ, Haraki K, Mobilio D, Walker G, Powers R, Tabei K, et al. MS/NMR: a structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry, and nuclear magnetic resonance spectroscopy. Anal Chem 2001;73:571–81.
[36] Swayze EE, Jefferson EA, Sannes-Lowery KA, Blyn LB, Risen LM, Arakawa S, et al. SAR by MS: a ligand based technique for drug lead discovery against structured RNA targets. J Med Chem 2002;45:3816–19.
[37] Carr R, Congreve M, Murray C, Rees D. Fragment-based lead discovery: leads by design. Drug Disc Today 2005;10:987–92.
[38] Hajduk PJ. Fragment-based drug design: how big is too big? J Med Chem 2006;49:6972–6.
[39] Siegel MG, Vieth M. Drugs in other drugs: a new look at drugs as fragments. Drug Disc Today 2007;12:71–9.
[40] Gibbs AC, Abad MC, Zhang X, Tounge BA, Lewandowski FA, Struble GT, et al. Electron density guided fragment-based lead discovery of keto hexokinase inhibitors. J Med Chem 2010;53:7979–91.
[41] Chen IJ, Hubbard RE. Lessons for fragment library design: analysis of output from multiple screening campaigns. J Comput Aided Mol Des 2009;23:603–20.
[42] Venhorst J, Nunez S, Kruse CG. Design of a high fragment efficiency library by molecular graph theory. ACS Med Chem Lett 2010;1:499–503.
[43] Hung AW, Ramek A, Wang Y, Kaya T, Wilson JA, Clemons PA, et al. Route to three-dimensional fragments using diversity-oriented synthesis. Proc Natl Acad Sci USA 2011;108:6799–804.
[44] Schade M. NMR fragment screening: advantages and applications. IDrugs 2006;9:110–13.
[45] Meanwell NA. Synopsis of some recent tactical application of bioisosteres in drug design. J Med Chem 2011;54:2529–91.
[46] Riley R. The potential pharmacological and toxicological impact of P450 screening. Curr Opin Drug Disc Dev 2001;4:45–54.
[47] Pellecchia M, Bertini I, Cowburn D, Dalvit C, Giralt E, Jahnke W, et al. Perspectives on NMR in drug discovery: a technique comes of age. Nat Rev Drug Disc 2008;7:738–45.
[48] Pellecchia M, Becattini B, Crowell KJ, Fattorusso R, Forino M, et al. NMR-based techniques in the hit identification and optimization processes. Expert Opin Ther Targets 2004;8:597–611.
[49] Zartler ER, Shapiro MJ. Protein NMR-based screening in drug discovery. Curr Pharm Des 2006;12:3963–72.
[50] Sem DS. NMR-guided fragment assembly. In: Jahnke W, Erlanson DA, editors. Fragment-based approaches in drug discovery. Weinheim: Wiley-VCH; 2006. p. 149–81.
[51] Hajduk PJ, Huth JR, Sun C. SAR by NMR: an analysis of potency gains realized through fragment-linking and fragment-elaboration strategies for lead generation. In: Jahnke W, Erlanson DA, editors. Fragment-based approaches in drug discovery. Weinheim: Wiley-VCH; 2006. p. 181–93.
[52] Zartler ER, Mo H. Practical aspects of NMR-based fragment discovery. Curr Top Med Chem 2007;7:1592–9.
[53] Dalvit C. NMR methods in fragment screening: theory and a comparison with other biophysical techniques. Drug Disc Today 2009;14:1051–7.
[54] Kobayashi M, Retra K, Figaroa F, Hollander JG, Ab E, Heetebrij RJ, et al. Target immobilization as a strategy for NMR based fragment screening: comparison of TINS, STD, and SPR for fragment hit identification. J Biomol Screen 2010;15:978–89.
[55] Gozalbes R, Carbajo RJ, Pineda-Lucena A. Contributions of computational chemistry and biophysical techniques to fragment-based drug discovery. Curr Med Chem 2010;17:1769–94.
[56] Boehm HJ, Boehringer M, Bur D, Gmuender H, Huber W, Klaus W, et al. Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization, a promising alternative to random screening. J Med Chem 2000;43:2663–74.
[57] Mayer M, Meyer B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angew Chem Int Ed 1999;38:1784–8.
[58] Kalk A, Berendsen HJC. Proton magnetic relaxation and spin diffusion in proteins. J Magn Reson 1976;24:343–66.
[59] Taldone T, Sun W, Chiosis G. Discovery and development of heat shock protein 90 inhibitors. Bioorg Med Chem 2009;17:2225–35.
[60] Brough PA, Barril X, Borgognoni J, Chene P, Davies NG, Davis B, et al. Combining hit identification strategies: fragment-based and in silico approaches to orally active 2-aminothieno[2,3-d]pyrimidine inhibitors of the Hsp90 molecular chaperone. J Med Chem 2009;52:4794–809.
[61] Hajduk PJ, Shuker SB, Nettesheim DG, Craig R, Augeri DJ, Betebenner D, et al. NMR-based modification of matrix metalloproteinase inhibitors with improved bioavailability. J Med Chem 2002;45:5628–39.
[62] Wada CK. The evolution of the matrix metalloproteinase inhibitor drug discovery program at Abbott Laboratories. Curr Top Med Chem 2004;4:1255–67.
[63] Oltersdorf T, Elmore SW, Shoemaker AR, Armstrong RC, Augeri DJ, Belli BA, et al. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. Nature 2005;435:677–81.
[64] Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promotes protein stability. Nat Protoc 2007;2:2212–21.

[65] Eakin AE, Green O, Hales N, Walkup GK, Bist S, Singh A, et al. Pyrrolamide DNA gyrase inhibitors: fragment-based nuclear magnetic resonance screening to identify antibacterial agents. Antimicrob Agents Chemother 2012;56:1240–6.
[66] Basse N, Kaar JL, Settanni G, Joerger AC, Rutherford TJ, Fersht AR. Toward the rational design of p53-stabilizing drugs: probing the surface of the oncogenic Y220C mutant. Chem Biol 2010;17:46–56.
[67] Rondeau J-M, Schreuder H. Protein crystallography and drug discovery. In: Wermuth CA, editor. The practice of medicinal chemistry. 2nd ed. New York: Elsevier Academic Press; 2003. p. 417–44.
[68] Hartshorn MJ, Murray CW, Cleasby A, Frederickson M, Tickle IJ, Jhoti H. Fragment-based lead discovery using X-ray crystallography. J Med Chem 2005;48:403–13.
[69] Davis AM, St-Gallay SA, Kleywegt GJ. Limitations and lessons in the use of X-ray structural information in drug design. Drug Disc Today 2008;13:831–41.
[70] Sondergaard CR, Garrett AE, Carstensen T, Pollastri G, Nielsen JE. Structural artifacts in protein-ligand X-ray structures: implications for the development of docking scoring functions. J Med Chem 2009;52:5673–84.
[71] Wyatt PG, Woodhead AJ, Berdini V, Boulstridge JA, Carr MG, Cross DM, et al. Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure based drug design. J Med Chem 2008;51:4986–99.
[72] Hubbard RE, Davis B, Chen I, Drysdale MJ. The SeeDs approach: integrating fragments into drug discovery. Curr Top Med Chem 2007;7:1568–81.
[73] Jhoti H, Cleasby A, Verdonk M, Williams G. Fragment-based screening using X-ray crystallography and NMR spectroscopy. Curr Opin Chem Biol 2007;11:485–93.
[74] Gozalbes R, Carbajo RJ, Pineda-Lucena A. Contributions of computational chemistry and biophysical techniques to fragment-based drug discovery. Curr Med Chem 2010;17:1769–94.
[75] Murray CW, Blundell TL. Structural biology in fragment-based drug design. Curr Opin Struct Biol 2010;20:497–507.
[76] Wood WJL, Patterson AW, Tsuruoka H, Jain RK, Ellman JA. Substrate activity screening: a fragment-based method for the rapid identification of nonpeptidic protease inhibitors. J Am Chem Soc 2005;127:15521–7.
[77] Salisbury CM, Ellman JA. Rapid identification of potent nonpeptidic serine protease inhibitors. Chem Bio Chem 2006;7:1034–7.
[78] Patterson AW, Wood WJL, Hornsby M, Lesley S, Spraggon G, Ellman JA. Identification of selective, nonpeptidic nitrile inhibitors of cathepsin S using the substrate activity screening method. J Med Chem 2006;49:6298–307.
[79] Navratilova I, Hopkins AL. Fragment screening by surface plasmon resonance. ACS Med Chem Lett 2010;1:44–8.
[80] Danielson UH. Fragment library screening and lead characterization using SPR biosensors. Curr Top Med Chem 2009;9:1725–35.
[81] Neumann T, Junker HD, Schmidt K, Sekul R. SPR-based fragment screening: advantages and applications. Curr Top Med Chem 2007;7:1630–42.
[82] Congreve M, Rich RL, Myszka DG, Figaroa F, Siegal G, Marshall FH. Fragment screening of stabilized G protein-coupled receptors using biophysical methods. Methods Enzymol 2011;493:115–36.
[83] Xiang Y, Hirth B, Asmussen G, Biemann HP, Bishop KA, Good A, et al. The discovery of novel benzofuran-2-carboxylic acids as potent Pim-1 inhibitors. Bioorg Med Chem Lett 2011;21:3050–6.
[84] Rich RL, Myszka DG. Kinetic analysis and fragment screening with Fujifilm AP-3000. Anal Biochem 2010;402:170–8.
[85] Kreatsoulas C, Narayan K. Algorithms for the automated selection of fragment-like molecules using single-point surface plasmon resonance measurements. Anal Biochem 2010;402:179–84.
[86] Erlanson DA, Wells JA, Braisted AC. Tethering: fragment based drug discovery. Annu Rev Biophys Biomol Struct 2004;33:199–223.
[87] Oslob JD, Erlanson DA. Tethering in early target assessment. Drug Disc Today: Targets 2004;3:143–50.
[88] Proll F, Fechner P, Proll G. Direct optical detection in fragment-based screening. Anal Bioanal Chem 2009;393:1557–62.
[89] Concepcion J, Witte K, Wartchow C, Choo S, Yao D, Persson H, et al. Label-free detection of biomolecular interactions using biolayer interferometry for kinetic characterization. Comb Chem High Throughput Screen 2009;12:791–800.
[90] Ladbury JE, Klebe G, Freire E. Adding calorimetric data to decision making in lead discovery: a hot tip. Nat Rev Drug Disc 2010;9:23–7.
[91] Zoete V, Grosdidier A, Michielin O. Docking, virtual high throughput screening, and in silico fragment-based drug design. J Cell Mol Med 2009;13:238–48.
[92] Law JMS, Fung DYK, Zsoldos Z, Simon A, Szabo Z, Csizmadia IG, et al. Validation of the SPROUT de novo design program. THEOCHEM 2003;666:651–7.
[93] Edink E, Rucktooa P, Retra K, Akdemir A, Nahar T, Zuiderveld O, et al. Fragment growing induces conformational changes in acetylcholine binding protein: a structural and thermodynamic analysis. J Am Chem Soc 2011;133:5363–71.
[94] Hughes SJ, Millan DS, Kilty IC, Lewthwaite RA, Mathias JP, O'Reilly MA, et al. Fragment based discovery of a novel and selective PI3 kinase inhibitor. Bioorg Med Chem Lett 2011;21:6586–90.
[95] Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45:2615–23.
[96] Howard N, Abell C, Blakemore W, Chessari G, Congreve M, Howard S, et al. Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. J Med Chem 2006;49:1346–55.
[97] Verdonk ML, Rees DC. Group efficiency: a guideline for hits-to-leads chemistry. Chem Med Chem 2008;3:1179–80.
[98] Potter A, Oldfield V, Nunns C, Fromont C, Ray S, Northfield CJ, et al. Discovery of cell-active phenyl-imidazole Pin1 inhibitors by structure-guided fragment evolution. Bioorg Med Chem Lett 2011;20:6483–8.
[99] Guertin KR, Setti L, Qi L, Dunsdon RM, Dymock BW, Jones PS, et al. Identification of a novel class of orally active pyrimido [5,4-3] [1,2,4]triazine-5,7-diamine-based hypoglycemic agents with protein tyrosine phosphatase inhibitory activity. Bioorg Med Chem Lett 2003;13:2895–8.
[100] Tjernberg A, Hallen D, Schultz J, James S, Benkestock K, Bystrom S, et al. Mechanism of action of pyridazine analogues on protein tyrosine phosphatase 1B (PTP1B). Bioorg Med Chem Lett 2004;14:891–5.

[101] Yi F, Regan L. A novel class of small molecule inhibitors of Hsp90. ACS Chem Biol 2008;3:645–54.
[102] Lor LA, Schneck J, McNulty DE, Diaz E, Brandt M, Thrall SH, et al. Simple assay for detection of small-molecule redox activity. J Biomol Screen 2007;12:881–90.
[103] Johnston PA, Soares KM, Shinde SN, Foster CA, Shun TY, Takyi HK, et al. Development of a 384-well colorimetric assay to quantify hydrogen peroxide generated by the redox cycling of compounds in the presence of reducing agents. Assay Drug Dev Technol 2008;6:505–18.
[104] Soares KM, Blackmon N, Shun TY, Shinde SN, Takyi HK, Wipf P, et al. Profiling the NIH small molecule repository for compounds that generate H_2O_2 by redox cycling in reducing environments. Assay Drug Dev Technol 2012;8:152–74.
[105] Seidler J, McGovern SL, Doman TN, Shoichet BK. Identification and prediction of promiscuous aggregating inhibitors among known drugs. J Med Chem 2003;46:4477–86.
[106] Feng BY, Shoichet BK. A detergent-based assay for the detection of promiscuous inhibitors. Nat Protocols 2006;1:550–3.
[107] Tsai J, Lee JT, Wang W, Zhang J, Cho H, Mamo S, et al. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. Proc Natl Acad Sci USA 2008;105:3041–6.
[108] Bollag G, Hirth P, Tsai J, Zhang J, Ibrahim PN, Cho H, et al. Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. Nature 2010;467:596–9.
[109] Babaoglu K, Shoichet BK. Deconstructing fragment-based inhibitor discovery. Nat Chem Biol 2006;2:720–3.
[110] Chen D, Errey JC, Heitman LH, Marshall FH, Ijzerman AP, Siegal G. Fragment screening of GPCRs using biophysical methods: identification of ligands of the adenosine A(2A) receptor with novel biological activity. ACS Chem Biol 2012;7(12):2064–73.

CHAPTER

8

Molecular Variations Based on Isosteric Replacements

Paola Ciapetti[1] *and Bruno Giethlen*[2]

[1]Head of Medicinal Chemistry, Novalyst Discovery, Illkirch, Cedex, France; [2]Novartis Pharma AG, Basel, Switzerland

OUTLINE

Things are not always what they seem; the first appearance deceives many; the intelligence of a few perceives what has been carefully hidden. . . ***Anonymous***

I. INTRODUCTION

In a biologically active molecule, the replacement of an atom or a group of atoms by another one presenting the same physiochemical properties is based on the concept of isosterism. The notion of isosterism was

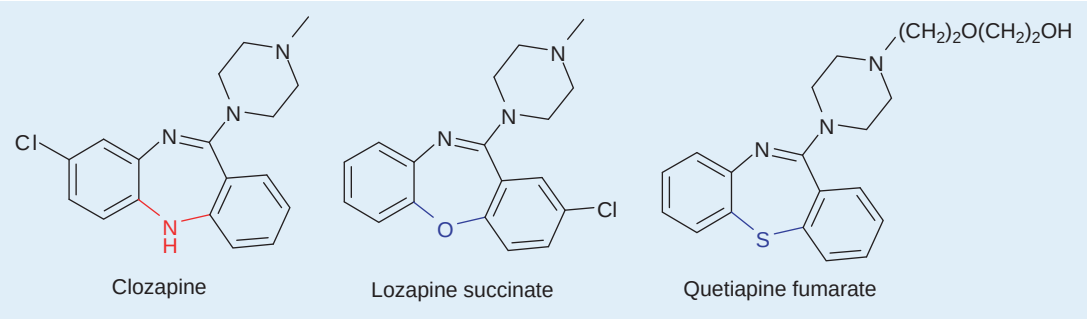


FIGURE 8.1 Isosteric replacements of −NH− with −O− and −S− in clozapine analogs.

introduced in 1919 by Langmuir [1], who was mainly focused on the similarities of electronic and steric arrangement of atoms, groups, radicals, and molecules.

The concept of isosteres was then broadened by Grimm in 1925 with the statement of Hydride Displacement Law, and—further on—Erlenmeyer extended Grimm's classification defining isosteres as atoms, ions, and molecules in which the peripheral layers of electrons can be considered identical. The extensive application of isosterism to modify a part of a biologically active molecule to get another one of similar activity has given rise to the term of "bioisosterism" or nonclassical isosterism. As initially defined by Friedman [2], bioisosteres include all atoms and molecules that fit the broadest definition for isosteres and that elicit the similar biological activity. In medicinal chemistry, the concept of bioisosterism is a research tool of the utmost importance, widely used in analogs design.

The term "analog" is derived from the Greek word αηαλσγια (analogia) and has been used to point out structural and functional similarity. Applied to drugs, an analog of an existing drug shares chemical and therapeutic similarity with the parent compound. This definition implies that three categories of drug analogs can be listed: (a) those presenting chemical and therapeutic similarity; (b) those presenting only similar chemical features; and (c) those eliciting the same pharmacological effect but displaying a completely different chemical scaffold. Isosteres have been classified as either classical or nonclassical according to the degree of electronic and steric alikeness. Bioisosterism is used interchangeably with the term nonclassical isosteres. Figure 8.1 reports clozapine and two analogs as an example of classic isosteric replacement of NH with O and S.

An illustration of bioisosteric change is given in Figure 8.2, where the amide moiety of diazepam has been replaced by a triazole ring to give alprazolam.

Taken in its broadest meaning, bioisosterism include the replacement of the initial molecular scaffold by a different one, keeping the same biological activity. This approach, called "scaffold hopping," is well illustrated by molecules like diazepam, zolpidem, zaleplon, and zopiclone, which exert the same biological response acting as full agonists of GABA-A (γ-aminobutyric acid) receptor at the benzodiazepine site though being structurally different (see Figure 8.3).

Isosterism, bioisosterism, and scaffold hopping will be discussed in more details in the following paragraphs.

II. HISTORY: DEVELOPMENT OF THE ISOSTERISM CONCEPT

The development of the concept of isosterism takes its roots in the attempts to extend to whole molecules the knowledge acquired for elements—namely that two elements possessing an identical peripheral electronic distribution do also possess similar chemical properties.

A. The Molecular Number

In 1918 [3], Allen defined the molecular number of a compound in a way similar to the atomic number:

$$N = aN_1 + bN_2 + cN_3 + \ldots + zN_i$$

where N = molecular number N_1, N_2, N_3, . . . N_i = respective atomic numbers of each element of the molecule.

a, b, . . . z = number of atoms of each element present in the molecule.

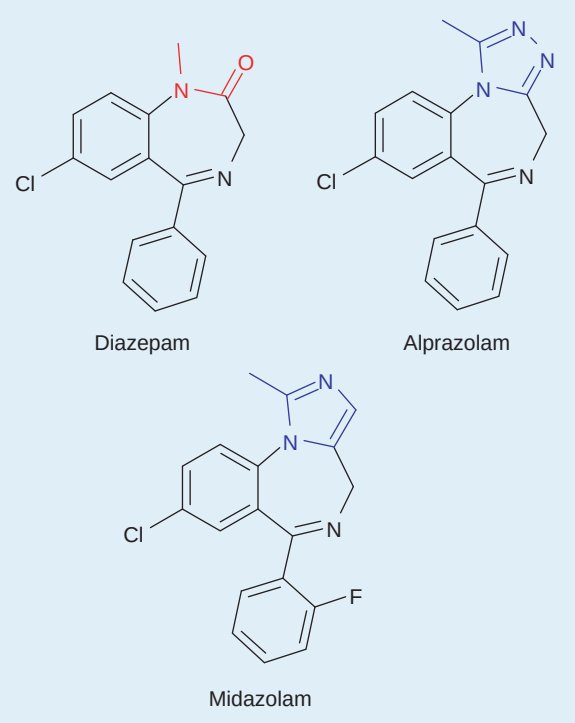


FIGURE 8.2 Example of bioisosteric replacement in the benzodiazepine series.

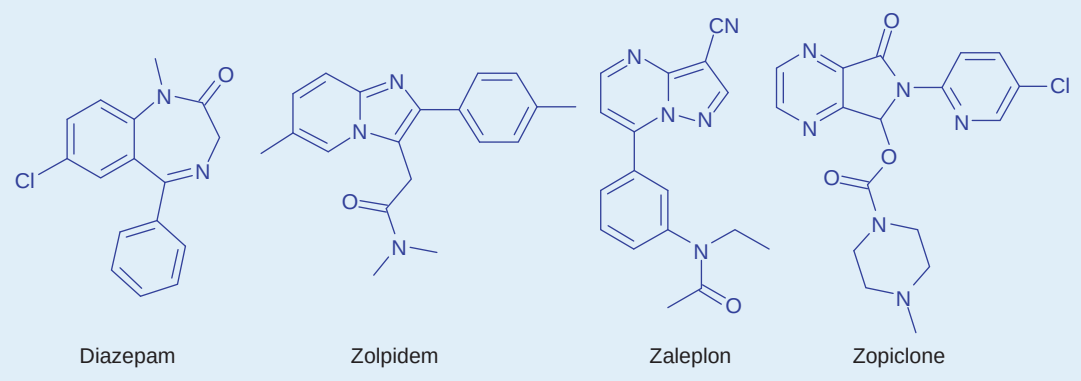


FIGURE 8.3 Scaffold hopping can be considered as the broadest example of bioisosterism: structurally different molecules elicit the same biological activity.

Example: Comparison of the ammonium and sodium cations. The atomic number of nitrogen is 7 and that of hydrogen is 1. Thus, the molecular number of the ammonium cation can be calculated and compared to that of the sodium ion:

	Atomic number		Molecular number
NH_4^+	$7 + (4 \times 1)$	=	11
Na^+	11	=	11

TABLE 8.1 Groups of Isosteres as Identified by Langmuir

Groups	Isosteres
1	H^-, He, Li^+
2	O^{2-}, F^-, Ne, Na^+, Mg^{2+}, Al^{3+}
3	S^{2-}, Cl^-, Ar, K^+, Ca^{2+}
↓	↓
8	N_2, CO, CN^-
9	CH_4, NH_4^+
10	CO_2, N_2O, N^{3+}, CNO^-
↓	↓
21	SeO_4^{2-}, AsO_4^{3-}

Possessing the same molecular number, the ammonium cation should resemble the sodium cation. This is roughly true. More generally, two compounds with identical molecular numbers present at least some similar physical properties (e.g., specific heat).

B. The Isosterism Concept

Independently of Allen, Langmuir in 1919 [4] defined the concept of isosterism:

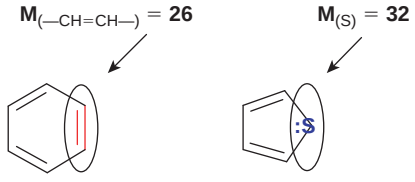
> Comolecules are thus isosteric if they contain the same number and arrangement of electrons. The comolecules of isosteres must, therefore, contain the same number of atoms. The essential differences between isosteres are confined to the charges on the nuclei of the constituent atoms.

On the basis of these similarities, Langmuir identified a list of twenty-one groups of isosteres [5]. Some of these are listed in Table 8.1. He further deduced from the octet theory that the number and arrangement of electrons in these molecules are the same. Thus, isosteres were initially defined as those compounds or groups of atoms that have the same number and arrangement of electrons. Then, he defined other relationships in a similar manner. Argon was viewed as an isostere of K^+ ion and methane as an isostere of NH_4^+ ion. He deduced that K^+ ions and NH_4^+ ions must be similar because argon and methane are very similar in physical properties. The biological similarity of molecules such as CO_2 and N_2O was later coincidentally acknowledged, as both compounds were capable of acting as reversible anesthetics to the slime mold *Physarum plycephalum* [6].

The first example clearly demonstrates that isosterism does not inevitably imply "isoelectric" structures (having the same total electric charge), but it becomes evident that isoelectronic isosteres show the closest analogies:

$$C{=}O \text{ and } N{=}N;\ CO_2 \text{ and } NO_2;\ N{=}N{=}N^- \text{ and } N{=}C{=}O^-$$

In the field of organic chemistry, Langmuir predicted the analogy between diazomethane and ketene, which was only discovered later.



C. The Notion of Pseudoatoms and Grimm's Hydride Displacement Law

Later on, in 1925, Grimm [7] formulated the "hydride displacement law," according to which the addition of hydrogen to an atom confers on an aggregate the properties of the atom of next highest atomic number. An isoelectronic relationship [8] exists among such aggregates, which were named "pseudoatoms." Thus, when a proton is "added" to the O^{2-} ion in the nuclear sense, an isotope of fluorine is obtained (Figure 8.4).

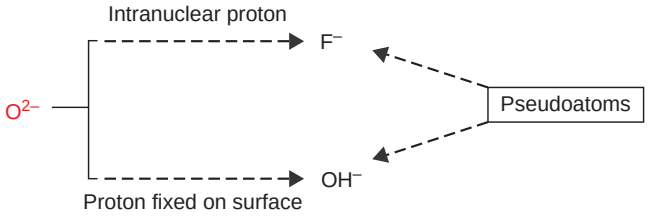


FIGURE 8.4 The notion of pseudoatoms.

TABLE 8.2 Hydride Displacement Law: In Each Vertical Column the Atom is Followed by its Pseudoatoms

Number of electrons					
6	7	8	9	10	11
C^{4-}	N^{3-}	–O–	–F	Ne	Na^+
	CH^{3-}	–NH–	–OH	FH	
		$-CH_2-$	$-NH_2$	OH_2	
			$-CH_3$	NH_3	OH_3^+
				CH_4	NH_4^+

When the same proton is introduced at the peripheral electronic level, a "pseudo-F"—in other words, an OH^-—is created. In this context, the H^+ ion having penetrated the electronic shell of the oxygen is assumed to be masked by the greater atom and to exert only negligible effect toward the outside. The fluoride anion F^- and the hydroxyl anion OH^- show therefore some analogies. The generalization of the pseudoatom concept represents the so-called "hydride displacement law" proposed in a tabular form by Grimm [8] in each vertical column, the original atom is followed by its isosteric pseudoatoms (Table 8.2).

D. Erlenmeyer's Expansion of the Isosterism Concept

Starting in 1932, Erlenmeyer [9] published a series of detailed studies about the isosterism concept, and particularly, about its first applications to biological problems. Erlen-meyer proposed his own definition of isosteres as "elements, molecules or ions in which the peripheral layers of electrons may be considered identical."

Erlenmeyer also proposed three expansions of the isosterism concept:

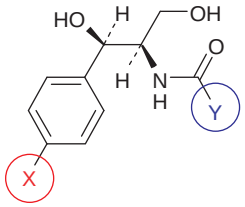
1. To the whole group of elements present in a given column of the periodic table. Thus, silicon becomes isosteric to carbon, sulfur to oxygen, etc.
2. To the pseudoatoms, with the aim of including groups which at a first glance seem totally different, but which in practice possess rather similar properties. This is the case for the pseudohalogens (e.g., $Cl \cong CN \cong SCN$, etc.)
3. To the ring equivalents: the equivalence between $-CH=CH-$ and $-S-$ explaining the well-known analogy between benzene and thiophene (Table 8.3).

E. Isoserism Criteria: Present Conceptions

The main criterion for isosterism is that two isosteric molecules must present similar, if not identical, volumes and shapes. Ideally, isosteric compounds should be isomorphic and able to co-crystallize. Among the other physical properties that isosteric compounds usually share, one can cite boiling point, density, viscosity, and thermal conductivity. However, certain properties must be different: dipolar moments, polarity, polarization, size, and shape (e.g., in comparing F^- and OH^-, the size and the shape of H cannot be totally neglected). After all, the external orbital may be hybridized differently.

In conclusion, it became evident to the physicists that the concept of isosterism, developed before quantum-mechanical theories, could not provide at the molecular level the same results as those that the periodic classification had provided for the elements, namely a correlation between electronic structure and physical and chemical properties. In the field of medicinal chemistry, the isosterism concept, taken in its broadest sense, has proven to

TABLE 8.3 The Sulphur Atom is Approximately Equivalent to an Ethylenic Group (Size, Mass, Capacity to provide an Aromatic Lone Pair)



Compound	E°C	Isostere	E°C
Benzene	80°	Thiophene	84°
Methylbenzene	110°	2-Methyl-thiophene	113°
Chlorobenzene	132°	2-Chloro-thiophene	130°
Acetylbenzene	200°	2-Acetyl-thiophene	214°

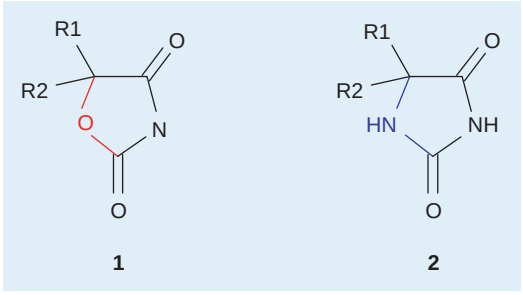


FIGURE 8.5 5,5′-Disubstituted oxazolidine-diones **1** and hydantoins **2** show similar antiepileptic profiles.

be a research tool of the utmost importance. The main reason for this is that isosteres are often much more alike in their biological than in their physical and chemical properties. An illustrative example is found in the comparison of oxazolidine-diones (**1**) and hydantoins (**2**), which possess different chemical reactivities but present a similar antiepileptic profile (Figure 8.5).

F. The Bioisosterism Concept: Friedman's and Thornber's Definitions

Recognizing the usefulness of the isosterism concept in the design of biologically active molecules, Friedman [2] proposed to call bioisosteres compounds "which fit the broadest definition of isosteres and have the same type of biological activity." This definition received rapid acceptance and is now commonly used. Moreover, Friedman considers that isosteres that exhibit opposite properties (antagonists) have also to be considered as bioisosteres, since usually they interact with the same recognition site. This is the case for para-aminobenzoic acid and para-aminobenzene-sulfonamide [10,11] and also for glutamic acid and its phosphonic analogs [12,13].

The use of the word isosterism has been largely taken beyond its original meaning when employed in medicinal chemistry, and Thornber [14] proposes a loose and flexible definition of the term bioisostere: "Bioisosteres are groups or molecules which have chemical and physical similarities producing broadly similar biological effects."

III. CURRENTLY ENCOUNTERED ISOSTERIC AND BIOISOSTERIC MODIFICATIONS

In 1970, Burger [15] classified and subdivided bioisosteres into two broad categories: classic and nonclassic. Grimm's Hydride Displacement Law and Erlenmeyer's definition of isosteres outline a series of replacement that have been termed "classical bioisosteres." Classical bioisosteres [5,16] have been traditionally divided into several

TABLE 8.4 Classic Bioisostere Atoms and Groups

Monovalent	Divalent	Trivalent	Tetravalent
–OH, $–NH_2$, $–CH_3$, –OR	$–CH_2–$	=CH–	=C=
–F, –Cl, –Br, –I, –SH, $–PH_2$	–O–	=N–	=Si=
$–Si_3$, –SR	–S–	=P–	$=N^+=$
	–Se–	=As–	$=P^+=$
	–Te–	=Sb–	$=As^+=$
			$=Sb^+=$

TABLE 8.5 Non-Classical Isosteres

–CO–	–COOH	$–SO_2NH_2$	–H	–CONH–	–COOR	$–CONH_2$
$–CO_2–$	$–SO_3H$	$–PO(OH)NH_2$	–F	–NHCO–	–ROCO–	$–CSNH_2$
$–SO_2–$	–tetrazole					
$–SO_2NR–$	$–SO_2NHR$		–OH		–catechol	
	$–SO_2NH_2$		$–CH_2OH$			
–CON–	–3-hydroxyisoxazole				–benzimidazole	
–CH(CN)–	–2-hydroxychromones		$–NHCONH_2$			C_4H_4S
R–S–R			$–NH–CS–NH_2$			$–C_5H_4N$
(R–O–R)	=N–					$–C_6H_5$
R–N(CN)–	C(CN)=R′		$–NH–C(=CHNO_2)–NH_2$			
			$–NH–C(=CHCN)–NH_2$			
–halide						$–C_4H_4NH$
	$–CF_3$					
	–CN					
	$–N(CN)_2$					
	$–C(CN)_3$					

distinct categories: (a) monovalent atoms or groups; (b) divalent atoms or groups; (c) trivalent atoms or groups; (d) tetravalent atoms, and (e) ring equivalents (Table 8.4).

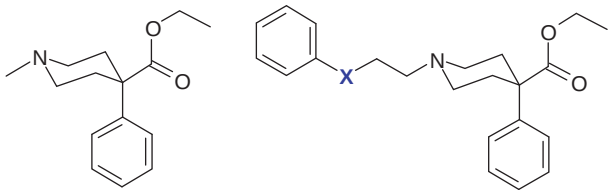
Some nonclassical isosteres are reported in Table 8.5 and from a brief glance it can be noticed that they do not obey the steric and electronic definition of classical isosteres. A second notable characteristic of nonclassical bioisosteres is that they do not have the same number of atoms as the substituent or moiety for which they are used as a replacement.

As the distinction between isosteres and bioisosteres is rather of academic interest, it is preferred in this chapter to treat both categories together. Consequently, for example, divalent series such as O=, HN=, and $H_2C=$ can be discussed together with S=. However, the correct nomenclature will be used as much as possible, keeping in mind that "isosteric replacement" embraces both true isosteres and bioisosteres.

A. Replacement of Univalent Atoms or Groups

Halogens (particularly chlorine) can be replaced by other electron-attracting functions such as trifluoromethyl or cyano groups. In the antibiotic chloramphenicol, both the chlorine atoms of the dichloroacetic moiety and of the *para*-nitro-phenyl group yielded productive isosteric replacements (Table 8.6). Many other examples of univalent atoms or groups replacements are found in the chapter dealing with substituent effects (Chapter 20) and with quantitative structure–activity relationships (Chapter 23).

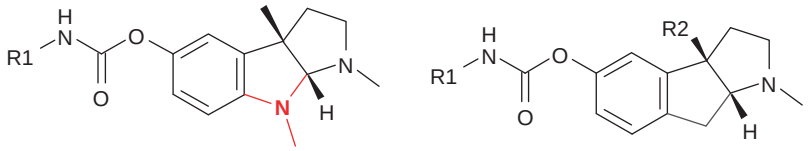
TABLE 8.6 Isosteric Replacements in the Amphenicol Family



Meperidine

Compound	X	Y
Chloramphenicol	$-NO_2$	$-CH-Cl_2$
Thiamphenicol	CH_3-SO_2-	$-CH-Cl_2$
Cetophenicol	CH_3-CO-	$-CH-Cl_2$
Azidamphenicol	$-NO_2$	$-CH_2-N_3$

TABLE 8.7 Meperidine Analogs [17]



(a) (b)

X	**Analgesic potency (Meperidine = 1)**
O	12
NH	80
CH_2	20
S	1,5

B. Interchange of Divalent Atoms and Groups

A first series of frequently interchanged divalent atoms or groups is represented by O, S, NH, and CH_2, and many interesting examples are found in the literature. In a study on meperidine analogs (Table 8.7), potent analgesic compounds were found for X = O, NH, and CH_2 [17]. Surprisingly, the sulfur analog showed only moderate activity. As an *in vivo* test was used to assess the activity, the weaker effect may be attributable to a faster metabolism (sulfoxide or sulfone formation?).

Similar changes can be applied to cyclic series, for example, to a series such as piperidine–morpholine–thiomorpholine–piperazine or in introducing oxygen or sulfur atoms into cyclic ketoprofen analogs [18].

Nice isosteric variations were observed in a series of thermolysin inhibitors [19]. For these isosteres the replacement of the phosphonamide (X = NH) function by a phosphonate (X = O) or a phosphinate ($X = CH_2$) function demonstrated clearly that the maximal activity was associated with the phosphonamide, which is able to establish a hydrogen bond with alanine 113 (Figure 8.6).

In order to have a more stable analog of the acetylcholinesterase inhibitor alkaloid physostigmine, Chen et al. [20] prepared some 8-carbaisosteres of physostigmine (Table 8.8). The authors envisioned that replacing the *N*-methyl group at N_8 of the physostigmine nucleus by a methylene group would increase its chemical and metabolic stability, thanks to the change of the less stable aminal group to a more stable amino group.

The carbaisosteres are as potent as or even more potent than the corresponding physostigmines. In addition, the (−)-enantiomers which possess the same absolute configuration at C_{3a} and C_{8a} as that of physostigmine, are generally six to twelve times more potent in inhibiting acetylcholinesterase than the corresponding (+)-enantiomers.

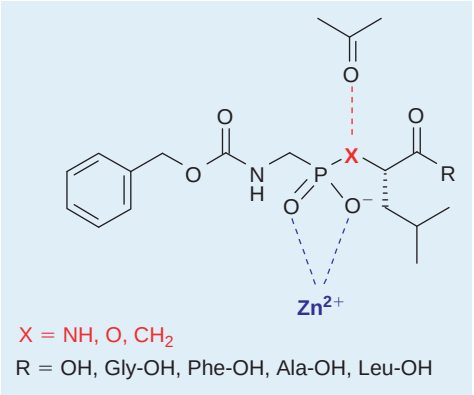
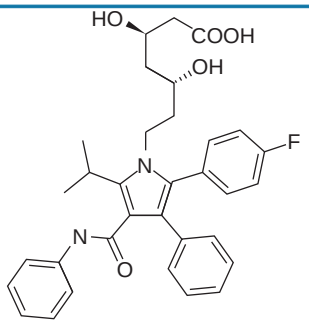


FIGURE 8.6 Isostery in thermolysin inhibitors [19].

TABLE 8.8 Physostigmines (a) and Carbaisosteres (b) [20]



Compound	R1	R2	IC_{50} (nM)	LD_{50} (mg/kg)
(−)-Physostigmine	CH_3	CH_3	128	0.88
(−)-Heptyl physostigmine	*n*-C_7H_{13}	CH_3	110	24
(±)-Carba-isostere 1	*n*-C_7H_{13}	CH_3	114	21
(−)-Carba-isostere 2a	*n*-C_7H_{13}	C_2H_5	36	6
(+)-Carba-isostere 2b	*n*-C_7H_{13}	C_2H_5	211	18

Other interesting bioisosteric replacements of oxygen atoms were devised during the search of non-hydrolyzable phosphotyrosyl (pTyr) mimetics. The phosphoryl ester oxygen of pTyr has been replaced either by a methylene (Pmp) [21] or by a difluoromethylene (F_2Pmp) [22].

C. Interchange of Trivalent Atoms and Groups

The substitution of −CH= by −N= in aromatic rings has been one of the most successful applications of classical isosterism (see Section III.D.). Interchange of trivalent atoms are found also in nonaromatic rings. For example the 4-dimethylamino-antipyrine and its carba-isostere are about as equally active as antipyretics [23] (Figure 8.7).

Similar interchanges are found in the proceeding from desipramine to nortriptyline and protriptyline (Figure 8.7), or among the antihistaminics when comparing tripelennamine with chlorpheniramine (Figure 8.7).

D. Ring Equivalents

The importance of chemical rings in drug discovery is never overstated. Looking at the best selling drugs in 2006 (reported in Table 8.9), eight out of ten molecules are small molecules and that all of them have a chemical ring in their structure. This trend is reflected also by the other marketed drugs and by the ensemble of bioactive molecules. In bioactive molecules, rings have a multiple role. Rings are responsible of the basic shape of the

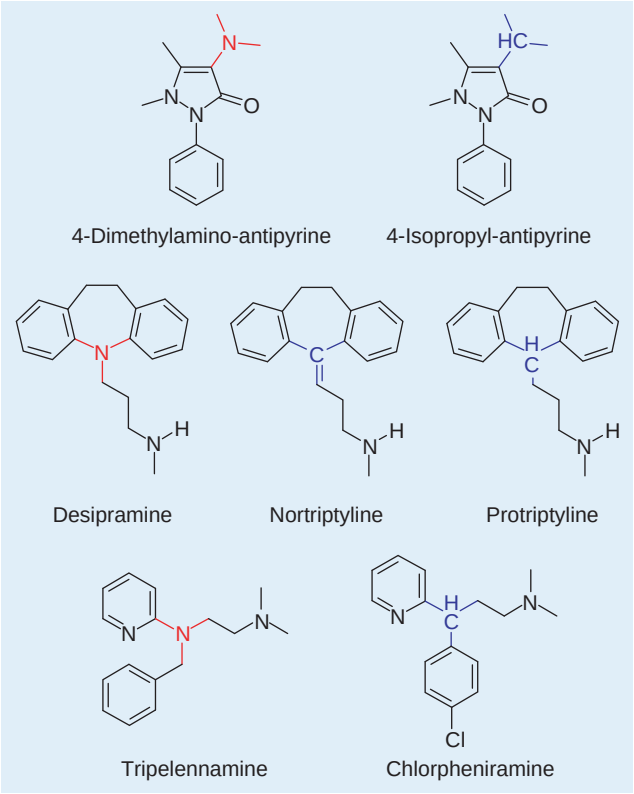


FIGURE 8.7 Interchange of trivalent atoms and groups.

molecule, rendering it rigid or flexible, and are also responsible for a suitable spatial orientation of the pharmacophoric groups. In other molecules, chemical rings are directly responsible of the biological activity, as they interact directly with the receptors either *via* heteroatom forming hydrogen bonding or *via* hydrophobic interactions.

The ADME (absorption, distribution, metabolism, excretion) profile as well as the toxicity of biologically active molecules can rely on the nature of a given ring, such as hydrophobicity, polarity, and electronic properties. Medicinal chemists use to change one ring with a bioisosteric one in order to improve the ADME-tox profile of a given molecule while keeping the desired activity and selectivity.

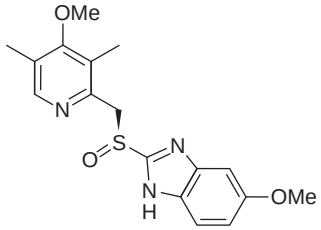
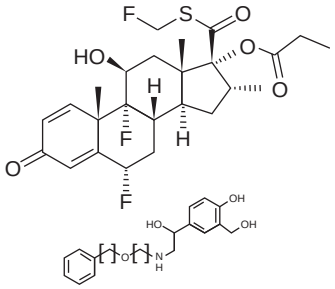
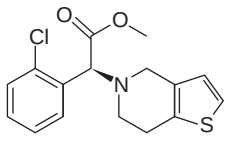
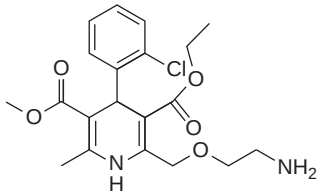
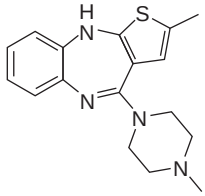
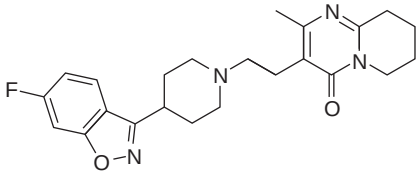
In the conception of "me too" drugs as well as in getting a stronger patent position, bioisosteric replacement of ring systems has played a master role in drug design. The substitution of –CH═ by –N═ or –CH═CH– by –S– in aromatic rings has been one of the most successful applications of classical isosterism. Early examples are found in the sulfonamide antibacterial with the development of sulphapyridine, sulphapyrimidine, sulphathiazole, etc. (Figure 8.8).

Other examples are found in the neuroleptic or antidepressant tricyclics, in the benzodiazepine tranquilizers and antiepileptics, and in the development of semi-synthetic penicillins and cephalosporins with broader spectra of activity and greater stability toward β-lactamases. In some instances. an aromatic ring can be replaced by an ethynyl group. In some instances an aromatic ring can a replaced by an ethyln group. Such a replacement was reported by Wallace et al. [24] in a series of inhibitors of Endothelin-converting Enzyme (ECE-1) deriving from the biphenyl compound CGS 26 303 and its analogs (Figure 8.8) [25].

In all these cases no essential activity difference is found between the original drug and its isostere. However, it can happen that the procedure fails. Binder et al, [26], for example, reported that thieno[2,3-*d*]isoxazole-3-methanesulfonamide, the thiophene analog of the anticonvulsant drug zonisamide (Figure 8.9) [27], was practically inactive against pentetrazole- or electroshock-induced convulsions in mice, even at high doses.

Due to the paramount role of ring systems in the drug discovery process, a lot of attention has been given to the equivalence between rings. Indeed the efforts to replace rings or other functional groups in an active

TABLE 8.9 Best Selling Drugs in 2006

Structure	Leading brands	Indication	Company	2006 sales (US$BN)
	Lipitor (atorvastatin)	High cholesterol	Pfizer	13.6
	Nexium (esomeprazole)	Gastroesophageal reflux	Astra-Zeneca	6.7
	Seretide/Advair (fluticasone + salmeterol)	Asthma	GlaxoSmithKlin	6.3
	Plavix (clopidogrel)	Coronary artery disease	Sanofi-Aventis Brustol-Myers Squibb	5.8
	Norvasc (amlodipine)	Hypertension	Pfizer	5
165-Amino-acid-glycosylated protein hormone which contains 5 *N*-linked oligosaccharide chains	Aranesp (darbepoetin alfa)	Anemia	Amgen	5
	Zyprexa (olanzapine)	Schizophrenia	Eli Lilly	4,7
	Ripserdal (risperidone)	Antipsychotic	Janssen Pharmaceutica	4.6
Recombinant human soluble tumor necrosis factor-alpha (TNF-α) receptor protein	Enbrel (etanercept)	Autoimmune diseases	Amgen-Wyeth	4.5
	Effexor (venlafaxine)	Depression	Wyeth	4

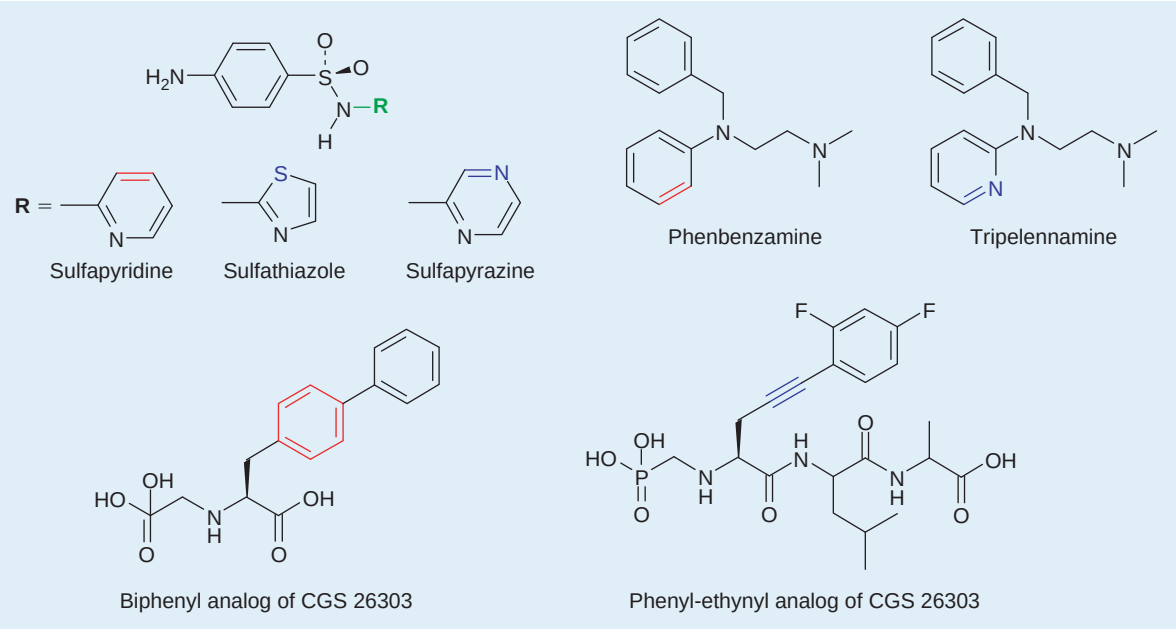


FIGURE 8.8 Classical ring equivalents.

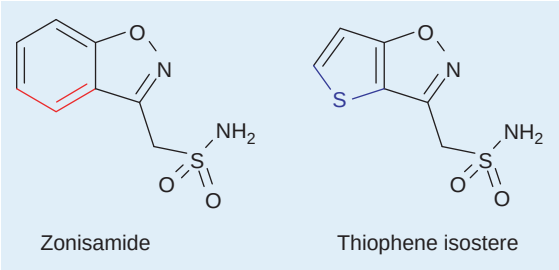


FIGURE 8.9 The thiophene isostere of zonisamide is practically inactive as an anticonvulsant.

molecule form a large part of medicinal chemistry practice devoted also to extend the space of bioactive molecules in the chemistry universe. A typical lead optimization program often involves a trial-and-error process of replacing chemical functionality, including rings, and in most cases this intellectual exercise is based on a chemist's knowledge and recollection of his or her past experience. As a source of inspiration, the following paragraphs list some examples of bioisosteres of the most encountered rings in biologically active molecules.

1. *Bioisosteres of Pyridine*

One of the most used cycles in medicinal chemistry is the pyridine ring. The bioisosteres of this heterocycle are well known as ligands for the central nicotinic cholinergic receptors. The pyridine ring of nicotine can be replaced by different other rings like methyl-isoxazole or methyl-isothiazole [28–30] (Figure 8.10a). A novel series of nicotinic agonists was described by Olesen et al. [30] In their paper, the bioisosteric replacement of the isoxazole ring in the (3-methyl-5-isoxazoly)methylene-azacyclic compound (**3**, Figure 8.10b) with pyridine, pyrazine, oxadiazole, or an acyl group resulted in ligands with moderate to high affinity for the central nicotinic cholinergic receptors ($IC_{50} = 2.0$ to $IC_{50} > 1{,}000$ nM) (Figure 8.10b).

Two other publications on the bioisosteric replacement of pyridine show similar results. The first paper reports the study of bioisosteric potential of diazines in the field of combined antithrombic thromboxane A_2 synthetase inhibitors and receptor antagonists [31,32]. On the basis of the structure–activity relationships (SAR) observed in this study, it turned out that only the 2-pyrazinyl, 4-pyridazinyl, and 5-pyrimidinyl systems are appropriate bioisosteric moieties for the 3-pyridyl system in the dual active platelet antiaggregatory compound Ridogrel

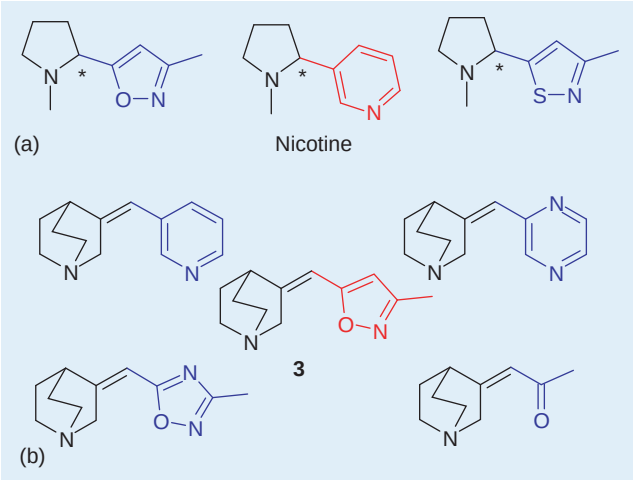


FIGURE 8.10 Ligands for central cholinergic receptors with different non-classical bioisosteres of the pyridine ring.

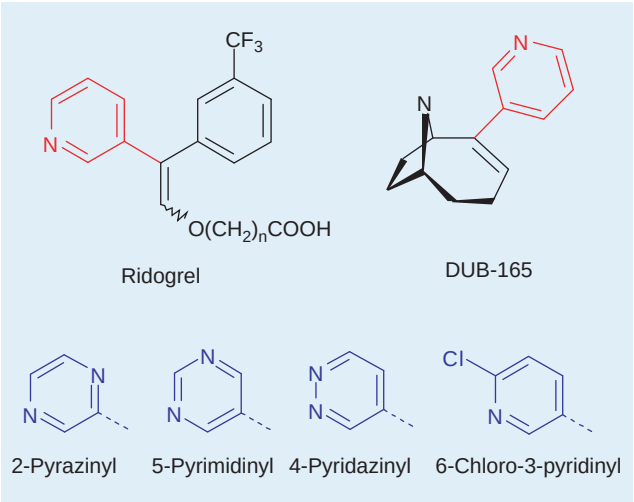


FIGURE 8.11 Bioisosteric moieties for the 3-pyridyl ring.

(Figure 8.11). Gohlke et al. [31] also observed the bioisosteric potential of diazines in the SAR of the compound DUB-165. The replacement of the 3-pyridyl group of DUB-165 by a 4-pyridazinyl, 5-pyrimidinyl, or 2-pyrazinyl moiety resulted in ligands retaining affinity for nAChRs subtypes, thus demonstrating that the three isomeric diazines are appropriate bioisosteres of the 3-pyridyl moiety (Figure 8.11).

N-(6-Chloronaphthalen-2-)sulfonylpiperazine derivatives **4** and **5** (Figure 8.12) are potent factor Xa inhibitors. Haginoya et al. [33] proposed replacing the pyridine-phenyl or the pyridine-piperidine residue by a fused-bicyclic ring that contains an aliphatic amine and a pyridine to yield the compound **6**, which has an interesting factor Xa inhibitor activity. The bioisosteric replacement of the pyridine moiety of the 6-methyl-5,6,7,8-tetrahydro-[1,6]naphthyridine by phenyl, thiophene, or thiazole analogs yielded analogs with similar or better antifactor Xa activity, but also to conserve a moderate bioavailability.

A nice example of pharmacophore equivalent of the 2,3-diaminopyridine with cyclopropylamino acid amide has been reported recently [34]. A series of 2,3-diaminopyridines represented by **7** acting as bradykinin (BK) B_1 antagonists was thoroughly investigated with the aim to replace the metabolically labile diaminopyridine ring with a more bioavailable moiety. An accurate design strategy led researchers at Merck to replace the 2,3-diamino pyridine with a series of alkylamino acid amides. Among this, series compound **8**, though less active than **7** as BK B1 antagonist, showed a better pharmacokinetic profile with an improved bioavailability, an increased half-life, and a decreased clearance compared to **7** (Figure 8.13).

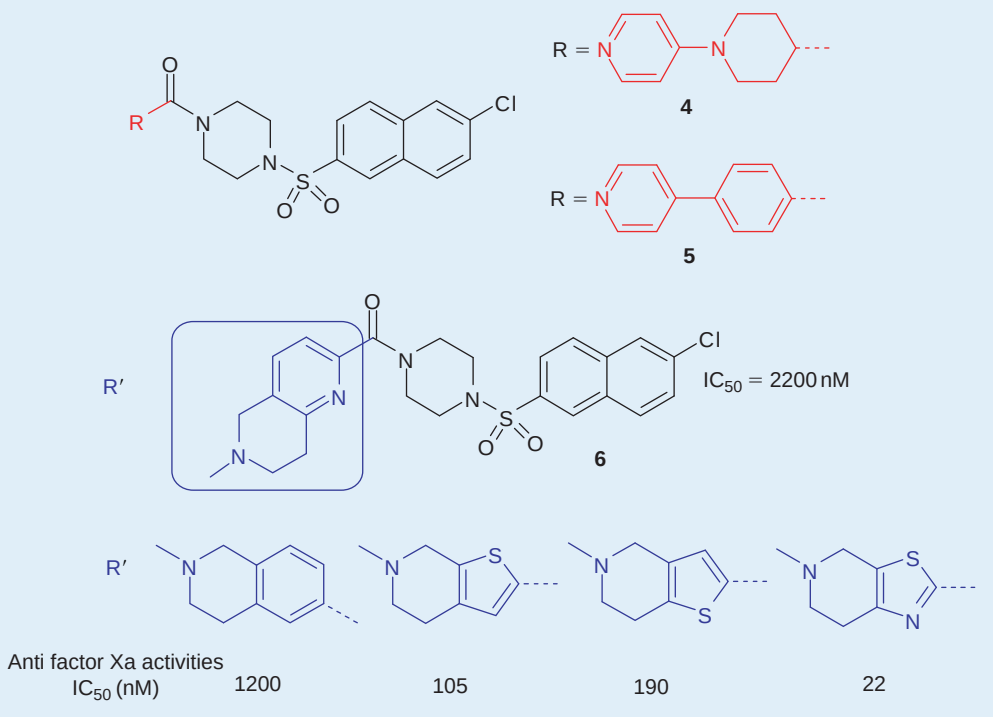


FIGURE 8.12 Bioisosteric replacements of the pyridine ring in a series of factor Xa inhibitors.

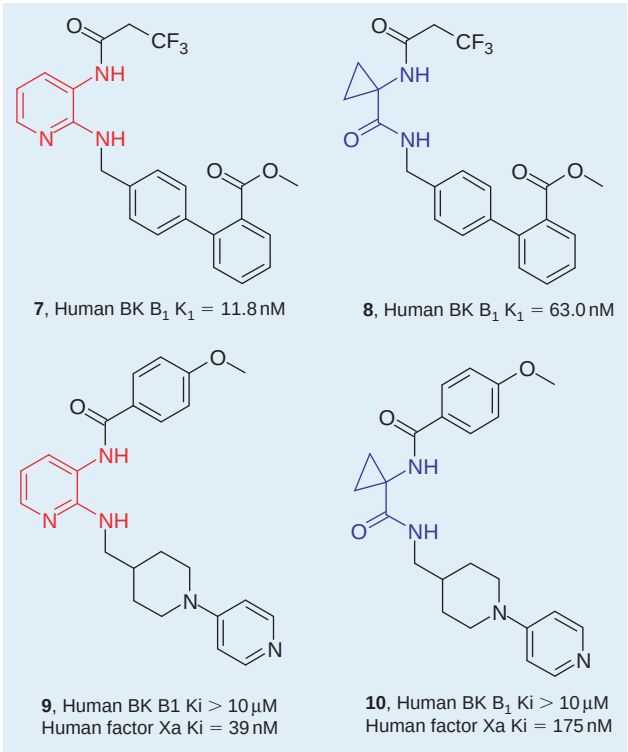


FIGURE 8.13 Cyclopropylamino acid amide as pharmacophoric replacement for 2,3-diaminopyridine. Application to the design of novel bradykinin B_1 receptor antagonist and factor Xa inhibitor.

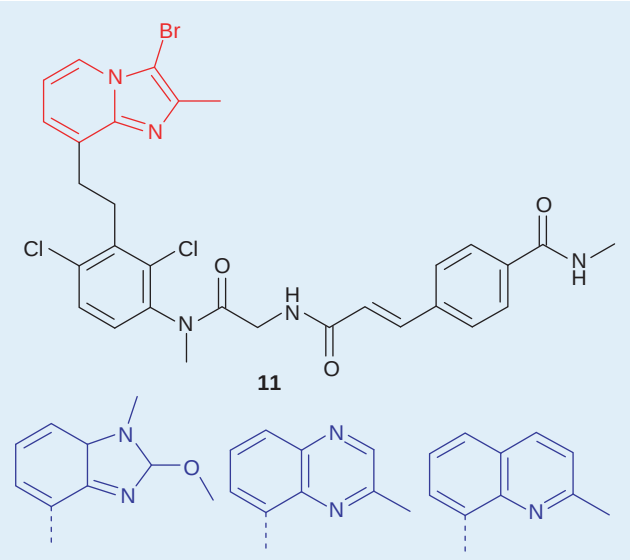


FIGURE 8.14 Novel class of orally active non-peptide bradykinin B_2 receptor antagonists showing valuable examples of imidazo[1,2-a]pyridine analogs.

To demonstrate the generality of this bioisosteric equivalence, the same group applied the same substitution to **9**, a factor Xa inhibitor that yielded **10**. Compound **10** reflects the same loss in potency that was observed for **8** towards **7**, but reoptimization of the *N*-acyl group afforded quickly a significant gain of potency.

2. *Imidazo[1,2-a]pyridine Bioisosteres*

Abe et al. [35] reported that the imidazol [1,2-a]pyridine moiety of the basic framework of a class of the non-peptide bradykinin B_2 receptor antagonists (**11**; Figure 8.14) could be successfully replaced by several heterocyclic bioisosteres. Among those, the 1-methyl-2-methoxy-1*H*-benzimidazole, 2-methylquinoxaline and 2-methylquinoline derivatives showed potent B2 binding affinities against both human and guinea pig B2 receptors (Figure 8.14).

3. *Pyridazine Bioisosteres*

Since the antidepressant minaprine has been launched, several pyridazine analogs were proposed and synthesized. Most of them showed different activities on the central nervous system. A series of these compounds (Figure 8.15) are acetylcholinesterase inhibitors with variable bioisosteres of the central pyridazine [36,37]. The replacement by pyridine, 1,2,4-thiadiazole, and triazines yields compounds with weaker but still acceptable activity.

Examples from literature show that there are several nonclassical bioisosteres for the pyridine system or the pyridazine system. What is the basis for determining the most appropriate analog with which to start? An indication could be given by the comparison of the boiling points of these heterocycles, assuming that the more similar the boiling point is the most appropriate the bioisostere is (Figure 8.16) [38]. For example, while searching for a bioisostere of the pyridine ring, judging by the boiling points it appears that the best candidates are the pyrimidine, the pyrazine, and the 1,2,4-thiadiazole systems (Figure 8.16). The same could be extended to the pyridazine ring where this comparison permits the selection of the 1,2,4-triazine or the 1,3,4-thiadiazole rings (Figure 8.16). In fact, those findings confirm the examples found in the literature and reported in the previous part of this chapter about isosteres or bioisosteres of the pyridine and pyridazine rings.

A possible interpretation of these results can be that in the heterocyclic series, the boiling point is correlated to the dipolar moment of the molecule and that for two heterocyclic rings having the same aromatic geometry, the similarity of the dipolar moments may represent the dominant feature.

4. *Bioisosteres of Other Heterocycles*

Chemical rings are fundamental for drug molecules as a typical drug molecule consists of a combination of chemical rings, chains, and functional groups. The high proportion of chemical rings in drug molecules speaks in

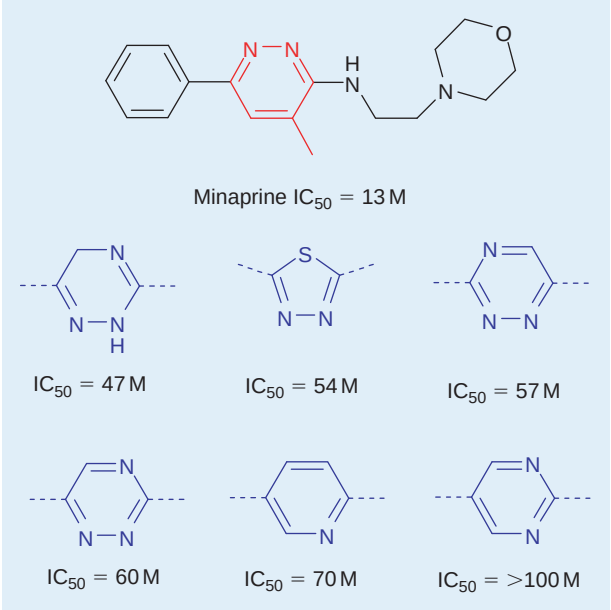


FIGURE 8.15 *In vitro* inhibition of acetylcholinesterase in rat striatum homogenates: pyridine, thiadiazole, and triazine replacement of the pyridazine ring.

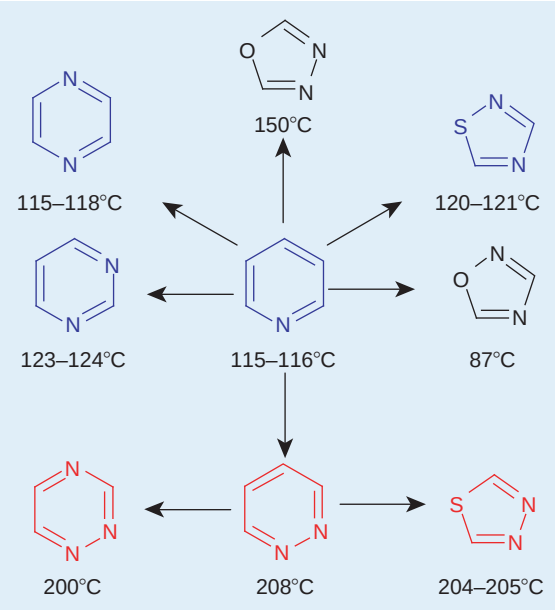


FIGURE 8.16 Structures and boiling points of pyridine and pyridazine isosteres.

favor of the large effort to synthesize and find new chemical ring equivalents, whether it is to design a new platform for a better patent position, replace a metabolically unstable moiety, look for a more favorable receptor interaction, or have a better pharmacodynamic profile.

There are many examples of ring equivalents, and so far it seems impossible to rationalize a method that could give a reliable indication of which heterocycle could be the best substitute of another one. The best approach is to look at the examples reported in the literature and learn from those.

Selective cyclooxygenase-2 inhibitors (COX-2 inhibitors) give a nice example of bioisosteres of heterocycles. The comparison of the most potent selective COX-2 inhibitors (Figure 8.17) suggests that isoxazoles, pyridines, and pyrazoles are good bioisosteres of each other as well as nitrophenol and indanones [39].

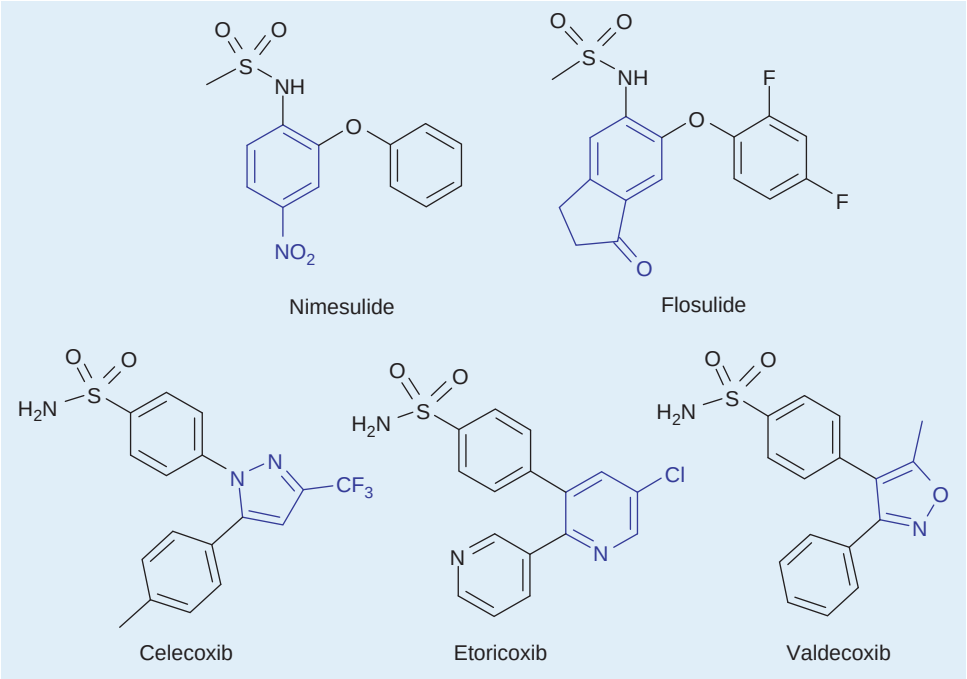


FIGURE 8.17 Examples of selective cyclooxygenase-2 inhibitors showing different ring equivalents.

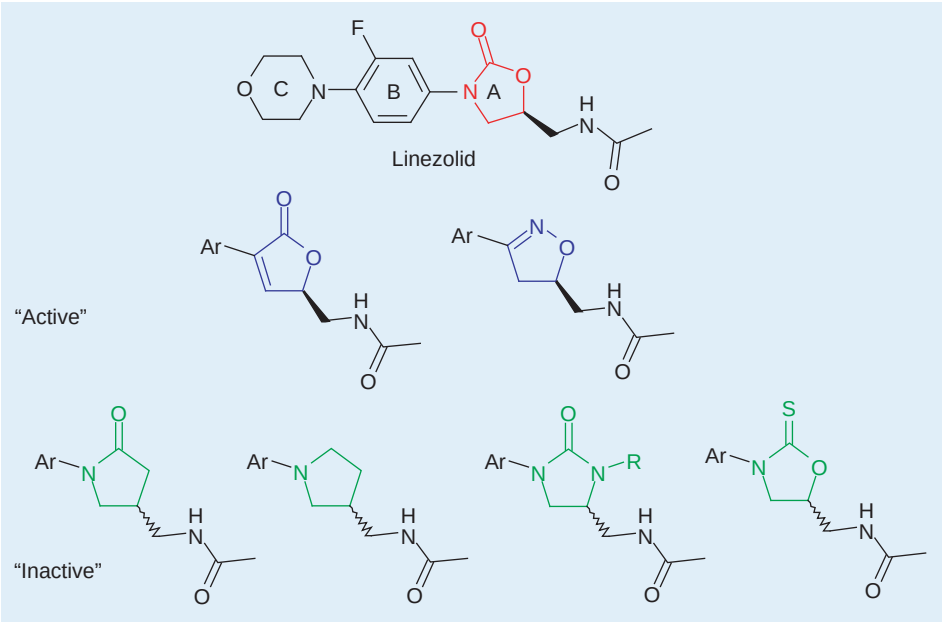


FIGURE 8.18 Oxazolidinone bioisosteres synthesized as antibacterial agents.

Another example of heterocycle bioisosteres is given by the oxazolidinone antibacterials. Since the discovery of linezolid as potent antibacterial against Gram positive organisms [40], many researchers have taken inspiration from the oxazolidinone scaffold in order to create new and potentially improved antibacterial agents. The structure of Lienzolid (Figure 8.18) is divided in three parts: the A ring corresponding to the oxazolidinone moiety, the B ring coresponding to the Phenyl ring, and the C ring corresponding to the morpholine. Some research groups have worked around the modification of the B and C rings [41,42].

The work of Snyder et al, together with other examples found in the literature, shows that several levels of activity can be identified among the different isosteres. It has to be kept in mind that the replacement of a ring

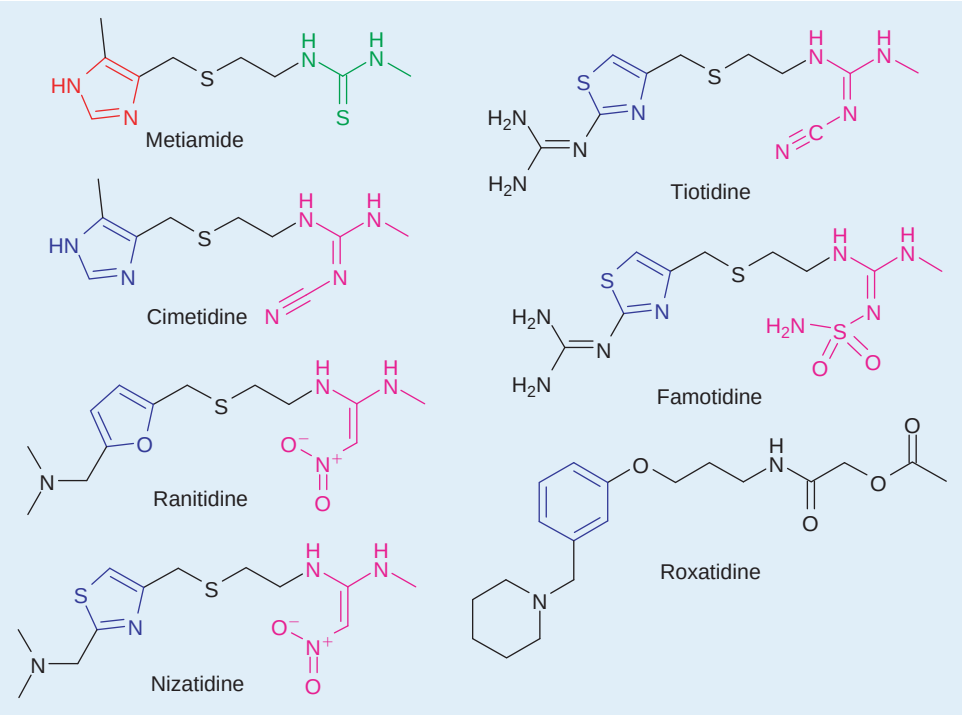


FIGURE 8.19 Antiulcer H_2-receptor histamine antagonists: evolution of structures in the course of the time. Note the progressive use of a furan, a thiazole, and finally a phenyl ring in place of the original imidazole ring. On the same series it is noteworthy the bioisosterism of the urea moiety with different surrogates: thiourea, *N*-cyanoguanidine, *N*-nitro-ethene diamine, and *N*-aminosulfonil guanidine.

by its isosteres does not always lead to an iso-active or more active compound (Figure 8.18), but in general it is possible to obtain at least one or more equivalent systems [43].

Another particularly interesting example of ring bioisosterism is found in the development of the antiulcer H2-receptor histamine antagonists in which the initial imidazole ring was changed to various other "equivalents" such as a furan, a thiazole, and finally a phenyl ring (Figure 8.19). A detailed and interesting account of the discovery and the development of these compounds is found in Ganellin and Roberts' book [44].

Better bioisosteric design possibilities are provided by quantum-chemical calculations. Mallamo et al. [45] made use of electrostatic potential surface maps complementarity in defining sulfonyl heterocycles bioisosteric to the steroidal antiandrogenic drug zanoterone (Figure 8.20). Striking differences in the electrostatic potential surfaces accounted for the observed variability in the furan (active) and the thiophene (inactive) analogs of zanoterone. Good androgen receptor affinity was then anticipated and effectively found for the oxazole and the thiazole analogs of zanoterone.

The apparent failure of the isosterism concept for the inactive thiophene, inversed furan, and pyrimidine is thus interpretable on a rational basis. Table 8.10 lists some more "exotic" examples of bioisosteric replacements of cyclic systems.

E. Groups with Similar Polar Effects: Functional Equivalents

1. *Carboxylic Acid Bioisosteres*

Medicinal chemists have frequently faced the problem of developing surrogates for carboxylic acid groups. The availability of bioisosteric replacement for the carboxylic acid group has been critical to the development of novel medicinal agents, especially in the area of neurochemistry. Less polar and more hydrophobic groups were needed to replace the carboxylic group of GABA in order to cross the blood–brain barrier while keeping some of the physiochemical properties of the parent functional group like pK_a and partition coefficient.

Thus, the two GABAergic agonists, isoguvacine and THIP (Figure 8.21), have similar pharmacological properties to GABA itself. The key parameters in these compounds are the acidic (pK_a <4) and the basic (protonated nitrogen) functions with a <5.1 Å intercharge distance.

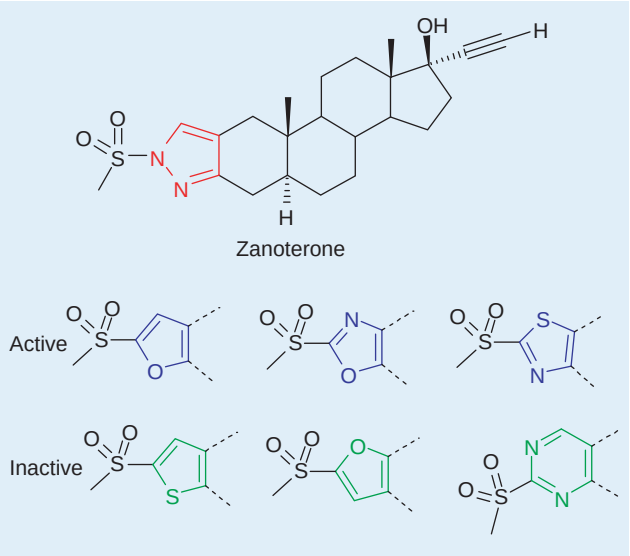
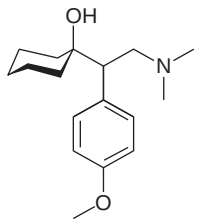
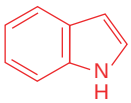
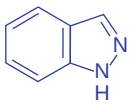
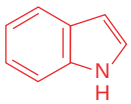
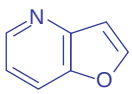
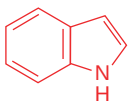
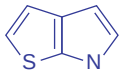
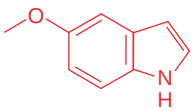
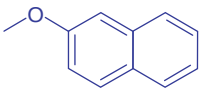
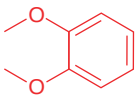


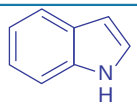
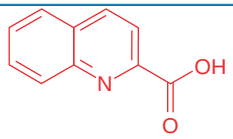
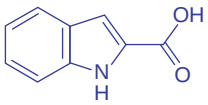
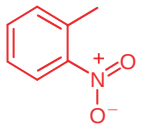
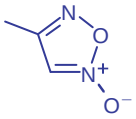
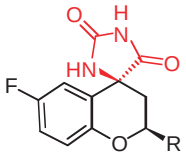
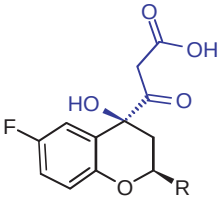
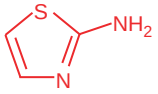
FIGURE 8.20 Zanoterone isosteres [45].

TABLE 8.10 Other Ring Bioisosteres

Original ring	Bioisostere	Activity	Reference
	 Indole	5-HT_3 antagonists	Fludzinski et al. [46]
 Indazole	 Indole	5-HT_{1F} agonist	Mathes et al. [47]
 Furo[3,2-b] pyridine	 Indole	Hallucinogen serotonin agonist	Blair et al. [48]
 6H-thieno[2,3-b] pyrrole	 Indole	Agonist at melatonin receptors and antagonist at $5HT_{2c}$	Yous et al. [49]; Depreux et al. [50]
 Napthalene	 3,4-Dialkoxy-phenyl	Phospho diesterase inhibitors GABA uptake inhibitors	Blaskó et al. [51]; Kardos et al. [52]

(*Continued*)

TABLE 8.10 (Continued)

Original ring	Bioisostere	Activity	Reference
 Indole	 Quinoline-2-carboxylate	Glycine antagonists	Salituro et al. [53]
 Indole2-carboxylate	 *o*-Nitro-phenyl	Calcium antagonist	Calvino et al. [54]
 Furoxane	 Spiro-hydantoin	Aldose reductase inhibitor	Lipinski et al. [55]
 Spiro hydroxyacetic acid unit	 Thiazol-2-ylamine	Dopamine D_3 receptor agonist	Löber et al. [56]

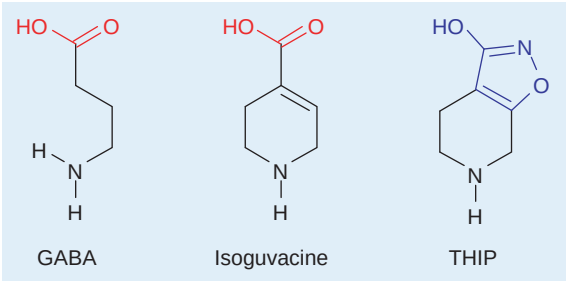


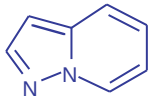
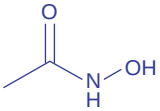
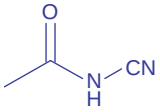
FIGURE 8.21 An example of bioisosterism, or non-classical isosterism: GABA, isoguvacine, and THIP are all agonists for the GABA-A receptor. The 3-hydroxy-isoxazole ring has a comparable acidity to that of a carboxylic acid function [57].

For the carboxylic function of active compounds, three different classes of bioisosteres can be distinguished: (a) the direct derivatives (Table 8.11); (b) the planar acidic heterocycles (Table 8.12); and (c) the nonplanar sulfur- or phosphorus-derived acidic functions (Table 8.5).

A. DIRECT DERIVATIVES

Direct derivatives of carboxylic acids are different functional groups that maintain an acidic proton and a hydrogen bond acceptor (the carbonyl) in order to have similar specific interactions with the receptor. Among the most common direct derivatives can be listed the hydroxamic acids (R—CO—NH—OH), the acylcyanamides (R—CO—NH—CN), and the acylsulfonamides (R—CO—NH—SO_2—R′) (Table 8.11).

TABLE 8.11 Carboxylic Acid Isosteres: Direct Derivatives

 Pyrazolo[1,5-a] pyridine	Hydroxamic acids	High chelating power	Almquist et al. [58]
		Histone deacetylase inhibitors	Massa et al. [59]; Lu et al. [60]
		Matrix metalloproteinases inhibitors	Remiszewski et al. [61]; Plumb et al. [62]; Kelly et al. [63]; Buggy et al [64]
		Tumor necrosis factor α converting enzyme	Hanessian et al. [65]; Aranapakam et al. [66,67]; Noe et al. [68]; Duan et al. [69]
	Acyl-cyanamides	Mainly academic interest	von Kohler et al. [70]; Kwon et al. [71]
	Acyl-sulfonamides	Glycine, GABA, and β-alanine analogs	Drummond and Johnson [72]
		Antiatherosclerotics pK_a # 4,5	Albright et al [73]
		β_3 Adrenergic receptor agonist hepatitis C virus	Uehling et al. [74]; Johansson et al [75]

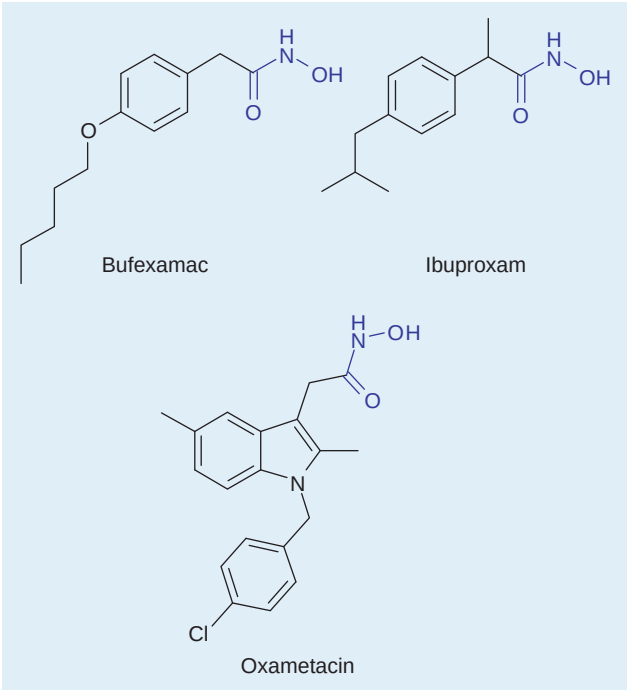
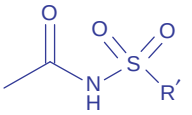
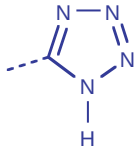
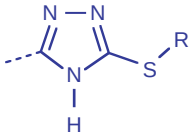
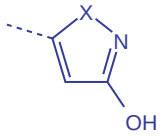
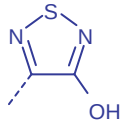
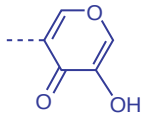
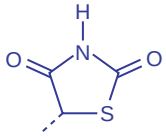
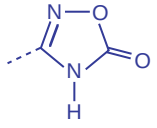
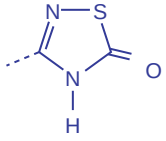


FIGURE 8.22 Hydroxamate isosteres of antiinflammatory drugs.

Hydroxamic acids: The hydroxamic acids have a high chelating power, and they have found a wide application as inhibitors of enzymes having a metal ion like Zn^{2+} in the active site. Noteworthy examples are found in the inhibitors of matrix metalloproteinases (MMP) [65–67,86], tumor necrosis factor α converting enzyme (TNFα) [69], and histone deacetylase (HDAC) [59,60] where the hydroxamic acid derivatives have given potent and bioavailable compounds, despite the fact that this group is particularly prone to hydrolysis, reduction, and glucoronidation [68]. Another interesting example of exploitation of the hydroxamic acid is given by the anti-inflammatory hydroxamates bufexamac [87], ibuproxam [88], and oxametacin [89] (Figure 8.22), where the hydroxamates can act either as a prodrug-like ibuproxam being metabolized to ibuprofen (CONHOH → COOH) in humans [90] or as a true bioisostere like oxametacin being metabolically stable in humans [91,92].

TABLE 8.12 Carboxylic Acid Isosteres: Planar Acidic Heterocycles and Aryl Derivatives

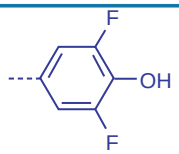
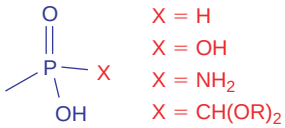
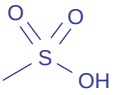
	Tetrazoles	Very popular, great number of publications. $pK_a = 6.6$ to 7.2	Bovy et al. [76]; Marshall et al [77]
	Mercaptoazoles + Sulfinylazoles + Sulfonylazoles	Phosphonate isosteres pK_a mercapto: 8.2–11.5 pK_a sulfinyl: 5.2–9.8 pK_a sulfonyl: 4.8–8.7	Chen et al. [20]
	Isoxazoles Isothiazoles	GABA and glutamic acid analogs	Krogsgaard-Larsen et al. [57] Krogsgaard-Larsen [78]
 X = O or S	Hydroxy-thiadiazole	Isoxazole isostere pK_a # 5	Lunn et al. [79]
	Hydroxy-chromones	Kojic acid derivatives (as GABA agonists)	Atkinson et al. [80]
	Thiazolidinediones	Dual PPAR α/γ agonists	Hulin et al. [81] Henke [82]
	1,2,4-Oxadiazole-5(4*H*)-ones	Antimycobacterial	Gezginci et al. [83]
	1,2,4-Thiadiazole-5(4*H*)-ones	Antimycobacterial	Gezginci et al. [83]
	3,5-Difluoro-4-hydroxyphenyl	Aldose reductase inhibitor GABA analog	Nicolaou et al. [84] Qiu et al. [85]

Acylsulfonamides: Acylsulfonamides are also used in drug discovery and have been employed in several therapeutic areas such as β_3 adrenergic receptor agonist [74], chemokine receptor 1 (CXCR1) inhibitors [93], hepatitis C virus (HCV) inhibitors [75], angiotensin receptor agonists [94], antitumoral compounds with antiproliferatives properties [95,96], and prostaglandin EP4 receptor antagonists [97]. Acylcyanamides are mainly of academic interest.

B. PLANAR ACIDIC HETEROCYCLES AND ARYL DERIVATIVES

Table 8.12 reports the planar acidic heterocycles and aryl derivatives that are commonly used as carboxylic acid bioisosteres. The most used is the tetrazole ring, which has found wide application in different therapeutic

TABLE 8.13 Carboxylic Acid Isosteres: Non Planar Sulfur- or Phosporus-Derived Acidic Functions

	Phosphinates Phosphonates Phosphonamides	Many examples in glutamate and in $GABA_B$ antagonist series	Froestl et al. [112]
	Sulfonates	Sulphonic analogs of GABA and glutamic acid	Rosowski et al. [113]
	Sulfonamides	Weak acids, used rather as equivalents of phenolic hydroxyls in the design of catecholamine analogs	von Kohler et al. [70]

fields spanning type 2 diabetes [98,99], hepatitis C virus (HCV) [75], malaria [100], Alzheimer's disease (AD) [101], anxiety treatment [102], and pain management [103]. The tetrazole has been employed to fix different issues, for example, to improve bioavailability [100], to enhance the blood–brain barrier penetration [102,103], to increase potency [99,104], to get a better chemical stability [101], to bring some selectivity (the GABA tetrazole analog inhibits GABA-transaminase, but not succinic semialdehyde dehydrogenase) [105], or to be used as a pro-drug [106]. However, in some instances, tetrazole analogs are poorly active [107].

Hydroxy-isoxazoles and other cognate heterocyclic phenols encompassing an acidity range from 3.0 to 7.1 were incorporated in GABA agonists, antagonists, and uptake inhibitors [108,109]. The experience gained with 3-hydroxy-isoxazoles in the GABA field was also transferable to glutamate receptor ligands and led to selective antagonists for glutamic acid receptor subtypes [108].

Thiazolidinediones are another class of heterocyclic carboxylic acid surrogates commonly used for peroxisome proliferator-activated receptors (PPAR) agonists [81,82], as potent antihyperglycemic and lipid activity modulators. Other interesting but less studied heterocyclic surrogates are: 3,5-dioxo-1,2,4-oxadiazolidine [110], 3-hydroxy-1,2,5-thiadiazoles [79], 1,2,4-oxadiazole-5(4*H*)-ones [83], 1,2,4-thiadiazole-5(4*H*)-ones [83], 3,5-difluoro-4-hydroxyphenyl [84,85], and 3-hydroxy-γ-pyrones [80,111].

C. NONPLANAR SULFUR- OR PHOSPHOROUS-DERIVED ACIDIC FUNCTIONS

The most extensive use of phosphonates (Table 8.13) occurred in the design of amino acid neurotransmitter antagonists such as glutamate [114] and GABA-B antagonists [112].

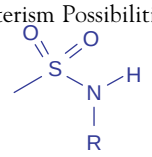
Among a set of cholecystokinin (CCK) antagonists derived from the nonpeptide CCK-B selective antagonist CI-988, Drysdale et al. [115] prepared a series of carboxylate surrogates spanning a pK_a range from <1 (sulfonic acid) to >9.5 (thio-1,2,4-triazole). The affinity and the selectivity of the compounds were rationalized by considering pK_a values, charge distribution, and geometry of the respective acid mimics (Table 8.14).

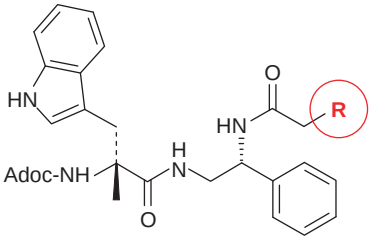
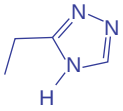
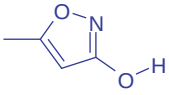
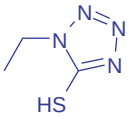
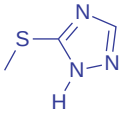
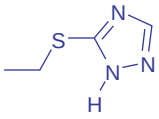
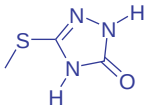
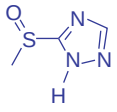
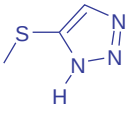
In order to choose the best carboxylic acid bioisoster, pK_a and log P are important parameters to examine before proceeding with one surrogate or another. An interesting study comparing pK_a and log P values of some aryl phosphonic acids, aryl tetrazoles, and aryl sulfonamides has been published by Franz [116]. The values of pK_a and log P are very important parameters in drug design. Several published examples show that the interchange of carboxylic acid for tetrazole and sulfonamide often results in useful drugs. This study indicates that the log P will be lowered by about 1 log unit by substituting a phosphonic acid by a carboxylic acid group. There is an increase in acidity that may be a limiting factor in the absence of active transport. Replacing a carboxylic acid group with a phosphonic acid group gives a much more acidic compound; replacing it with a sulfonamide group results in a much less acidic compound; and with a tetrazole replacement, acidity is essentially unchanged (Figure 8.23). If one wants to lower the log P, and if increased acidity of the compound is not a limiting factor, substitution of a phosphonic acid group for a carboxylic acid would be a viable approach (Figure 8.24).

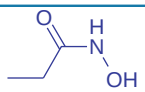
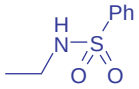
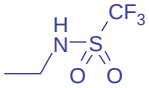
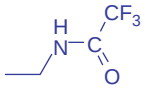
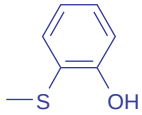
Diamino-cyclobutene-diones as α-amino carboxylic acid surrogates: Diamino-cyclobutene-dione was proposed by Kinney et al. [117] as an original surrogate of the α-amino carboxylic acid function (Figure 8.25).

Carboxylic functions as phosphonates surrogates: pTyr mimetics serve as important components of many competitive protein-tyrosine kinases inhibitors. To date, the most potent of these inhibitors have relied on phosphonate-based structures to replace the 4-phosphoryl group of the parent pTyr residue (Figure 8.26). Interestingly, it was

TABLE 8.14 Exploration of the Carboxyl Isosterism Possibilities in a Series of CCK Antagonists



R	IC_{50} (nM) CCK-B	IC_{50} (nM) CCK-A	A/B ratio	pK_a
R–CH_2–COOH	1.7	4500	2500	5.6
CHARGE DISTRIBUTED MONOANIONIC ACID MIMICS				
	6.0	970	160	5.4
	2.6	1700	650	6.5
	2.4	620	260	4.3
	2.5	680	270	>9.5
	16	850	53	>9.5
	4.3	660	150	7.7
	1.7	940	550	7.0
	6.3	1300	200	5.2
	18	600	33	>8.2
	14	1300	93	>9.5

R	IC_{50} (nM) CCK-B	IC_{50} (nM) CCK-A	A/B ratio	pK_a
POINT CHARGE MONOANIONIC ACID MIMICS				
	70	300	4.3	>9.5
	77	680	9	7.9
	110	790	7	>9.5
	80	510	6.4	>9.5
	21	1500	71	>9.5
TETRAHEDRAL ACID MIMICS				
$P(O)(OH)_2$	27	5200	190	3.4; 7.7
CH_2—$P(O)(OH)_2$	23	2700	120	3.4; 7.8
P(O)(OH)(OEt)	12	480	40	6.5
P(O)(OH)Me	12	1700	140	3.8
CH_2—P(O)(OH)Me	23	4400	190	3.7
CH_2—SO_3Na	1.3	1010	780	–

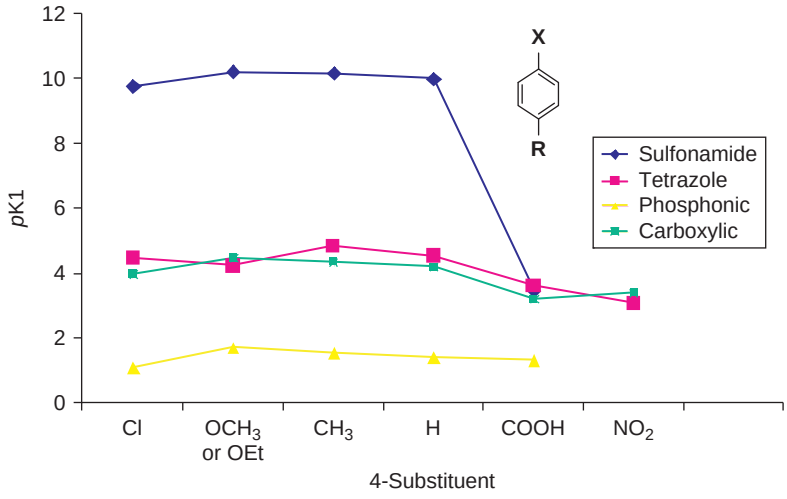


FIGURE 8.23 pK_1 values for aromatic compounds. *Redrawn after Franz [116].*

found that carboxy-based pTyr analogs can be utilized to introduce the anionic oxygen functionality of the parent phosphate. Particularly, when *p*-(2-malonyl)phenylalanine (Pmf) was incorporated as pTyr replacement in the high-affinity Grb2 SH2 domain binding sequence, potencies approaching one of phosphonate mimetics were obtained [118].

The above example is an elegant illustration of the possibility to mimic the pyramidal structure of phosphates or phosphonates by means of two planar carboxylic groups, the three-dimensionality originating from the malonic methylenic carbon atom.

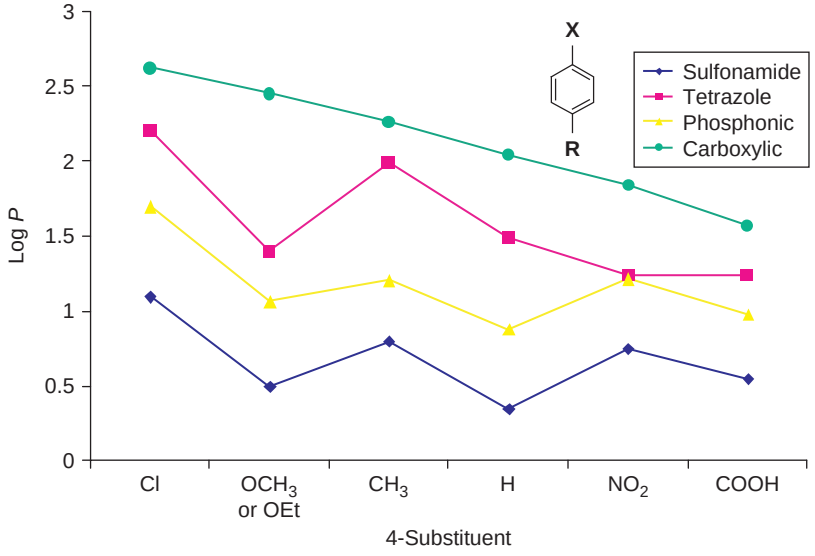


FIGURE 8.24 Log *P* values for aromatic compounds. *Redrawn after Franz [116].*

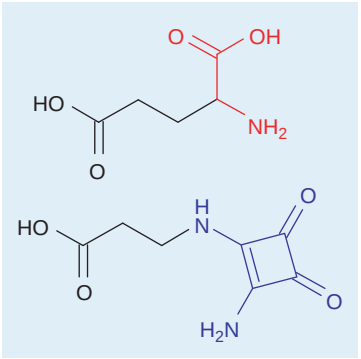


FIGURE 8.25 3,4-Diamino-3-cyclobutene-1,2-dione as surrogate of the α-amino carboxylic acid function.

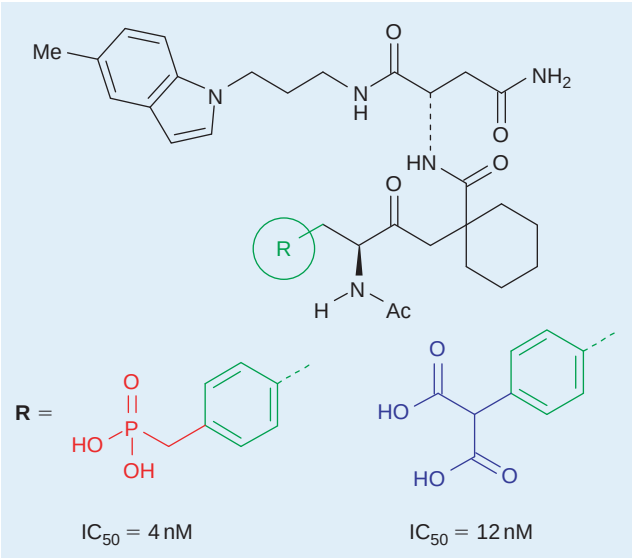


FIGURE 8.26 Malonates as surrogates of phosphonates.

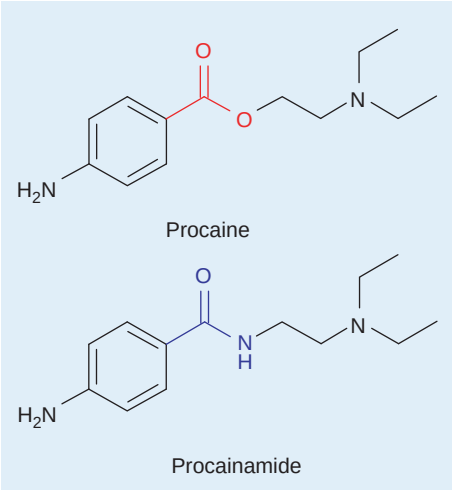


FIGURE 8.27 Amide as an ester surrogate of the local anesthetic procaine.

2. *Carboxylic Esters Bioisosteres*

Carboxylic esters are nice tools for SAR studies investigation but they are not a valuable group to be left in drug molecules because they lack the proper of *in vivo* stability. As a matter of fact, they are hydrolized quite quickly by the esterases that are quite ubiquitous and are present in the blood, the liver, the kidneys, and other organs. To overcome this problem, the carboxylic ester in the most promising compounds has been replaced by bioisosteric groups leading to molecules with improved pharmacokinetic profile.

One of the first examples of ester bioisoster is the amide bond as it is illustrated by the local anesthetic procaine, where the readily hydrolyzable ester group was replaced with an amide to give procainamide as a more stable analog (Figure 8.27).

Similarly, the lactone ring of the muscarinic agonist pilocarpine was changed into various still active isosteres, such as the corresponding thiolactone, lactam, lactol, and thiolactol [119]. A series of aspirin isosteres has been prepared by replacing the carboxylic ether oxygen successively by a nitrogen, sulfur, or carbon isosteric equivalent [120]. None of the isosteric compounds showed any activity. This result is readily understood since the particular role of aspirin as an acylating agent of the enzyme cyclooxygenase has been demonstrated [121].

Pilocarpine is widely used as a topical miotic for controlling the elevated intraocular pressure associated with glaucoma. Besides its low lipophilicity, which stimulated the search for prodrugs [122], pilocarpine has a short duration of action, its lactonic ring being rapidly opened to yield pilocarpic acid. In pilocarpine, by substituting the lactonic ester function with its carbamate equivalent, a much more stable analog—which is as effective as pilocarpine—was obtained [123].

In addition to these classical changes, more exotic groups and radicals have been proposed and tested as valuable carboxylic ester surrogates, like five-membered heterocycles and alkyl oximes. Studies about the design and the synthesis of more metabolically stable muscarinic ligands have shown that it is possible to replace the methyl ester of arecoline **12** and of the quinuclidine **13** with a series of five-membered rings like oxadiazoles, thiadiazoles, triazoles, and tetrazoles to give potent muscarinic agonist (Figure 8.28) [124–130].

Worth noting is the successful replacement of the methyl ester with oxime ethers and *N*-methoxy imidoyl nitrile in the design of muscarinic agonists (Figure 8.28) [131,132].

Several 1,2,4-oxadiazoles and the other five-membered heterocycles are employed for other therapeutic indications, like monoamine transporter and opioid receptors [133,134], 5-HT agonists [135], bradykinin B_1 receptor antagonists [136], and vascular endothelial growth factor receptor (VEGFR-2) inhibitors [137].

Recently, a nice example of the use of the methyl ester bioisosteres in the design of a series of metabolically stable HIV inhibitors has been published (Figure 8.29) [138,139]. Starting from an alkenyldiarylmethane containing three methyl esters, the systematic substitution of an ester at the time with the appropriate biomimetic group led to a molecule with an enhanced metabolic stability in rat plasma ($t_{1/2} = 61$ hour) along with the ability to inhibit HIV-1 reverse transcriptase (Figure 8.29).

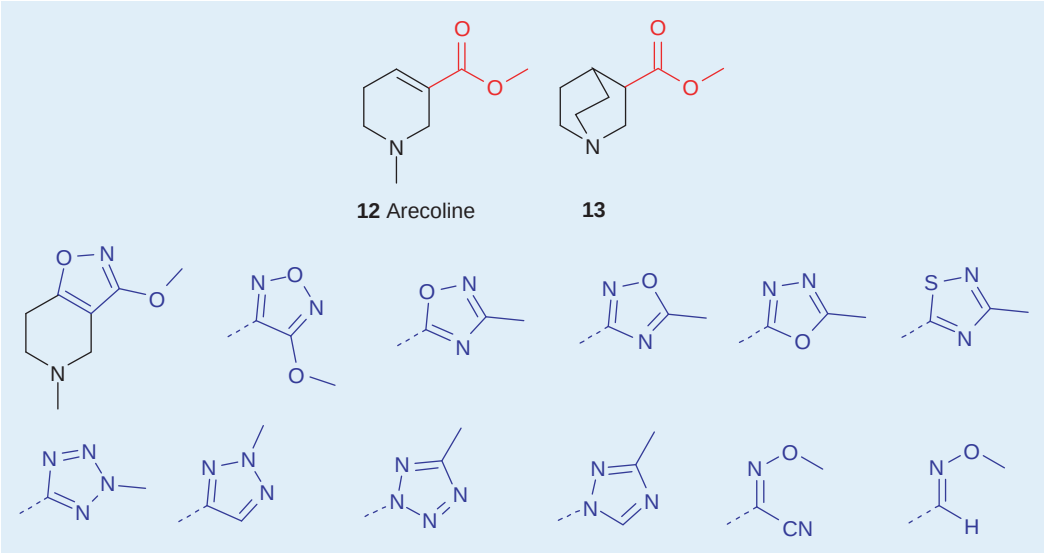


FIGURE 8.28 Muscarinic ligands derived from the ester bioisoster replacement of arecoline **12** and quinuclidine methyl ester **13**.

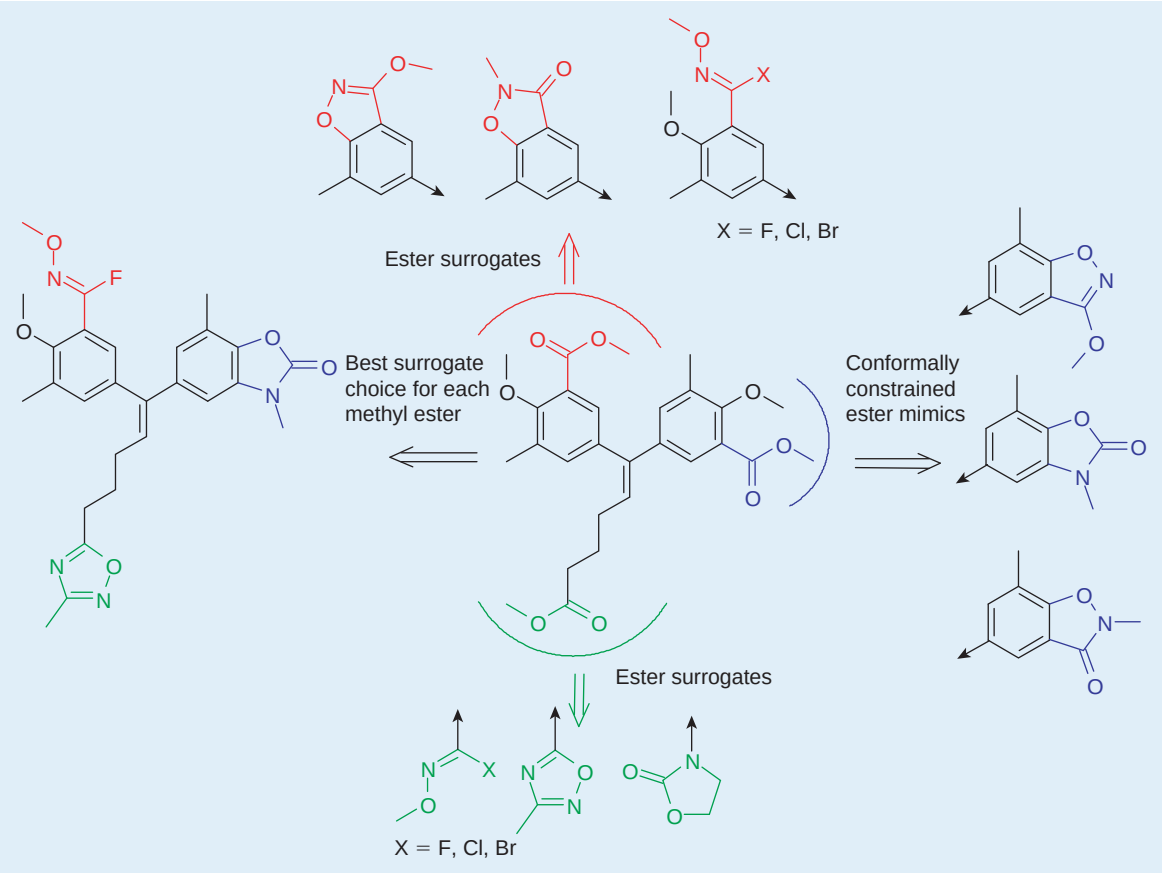


FIGURE 8.29 Design of new metabolically stable HIV inhibitor.

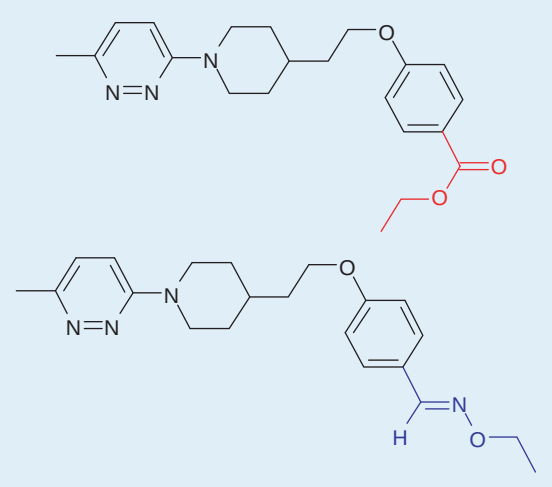


FIGURE 8.30 Ethyl oxime ether as ester replacement gave an improved bioavailability.

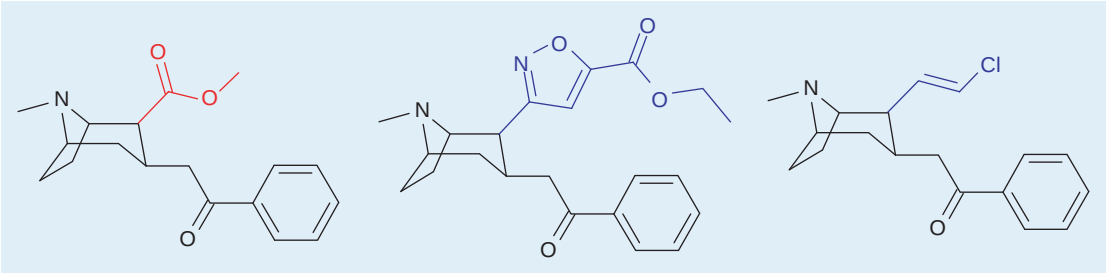


FIGURE 8.31 Replacement in (−)-cocaine of the carbomethoxy group by a carbethoxyisoxazole and a chlorovinyl moiety.

A confirmation of the alkyl oxime ether as metabolically stable isoster of an ester has been reported by Watson et al. in the synthesis of a capsid binder active against Human Rhinovirus [140] (Figure 8.30).

The change in (−)-cocaine of the carbomethoxy substituent into carbethoxyisoxazole doubles the potency in [^{3}H] mazindol binding and [^{3}H] dopamine (DA) uptake (Figure 8.31). Astonishingly, the replacement of the carbomethoxy group by a chlorovinyl moiety produces a comparable gain in potency, thus arguing against the involvement of the carbomethoxy group in H-bonding [141].

Another rather unusual example of ester isosterism is the replacement of the ether oxygen by a fluoro-nitrogen (Figure 8.32a) as mentioned by Lipinski [142]. Other uncommon examples are found in the replacement of the ester function of acetylcholine by exo−endo amidinic functions of 3-aminopyridazines in muscarinic agonists (Figure 8.32b) [143] and of the carbomethoxy group of α-yohimbine (rauwolscine) by a *N*-methylsulfonamide function (Figure 8.32c) [144].

3. *Carboxamide Bioisosteres*

Biologically active molecules containing amide bonds usually suffer from pharmacokinetic liability. In order to increase their stability, bioisosteric transformation of the carboxamide have been performed and yielded a lot of successful examples, especially in the area of peptidomimetic. The isosteric replacements for peptidic bonds have been summarized by Spatola [145] and by Fauchère [146]. The most used and well-established modifications are: *N*-methylation, configuration change (D-configuration at Cα), formation of a retroamide or an α-azapeptide, use of aminoisobutyric or dehydroamino acids, replacement of the amidic bond by an ester [depsipeptide], ketomethylene, hydroxyethylene or thioamide functional group, carba replacement of the amidic carbonyl, and use of an olefinic double bond (Figure 8.33).

In other cases, carboxamides are converted to sulfonamides, as illustrated by the synthesis of the hypoglycemic sulfonyl isostere of glybenclamide [147].

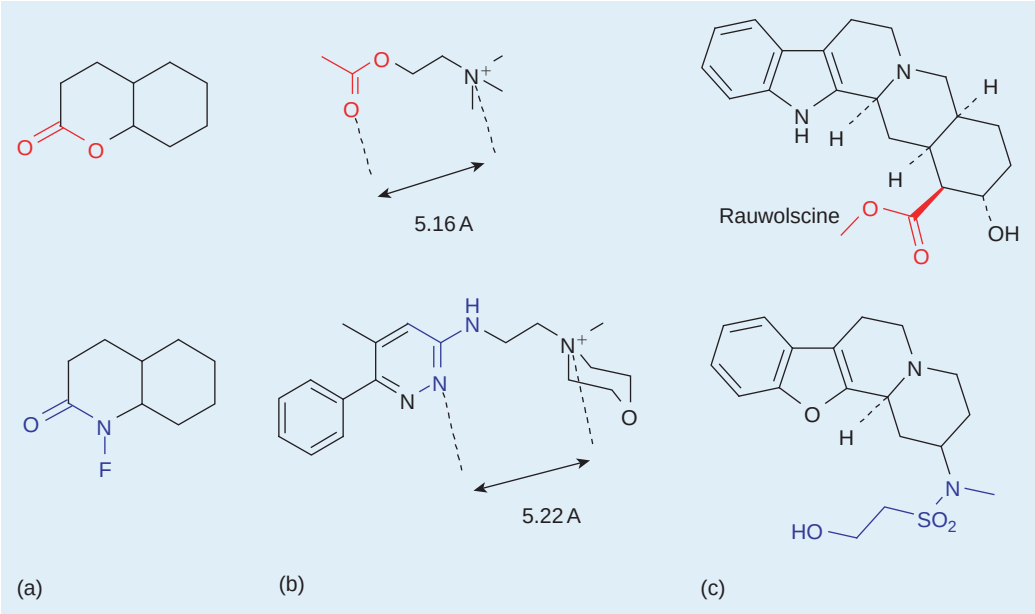


FIGURE 8.32 (a) Replacement of ester ether oxygen by a fluoro-nitrogen; (b) Exo–endo amidine in place of a carboxylic ester functionality; (c) *N*-methylsulfonamide analog of α-yohimbine (rauwolscine).

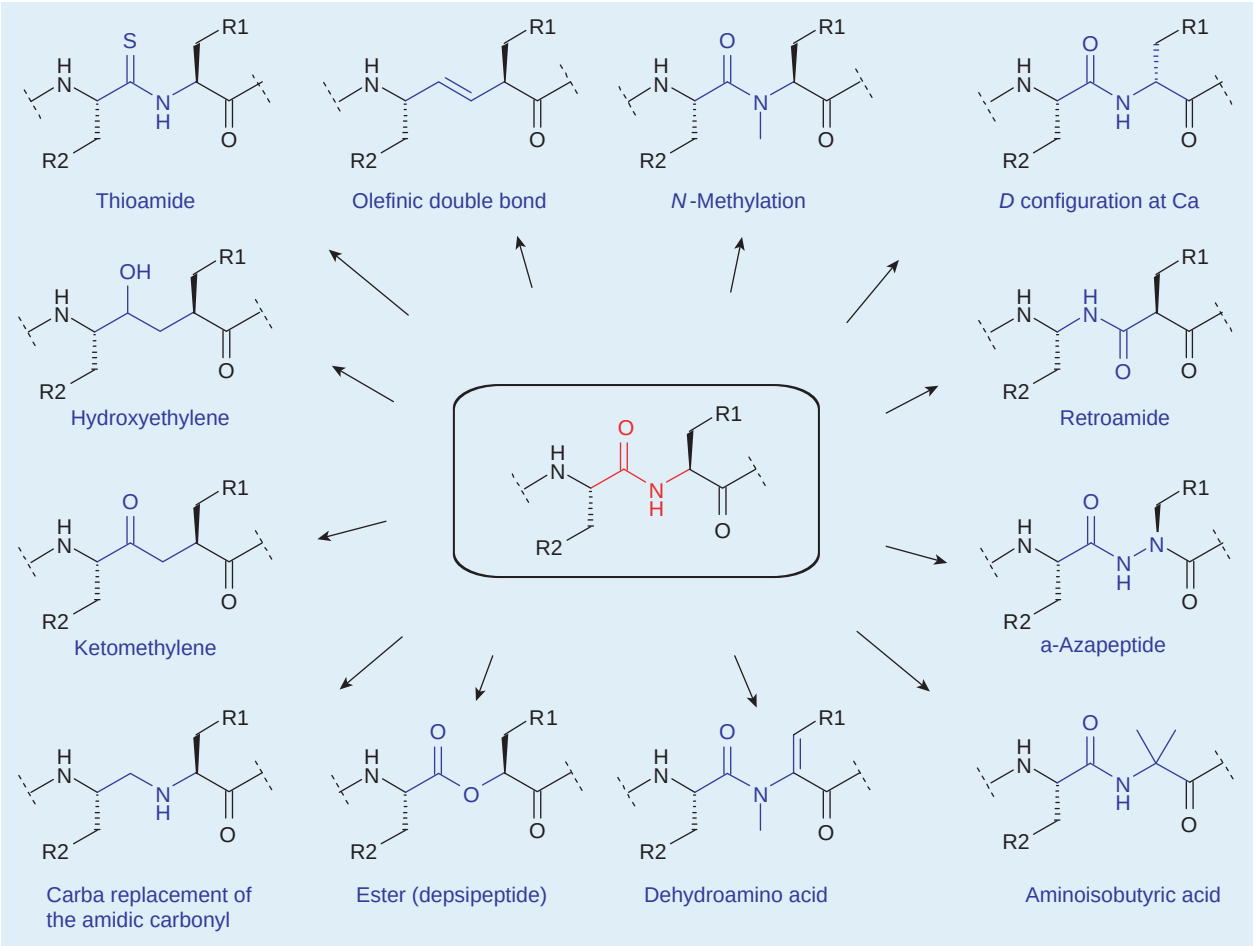


FIGURE 8.33 Well-established isosteric replacements for peptidic bonds.

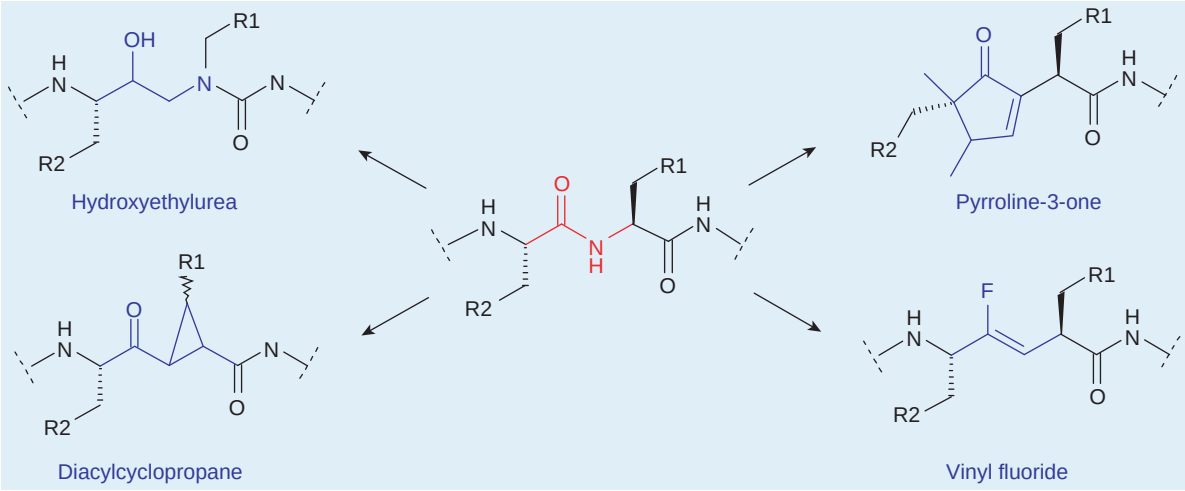
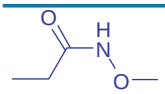
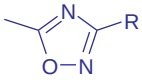
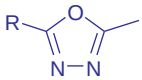
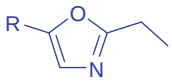
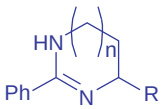
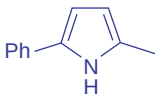


FIGURE 8.34 Unusual isosteric replacements for the peptidic bond.

TABLE 8.15 Heterocyclic Surrogates of the Amide Bond

	1,2,4-Oxadiazoles	β_3 adrenergic receptor agonist	Naylor et al. [152,153]; Feng et al. [154]; Biftu et al. [155]; Parmee et al. [156]
		Fatty acid oxidation inhibitors	Elzein et al. [157]; Koltun et al. [158]
	1,3,4-Oxadiazoles	Phe-Gly peptidomimetics NK1 receptor antagonists	Borg et al. [159]; Ladduwahetty et al. [160]
	Oxazoles	Fatty acid oxidation inhibitors	Elzein et al. [157]; Koltun et al. [158]
	Cyclic amidines	Dopamine D4 receptor agonists	Einsiedel et al. [161,162]
	Pyrroles	D3 antagonists	Einsiedel et al. [163]
	Phenylimidazoles	D4 receptor ligand	Thurkauf et al. [164]

More unusual isosteric replacements for the peptidic bond were proposed and are reported in Figure 8.34. Among them, hydroxyethylureas were used in the design of a novel class of potent HIV-1 protease inhibitors, diacylcyclopropanes in the design of novel renin inhibitors, and pyrroline-3-ones for various proteolytic enzyme inhibitors [148,149]. Vinyl fluorides can probably be considered as representing the closest possible bioisosteres of the peptide bond. The synthetic methods available allow—by an appropriate selection of the precursors—the preparation of analogs of dipeptidic combinations of amino acids bearing no other functionalities in their side chains, for example, Gly, Ala, Val, Phe, and Pro [150]. Vinyl fluorides have been used in the design of bioisosteres of peptide bonds as in the case of the analgesic dipeptide 2, 6,-dimethyl-L-tyrosyl-D-alanine-phenylpropionamide [151].

Other structural modifications of the amide bond proposed to create more chemically stable and orally available molecules include heterocyclic rings, such as 1,2,4-oxadiazoles and 1,3,4-oxadiazoles (Table 8.15). 1,2,4-Oxadiazoles have been used as bioisosteres of the carboxamide moiety in SH2 inhibitors of tyrosine kinase ZAP-70 [165], 5-HT_{1D} receptor agonists [135], histamine H_3 receptor antagonists [166], and 5-HT_3 receptor

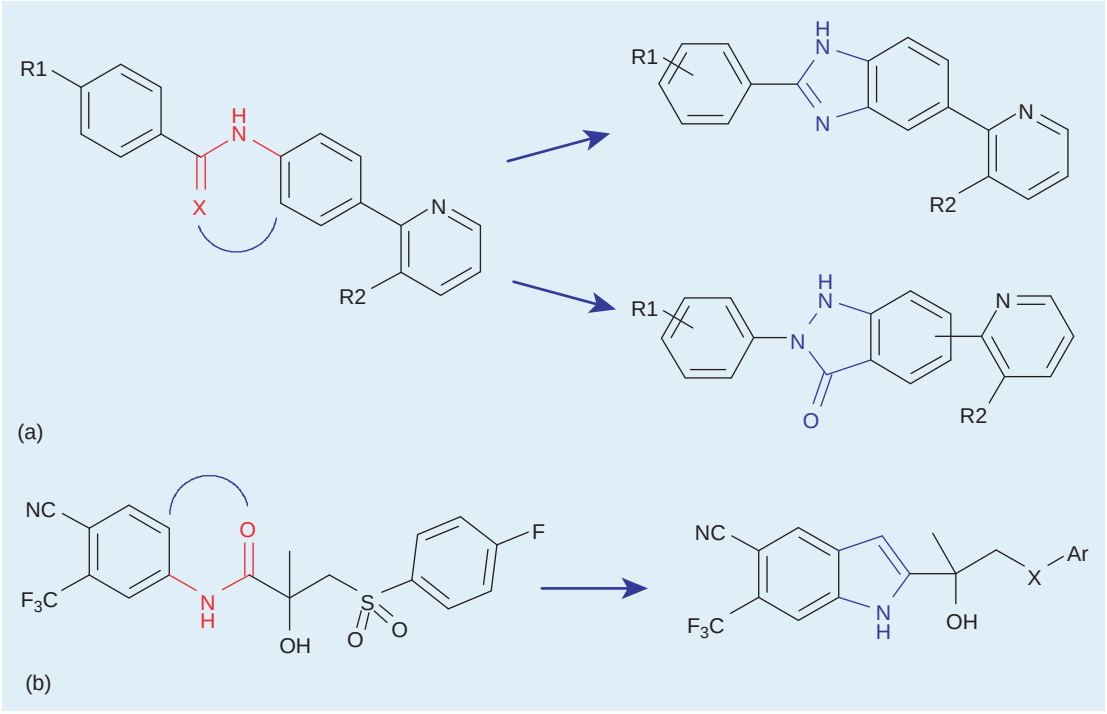


FIGURE 8.35 Benzimidazoles, indazolones, and indoles as carboxamides bioisosteres.

antagonists [167]. An example worthy of note is given by the β_3 adrenergic receptor agonists, where the replacement of the amide bond with 1,2,4-oxadiazoles led to compounds with improved oral bioavailability retaining the β_3AR agonist activity [152–156]. 1,3,4-Oxadiazoles have also been used as amide bond surrogates in several therapeutic areas like benzodiazepine receptor agonists [168], muscarinic receptor agonists [169], and NK1 receptor antagonists [160], and they have also been used as Phe-Gly peptidomimetics [159].

Other heterocyclic replacements of the amide bond include oxazole rings [157,158], cyclic amidines [161,162], pyrrole rings [163], and phenylimidazole [164] (Table 8.15).

In the case of aryl amides, heterocyclisation was explored in order to overcome the pharmacokinetic liability. The search for transient receptor potential vanilloid 1 (TRPV1) antagonists as potential analgesic has led to the discovery of a series of biaryl amides (such as those reported in Figure 8.35a) that suffer from a pharmacokinetic liability [170]. In order to overcome this problem, heterocyclization was explored and has yielded benzimidazoles and indazolones as promising ǀcarboxamides bioisosteres [170] (Figure 8.35a). Other examples of benzimidazoles as aryl amides isosteres has also been reported as melanin-concentrating hormone receptor 1 (MCH R1) antagonists [171,172].

The indole ring has also been reported as a valuable carboxamide bioisoster (Figure 8.35b) of the nonsteroidal antiandrogen bicalutamide [173].

4. *Urea and Thiourea Bioisosteres*

Chemical modifications of the thiourea group of the metiamide led to the discovery of *N*-cyanoguanidine, *N*-nitro-ethene diamine, and *N*-aminosulfonil guanidine as successful bioisosteres of the urea group [44] (Figure 8.19). This successful example of bioisosterism has been applied to the sulfonyl-urea function of torsemide [174], a diuretic loop (Figure 8.36). This investigation led to the design and synthesis of the corresponding thiourea, cyanoguanidine, and diaminonitroethylene group. The structure, the electronic features (charges, molecular electrostatic potential, molecular orbitals), and the lipophilic properties of these three analogs were determined. In conclusion, cyanoguanidine and diaminonitroethylene turned out to be valuable isosteres of the (thio) urea function, as they share common geometric and electronic properties. In addition, they cover a range of lipophilicity that makes them suitable for compound pharmacomodulation, as illustrated by the diuretic and antiepileptic properties of analogs of torsemide.

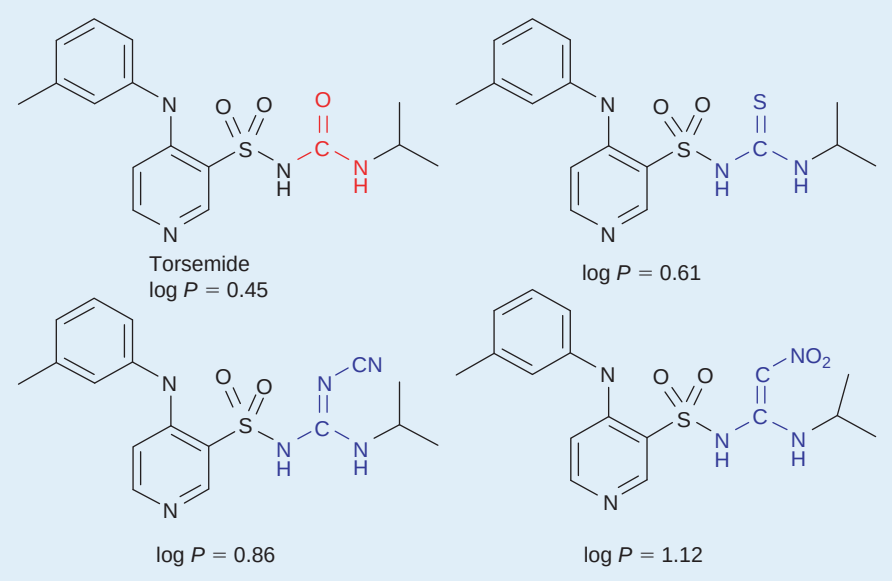


FIGURE 8.36 Structure of torsemide and its urea bioisosteres.

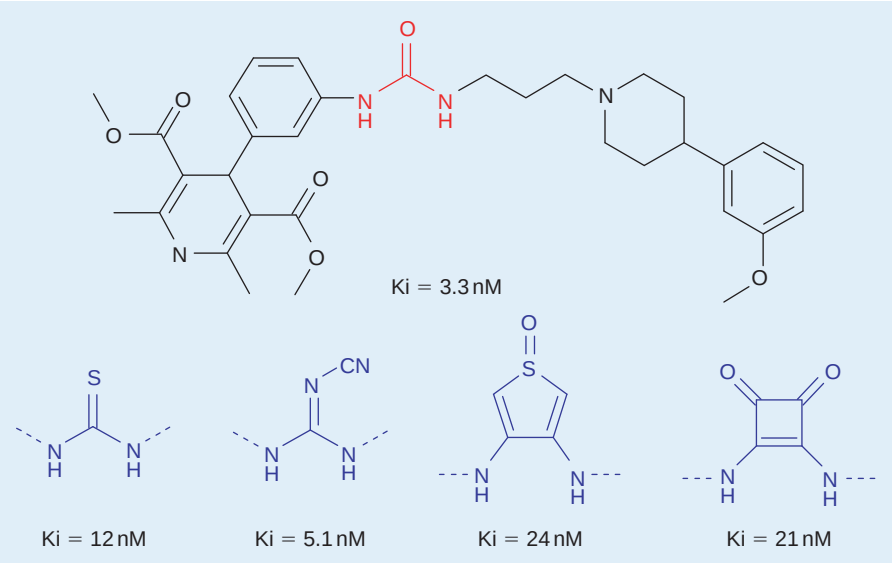


FIGURE 8.37 Binding affinities for various urea bioisosteres acting as antagonists at the NPY Y_1 receptor.

In a study undertaken on antagonists of the dihydropyridine neuropeptide Y_1 receptor, common bioisosteres such as thiourea and cyanoguanidine were examined as urea replacement (Figure 8.37) and also more uncommon derivatives like squaric acid and thiadiazole oxide [175,176]. Both cyanoguanidine and thiourea derivatives demonstrated potent binding affinity at the Y_1 receptor. The two heterocycle replacements, squaric acid and thiadiazole oxide, demonstrated good binding affinity, although both were ten times less active compared to parent urea.

Among more exotic surrogates, the 3,4-diamino thiadiazole dioxide moiety was proposed as a weakly acidic urea equivalent (Figure 8.38) [177]. The similar thiatriazole dioxide is found in the H_2 antagonist tuvatidine (HUK 978). Other bioisosteres are exo–endo amidinic heterocyles bearing an electron-attracting function in the α position [178,179] (Figure 8.38).

5. *Bioisosteres of the Phenol Function*

The optimal bioisosteres of the phenolic function should have approximately the same size as the hydroxyl itself and should have approximately the same acidity range (weak acid) and be able to form hydrogen bonds.

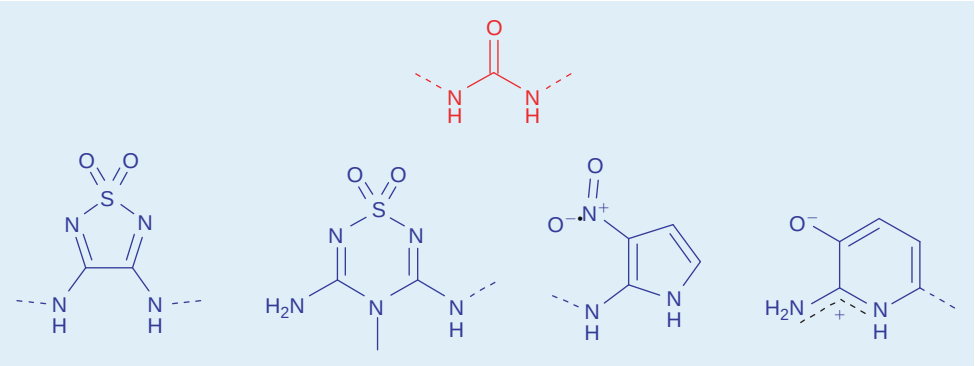


FIGURE 8.38 Less common urea equivalents.

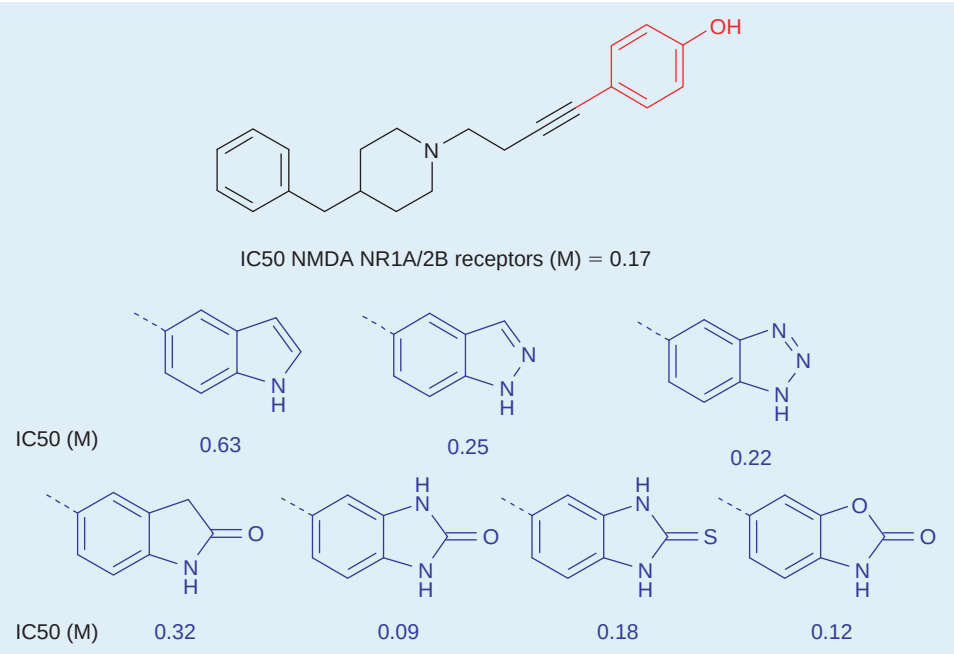


FIGURE 8.39 Bioisosteric replacements of the phenol function in the design of *N*-methyl-D-aspartate (NMDA) receptor antagonists.

The most popular surrogates for phenolic functions are NH groups rendered acidic through the presence of an electron-attracting group. Figure 8.39 shows an application of this bioisostery in the design of *N*-methyl-D-aspartate (NMDA) receptor antagonists [180]. In this case, the replacement of the phenol by heterocyclic NH-containing rings was performed in order to slow metabolism and hence to improve oral bioavailability. Indeed, the potent and NR1A/2B-receptor selective benzimidazolone analog was obtained, demonstrating oral activity in a rodent model of Parkinson's disease at 10 and 30 mg/kg.

Using this analogy, Wu et al. [181] prepared a series of phenolic bioisosteres of benzazepine D_1/D_5 antagonists (Figure 8.40).

Compared to the reference compound SCH 38393, they exhibit similar or more potent activities, high selectivity over D_2-D_4 receptors, and improved *in vivo* pharmacokinetics. The optimization of the hydrogen bond donating capacity of various heterocycles allowed the identification of several potent D_1/D_5 antagonists with high selectivity D_2-D_4, α_{2a} adrenergic receptors, and the 5-HT transporter, keeping at the same time excellent pharmacokinetic profiles.

Other popular phenol bioisoteric substituents [182,183] include methanesulfamide (CH_3SO_2NH-), hydroxymethyl ($HOCH_2-$) or hydroxyisopropyl ($HOC(CH_3)_2-$), various amide groups ($-NHCHO$, $-NHCOCH_3$, $-NHCOC_6H_5$),

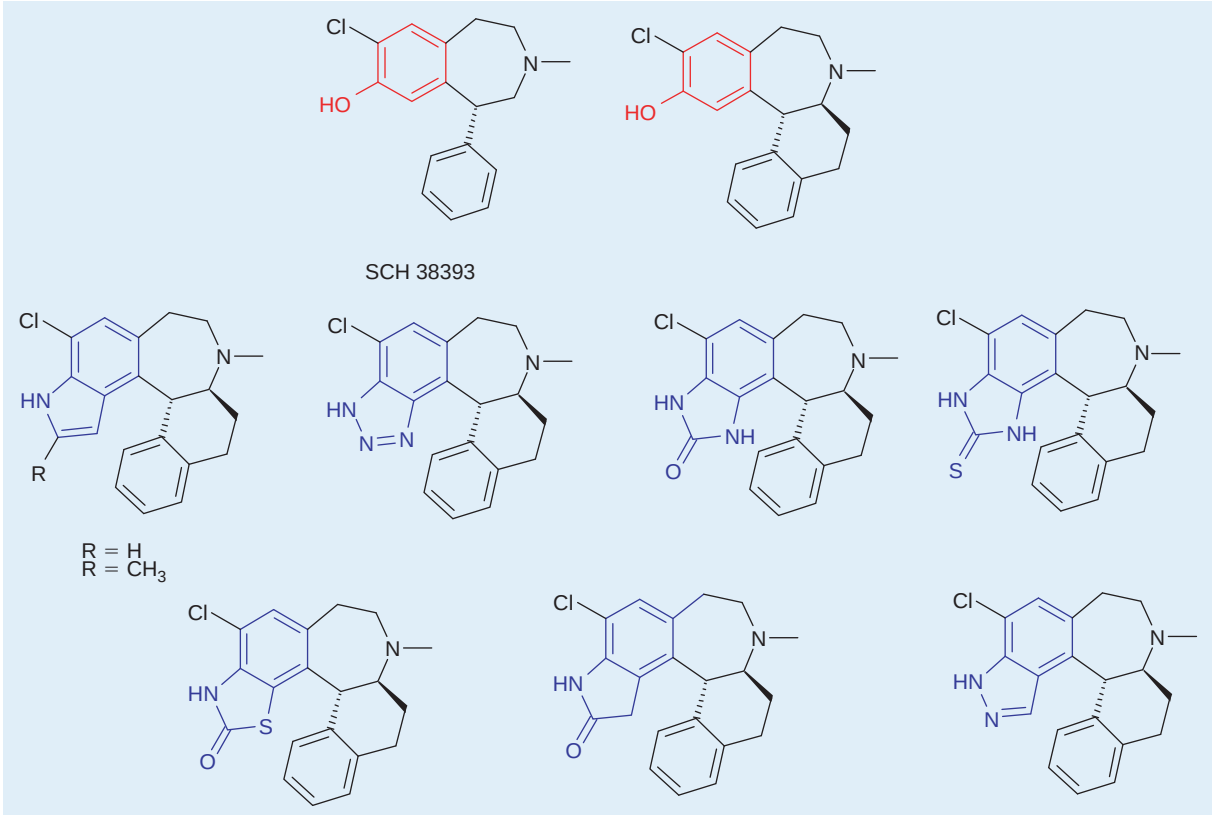


FIGURE 8.40 Phenol bioisosteres of benzazepine D_1/D_5 antagonists.

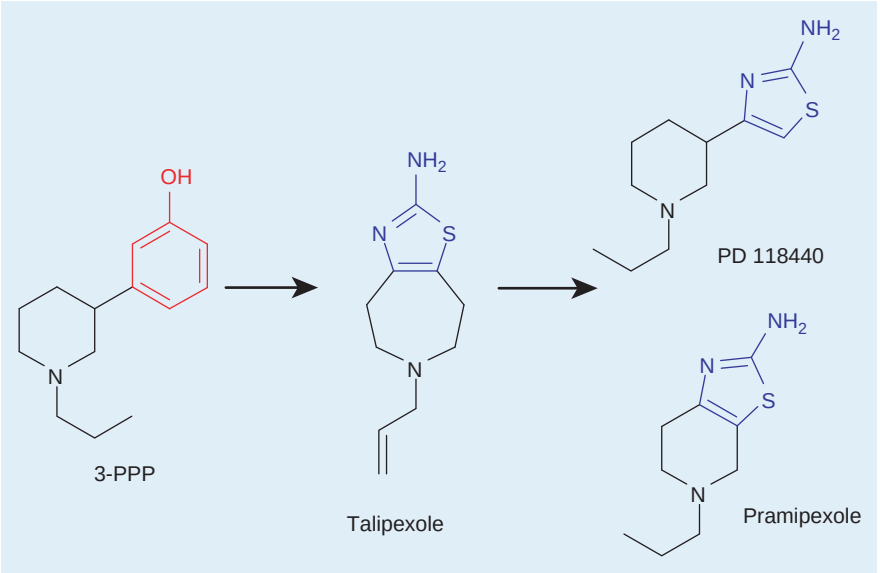


FIGURE 8.41 2-Aminothiazole bioisostery applied to the dopamine agonist 3-PPP.

methanesulfamidomethyl ($CH_3SO_2NHCH_2$—), dimethylaminosulfonamide ($(CH_3)_2NSO_2NH$—), and others groups with an ionizeable proton next to or near an aromatic ring.

Hacksell et al. [184] described the compound (±)-3-(1-propyl-3-piperidinyl)phenol [(±)-3-PPP] as highly selective for presynaptic brain DA receptors. However, the clinical potential of these molecules (Figure 8.41) as antischizophrenic or antiparkinsonian agents may be limited by their relatively low oral bioavailability and their short duration of action [185]. Thus, a heterocyclic analog of 3-PPP might retain some of its pharmacological

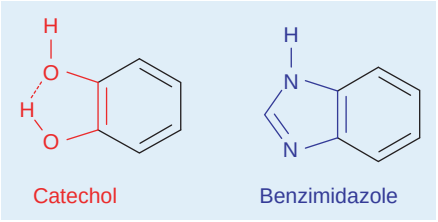


FIGURE 8.42 The bioisostery catechol/benzimidazole.

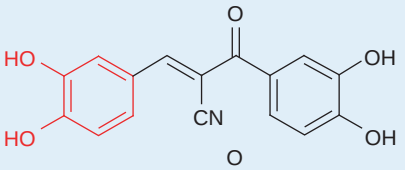
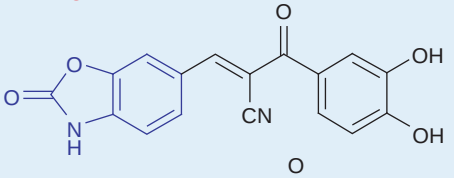
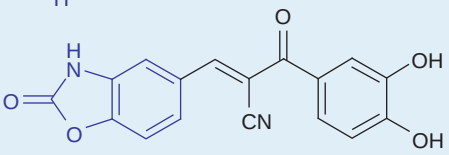
Compounds	Name	IC_{50} Values [μM]			
		IGF-1R	IR	SRC	PKB
	AG 538	0.06	0.12	2	76
	GB 19	0.37	0.6	1.5	21
	AGL 2263	0.43	0.4	2.2	55

FIGURE 8.43 Inhibition of IGF-1R and other kinases by AG 538 bioisosteres and analogs.

activity while having improved oral bioavailability and duration of action [186]. In the case of talipexole (B-HT 920) (Figure 8.41), a DA agonist that has been reported to be selective for DA autoreceptor, 2-aminothiazolyl moiety was used as a replacement for the phenol ring.

This successful example led successively to the discovery of PD 118440, a dopamine-autoreceptor agonist [186], and pramipexole, a DA agonist with preference for the DA D_3 (over DA D_2) receptor [187–189].

6. *Catechol Bioisosteres*

All the previous examples can be applied to the catechol family. Many of these bioisosteres share with the catechols the ability to chelate metal atoms and to form hydrogen-bonded second rings. In benzimidazole, this H-bond ring is mimicked by way of a covalent ring structure (Figure 8.42).

A recent case described the synthesis of more stable bioisosteres of inhibitors of the insulin-like growth factor-1 receptor kinase (Figure 8.43). Based on the structure of AG 538 [190], which contains two catechol rings and is sensitive to oxidation in cell, a new series of kinase inhibitors were developed. The catechol moiety was replaced by a benzoxazolone ring resistant to oxidation yielding two compounds, GB19 and AGL2263, which maintain the same potency as AG 538.

Until now, 2-aminothiazole derivatives or other aromatic heterocyclic systems have been used as a bioisosteric surrogate for the catechol nucleus (cf. studies around the pramipexole) [187]. Hübner et al. [191–193] evaluated nonaromatic catechol bioisosteres (Figure 8.44). They were able to demonstrate that the π-electronic system of the nonaromatic endiyne FAUC 88 and 73 can be used as an efficient bioisostere for the catechol fragment of dopamine. Combined with conformational rigidization, this approach led to dopamine-receptor agonists with high affinity for the subtypes of D_2-family and especially for D_3.

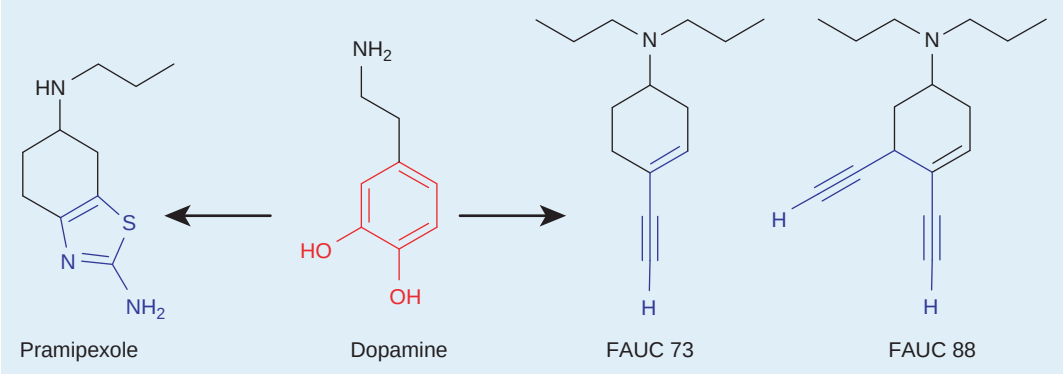


FIGURE 8.44 Aromatic and non-aromatic dopamine isosteres.

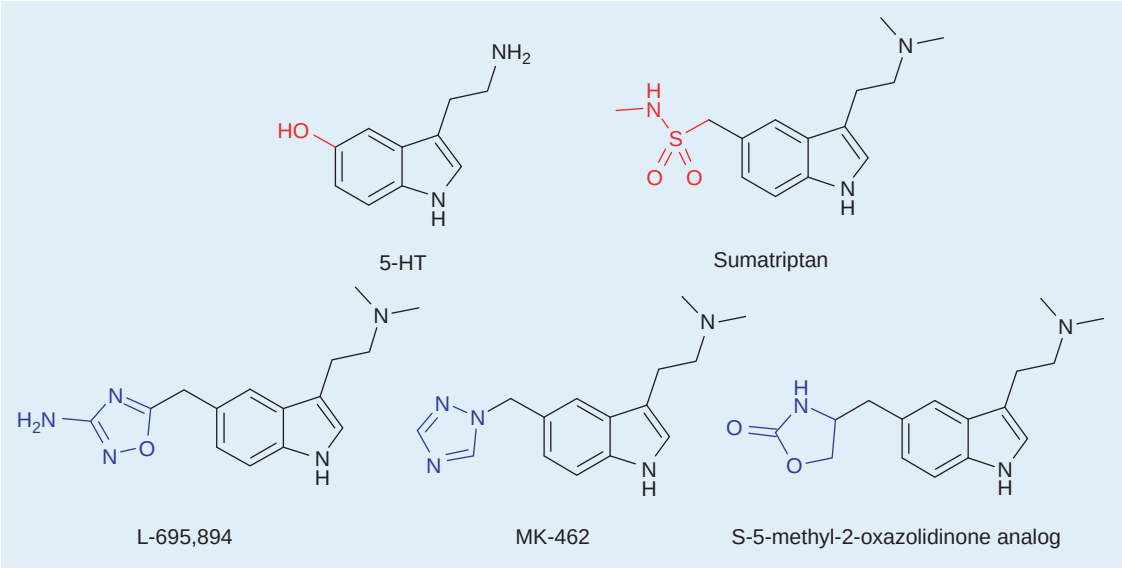


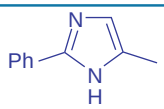
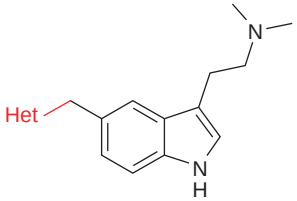
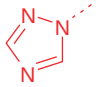
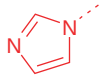
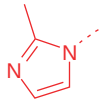
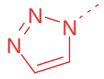
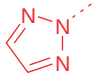
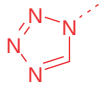
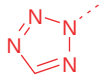
FIGURE 8.45 Replacement of the phenolic function of serotonine by various functions able to serve as H-bond acceptors.

7. *Sulfonamides Bioisosteres*

The introduction of the 5-HT_{1D} receptor agonist sumatriptan for the acute treatment of migraine has marked an intense research effort to discover more potent and selective 5-HT_{1D} receptor agonists with improved pharmacokinetic profiles [194]. The H-bond acceptor ability of the 5-position is crucial for 5-HT_{1D} receptor affinity and selectivity as was revealed by reference compounds (5-HT and sumatriptan). The search for other analogs with better bioavailability led to the replacement of the methyl sulfonamide with different heterocycles that are able to be H-bond acceptors. The amino-oxadiazole analog, L-695,894 (Figure 8.45), has a better oral bioavailability but a lack of good selectivity. In fact, it has a significant affinity for 5-HT_{2A} and 5-HT_{2C} receptors [135]. Alternative five-membered heteroaromatic rings were explored to find analogs with log Ds < -0.5 in order to minimize the central nervous system penetration [194]. Simple unsubstituted imidazoles, triazoles, and tetrazoles (Table 8.16) have high affinity and good selectivity for the 5-HT_{1D} receptor. The 1,2,4-triazole 1, MK-462, was shown to have the optimal pharmacokinetic profile with rapid oral absorption and high bioavailability.

Using a different approach, Glen et al. [195] built a pharmacophore based on the pharmacological activities of a series of novel C- and *N*-linked hydantoin analogs. Their model was used as a framework for the design of a diverse series of analogs with good affinity and selectivity for 5-HT_{1D}. Suitable with the constraints required for good oral absorption, a potent selective 5-HT_{1D} agonist, S-5-methyl-2-oxazolidinone analog, has been described.

TABLE 8.16 Displacement of [3H]-5-HT to 5-HT1D Recognition Sites in Pig Caudate Membranes by *N*-Linked Imidazoles, Triazoles and Tetrazoles and Standard 5-HT1D Agonists

Compound	Het	pIC_{50}	log D[a]
5-HT		8.0	
Sumatriptan		7.7	− 1.17
L-695,894		7.6	− 0.67
1, MK-462		7.3	− 0.74
2		7.5	− 0.53
3		7.2	− 0.74
4		7.3	− 0.70
5		6.6	− 0.20
6		7.4	− 0.64
7		7.4	− 0.34

[a] *log* P *measured at pH 7.4.*

The dipolar azido group is a bioisostere [196] of the SO_2NH_2 and SO_2Me hydrogen bonding pharmacophores present in many selective COX-2 inhibitors, and the azido analogs **14** and **15** are useful biochemical agents for photoaffinity labeling of the COX-2 enzyme (Figure 8.46).

F. Reversal of Functional Groups

The reversal of the peptidic functional groups is often used in peptide chemistry. The obtained retropeptides are generally more resistant to enzymatic attacks (Figure 8.33) [146,197]. For thiorphan and *retro*-thiorphan, an identical binding mode to the zinc protease thermolysin was demonstrated [198]. Similar inhibition values for thermolysin and neutral endopeptidase were observed, whereas, for another zinc protease, angiotensin-converting enzyme (ACE), noticeable differences for inhibition were found (Figure 8.47).

But the strategy of functional inversion can also be applied to nonpeptidic compounds. A historical example is the change from orthoform to neo-orthoform (orthocaine; Figure 8.48). The unwanted side effects often encountered with aromatic para-amino substituted compounds ("para effects," essentially of allergic origin) are

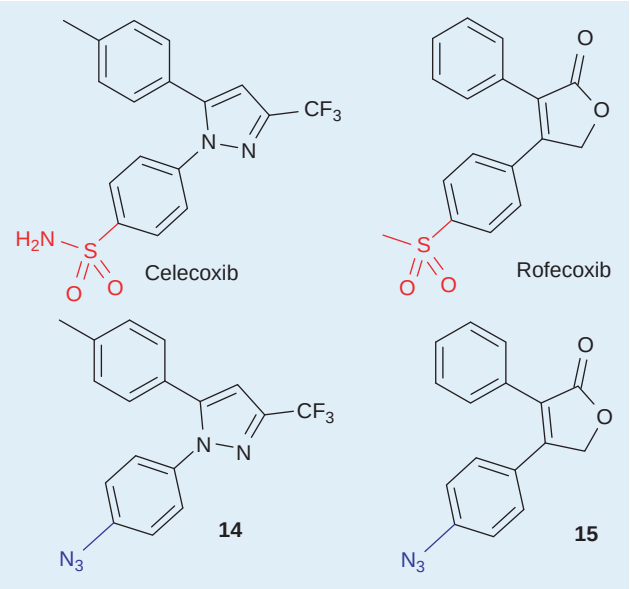
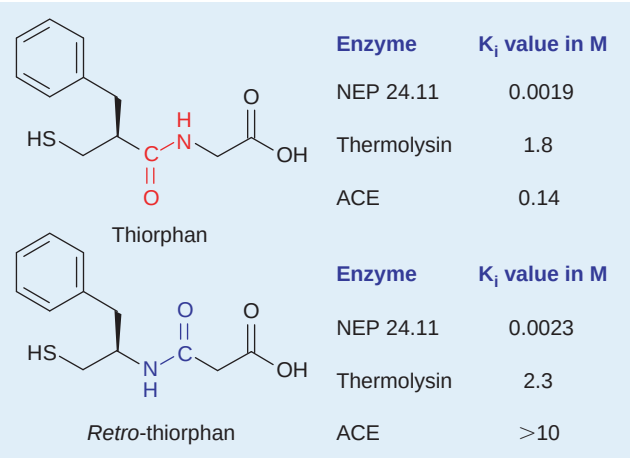


FIGURE 8.46 Selective COX-2 inhibitors and their azido bioisosteres.



Enzyme	K_i value in M
NEP 24.11	0.0019
Thermolysin	1.8
ACE	0.14

Enzyme	K_i value in M
NEP 24.11	0.0023
Thermolysin	2.3
ACE	>10

FIGURE 8.47 Inhibition values of thiorphan and *retro*-thiorphan for three zinc proteases.

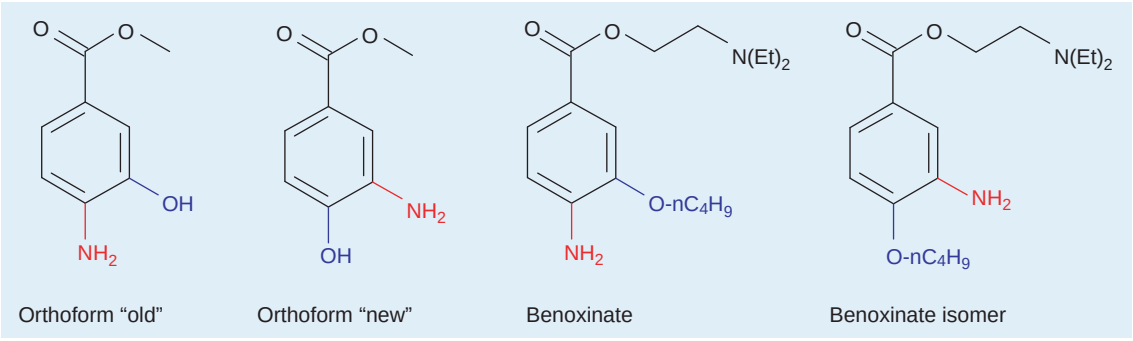


FIGURE 8.48 Positional isomery in local anesthetics [199].

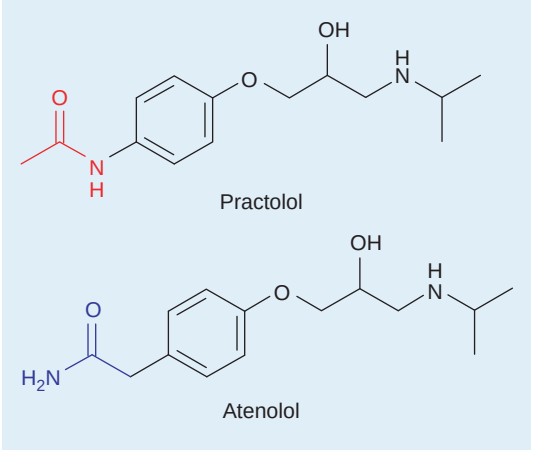


FIGURE 8.49 Reversal of functional goups in practolol yields atenolol.

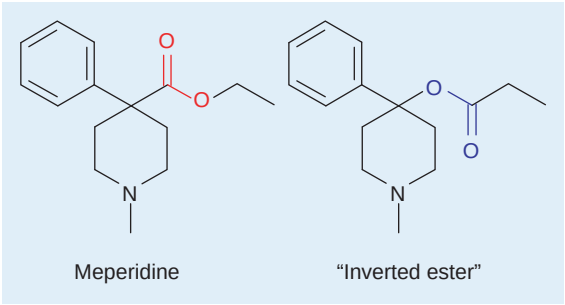


FIGURE 8.50 Meperidine and the corresponding inverted ester [17].

abolished in the meta amino isomer, whereas the local anesthetic activity is maintained. Similarly the "meta" isomer of benoxinate has a local anesthetic activity identical to that benoxinate itself [199].

The β-blocking agent practolol, which was one of the first cardioselective β-blockers, was rapidly replaced by its isomeric analog atenolol, which presents fewer side effects (Figure 8.49).

The inversion of the ester function of *meperidine* leads to 1-methyl-4-phenyl-4-propionoxy piperidine (Figure 8.50), which is five times more potent as an analgesic drug than meperidine and represents the model compound of the series of inverted esters [17].

The change from indomethacin to clometacin, although representing a clean example of functional group reversal, causes more profound alterations than that shown in the previous examples (Figure 8.51). At a first glance, this change can even seem too drastic; however, in rotating the molecule of clometacin by 180°, the resemblance with the parent molecule becomes evident.

Indomethacin is mainly used as a nonsteroidal anti-inflammatory agent and occasionally as an analgesic; clometacin on the other hand is usually recommended as an analgesic and shows weak anti-inflammatory properties. Applied to serotonin, a similar reversal of a functional moiety yielded $5HT_{2C}$ receptor selective agonists [200] (Figure 8.52).

IV. SCAFFOLD HOPPING

Scaffold hopping is a useful technique used by medicinal chemists to discover structurally novel compounds starting from known active molecules by changing the molecular backbone of the template. Medicinal chemists quite often have the need of "jumping" in a different chemical space: new chemical entities offer the choice of easier chemical accessibility and so possibility of speeding up the lead optimization process. Switching to a different molecular backbone can help avoid some undesirable ADME-tox properties. Scaffold hopping also comes

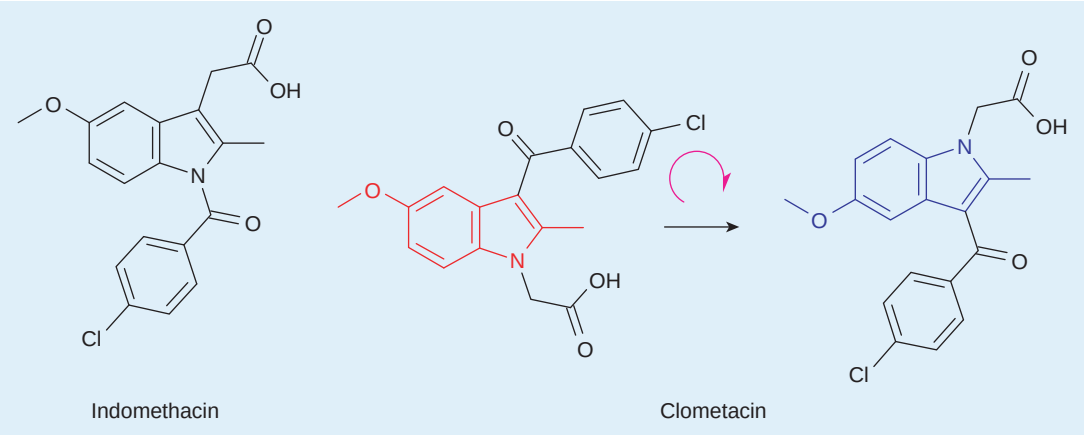


FIGURE 8.51 Functional inversion applied to indomethacin.

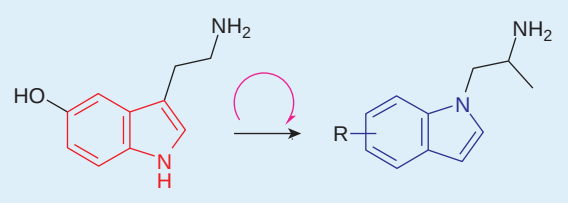


FIGURE 8.52 Serotonin analogs resulting from a functional reversal [200].

to a rescue to move from complex natural molecules to easier synthesizable small molecules. Moreover "jumping" in a new chemical series isofunctional with a given one is fundamental when new IP position is sought.

A. Successful Examples of Serendipitous Scaffold Hopping

The concept of scaffold hopping relies upon the assumption that structurally and chemically distinct templates interact with the same receptor, inducing equivalent biological activity. This seems to contradict the leading principle of the lead optimization process and drug design that is based on the concept that chemical similarity is reflected by similar biological activity. On the other hand, this concept is not always true. Structural and chemical analogy does not always predict similar biological affinity, as it is shown in Figure 8.53. In fact, promethazine, imipramine, and chlorpromazine, which are strictly similar from a structural point of view, interact with different receptors and present completly different therapeutic profiles acting respectively as H1 antagonist, uptake inhibitor, and dopamine antagonist.

Indeed scaffold hopping can be interpreted as the broadest expansion of the bioisosterism concept. The idea of scaffold hopping is indeed not new; it is rather a new term to designate known experimental evidences. A retrospective analysis of the marketed drugs and/or of the ligands acting on the same receptors brings up some interesting examples of structurally different molecules eliciting the same biological activity. A nice example is given by the nonsteroidal anti-inflammatory drugs (NSAIDs) COX-2 inhibitors, as illustrated in Figure 8.54, which lists four known traditional NSAIDs: indomethacin, nimesulide, diclofenac, and celecoxib.

Actually, the GABA-A ligands at the benzodiazepine site give a nice picture of matching and mismatching of scaffold hopping and bioisosterism (Figure 8.55). Looking at diazepam, zolpidem, zopiclone, and zaleplon, a structural analogy is barely found. Despite this, they are all GABA-A ligands acting as full agonists. The same can be said about the inverse agonists DMCM, **16**, **17**, FG 8094, and **18**: chemically different compounds have comparable biological activity. On the other hand, similar compounds like FG 8094, Bretazenil, and Ro 15-1788 do not give the same biological response.

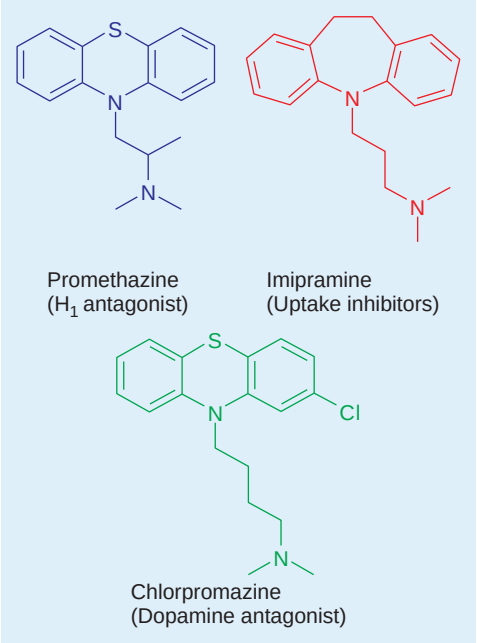


FIGURE 8.53 Structure and chemical similarity are not always synonymous of similar biologically active compounds.

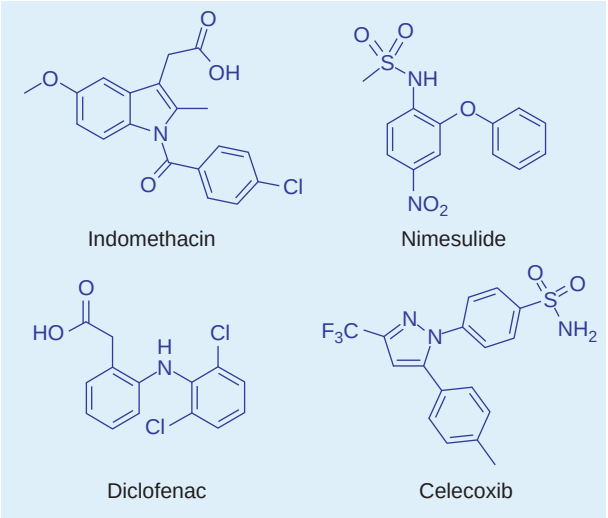


FIGURE 8.54 An example of scaffold hopping: four structurally different molecules have similar therapeutic properties by interacting with the same receptor.

B. Scaffold Hopping and Virtual Screening

The examples cited above illustrate that similar biological activity can be obtained with structurally different molecules, which seems to contradict the bioisosterism principle discussed in this chapter. The serendipitous examples of scaffold hopping described above have been generated either by high-throughput screening (HTS) or by selective optimization of side activity (the SOSA approach) [201]. In other studies, the "hopping" in a different molecular scaffold has relied upon the intuition and/or the experience of the medicinal chemist working on the project. Most of the virtual screening studies aiming to hop in isofunctional molecules are based on the bioisosterism concept that structural likeness predicts analogous biological activity. At a first glance, it can be concluded that a rational approach to scaffold hopping is not possible. Indeed molecular modeling and

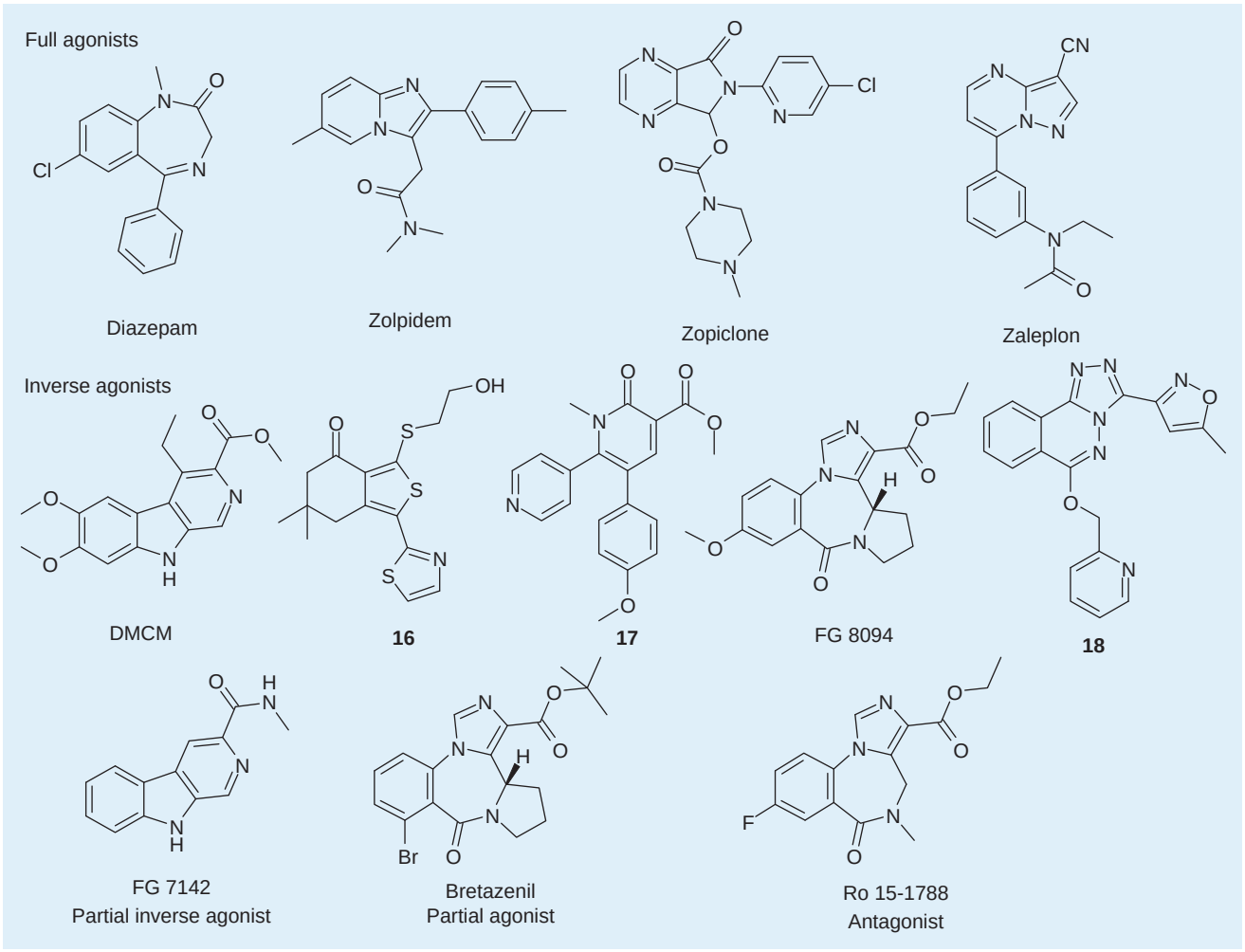


FIGURE 8.55 GABA-A ligands binding at benzodiazepine site.

chemoinformatics experts have set up a lot of software and methods to address the paradox of scaffold hopping and bioisosterism to find a logical process [202,203].

Among the virtual screening techniques reported so far, four types of computational approaches for scaffold hopping can be distinguished: (a) shape matching, (b) pharmacophore searching, (c) fragment replacement, and (d) similarity searching. Most of the programs offer combinations of several approaches, and they are used—according to the case scenario of a given project—either as on a ligand-based virtual screening base or on a structure (receptor)-based screening. Other studies have reported a combination of the two methods to increase the performance of the virtual screening. In a case where more active ligands are known but no information is known about the receptor structure, it will be more useful to proceed through shape matching combined with a pharmacophore searching. By contrast, when the receptor structure is known, a virtual screening based on molecular docking is a powerful and successful technique. A detailed description of the computational approaches of scaffold hopping goes beyond the scope of this chapter and only a few illustrative examples will be reported. Figure 8.56 reports an example of scaffold hopping related to the design of a novel cholecystokinin-2 (CCK_2) antagonist [204]. Starting from a known series of indoles (**19**) with high *in vitro* affinity and selectivity but poor bioavailability, a CCK_2 pharmacophore receptor has been identified and used to design new scaffolds to overcome the problems associated with the high level of biliary elimination of the compound **19**. Two news series of pyrroles and imidazoles (**20**) were identified, and although they are not as potent as the original indole derivative *in vitro*, they have greater potential as drug candidates because of their reduced tendency to biliary excretion.

Scaffold hopping is an extremely useful technique to switch from peptide ligands to small molecule ligands. A fragment-based docking procedure followed by a substructure search were used to identify a set of small

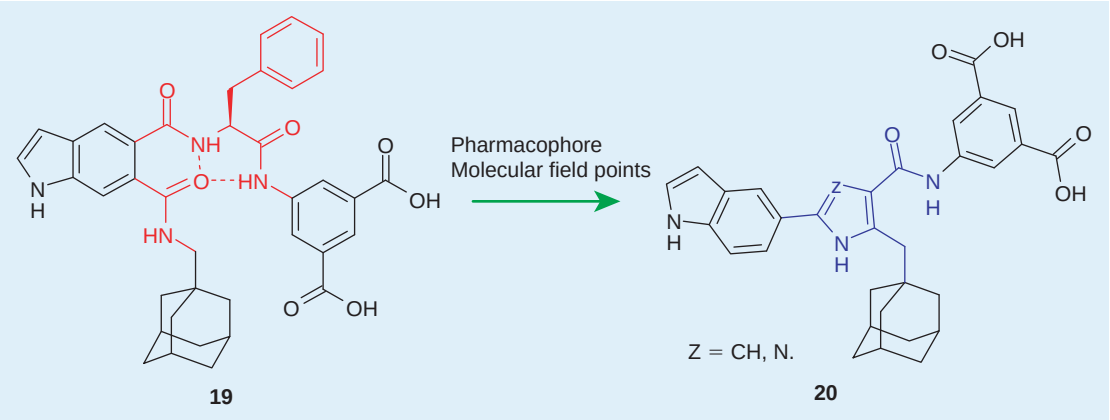


FIGURE 8.56 Design of a novel CCK_2 antagonist via a scaffold hopping approach from a known series of indoles.

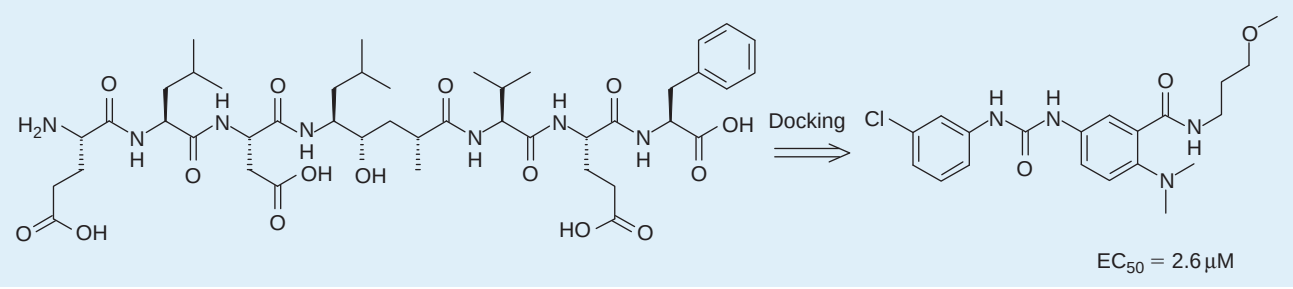


FIGURE 8.57 Discovery of non-peptide inhibitor of β-secretase by high-throughput docking.

molecules as β-secretase inhibitors [205]. Figure 8.57 shows one ligand identified with virtual screening that may serve as a starting point for further optimization for Alzheimer's disease.

Most of the studies based on virtual screening use a hierarchical approach combining several conceptually different techniques [206]. The most common combination used is a 3D pharmacophore modeling followed by a docking study. Figure 8.58 contains an illustrative example of this combined use: a set of ten known inhibitors of Human Rhinovirus Coat protein was used to design a 3D pharmacophore that was subsequently used for a docking study of the Maybridge DB [206].

A set of ten compounds was retrieved according to the score fit of the molecules. The ten molecules were tested *in vitro*, and all of them displayed antirhinoviral activity in the micromolar range. Particularly promising results were achieved with the compound reported in Figure 8.58, which shows an EC_{50} of 4.3 μM.

V. ANALYSIS OF THE MODIFICATIONS RESULTING FROM ISOSTERISM

It is rare that the replacement of a part of a molecule by an isosteric or bioisosteric group leads to a strictly identical active principle. In practice, that is not even sought, and one prefers that the new compound produces a change as compared to the parent molecule. In general, the isosteric replacement, even though it represents a subtle structural change, results in a modified profile: some properties of the parent molecule will remain unaltered and others will be changed. Bioisosterism will be productive if it increases the potency, the selectivity, and the bioavailability, or decreases the toxicity and undesirable effects of the compound. In proceeding to isosteric modifications, one will focus predominantly on a given parameter (e.g., structural, electronic, hydrophilic), but it is all but impossible not to alter several parameters simultaneously.

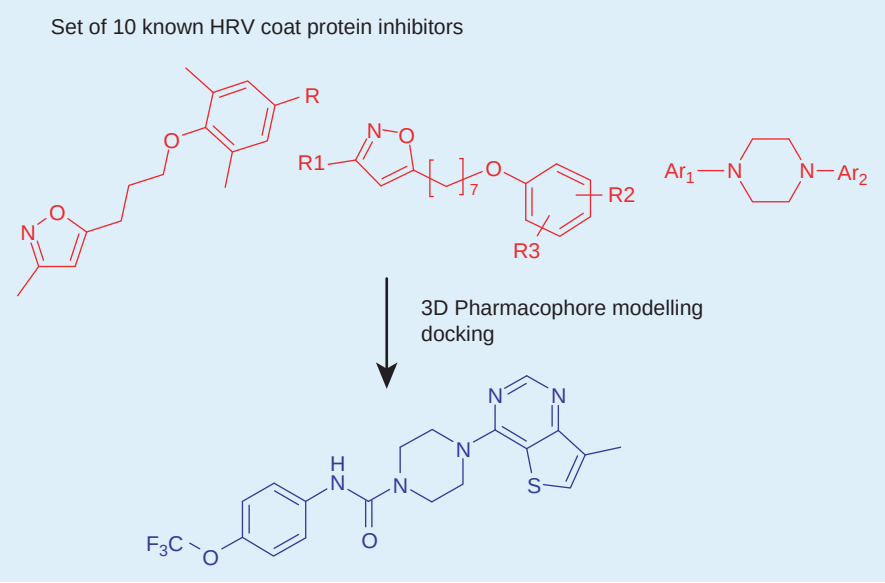


FIGURE 8.58 A set of ten known HRV coat protein inhibitors was used for 3D pharmacophore modeling followed by docking study afforded new candidate structure suitable for an optimization process.

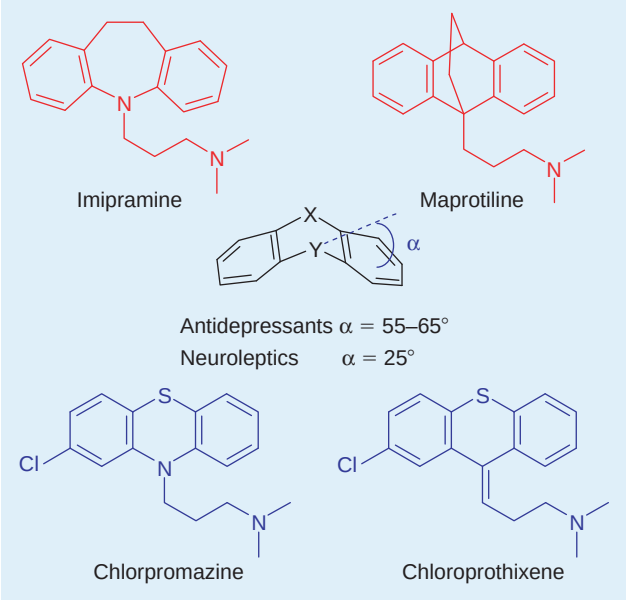


FIGURE 8.59 The tricyclic antidepressants (imipramine and maprotiline) are characterized by an dihedral angle of 55° to 65° between the two benzo rings, this angle is only 25° for the tricyclic neuroleptics (chlorpromazine, chlorprothixene) [207].

A. Structural Parameters

These will be important when the portion of the molecule involved in the isosteric change serves to maintain other functions in a particular geometry. That is the case for tricyclic psychotropic drugs (Figure 8.59).

In the two antidepressants (imipramine and maprotiline), the bioisosterism is geometrical insofar as the dihedral angle α formed by the two benzo rings is comparable: $\alpha = 65°$ for the dibenzazepine and $\alpha = 55°$ for the dibenzocycloheptadiene [207]. This same angle is only 25° for the neuroleptic phenothiazines and the thioxanthenes. In these examples, the part of the molecule modified by isosterism is not involved in the interaction with the receptor. It serves only to position the other elements of the molecule correctly.

The structure of various bioisosteric retinoic acid receptor agonists highlights the dominantly geometric parameter of this bioisostery (Figure 8.60) [208].

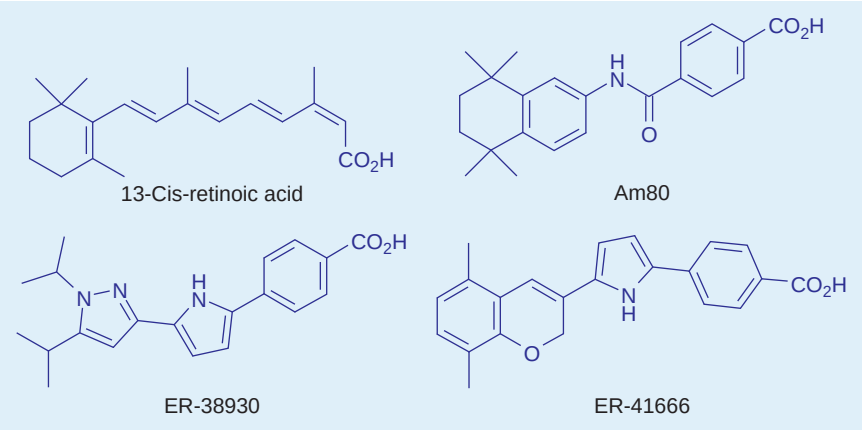


FIGURE 8.60 Bioisostery in its broadest sense.

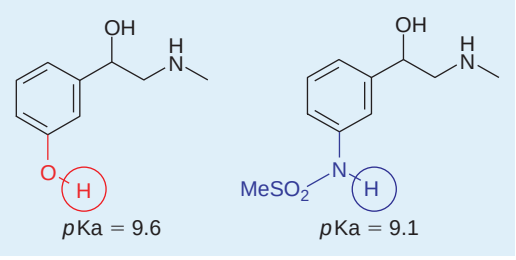


FIGURE 8.61 An example of bioisosterism, or nonclassical isosterism, the methylsulfonamide substituent has comparable acidity to the phenolic hydroxyl group [197].

B. Electronic Parameters

Electronic parameters govern the nature and the quality of ligand–receptor or ligand–enzyme interactions. The relevant parameters will be inductive or mesomeric effects, polarizability, pK_a, capacity to form hydrogen bonds, etc. Despite their very different substituents in the *meta* position, the two epinephrine analogs (Figure. 8.61) exert comparable biological effects: they are both β-adrenergic agonists. In fact, the key parameter resides in the very close pK_a values [197].

C. Solubility Parameters

When the functional group involved in the isosteric change plays a role in the absorption, the distribution, or the excretion of the active molecule, the hydrophilic–lipophilic parameters become important.

Imagine in an active molecule the replacement of $-CF_3$ ($\pi = +0.88$) by $-CN$ ($\pi = -0.57$) (Figure 8.62). The electron-attracting effect of the two groups will be comparable, but the molecule with the cyano function will clearly be more hydrophilic. This loss in lipophilicity can then be corrected by attaching elsewhere on the molecule a propyl, isopropyl, or cyclopropyl group.

D. Anomalies in Isosterism

In this section, two applications of the bioisosterism concept that imply unusual behaviors of commonly encountered atoms or groups are discussed.

1. Fluorine–Hydrogen Isosterism

It is an anomaly that fluorine does not resemble other halogens—notably chlorine—while on the other hand it often mimics an atom of hydrogen [209].

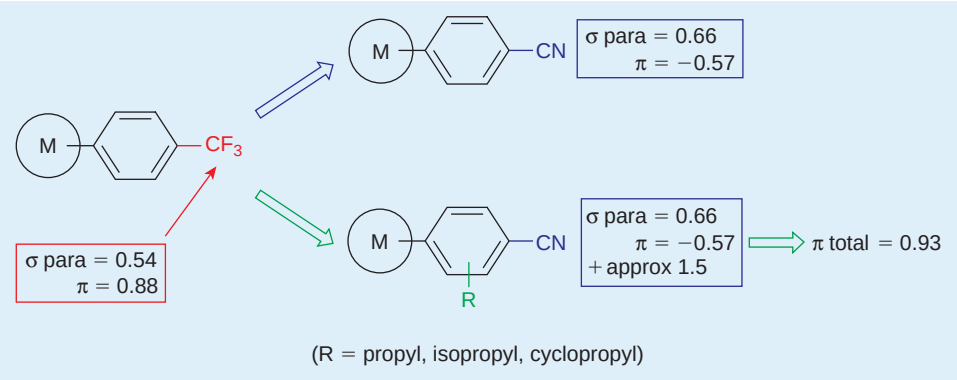


FIGURE 8.62 The loss in lipophilicity resulting from the bioisosteric exchange of a CF_3 for a CN has to be compensated by the equivalent of a three carbon residue.

TABLE 8.17 Fluorine–Hydrogen Isosterism. Observe the Comparable Sizes of the Two Atoms, whereas Chlorine is Close to the Methyl and Trifluoromethyl

Parameter	H	F	Cl	CH_3	CF_3
Atomic radius	0.29	0.64	0.99	–	–
Van der Waals radius	1.2	1.35	1.80	<2	<2
Molecular refractivity	1.03	0.92	6.03	5.65	5.02
Electronic effect (*para* σ)[a]	0.00	0.06	0.23	−0.17	0.54
Resonance effect (*R*)[a]	0.00	−0.34	−0.15	−0.13	0.19
Electronic effect (σ^*)[b]	–	3.08	2.68	0.00	2.85

[a]*For aromatic systems.*
[b]*for aliphatic systems.*

1. *Steric aspects*: The fluorine atom is considerably smaller than the rest of the halogen atoms. Seen from the steric point of view, it resembles more hydrogen than chlorine (Table 8.17). Effectively, fluoro-derivatives differ from the other halogenated derivatives because fluorine forms particularly stable bonds with carbon and—in contrast to other halogens—is only rarely ionized or displaced. Because it is both chemically inert and of small size, organic fluorine is often compared to hydrogen.

   One understands especially the incorporation by living organisms of fluoroacetic acid in place of acetic [210] acid or of 5-fluoro-nicotinic acid and 5-fluoro-uracil as antimetabolites. This "fraudulent" incorporation leads to lethal syntheses [211]. This is generally not the case with the corresponding chlorinated, brominated, or iodinated analogs.
2. *Electronic aspects*: Fluorine is the most electronegative of the halogens (Table 8.17) and forms particularly stable bonds with carbon atoms. This chemical inertia explains why fluoro-derivatives are more resistant to metabolic degradation (Figure 8.63). Thus, for the β-haloalkylamines (nitrogen mustards), the alkylating activity is lost when chlorine or bromine are replaced by fluorine or by hydrogen [212].

   The isosterism between H and F will therefore often serve to give analogs that are more resistant to metabolic degradation (obstructive halogenation: flunarizine and in flufenisal; Figure 8.63). Similarly the CF_3 group is biostable, whereas CH_3 is easily oxidized [209].
3. *Absence of d orbitals*: Another difference between fluorine and the other halogens comes from the absence of a d orbital for fluorine, and thus its incapacity to participate in resonance effects with a donor of π electrons (Figure 8.64).

   This explains why *para*-fluorophenol is slightly less acidic than phenol, while for other *para*-halogenated phenols the acidity changes in parallel with the atomic number (Table 8.18 [209]).
4. *Case study*: A good example of continuous variation of activity in halogenated compounds is provided by a series of antihistaminic drugs related to tripelennamine (Figure 8.65; X = H).

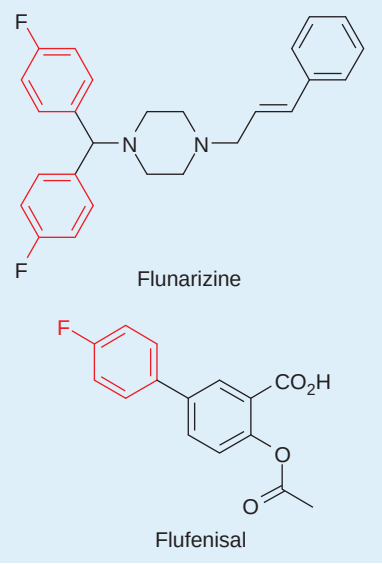


FIGURE 8.63 In flunarizine and in flufenisal fluorine atoms in para position prevent metabolic hydroxylation.

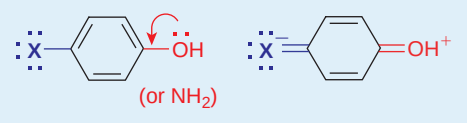


FIGURE 8.64 The resonance between the OH lone pair and the X group is not possible if X = F.

TABLE 8.18 Dissociation Constants of *p*-Halogenated Phenols [209]

Compound	Dissociation constant K_a 3 10 [210]
Phenol	0.32
p-Fluorophenol	0.26
p-Chlorophenol	1.32
p-Bromophenol	1.55
p-Iodophenol	2.19

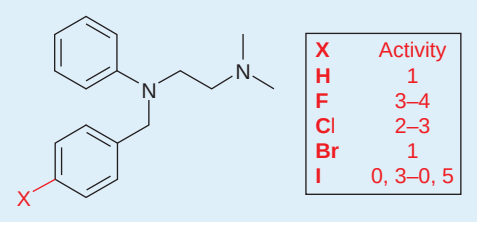


FIGURE 8.65 Variation of activity in a series of antihistaminic compounds as a function of the halogated para-substituent [213].

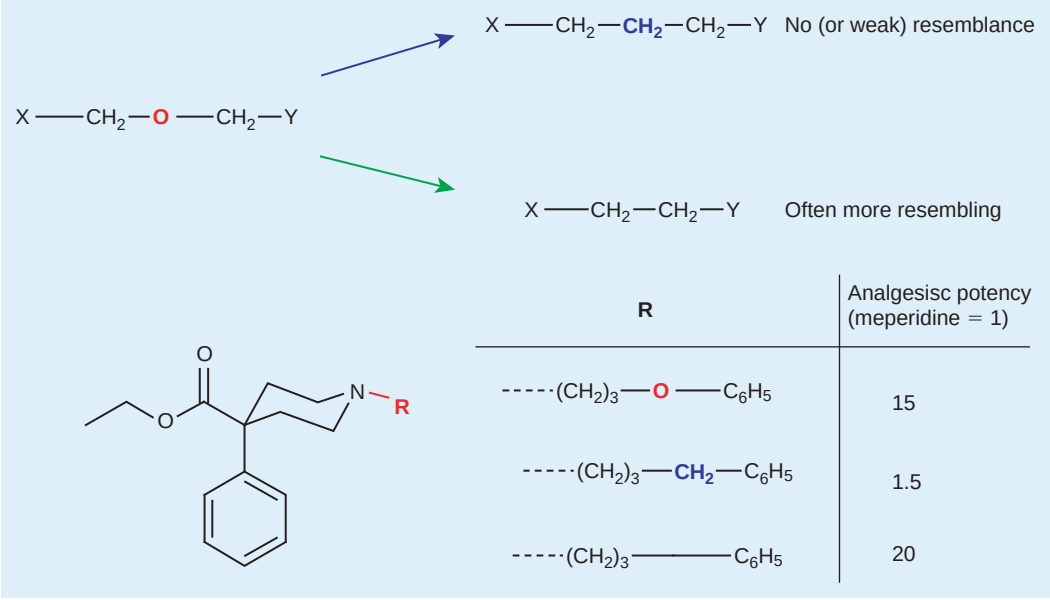


FIGURE 8.66 Friedman's ether oxygen-methylene group paradox [2].

Apparently we are dealing here with a classical isosteric series: F, Cl, Br, I, but one sensitive to steric hindrance in the para position. Probably what happens *in vivo* is para-hydroxylation of the benzene ring. The best candidate becomes then the para-fluoro compound, since it is not bulkier than the unsubstituted compound, while being biostable.

2. *Exchange of Ether Oxygen and Methylene Group*

Ether oxygen atoms and methylene groups possess a similar tetrahedral structure and should normally be isosteric. In fact, the isosterism between O and CH_2 very often yields anomalous results, which brought Friedman [214] to the interesting observation "that the omission of the ether oxygen changes biological activity much less in some cases than the replacement by the isosteric methylene group" (Figure 8.66). In the meperidine series, for example, the change from the *N*-phenoxypropyl derivative to the isosteric phenoxybutyl decreases the analgesic potency by a factor of ten, whereas the omission of the ether oxygen yields a slightly more potent compound [17]. A list of seven other examples is given by Schatz in the second edition of Burger's *Medicinal Chemistry* [215].

The explanation for this anomalous behavior may be that the omission of the ether oxygen yields a closer compound in terms of lipophilicity than its replacement by a methylene. An example that can be compared to Friedman's paradox is found in the resemblance of the phenylethyl type β-blockers (e.g., dichloroisoprenaline, sotalol) to the phenoxypropanol type (e.g., practolol, acebutolol).

VI. MINOR METALLOIDS-TOXIC ISOSTERES

In this section we describe some "exotic" applications of the bioisostery concept, implying the utilization of unusual elements such as silicon, boron, selenium, arsenic, and antimony. The use of those elements as bioisosteres of carbon in existing drugs is a different approach, enabling the introduction of a new drug-like chemical space into the drug discovery and development process.

A. Carbon–Silicon Bioisosterism

Silicon is directly below carbon in the periodic table, so according to the Erlenmeyer's expansion of the isosterism concept, carbon and silicon can be considered as true isosteres. Sila-substitution (C/Si exchange) of biologically active substances is an approach to search for new drug-like candidates with improved pharmacological properties and stronger IP position. The application of this isosterism remains, however, limited. For reviews on the subject, see Fessenden and Fessenden [216], Tacke and Zilch [217,218], Ricci et al, [219], and Showell and Mills [220].

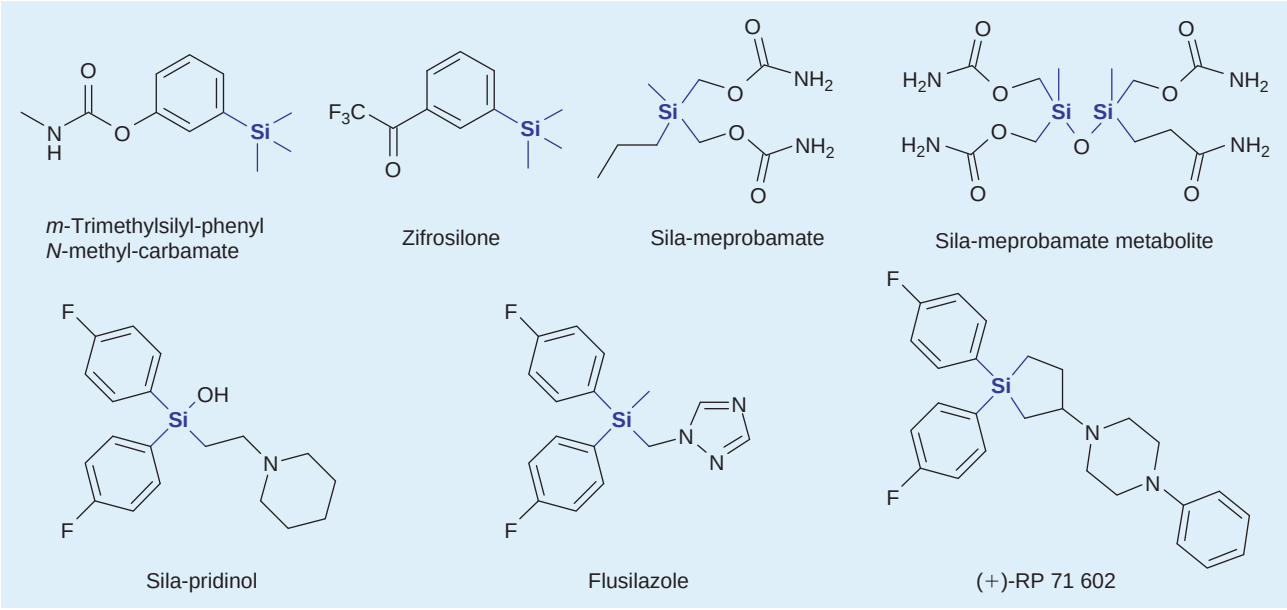


FIGURE 8.67 Organosilicon active substances.

Silicon is more electropositive than carbon (and even more if compared to oxygen and nitrogen) and the covalent silicon–carbon bonds in the sp^3 hybridization state are 20 percent longer than the corresponding carbon–carbon bond. This difference is reflected on the properties of the silaisosteres compared to the corresponding carbon analogs. For example, acidity, the hydrogen bond strength of the silanol is more favorable as a donor than that of carbinol. In this respect, the silanol isosteres may be beneficial in those pharmacophores where the H-bonding has a predominant role. From a lipophilicity point of view, the silicon-containing analogs are more lipophilic than their carbon analogs, and when an increase in lipophilicity is sought this stratagem could be used. The C = O double bond of a ketone is quite stable, whereas the formation of a Si = O double bond is disfavored over its hydrate form, the silicon diol. The chemistry of silicon relies mainly on the chemistry of single bonds, which have led to its appropriate use as a tetrahedral bioisostere of carbon. In Figure 8.67, some of examples of the sila-substitution of existing drugs are shown. Among these, m-trimethylsilyl-phenyl *N*-methylcarbamate and m-trimethylsilyl-α-trifluoroacetophenone (zifrosilone) are acetylcholinesterase inhibitors [221–223], sila-meprobamate is a CNS depressant [224], sila-pridinol is an anticholinergic [225], flusilazole is a fungicide for agricultural use [226], and (+)-RP 71,602 is a potent and selective 5-HT_{2A} antagonist [227].

Figure 8.68 shows a silicon-containing hypocholesterolemic squalene epoxidase inhibitor; [228,229] the silicon analog of the α_2-adrenergic antagonist atipamezole; [230] a highly potent, stable, and CNS-penetrating silatecan; [231] some ACE inhibitors; [232] HIV protease inhibitors; [233] the (*R*)-sila-analog of the antidepressant venlafaxine; [234] and a trimethylsilylpyrazole as novel inhibitor of p38 MAP kinase [235] (mitogen-activating protein). As for the preceding molecules, the silicon atom of these compounds is quaternary and thus expected to be less sensitive to metabolic degradation (Figure 8.68).

Compared to their carbon bioisosteres, silicon-containing molecules are more sensitive to hydrolysis and to nucleophilic attack in general. Even when located in the center of a quaternary structure, the risk exists that the silicon atom will be attacked. Thus, 1-chloro-1-sila-bicyclo-(2,2,1)-heptane can still be hydrolyzed by an attack on the vacant d orbital [236]. This attack is lateral and therefore possible even in cases where the corresponding carbon derivative would have been inert toward a S_N2 reaction (Figure 8.69).

This sensitivity toward lateral attacks explains the four times shorter duration of action of sila-meprobamate compared to its carbon isostere on a model of tranquillizing activity in mice (rotarod test, potentiation of hexobarbital-induced sleep, and intraperitoneal injection) [224]. On the other hand, when given orally, sila-meprobamate is practically inactive. One of the first metabolites formed has been characterized as being a di-siloxane [216] (Figure 8.67). For the two phenyl-trimethylsilyl-derived AChE inhibitors, the

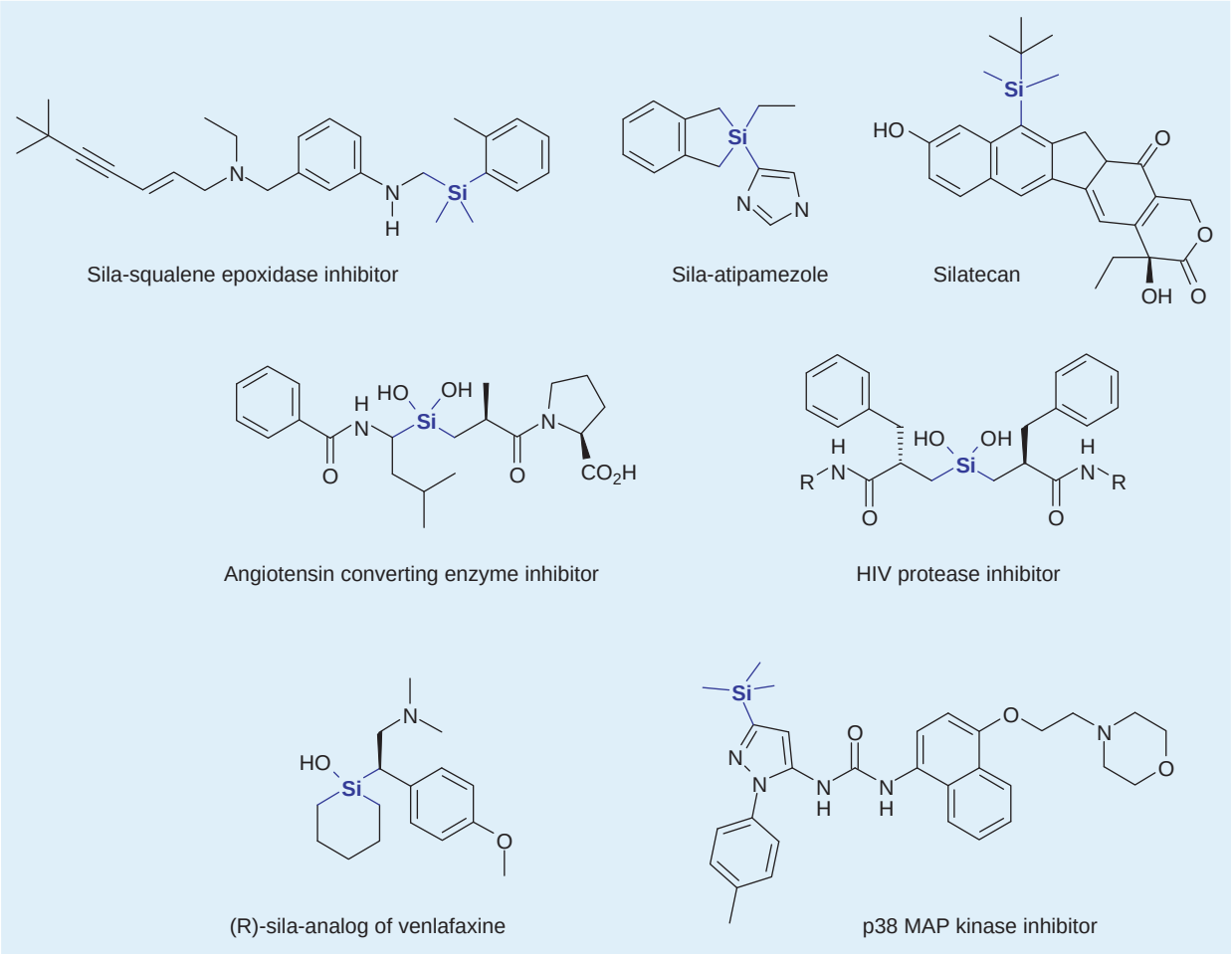


FIGURE 8.68 Structures of silicon-containing compounds which are biologically active.

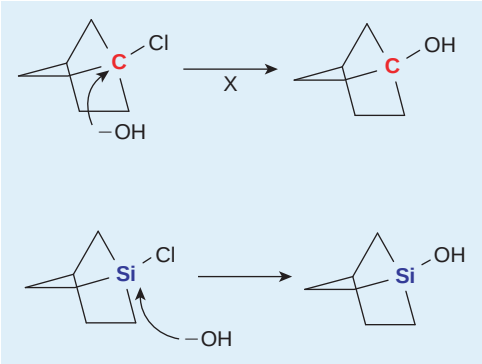


FIGURE 8.69 Due to the presence of a vacant d orbital, a lateral attack can substitute for dorsal attacks in organo-silicon derivatives [236].

rather positively charged trimethyl-silyl group mimics the trimethyl-ammonium function present in acetylcholine. For these compounds, metabolic oxidation does not take place on the silicon but on one of methyl groups (Si-CH_3 → Si−CH_2−OH) [221].

B. Carbon−Boron Isosterism

Boron is essential for plant growth and development. In medicinal chemistry, its main use is related to that of coupling reagent. The most important employment of boron as a drug is in the treatment of certain tumors by

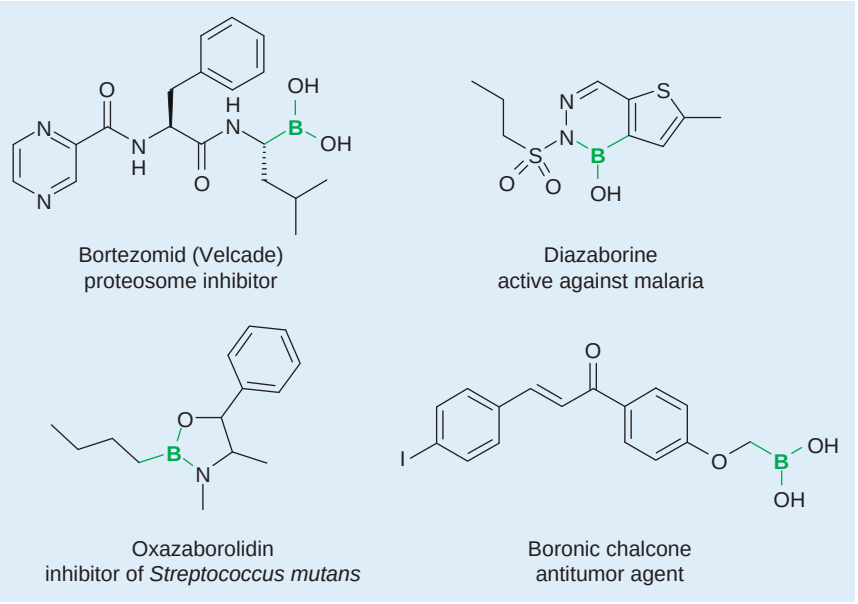


FIGURE 8.70 Boron-containing molecules with a biological activity.

Boron Neutron Capture Therapy (BNCT) [237–239]. BNCT is a binary treatment modality that can selectively irradiate tumor tissue. The delivery to the tumor cells of a drug containing the ^{10}B isotope followed by a low energy irradiation (neutrons) causes the ^{10}B to split, releasing an alpha particle a lithium nucleus. These products of the $^{10}B(n,\alpha)^{7}Li$ reaction are very damaging to cells but have a combined path length in tissue of approximately 14 μm, or roughly the diameter of two cells. Thus, most of the ionizing energy imparted to tissue is localized to 10B-loaded cells. The problem here is to ensure a sufficient concentration of the product in the tumor being treated.

Few medicines based on boron are known. In general, boric acid or a boronic acid serve to esterify an α-diol or an ortho-diphenol. This is the case for the emetic antimony borotartrates of the ancient pharmacopoeias; for the injectable catecholamine solutions; for tolboxane [240], which is close to meprobamate and was commercially available as a tranquillizer some decades ago; and for the phenylboronic esters of chloramphenicol [241]. Boromycine was the first natural product containing boron isolated [242]. It is a complex between boric acid and a polyhydroxylated tetradentate macrocycle [243]. Another natural product is aplasmomycin with antibiotic properties [244].

Some of the boron-containing molecules biologically active are reported in Figure 8.70. Some boronic analogs of amino acids were prepared as chymotrypsine and elastase inhibitors [245], and more recently as an antineoplastic agent (i.e., Velcade, a proteosome inhibitor reported in Figure 8.70) [246]. Carboxyboranes complexed with a tertiary amine ($R_3N.H_2B$-COOH) are considered boron amino acid isosteres due to the isoelectronic features. Compounds-containing carboxy boranes have shown anticancer, hypolidemic, and antifungal activity [247]. Diazaborines are active against malaria [248], and oxazaborolidines possess antibacterial activity (Figure 8.70) [249]. Boronic chalcones are reported to be antitumor agents (Figure 8.70) [250].

Organoboron derivatives—even more than organosilicon compounds—are sensitive to hydrolytic degradation, which always leads to the final formation of boric acid. But boric acid has teratogenic properties in chickens. It produces the same malformations as those produced by a riboflavin (vitamin B_2) deficiency, and the administration of riboflavin prevents these toxic effects [251,252]. The mechanism by which boric acid produces a deficiency in riboflavin is not known. In humans, the chronic utilization of boron derivatives results in cases of borism (dry skin, cutaneous eruptions, and gastric troubles) [253].

C. Bioisosteries Involving Selenium

Selenium can be considered the best isoster of sulfur, as it is just below it in the periodic table. These two atoms have very similar physical properties: the radius of selenium is only 12.5 percent bigger than that of sulfur, and their electronegativity is rather similar.

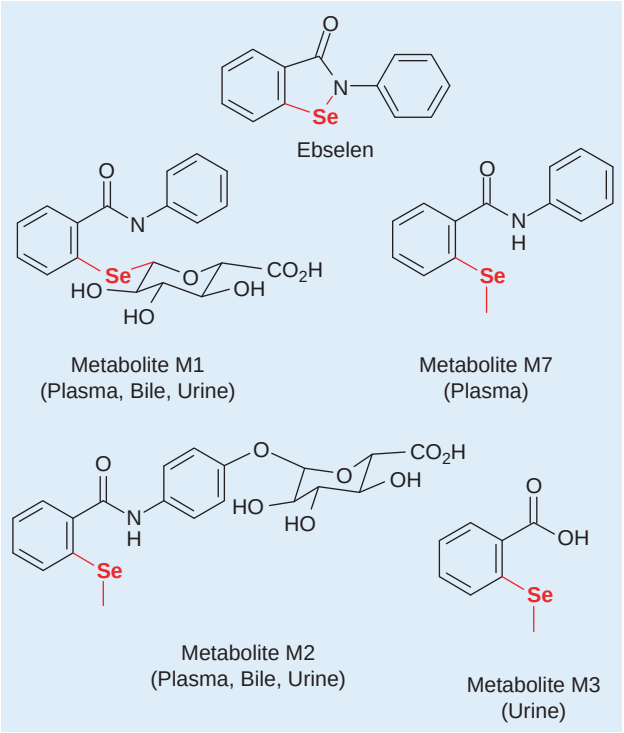


FIGURE 8.71 Ebselen and its main metabolites [256].

Selenium and its derivatives are highly toxic and, with the exception of ^{75}Se derivatives that serve diagnostic purposes (e.g., ^{75}*Se-selenomethionine*, used as a radioactive imaging agent in pancreatic scanning), there is no chemically defined seleno-organic drug on the market. Klayman reviewed a large number of selenium derivatives as chemotherapeutic agents in 1973 [254]. Selenium bioisosteres of sulfur compounds are mainly used as research tools (e.g., bis [2-chloroethyl] selenide as selenium bioisostere of the classical sulfur mustards [255]). Selenocysteine is present in the catalytic site of mammalian glutathione-peroxidase, and this explains the importance of selenium as an essential trace element.

The only selenium-containing drug candidate is *ebselen* (Figure 8.71) which owes its antioxidant and anti-inflammatory properties to its interference with the selenoenzyme glutathione-peroxidase [257]. Due to its strongly bound selenium moiety, only metabolites of low toxicity are formed [256].

References

[1] Langmuir I. The structure of atoms and the octet theory of valence. Proc Natl Acad Sci USA 1919;5(7):252–9.
[2] Friedman HL. Influence of isosteric replacements upon biological activity. NASNRS 1951;206:295–358.
[3] Allen HS. Molecular frequency and molecular number. Parts I–III. London: n.p.; 1918.
[4] Langmuir I. Isomorphism, isosterism, and covalence. J Amer Chem Soc 1919;41:1543–59.
[5] Patani GA, LaVoie EJ. Bioisosterism: a rational approach in drug design. Chem Rev 1996;96(8):3147–76.
[6] Seifriz W. Pathogenicity and isosterism. Science 1948;107:15–16.
[7] Grimm HG. Structure and size of the non-metallic hybrids. Z Electrochem 1925;31:474–80.
[8] Grimm HG. On the systematic arrangement of chemical compounds from the perspective of research on atomic composition; and on some challenges in experimental chemistry. Naturwissenschaften 1928;17:557–64.
[9] Erlenmeyer H, Leo M. On pseudoatoms. Helv Chim Acta 1932;15:1171–86.
[10] Gelmboldt VO, Ennan AA, Ganin EV, Simonov YA, Fonari MS, Botoshansky MM. Synthesis and structure of fluorosilicic acid compounds with 4-aminobenzoic acid and with 4-aminobenzenesulfonamide: the role of H-bonding in crystal structure formation. J Fluorine Chem 2004;125:1951–7.
[11] McLeod JW, Mayr-Harting A, Walker N. Observations on the bactericidal and bacteriostatic actions of p-aminobenzenesulfonamide and p-hydroxylaminobenzenesulfonamide, with special reference to their suppression by p-aminobenzoic acid. Br J Exp Pathol 1944;25:27–37.
[12] Brauner-Osborne H, Krogsgaard-Larsen P. Pharmacology of (S)-homoquisqualic acid and (S)-2-amino-5-phosphonopentanoic acid [(S)-AP5] at cloned metabotropic glutamate receptors. Br J Pharmacol 1998;123(2):269–74.

[13] Malachowski WP, Coward JK. The chemistry of phosphapeptides: investigations on the synthesis of phosphonamide, phosphonate, and phsophinateanalogues of glutamyl-γ-glutamate. J Org Chem 1994;59(25):7625–34.
[14] Thornber CW. Isosterism and molecular modification in drug design. Chem Soc Rev 1957;8:563–80.
[15] Burger A, editor. Medicinal chemistry. 3rd ed. New York: Wiley-Interscience ed; 1970.
[16] Lima LM, Barreiro EJ. Bioisosterism: a useful strategy for molecular modification and drug design. Curr Med Chem 2005;12(1):23–49.
[17] Janssen PAJ, Van der Eycken CAM. The chemical anatomy of potent morphine-like analgesics (p. 42). In: Burger A, editor. Drugs affecting the central nervous system. New York: Marcel Dekker, Inc.; 1968. p. 25–60.
[18] Boyle EA, Mangan FR, Markwell RE, Smith SA, Thomson MJ, Ward RW, et al. 7-Aroyl-2,3-dihydrobenzo[b]furan-3-carboxylic acids and 7-benzoyl-2,3-dihydrobenzo[b]thiophene-3-carboxylic acids as analgesic agents. J Med Chem 1986;29:894–8.
[19] Morgan BP, Scholtz JM, Ballinger MD, Zipkin ID, Bartlett PA. Differential binding energy: a detailed evaluation of the influence of hydrogen bonding and hydrophobic groups on the inhibition of thermolysin by phosphorous-containing inhibitors. J Am Chem Soc 1991;113:297–307.
[20] Chen YL, Nielsen J, Hedberg KD, Dunaiskis A, Jones S, Russo L, et al. Syntheses, resolution, and structure-activity relationships of potent acetylcholinesterase inhibitors: 898-carbaphysostigmine analogues. J Med Chem 1992;35:1429–34.
[21] Marseigne I, Roques BP. Synthesis of new amino-acids mimicking sulfated and phosphorylated tyrosine residues. J Org Chem 1988;53:3621–4.
[22] Burke Jr TR, Smyth M, Nomizu M, Otaka A, Roller PP. Preparation of fluoro- and hydroxy-4-phosphonomethyl-d,l-phenylalanine suitably protected for solid phase synthesis of peptides containing hydrolytically stable analogues of O-phosphotyrosine. J Org Chem 1993;58:1336–40.
[23] Erlenmeyer H, Willi E. Zusammenhänge zwischen Konstitution und Wirkung bei Pyrazolonderivaten [Relationships between structure and effect in pyrazolone derivatives]. Helv Chim Acta 1935;18:740–3 [in German].
[24] Wallace EM, Moliterni JA, Moskal MA, Neubert AD, Marcopoulos N, Stamford LB, et al. Design and synthesis of potent selective inhibitors of endothelin-converting enzyme. J Med Chem 1998;41:1513–23.
[25] De Lombaert S, Stamford LB, Blanchard L, Tan J, Hoyer D, Diefenbacher CG, et al. Potent non-peptidic dual inhibitors of endothelin-converting enzyme and neutral endopeptidase 24.11. Bioorg Med Chem Lett 1997;7:1059–64.
[26] Binder D, Noe CR, Holzer W, Baumann K. Thiophen als Strukturelement physiologisch aktiver Substanzen, 16. Mitt. Thienoisoxazole durch Substitution am Oximstickstoff [Thiophene as a structural element of physiologically active compounds, XVI: thienoisoxazoles by substitution at the oxime nitrogen]. Arch Pharm 1987;320:837–43 [in German].
[27] Uno H, Kurokawa M, Masuda Y, Nishimura H. Studies on 3-substituted 1,2-benzisoxazole derivatives. 6. Syntheses of 3-(sulfamoylmethyl)-1,2-benzisoxazole derivatives and their anticonvulsant activities. J Med Chem 1979;22:180–3.
[28] Garvey DS, Wasicak JT, Decker MW, Brioni JD, Buckley MJ, Sullivan JP, et al. Novel isoxazoles which interact with brain cholinergic channel receptors have intrinsic cognitive enhancing and anxiolytic activities. J Med Chem 1994;37(8):1055–9.
[29] Garvey DS, Wasicak JT, Elliott RL, Lebold SA, Hettinger AM, Carrera GM, et al. Ligands for brain cholinergic channel receptors: synthesis and in vitro characterization of novel isoxazoles and isothiazoles as bioisosteric replacements for the pyridine ring in nicotine. J Med Chem 1994;37(26):4455–63.
[30] Olesen PH, Tonder JE, Hansen JB, Hansen HC, Rimvall K. Bioisosteric replacement strategy for the synthesis of 1-azacyclic compounds with high affinity for the central nicotinic cholinergic receptors. Bioorg Med Chem 2000;8(6):1443–50.
[31] Gohlke H, Gundisch D, Schwarz S, Seitz G, Tilotta MC, Wegge T. Synthesis and nicotinic binding studies on enantiopure diazine analogues of the novel (2-chloro-5-pyridyl)-9-azabicyclo[4.2.1]non-2-ene UB-165. J Med Chem 2002;45(5):1064–72.
[32] Heinisch G, Holzer W, Kunz F, Langer T, Lukavsky P, Pechlaner C, et al. On the bioisosteric potential of diazines: diazine analogues of the combined thromboxane A2 receptor antagonist and synthetase inhibitor ridogrel. J Med Chem 1996;39(20):4058–64.
[33] Haginoya N, Kobayashi S, Komoriya S, Yoshino T, Suzuki M, Shimada T, et al. Synthesis and conformational analysis of a non-amidine factor Xa inhibitor that incorporates 5-methyl-4,5,6,7-tetrahydrothiazolo[5,4-c]pyridine as S4 binding element. J Med Chem 2004;47(21):5167–82.
[34] Wood MR, Schirripa KM, Kim JJ, Wan BL, Murphy KL, Ransom RW, et al. Cyclopropylamino acid amide as a pharmacophoric replacement for 2,3-diaminopyridine. Application to the design of novel bradykinin B1 receptor antagonists. J Med Chem 2006;49(4):1231–4.
[35] Abe Y, Kayakiri H, Satoh S, Inoue T, Sawada Y, Inamura N, et al. A novel class of orally active non-peptide bradykinin B2 receptor antagonists. 3. Discovering bioisosteres of the imidazo[1,2-a] pyridine moiety. J Med Chem 1998;41(21):4062–79.
[36] Contreras JM, Parrot I, Sippl W, Rival YM, Wermuth CG. Design, synthesis, and structure-activity relationships of a series of 3-[2-(1-benzylpiperidin-4-yl)ethylamino]pyridazine derivatives as acetylcholinesterase inhibitors. J Med Chem 2001;44(17):2707–18.
[37] Contreras JM, Rival YM, Chayer S, Bourguignon JJ, Wermuth CG. Aminopyridazines as acetylcholinesterase inhibitors. J Med Chem 1999;42(4):730–41.
[38] Wermuth CG. The practice of medicinal chemistry. 2nd ed. London: Academic Press; 2003. p. 768.
[39] Biava M, Porretta GC, Cappelli A, Vomero S, Manetti F, Botta M, et al. 1,5-Diarylpyrrole-3-acetic acids and esters as novel classes of potent and highly selective cyclooxygenase-2 inhibitors. J Med Chem 2005;48(9):3428–32.
[40] Barbachyn MR, Ford CW. Oxazolidinone structure-activity relationships leading to linezolid. Angew Chem Int Ed Engl 2003;42(18):2010–23.
[41] Renslo AR, Luehr GW, Gordeev MF. Recent developments in the identification of novel oxazolidinone antibacterial agents. Bioorg Med Chem 2006;14(12):4227–40.
[42] Snyder LB, Meng Z, Mate R, D'Andrea SV, Marinier A, Quesnelle CA, et al. Discovery of isoxazolinone antibacterial agents. Nitrogen as a replacement for the stereogenic center found in oxazolidinone antibacterials. Bioorg Med Chem Lett 2004;14(18):4735–9.
[43] Nilius AM. Have the oxazolidinones lived up to their billing? Future perspectives for this antibacterial class. Curr Opin Investig Drugs 2003;4(2):149–55.
[44] Ganellin CR. Discovery of cimetidine, ranitidine and other H_2-receptor histamine antagonists. In: Ganellin CR, Roberts SM, editors. Medicinal chemistry: The role of organic chemistry in drug research. London: Academic Press; 1993. p. 227–55.

[45] Mallamo JP, Pilling GM, Wetzel JR, Kowalczik PJ, Bell MR, Kullnig RK, et al. Antiandrogenic steroidal sulfonyl heterocycles. Utility of electrostatic complementarity in defining bioisosteric sulfonyl heterocycles. J Med Chem 1992;35:1663–70.
[46] Fludzinski P, Evrard DA, Bloomquist WE, Lacefield WB. Indazoles as indole bioisosteres: synthesis and evaluation of the tropanyl ester and amide of indazole-3-carboxylate as antagonists to the serotonin $5HT_3$ receptor. J Med Chem 1987;30:1535–7.
[47] Mathes BM, Hudziak KJ, Schaus JM, Xu YC, Nelson DL, Wainscott DB, et al. Substituted furo[3,2-b]pyridines: novel bioisosteres of 5-HT 1F receptor agonists. Bioorg Med Chem Lett 2004;14(1):167–70.
[48] Blair JB, Marona-Lewicka D, Kanthasamy A, Lucaites VL, Nelson DL, Nichols DE. Thieno[3,2-b]- and thieno[2,3-b]pyrrole bioisosteric analogues of the hallucinogen and serotonin agonist N, N-dimethyltryptamine. J Med Chem 1999;42(6):1106–11.
[49] Yous S, Andrieux J, Howell HE, Morgan PJ, Renard P, Pfeiffer B, et al. Novel naphthalenic ligands with high affinity for the melatonin receptor. J Med Chem 1992;35(8):1484–6.
[50] Depreux P, Lesieur D, Mansour HA, Morgan P, Howell HE, Renard P, et al. Synthesis and structure-activity relationships of novel naphthalenic and bioisosteric related amidic derivatives as melatonin receptor ligands. J Med Chem 1994;37(20):3231–9.
[51] Blaskó G, Major E, Blaskó G, Rózsa I, Szántay C. Pyrimido[1,6-a]pyrido[3,4-b] indoles as new platelet inhibiting agents. Eur J Med Chem 1986;21:91–5.
[52] Kardos J, Blaskó G, Simonyi M, Szántay C. Octahydroindolo[2,3-a]quinolizin-2-one, a novel structure for γ-aminobutyric acid (GABA) uptake inhibition. Eur J Med Chem 1985;21:151–4.
[53] Salituro FG, Harrison BL, Baron BM, Nyce PL, Stewart KT, Kehne JH, et al. 3-(2-Carboxyindol-3-yl)propionic acid-based antagonists of the N-methyl-d-aspartic receptor associated glycine binding site. J Med Chem 1992;35:1791–9.
[54] Calvino R, Stilo A, Di Fruttero R, Gasco AM, Sorba G, Gasco A. Pharmacochemistry of the furoxan ring: recent developments. Il Farmaco 1993;48:321–34.
[55] Lipinski CA, Aldinger CE, Beyer TA, Bordner J, Burdi DF, Bussolotti DL, et al. Hydantoin isosteres: in vivo active spiro hydroxy acetic aldose reductase inhibitors. J Med Chem 1992;35:2169–77.
[56] Lober S, Hubner H, Gmeiner P. Fused azaindole derivatives: molecular design, synthesis and in vitro pharmacology leading to the preferential dopamine D3 receptor agonist FAUC 725. Bioorg Med Chem Lett 2002;12(17):2377–80.
[57] Krogsgaard-Larsen P, Hjeds H, Falch E, Jørgensen FS, Nielsen L. Recent advances in GABA agonists, antagonists and uptake inhibitors: structure-activity relationships and therapeutic potential. In: Testa B, editor. Advances in drug research, vol. 17. London: Academic Press; 1988. p. 381–456.
[58] Almquist RG, Chao WR, Jennings-White C. Synthesis and biological activity of carboxylic acid replacement analogues of the potent angiotensin converting enzyme inhibitor 5(S)-benzamido-4-oxo-6-phenylhexanoyl-l-proline. J Med Chem 1985;28:1067–71.
[59] Massa S, Mai A, Sbardella G, Esposito M, Ragno R, Loidl P, et al. 3-(4-Aroyl-1H-pyrrol-2-yl)-N-hydroxy-2-propenamides, a new class of synthetic histone deacetylase inhibitors. J Med Chem 2001;44(13):2069–72.
[60] Lu Q, Yang YT, Chen CS, Davis M, Byrd JC, Etherton MR, et al. Zn21-chelating motif-tethered short-chain fatty acids as a novel class of histone deacetylase inhibitors. J Med Chem 2004;47(2):467–74.
[61] Remiszewski SW, Sambucetti LC, Bair KW, Bontempo J, Cesarz D, Chandramouli N, et al. N-hydroxy-3-phenyl-2-propenamides as novel inhibitors of human histone deacetylase with in vivo antitumor activity: discovery of (2E)-N-hydroxy-3-[4-[[(2-hydroxyethyl)[2-(1H-indol-3-yl)ethyl]amino]methyl]phenyl]-2-propenamide (NVP-LAQ824). J Med Chem 2003;46(21):4609–24.
[62] Plumb JA, Finn PW, Williams RJ, Bandara MJ, Romero MR, Watkins CJ, et al. Pharmacodynamic response and inhibition of growth of human tumor xenografts by the novel histone deacetylase inhibitor PXD101. Mol Cancer Ther 2003;2(8):721–8.
[63] Kelly WK, O'Connor OA, Krug LM, Chiao JH, Heaney M, Curley T, et al. Phase I study of an oral histone deacetylase inhibitor, suberoylanilide hydroxamic acid, in patients with advanced cancer. J Clin Oncol 2005;2(17):3923–31.
[64] Buggy JJ, Cao ZA, Bass KE, Verner E, Balasubramanian S, Liu L, et al. CRA-024781: a novel synthetic inhibitor of histone deacetylase enzymes with antitumor activity in vitro and in vivo. Mol Cancer Ther 2006;5(5):1309–17.
[65] Hanessian S, Moitessier N, Gauchet C, Viau M. N-aryl sulfonyl homocysteine hydroxamate inhibitors of matrix metalloproteinases: further probing of the S(1), S(1)9, and S(2)9 pockets. J Med Chem 2001;44(19):3066–73.
[66] Aranapakam V, Davis JM, Grosu GT, Baker J, Ellingboe J, Zask A, et al. Synthesis and structure-activity relationship of N-substituted 4-arylsulfonylpiperidine-4-hydroxamic acids as novel, orally active matrix metalloproteinase inhibitors for the treatment of osteoarthritis. J Med Chem 2003;46(12):2376–96.
[67] Aranapakam V, Grosu GT, Davis JM, Hu B, Ellingboe J, Baker JL, et al. Synthesis and structure-activity relationship of alpha-sulfonylhydroxamic acids as novel, orally active matrix metalloproteinase inhibitors for the treatment of osteoarthritis. J Med Chem 2003;46(12):2361–75.
[68] Noe MC, Natarajan V, Snow SL, Mitchell PG, Lopresti-Morrow L, Reeves LM, et al. Discovery of 3,3-dimethyl-5-hydroxypipecolic hydroxamate-based inhibitors of aggrecanase and MMP-13. Bioorg Med Chem Lett 2005;15(11):2808–11.
[69] Duan JJ, Chen L, Wasserman ZR, Lu Z, Liu RQ, Covington MB, et al. Discovery of gamma-lactam hydroxamic acids as selective inhibitors of tumor necrosis factor alpha converting enzyme: design, synthesis, and structure-activity relationships. J Med Chem 2002;45(23):4954–7.
[70] Kohler HV, Eichler B, Salewski R. Untersuchungen zum sauerstoffanalogen Charakter der $C(CN)_2$- und NCN- Gruppen [Studies in the oxygen analog character of the $C(CN)_2$- and NCN- groups]. Z Anorg Chem 1970;379:183–92 [in German].
[71] Kwon C-H, Nagasawa HT, DeMaster EG, Shirota FN. Acyl, N-protected α-aminoacyl, and peptidyl derivatives as prodrug forms of the alcohol deterrent agent cyanamide. J Med Chem 1986;29:1922–9.
[72] Drummond JT, Johnson G. Convenient procedure for the preparation of alkyl end aryl substituted N-(aminoalkylacyl)sulfonamides. Tetrahedron Let 1988;29:1653–6.
[73] Albright JD, DeVries VG, Du MD, Largis EE, Miner TG, Reich MF, et al. Potential antiatherosclerotic agents. 3. Substituted benzoic and non benzoic analogues of cetabon. J Med Chem 1983;26:1393–411.
[74] Uehling DE, Donaldson KH, Deaton DN, Hyman CE, Sugg EE, Barrett DG, et al. Synthesis and evaluation of potent and selective beta(3) adrenergic receptor agonists containing acylsulfonamide, sulfonylsulfonamide, and sulfonylurea carboxylic acid isosteres. J Med Chem 2002;45(3):567–83.

[75] Johansson A, Poliakov A, Akerblom E, Wiklund K, Lindeberg G, Winiwarter S, et al. Acyl sulfonamides as potent protease inhibitors of the hepatitis C virus full-length NS3 (protease-helicase/NTPase): a comparative study of different C-terminals. Bioorg Med Chem 2003;11(12):2551–68.
[76] Bovy PR, Reitz DB, Collins JT, Chamberlain TS, Olins GM, Corpus VM, et al. Nonpeptide angiotensin II antagonists: N-phenyl-1-H-pyrrole derivatives are angiotensin II receptor antagonists. J Med Chem 1993;36:101–10.
[77] Marshall WS, Goodson T, Cullinan GJ, Swanson-Bean D, Haisch KD, Rinkema LE, et al. Leukotriene receptor antagonists. 1. Synthesis and structure-activity relationships of alkoxyacetophenone derivatives. J Med Chem 1987;30:682–9.
[78] Krogsgaard-Larsen P. In: Hansch C, Sammes PG, Taylor JB, Emmet JC, editors. Comprehensive medicinal chemistry, vol. 3. Oxford: Pergamon Press; 1990. p. 493–537.
[79] Lunn WHW, Schoepp DD, Lodge D, True RA, Millar JD. Poster N P-041.A. LY262466, DL-2-amino-3-(4-hydroxy-1,2,5-thiazol-3-yl) propanoic acid hydrochloride, a novel and selective agonist at the AMPA excitatory amino acid receptor. 12th International Symposium on Medicinal Chemistry 1992:pp. 113–17.
[80] Atkinson JG, Girard Y, Rokach J, Rooney CS, McFarlane CS, Rackham A, et al. Kojic amine-A novel γ-aminobutyric acid analogue. J Med Chem 1979;22:90–106.
[81] Hulin B, McCarthy PA, Gibbs EM. The glitazone family of antidiabetic agents. Curr Pharm Des 1996;2:85–102.
[82] Henke BR. Peroxisome proliferator-activated receptor alpha/gamma dual agonists for the treatment of type 2 diabetes. J Med Chem 2004;47(17):4118–27.
[83] Gezginci MH, Martin AR, Franzblau SG. Antimycobacterial activity of substituted isosteres of pyridine- and pyrazinecarboxylic acids. 2. J Med Chem 2001;44(10):1560–3.
[84] Nicolaou I, Zika C, Demopoulos VJ. [1-(3,5-Difluoro-4-hydroxyphenyl)-1H-pyrrol-3-yl]phenylmethanone as a bioisostere of a carboxylic acid aldose reductase inhibitor. J Med Chem 2004;47(10):2706–9.
[85] Qiu J, Stevenson SH, O'Beirne MJ, Silverman RB. 2,6-Difluorophenol as a bioisostere of a carboxylic acid: bioisosteric analogues of gamma-aminobutyric acid. J Med Chem 1999;42(2):329–32.
[86] Noe MC, Natarajan V, Snow SL, Wolf-Gouveia LA, Mitchell PG, Lopresti-Morrow L, et al. Discovery of 3-OH-3-methylpipecolic hydroxamates: potent orally active inhibitors of aggrecanase and MMP-13. Bioorg Med Chem Lett 2005;15(14):3385–8.
[87] Buu-Hoï NP, Lambelin G, Lepoivre C, Gillet C, Gautier M, Thiriaux J. Un nouvel agent antiinflammatoire de structure non stéroïdique: l'acide p-butoxyphénylacéthydroxamique [A new anti-inflammatory agent of non-steroidal structure: P-butoxyphenylacethydroxamic acid]. C R Acad Sci (Paris) 1965;261:2259–62.
[88] Orzalesi G., Selleri R. Pharmaceutical 2-(4-isobutylphenyl) propionohydroxamic acid. Ger. Offen. 2,400,531 (24 Jul. 1974, to Societa Italo-Britannica L. Manetti & H.Roberts e C.). 1974;C.A. 1974; 81; 120272i.
[89] De Martiis F, Franzone JS, Tamietto T. Sintesi e proprieta antiflogistiche di alcuni acidici indolil-acetoidrossammici [Synthesis and anti-inflammatory properties of various indolyl-acetohydroxamic acids]. Bol Chim Farm 1975;114:309–18 [in Italian].
[90] Orzalesi G, Mari F, Bertol E, Selleri R, Pisaturo G. Anti-inflammatory agents: determination of ibuproxam and its metabolite in humans. Arzneim-Forsch 1980;30:1607–9.
[91] Demay F, De Sy J. A new non-steroidal anti-inflammatory drug (NSAID) in current rheumatologic practice (oxamethacin). Curr Ther Res 1982;31:113–18.
[92] Vergin H, Ferber H, Brunner F, Kukovetz WR. Pharmakokinetik und Biotransformation von Oxametacin bei gesunden Probanden [Pharmacokinetics and biotransformation of oxametacin in healthy volunteers]. Arzneim Forsch 1981;31:513–18 [in French].
[93] Allegretti M, Bertini R, Cesta MC, Bizzarri C, Di Bitondo R, Di Cioccio V, et al. 2-Arylpropionic CXC chemokine receptor 1 (CXCR1) ligands as novel noncompetitive CXCL8 inhibitors. J Med Chem 2005;48(13):4312–31.
[94] Wan Y, Wallinder C, Johansson B, Holm M, Mahalingam AK, Wu X, et al. First reported nonpeptide AT1 receptor agonist (L-162,313) acts as an AT2 receptor agonist in vivo. J Med Chem 2004;47(6):1536–46.
[95] Lobb KL, Hipskind PA, Aikins JA, Alvarez E, Cheung YY, Considine EL, et al. Acyl sulfonamide anti-proliferatives: benzene substituent structure-activity relationships for a novel class of antitumor agents. J Med Chem 2004;47(22):5367–80.
[96] Mader MM, Shih C, Considine E, Dios AD, Grossman CS, Hipskind PA, et al. Acyl sulfonamide anti-proliferatives. Part 2: activity of heterocyclic sulfonamide derivatives. Bioorg Med Chem Lett 2005;15(3):617–20.
[97] Hattori K, Tanaka A, Fujii N, Takasugi H, Tenda Y, Tomita M, et al. Discovery of diphenyloxazole and Ndelta-Z-ornithine derivatives as highly potent and selective human prostaglandin EP(4) receptor antagonists. J Med Chem 2005;48(9):3103–6.
[98] Liljebris C, Larsen SD, Ogg D, Palazuk BJ, Bleasdale JE. Investigation of potential bioisosteric replacements for the carboxyl groups of peptidomimetic inhibitors of protein tyrosine phosphatase 1B: identification of a tetrazole-containing inhibitor with cellular activity. J Med Chem 2002;45(9):1785–98.
[99] Momose Y, Maekawa T, Odaka H, Ikeda H, Sohda T. Novel 5-substituted-1H-tetrazole derivatives as potent glucose and lipid lowering agents. Chem Pharm Bull (Tokyo) 2002;50(1):100–11.
[100] Biot C, Bauer H, Schirmer RH, Davioud-Charvet E. 5-Substituted tetrazoles as bioisosteres of carboxylic acids: bioisosterism and mechanistic studies on glutathione reductase inhibitors as antimalarials. J Med Chem 2004;47(24):5972–83.
[101] Kimura T, Shuto D, Hamada Y, Igawa N, Kasai S, Liu P, et al. Design and synthesis of highly active Alzheimer's beta-secretase (BACE1) inhibitors, KMI-420 and KMI-429, with enhanced chemical stability. Bioorg Med Chem Lett 2005;15(1):211–15.
[102] Roppe J, Smith ND, Huang D, Tehrani L, Wang B, Anderson J, et al. Discovery of novel heteroarylazoles that are metabotropic glutamate subtype 5 receptor antagonists with anxiolytic activity. J Med Chem 2004;47(19):4645–8.
[103] Kozikowski AP, Zhang J, Nan F, Petukhov PA, Grajkowska E, Wroblewski JT, et al. Synthesis of urea-based inhibitors as active site probes of glutamate carboxypeptidase II: efficacy as analgesic agents. J Med Chem 2004;47(7):1729–38.
[104] Valgeirsson J, Nielsen EO, Peters D, Mathiesen C, Kristensen AS, Madsen U. Bioisosteric modifications of 2-arylureidobenzoic acids: selective noncompetitive antagonists for the homomeric kainate receptor subtype GluR5. J Med Chem 2004;47(27):6948–57.
[105] Kraus JL. Isosterism and molecular modification in drug design: tetrazole analogue of GABA: effects on enzymes of the gamma-aminobutyrate system. Pharmacol Res Commun 1983;15:183–9.

[106] Hallinan EA, Tsymbalov S, Dorn CR, Pitzele BS, Hansen Jr DW, Moore WM, et al. Synthesis and biological characterization of L-N(6)-(1-iminoethyl)lysine 5-tetrazole-amide, a prodrug of a selective iNOS inhibitor. J Med Chem 2002;45(8):1686–9.
[107] Schlewer G, Wermuth CG, Chambon J-P. Analogues tétrazoliques d'agents GABA-mimétiques [Tetrazole analogues of GABA-mimetics]. Eur J Med Chem 1984;19:181–6 [in French].
[108] Krogsgaard-Larsen P, Ferkany JW, Nielsen EO, Madsen U, Ebert B, Johansen JS, et al. Novel class of amino acid antagonists at non-N-methyl-d-aspartic acid excitatory amino acid receptors. Synthesis, in vitro and in vivo pharmacology, and neuroprotection. J Med Chem 1991;34:123–30.
[109] Krogsgaard-Larsen P, Rodolskov-Christiansen T. GABA agonists: synthesis and structure-activity studies on analogues of isoguvacine and THIP. Eur J Med Chem 1979;14:157–64.
[110] Kraus JL. Isosterism and molecular modification in drug design: new n-dipropylacetate analogs as inhibitors of succinic semi aldehyde dehydrogenase. Pharmacol Res Commun 1983;15:119–29.
[111] Lichtenthaler FW, Heidel P. Intermediates in the formation of γ-pyrones from hexose derivatives: a simple synthesis of kojic acid and hydroxymaltol. Angew Chem Int Ed 1969;8:978–9.
[112] Froestl W, Furet P, Hall RG, Mickel SJ, Strub D, Sprecher GV, et al. $GABA_B$ antagonists: novel CNS-active compounds. In: Testa B, Kyburz E, Fuhrer W, Giger R, editors. Perspectives in medicinal chemistry. Weinheim: VHC; 1993. p. 259–72.
[113] Rosowski A, Forsch RA, Freisheim JH, Moran RG, Wick M. Methotrexate analogues. 19. Replacement of the glutamate side chain in classical antifolates by l-homocysteic acid and l-cysteic acid: effect on enzyme inhibition and antitumor activity. J Med Chem 1984;27:600–4.
[114] Watkins JC, Krogsgaard-Larsen P, Honoré T. Structure-activity relationships in the development of excitatory amino acid receptor agonists and competitive antagonists. Trends Pharm Sci 1990;11:25–33.
[115] Drysdale MJ, Pritchard MC, Horwell DC. Rationally designed "dipeptoid" analogues of CCK. Acid mimics of the potent and selective non peptide CCK-B receptor antagonist CI-988. J Med Chem 1992;35:2573–81.
[116] Franz RD. Comparisons of pKa and log P values of some carboxylic and phosphonic acids: synthesis and measurement. AAPS PharmSci 2001;3(2):E10.
[117] Kinney WA, Lee NE, Garrison DT, Podlesny Jr EJ, Simmonds JT, et al. Bioisosteric replacement of the α-amino carboxylic functionality in 2-amino-5-phosphonopentanoic acid yields unique 3,4-diamino-3x-cyclobutene-1,2-dione containing NMDA antagonists. J Med Chem 1992;35:4720–6.
[118] Gao Y, Luo J, Yao Z-J, Guo R, Zou H, Kelley J, et al. Inhibition of Grb2 SH2 domain binding by non-phosphate-containing ligands. 2. 4-(2-Malonyl)phenylalanine as a potent phosphotyrosyl mimetic. J Med Chem 2000;43:911–20.
[119] Shapiro G, Floersheim P, Boelsterli J, Amstutz R, Bolliger G, Gammenthaler H, et al. Muscarinic activity of the thiolactone, lactam, lactol, and thiolactol analogues of pilocarpine and a hypothetical model for the binding of agonists to the m_1 receptor. J Med Chem 1992;35:15–27.
[120] Thompkins L, Lee KH. Comparison of analgesic effects of isosteric variations of salicylic acid and aspirin (acetylsalicylic acid). J Pharm Sci 1975;64:760–3.
[121] Roth GJ, Stanford N, Majerus PW. Acetylation of prostaglandine synthase by aspirin. Proc Nat Acad Sci USA 1975;72:3073–6.
[122] Bundgaard H. Design of prodrugs. Amsterdam: Elsevier; 1985. pp. 55–61.
[123] Sauerberg P, Chen J, WoldeMussie E, Rapoport H. Cyclic carbamate analogues of pilocarpine. J Med Chem 1989;32:1322–6.
[124] Sauerberg P, Kindtler JW, Nielsen L, Sheardown MJ, Honoré T. Muscarinic cholinergic agonists and antagonists of the 3-(3-alkyl-1,2,4-oxadiazol-5-yl)1,2,5,6-tetrahydropyridine type synthesis and structure-activity relationships. J Med Chem 1991;34:687–92.
[125] Sauerberg P, Olesen PH, Nielsen S, Treppendahl S, Sheardown MJ, Honoré T, et al. Novel functional M_1selective muscarinic agonists. Synthesis and structure-activity relationships of 3-(1,2,5-thiadiazolyl)-1,2,5,6-tetrahydro-1-methylpyridines. J Med Chem 1992;35:2263–74.
[126] Sauerberg P, Larsen JJ, Falch E, Krogsgaard-Larsen P. A novel class of conformationally restricted heterocyclic muscarinic agonists. J Med Chem 1986;29(6):1004–9.
[127] Saunders J, Cassidy M, Freedman SB, Harley EA, Iversen LL, Kneen C, et al. Novel quinuclidine-based ligands for the muscarinic cholinergic receptor. J Med Chem 1990;33:1128–38.
[128] Street LJ, Baker R, Book T, Reeve AJ, Saunders J, Willson T, et al. Synthesis and muscarinic activity of quinuclidinyl- and (1-azanorbornyl)pyrazine derivatives. J Med Chem 1992;35:295–305.
[129] Wadsworth HJ, Jenkins SM, Orlek BS, Cassidy F, Clark MSG, Brown F, et al. Synthesis and muscarinic activities of quinuclidin-3-yltriazole and -tetrazole derivatives. J Med Chem 1992;35:1280–90.
[130] Jenkins SM, Wadsworth HJ, Bromidge S, Orlek BS, Wyman PA, Riley GJ, et al. Substituent variation in azabicyclic triazole- and tetrazole-based muscarinic receptor ligands. J Med Chem 1992;35(13):2392–406.
[131] Bromidge SM, Brown F, Cassidy F, Clark MS, Dabbs S, Hadley MS, et al. Design of [R-(Z)]-(1)-alpha-(methoxyimino)-1-azabicyclo[2.2.2] octane-3-acetonitri le (SB 202026), a functionally selective azabicyclic muscarinic M1 agonist incorporating the N-methoxy imidoyl nitrile group as a novel ester bioisostere. J Med Chem 1997;40(26):4265–80.
[132] Toja E, Bonetti C, Butti A, Hunt P, Fortin M, Barzaghi F, et al. 1-Alkyl-1,2,5,6-tetrahydropyridine-3-carboxaldehyde-O-alkyl-oximes: a new class of potent orally available active muscarinic agaonists related to arecoline. Eur J Med Chem 1991;22:853–68.
[133] Carroll FI. 2002 Medicinal chemistry division award address: monoamine transporters and opioid receptors. Targets for addiction therapy. J Med Chem 2003;46(10):1775–94.
[134] Petukhov PA, Zhang M, Johnson KJ, Tella SR, Kozikowski AP. Sar studies of piperidine-based analogues of cocaine. Part 3: oxadiazoles. Bioorg Med Chem Lett 2001;11(16):2079–83.
[135] Street LJ, Baker R, Castro JL, Chambers MS, Guiblin AR, Hobbs SC, et al. Synthesis and serotonergic activity of 5-(oxadiazolyl)tryptamines: potent agonists for 5-HT1D receptors. J Med Chem 1993;36(11):1529–38.
[136] Kuduk SD, Ng C, Feng DM, Wai JM, Chang RS, Harrell CM, et al. 2,3-Diaminopyridine bradykinin B1 receptor antagonists. J Med Chem 2004;47(26):6439–42.

[137] Borzilleri RM, Zheng X, Qian L, Ellis C, Cai ZW, Wautlet BS, et al. Design, synthesis, and evaluation of orally active 4-(2,4-difluoro-5-(methoxycarbamoyl)phenylamino)pyrrolo[2,1-f][1,2,4]triaz ines as dual vascular endothelial growth factor receptor-2 and fibroblast growth factor receptor-1 inhibitors. J Med Chem 2005;48(12):3991–4008.
[138] Sakamoto T, Cullen MD, Hartman TL, Watson KM, Buckheit RW, Pannecouque C, et al. Synthesis and anti-HIV activity of new metabolically stable alkenyldiarylmethane non-nucleoside reverse transcriptase inhibitors incorporating N-methoxy imidoyl halide and 1,2,4-oxadiazole systems. J Med Chem 2007;50(14):3314–21.
[139] Deng BL, Hartman TL, Buckheit Jr RW, Pannecouque C, De Clercq E, Cushman M. Replacement of the metabolically labile methyl esters in the alkenyldiarylmethane series of non-nucleoside reverse transcriptase inhibitors with isoxazolone, isoxazole, oxazolone, or cyano substituents. J Med Chem 2006;49(17):5316–23.
[140] Watson KG, Brown RN, Cameron R, Chalmers DK, Hamilton S, Jin B, et al. An orally bioavailable oxime ether capsid binder with potent activity against human rhinovirus. J Med Chem 2003;46(15):3181–4.
[141] Kozikowski AP, Roberti M, Xiang L, Bergmann JS, Callahan PM, Cunningham KA, et al. Structure-activity relationship studies of cocaine: replacement of the C-2 ester group by vinyl argues against H-bonding and provides an esterase-resistant, high-affinity cocaine analogue. J Med Chem 1992;35:4764–6.
[142] Lipinski CA. Bioisosterism in drug design. In: Bailey DM, editor. Annual report in nedicinal chemistry, vol. 21. San Diego: Academic Press; 1986. p. 283–91.
[143] Wermuth CG. Aminopyridazines: an alternate route to potent muscarinic agonists with no cholinergic syndrome. Il Farmaco 1993;48:253–74.
[144] Huff JR, Anderson PS, Baldwin JJ, Clineschmidt BV, Guare JP, Lotti VJ, et al. N-(1,3,4,6,7,12b-hexahydro-2H-benzo[b]furo[2,3-a]quinolizin-2-yl)-N-methyl-2-hydroxyethane-sulfonamide: a potent and selective α_2-adrenoreceptor antagonist. J Med Chem 1985;28:1756–9.
[145] Spatola AF. Peptide backbone modifications: structure-activity analysis of peptides containing amide bond surrogates. In: Weinstein B, editor. Chemistry and biochemistry of amino acids, peptides, and proteins, vol. 7. New York: Marcel Dekker; 1983. p. 267–357.
[146] Fauchère J-L. Elements for the rational design of peptide drugs. In: Testa B, editor. Advances in drug research, vol. 15. London: Academic Press; 1986. p. 29–69.
[147] Fournier J-P, Moreau RC, Narcisse G, Choay P. Synthèse et propriétés pharmacologiques de sulfonylurées isostères du glibenclamide [Synthesis and pharmacological properties of isosteric sulfonylureas glibenclamide]. Eur J Med Chem 1982;17:81–4 [in French].
[148] Smith III AB, Holcomb RC, Guzman MC, Keenan TP, Sprengeler PA, Hirschmann R. An effective synthesis of scalemic 3,5,5-trisubstituted pyrrolin-4-ones. Tetrahedron Lett 1993;34:63–6.
[149] Smith III AB, Keenan TP, Holcomb RC, Sprengeler PA, Guzman MC, Wood JL, et al. Design, synthesis, and crystal structure of a pyrrolinone-based peptidomimetic possessing the conformation of a β-strand: potential application to the design of novel inhibitors of proteolytic enzymes. J Amer Chem Soc 1992;114:10672–4.
[150] Allmendinger T, Felder E, Hungerbuehler E. Fluoroolefin dipeptide isosteres. In: Weldi JT, editor. Selective fluorination in organic and bioorganic chemistry [Vol ACS Symp 456]. Washington: American Chemical Society; 1991. p. 186–95.
[151] Chandrakumar NS, Yonan PK, Stapelfeld A, Svage M, Rorbacher E, Contreras PC, et al. Preparation and opioid activity of analogues of the analgesic dipeptide 2,6-dimethyl-l-tyrosyl-N-(3-phenylpropyl)-d-alanylamide. J Med Chem 1992;35:223–33.
[152] Naylor EM, Colandrea VJ, Candelore MR, Cascieri MA, Colwell Jr LF, Deng L, et al. 3-Pyridylethanolamines: potent and selective human beta 3 adrenergic receptor agonists. Bioorg Med Chem Lett 1998;8(21):3087–92.
[153] Naylor EM, Parmee ER, Colandrea VJ, Perkins L, Brockunier L, Candelore MR, et al. Human beta3 adrenergic receptor agonists containing imidazolidinone and imidazolone benzenesulfonamides. Bioorg Med Chem Lett 1999;9(5):755–8.
[154] Feng DD, Biftu T, Candelore MR, Cascieri MA, Colwell Jr LF, Deng L, et al. Discovery of an orally bioavailable alkyl oxadiazole beta3 adrenergic receptor agonist. Bioorg Med Chem Lett 2000;10(13):1427–9.
[155] Biftu T, Feng DD, Liang GB, Kuo H, Qian X, Naylor EM, et al. Synthesis and SAR of benzyl and phenoxymethylene oxadiazole benzenesulfonamides as selective beta3 adrenergic receptor agonist antiobesity agents. Bioorg Med Chem Lett 2000;10(13):1431–4.
[156] Parmee ER, Naylor EM, Perkins L, Colandrea VJ, Ok HO, Candelore MR, et al. Human beta3 adrenergic receptor agonists containing cyclic ureidobenzenesulfonamides. Bioorg Med Chem Lett 1999;9(5):749–54.
[157] Elzein E, Ibrahim P, Koltun DO, Rehder K, Shenk KD, Marquart TA, et al. CVT-4325: a potent fatty acid oxidation inhibitor with favorable oral bioavailability. Bioorg Med Chem Lett 2004;14(24):6017–21.
[158] Koltun DO, Marquart TA, Shenk KD, Elzein E, Li Y, Nguyen M, et al. New fatty acid oxidation inhibitors with increased potency lacking adverse metabolic and electrophysiological properties. Bioorg Med Chem Lett 2004;14(2):549–52.
[159] Borg S, Vollinga RC, Labarre M, Payza K, Terenius L, Luthman K. Design, synthesis, and evaluation of Phe-Gly mimetics: heterocyclic building blocks for pseudopeptides. J Med Chem 1999;42(21):4331–42.
[160] Ladduwahetty T, Baker R, Cascieri MA, Chambers MS, Haworth K, Keown LE, et al. N-heteroaryl-2-phenyl-3-(benzyloxy)piperidines: a novel class of potent orally active human NK1 antagonists. J Med Chem 1996;39(15):2907–14.
[161] Einsiedel J, Hubner H, Gmeiner P. Benzamide bioisosteres incorporating dihydroheteroazole substructures: EPC synthesis and SAR leading to a selective dopamine D4 receptor partial agonist (FAUC 179). Bioorg Med Chem Lett 2001;11(18):2533–6.
[162] Einsiedel J, Hubner H, Gmeiner P. Cyclic amidines as benzamide bioisosteres: EPC synthesis and SAR studies leading to the selective dopamine D4 receptor agonist FAUC 312. Bioorg Med Chem Lett 2003;13(5):851–4.
[163] Einsiedel J, Thomas C, Hubner H, Gmeiner P. Phenyloxazoles and phenylthiazoles as benzamide bioisosteres: synthesis and dopamine receptor binding profiles. Bioorg Med Chem Lett 2000;10(17):2041–4.
[164] Thurkauf A, Yuan J, Chen X, He XS, Wasley JW, Hutchison A, et al. 2-Phenyl-4(5)-[[4-(pyrimidin-2-yl)piperazin-1-yl]methyl]imidazole. A highly selective antagonist at cloned human D4 receptors. J Med Chem 1997;40(1):1–3.
[165] Vu CB, Corpuz EG, Merry TJ, Pradeepan SG, Bartlett C, Bohacek RS, et al. Discovery of potent and selective SH2 inhibitors of the tyrosine kinase ZAP-70. J Med Chem 1992;42(20):4088–98.
[166] Clitherow JW, Beswick P, Irving WJ, Scopes DIC, Barnes JC, Clapham J, et al. Novel 1,2,4-oxadiazoles as potent and selective histamine H_3 receptor antagonists. Bioorg Med Chem Lett 1996;6:833–8.

[167] Swain CJ, Baker R, Kneen C, Moseley J, Saunders J, Seward EM, et al. Novel 5-HT3 antagonists. Indole oxadiazoles. J Med Chem 1991;34(1):140–51.
[168] Tully WR, Gardner CR, Gillespie RJ, Westwood R. 2-(Oxadiazolyl)- and 2-(thiazolyl)imidazo[1,2-a]pyrimidines as agonists and inverse agonists at benzodiazepine receptors. J Med Chem 1991;34(7):2060–7.
[169] Orlek BS, Blaney FE, Brown F, Clark MS, Hadley MS, Hatcher J, et al. Comparison of azabicyclic esters and oxadiazoles as ligands for the muscarinic receptor. J Med Chem 1991;34(9):2726–35.
[170] Fletcher SR, McIver E, Lewis S, Burkamp F, Leech C, Mason G, et al. The search for novel TRPV1-antagonists: from carboxamides to benzimidazoles and indazolones. Bioorg Med Chem Lett 2006;16(11):2872–6.
[171] Wu WL, Burnett DA, Caplen MA, Domalski MS, Bennett C, Greenlee WJ, et al. Design and synthesis of orally efficacious benzimidazoles as melanin-concentrating hormone receptor 1 antagonists. Bioorg Med Chem Lett 2006;16(14):3674–8.
[172] Wu WL, Burnett DA, Spring R, Qiang L, Sasikumar TK, Domalski MS, et al. Synthesis and structure-activity relationships of piperidine-based melanin-concentrating hormone receptor 1 antagonists. Bioorg Med Chem Lett 2006;16(14):3668–73.
[173] Lanter JC, Fiordeliso JJ, Allan GF, Musto A, Hahn do W, Sui Z. A bioisosteric approach to the discovery of indole carbinol androgen receptor ligands. Bioorg Med Chem Lett 2006;16(21):5646–9.
[174] Wouters J, Michaux C, Durant F, Dogne JM, Delarge J, Masereel B. Isosterism among analogues of torasemide: conformational, electronic and lipophilic properties. Eur J Med Chem 2000;35(10):923–9.
[175] Luo G, Mattson GK, Bruce MA, Wong H, Murphy BJ, Longhi D, et al. Isosteric N-arylpiperazine replacements in a series of dihydropyridine NPY1 receptor antagonists. Bioorg Med Chem Lett 2004;14(24):5975–8.
[176] Poindexter GS, Bruce MA, Breitenbucher JG, Higgins MA, Sit SY, Romine JL, et al. Dihydropyridine neuropeptide Y Y1 receptor antagonists 2. Bioisosteric urea replacements. Bioorg Med Chem 2004;12(2):507–21.
[177] Lumma Jr WC, Anderson PS, Baldwin JJ, Bolhofer WA, Habecker CN, Hirshfield JM, et al. Inhibitors of gastric acid secretion: 3,4-diamino-1,2,5-thiadiazole 1-oxides and 1,1-dioxides as urea equivalents in a series of histamine H_2-receptor antagonists. J Med Chem 1982;25:207–10.
[178] Young RC, Ganellin CR, Graham MJ, Grant EH. The dipole moments of 1,3-dimethylthiourea, 1,3-dimethyl-2-cyanoguanidine and 1,1-bis-methylamino-2-nitroethene in aqueous solution. Tetrahedron 1982;38:1493–7.
[179] Young RC, Ganellin CR, Graham MJ, Roantree MJ, Grant EH. The dielectric properties of seven polar amidine-containing compounds of biological interest. Tetrahedron Lett 1985;26:1897–900.
[180] Wright JL, Gregory TF, Kesten SR, Boxer PA, Serpa KA, Meltzer LT, et al. Subtype-selective N-methyl-d-aspartate receptor antagonists: synthesis and biological evaluation of 1-(heteroarylalkynyl)-4-benzylpiperidines. J Med Chem 2000;43:3408–19.
[181] Wu WL, Burnett DA, Spring R, Greenlee WJ, Smith M, Favreau L, et al. Dopamine D1/D5 receptor antagonists with improved pharmacokinetics: design, synthesis, and biological evaluation of phenol bioisosteric analogues of benzazepine D1/D5 antagonists. J Med Chem 2005;48(3):680–93.
[182] Burger A. Isosterism and bioisosterism in drug design. Progr Drug Res 1991;37:287–371.
[183] Uloth RH, Kirk JR, Gould WA, Larsen AA. Sulfonanilides. I. Monoalkyl- and arylsulfonamidophenethanolamines. J Med Chem 1966;9(1):88–97.
[184] Hacksell U, Arvidsson LE, Svensson U, Nilsson JL, Sanchez D, Wikstrom H, et al. 3-Phenylpiperidines. Central dopamine-autoreceptor stimulating activity. J Med Chem 1981;24(12):1475–82.
[185] Bhaird NN, Fowler CJ, Thorberg O, Tipton KF. Involvement of catechol-O-methyl transferase in the metabolism of the putative dopamine autoreceptor agonist 3-PPP(3-(3-hydroxyphenyl)-N-n-propylpiperidine). Biochem Pharmacol 1985;34(19):3599–601.
[186] Jaen JC, Wise LD, Caprathe BW, Tecle H, Bergmeier S, Humblet CC, et al. 4-(1,2,5,6-Tetrahydro-1-alkyl-3-pyridinyl)-2-thiazolamines: a novel class of compounds with central dopamine agonist properties. J Med Chem 1990;33(1):311–17.
[187] Mierau J, Schneider FJ, Ensinger HA, Chio CL, Lajiness ME, Huff RM. Pramipexole binding and activation of cloned and expressed dopamine D2, D3 and D4 receptors. Eur J Pharmacol 1995;290(1):29–36.
[188] Schneider CS, Mierau J. Dopamine autoreceptor agonists: resolution and pharmacological activity of 2,6-diaminotetrahydrobenzothiazole and an aminothiazole analogue of apomorphine. J Med Chem 1987;30(3):494–8.
[189] van Vliet LA, Rodenhuis N, Wikstrom H, Pugsley TA, Serpa KA, Meltzer LT, et al. Thiazoloindans and thiazolobenzopyrans: a novel class of orally active central dopamine (partial) agonists. J Med Chem 2000;43(19):3549–57.
[190] Blum G, Gazit A, Levitzki A. Development of new insulin-like growth factor-1 receptor kinase inhibitors using catechol mimics. J Biol Chem 2003;278(42):40442–54.
[191] Hubner H, Haubmann C, Utz W, Gmeiner P. Conjugated enynes as nonaromatic catechol bioisosteres: synthesis binding experiments, and computational studies of novel dopamine receptor agonists recognizing preferentially the D subtype. J Med Chem 2000;43(4):756–62.
[192] Lenz C, Boeckler F, Hubner H, Gmeiner P. Analogues of FAUC 73 revealing new insights into the structural requirements of nonaromatic dopamine D3 receptor agonists. Bioorg Med Chem 2004;12(1):113–17.
[193] Lenz C, Haubmann C, Hubner H, Boeckler F, Gmeiner P. Fancy bioisosteres: synthesis and dopaminergic properties of the endiyne FAUC 88 as a novel non-aromatic D3 agonist. Bioorg Med Chem 2005;13(1):185–91.
[194] Street LJ, Baker R, Davey WB, Guiblin AR, Jelley RA, Reeve AJ, et al. Synthesis and serotonergic activity of N,N-dimethyl-2-[5-(1,2,4-triazol-1-ylmethyl)-1H-indol-3-yl]ethylamine and analogues: potent agonists for 5-HT1D receptors. J Med Chem 1995;38(10):1799–810.
[195] Glen RC, Martin GR, Hill AP, Hyde RM, Woollard PM, Salmon JA, et al. Computer-aided design and synthesis of 5-substituted tryptamines and their pharmacology at the 5-HT1D receptor: discovery of compounds with potential anti-migraine properties. J Med Chem 1995;38(18):3566–80.
[196] Habeeb AG, Praveen Rao PN, Knaus EE. Design and synthesis of celecoxib and rofecoxib analogues as selective cyclooxygenase-2 (COX-2) inhibitors: replacement of sulfonamide and methylsulfonyl pharmacophores by an azido bioisostere. J Med Chem 2001;44(18):3039–42.
[197] Larson AA, Lish PM. A new bio-isostere: alkylsulphonamidophenethanolamines. Nature (London) 1964;203:1283–5.

[198] Roderick SL, Fournié-Zaluski MC, Roques BP, Matthews BW. Thiorphan and retro-thiorphan display equivalent interactions when bound to crystalline thermolysin. Biochemistry 1989;28:1493–7.
[199] Büchi J, Stünzi E, Flury M, Hirt R, Labhart P, Ragaz L. Über lokalanästhetisch wirksame basische ester und amide verschiedener alkoxy-amino-benzoesäuren [Concerning local anesthetic effective basic esters and amides of different alkoxy- amino- benzoic acids]. Helv Chim Acta 1951;34:1002–13 [in German].
[200] Bös M, Jenck F, Martin JR, Moreau J-L, Sleight AJ, Wichmann J, et al. Novel agonists of 5HT2C receptors. Synthesis and biological evaluation of substituted 2-(indol-1-yl)-1-methylethylamines and 2-(indeno[1,2-b]pyrrol-1-yl)-1-methylethylamines. Improved therapeutics for obsessive compulsive disorder. J Med Chem 1997;40:2762–9.
[201] Wermuth CG. Selective optimization of side activities: another way for drug discovery. J Med Chem 2004;47(6):1303–14.
[202] Böhm H-J, Flohr A, Stahl M. Scaffold hopping. Drug Disc Today Technol 2004;1(3):217–24.
[203] Schneider G, Schneider P, Renner S. Scaffold hopping: how far can you jump? QSAR Comb Sci 2006;25(12):1162–71.
[204] Low CM, Buck IM, Cooke T, Cushnir JR, Kalindjian SB, Kotecha A, et al. Scaffold hopping with molecular field points: identification of a cholecystokinin-2 (CCK2) receptor pharmacophore and its use in the design of a prototypical series of pyrrole- and imidazole-based CCK2 antagonists. J Med Chem 2005;48(22):6790–802.
[205] Huang D, Luthi U, Kolb P, Edler K, Cecchini M, Audetat S, et al. Discovery of cell-permeable non-peptide inhibitors of beta-secretase by high-throughput docking and continuum electrostatics calculation. J Med Chem 2005;48(16):5108–11.
[206] Steindl TM, Crump CE, Hayden FG, Langer T. Pharmacophore modeling, docking, and principal component analysis based clustering: combined computer-assisted approaches to identify new inhibitors of the human rhinovirus coat protein. J Med Chem 2005;48 (20):6250–60.
[207] Wilhelm M. The chemistry of polycyclic psycho-active drugs: serendipity or systematic investigation? Pharm J 1975;214:414–16.
[208] Yoshimura H, Kikuchi K, Hibi S, Tagami K, Satoh T, Yamauchi T, et al. Discovery of novel and potent retinoic acid receptor alpha agonists: syntheses and evaluation of benzofuranyl-pyrrole and benzothiophenyl-pyrrole. J Med Chem 2000;43:2929–37.
[209] Chenoweth MB, McCarthy LP. On the mechanism of the pharmacophoric effect of halogenation. Pharmacol Rev 1963;15:673–707.
[210] Goldman P. The carbon–florine bond in compounds of biological interest. Science 1969;164:1123–30.
[211] Peters RA. Biochemical lesions and lethal synthesis. Oxford: Pergamon Press; 1963.
[212] Chapman NB, James JW, Graham JDP, Lewis GP. Chemical reactivity and pharmacological activity among 2-haloethylamine derivatives with a naphtylmethyl group. Chem Ind (London) 1952;805–7.
[213] Vaughan JRJ, Anderson GW, Clapp RC, Clark JH, English JP, Howard KL, et al. Antihistamine agents. IV. Halogenated N,N-dimethyl-N9-benzyl-N-(2-pyridyl)-ethylenediamines. J Org Chem 1949;14:228–34.
[214] Friedman HL. In: Influence of isosteric replacements upon biological activity, Symposium on chemical–biological correlation, Washington, D.C.: National Research Council; 1951. p. 295.
[215] Schatz VB. Isosterism and bioisosterism. In: Burger A, editor. Medicinal chemistry. New York: Interscience Publishers, Inc.; 1963. p. 72–88.
[216] Fessenden RJ, Fessenden JS. The biological properties of silicon compounds. In: Harper NJ, Simmonds AB, editors. Advances in drug research, vol. 4. London: Academic Press; 1967. p. 95–132.
[217] Tacke R, Zilch H. Sila-substitution-a useful strategy for drug design? Endeavour New Series 1986;10:191–7.
[218] Tacke R, Zilch H. Drug-design by sila-substitution and microbial transformations of organosilicon compounds: some recent results. L'Actualité Chimique 1986;3:75–82.
[219] Ricci A, Seconi G, Taddei M. Bioroganosilicon chemistry: trends and perspectives. Chimica Oggi-Chemistry Today 1989;7:15–21.
[220] Showell GA, Mills JS. Chemistry challenges in lead optimization: silicon isosteres in drug discovery. Drug Disc Today 2003;8(12):551–6.
[221] Anonymous. Zifrosilone. Drugs Fut 1994;19:854–5.
[222] Hornsperger J-M, Collard J-N, Heydt JG, Giacobini E, Funes S, Dow J, et al. Trimethylsilylated trifluoromethyl ketones, a novel class of acetylcholinesterase inhibitors: biochemical and pharmacological profile of MDL 73,745. Biochem Soc Trans 1994;22:758–63.
[223] Metcalf RL, Fukuto TR. Silicon-containing carbamate insecticides. J Econ Entomol 1965;58:1151.
[224] Fessenden RJ, Coon MD. Silicon-substituted medicinal agents. Silacarbamates related to meprobamate. J Med Chem 1965;8:604–8.
[225] Tacke R. Sila-Pharmaka, XIX. Sila-Pridinol und Pridinol: darstellung und Eigenschaften sowie Strukturen im kristallinen und gelösten Zustand [Sila-pharmaceutics, XIX. Sila pridinol and pridinol: preparation and properties and structures in the crystalline and dissolved state]. Chem Ber 1980;113:1962–80 [in German].
[226] Moberg WK. Fungicidal 1,2,3-triazole derivatives. US Patent 4,510,136 (Apr 9, 1985 to EI Du Pont de Nemours & Co); 1985.
[227] Damour D, Barreau M, Dutruc-Rosset G, Doble A, Piot O, Mignani S. 1,1-Diphenyl-3-dialkylamino-1-silacyclopentane derivatives: a new class of potent and selective 5-HT_{2A} antagonists. Bioorg Med Chem Lett 1994;4:415–20.
[228] Gotteland J-P, Brunel I, Gendre F, Désiré J, Delhon A, Junquéro D, et al. (Aryloxy)methylsilane derivatives as new cholesterol biosynthesis inhibitors: synthesis and hypocholesterolemic activity of a new class of squalene epoxidase inhibitors. J Med Chem 1995;38:3207–16.
[229] Gotteland J-P, Delhon A, Junquéro D, Oms P, Halazy S. Design and synthesis of new hypocholesteremic organosilanes with antioxidant properties. Bioorg Med Chem Lett 1996;6:533–8.
[230] Heinonen P, Sipilä H, Neuvonen K, Lönnberg H, Cockroft VB, Wurster S, et al. Synthesis and pharmacological properties of 4(5)-(2-ethyl-2,3-dihydro-2-silainden-2-yl)imidazole, a silicon analogue of atipamezole. Eur J Med Chem 1996;31:725–9.
[231] Bom D, Curran DP, Kruszewski S, Zimmer SG, Strode JT, Kohlhagen G, et al. The novel silatecan 7-tert-butyldimethylsilyl-10-hydroxy-camptothecin displays high lipophilicity, improved human blood stability, and potent anticancer activity. J Med Chem 2000;43:3970–80.
[232] Mutahi M, Nittoli T, Guo L, Sieburth SM. Silicon-based metalloprotease inhibitors: synthesis and evaluation of silanol and silanediol peptide analogues as inhibitors of angiotensin-converting enzyme. J Am Chem Soc 2002;124(25):7363–75.
[233] Chen CA, Sieburth SM, Glekas A, Hewitt GW, Trainor GL, Erickson-Viitanen S, et al. Drug design with a new transition state analog of the hydrated carbonyl: silicon-based inhibitors of the HIV protease. Chem Biol 2001;8(12):1161–6.
[234] Showell GA, Barnes MJ, Daiss JO, Mills JS, Montana JG, Tacke R, et al. (R)-sila-venlafaxine: a selective noradrenaline reuptake inhibitor for the treatment of emesis. Bioorg Med Chem Lett 2006;16(9):2555–8.

[235] Barnes MJ, Conroy R, Miller DJ, Mills JS, Montana JG, Pooni PK, et al. Trimethylsilylpyrazoles as novel inhibitors of p38 MAP kinase: a new use of silicon bioisosteres in medicinal chemistry. Bioorg Med Chem Lett 2007;17(2):354–7.
[236] Sommer LH, Bennet OF, Campbell PG, Weyenberg DR. Stereochemistry of hydride ion displacement from silicon. Enhanced rates at bridgehead and 4-ring silicon atoms. J Amer Chem Soc 1957;79:3295–6.
[237] Alam F, Soloway AH, Bapat BV, Barth RF, Adams DM. Boron compounds for neutron capture therapy. Basic Life Sci 1989;50:107–11.
[238] Gabel D. Tumor-seeking for boron neutron capture therapy: synthesis and biodistribution. Basic Life Sci 1989;50:233–41.
[239] Kahl SB, Joel DD, Finkel GC, Micca PL, Nawrocky MM, Coderre JA. A carboranyl porphyrin for boron neutron capture therapy of brain tumors. Basic Life Sci 1989;50:193–203.
[240] Caujolle F, Chan PH. Structure chimique et activité spasmolytique des organoboriques [Chemical structure and spasmolytic activity of organo-boric compounds]. Arch Int Pharmacodyn Ther 1968;172:467–74 [in French].
[241] Mubarak SIM, Stanford JB, Sugden JK. Some aspects of the antimicrobial and chemical properties of phenyl boronate esters of chloramphenicol. Drug Dev Ind Pharm 1984;10:1131–60.
[242] Hutter R, Keller-Schierlein W, Knusel F, Prelog V, Rodgers Jr GC, Suter P, et al. The metabolic products of microorganisms. Boromycin Helv Chim Acta 1967;50(6):1533–9.
[243] Dünitz JD, Hawley DM, Miklos D, White DNJ, Berlin Y, Marusik R, et al. Structure of boromycin. Helv Chim Acta 1971;54:1709–13.
[244] Okami Y, Okazaki T, Kitahara T, Umezawa H. Studies on marine microorganisms. V. A new antibiotic, aplasmomycin, produced by a streptomycete isolated from shallow sea mud. J Antibiot (Tokyo) 1976;29(10):1019–25.
[245] Kinder DH, Katzenellenbogen JA. Acylamino boronic acids and difluoroborane analogues of amino acids: potent inhibitors of chymotrypsine and elastase. J Med Chem 1985;28:1917–25.
[246] Adams J, Behnke M, Chen S, Cruickshank AA, Dick LR, Grenier L, et al. Potent selective inhibitors of the proteasome: dipeptidyl boronic acids. Bioorg Med Chem Lett 1988;8(4):333–8.
[247] Dembitsky VM, Srebnik M. Synthesis and biological activity of α-aminoboronic acids, amine-carboxyboranes, and their derivatives. Tetrahedron 2003;59(5):579–93.
[248] Surolia N, RamachandraRao SP, Surolia A. Paradigm shifts in malaria parasite biochemistry and anti-malarial chemotherapy. Bioessays 2002;24(2):192–6.
[249] Jabbour A, Steinberg D, Dembitsky VM, Moussaieff A, Zaks B, Srebnik M. Synthesis and evaluation of oxazaborolidines for antibacterial activity against streptococcus mutans. J Med Chem 2004;47(10):2409–10.
[250] Kumar SK, Hager E, Pettit C, Gurulingappa H, Davidson NE, Khan SR. Design, synthesis, and evaluation of novel boronic-chalcone derivatives as antitumor agents. J Med Chem 2003;46(14):2813–15.
[251] Landauer W. On the chemical production of developmental abnormalities and of phenocopies in chicken embryos. J Cell Comp Physiol 1954;43(1):261–305.
[252] Landauer W, Clark EM. On the role of riboflavin in the teratogenic activity of boric acid. J Expt Zool 1964;156:307–12.
[253] Browning E. Toxicity of industrial metals. 2nd ed. New York: Appleton-Century-Crofts; 1969. p. 90–97.
[254] Klayman DL, Günther WHH. Organic selenium compounds: their chemistry and biology. New York: Wiley-Interscience; 1973.
[255] Kang S-I, Spears CP. Linear free energy relationships and cytotoxicities of para-substituted 2-haloethyl aryl selenides and bis(-chloroethyl) selenides. J Med Chem 1987;30:597–602.
[256] Fischer H, Terlinden R, Löhr JP, Römer A. A novel biologically active selenoorganic compound. VIII. Biotransformation of ebselen. Xenobiotica 1988;18:1347–59.
[257] Parnham MJ, Graf E. Seleno-organic compounds and the therapy of hydroperoxide-linked pathological conditions. Biochem Pharmacol 1987;36:3095–102.

CHAPTER

9

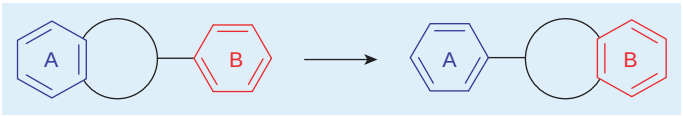
Ring Transformations

Christophe Morice[1] *and Camille G. Wermuth*[2]

[1]Prestwick Chemical Inc., Boulevard Gonthier, Illkirch, France; [2]Founder of Prestwick Chemical, Strasbourg, France

OUTLINE

Rien ne se perd, rien ne se crée, tout se transforme. ***Antoine Lavoisier (1743–1794)***



I. INTRODUCTION

When active molecules contain cyclic systems, these can be opened, expanded, contracted, and modified in many other ways, or even abolished. Conversely, noncyclic molecules can be cyclized, attached to, or included in, ring systems.

In the daily practice of medicinal chemistry, three kinds of approaches are currently used. The first approach does not affect the global complexity of the cyclic system and yields generally close analogs (or "me-too" compounds) of the original active principle. For this, we propose the term "analogical approach." It consists of ring-chain transformations, ring contractions or expansions, and various other ring transformations. The second strategy, called the "disjunctive approach,"[1] aims at the progressive simplification of the original active principle (which is often a natural compound). The objective is to extract information about the minimal structure that is required for activity. Finally, the "conjunctive approach"[1] is based on the creation or addition of

supplementary rings. The objective is to constrain an originally flexible compound and to impose precise conformations and configurations. The preparation of such molecules is of prime importance in the exploration of ligand–receptor interaction and for molecular modeling studies.

II. ANALOGICAL APPROACHES

A. Analogy by Ring Opening: Open-chain Analogs

Open ring analogs of cyclic active principles (open drugs, open-chain analogs) can be designed and synthesized. The usefulness of such compounds is rather questionable, and it appears that most of them were prepared for me-too purpose. Actually, two possibilities can be foreseen:

1. The open analog is again cyclized after oxidation or dehydration by a metabolic enzyme. We deal here with potential rings, which represent nothing more than metabolic precursors of the active species.
2. The open analog does not cyclize *in vivo* but can present some conformational analogy with the ring-containing active principle. These kinds of analogs are known as pseudocycles.

1. Potential rings: In vivo Return to the Cyclic Derivative

Compounds generating the active form after *in vivo* cyclization are in fact prodrugs and will be discussed in a more detailed manner in Chapter 38.

A. PROGUANIL

A historical example is the antimalarial drug proguanyl [2] (Figure 9.1). It was observed that this compound is inactive in *in vitro* cultures of *Plasmodium gallinaceum* but that the serum of animals treated with proguanyl is active in these cultures.

It was concluded that the actual active principle was a metabolite [4] that was subsequently identified as cycloguanyl [3]. In tropical medicine, proguanyl is preferred to cycloguanyl, the latter compound being too rapidly eliminated by the kidneys.

B. POTASSIUM CANRENOATE

This is a water-soluble prodrug that can be administered parenterally (Figure 9.2). It has no intrinsic activity, but it can exert its diuretic activity (as an aldosterone antagonist) because of its interconversion with canrenone. Canrenone is itself the major metabolite of spironolactone [5].

2. Irreversible Open Compounds: Pseudocycles

The open analog is assumed to present a similar conformation to that of the cyclic one. Before adopting the term pseudocycles, it must be ascertained by nuclear magnetic resonance (NMR) or X-ray crystallography that they really mimic the ring-closed analog.

A. DIETHYLSTILBESTROL

The theory of pseudocycles was elaborated to account for the estrogenic activity of compounds such as bisdehydrodoisynolic acid, allenestrol, and diethylstilbestrol (Figure 9.3).

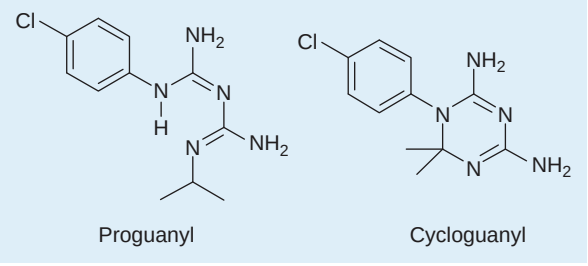


FIGURE 9.1 Cycloguanyl is the active metabolite of proguanyl [3].

The similarity with the natural hormone estradiol is striking. Nevertheless, it is highly probable that for receptor binding, the general shape of the molecules and the distances separating the functional groups are more important than their degree of cyclization [6].

B. NONPHENOLIC ESTRADIOL ANALOGS

Aromatic open ring analogs are much less described in the literature. The difficulty of performing a pseudocycle that mimics the aromaticity and the planarity of the original ring limits its use. For example, salicylaldoxime **1** [7] and anthranylaldoxime **2** [8] generate a pseudocycle that reproduces the naphthol structure **3** [7] (Figure 9.4). The hydrogen bond between the oxime group and the OH (in **1**) or the NH (in **2**) groups are essential to create a stable pseudo-ring that ensures a high-affinity interaction with the estrogen receptors via the hydroxyl group. These pseudocycles represent bioisosteres of the phenolic A group of estradiol.

C. CLONIDINE

Open analogs of the centrally acting hypotensive agent clonidine (Figure 9.5) have a similar activity profile but with a 30 to 100-fold loss in potency [9]. Surprisingly, seco-clonidine, which is the closest analog of clonidine, was found to be less active than the corresponding monomethyl derivative.

D. CROMAKALIM

A more recent example is the open-chain analog of cromakalim that was prepared as a more flexible pyrrolidone replacement (Figure 9.6). It retains about a third of the potency of cromakalim [10].

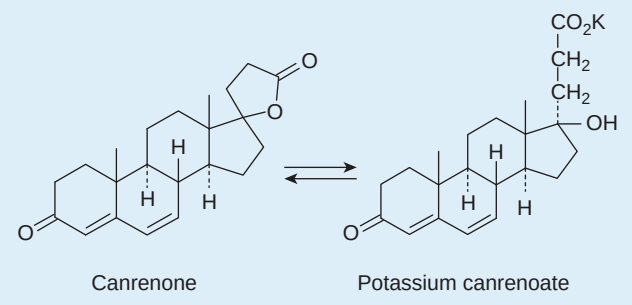


FIGURE 9.2 *In vivo* the inactive potassium canrenoate cyclizes to canrenone [5].

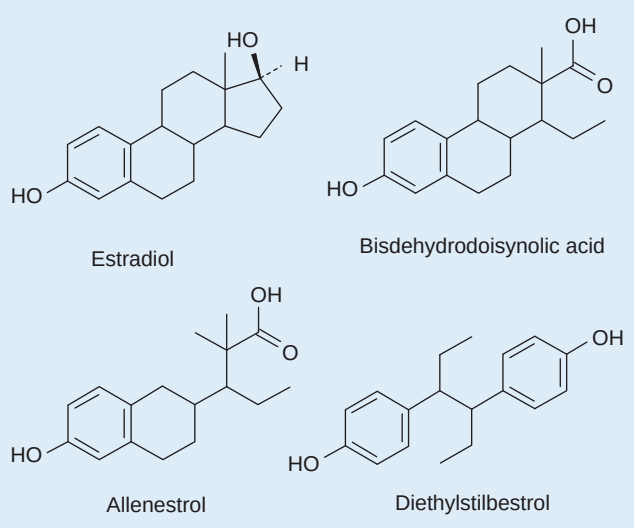


FIGURE 9.3 Open analogs of estradiol.

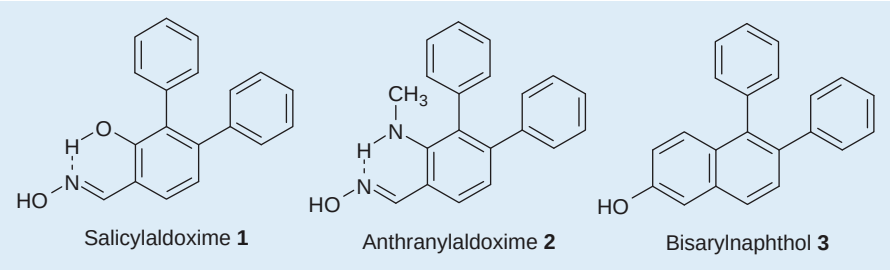


FIGURE 9.4 Aldoximes: as bioisosteres of the phenol ring.

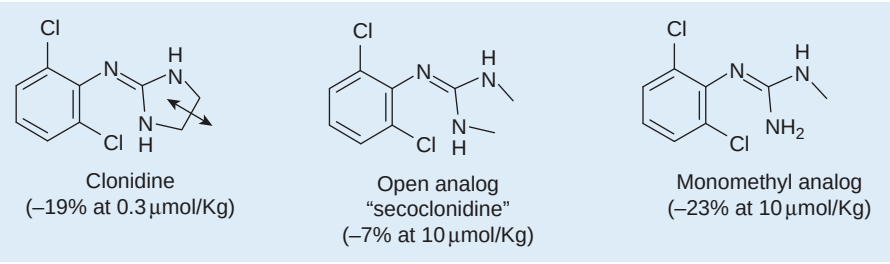


FIGURE 9.5 Open analogs of clonidine (the hypotensive activity is expressed as variation of the arterial blood pressure 30 min after i.v. injection in pentobarbital anesthetized rats) [9].

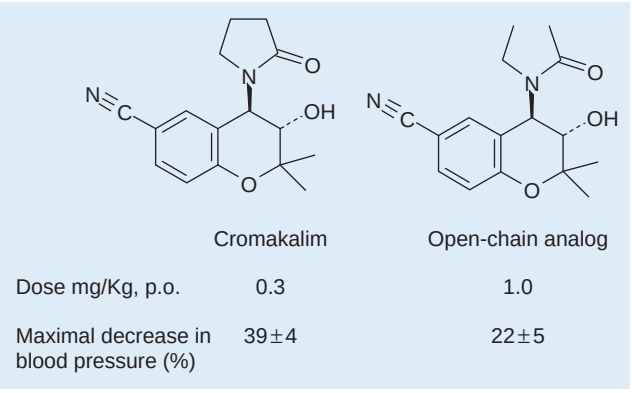


FIGURE 9.6 (3*S*,4*R*) cromakalim and its open-chain analog [10].

B. Analogy by Ring Closure

Cyclizing open structures or creating an additional ring system in a given structure represents one of the useful methods in the search for biologically active conformations. The end result is a more constrained molecule with an imposed conformation. This strategy is related to the conjunctive approach developed further on in this chapter.

The inconvenience is that additional isomeric centers may be introduced and that the selected cyclization mode might not lead to the active conformation adopted by the open-chain drug. A particularly convincing example is given by the ring-closed analog of the thrombin inhibitor NAPAP, which is 100 times more potent than the corresponding open-chain drug [11] (Figure 9.7).

1. *Inhibition of Gastric H^+/K^+ ATPase by Substituted Imidazo[1,2-a]Pyridines*

Apparently, the substitution in the *para* position of the pyridine ring is detrimental (R = H → R = Me; Figure 9.8). However, the ring closure achieving a conformational restriction yields a highly potent compound [13].

2. *Mevinolin and Compactin*

An example of reversible ring closure is found with mevinolin and compactin that are both potent inhibitors of hydroxy-methyl-glutaryl-coenzyme A reductase (HMG-CoA reductase), the rate-determining enzyme in the *de novo* biosynthesis of cholesterol.

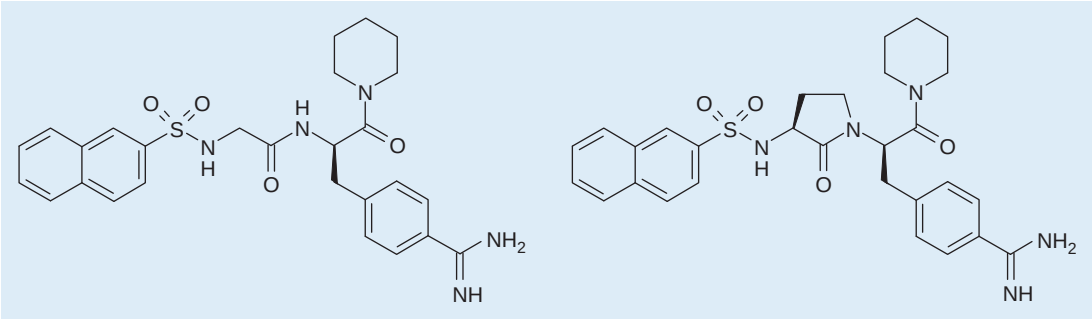


FIGURE 9.7 The ring-closed analog of the thrombin inhibitor NAPAP is 100 times more potent than the corresponding open-chain drug [12].

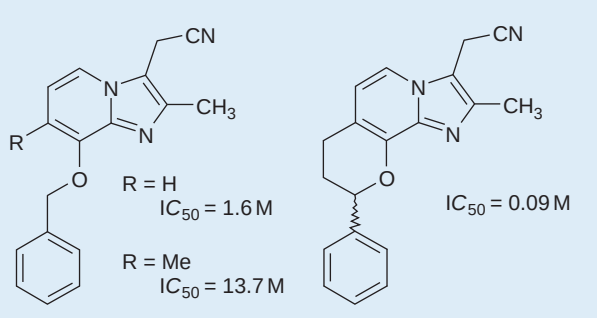


FIGURE 9.8 Ring closure achieves a conformational restriction and yields a highly potent compound.

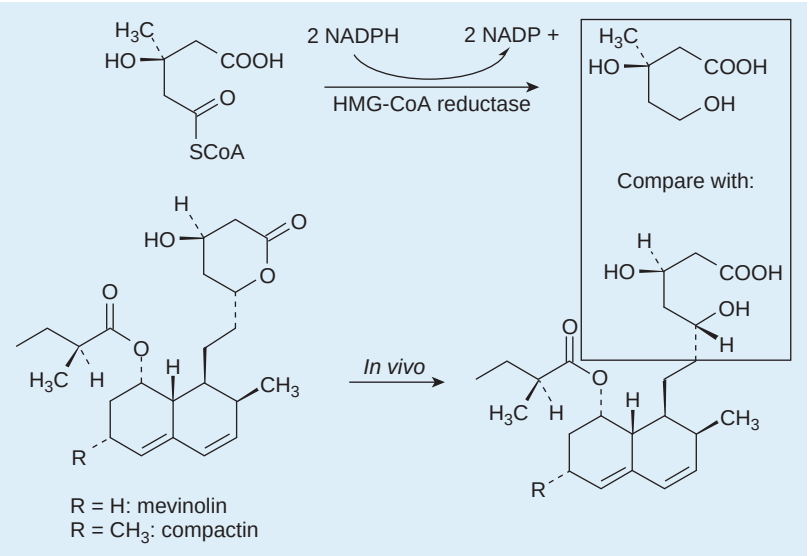


FIGURE 9.9 The chemically stable lactones mevinolin and compactin represent the ring-closed forms of the *in vivo* active parent substituted 3,5-dihydroxyvaleric acid.

In vivo, these ring-closed derivatives (Figure 9.9) are hydrolyzed to the open-chain 3,5-dihydroxyvaleric acid form that mimics the structure of the proposed intermediate in the reduction of HMG-CoA by HMG-CoA reductase [14].

3. *Arylpropionic Analgesic and Anti-inflammatory Drugs*

The potent analgesic benzoyl-indane carboxylic acid TAI-901 [15] is the cyclized analog of the well-known anti-inflammatory analgesic agent ketoprofene (Figure 9.10). The corresponding heterocyclic analogs were also prepared [16]. The compounds show potent analgesic activities with low gastric irritation.

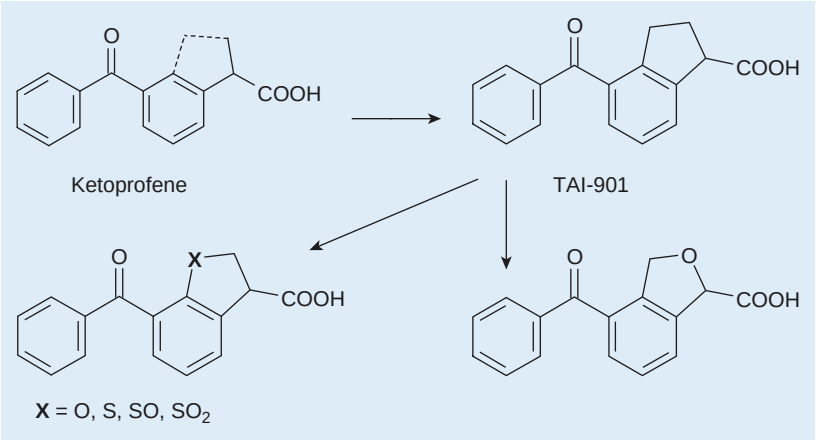


FIGURE 9.10 Ring-closed ketoprofene analogs.

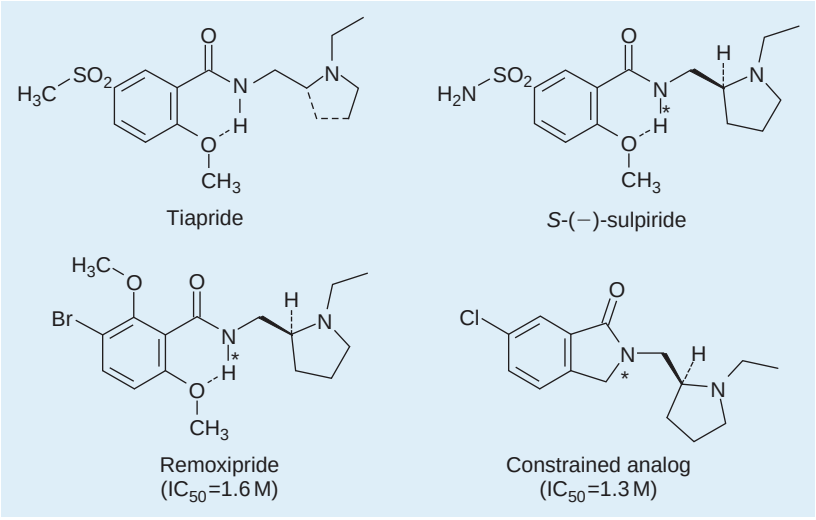


FIGURE 9.11 The typical sulpiride side-chain results formally from a ring closure of the diethylaminoethyl side-chain of earlier prepared derivatives such as tiapride. Observe the intramolecular hydrogen bond that creates a pseudocycle[15] and which can be mimicked by a covalent and constraint analog. [16]

4. *The Sulpiride Side-chain*

Among the numerous benzamide drugs developed by the Delagrange scientists, some have diethylaminoethyl side-chains (e.g., tiapride), whereas others, such as sulpiride, have *N*-ethyl-pyrrolidinyl-methyl side-chains that can be considered as the corresponding ring-closed analogs (Figure 9.11).

Both compounds are dopaminergic antagonists with neurotropic and antiemetic activity. Note that the ring closure creates an asymmetric center; the commercial form is the racemate, the slightly more active isomer being the *S*-(−)-sulpiride [17]. An additional constraining factor results from the establishment of a hydrogen bond between the amidic N-H hydrogen and the methoxy oxygen [18]. A conformationally restricted remoxipride analog in which the intramolecular hydrogen bond is replaced by a covalent bond (Figure 9.11) is equipotent in D2 receptor preparations [19].

5. *Cyclized Dopamine: The ADTN's*

The 2-amino-5,6-dihydroxy- and the 2-amino-6,7-dihydroxy-1,2,3,4-tetrahydro-naphtalenes are cyclized analogs of dopamine, corresponding to the α- and the β-rotamer, respectively (Figure 9.12).

As the cyclization generates a chiral center, four different ADTN's are possible, showing differential affinities for the dopamine receptors [20]. An extensive study of the aminotetralins and analogs containing additional rings (octahydrobenzo[g]quinolines) and compounds resulting from ring enlargements was published by Seiler et al [21].

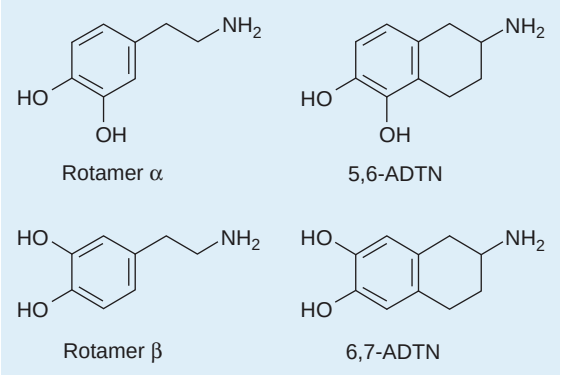


FIGURE 9.12 Ring-closed analogs of the two rotamers of dopamine.

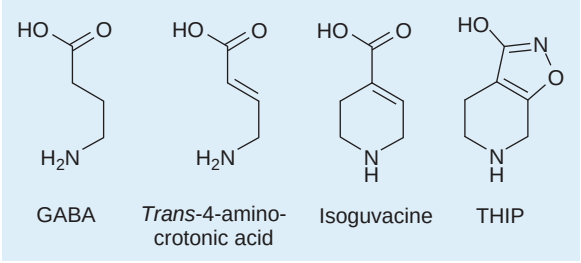


FIGURE 9.13 GABAergic agonists [22].

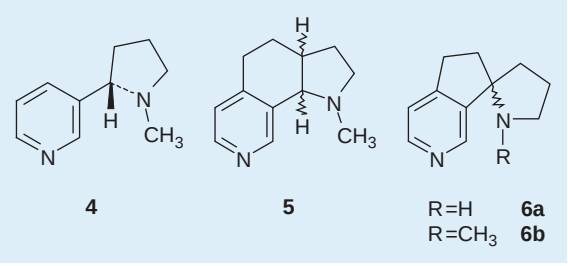


FIGURE 9.14 Natural (*S*)-(−)-nicotine and its bridged [25] and spiro[26] analogs.

6. *GABAergic Agonists*

The transition from γ-aminobutyric-acid (GABA) to *trans*-4-amino-crotonic acid, followed by cyclization into isoguvacine, and finally into THIP (Figure 9.13), simultaneously achieves the rigidification of the flexible GABA molecule and the production of THIP, a metabolically stable and still potent GABA agonist [22].

7. *Ring-Closed Analog of Nicotine*

Alzheimer's disease has received the most attention as a therapeutic target for nicotinic drugs, as nicotinic receptor binding was found to be significantly reduced in distinct brain regions of Alzheimer's patients [23]. Because the identification of various adverse physiological effects [23,24] of the alkaloid (*S*)-nicotine (**4**; Figure 9.14), synthetic modifications of its structure have been performed in order to improve potency and selectivity while reducing the toxicity.

Several groups [25,27] pursued the concept of conformational restriction in nicotine with the objective of forcing both pharmacophore nitrogen atoms into well-defined angles ("up" and "down" conformations), which would help in elucidating the "active" conformation of nicotine on the one hand and obtaining a tool to study receptor subtype specificity on the other hand. These endeavors led to, among others, the class of

hexahydropyrroloisoquinolines (**5**; Figure 9.14) [25]. The correct stereochemistry began to play an increasingly decisive role. While the relatively low enantioselectivity of nicotine (the affinity of (*S*)-nicotine is 10−100 times higher than that of (*R*)-nicotine) [28] has been an intriguing phenomenon for many years, the premise that conformational restraint of nicotine should enhance enantioselectivity has been well established [29].

The (−)-enantiomer of **5** shows low relative affinity ($K_i = 605$ nM) for the nAChR [3*H*]nicotine binding site, whereas the (+)-enantiomer fails to displace the radioligand even at 10 μM concentrations [27].

Spiro-annulated compounds **6a**,**b** (Figure 9.14) exhibited much higher binding affinities [26]. *nor*-Derivative (+)-**6a** bound at $K_i = 53.1$ nM with a 10-fold higher affinity than its enantiomer ($K_i = 533$ nM). Introduction of the *N*-methyl substituent resulted in a significant improvement, probably due to the gain of fortified receptor interaction: (+)-**6b** ($K_i = 4.79$ nM) appeared to be the most interesting ligand of this set as it not only bounded in the low nanomolar range but also exhibited a 30-fold higher affinity than its enantiomer ($K_i = 148$ nM) [26].

8. Cyclic Analogs of β-Blockers

Conventional β-blockers possess a number of pharmacological properties, such as β-blocking, quinidine-like, local anesthetic, and hypotensive. With the hope of achieving some specificity, Basil et al [30] considered the possibility of synthesizing ring-closed analogs (Figure 9.15). One of the compounds prepared, 3,4-dihydro-3-hydroxy-6-methyl-1,5-benzoxazocine, was a potent β-blocker. This activity is unlikely to be due to hydrolysis to the open-chain derivative, since the corresponding primary amine formed by hydrolysis of the benzoxazocine ring has less than 0.25 the activity of the latter. On the other hand, it is difficult to reconcile the benzoxazocine configuration with the structural requirements associated with the occupation of β-receptors.

Later studies by Evans et al also envisaged the synthesis of cyclized analogs of the phenylpropanolamine type of β-blockers [31,32]. The authors hoped that by restricting the conformation (by cyclizing the carbon atom bearing the terminal amino group to the aromatic ring; see Figure 9.15), β-blocking activity would be lost but antihypertensive activity might be retained. This turned out to be true in animal tests and in double-blind clinical studies, so the potassium channel activator cromakalim was developed [32].

9. Cyclized Diphenhydramine

Nefopam (Figure 9.16) is the representative of a new class of centrally acting skeletal muscle relaxants, also possessing a benzoxazocine structure [33]. Formally, nefopam is a cyclized analog of orphenadrine and diphenhydramine.

In contrast to the parent molecules, nefopam has no antihistaminic activity, keeping only the muscle relaxant effects. Clinically, nefopam is used as muscle relaxant, but—and this was originally not anticipated—it is also an antidepressant and an analgesic. These clinical indications can be explained by the interference of nefopam with the serotonergic transmission. More precisely, studies of the separated stereoisomers of nefopam explain its serotonin uptake properties and suggest that descending serotonergic pathways are involved in its antinociceptive activity [34].

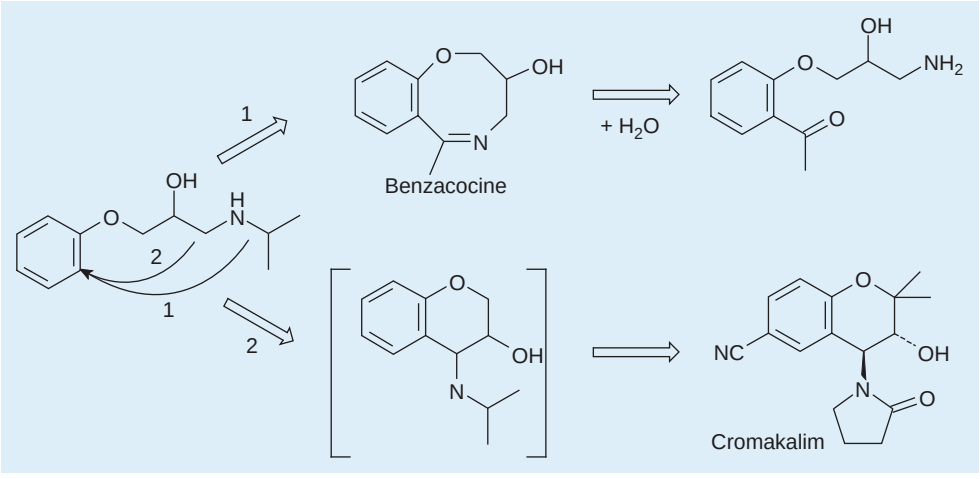


FIGURE 9.15 Cyclized analogs of β-blocking phenylpropanolamines [31,32].

10. *Rimonabant Analogs*

Several groups have recently reported constrained me-too analogs of rimonabant (compound **7**; Figure 9.17), the most advanced antagonist of cannabinoid type 1 receptor (CB-1), which was recently approved in the European Union for the treatment of obesity [36]. Compounds that incorporate conformational constraints into the tetrasubstituted pyrazole structure have the potential for increased binding activity and subtype selectivity (toward CB-2 receptor), resulting from preferential bioactive conformation (Figure 9.17). Tricycles analogs have been described including the central pyrazole ring and the 5-aryl group (compounds **8a**–**c**) [35]. The orientation of the 5-aryl group is imposed by the size of the central cycle. This leads to dramatic differences of affinities for the mouse CB-1 receptor, from micromolar to subpicomolar values. Surprisingly, the selectivity for mCB-1 versus mCB-2 varies from 60,000:1 to 1:6,000.

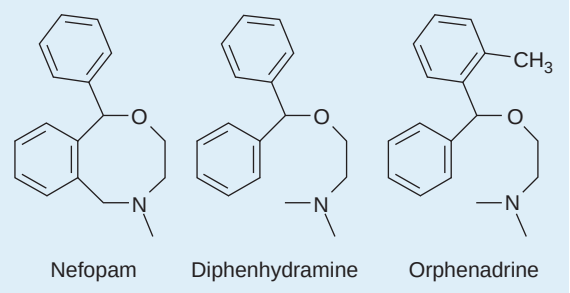


FIGURE 9.16 Nefopam is a cyclized analog of orphenadrine and diphenhydramine [33].

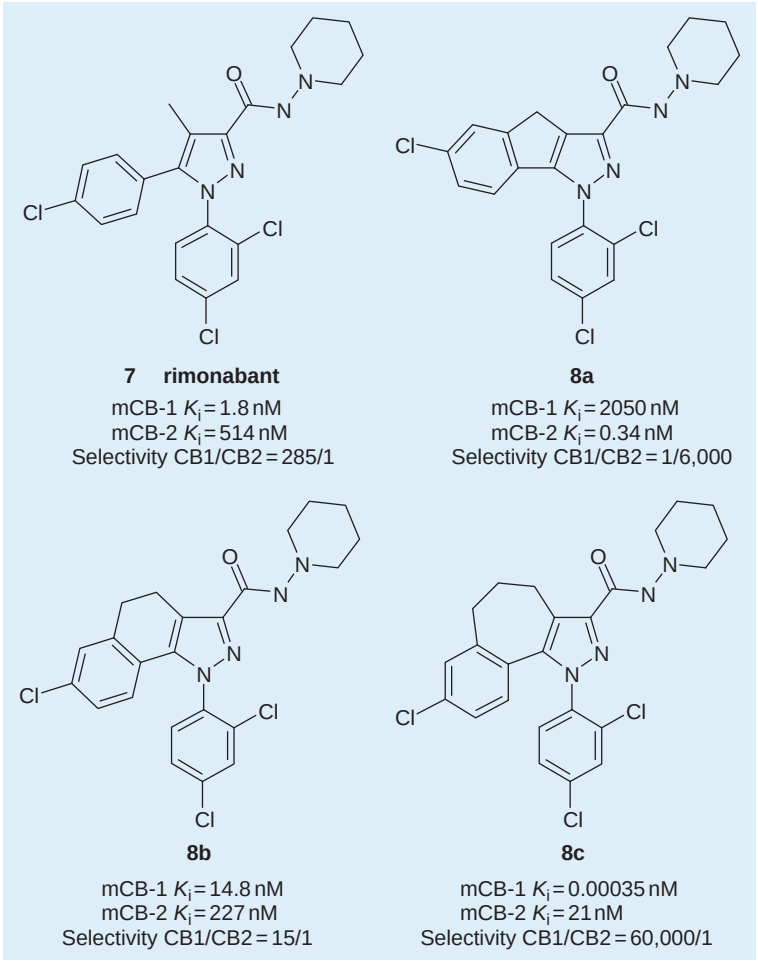


FIGURE 9.17 Cannabinoid mouse receptor (mCB-1 and mCB-2) affinity and selectivity for rimonabant **7** and compounds **8a**–**c** [35].

Compounds having a constraint between the pyrazole core C-4 position and the carboxamide moiety have also been reported (Figure 9.18, compound **9**). These pyrazolopirimidinone derivatives are slightly less active in the human CB-1 binding assay than rimonabant ($K_i = 12$ versus 2.1 nM)) [37]. This could be explained by the splitting of carboxamide group in a nonfavorable orientation.

To constrain this carboxamide hydrogen donor/acceptor group to exhibit the preferential bioactive conformation, a novel series of analogs has been prepared [38]. In this case, an ethylene bridge is incorporated between the pyrazole N-2 position and the carboxamide nitrogen (Figure 9.18, compound **10**). The potency of compound **10** is similar to that observed for compound **7** in the same *in vitro* human CB-1 assay ($K_i = 2.2$ versus 1.1 nM).

The slight structural variation of constrained analogs compared to rimonabant is shown in Figure 9.19. The superposition of low-energy conformations [39] of **7** with the tricyclic analogs (**8c**, **9**, and **10**) are in accordance with the results obtained in affinity binding assays regarding the position of the three substitutions around the central pyrazole core.

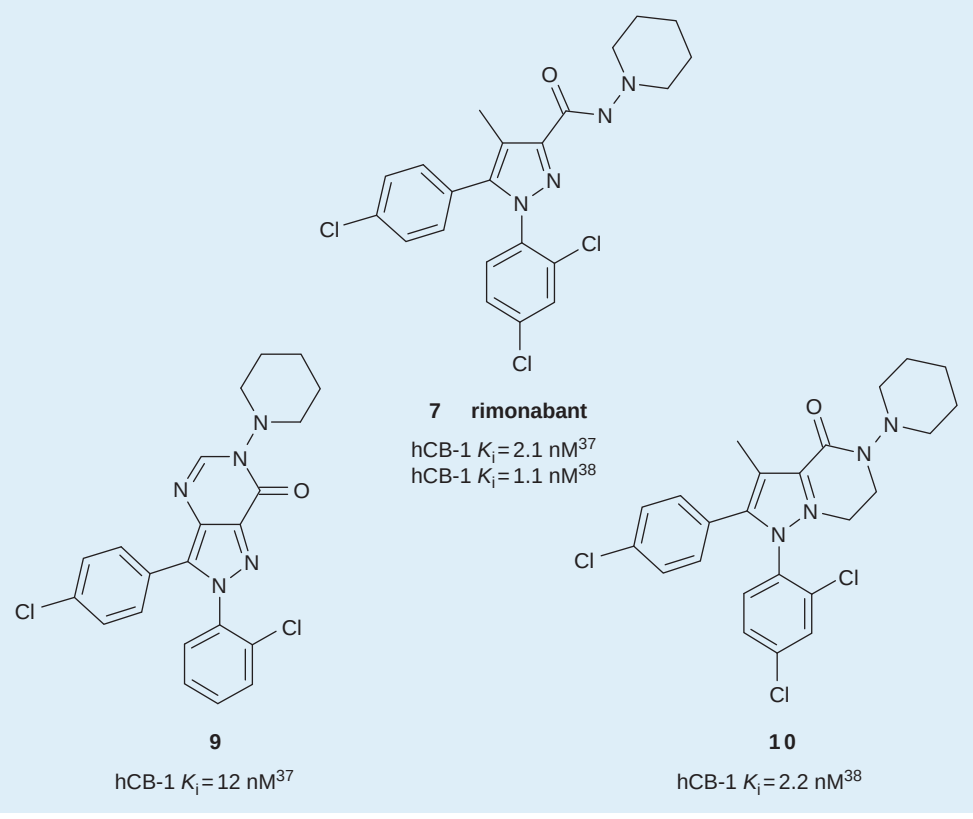


FIGURE 9.18 Cannabinoid human receptor (hCB-1) affinity for rimonabant **7**, compounds **9** and **10** [37,38].

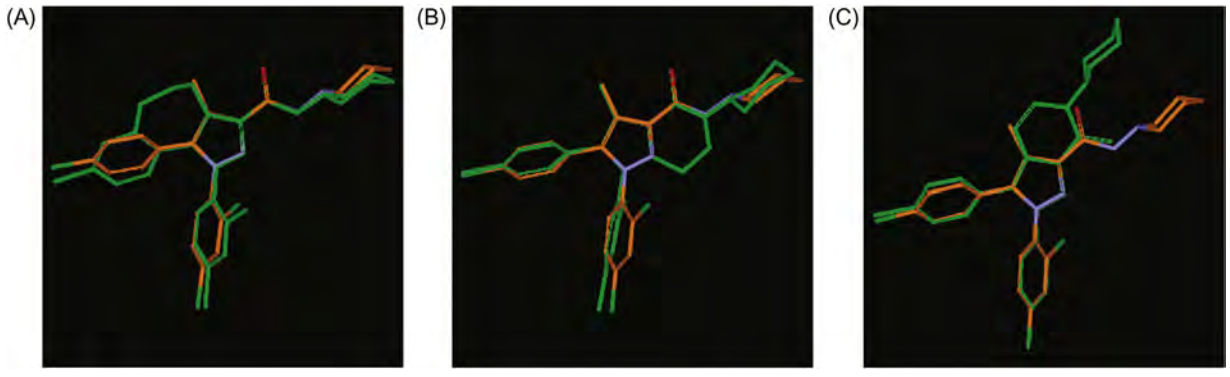


FIGURE 9.19 Overlay of the low-energy conformation of rimonabant 7 successfully with compound 8c (A), compound 10 (B), and compound 9 (C) [39].

11. Ring Variations Around Phenylbutazone

The anti-inflammatory drug phenylbutazone (compound **11**; Figure 9.20) led to many me-too copies, such as the ring-opened analog bumadizon **12** (Ca^{++} salt = Eumotol® [40]) or the ring-closed analog apazone **13** (Prolixan®) [41]. The cinnoline derivative **14** [42] results again from a ring closure and served as model for the design of its open counterpart, the styrylbutazone **15** [43]. The quinolinyl-3,5-dioxo-pyrazolidine **16** [44] represents another interesting ring variation with a 7-chloro-quinoline moiety in the butazone portion.

C. Other Analogies

Applied to ring systems, the following molecular modifications seem to be conducted mainly with the objective of bypassing patent protections and allowing the synthesis of me-too products.

1. Ring Enlargement and Ring Contraction

Ring enlargement and ring contraction can be considered homologous variations in the cyclic series (mentioned in Chapter 14).

A. BARBITURICS AND OPIOIDS

In Figure 9.21, two additional examples, taken from the barbituric and from the opiate series respectively, are shown. In the case of the change of the barbiturics to hydantoins, the contraction is accompanied by the loss of a carbonyl group (Figure 9.22). Nevertheless, potent antiepileptics are found in both series.

B. INOGATRAN AND MELAGATRAN

The classical motif of thrombin inhibitors is the D-Phe-Pro-Arg sequence mimicking thrombin's natural substrate, fibrinogen. In development candidates such as inogatran [45] and melagatran [46] (Figure 9.23), the proline unit was replaced by its ring-expanded and its ring-contracted equivalents respectively.

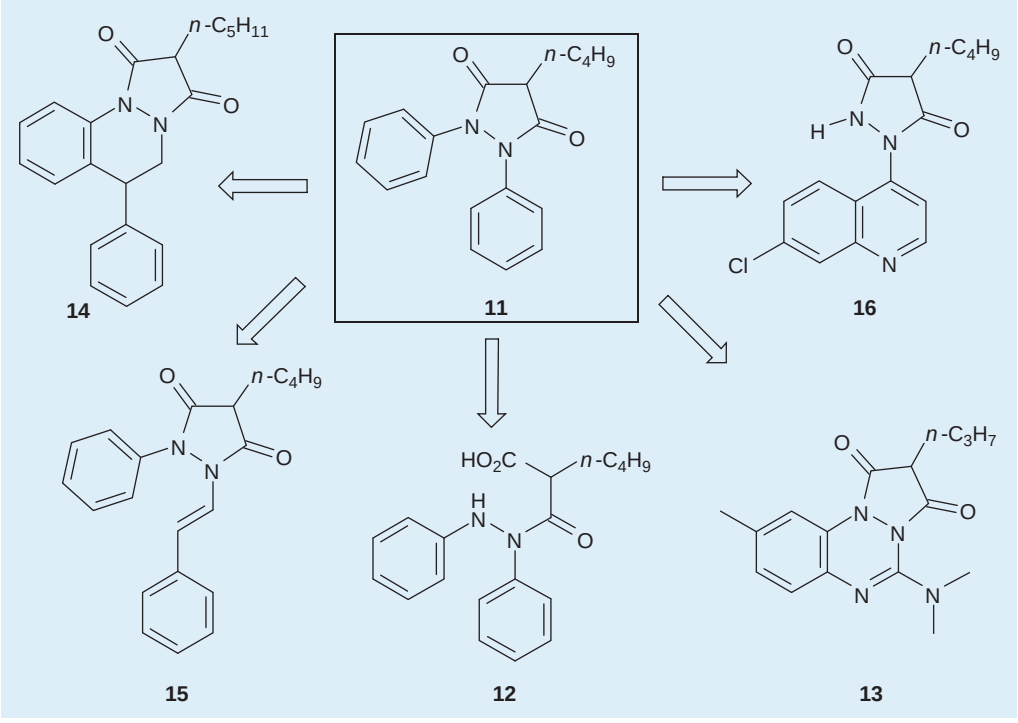


FIGURE 9.20 Successive ring openings and closures in phenylbutazone-derived anti-inflammatory drugs.

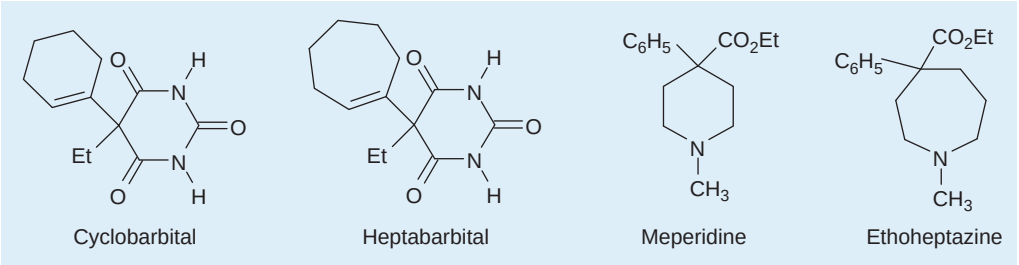


FIGURE 9.21 Six-membered rings exchanged for seven-membered rings.

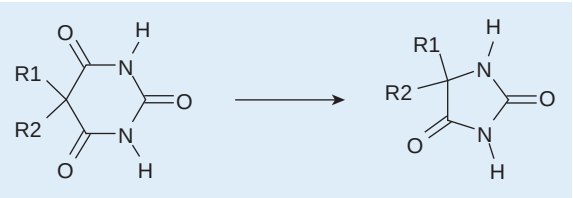


FIGURE 9.22 (a) Barbiturates (left) and hydantoins (right).

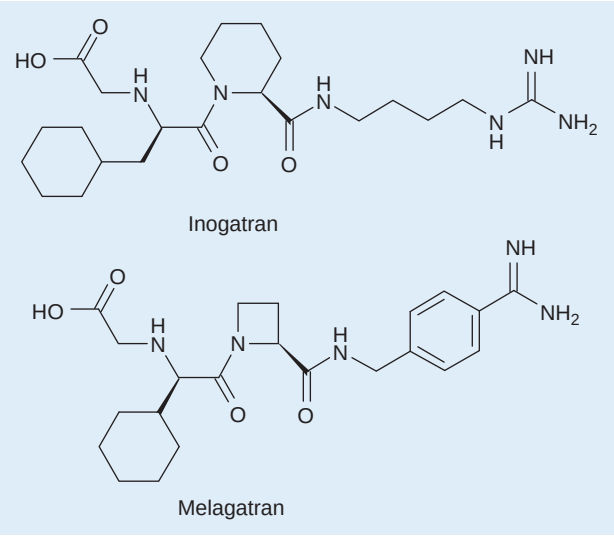


FIGURE 9.23 Structures of the thrombin inhibitors inogatran and melagatran [45,46].

C. AMPA RECEPTOR ANTAGONISTS

Several 2,3-benzodiazepin-4-ones such as compound GYKI 53 655 (Figure 9.24) are AMPA receptor antagonists and possess noteworthy anticonvulsant and neuroprotective properties. The corresponding ring-contracted analogs, 6,7-methylene-dioxydihydrophtalazines (compound SYM 2207) and 6,7-methylene-dioxyphtalazin-1(2*H*)-ones, possess a similar activity profile (compound **17**) [47,48].

D. OXOTREMORINE

Oxotremorine and its ring-opened analog Oxo-2 (Figure 9.25) are partial muscarinic agonists producing large guanine nucleotide shifts in the heart (32 and 23 respectively), suggesting strong M2 agonist-like effects [49].

The corresponding piperidinic analog Oxo-Pip [50], having a predicted antagonist [^{3}H]QNB/[^{3}H] CD ratio of 2.2, produced only a weak shift (5.0) in the concentration–response curve with the addition of the stable guanine nucleotide analog [49]. Thus, the change from a pyrrolidine to a piperidine ring is able to change a partial agonist into an antagonist.

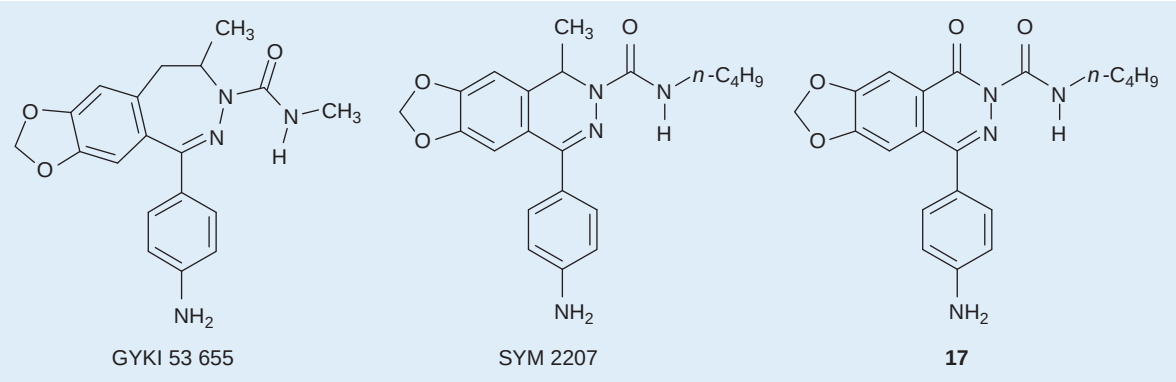


FIGURE 9.24 The heptacyclic AMPA antagonist GYKI 53 655 and ring-contracted analogs [47,48].

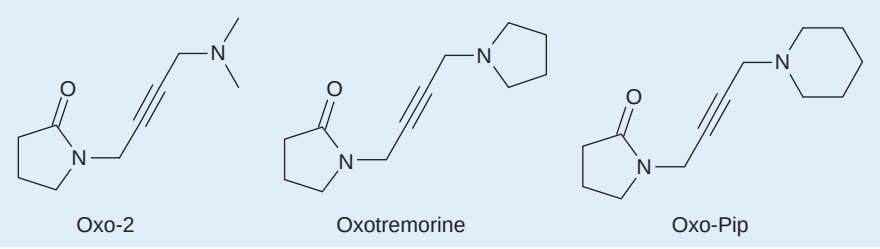


FIGURE 9.25 Oxotremorine yields ring-opened and ring-extended analogs [49].

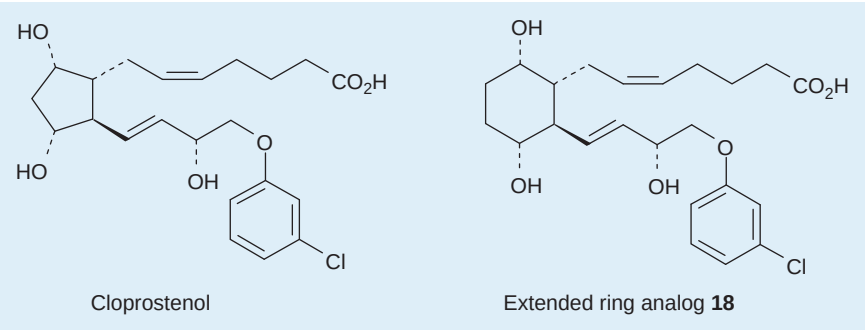


FIGURE 9.26 Cloprostenol and ring-extended analog [51].

The replacement of the core cyclopentane ring of the prostaglandin FP agonist cloprostenol with a cyclohexane ring (compound **18**; Figure 9.26) yielded a clearly less active agent (EC_{50} = 319 nM instead of 1 nM) [51].

2. *Reorganization of Ring Systems*

The four molecular variations described below represent some more "exotic" approaches to the design and manipulation of the original ring systems. They may bring useful alternatives allowing escape from overcrowded avenues of research.

A. TRANSFORMING SIMPLE RINGS INTO SPIRO DERIVATIVES OR INTO BI- OR TRICYCLIC SYSTEMS

A first example is in the guanethidine analogs [52]. As the original guanethidine patents covered ring sizes varying from five-membered to ten-membered rings, a possible way to get round them was the design of isolipophilic *spiro* systems (Dausse compounds **a** and **b** [53,54]; Figure 9.27). Another possibility, originating from Takeda scientists, implies the use of an azetidine surrogate for the ethylene-diamine chain [55]. Finally, polycyclic systems can replace the octahydroazocine ring, as illustrated by the bicyclic compounds **c** and **d** from Dausse [56] or by the tricyclic compound from Lumière Laboratories [57,58]. Many other imaginative solutions were proposed; they are well reviewed by Mull and Maxwell [52]. More recently, similar variations were applied to the design of an impressive number of analogs of the anticonvulsant drug gabapentin (Figure 9.28) [59,60].

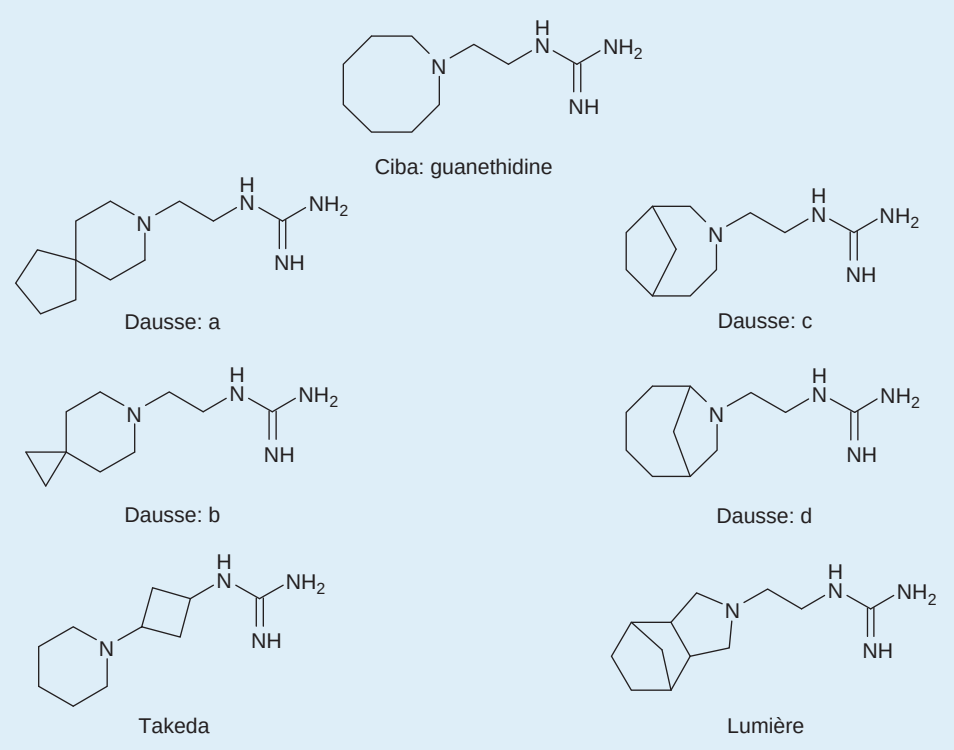


FIGURE 9.27 Alternative possibilities in the design of guanethidine analogs.

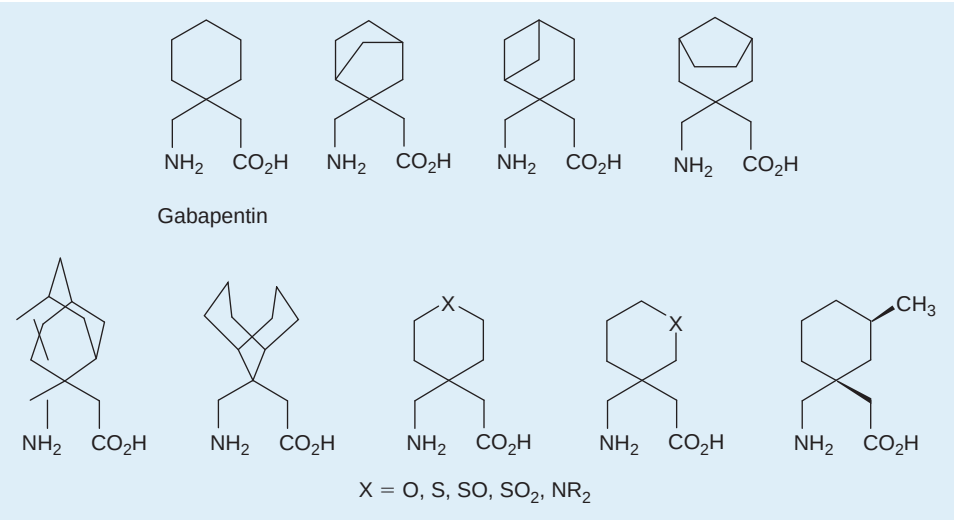


FIGURE 9.28 Alternative possibilities in the design of gabapentin analogs.

B. SPLITTING BENZO COMPOUNDS ("BENZO CRACKING")

Dissociation of a fused ring system (Figure 9.29), particularly by splitting a benzo compound, can sometimes improve its solubility and nevertheless only alter slightly its pharmacokinetic profile and its long-term toxicity.

C. RESTRUCTURING RING SYSTEMS

Among the above-mentioned molecular variations on ring systems, some can be used simultaneously. Thus, the splitting of the benzimidazole heterocycle in the anthelmintic thiabendazole and the concomitant association of the two five-membered rings yield tetramizole (Figure 9.30). One of the two enantiomers of tetramizole, the L-(−)-form, or levamizole, is also a potent anthelmintic.

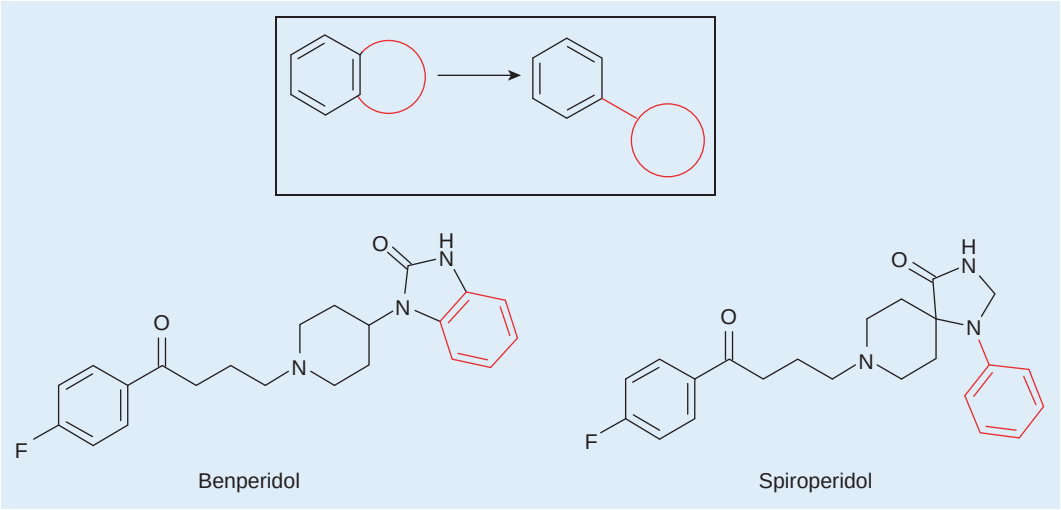


FIGURE 9.29 Splitting of fused rings often yields drugs with similar activity, sometimes with improved solubility and/or less toxicity.

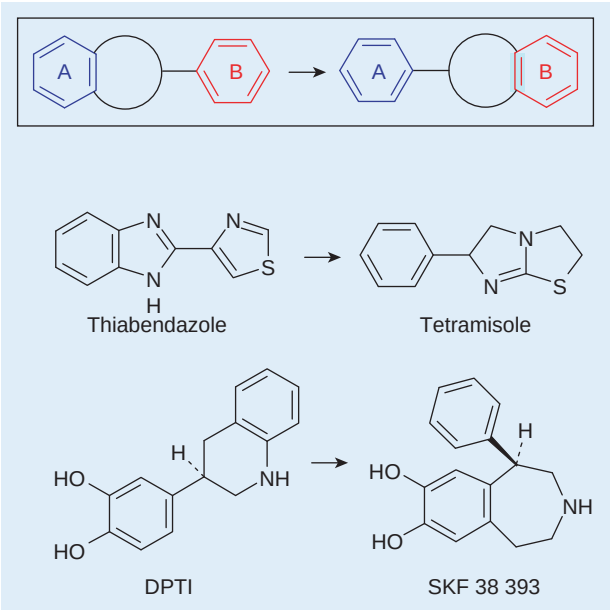


FIGURE 9.30 Restructured ring systems.

The change from the D_1-selective dopaminergic agonist **DPTI** (3-(3,4-dihydroxyphenyl)-1,2,3,4-tetrahydroisoquinoline) to the equally D_1-selective compound **SKF 38 393** is a combination of benzo cracking, a new benzo fusion, and ring enlargement (Figure 9.30). As a result, the compounds still resemble each other and are both recognized by the dopamine D_1-receptor [21].

An interesting example of a restructured ring system was designed by Meyer et al [61] for the design of 5-methoxy-hexahydro-1*H*-benz[f]isoindole as a surrogate of the very frequently used *ortho*-methoxy-*N*-arylpiperazine. It led both to an enhancement of affinity and improved selectivity for α_{1A} receptor antagonism (Figure 9.31).

D. RING DISSOCIATION

The natural compound khellin generated two families of cardioactive drugs: on one side the benzopyrones, illustrated by the 3-methyl-chromone [62] and chromonar (carbochromen); on the other side the benzofurans, illustrated by amiodarone (Figure 9.32). Both families possess anti-arrhythmic and anti-anginal properties.

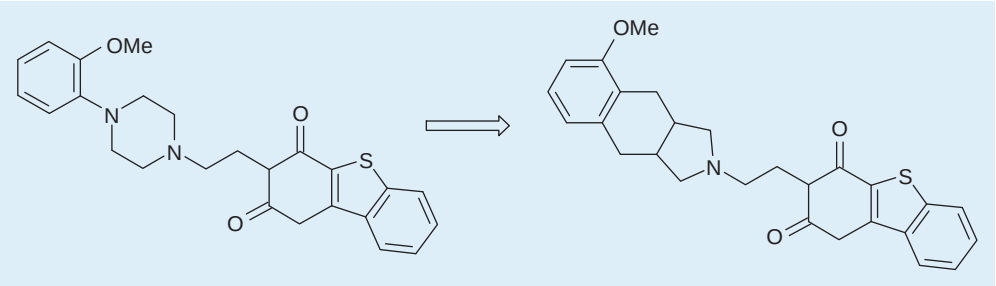


FIGURE 9.31 Hexahydrobenz[f]isoindole as a surrogate of the very frequently used *ortho*-methoxy-*N*-arylpiperazine.

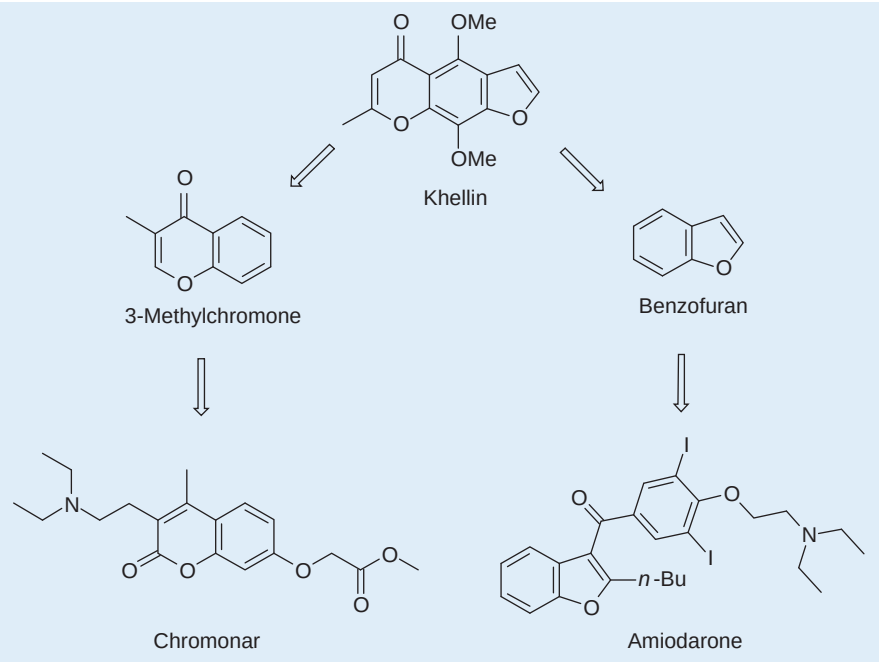


FIGURE 9.32 Cardioactive drugs obtained by dissociation of the khellin molecule into benzopyrones and benzofurans.

III. DISJUNCTIVE APPROACHES

Starting from a polycyclic structure (which is often of natural origin), the chemist proceeds to progressive pruning of the molecule ("molecular strip tease"). Sometimes very simple reasoning guides the medicinal chemist, and the final compound has only a remote resemblance to the model compound. Such an exercise led to the transformation of the natural compound asperlicin to a totally synthetic simplified benzodiazepine derivative (see below, Figure 9.36) [63].

A. Cocaine-Derived Local Anesthetics

Figure 9.33 illustrates how simplified synthetic copies of cocaine were designed. The change from cocaine to procaine retains the local anesthetic effects without keeping the narcotic properties.

B. Morphinic Analgesics

Probably more than a thousand of simplified analogs of the alkaloid morphine were investigated [64]. Many of them were just inactive, but it was soon recognized that the phenylpiperidine unit was crucial for the central analgesic properties (Figure 9.34). In contrast to what was observed for cocaine, no clear discrimination between the analgesic and the narcotic properties could be achieved.

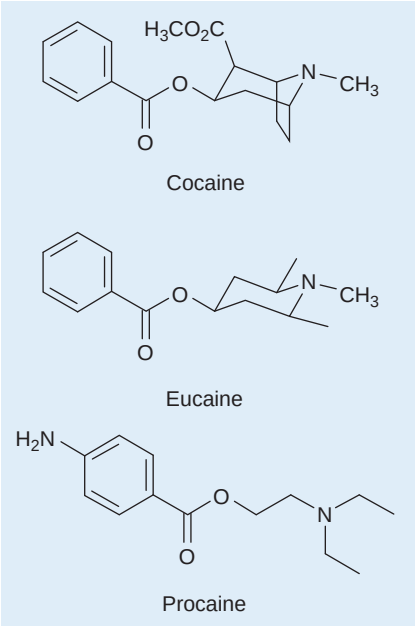


FIGURE 9.33 Progressive simplification of the cocaine molecule.

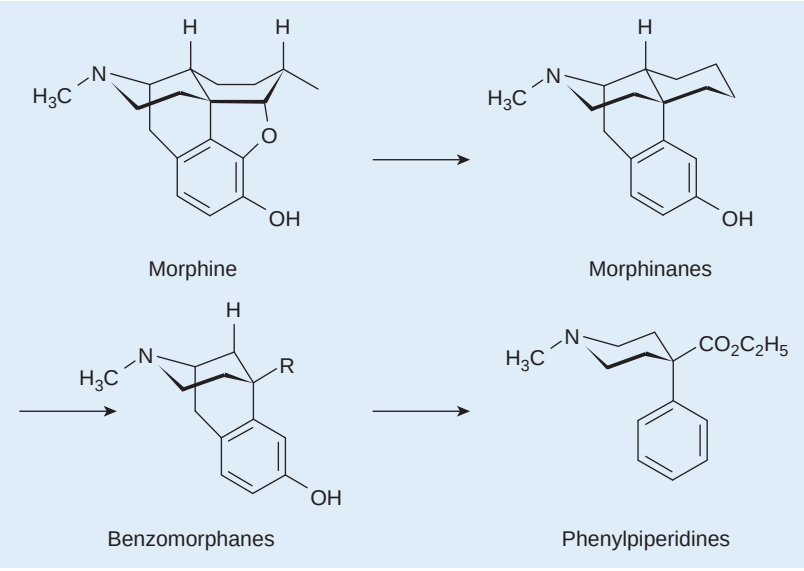


FIGURE 9.34 Progressive simplification of the morphine molecule.

C. Dopamine Autoreceptor Agonists

The discovery of 3-(3-hydroxyphenyl)-*N*-*n*-propyl piperidine ((±)-3-PPP), a centrally acting dopamine receptor agonist with selectivity for dopaminergic autoreceptors [65], offers a potential alternative to neuroleptics in the treatment of schizophrenia. The structure of (±)-3-PPP (Figure 9.35) can be considered a product resulting from a disjunctive approach applied to pergolide [67,68]. Surprisingly, an increase in the pergolide-like character of 3-PPP, through incorporation of a methylmercaptomethyl group (compound **19**), did not improve the potency [66].

D. CCK Antagonists

After the discovery of the potent CCK antagonistic activity of the natural compound asperlicin [69], the scientists from the Merck group first prepared some simple semi-synthetic derivatives [70]. Then, recognizing in

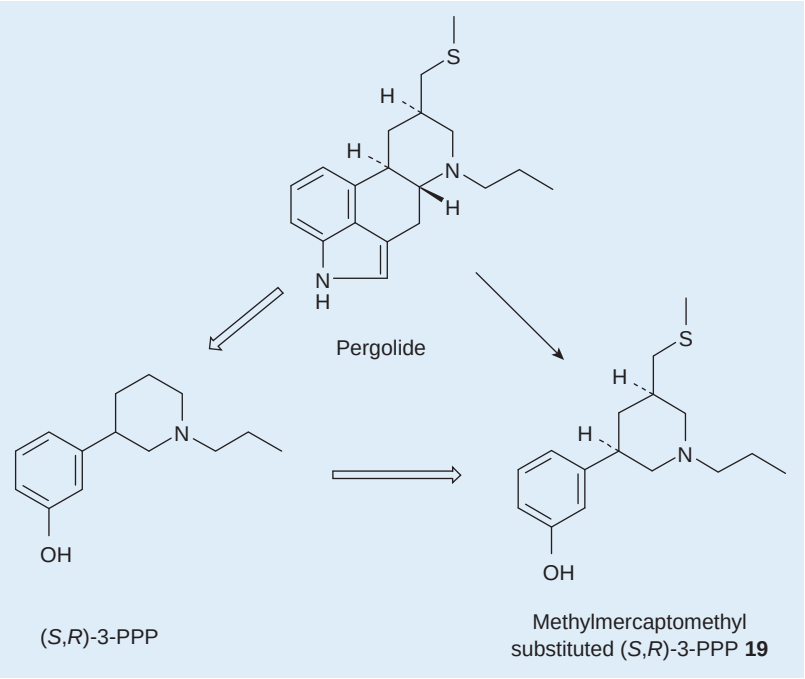


FIGURE 9.35 3-PPP is a result of the disjunctive approach applied to pergolide [65,66].

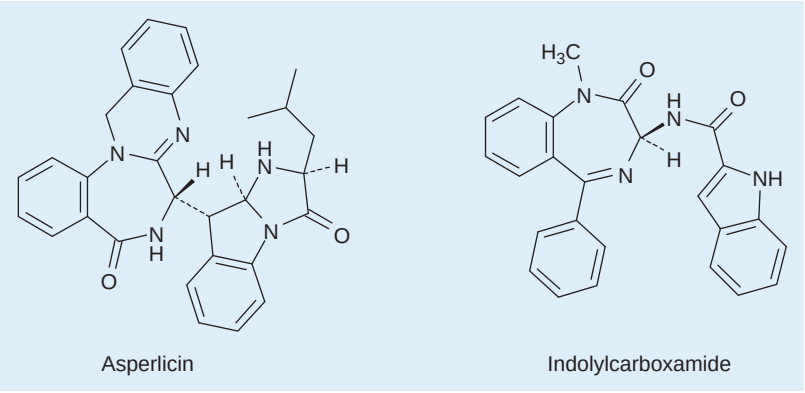


FIGURE 9.36 Productive disjunction of the asperlicin molecule [63].

asperlicin the elements of a benzodiazepinone and a tetrahydroindole, they followed the hunch that these elements alone may confer some CCK antagonistic activity [63]. This reasoning proved to be valid (Figure 9.36).

IV. CONJUNCTIVE APPROACHES

As already mentioned at the beginning of this chapter, the purpose of the conjunctive method lies in the design of compounds structurally more complex than the lead compound. In practice, this is generally achieved by creating and/or adding supplementary ring systems to constrain the molecule and to impose specific conformations.

A. Dopaminergic Antagonists

Starting from the flexible haloperidol molecule, Humber et al [71] designed the rigid (+)-butaclamol, which contains three clearly defined stereocenters (Figure 9.37). The same stereochemical requirements as in (+)-butaclamol are found in compound Ro 14-8625 prepared by Imhof et al [72].

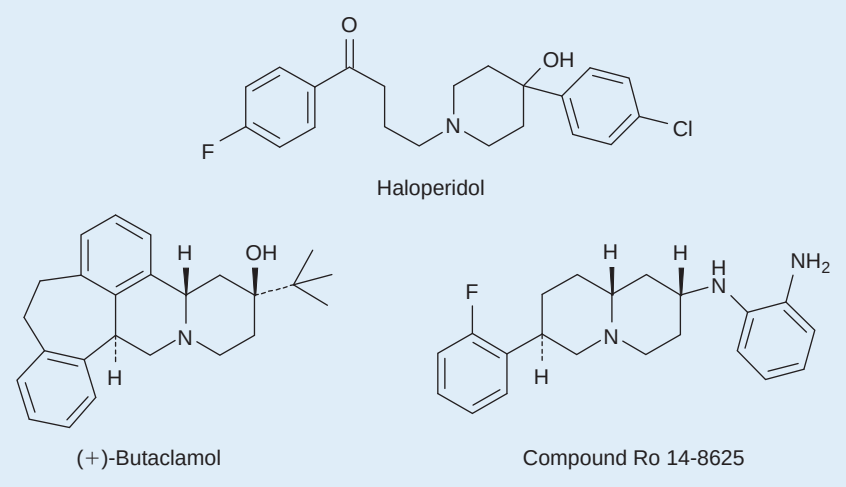


FIGURE 9.37 The conjunctive method applied to the design of haloperidol-derived dopaminergic antagonists.

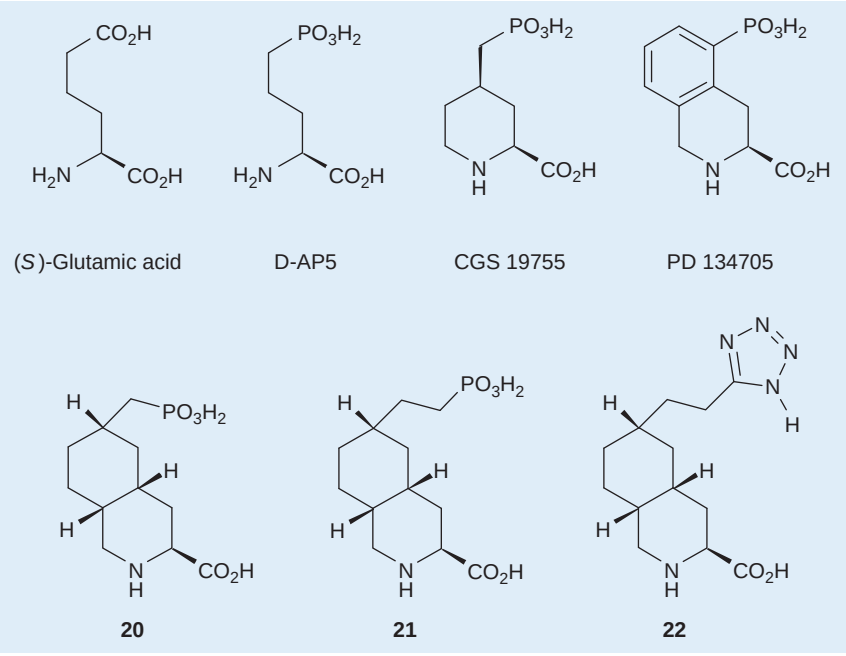


FIGURE 9.38 The conjunctive method applied to the design of NMDA and AMPA receptor antagonists.

B. Glutamate NMDA and AMPA Receptor Antagonists

The progressive change of glutamic acid to D-AP5 [73], then to the piperidine analog CGS 19755 [74], and finally to the tetrahydroisoquinoline PD 134705 [75] led to NMDA receptor antagonists. Similarly, rigidification into the perhydroquinolines **20**, **21**, and **22** [76] (Figure 9.38) illustrates another application of the conjunctive approach that led to potent AMPA antagonists. In addition to the elements enhancing structural rigidity, these latter compounds contain three new chiral centers.

C. Norfloxacin Analogs

Since the development in 1980 of *norfloxacin* [77] as a useful antibacterial agent, a large number of analogs have been synthesized. Among them, the conjunctive approach led to highly potent tetracyclic analogs (Figure 9.39) [78].

In a comparable way, a number of annelated analogs of the 5-HT_3 antagonist ondansetron were investigated (Figure 9.40). Among them, cilansetron ($n = 1$) was found to be about ten times more potent without loss in selectivity [79].

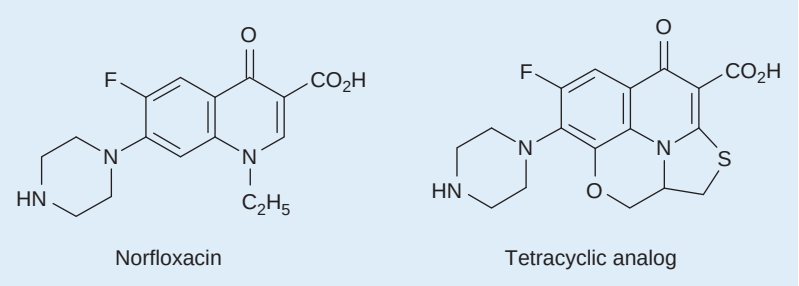


FIGURE 9.39 Norfloxacin and its tetracyclic analog [78].

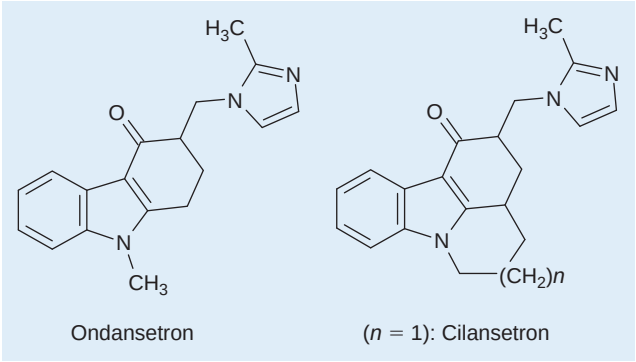


FIGURE 9.40 Cilansetron, an annelated analog of ondansetron [79].

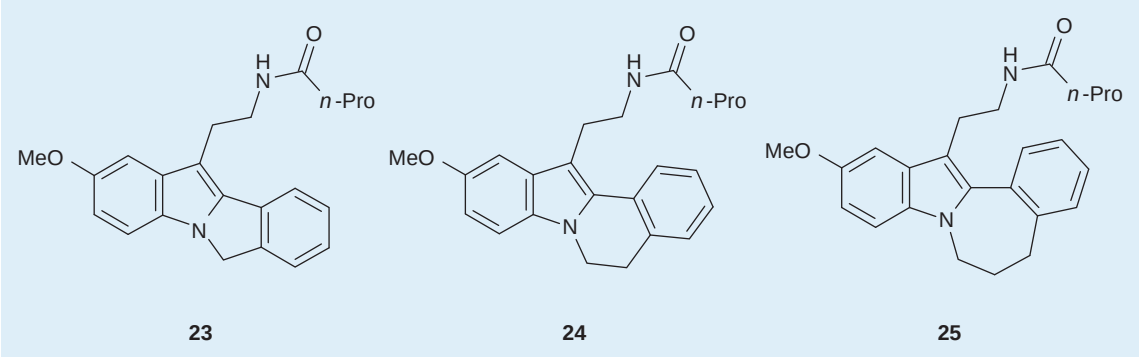


FIGURE 9.41 Ring-fused melatonin analogs [80].

D. Melatonin Analogs

Tetracyclic analogs in which the melatonin indole ring is fused to a dihydroindole, a tetrahydroisoquinoline, and a benzazepine system, respectively (compounds **23**, **24** and **25**; Figure 9.41), represent ligands for the melatonin MT2 receptor as potent as melatonin but selective with regard to the MT1 receptor. Interestingly, the passage from the six-membered (compound **24**) to the seven-membered fused ring system (compound **25**) produced a switch from agonist to antagonist activity [80].

The well-known angiotensin converting enzyme (ACE) inhibitor captopril has a relatively simple structure and therefore represents an excellent starting material for the conjunctive approaches (Figure 9.42).

Modeling studies based on a template structure constructed from the superposition of the energy-minimized benzo-fused ACE inhibitors shown in Figure 9.42 (compounds **27** and **29**) suggested the synthesis of the 13-membered heterocyclic lactam analog (compound **28**) [81].

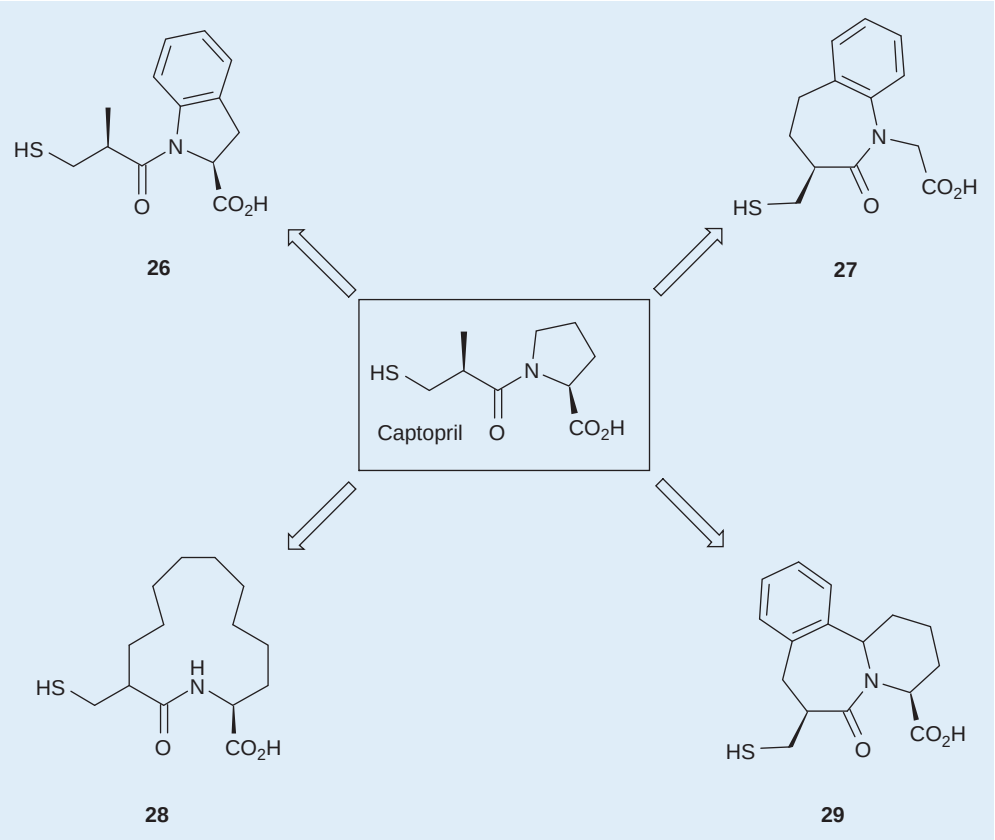


FIGURE 9.42 ACE inhibitors derived from captopril.

V. CONCLUSION

Molecular variations involving the study of homologous series or the application of the vinylogy concept induce relatively minor changes to the pharmacological profile and rather result in optimizing the potency. Modifying ring systems—ring-chain transformation, ring contractions and expansions, and reorganization of cyclic systems—represents a highly productive approach in the design of new drug analogs and in the exploration of the drug–receptor interactions.

References

[1] Schueller FW. Chemobiodynamics and drug design. New York: McGraw-Hill; 1960.
[2] Curd FHS, Davey DG, Rose FL. Studies on synthetic antimalarial drugs-X: some biguanide derivatives as new types of antimalarial substances with both therapeutic and causal prophylactic activity. Ann Trop Med 1945;39:208–14.
[3] Crowther AF, Levi AA. Proguanyl: the isolation of a metabolite with high antimalarial activity. Br J Pharmacol 1953;8:93–101.
[4] Hawking F, Perry WLM. Activation of paludrine. Br J Pharmacol 1948;3:320–31.
[5] Weiner IM. Drugs affecting renal function and electrolyte metabolism. In: Goodman-Gilman A, Rall TW, Nies AS, Taylor P, editors. Goodman and Gilman's the pharmacological basis of therapeutics. 8th ed. New York: Pergamon Press; 1990. p. 708–31.
[6] Buzetta B, Hospital M. Relations structure-activité [Structure-activity relationships]. In: Cohen Y, editor. Pharmacologie moléculaire [Molecular pharmacology]. Paris: Masson & Cie; 1978. p. 27–36. [in French].
[7] Minutolo F, Bertini S, Papi C, Carlson KE, Katzenellenbogen JA, Macchia M. Salicylaldoxime moiety as a phenolic "A-Ring" substitute in estrogen receptor ligands. J Med Chem 2001;44(24):4288–91.
[8] Minutolo F, Antonello M, Bertini S, Ortore G, Placanica G, Rapposelli S, et al. Novel estrogen receptor ligands based on an anthranylaldoxime structure: role of the phenol-type pseudocycle in the binding process. J Med Chem 2003;46(19):4032–42.
[9] Rouot B, Leclerc G, Wermuth CG, Miesch F, Schwartz J. Synthèse et essais pharmacologiques d'arylguanidines, analogues ouverts de la clonidine [Synthesis and pharmacological tests on arylguanidines, open analogs of clonidine]. Eur J Med Chem 1978;13:337–42. [in French].

[10] Ashwood VA, Cassidy F, Coldwell MC, Evans JM, Hamilton TC, Howlett DR, et al. Synthesis and antihypertensive activity of 4-(substituted-carbonylamino)-2H-1-benzopyrans. J Med Chem 1990;33:2667–72.
[11] Mack H, Pfeiffer T, Hornberger W, Bohm HJ, Hoffken HW. Design, synthesis, and biological activity of novel rigid amidinophenylalanine derivatives as inhibitors of thrombin. J Enzyme Inhib 1995;9(1):73–86.
[12] Kubinyi H. Molecular similarity. 2. The structural basis of drug design. Pharm Unserer Zeit 1998;27(4):158–72.
[13] Kaminski JJ, Wallmark B, Briving C, Andersson BM. Antiulcer agents. 5. Inhibition of gastric H1/K(1)-ATPase by substituted imidazo [1,2-a]pyridines and related analogues and its implication in modeling the high affinity potassium ion binding site of the gastric proton pump enzyme. J Med Chem 1991;34(2):533–41.
[14] Nakamura CE, Abeles RH. Mode of interaction of β-hydroxy-β-methylglutaryl coenzyme A reductase with strong binding inhibitors: compactin and related compounds. Biochemistry 1985;24:1364–76.
[15] Kawai K, Tamura S, Morimoto S, Ishii H, Kuzuna S. Pharmacology of 4-benzoyl-1-indancarboxylic acid (TAI-901) and 4-(4-methylbenzoyl)-1-indancarboxylic acid (TAI-908). Arzneim Forsch 1982;32:113–17.
[16] Boyle EA, Mangan FR, Markwell RE, Smith SA, Thomson MJ, Ward RW, et al. 7-Aroyl-2,3-dihydrobenzo[b]furan-3-carboxylic acids and 7-benzoyl-2,3-dihydrobenzo[b]thiophene-3-carboxylic acids as analgesic agents. J Med Chem 1986;29:894–8.
[17] Jenner P, Clow A, Reavill C, Theodoru A, Marsden CD. Stereoselective actions of substituted benzamide drugs on cerebral dopamine mechanisms. J Pharm Pharmacol 1980;32:39–44.
[18] Waterbeemd van de H, Testa B. Theoretical conformational studies of some dopamine antagonistic benzamide drugs: 3-pyrrolidyl- and 4-piperidyl derivatives. J Med Chem 1983;26:203–7.
[19] Norman MH, Kelley JL, Hollingsworth EB. Conformationally restricted analogues of remoxipride as potential antipsychotic agents. J Med Chem 1993;36:3417–23.
[20] Cannon JG, Lee T, Goldman HD, Costall B, Naylor RJ. Central dopamine agonist properties of some 2-aminotetralin derivatives after peripheral and intracerebral administration. J Med Chem 1977;20:1111–16.
[21] Seiler MP, Bölsterli JJ, Floersheim P, Hagenbach A, Markstein R, Pfäffli P, et al. Recognition at dopamine receptor subtypes. In: Testa B, Kyburz E, Fuhrer W, Giger R, editors. Perspectives in medicinal chemistry. Basel and Weinheim: Verlag Helvetica Chemica Acta and VCH; 1993. p. 221–37.
[22] Krogsgaard-Larsen P, Hjeds H, Falch E, Jørgensen FS, Nielsen L. Recent advances in GABA agonists, antagonists, and uptake inhibitors: structure–activity relationships and therapeutic potential. In: Testa B, editor. Adv drug res, vol. 17. London: Academic Press; 1988. p. 381–456.
[23] Levin ED, Simon BB. Nicotinic acetylcholine involvement in cognitive function in animals. Psychopharmacology (Berl) 1998;138 (3–4):217–30.
[24] Williams M, Arneric SP. Beyond the tobacco debate: dissecting out of the therapeutic potential of nicotine. Expert Opin Invest Drugs 1996;5:1035–45.
[25] Glassco W, Suchocki J, George C, Martin BR, May EL. Synthesis, optical resolution, absolute configuration, and preliminary pharmacology of (1)- and (2)-cis-2,3,3a,4,5,9b-hexahydro-1-methyl-1H-pyrrolo-[3,2-h]isoquinoline, a structural analog of nicotine. J Med Chem 1993;36:3381–5.
[26] Ullrich T, Krich S, Binder D, Mereiter K, Anderson DJ, Meyer MD, et al. Conformationally constrained nicotines: polycyclic, bridged, and spiro-annulated analogues as novel ligands for the nicotinic acetylcholine receptor. J Med Chem 2002;45 (18):4047–54.
[27] Glennon RA, Dukat M. Nicotine receptor ligands. Med Chem Res 1996;465–86.
[28] Holladay MW, Lebold SA, Lin NH. Structure–activity relationships of nicotinic acetylcholine receptor agonists as potential treatments for dementia. Drug Dev Res 1995;35:191–213.
[29] Damaj MI, Glassco W, Marks MJ, Slobe B, James JR, May EL, et al. Pharmacological investigation of (1)- and (2)-cis-2,3,3a,4,5,9b-hexahydro-1-methyl-1H-pyrrolo-[3,2-h]isoquinoline, a bridged-nicotine analog. J Pharmacol Exp Ther 1997;282(3):1425–34.
[30] Basil B, Coffee ECJ, Gell DL, Maxwell DR, Sheffield DJ, Wooldridge KRH. A new class of sympathic β-receptor blocking agents. 3,4-Dihydro-3-hydroxy-1,5-benzoxazocines. J Med Chem 1970;13:403–6.
[31] Evans JM, Fake CS, Hamilton TC, Poyser RH, Watts EA. Synthesis and antihypertensive activity of substituted trans-4-amino-3,4-dihydro-2,2-dimethyl-2H-1-benzopyran-3-ols. J Med Chem 1983;26:1582–9.
[32] Stemp G, Evans JM. Discovery and development of cromokalim and related potassium channel activators. In: Ganellin CR, Roberts SM, editors. Med chem. London: Academic Press; 1993. p. 141–62.
[33] Heel RC, Brogden RN, Pakes GE, Speight TM, Avery GS. Nefopam: a review of its pharmacological properties and therapeutic efficacy. Drugs 1980;19:249–67.
[34] Glaser R, Donnel D. Stereoisomer differentiation for the analgesic drug nefopam hydrochloride using modeling studies of serotonin uptake area. J Pharm Sci 1989;78:87–90.
[35] Murineddu G, Ruiu S, Loriga G, Manca I, Lazzari P, Reali R, et al. Tricyclic pyrazoles. 3. Synthesis, biological evaluation, and molecular modeling of analogues of the cannabinoid antagonist 8-chloro-1-(29,49-dichlorophenyl)-N-piperidin-1-yl-1,4,5,6-tetrahydrobenzo [6,7] cyclohepta[1,2-c]pyrazole-3-carboxamide. J Med Chem 2005;48(23):7351–62.
[36] Antel J, Gregory PC, Nordheim U. CB1 cannabinoid receptor antagonists for treatment of obesity and prevention of comorbid metabolic disorders. J Med Chem 2006;49(14):4008–16.
[37] Carpino PA, Griffith DA, Sakya S, Dow RL, Black SC, Hadcock JR, et al. New bicyclic cannabinoid receptor-1 (CB1-R) antagonists. Bioorg Med Chem Lett 2006;16(3):731–6.
[38] Smith RA, Fathi Z, Brown SE, Choi S, Fan J, Jenkins S, et al. Constrained analogs of CB-1 antagonists: 1,5,6,7-Tetrahydro-4H-pyrrolo[3,2-c]pyridine-4-one derivatives. Bioorg Med Chem Lett 2007;17(3):673–8.
[39] 3D structures were generated and minimized using DS viewerPro 6.0, © 2005 by Accelrys Software, Inc.
[40] Pfister R, Sallmann A, Hammerschmidt W. Substituted malonic acid hydrazides. US Patent 3,455,999 (July 15, 1969; Geigy Chemical Corporation).

[41] Mixich G. Zum chemischen Verhalten des Antiphlogisticums "Azapropazon" (Mi 85) 5 3-dimethylamino-7-methyl-1,2-(n-propylmalonyl)-1,2-dihydro-1,2,4-benzotriazin [Concerning the chemical behavior of the antiphlogistic "Azapropazone" (Mi 85) 5 3-dimethylamino-7-methyl-1,2-(n-propylmalonyl)-1,2-dihydro-1,2,4-benzotriazin]. Helv Chim Acta 1968;51:532–8. [in German].
[42] Jahn U, Wagner-Jauregg T. Vergleich von Zwei neuen Klassen Antiphlogisticher Substanzen im Collier-Test [A comparison of two new classes of antiphlogic substances in a Collier test]. Arzneim-Forsch 1968;18:120–1. [in German].
[43] Yamamoto H, Kaneko S-I. Synthesis of 1-phenyl-2-styryl-3,5-dioxopyrazolidines as antiinflammatory agents. J Med Chem 1970;13:292–5.
[44] Wermuth CG, Choay J. Nouveaux dérivés de la pyrazolidine, leur procédé de fabrication et médicaments contenant ces nouveaux dérivés [New derivatives of pyrazolidine, their method of manufacture, and pharmaceuticals containing these new derivatives]. Fr Patent 2,244,513 (July 12, 1973;Choay SA) [in French].
[45] Teger-Nilsson AC, Bylund R, Gustafsson D, Gyzander E, Eriksson U. In vitro effects of inogatran, a selective low molecular weight thrombine inhibitor. Thromb Res 1997;85:133.
[46] Gustafsson D, Antonsson T, Bylund R, Eriksson U, Gyzander E, Nilsson I, et al. Effects of melagatran, a low-molecular-weight thrombin inhibitor, on thrombin and fibrinolytic enzymes. Thromb Haemostasis 1998;79:110.
[47] Pelletier JC, Hesson D, Jones KA, Costa AM. Substituted 1,2-dihydrophtalazines: potent, selective, and noncompetitive inhibitors of the AMPA receptor. J Med Chem 1996;39:343–6.
[48] Grasso S, De Sarro G, De Sarro A, Micale N, Zappala M, Puja G, et al. Synthesis and anticonvulsant activity of novel and potent 6,7-methylenedioxyphtalazin-1 (2H)-ones. J Med Chem 2000;43:2851–9.
[49] Trybulski EJ, Zhang J, Kramss RH, Mangano RM. The synthesis and biochemical pharmacology of enantiomerically pure methylated oxotremorine derivatives. J Med Chem 1993;36:3533–41.
[50] Ringdahl B, Jenden DJ. Pharmacological properties of oxotremorine and its analogues. Life Sci 1983;32:2401–13.
[51] Klimko PG, Davis TL, Griffin BW, Sharif NA. Synthesis and biological activity of a novel 11a-homo (cyclohexyl) prostaglandin. J Med Chem 2000;43:3400–7.
[52] Mull RP, Maxwell RA. Guanethidine and related adrenergic neuronal blocking agents. In: Schlittler E, editor. Antihypertensive agents. New York: Academic Press; 1967. p. 115–49.
[53] Giudicelli R, Najer H, Lefèvre F. Comparaison des durées d'action du N-β-guanidino-éthyl-aza-6 spiro[2,5] octane (LD 3598) et de la guanéthidine [A comparison of the duration of the action of N-β-guanidino-ethyl-aza-6 spiro[2,5] octane (LD 3598) and guanethidine]. Compt Rend Acad Sci 1965;260:726–9. [in French].
[54] Najer H, Giudicelli R, Sette J. Guanidines douées d'action anti-hypertensive, 4e mémoire: N-β-guanidinoéthyl azaspiro alcanes (1ère partie) [Guanidines endowed with anti-hypertensive action, 4th report: N-β-guanidinoethyl azaspiro alkane (Part 1)]. Bull Soc Chim Fr 1964;2572–81. [in French].
[55] Toda N, Usui H, Shimamoto K. Modification by AZ-55, guanethidine and bretylium of responses of atria and aortic strips to transmural stimulation. Jpn J Pharmacol 1972;22:125–35.
[56] Najer H, Giudicelli R, Sette J. Guanidines douées d'activité antihypertensive, 3e mémoire: N-β-guanidinoéthyl azabicyclo alcanes [Guanidines endowed with anti-hypertensive activity, 3rd report: N-β-guanidinoethyl azabicyclo alkanes]. Bull Soc Chim Fr 1962;1593–7. [in French].
[57] Anonymous. New basic derivatives of 6,9-endoxo-3-azabicyclo-[4.3.0]-nonane. Br. Patent 973,533 (February 21, 1962; Laboratoire Lumière, SA).
[58] Anonymous. New basic derivatives of 6,9-endomethylene-3-azabicyclo-[4.3.0]-nonane. Br. Patent 972,088 (February 19, 1962; Laboratoire Lumière, SA).
[59] Bryans JS, Morrell AI. Novel stereoselective processes for preparation of gabapentin analogues W09914184 (1999; Warner-Lambert Co.).
[60] Horwell DC, Bryans JS, Kneen CO, Morrell AI, Ratcliffe GS. Preparation of novel bridge cyclic amino acids as pharmaceutical agents W09733859 (1997; Warner-Lambert Co.).
[61] Meyer MD, Altenbach RJ, Basha FZ, Carroll WA, Condon S, Elmore SW, et al. Structure–activity studies for a novel series of tricyclic substituted hexahydrobe-nz[e]isoindole alpha(1A) adrenoceptor antagonists as potential agents for the symptomatic treatment of benign prostatic hyperplasia (BPH). J Med Chem 2000;43(8):1586–603.
[62] Jongebreur G. Relation between the chemical constitution and the pharmacological action, especially on the coronary vessels of the heart, of some synthesized pyrones and khellin. Arch Intern Pharmacodynamie 1952;90:384–411.
[63] Evans BE, Rittle KE, Bock MG, DiPardo RM, Freidinger RM, Whitter WL, et al. Methods for drug discovery: development of potent, selective, orally effective cholecystokin antagonists. J Med Chem 1988;31:2235–46.
[64] Janssen PAJ, Van der Eycken CAM. The chemical anatomy of potent morphine-like analgesics. In: Burger A, editor. Drugs affecting the central nervous system. New York: Marcel Dekker, Inc.; 1968. p. 25–60.
[65] Hjorth S, Carlsson A, Wikström H, Lindberg P, Sanchez D, Hacksell U, et al. 3-PPP, a new centrally acting DA-receptor agonist with selectivity for auto-receptors. Life Sci 1981;28:1225–38.
[66] Kelly TR, Howard HR, Koe K, Sarges R. Synthesis and dopamine autoreceptor activity of a 5-(methylmercapto)methyl-substituted derivative of (6)-3-PPP (3-(3-hydroxyphenyl)-1-n-propylpiperidine). J Med Chem 1985;28:1368–71.
[67] Bach NJ, Kornfeld EC, Jones ND, Chaney MO, Dorman DE, Paschal JW, et al. Bicyclic and tricyclic ergoline partial structures. Rigid 3-(2-aminoethyl)pyrazoles as dopamine agonists. J Med Chem 1980;23:481–91.
[68] Fuller RW, Clemens JA, Kornfeld EC, Snoddy HD, Smalstig EB, Bach NJ. Effects of (8β)-8-[(methylthio)methyl]-6-propylergoline on dopaminergic function and brain dopamine turnover in rats. Life Sci 1979;24:375–82.
[69] Chang RSL, Lotti VY. Biochemical and pharmacological characterization of an extremely potent and selective nonpeptide CCK antagonist. Proc Natl Acad Sci U S A 1986;83:4923–6.
[70] Bock MG, DiPardo RM, Rittle KE, Evans BE, Freidinger RM, Veber DF, et al. Cholecystokinin antagonists. Synthesis of asperlicin analogues with improved potency and water solubility. J Med Chem 1986;29:1941–5.
[71] Bruderlein FT, Humber LG, Voith K. Neuroleptic agents of the benzocycloheptapyridoisoquinoline series. 1. Syntheses and stereochemical and structural requirements for activity of butaclamol and related compounds. J Med Chem 1975;18:185–91.

[72] Imhof R, Kyburz E, Daly JJ. Design, synthesis, and X-ray data of novel potential antipsychotic agents. Substituted 7-phenylquinolizidines: stereospecific, neuroleptic, and antinociceptive properties. J Med Chem 1984;27:165–75.
[73] Evans RH, Francis AA, Jones AW, Smith DAS, Watkins JC. The effects of a series of ω-phosphonic α-carboxylic amino acids on electrically evoked and excitant amino acid-induced responses in isolated spinal cord preparations. Br J Pharmacol 1982;75:65–75.
[74] Hutchinson AJ, Williams M, Angst C, de Jesus R, Blanchard L, Jackson RH, et al. 4-(Phosphonoalkyl)- and 4-(phosphonoalkenyl)-2-piperidinecarboxylic acids: synthesis, activity at N-methyl-d-aspartic acid receptors, and anticonvulsant activity. J Med Chem 1989;32:2171–8.
[75] Humblet C, Johnson G, Malone T, Ortwine DF. Design, synthesis, and molecular modeling of phosphonoalkyl-substituted tetra-hydroisoquinolines as competitive NMDA antagonists. In: Proceedings of the XIth international symposium on medicinal chemistry 1990; p. 26.
[76] Ornstein PL, Arnold MB, Augenstein NK, Lodge D, Leander JD, Schoepp DD. (3SR,4aRS,6RS,8aRS)-6-[2-(1H-tetrazol-5-yl)ethyl]decahydroisoquinoline-3-carboxylic acid: a structurally novel, systematically active, competitive AMPA receptor antagonist. J Med Chem 1993;36:2046–8.
[77] Koga H, Itoh A, Murayama S, Suzue S, Irikura T. Structure–activity relationships of antibacterial 6,7- and 7,8-disubstituted 1-alkyl-1,4-dihydro-4-oxoquinoline-3-carboxylic acids. J Med Chem 1980;23:1358–63.
[78] Jinbo Y, Taguchi M, Inoue Y, Kondo H, Miyasaka T, Tsujishita H, et al. Synthesis and antibacterial activity of a new series of tetracyclic pyridone carboxylic acids. 2. J Med Chem 1993;36:3148–53.
[79] Wijngaarde van I, Hamminga D, Hes Rv, Standaar PJ, Tipker J, Tulp MTM, et al. Development of high-affinity 5-HT3 receptor antagonists. Structure-affinity relationships of novel 1,7-annelated indole derivatives. 1. J Med Chem 1993;36:3693–9.
[80] Faust R, Garrat PJ, Yeh L-K, Tsotinis A, Panoussopoulou M, Calogeropoulou T, et al. Mapping the melatonin receptor. 6. Melatoonin agonists and antagonists derived from 6H-isoindolo[2,1-a]indoles, 5,6-dihydroindolo[2,1-a]isoquinolines, and 6,7-dihydro-5H-benzo[c]azepino[2,1-a]indoles. J Med Chem 2000;43:1050–61.
[81] Stanton JL, Sperbeck DM, Trapani AJ, Cote D, Sakane Y, Berry CJ, et al. Heterocyclic lactam derivatives as dual angiotensin converting enzyme and neutral endopeptidase 24.11 inhibitors. J Med Chem 1993;36:3829–33.

CHAPTER

10

Macrocycles: Under-Explored and Poorly Exploited Drug Class Despite the Proven Therapeutic Potential

Pierre Raboisson

Janssen Infectious Diseases and Diagnostics BVBA, Beerse, Belgium

OUTLINE

There is something of the marvelous in all things of nature. ***Aristotle***

I. NATURE AS A SOURCE OF MACROCYCLES

Macrocyclic drugs are almost exclusively originate from natural sources, primarily microorganisms [1,2]. Approximately 3 percent of natural products are macrocycles. Despite this modest representation, they are often involved in key biological functions, like the stimulation of growth, formation of myceliae and spores, cell−cell communication via autocrine hormones and mediators, and chemical warfare between micro-organisms competing for scarce resources [1,2]. In the arsenal of drugs, macrocycles have been used predominantly as anti-infectives. Representative examples are highlighted in Figure 10.1. Important antibiotics such as erythromycin (**1**), vancomycin (**2**), josamycin (**3**), and rifampicin (**4**), and anti-fungals such as amphotericin B (**5**) belong to the natural macrocyclic pharmacopeia.

Macrocyclic derivatives like epothilone B (**6**) and rapamycin (**7**) have been used in oncology, and others like cyclosporine A (**8**) and tacrolimus (**9**) have found therapeutic use in modulating the immune-system responses (Figure 10.2).

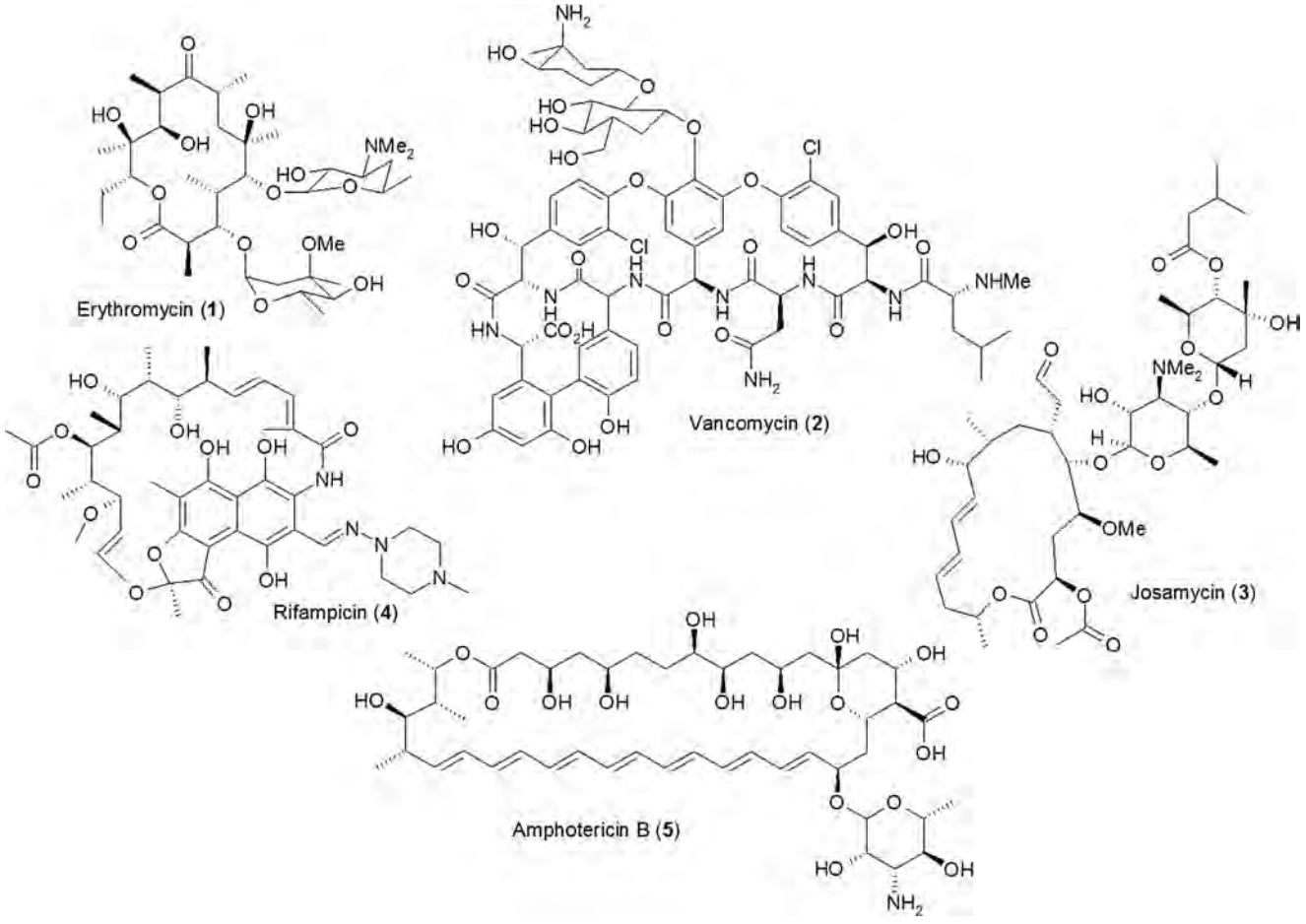


FIGURE 10.1 Example of natural macrocyclic anti-infective drugs.

Despite their unique characteristics, macrocycles have been under-explored by medicinal chemists. They are often seen as too complex to optimize against a given target and difficult to scale up while maintaining a reasonable price for the active pharmaceutical ingredient (API). Thanks to the evolution of synthetic macrocyclization procedures (e.g., olefin metathesis), synthetic macrocyclic molecules have recently been revisited and are stimulating growing interest in modern drug discovery approaches.

As an illustrative example, the fully synthetic complex macrocycle Simeprevir (**10**; Figure 10.3), [3] which breaks almost all the "rule of 5" criteria, has recently been approved for the treatment of chronic hepatitis C (HCV).

II. IDENTIFICATION OF MACROCYCLIC DRUGS USING EITHER PHENOTYPIC SCREEN OR TARGET-BASED APPROACH

Interestingly, in an era in which the major focus of drug discovery is on target-based approaches, the phenotypic screen remains the primary approach to identifying first-in-class drugs, as illustrated in Figure 10.4 [4]. Indeed, one disadvantage of the target-based approach is that the strategy deployed to target specific molecular hypotheses may be too artificial and may not translate to the disease pathogenesis. Another one is the lack of therapeutic window often observed in the clinic, with small-molecule drug candidates targeting host proteins. Thus, even if the optimization of desired molecular properties might be more challenging for compound series identified by phenotypic screening, because of the lack of structure-guided design, the activities observed in cell-lines are often translated into therapeutic impact in a given disease. This is especially true for infectious diseases for which the activities measured *in vitro* in cellular screening often translate to a high efficacy *in vivo*. Although target-based approaches mainly use the "associative approach" (i.e., start with small molecules or fragments and

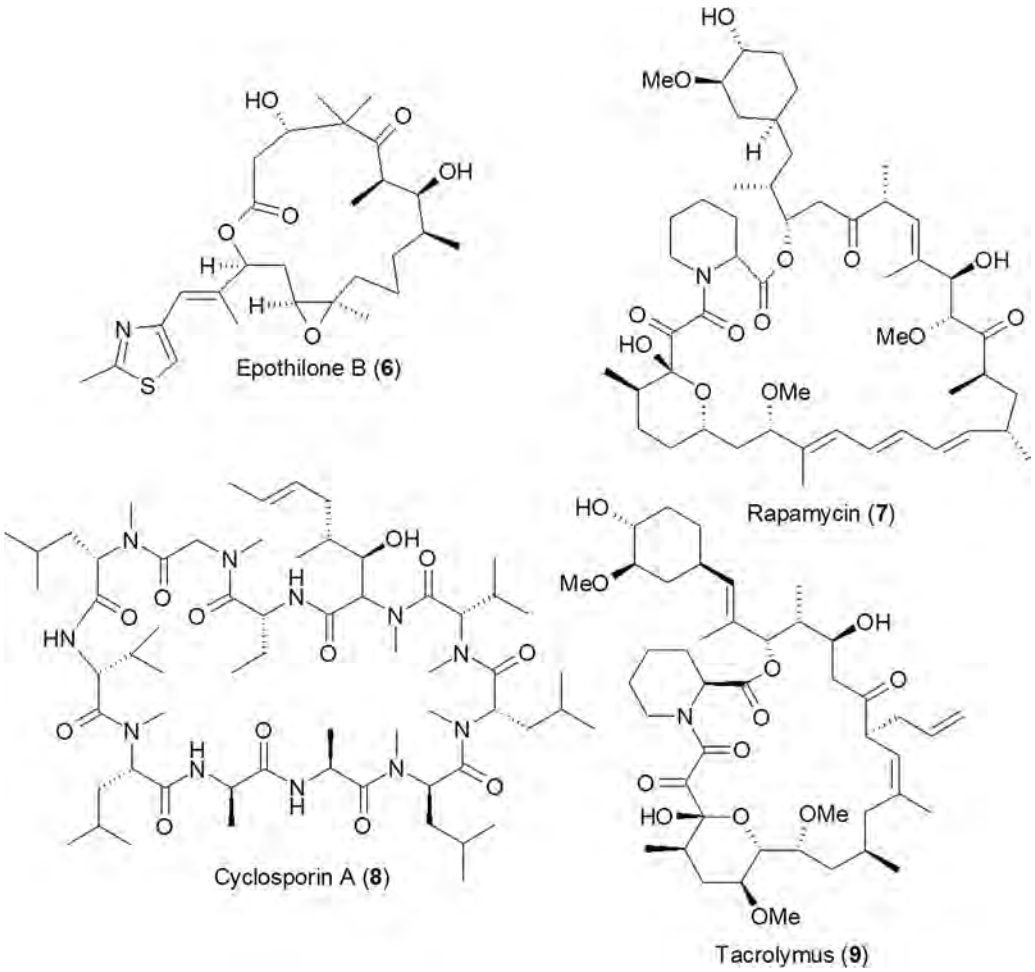


FIGURE 10.2 Examples of drugs modulating oncology or immunology targets.

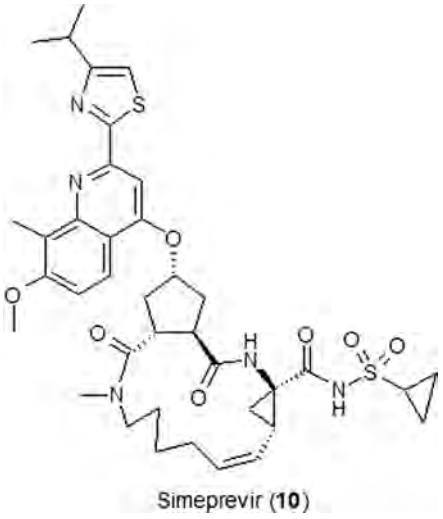


FIGURE 10.3 Simeprevir (Olysio®) [3].

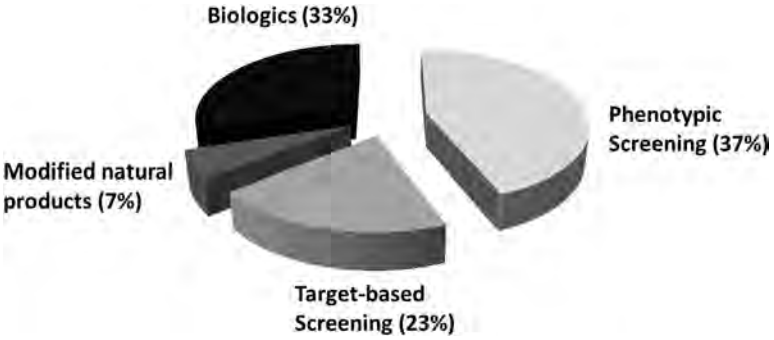


FIGURE 10.4 The distribution of first-in-class drugs discovered between 1999 and 2008 [4].

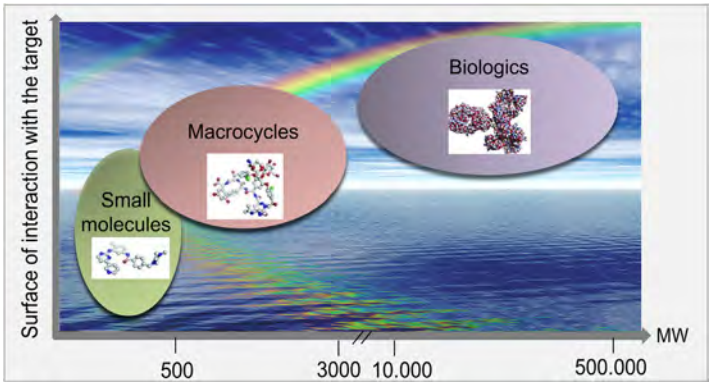


FIGURE 10.5 Middle space occupied by macrocycles.

then make them bigger), the phenotypic screen allows often a "dissociative approach," with successive truncations of undesired molecular groups and functions to reach the desired profile.

In this context, it can be hypothesized that the target-centric approach—together with the scarcity of larger molecules in diversity libraries and the reluctance of some scientists to assess molecules with a high degree of complexity—contributes to the attrition rate and low productivity in pharmaceutical R&D.

III. MACROCYCLES: THE DRUGS IN THE MIDDLE SPACE

The macrocyclic drugs cover a unique chemical space: they might be considered bigger versions of small molecules or as the smallest examples of biomolecules with functional subdomains (Figure 10.5) [1,5].

Due to the unique positioning of macrocyclic drugs between small and large drugs, they also display unique features. Macrocycles should be seen as pre-organized but not rigid molecules. Thanks to their larger size compared to small molecules, macrocycles can better interact with target proteins when a larger surface of contact is required to produce a pharmacological effect. These targets are often difficult to tackle with regular small molecules and are commonly modulated by biologics, for which drawbacks such as high cost and reduced patient compliance—linked to their use as injectables—have often limited their broad market penetration.

IV. EFFECT OF THE MACROCYCLIZATION ON DRUG-LIKE PROPERTIES

Despite their high molecular weights and polar surface areas, macrocycles can adopt different conformation in various media, thus retaining good cell permeability [5]. This characteristic allows some of them to cross cell membranes and interact with intracellular targets. To control vital functions, nature has used the "macrocyclization trick" to allow larger molecules to gain cell-permeability via a process so-called "selective evolution". [2]. This can be illustrated with cyclic peptides, which—despite their high molecular weight—can still permeate membranes via passive diffusion, even though the corresponding open versions generally totally lack

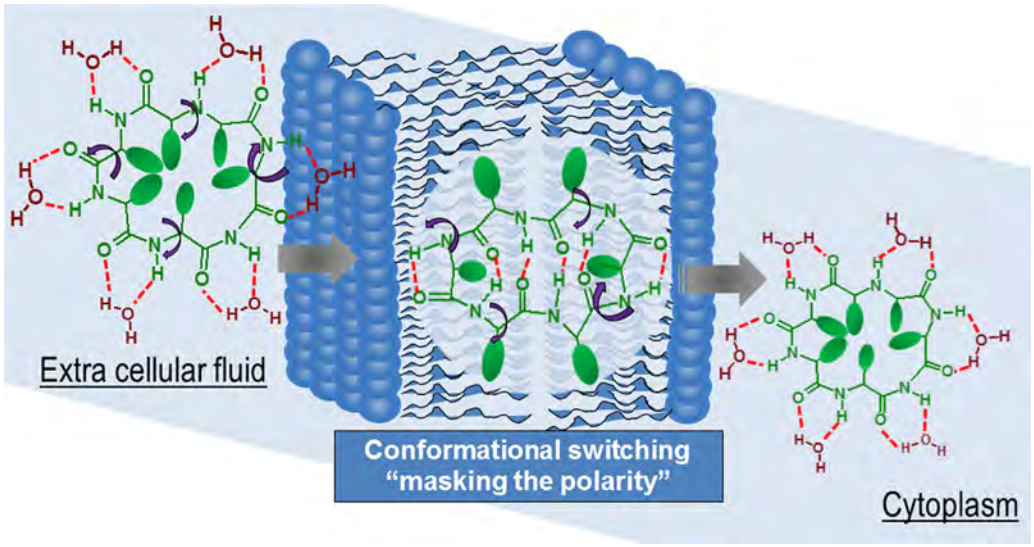


FIGURE 10.6 Flip-Flop conformational change allowing macrocyclic peptids to cross cellular membranes.

permeability [5]. This increased cell permeability of the macrocyclic peptides over linear counterparts is the result of several factors, such as the elimination of charge termini and the facilitated network of intra-molecular hydrogen bonds, which masks to a large extent the polarity of the macromolecule [5]. The latter is illustrated in Figure 10.6. When the cyclic peptide is dissolved in biological fluid, the polar groups of the molecule are turned outwards toward the biological fluid, resulting in higher water solubility. When the macrocycle enters the lipid bilayer of the membrane, the polarity of the molecule is masked by a rearrangement. The rearrangement establishes a series of intramolecular hydrogen bonds involving the backbone and the side chains, effectively masking the molecules' polarity and presenting lipophilic residues to the lipid bilayer of the membrane. Eventually, the cyclic peptide reaches the cytosol via the same "flip-flop" mechanism. As the permeation of the membrane is at least partially explained by some molecular conformational changes, the diffusion of macrocycles cannot be predicted by the regular calculated parameters often used to predict the permeability for small molecules (logP, MW, PSA…) and might even be dramatically different for some diastereomers [5]. Nevertheless, many macrocycles retain sufficient permeability to be screened in various cell-based phenotypic screens, and often exhibit desired pharmacokinetic profiles, including oral bioavailability. Indeed, it has been shown that oral bioavailability can be obtained for macrocycles at molecular weights above 1 kDa and polar surface areas ranging toward 250 $Å^2$ [6,7].

Another advantage offered by the macrocyclization is the increased metabolic stability compared to the corresponding open forms. This characteristic can be illustrated with cyclic peptides, which are known to be more stable toward proteolytic degradation.

V. INTERACTION OF MACROCYCLES WITH THEIR TARGETS

The macrocyclic drugs can interact with their target via a classical protein−small molecule interaction (e.g., enzyme inhibitors or receptor antagonists) or can establish more complex interactions with the target protein involving a large surface of contact, similar to the biologics [3]. This more complex macrocyclic-dependent gain-of-function for the complex drug-target can either take place at the surface of a single protein or at the interface between different proteins, which represents a unique opportunity to target protein−protein interactions [8].

A. Macrocycle-Dependent Gain-of-Function for the Complex Drug-Target

Another key feature of the macrocycles is the high degree of structural molecular pre-organization [2,8]. This allows the functional groups to interact with the target, often involving an extended binding site without major entropic penalty upon binding. This feature of the macrocycle makes them often very potent and selective for

their targets. The macrocycles are in the middle chemical space. Hence, they can interact either with binding pockets like the small molecules or with an extended surface of the target in the same way as biologicals [2,8]. Often, macrocycles can engage their target by creating new protein surface interactions, thereby inducing a macrocycle-dependent gain-of-function for the complex. This mechanism has been reported for various drugs like Epothilone B (**6**; Figure 10.2) and Erythromycin (**1**; Figure 10.1). Epothilone B, an approved drug for the treatment of metastatic breast cancer, has been shown to bind at the interface of two tubulin subunits in a region overlapping with the taxol binding site, leading to an extensive reconfiguration of the $\alpha-\beta$ interface, stabilizing the dimer and disturbing the overall microtubule dynamics [9]. The macrolide antibiotic erythromycin functions by binding to the inner surface of the ribosomal tunnel effectively to create a new surface, thereby physically impeding the exit of nascent peptides from the ribosome by narrowing the passage [10]. Due to their larger contact surface—which can often not be achieved with standard small molecules—macrocycles can unlock novel targets that have been typically difficult to address with standard small molecules.

Other macrocyclic drugs can modulate some protein−protein interactions. This is illustrated with the immunosuppressant agents cyclosporine A (**8**; Figure 10.2) and sanglifehrins, which form ternary complexes with cyclophilin A and calcineurin (Figure 10.7). The complex leads to the inhibition of calcineurin Serine / Threonine phosphatase activity, the release of dephosphorylated Nuclear Factor of Activated T cell (NFAT), and the downstream inhibition of IL2 release, which leads to the silencing of the immune system via the blocking of T-cell activation [11]. Interestingly, cyclosporine A was shown to inhibit different virus replication (HIV, HCV, HBV) using phenotypic screening.

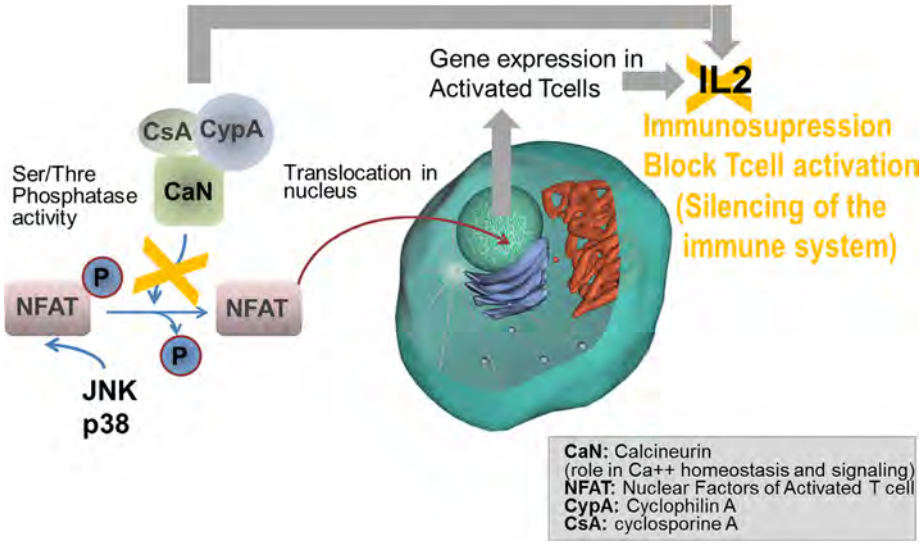


FIGURE 10.7 Immunosuppressive function of the cyclophilin A-calcineurin-cyclosporine A ternary complex.

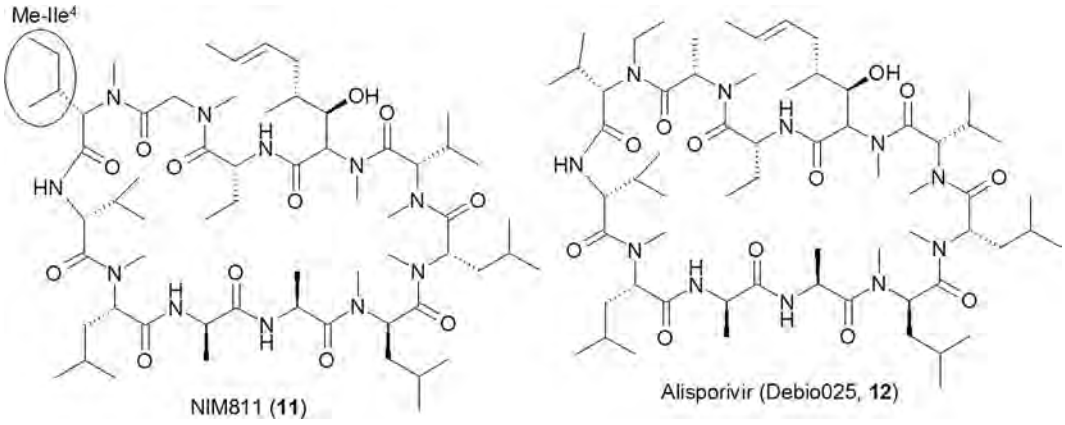


FIGURE 10.8 Anti-Hepatitis C drugs derived from cyclosporin A [4,12].

Different pharmaceutical groups have modified the structure of CypA, leading to the very potent Hepatitis C Virus (HCV) inhibitors NIM-811 (**11**; Figure 10.8) [4] and Alisporivir (Debio-025, **12**) [12]. In the case of Alisporivir, a single *N*-methyl-Leucine to *N*-ethyl-Valine amino-acid replacement delivered a compound with a totally different profile. Debio-025 is ten-fold more potent as a HCV inhibitor in the replicon cell-based assay versus Cyclosporine and lacks the immunosuppressive character [12]. This lack of immunosuppressive character could be rationalized by a steric clash between the *N*-ethyl-Valine of Debio-025 and the Tryptophan 352 of the calcineurin. Debio-025 has been reported in late-stage clinical trials for patients chronically infected with HCV. This mechanism tackling HCV replication is associated with an advantageous virological resistance profile over other direct-acting antivirals (DAAs) [12].

B. Macrocyclization as a Strategy to Enhance Pharmacokinetic Parameters

Other macrocycles have been reported to act as "classical" enzyme inhibitors. In the CNS field, the P1-P3 cyclisation of **13** (Figure 10.9; $IC_{50} > 100\ \mu M$) afforded the much more active derivative **14** ($IC_{50} = 2.9\ \mu M$) as a result of a lower entropic penalty upon binding to BACE1 [13]. Further lead optimization afforded **15**, which combined a high potency on BACE1 ($IC_{50} = 4$ nM), a high permeability despite a relatively high molecular weight and PSA, and a limited susceptibility toward PgP efflux mechanism (**13** suffers from that issue). Moreover, **15** was shown to permeate the brain and demonstrated some efficacy on the APP-YAC transgenic mouse Alzheimer's disease model with a reduction of the amyloid precursor protein fragment levels ABeta40 [13].

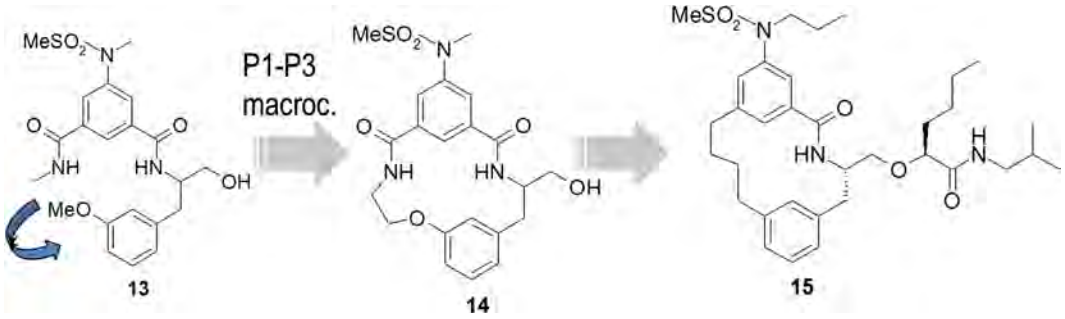


FIGURE 10.9 BACE-1 inhibitors [13].

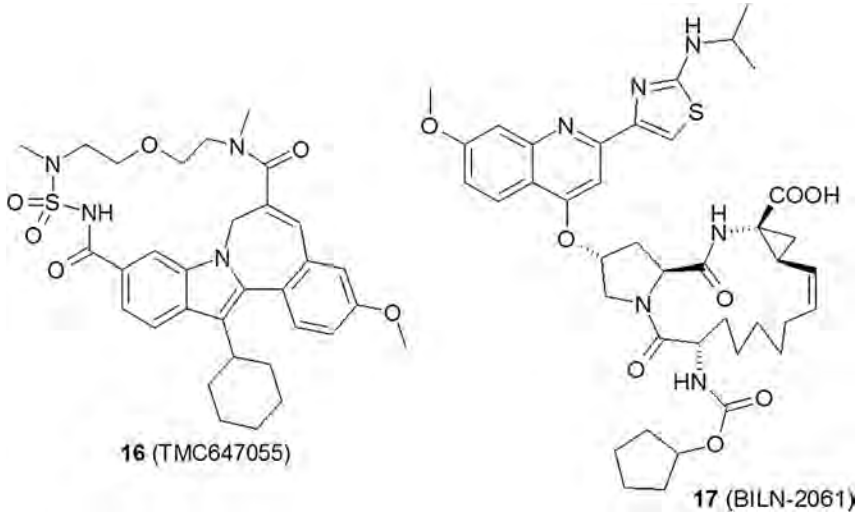


FIGURE 10.10 Hepatitis C virus inhibitors [14,15].

A recent success story for the synthetic macrocycles has been reported in the field of HCV protease inhibitors. Despite the large screening campaigns initiated by many pharmaceutical groups on the HCV NS3/4A protease, no valuable hits have been identified. However, researchers at Boehringer Ingelheim demonstrated that the end cleavage product of the NS3/4A was a weak inhibitor of the HCV protease. After an elegant lead-optimization effort, the macrocycle BILN-2061 (**17**; Figure 10.10) was identified as a potent, orally bioavailable HCV inhibitor [14]. Despite a very potent activity observed in the clinic, this asset was stopped in Ph1 due to the emergence of cardio safety alert in cynomologus monkeys. Multiple companies initiated some follow-on drug discovery programs to identify novel Hepatitis C NS3/4A inhibitors, with the rationale that the cardiosafety alert observed with BILN-2061 was compound specific. This culminated with the discovery of Simeprevir (Olysio®; Figure 10.3), [3] a potent one-pill once-a-day treatment for the HCV infection.

Another example of a macrocyclic HCV drug candidate has been reported with the discovery of TMC647055 (**16**; Figure 10.10), a potent HCV NS5B inhibitor [15]. In this case, the macrocycle was introduced in the molecule to improve the water solubility and PK characteristics of the very lipophilic 3-cyclohexylindole scaffold, resulting in a compound with better potency over the corresponding open forms and with excellent oral bioavailability. Furthermore, it was shown that the sulfone in the macrocyclic ring was making an additional key hydrogen bond in the NS5B binding domain.

VI. SYNTHESIS OF MACROCYCLES & LIBRARY ENRICHMENT

Despite this recent growing interest in synthetic macrocycles, the class remains poorly represented in most companies' libraries [6]. Macrocycles are often seen as too complex for parallel production of libraries. Library enrichment criteria defined by many companies (e.g., rule-of-five compliance, low price per compound) have left little room for macrocycle acquisition. Thus, the pharmaceutical industry should undertake the necessary effort to bring this additional structural diversity into their compound collections. Growing capacity to generate these molecules using novel macrocyclization procedures has been largely documented elsewhere [2,6,16,17]. Solid phase synthesis was, for instance, used to produce >10,000 cpd libraries [18].

VII. CONCLUSION

Macrocyclic ring structures cover a unique chemical space. They represent an opportunity to unlock novel—often complex—targets that are difficult to address with classical small molecules. Often, macrocyclic derivatives retain sufficient permeability to allow their use in phenotypic screens, by far the most powerful path toward the identification of first-in-class drugs. Although almost all macrocyclic drugs originated from natural products, recent success stories reported with synthetic macrocycles coupled with the growing capacity to generate these molecules using novel macrocyclization procedures should encourage scientists to enter this largely unexplored field of medicinal chemistry and extend their collection of larger molecule libraries. Together with other molecules belonging to the "parallel universe"(e.g., nucleosides/nucleotides, natural products), it is likely that the macrocycle will deliver attractive new drugs addressing important unmet medical needs.

References

[1] Wessjohann LA, Ruijter E, Garcia-Rivera D, Brandt W. What can a chemist learn from nature's macrocycles? A brief, conceptual view. Mol Divers 2005;9:171–86.
[2] Driggers EM, Hale SP, Lee J, Terrett NK. The exploration of macrocycles for drug discovery: an underexploited structural class. Nat Rev 2008;7:608–24.
[3] Raboisson P, de Kock H, Rosenquist Å, Nilsson M, Salvador-Oden L, Lin T-I, et al. Structure–activity relationship study on a novel series of cyclopentane-containing macrocyclic inhibitors of the Hepatitis C virus NS3/4A protease leading to the discovery of TMC435350. Bioorg Med Chem Lett 2008;18:4853–8.
[4] Weidmann B, Ryder NS, Cooreman MP, Lin K. NIM811, a cyclophilin inhibitor, exhibits potent in vitro activity against Hepatitis C virus alone or in combination with alpha interferon. Antimicrob Agents Chemother 2006;50:2976–82.
[5] Rezai T, Yu B, Millhauser GL, Jacobson MP, Lokey RS. Testing the conformational hypothesis of passive membrane permeability using synthetic cyclic peptide diastereomers. J Am Chem Soc 2006;128:2510–11.
[6] Marsault E, Peterson ML. Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery. J Med Chem 2011;54:1961–2004.

[7] Giordanetto F, Kihlberg J. Macrocyclic drugs and clinical candidates: what can medicinal chemists learn from their properties? J Med Chem 2014;57:278–95.
[8] Villar EA, Beglov D, Chennamadhavuni S, Porco JA, Kozakov D, Vajda S, et al. How proteins bind macrocycles. Nat Chem Biol 2014;10:723–32.
[9] Coderch C, Klett J, Morreale A, Diaz JF, Gago F. Comparative binding energy (COMBINE) analysis supports a proposal for the binding mode of epothilones to b-tubulin. Chem Med Chem 2012;7:836–43.
[10] Wilson DN. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nat Rev Microb 2014;12:35–48.
[11] Lei Jin L, Harrison SC. Crystal structure of human calcineurin complexed with cyclosporin A and human cyclophilin. PNAS 2002;99:13522–6.
[12] Coelmont L, Kaptein S, Paeshuyse J, Vliegen I, Dumont J-M, Vuagniaux G, et al. Debio 025, a cyclophilin binding molecule, is highly efficient in clearing hepatitis C virus (HCV) replicon-containing cells when used alone or in combination with specifically targeted antiviral therapy for HCV (STAT-C) inhibitors. Antimicrob Agents Chemother 2009;53:967–75.
[13] Stachel SJ, Coburn CA, Sankaranarayanan S, Price EA, Wu G, Crouthamel M, et al. Macrocyclic inhibitors of β-secretase: functional activity in an animal model. J Med Chem 2006;49:6147–50.
[14] Llinas-Brunet M, Bailey M, Bolger G, Brochu C, Faucher A-M, Ferland J-M, et al. Structure-activity study on a novel series of macrocyclic inhibitors of the hepatitis C virus NS3 protease leading to the discovery of BILN 2061. J MedChem 2004;47:1605.
[15] Vendeville S, Lin T-I, Hu L, Tahri A, McGowan D, Cummings MD, et al. Finger loop inhibitors of the HCV NS5b polymerase. Part II. Optimization of tetracyclic indole-based macrocycle leading to the discovery of TMC647055. Bioorg Med Chem Lett 2012;22:4437–43.
[16] Wessjohann LA, Rivera DG, Vercillo OE. Multiple multicomponent macrocyclizations (MiBs): a strategic development toward macrocycle diversity. Chem Rev 2009;109:796–814.
[17] White CJ, Yudin AK. Contemporary strategies for peptide macrocyclization. Nat Chem 2011;3:509–24.
[18] Marsault E. Discovery of a new class of macrocyclic antagonists to the human motilin receptor. J Med Chem 2006;49:7190–7.

SECTION THREE

Primary Exploration of Structure-Activity Relationships

CHAPTER

11

Conformational Restriction and Steric Hindrance in Medicinal Chemistry

Peter Wipf[1], *Erin M. Skoda*[1] *and André Mann*[2]

[1]Department of Chemistry, University of Pittsburgh, Pittsburgh PA, USA; [2]LAAS-CNRS, Université Louis Pasteur, Strasbourg, France

OUTLINE

I. INTRODUCTION

For the practice of medicinal chemistry, an understanding of the thermodynamics of molecular recognition processes is essential. Noncovalent interactions between a ligand and a protein are mediated by several parameters. Ion–ion, dipole–dipole, and ion-dipole forces, hydrophobicity, hydrogen bonding, and shape complementarity play significant roles. The magnitude of the binding equilibrium is determined by the thermodynamic contributions of each of these interactions. When two molecules bind, their binding energy must compensate for their desolvation— in most cases the removal of all or part of their water shell. The probability that a ligand is trapped in a receptor is dependent on the promotion of binding by favorable interactions of polar functional groups and the formation of hydrophobic contacts. In addition to desolvation, adverse factors for binding are mainly related to the restriction of translational and vibrational energies upon complex formation [1–4]. While the ligand generally loses conformational freedom, the entropy of the solvent increases. In order to overcome entropy penalties and improve the binding affinity of a ligand, a suitable strategy is to restrict its conformational flexibility. Therefore, once a lead compound has been identified for a targeted biological receptor, the optimization toward potency and/or selectivity usually involves the modification of its conformation. Depending on the chemical structure of the identified lead, conformational restriction and/or steric hindrance can be introduced by adding or removing key bonds and substituents. An optimization of the free energy released during the association of a ligand with a receptor can be achieved by modifying the spatial disposition of functional groups through steric strain, hybridization, or cyclization. The expected benefits are receptor selectivity, potency increase, pharmacophore optimization, and metabolic stabilization. As a bonus, this strategy can be expected to produce original compounds and innovative chemistry. Finally, there are also some risks: any structural change

in a compound alters more than one property and can have unexpected consequences. Indeed, this approach is valuable in cases where the lead compounds are exhibiting flexibility and where the low-energy conformations are not representative for a good fit to the receptor. In this chapter, some theoretical points related to the thermodynamic aspects of lead optimization will be addressed, and then several examples will be presented where the modification of rigidity has had an impact in drug design.

A. Theoretical Aspects of Ligand Binding

Molecular recognition is a key factor in all biological processes. Accordingly, many of the important issues in medicinal chemistry hinge on an understanding of the noncovalent interactions between a biomacromolecule (receptor) and a small ligand (drug). Under equilibrium conditions, the binding affinity can be expressed in terms of the difference in free energy (ΔG°) of free and bound states or the equilibrium constant (K), which are correlated by the Gibbs-van't Hoff equation (Eq. 11.1):

$$\Delta G^\circ = -RT \ln K \quad \{\text{with } \Delta G^\circ = \Delta H^\circ - T\Delta S^\circ; \mathrm{R} = 8.13\ \mathrm{J/mol/K}\} \tag{11.1}$$

The more negative ΔG°, the stronger is the binding of the ligand to the bioreceptor (the binding constant is large). As a reminder, if a noncovalent association is expected, ΔG° must be negative and the Gibbs equation usually contains a negative value for ΔH° and a positive value for $T\Delta S^\circ$. If $\Delta G^\circ = -30$ kJ/mol, the association constant is in the micromolar range. By using the Gibbs equation, we can distinguish between the effects of functional group interactions within the complex (ΔH°) and changes in the entropy upon binding ($T\Delta S^\circ$). The values of the standard free enthalpy ΔH° can serve as a quantitative indicator of the changes in intermolecular bond energies (mainly hydrogen bonding and van der Waals interactions) occurring during the binding, while the standard free entropy ΔS° is a good indicator of the reorganization in the solvation shell and the conformational changes during the same process. Only a negative ΔG° favors the equilibrium of the association. While in the formation of a covalent bond, ΔH° usually has the major contribution to ΔG°, in the case of a noncovalent interaction or equilibrium, the contributions by ΔH° and $T\Delta S^\circ$ to the free energy are often comparable. Given the large bond strengths of 300 to 500 kJ/mol for covalent bonds and the prevalence of the enthalpy term ($\Delta H^\circ >> T\Delta S^\circ$), the much smaller energy of 1 to 60 kJ/mol released in a reversible interaction is composed of roughly equivalent enthalpy and entropy terms ($\Delta H^\circ \cong T\Delta S^\circ$) [5–8] (Figure 11.1).

The binding energy values for noncovalent associations are small and comparable to the solvation energies that oppose them in the binding event in aqueous media. Binding constants have been estimated by taking advantage of the Gibbs equation. Böhm used a set of interactions arising from experimentally determined binding constants resulting from the association of small organic molecules with proteins [10]. Five types of ΔG° contributions were identified (Table 11.1) and their values were extracted from a regression analysis. These values are generally used for a semi-quantitative evaluation of binding constants and can be applied in ligand optimization.

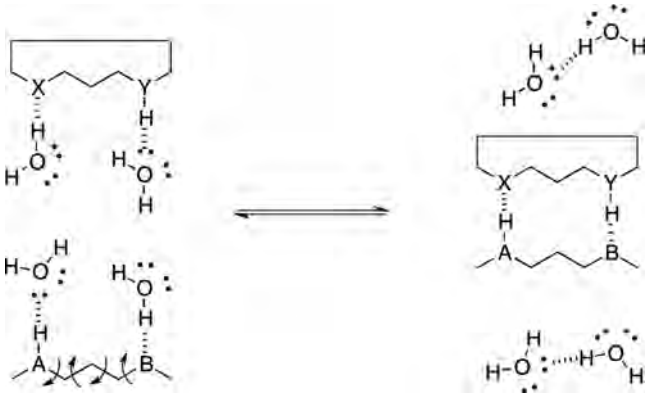


FIGURE 11.1 Noncovalent association between a ligand with two polar atoms (A and B) linked by a 3-carbon chain with a receptor containing complementary functions (X and Y), with the dashed line indicating hydrogen bonding [9].

TABLE 11.1 Average Values for ΔG° Contributions [9,10]

Physical process	ΔG° (kJ/mol)
Energy cost of bimolecular association	+5.4
Energy cost of restriction of an internal rotor	+2.0 (per rotor)
Hydrophobic effect (per Å of buried hydrocarbon)	−0.17/Å
Benefit of forming a neutral hydrogen bond of ideal geometry	−4.7
Benefit of forming an ionic hydrogen bond of ideal geometry	−8.3

1. Enthalpy–Entropy Compensation

A key factor that needs to be considered in the analysis of binding events is that a decrease in motion implies a decrease in entropy, since it results in fewer accessible arrangements. Consider the formation of a specific noncovalent bond (e.g., L···R for the transformation L + R → L···R). An increase in its strength (which corresponds to an additional negative contribution to ΔH°, favorable to the binding process) will be accompanied by an increasing restriction in the molecular motion of L and R in L···R (which corresponds to a decrease in ΔS°, unfavorable to the binding process). This opposing interplay between enthalpy and entropy is known as enthalpy–entropy compensation and represents a fundamental property of noncovalent interactions [11,12]. It arises because bonding opposes motion and, reciprocally, motion opposes bonding. The two effects can be compromised because the strength of noncovalent bonds at room temperature is comparable to kT, the product of the Boltzmann constant and the temperature. The thermodynamic parameters (ΔG°, ΔH°, ΔS°) for 136 ligands binding to ten biological receptors (mainly membrane receptors) have been analyzed [12]. It appeared that ΔH° and $T\Delta S^\circ$ values show a linear correlation (Eq. 11.2). This relationship seems remarkable in view of its high correlation and statistical relevance (n = 186). The observed enthalpy–entropy compensation seems to reflect general principles of the binding processes, irrespective of the nature of ligands and of their macromolecular targets:[12]

$$\Delta H^\circ = -40\ \text{kJ mol}^{-1} + 278\ \Delta S^\circ(\text{in kJ K}^{-1}\text{mol}^{-1}) \quad (n = 186, R^2 = 0.981, P \leq 0.001) \tag{11.2}$$

This linear correlation reflects another concept: any tightening of the intermolecular bonds (the enthalpic contribution) is compensated by a loss of degrees of freedom (the entropic contribution). The binding association can be enthalpy- or entropy-driven. The origin of the $\Delta H/\Delta S$ compensation is probably related to an intrinsic property of the hydrogen bond, which determines the association of the participants (water, drug, and binding site) in the drug-receptor binding equilibrium [11,12]. Accordingly, the thermodynamic laws teach us that three factors are involved in improving binding affinity: (1) improving ligand-protein interactions over those with the solvent in order to obtain a favorable (negative) enthalpy change; (2) increasing the hydrophobicity of the ligand in order to displace solvent molecules in the binding site and increase entropy; and (3) pre-shaping the ligand to complement the geometry of the binding site in order to minimize the loss of conformational entropy upon binding [13].

2. Enthalpy- Versus Entropy-Driven Binding

The following example is illustrative for the $\Delta H/\Delta S$ compensation. The binding of a series of agonists and antagonists of the β-adrenergic receptor displays thermodynamic differences between agonists and antagonists (Table 11.2) [14,15]. β-Adrenergic agonists and/or partial agonists are found to bind with large negative enthalpies (ΔH° from −79 to −17 kJ/mol), indicating strong electrostatic interactions between the bound conformation on the receptor, associated with a large loss of entropy (a tight complex is formed: $T\Delta S^\circ$ from −45.8 to + 8.6 kJ/mol). Antagonists, on the other hand, do not fulfill the requirements of a good complementarity as evidenced by the criterion that their enthalpies of binding are small (ΔH° from −21.3 to + 16.4 kJ/mol): The loss of entropy through conformational restriction is therefore less than that of the agonists, and—in fact—the entropy change is positive upon binding ($T\Delta S^\circ$ from +16.9 to + 52.5 kJ/mol). Overall, agonists and antagonists have similar association constants, supporting the $\Delta H^\circ/\Delta S^\circ$ compensation paradigm (Table 11.2).

In order to improve our understanding of the $\Delta H^\circ/\Delta S^\circ$ compensation, the data from Table 11.2 are represented with a plot of ΔH° versus $T\Delta S^\circ$. Indeed, in Figure 11.2, clusters clearly differentiate agonists from antagonists or partial agonists. An interesting explanation based on the dimerization of the membrane-bound receptors (G-protein-coupled receptors) has been given by Williams:[8] the agonists (●; activating the receptor after

TABLE 11.2 Thermodynamic Parameters of Ligand Binding to the β-Adrenergic Receptor of Turkey Erythrocytes and Ratio of Equilibrium Constants K_D at 0°C and 37°C [14]

	ΔG° (kJ/mol)	ΔH° (kJ/mol)	ΔS° (J/mol deg)	$T\Delta S^\circ$ (kJ/mol)	K_D (37°C)/ K_D (0 °C)
AGONISTS					
(−)-isoproterenol	− 39.3	− 56.0	− 54.0	− 16.7	23
(−)-norephedrine	− 33.1	− 78.9	− 147.7	− 45.8	55
PARTIAL AGONISTS					
soteronol	− 34.4	− 32.8	+ 5.3	+ 1.6	5.3
fenoterol	− 32.7	− 25.4	+ 23.6	+ 7.3	3.6
terbutaline	− 25.9	− 17.3	+ 27.8	+ 8.6	2.4
ANTAGONISTS					
(−)-propranolol	− 52.3	− 16.1	+ 116.7	+ 36.2	2.6
IPS-339	− 51.5	+ 1.1	+ 169.5	+ 52.5	0.95
pindolol	− 49.6	− 21.3	+ 91.2	+ 28.3	3.0
sotalol	− 34.4	− 9.0	+ 81.6	+ 25.3	1.6
atenolol	− 31.3	− 14.5	+ 54.4	+ 16.9	2.1
practolol	− 31.2	+ 16.4	+ 153.6	+ 47.6	0.45

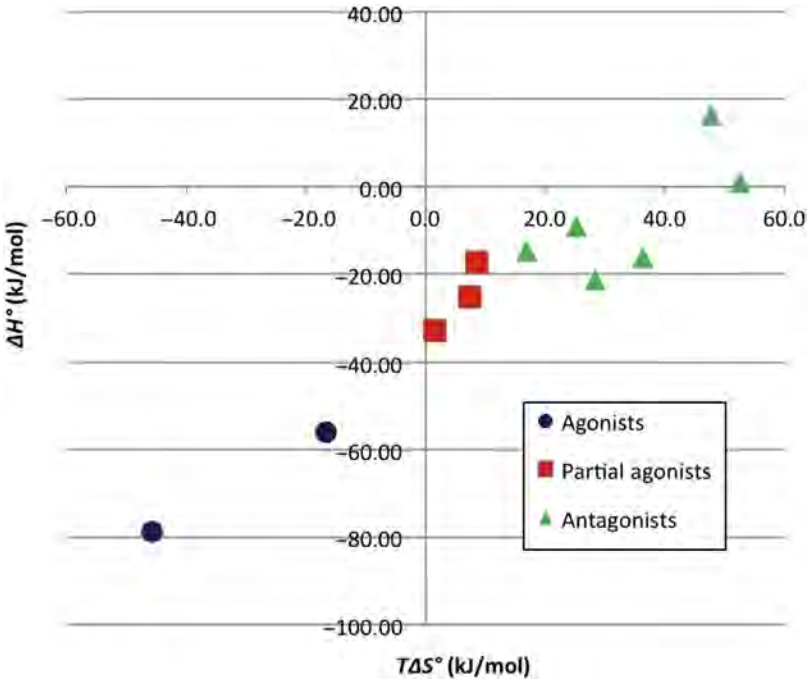


FIGURE 11.2 Plot of ΔH° versus $T\Delta S^\circ$ for the binding of agonists (●), antagonists (▲), and partial agonists (■) to the β-adrenergic receptors. *Data from Table 11.2 were used [14].*

binding) induce receptor aggregation, which is beneficial in bonding (negative contribution to ΔH°) but adverse in entropy (negative contribution to $T\Delta S^\circ$). The antagonist (▲; without activation after binding) should therefore be (in comparison to agonist binding) relatively disfavoring for dimerization and entropy. The consequences are seen in Figure 11.2. The agonists (●) are binding to the β-receptors by an enthalpy-driven process, whereas the antagonists (▲) are binding with entropy and/or enthalpy−entropy-driven processes.

In contrast, if agonists induce dissociation of receptor oligomers (e.g., the adenosine A1 or A2 receptors) [8,16], antagonists may induce the formation of oligomers. Such antagonist binding should therefore be (in comparison to agonist binding) relatively favorable in overall enthalpy and unfavorable in entropy, in contrast to the agonists with adverse enthalpy and favorable entropy.

Finally, an important point to stress is that hydrophobicity is a major source of binding in drug-receptor interactions. Based on a study of over 415 oral drugs, it appeared that on average drugs contain only one to two hydrogen donors and three to four acceptors, whereas the average number of hydrophobic atoms in a drug is 16 [17]. The contributions of polar and hydrophobic interactions in molecular recognition are related to the balance of enthalpy and entropy.

B. Steric Constraints

Before a small ligand interacts with a biological receptor, the two entities have their separate translational and rotational flexibility, which contribute to their respective entropy. Once the association occurs, some degrees of motion as well as internal rotations around single bonds are lost, and the consequence is an entropy cost of about -58 kJ/mol for a small ligand (MW $<$ 1,000 Da) at room temperature [4,7,8]. As a general consequence of the binding of a ligand to a receptor, entropy is decreased due to the loss of degrees of motion when two molecules are rigidly constrained within a complex. The torsional contribution to the entropy (the free rotation of a bond) is related to the number of rotatable bonds in the ligand, and freezing one of them has a cost of 2–3 kJ/mol at room temperature [4,7,9]. Considering Equation 11.1, the entropy penalty will render ΔS negative, reducing the binding energy. In order for $\Delta G^\circ < 0$, the costs in reducing conformational flexibility must be offset by favorable intermolecular interactions such as hydrogen bonds, van der Waals packing, hydrophobic interactions, and Coulombic interactions. One consequence of this effect is that the binding optimization of a flexible ligand can be accomplished by making ΔH° more negative, by making ΔS° more positive, or by an appropriate combination of both. Theoretically, the highest value for a given binding constant is accessible to a flexible ligand if the receptor recognizes the ligand in its low-energy conformation that presents an optimal orientation of the functional groups. Therefore, all information gained about the active conformation of the ligand will serve the chemical design of a better lead structure. Some illustrative examples can be obtained from the incorporation of constrained amino acids into bioactive compounds [18]. However, since the determination of the thermodynamic parameters of ligand/receptor interactions is still in its infancy, optimizations in structure-based drug design rely on semi-quantitative and empirical strategies and structure–activity relationships (SAR). Conformational constraints, steric hindrance, and rigidification are frequently used principles in medicinal chemistry to explore binding parameters. Phenylalanine, a lipophilic amino acid, has been a popular substrate for evaluating this concept of rigidification. Nonnatural phenylalanine analogs have been synthesized by the introduction of bulky substituents, by the application of carbocyclic constraints, or by the incorporation of a functional group. Examples of conformationally constrained phenylalanine analogs are shown in Figure 11.3 [19–22].

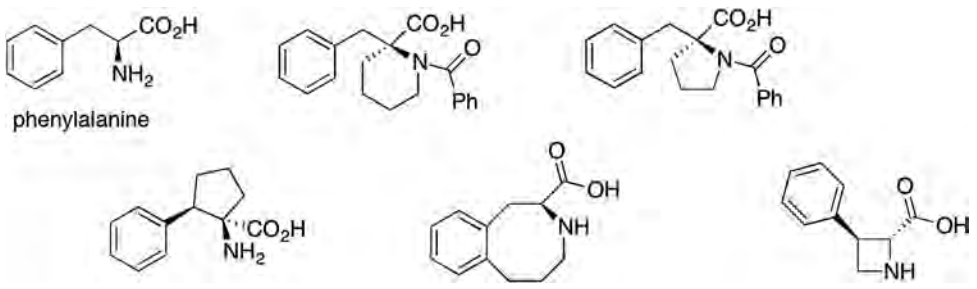


FIGURE 11.3 Conformational rigidification of phenylalanine analogs [19–22].

C. Conformational Analysis

Many computational methods are available to determine the low-energy conformations of a flexible ligand [23–25]. During the binding process, flexible ligands are refolded. This refolding is a general phenomenon and is presumably compensated by the search of a ligand for hydrogen bonds on the protein to replace the solute-solvent hydrogen bonds that are lost as the molecule enters the binding site (as shown in Figure 11.1) or in the unmasking of hydrophobic pockets on the receptor. A study has been performed on thirty-three ligands whose X-ray structures as well as conformations obtained by co-crystallization with their receptors have been recorded [26]. From this study, it appears that the degree of conformational change depends in part on the number of rotatable bonds in the ligand. For ligands with five or more rotatable bonds, the conformation in the crystal

structure does not represent the protein-bound conformation. Therefore, the solid-state structure of flexible ligands remains often of limited use for SAR analyses. Of course, in the more favorable cases, when a crystal structure of the complex-ligand−receptor can be obtained, the design of an improved ligand is facilitated because the main interactions can be recognized and then further optimized via synthesis.

D. Steric Effects

Steric effects may arise in a number of ways. Primary steric effects result from repulsions between valence electrons or nonbonded atoms. Such repulsions can only result in an increase in the energy of the system. In a chemical reaction, the overall steric effect may be either favorable or unfavorable. For example, if steric effects in the reactant are larger than in the product (or transition state), then the reaction is favored (steric augmentation); if the reverse case is true, the reaction is disfavored (steric diminution). The same arguments can be used in biological systems for the formation of a receptor−ligand complex. Comparing the binding of a ligand to a biological receptor with or without a subsequent chemical reaction—for example, an enzyme or a hormonal receptor—reveals some obvious differences. The enzymatic reactions involve only substructures in proximity to those atoms that are actually participating in bond making or breaking. Therefore, the enzyme tolerates structurally different ligands, provided the position on the substrate where the reaction should take place is accessible. Even if the direct affinity of a ligand for an enzyme is low, the subsequent chemical transformation can take place. In the formation of a ligand−receptor complex, any group of atoms that is in van der Waals contact with the receptor or the biomolecule can be or is involved in the binding event. If the receptor sits in a pocket that can adjust to any bioactive substance no matter its size or shape, then no steric effect will be observed. If, however, the parent biopolymer has limited conformational flexibility, and—as is likely—this flexibility is not equivalent in all directions, then a steric effect will be observed. Furthermore, the steric effect will be dependent on conformational states, and the minimal steric interaction principle will probably be observed. This principle states that a substituent whose steric effect is conformationally variable will prefer a conformation that minimizes steric repulsions and will give rise to the smallest steric strain. Finally, there are secondary steric effects on receptor binding that are produced by a substituent: (1) lowering the accessibility to an important group due to steric hindrance; (2) changing the population of a conformer due to steric effects; (3) shielding the active site from attack by a bulky group; and (4) variation in the electronic resonance of a π-bonded substituent by an out-of-plane repulsion.

E. Rigid Compounds and Bioavailability

Any effort toward improving bioavailability is of great importance in the drug discovery process. An intriguing correlation between the bioavailability of a compound and the number of its rotatable bounds was reported by Veber and coworkers using an empirical approach based on a set of 1,100 drug candidates and metabolism data collected from rats [27]. Remarkably, 10 or fewer rotatable bonds together with a polar surface area <140 $Å^2$ (or 12 or fewer H-bond donors and/or acceptors) irrespective of the molecular weight resulted in a high probability for good oral bioavailability in rats. This finding expanded the acceptable range of molecular weight values for a drug candidate beyond the threshold of 500 g/mol set by the Lipinski rules [28]. Veber suggested that by freezing some of the rotatable bounds, the molecular weight was no longer an essential parameter to be considered. The extent to which the number of rotatable bonds, as compared to other physicochemical parameters, affects oral bioavailability was recently also analyzed by Varma and coworkers [29]. Oral bioavailability (F) depends on the fraction of compound absorbed (F_a) as well as first-pass elimination that is determined by the fraction escaping gut-wall elimination (F_g) and the fraction escaping hepatic elimination (F_h), as shown in Equation 11.3. Each of the parameters evaluated (molecular weight, ionization state, lipophilicity, polar surface area, and rotatable bonds) affected at least one aspect of oral bioavailability. For example, while increasing molecular weight caused a decrease in F_a, increasing lipophilicity caused a decrease in F_g and F_h. Most notably, increasing the number of rotatable bonds had a negative effect on all three parameters and profoundly affected oral bioavailability [29]. Therefore, the introduction of conformational constraints in a drug candidate also has to be considered for resolving pharmacokinetic issues.

$$F = F_a \times F_g \times F_h \tag{11.3}$$

II. CASE STUDIES

A. The Study of Receptor Structure

A classic example of exploring the structure and binding requirements of a target enzyme by conformational restriction without the benefit of a co-crystal structure is presented by the μ-opioid receptor antagonists. Drugs that have been approved to target this receptor are used for pain management, and the name of the receptor family is derived from the prototypical antagonist of this family, morphine. Activation of the μ-opioid receptor family by endogenous opioid peptides in the body also affects behavioral and homeostatic functions and could therefore be used to treat indications in addition to pain. Zimmerman and coworkers first described the *trans*-3,4-dimethyl-4-(3-hydroxyphenyl)piperidine μ-opioid receptor antagonists in 1978 [30] and later clarified the structural requirements that lead to higher affinity binding (Figure 11.4, A) [31]. Subsequent to the discovery of these analogs, Le Bourdonnec and coworkers designed analogs of Zimmerman's antagonists in which the rotation about the N-substituent was constrained by the formation of a fused ring [32]. Because the stereochemical requirements at C3 and C4 had been established, the 4-stereoisomers of each of the regioisomers of the constrained analogs were synthesized. Interestingly both a potent antagonist (Figure 11.4, B) and agonist (Figure 11.4, C) were found among these analogs. An analysis of the lowest energy conformers of these three molecules showed good overlap of A and B, with the hydroxyphenyl substituent in the equatorial position. The lowest energy conformer of the agonist places the hydroxyphenyl substituent in the axial position. The crystal structure of the μ-receptor was not solved until 2012 [33], and the conformationally restricted analogs were useful in gaining information about the binding pocket of the receptor prior this discovery, providing a successful complement to the traditional bioactive structures based on morphine.

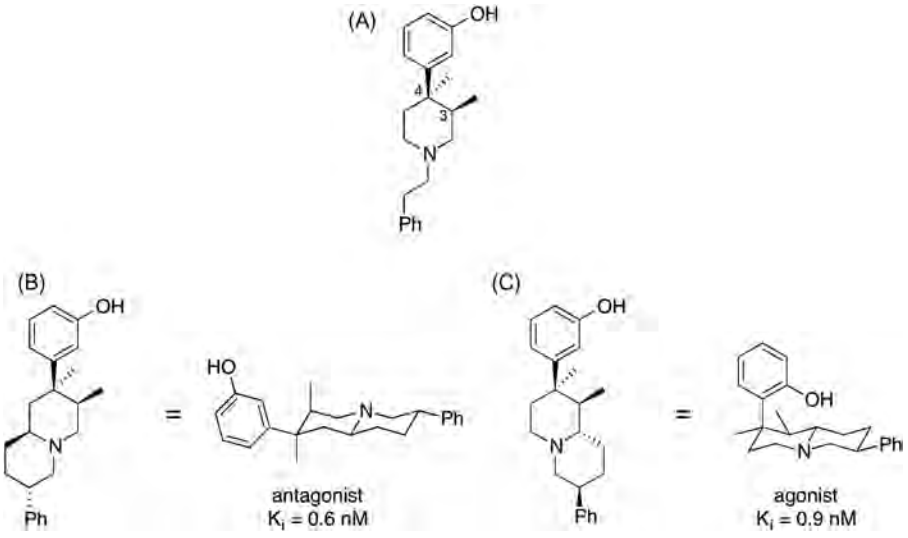


FIGURE 11.4 Structures of small molecules that interact with the μ-opioid receptor: (A) initial scaffold discovered;[30,31] (B) constrained, potent antagonist; (C) constrained, potent agonist [32].

B. Atropisomers

Atropisomers are chiral due to restricted rotations of σ-bonds with barriers exceeding 20 kcal/mol, which results in a half-life of >1000s at room temperature [34]. Atropisomers often result from a hindered rotation of bonds connecting two aromatic rings (as in Figure 11.5, A) or from a barrier in the ring flip of a medium-sized ring (as in Figure 11.5, B). Depending on the overall structure of the molecule, atropisomers can form either enantiomers or diastereomers. The first enantiomerically enriched atropisomers were isolated in 1922 [35], and since then differences in their physical properties and bioactivities have been noted [36–38].

Atropisomers racemize by a bond rotation, and the activation energy for this process depends on the electron distribution, solvent, temperature, and steric hindrance. Atropisomers that have a high barrier to racemization are easily separable and should be stable under physiological conditions. Conversely, atropisomers that have a low barrier to rotation will undergo interconversion under physiological conditions, and therefore it is unnecessary to separate them for biological activity differentiation and chemical characterization. Over the last two decades, the pharmaceutical industry has moved away from developing racemic mixtures as drug candidates; as a

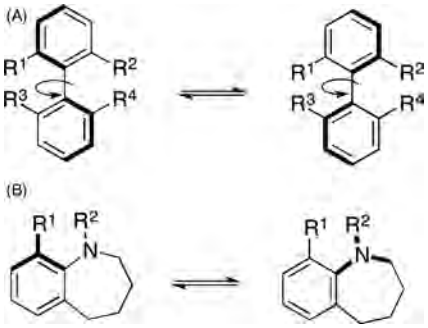


FIGURE 11.5 Generic atropisomers displaying (A) axial chirality due to hindered rotation of an aryl-aryl σ-bond and (B) planar chirality due to steric hindrance of the R groups and the dissymmetry of the bridging methylene and amine groups.

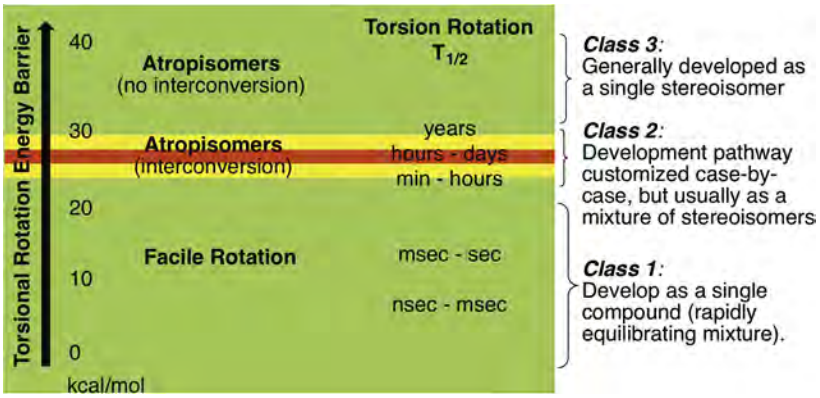


FIGURE 11.6 Table describing classes of atropisomers based on the barrier to rotation developed by LaPlante and coworkers for applications in drug development [36].

result, several reviews focus on the merits of developing a preclinical candidate as a single atropisomer rather than a mixture [36,37]. LaPlante and coworkers have established recommendations for the drug development of atropisomers based on US FDA guidelines summarized in Figure 11.6. [36,39], In short, if the barrier to rotation of an atropisomer is high ($\Delta E_{rot} \geq 30$ kcal/mol, $t_{1/2}$ in years), then the atropisomers should be resolved and the less active isomer should be treated as an impurity (Figure 11.6, Class 3). If the atropisomers rapidly equilibrate ($\Delta E_{rot} < 20$ kcal/mol, $t_{1/2}$ of minutes, Class 1), they should be treated as a single compound, since separation and individual bioactivity assessment would be challenging. The gray area of atropisomer development lies within the ΔE_{rot} 20−30 kcal/mol range ($t_{1/2}$ of hours to days, Class 2). In this area, development is challenging due to difficulties in isolation and stability, and the authors recommend increasing the ΔE_{rot} by adjusting steric hindrance, if possible, to move these compounds into Class 3.

1. *Bombesin Receptor Subtype-3 (BRS-3) agonists*

Obesity is a growing problem in our society and has been shown to cause a number of other diseases, including type-2 diabetes, cardiovascular disease, cancer, and hypertension [40]. As the rate of obesity is increasing, so too is the need for an effective anti-obesity drug. It was found that metabolic defects and obesity are present in mice that lack functional bombesin receptor subtype-3 (BRS-3) [41,42], and a BRS-3 agonist caused a decrease in food intake accompanied by an increase in metabolic rate in diet-induced obese mice [43]. In a BRS-3 HTS screen, Merck chemists discovered the compound shown in Figure 11.7 (A) [44,45]. Follow-up chemistry led to the unexpected discovery that many members of this compound class existed as a mixture of two separable enantiomers (Figure 11.7, B and C). Upon biological testing, it was determined that the *R*-enantiomer was much less potent than the *S*-enantiomer, with IC_{50} values of 169 nM versus 1.4 nM, respectively.

To determine the extent of steric bulk necessary to allow a room temperature resolution of the atropisomers, several analogs were analyzed [45]. It was found that when the C-7 substituent was either an H or F atom, no atropisomerism was observed (Figure 11.8, A). Bulkier halogens, such as Cl, a hydroxyl group (Figure 11.8, A), or a fused ring (Figure 11.8, B), led to atropisomerism. As expected, when a pyridine nitrogen atom lone pair occupied the space at the 7-position, no atropisomerism was observed (Figure 11.8, C). The addition of a large

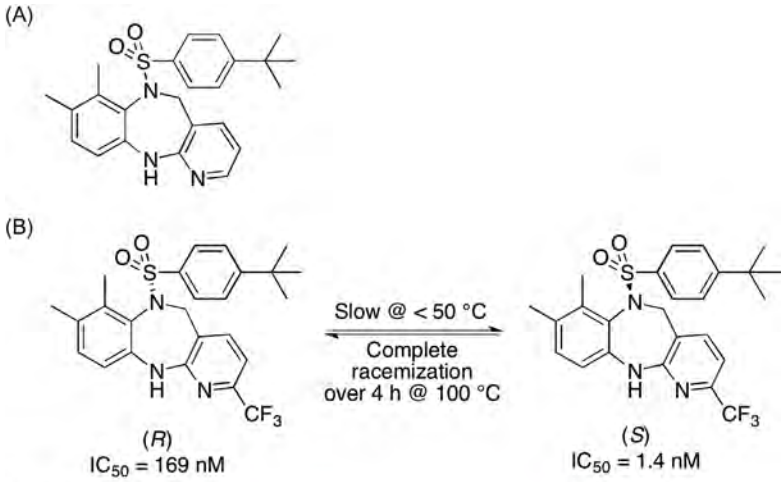


FIGURE 11.7 Structure of (A) an HTS hit from a BRS-3 screen, (B) more potent compounds resulting from follow-up chemistry that were separable atropisomers whose planar chirality was (C) confirmed by X-ray crystallography. *Adapted with permission from Liu P, Lanza TJ, Chioda M, Jones C, Chobanian HR, Guo Y, et al. Discovery of benzodiazepine sulfonamide-based bombesin receptor subtype 3 agonists and their unusual chirality. ACS Med Chem Lett 2011;2:933–7. Copyright 2011 American Chemical Society.*

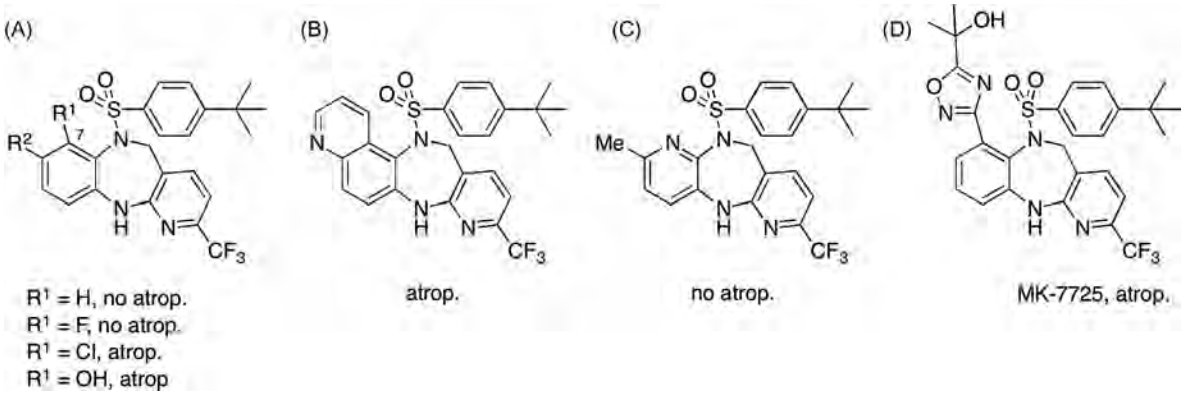


FIGURE 11.8 SAR study of atropisomers in which various substituents at C-7 were tested (no atrop. = no atropisomerism was observed; atrop. = atropisomerism was observed) [45].

heterocyclic substituent at C-7 led to the discovery of an inhibitor with comparable activity and greatly improved pharmacokinetics, MK-7725 (Figure 11.8, D). [46]

2. *Glycine transporter 1 (GlyT1) inhibitors*

Atropisomerism in the form of axial chirality is a more commonly observed occurrence, and enantiomers of this type were observed by Sugane and coworkers during their search for selective glycine transporter 1 (GlyT1) inhibitors [47]. GlyT1 is a glycine transporter that is present in forebrain areas where it terminates glycine activity by mediating its uptake [48,49]. It is believed that GlyT1 modulates NMDA receptor activity by regulating glycine levels, and a decrease in NMDA is a contributing factor to schizophrenia [50].

An HTS hit (Figure 11.9, A) from a GlyT1 screen was found to be a low micromolar inhibitor of GlyT1 but displayed no selectivity over the closely related GlyT2 [51]. It was noted that increasing the steric bulk at either the 3- or 4-position on the triazole ring increased the selectivity for GlyT1. Indeed, the *R*-enantiomer of the optimized

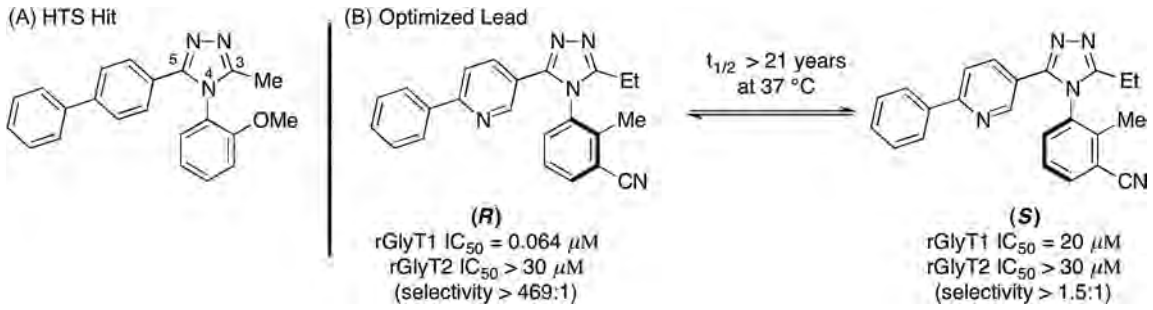


FIGURE 11.9 (A) Structure of an HTS hit from a GlyT1 screen; (B) optimized lead that resulted from increasing the steric bulk to impede axial rotation. [47].

lead (Figure 11.9, B) had an IC_{50} of 64 nM against GlyT1, while the *S*-enantiomer had an IC_{50} of 20 μM. These enantiomers were separable by chiral HPLC, and it was determined that the ΔE_{rot} was 31.4 kcal/mol (resulting in a $t_{1/2}$ of >21 years at 37°C), indicating that these compounds are Class 3 atropisomers (as described in Figure 11.6).

C. Peptidomimetics

1. *Gramicidin S Derivatives*

Gramicidin S (Figure 11.10) is a cyclic decapeptide natural product whose history spans more than seventy years. It displays antibiotic activity against both Gram-negative and Gram-positive bacteria by interacting with the bacterial lipid bilayer [52]. This interaction is closely linked to the structure of gramicidin S—its amphipathic properties are enhanced by its 3-dimensional scaffold. This enhancement is due to an antiparallel β-sheet that aligns hydrophilic and hydrophobic portions on opposite faces of the molecule. Two type II′ β-turns and four intramolecular hydrogen bonds stabilize the extended secondary structure, which has been extensively studied

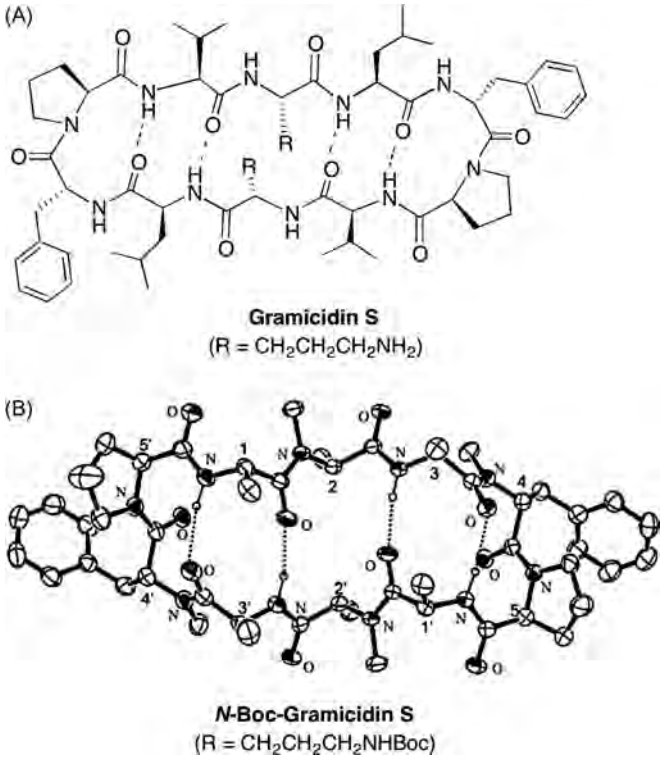


FIGURE 11.10 (A) Structure of gramicidin S showing the intramolecular hydrogen bonding that stabilizes the β-hairpin secondary structure; (B) X-ray structure of the Boc-derivative of gramicidin S [54].

by NMR [53], X-ray crystallography [54], and circular dichroism (CD) [55]. Perturbations to this 3-dimensional structure due to steric bulk or electronic variations that twist the molecule from its natural conformation cause a decrease or complete loss in function [56].

Because peptides are often plagued by poor bioavailability and rapid hydrolysis in vivo, bioisosteric replacements of the amide bond have been of great interest [57,58]. In order for such a rational design to be successful, the bioisostere must allow for a close match of the geometry of the peptide bond while remaining stable to peptidase cleavage. Among many isosteric substituents, the (*E*)-alkene peptide isosteres [58] have been quite successful, since they have a rigid geometry that is locked in place by a nonrotatable double bond and therefore mimic the sp^2 nature of the amide nitrogen particularly well. (*E*)-Alkene peptide isosteres are impervious to peptidase cleavage, but depending on the substituents attached to the double bond, any H-bonding in which an amide would participate can be lost. Furthermore, disubstituted alkenes are conformationally more flexible than the parent amide and display smaller dipole moments and greater lipophilicity. In the case of gramicidin S, loss of hydrogen bonds could be detrimental to the bioactivity, because they stabilize the molecule in its bioactive conformation. As such, this cyclopeptide antibiotic represents a useful probe substrate for the tolerance for bioisosteric group replacement and its effect on conformation and activity.

To explore the consequences of amide bond replacement on the properties of gramicidin S, more stable functional groups that should allow for the bioactive geometry of gramicidin S to remain intact were introduced into the backbone sequence. Specifically, (*E*)-alkene peptide isosteres with different substituents attached to the alkene were synthesized via a hydrozirconation/transmetalation/imine addition sequence. This synthetic sequence allowed for a stereoselective formation of the allylic amine portion of the isosteres from an alkyne and enantiomerically pure sulfinamide (Figure 11.11).

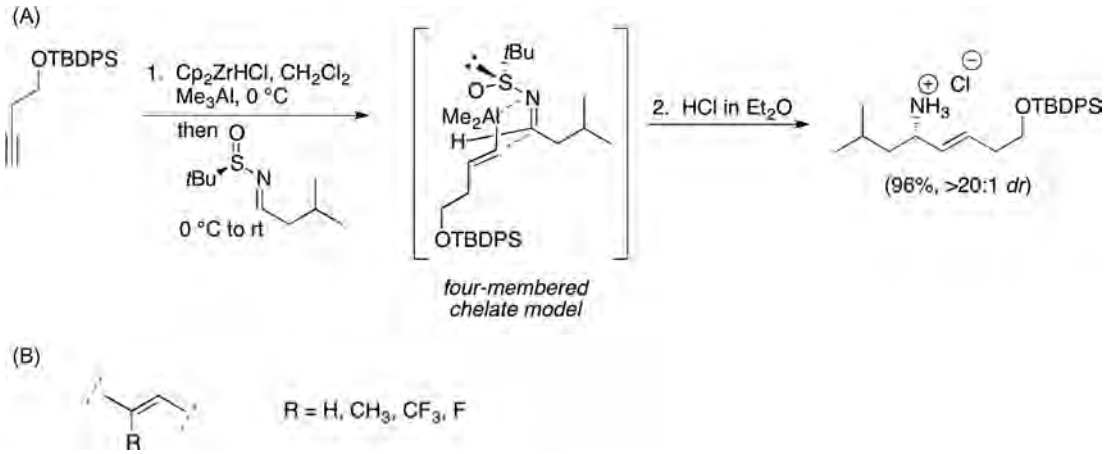


FIGURE 11.11 (A) Large-scale stereoselective synthesis of an allylic amine intermediate used in the synthesis of the (*E*)-alkene peptide isostere portion of the bioactive peptide mimic JP4-039 [59] and (B) examples of differentially substituted (*E*)-alkene peptide isosteres used in the gramicidin S analogs [60].

Trisubstituted alkenes with methyl, trifluoromethyl, and fluoro substituents as well as a disubstituted alkene were used as replacements for the Leu-Phe peptide bond. While the tertiary (*E*)-alkene peptide isostere containing the methyl substituent was of similar size as the corresponding trifluoromethyl analog, the methyl analog caused a greater perturbation of the secondary structure of the molecule, resulting in a CD spectrum that was indicative of a random coil secondary structure. The trifluoromethyl analog was a superior match of the parent compound in both solution and solid-state structures [55]. Even though the plane of the alkene was twisted 70° away from the interior of the β-turn due to the bulkiness and hydrophobicity of the CF_3 substituent, the overall geometry of the type II′ β-turn was highly preserved, and the β-hairpin structure of the cyclopeptide was unchanged.

A deprotected derivative of a trisubstituted (*E*)-alkene peptide isostere was also a successful structural mimic of the D-Phe-Pro moiety. Not only did this molecule have the desired 3-dimensional structure in solution (as indicated by the CD spectrum), but it also demonstrated equipotent biological activity to GS against *Bacillus subtilis* (20 and 15 μg/mL, respectively) [60]. The fluoro- and disubstituted (*E*)-alkene peptide isosteres also formed β-sheets, successfully mimicking the natural peptide. Therefore, these molecules demonstrate that peptide bond isostere replacements can preserve both the 3-dimensional structure as well as the biological profile of the parent structure.

2. XJB-5-131 and JP4-039 as GS-Mimics Targeting Mitochondria

Based on the properties of GS to interact with bacterial membranes and the knowledge that mitochondria evolved from bacteria, smaller and more drug-like GS-mimics were designed with the intent of using them to target the mitochondrial membrane. For the design of such a mimic, the ornithine side chains remained protected as the carbamate derivatives to decrease a potential hemolytic activity upon enrichment in erythrocytes. Furthermore, a segment of GS was used to decrease molecular weight, but the alkene peptide isostere and the D-Phe-Pro sequence were conserved to stabilize the type II′ β-turn structure, decrease the potential for proteolytic degradation, and diminish the number of hydrogen bond donors and acceptors. The (*E*)-alkene isostere XJB-5-131 was found to match a type II′ β-turn by CD analysis, and XJB-5-131 was enriched ca. 600-fold in mitochondria over the cytosol (Figure 11.12, A). The nitroxide "warhead" was introduced to react with reactive oxygen species (ROS) that are formed in the organelle, and the ROS scavenger XJB-5-131 proved to be effective in a number of disease and injury models involving cellular damage by reactive oxygen species. Such diseases include many acute and chronic degenerative disorders (e.g., hemorrhagic shock [61], hyperoxic acute lung injury [62], traumatic brain injury [63], microdialysis injury [64], and Huntington's disease [65]).

Following the success of the GS-nitroxide XJB-5-131 as a mitochondrial-targeted ROS scavenging agent, attempts to generate a lower molecular weight analog led to the discovery of JP4-039 (Figure 11.12, B). The structure of JP4-039 contains a nitroxide directly attached to the dipeptide isostere, and it also adopts a type II′ β-turn, as evidenced by its crystal structure [59]. A fluorescent derivative of JP4-039 was used to visualize the enrichment of this compound in mitochondria [66].

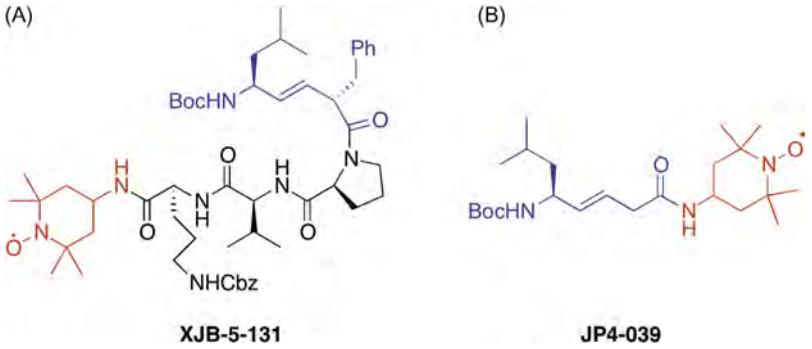


FIGURE 11.12 Structures of mitochondrial-targeted (A) XJB-5-131 and (B) JP4-039 (with the dipeptide mimetic in blue and the nitroxide in red).

D. Methyl Group Effects on Conformation

The seemingly innocuous methyl group is often added to small molecules to sterically block sites of metabolic activity and increase $t_{1/2}$. In some cases, its addition to a molecule also has a profound effect on biological potency [67,68]. Because the methyl group is a hydrophobic substituent, it can cause a slight increase in potency by reducing the desolvation energy that is required to remove the solvation by water molecules when the molecule enters a hydrophobic protein cavity from an aqueous environment [69,70]. This decrease in desolvation energy can increase the potency of a molecule by approximately 3.5-fold; any dramatic increase in potency beyond what is attributed to the increase in hydrophobicity is colloquially called the "magic methyl effect" [71] by medicinal chemists. Such "magic" effects have increased potency by over 1,000-fold [72] and even converted an agonist to an antagonist [73], but the addition of a methyl group is just as likely to decrease potency as it is to increase it [68].

1. Phospholipase D (PLD1/2) Inhibitors

Phospholipase D (PLD) catalyzes the hydrolysis of phosphatidylcholine into choline and phosphatidic acid, and suboptimal PLD function has been implicated in cancer [74] and CNS disorders [75,76]. By simply adding a methyl group to the ethylene diamine linker, Lindsley and coworkers were able to increase the activity of their PLD1 inhibitors from an IC_{50} of 11,800 nM to 20 nM (Figure 11.13) [77]. This dramatic increase in activity is likely due to a similar phenomenon observed in the factor Xa inhibitors [78]. The rotamer distribution of the dihedral angle around the amide backbone was analyzed, and it was found that when the methyl group was incorporated, fewer rotamers were formed and the rotational barriers increased. The more distinct conformational preferences favored the bioactive conformation [79].

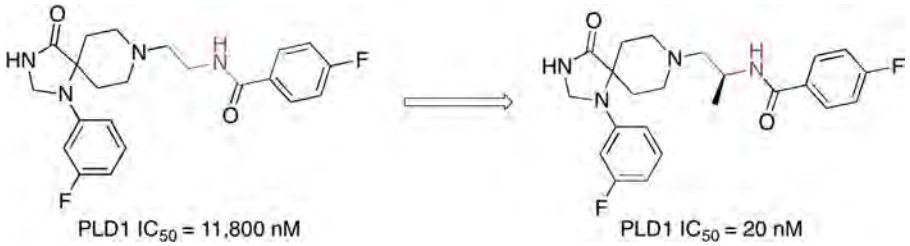


FIGURE 11.13 Dramatic effect on potency of PLD1 inhibitors by the addition of a methyl substituent, with the pertinent bonds shown in red [77].

2. *Phosphatidylinositol-3-Kinase (PI3K) Inhibitors*

While the PLD1 inhibitors experienced a dramatic increase in potency with the addition of a single methyl group, an analogous structural modification in Pfizer's phosphatidylinositol-3-kinase (PI3K) inhibitors had a much more variable effect on activity. The PI3K/mTOR (mammalian target of rapamycin) signaling pathway has been implicated in a number of disease pathologies including cancer [80] and aging [81]. Quinazolines, such as those shown in Figure 11.14 (A), were synthesized based on previously determined SAR [82]. It had been found that the C-2 methyl group was necessary for selectivity and was therefore incorporated into all new analogs. However, when a methyl group was added to the C-7 position (Figure 11.14, B), activity decreased substantially. Conversely, the addition of a methyl group at C-6 caused an almost 5-fold increase in potency over the 6-H derivative. These significant differences in potency with such seemingly small structural changes demonstrate the subtleties of relatively minor structural perturbations in medicinal chemistry. Computational studies of the 7-methyl and the 6-methyl derivatives showed that a planar orientation of the amide side-chain that was locked in place by the hydrogen bonding of the N-H to the quinazoline N-1 was optimal for the molecule to fit into the binding pocket [82]. The presence of the 7-methyl disrupts the planarity and introduces a twist to the N-C-C8-C9 dihedral angle calculated as $> 20^{\circ}$.

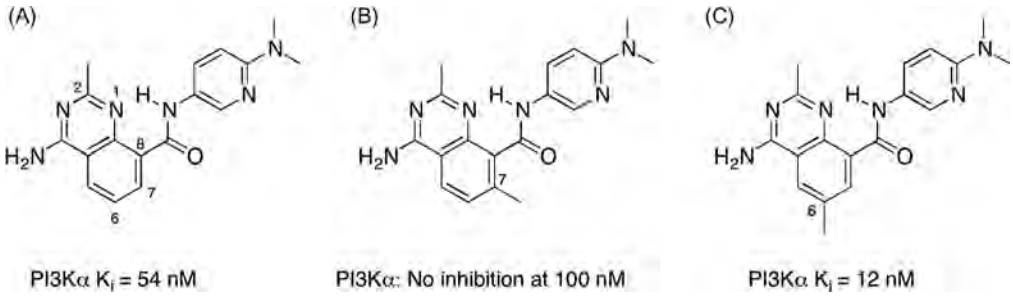


FIGURE 11.14 Structures of putative PI3K/mTOR inhibitors. The methyl group placement proved to be important for activity. (A) Absence of a methyl group at C-7 led to a fairly potent inhibitor with an IC_{50} of 54 nM, while (B) with a 7-methyl was inactive and (C) with a 6-methyl was the most potent inhibitor with an IC_{50} value of 12 nM [82].

E. Dihedral Angle Optimization

A thorough understanding of the receptor structure was used in the case of the detailed design of inhibitors of the hepatitis C virus (HCV). HCV is an infectious disease that primarily affects the liver with a chronic infection that eventually manifests itself by cirrhosis and possibly liver cancer or failure. The receptor structure of the virus was determined by a combination of X-ray crystallography of inhibitors bound to HCV NS5B, NMR spectroscopy, and docking [83]. This allowed for the detailed study of the bioactive conformation and specifically the requisite bond angles as shown in Figure 11.15 (A). Early in their studies, LaPlante and coworkers found that the bioactive conformation around dihedral angle α placed C-18 into the vicinity of NH-1. Several attempts to modulate this angle by adding a methyl group to C-16 or introducing a hydrogen-bonding pyridine N atom into the aromatic ring failed to improve potency, despite the bioactive conformation being supported by ROESY NMR correlations. Competing unfavorable forces such as electrostatic repulsion likely led to the undesirable potency. A more favorable outcome was obtained after modulating the rotation about the dihedral angles β and δ. Correlation of the cellular EC_{50} values with the substituents at C-2 showed that transitioning from the unsubstituted to a single methyl, a gem-dimethyl,

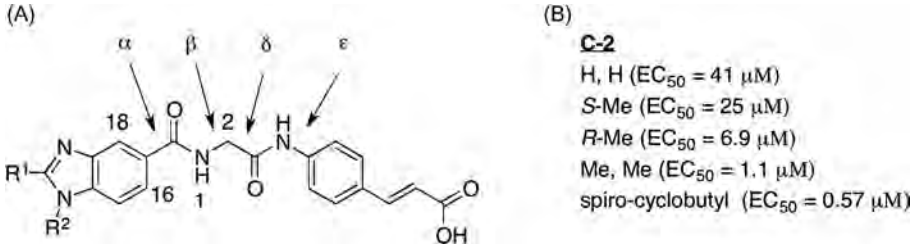


FIGURE 11.15 (A) HCV inhibitor with dihedral angles of interest highlighted and (B) EC_{50} values for varying substituents at C-2.

and finally a spiro-cyclobutyl group improved the cellular potency from 41 μM to 0.57 μM (Figure 11.15, B). Molecular mechanics was used to analyze the angle distributions of the unsubstituted C-2 compared to the spiro-cyclobutyl derivative, and the bioactive dihedral angle was more prevalent in the substituted analog.

The progression of molecular structure can be seen in Figure 11.16. The HTS hit (A) was active *in vitro* but had no cellular activity. Initial modification of the pertinent dihedral angle α led to the tryptophan series of compounds B. Unfortunately, this series had high micromolar cellular efficacy, which led to a transition to the diamide series (C). This series had the optimal dihedral angles for α, β, and δ, but still showed micromolar potency. Finally, dihedral angle ε was further restricted by the incorporation of the benzimidazole amide bond isostere as shown in D. This compound, with its fully optimized 3-dimensional structure, led to a dramatic improvement in cellular potency and served as the clinical candidate for the treatment of HCV infection.

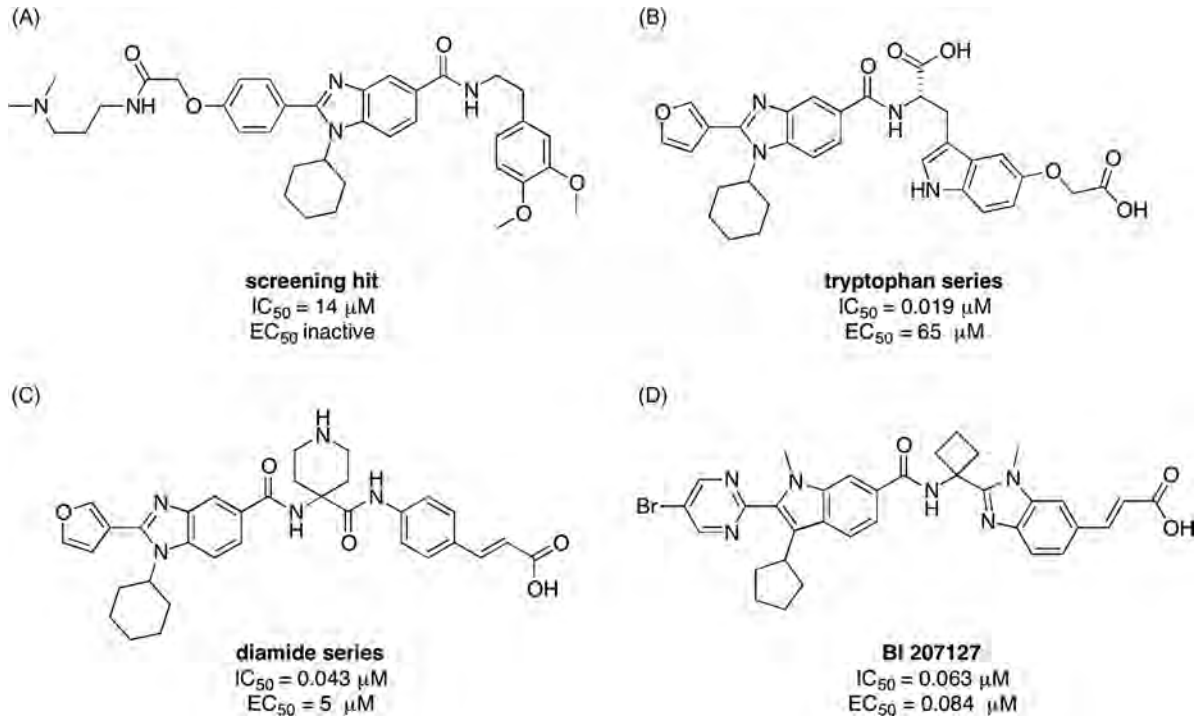


FIGURE 11.16 Structures of the (A) HTS hit, and (B) the lead structure initially used to determine bond angles, (C) the diamide series, and finally (D) the clinical candidate for HCV inhibition.

F. Diversity-Oriented Synthesis

Diversity-oriented synthesis (DOS) "involves the deliberate, simultaneous and efficient synthesis of more than one target compound in a diversity-driven approach to answer a complex problem." [84] It is a strategy that allows for the use of an efficient synthetic design to explore chemical space. This is an important as well as challenging undertaking, since chemical space is vast and significantly underexplored [85]. Estimates for the total number of synthetically tractable organic molecules containing carbon, hydrogen, nitrogen, oxygen, sulfur, and halogens limited to molecular weights below 500 g/mol vary between 10^{20} and 10^{60} molecules [86].

Two common approaches are used when applying DOS in medicinal chemistry. The first involves the expansion of a screening library so that a thorough study of functional groups, configurations, and conformations can be achieved during the primary HTS screening stage, as in the case of the β-cell apoptosis inhibitors [89]. This method offers an advanced starting point for the medicinal chemistry stage of a project, because the compounds included are so diverse yet methodical in stereochemistry and scaffold evaluation. The second approach involves the application of DOS to hit-to-lead development. Once a hit from a screen is identified, DOS can be used to synthesize a variety of structures in a targeted library rapidly to map out SAR information in a short amount of time, as demonstrated in the case of the Hsp70 inhibitors [95].

1. *Inhibitors of Cytokine-Induced β-cell Apoptosis Discovered Via HTS*

Type-1 diabetes is an autoimmune disease in which pancreatic insulin-producing β-cells undergo an increased rate of apoptosis due to an overproduction of cytokines such as IL-1β, IFN-γ, and TNF-α [87]. These cytokines activate transcription factors such as NFκB and STAT1, triggering the apoptotic pathway. An HTS screening campaign measuring cellular ATP levels as an indication of cell viability in a rat β-cell line treated with pro-inflammatory cytokines was used to find suppressors of this pathway [88]. The libraries tested in this HTS screen included several DOS libraries, among which one library was found to be particularly active. This library of fused medium-sized bicycles contained 6,488 compounds that were synthesized from stereochemically pure starting material (Figure 11.17) [89]. Starting with a chiral pool of desired amino alcohols and γ-amino acids, peptide coupling and reduction of the resulting amide were used to combine the building blocks. Acylation of the secondary amine followed by S_NAr cyclization provided the 8-membered lactam core. This synthesis utilized SynPhase solid support lanterns attached to a primary alcohol for ease of purification and handling. Capping of the aniline nitrogen by reaction with twenty-seven different building blocks, including sulfonyl chlorides, isocyanates, acyl chlorides, and aldehydes, provided a diverse set of intermediates (Figure 11.17, blue R^1). Subsequent deprotection of the secondary amine and conversion to sulfonamides, ureas, or tertiary amines yielded a diverse collection of analogs as a set of eight stereoisomers (Figure 11.17, red R^2).

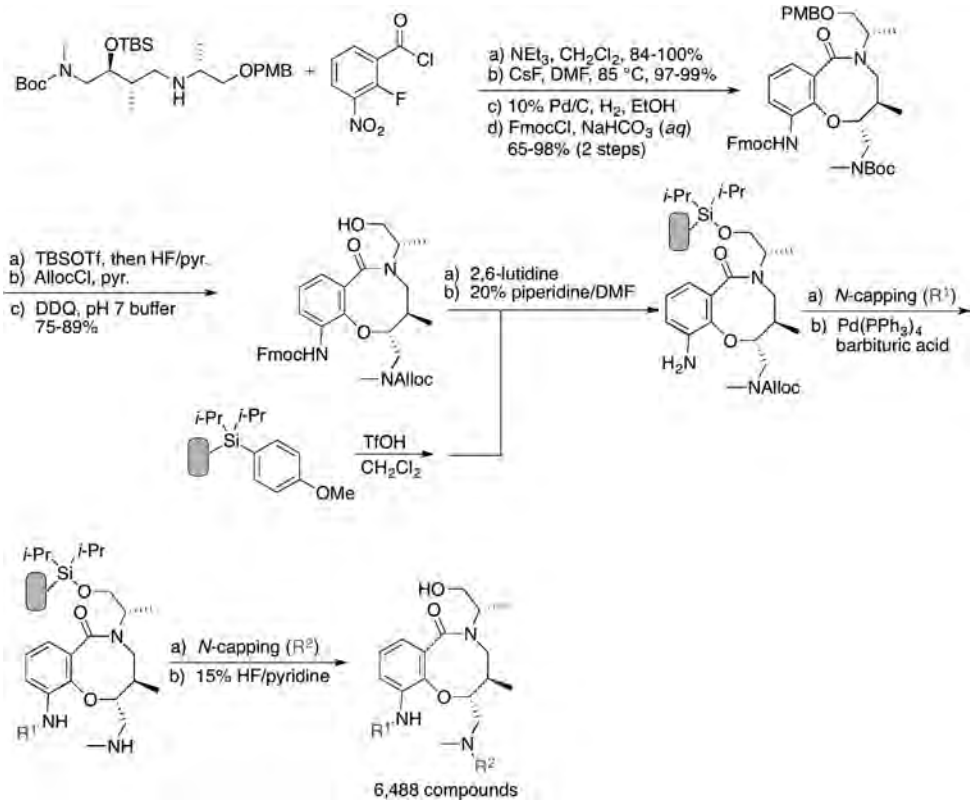


FIGURE 11.17 DOS library synthesis of 6-8 fused rings that were found to be hits in a β-cell apoptosis inhibitors screen. Only the *S,R,R* enantiomer is shown but each possible stereoisomer was synthesized. *This figure was adapted from Ref. [89].*

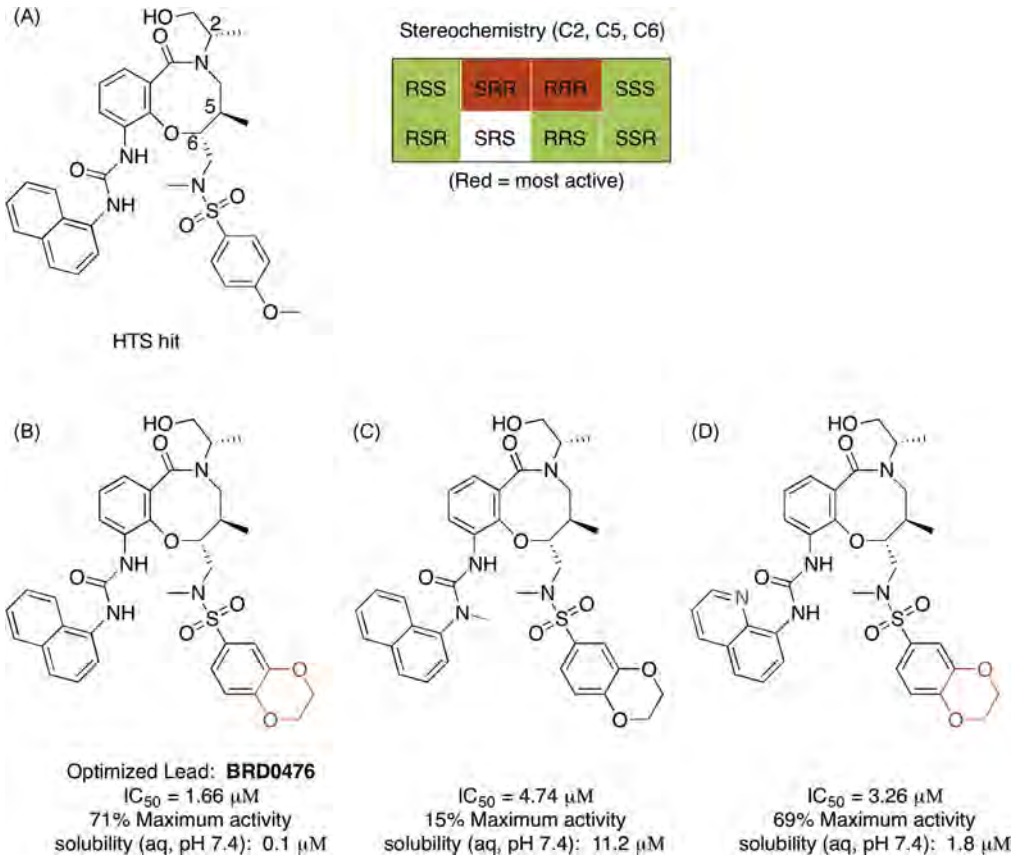


FIGURE 11.18 Structures of (A) an HTS hit and analogs with the incorporation of (B) a fused 1,4-dioxane, (C) a methyl urea, and (D) a quinoline, and their effects on potency and solubility [88,89].

The HTS screening hit is shown in Figure 11.18A. These data can be summarized in a heat map format with red boxes indicating the most active analogs and each 2×4 box showing all of the possible stereoisomers of a given analog. This heat map allows for a visualization of the activity and a graphical overview the SAR (structure-activity relationship) as well as SSAR (stereochemical structure-activity relationship) for all 6,488 analogs. The most potent compound in the original screen was the 2*S*,5*R*,6*R*-derivative with a naphthyl urea at R^1 and a *p*-methoxy phenyl sulfonamide at R^2 (Figure 11.18, A), exhibiting an EC_{50} of 4.89 μM. A comparison to other hits showed that the configuration within the 8-membered ring (C5 and C6) had a significant contribution to the activity, while the exocyclic stereocenter (C2) was inconsequential. For example, the (2*S*,5*S*,6*R*)-diastereomer of the most potent hit displayed an activity of <15% in the inhibition of apoptosis (Figure 11.18). Additionally, the bulky naphthyl urea was necessary for activity, as several analogs with smaller phenyl rings were found to be less potent.

Based on the SAR and SSAR that could be gained from the HTS, analog BRD0476 (Figure 11.18, B) was found to be the most potent suppressor of β-cell apoptosis with an IC_{50} of 1.66 μM, and it restored β-cell activity to 71 percent function. Mechanism of action studies of the optimized hit were challenging, due to the low aqueous solubility of the compound. Analogs, including a derivative with a methyl group on the urea (Figure 11.18, C), were synthesized to modify the dihedral angle between the urea and naphthyl groups. While the methyl urea analog was more soluble, its activity was only 15 percent, implying that the bulky naphthyl group was in the desired position for activity when the unsubstituted urea was used. The incorporation of a quinoline in place of the naphthyl group solved this problem, in that it allowed for the necessary steric bulk but also added a H-bonding group to increase solubility (Figure 11.18, D). Mechanism of action studies of this compound demonstrated potent inhibition of the STAT1 signal transduction induced by IFN-γ [90].

The utility of DOS in HTS screening is clearly demonstrated by this example. The information that could be gained by having so many different analogs of a core structure with all stereoisomers included allowed for the rapid determination of functional groups that were important for activity. Follow-up chemistry could focus on a single diastereomer to allow for an advanced probe in a short period of time.

2. *Molecular Probes of Heat Shock Protein 70 (Hsp70) Derived from a Focused Library*

The heat shock protein 70 (Hsp70) family forms a group of molecular chaperones that assist in the folding of nascent proteins. It also targets misfolded proteins for degradation and transport across biological membranes [91]. Hsp70s are antiapoptotic chaperones and are therefore implicated in cancer by acting at multiple points along the apoptotic pathway. Small molecules that modulate Hsp70 activity could find therapeutic applications as antitumor and antiviral agents as well as antibiotics, since viruses and bacteria also rely on the chaperones for survival. MAL3-101 [92] is an Hsp70 inhibitor that was prepared in the University of Pittsburgh Center for Chemical Methodologies and Library Development (UPCMLD) as part of a focused library based on the known Hsp70 modulator NSC 630668-R/1 [93] (Figure 11.19). MAL3-101 has fewer rotatable bonds than NSC 630668-R/1, but when the minimum energy conformations of the two compounds were superimposed via molecular modeling, the two structures overlaid well. More importantly, MAL3-101 selectively inhibited J-chaperone-stimulated Hsp70 ATPase activity in a concentration-dependent manner [92].

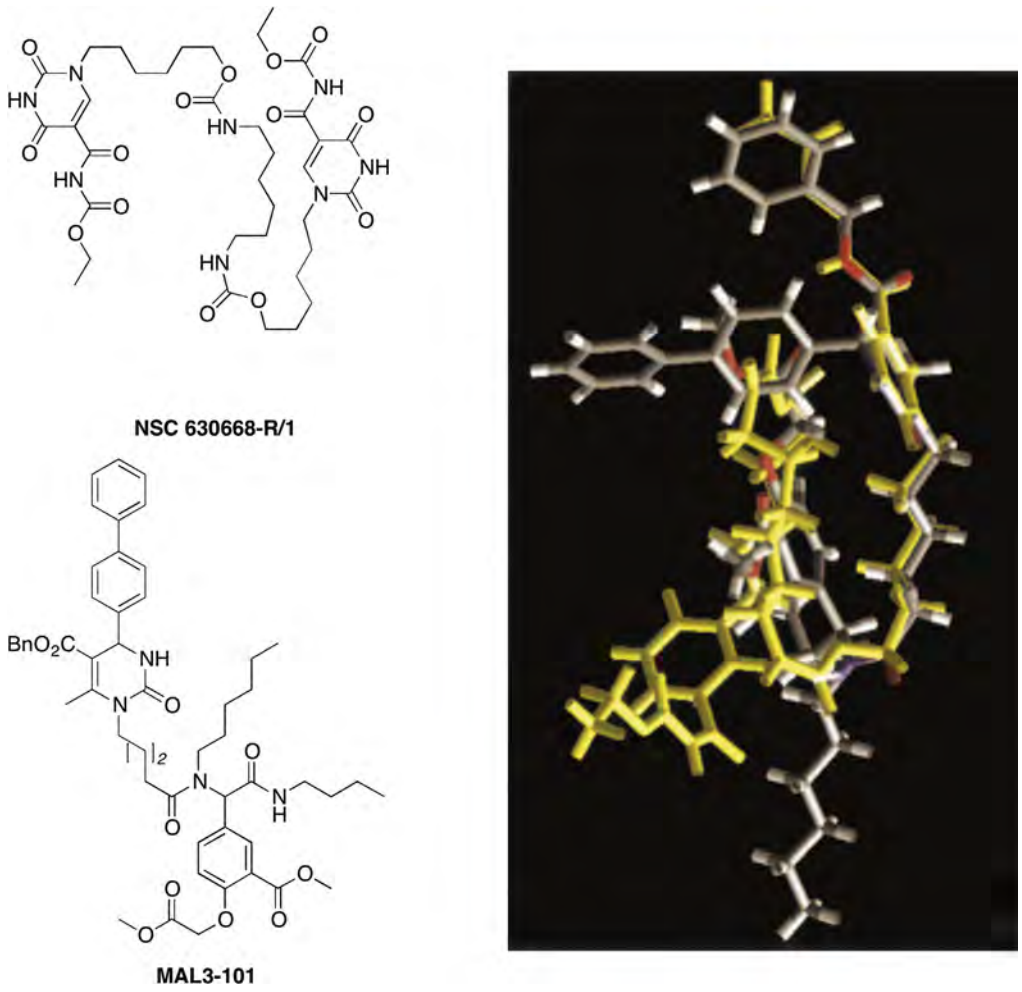


FIGURE 11.19 Superimposed structures of NSC 630668-R/1 and MAL3-101 after lowest energy conformer minimization. *This research was originally published in the Journal of Biological Chemistry: Fewell SW, Smith CM, Lyon MA, Dumitrescu TP, Wipf P, Day BW, et al. Small molecule modulators of endogenous and co-chaperone-stimulated Hsp70 ATPase activity. J Biol Chem 2004;279:51131–40. Copyright the American Society for Biochemistry and Molecular Biology.*

This library design relied on the use of two sequential multi-component reactions to synthesize efficiently analogs with seven points of diversification for screening purposes. The dihydropyrimidinones were synthesized in a one pot Biginelli reaction from the desired aldehyde, urea, and β-ketoester (Figure 11.20A), providing four points of diversification. An Ugi multi-component reaction then incorporated the appropriate amine, aldehyde, and alkyl cyanide building blocks, and contributed three additional points of diversification (Figure 11.20B).

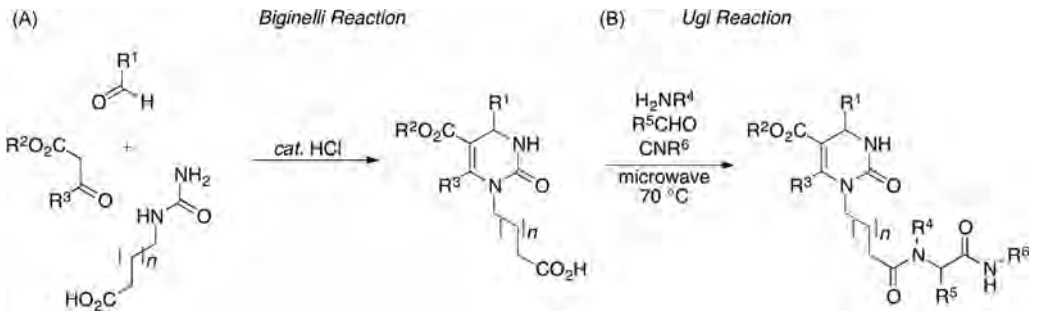


FIGURE 11.20 Diversity-oriented synthesis of the MAL3 library of Hsp70 probe molecules [94].

Following the successful discovery of MAL3-101, a second-generation library was synthesized using the same tandem Biginelli-Ugi reaction sequence [94]. In the first-generation library, it had been determined that a 5-carbon linker was optimal for Hsp70 inhibition (Figure 11.20, n = 2), and the importance of the biphenyl group in the R^1 position, the *n*-hexyl chain in the R^4 position, and the 5-formyl-2-methoxycarbonylmethoxybenzoic acid methyl ester at R^5 (Figure 11.20) were studied. This second-generation library led to the discovery of additional modulators of Hsp70 activity with diverging effects (Figure 11.21). Specifically, while MAL3-101 was an effective antimalarial, antiproliferative, and antitrypanosome agent, it had no effect on SV40. DMT3088 demonstrated improved antiproliferative activity over the first-generation analogs, and MAL2-11B showed SV40 activity (Figure 11.21), acting by inhibiting viral DNA synthesis. Interestingly, MAL2-213 had antimalarial activity but no antiproliferative activity [95]. The rapid access to conformationally more restricted NSC630668-R/1 analogs via a DOS strategy therefore allowed for the discovery of useful probes of Hsp70 with differing activities in cell proliferation, viral, and malaria assays.

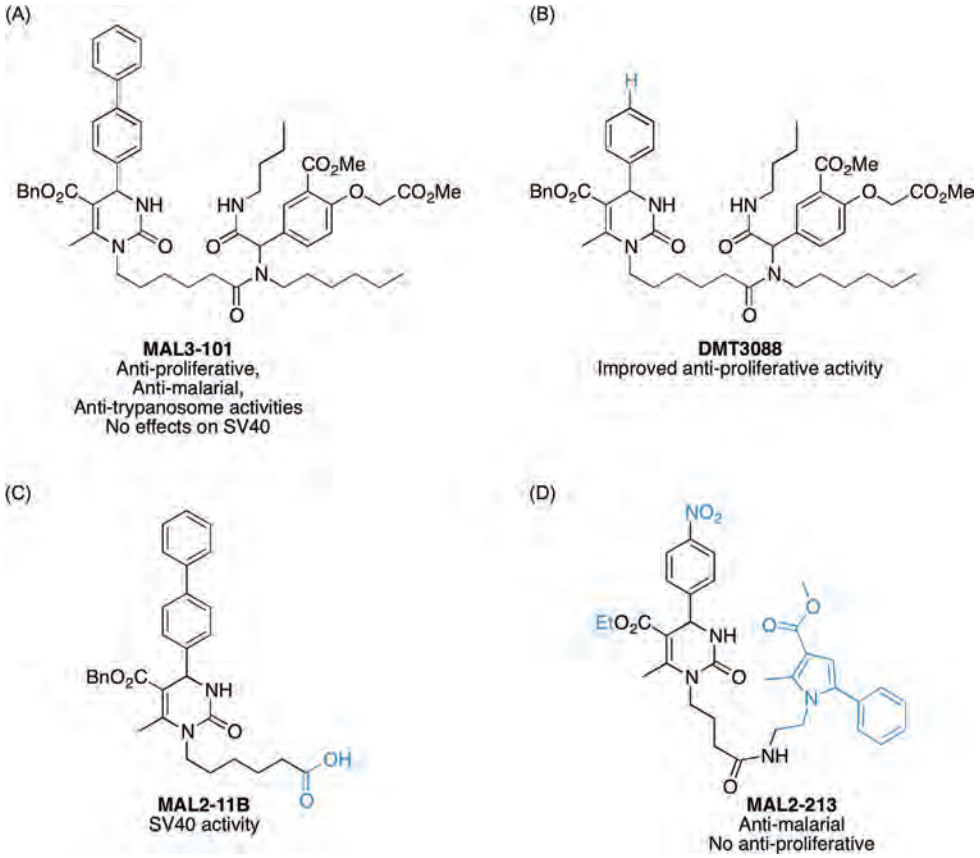


FIGURE 11.21 Differing activities of members of a focused Hsp70 library [95].

III. SUMMARY AND OUTLOOK

In biological processes, noncovalent interactions are orchestrated by complex thermodynamic principles. If a lead compound has been identified, one strategy for further optimization is to introduce structural elements that reduce the conformational flexibility, and, if possible, pre-organize the ligand in a conformation complementary to the bioreceptor. As a consequence, the energy penalty (entropic factor) associated with the binding can be diminished with respect to the flexible parent compound. This thermodynamic advantage can be leveraged into an improved value of the affinity constant. This valuable medicinal chemistry strategy based on thermodynamic considerations remains rather empirical in nature but has been successful, mainly for the optimization of preliminary pharmacophore models and the test of pharmacophore working hypotheses. Furthermore, selected structural modifications of the ligand scaffold may improve both potency and selectivity. However, the rigidification of flexible ligands without loss of potency is challenging, and sophisticated chemistry is needed to achieve the syntheses of constrained compounds with complex architectures. In this respect, the use of chemistry tools such as annulation of substituents and DOS will continue to have an important impact.

References

[1] Jencks WP. On the attribution and additivity of binding energies. Proc Natl Acad Sci 1981;78:4046–50.
[2] Raffa RB, Porreca F. Thermodynamic analysis of the drug-receptor interaction. Life Sci 1989;44:245–58.
[3] Andrews PR, Craik DJ, Martin JL. Functional group contributions to drug-receptor interactions. J Med Chem 1984;27:1648–57.
[4] Page MI. Entropy, binding energy, and enzymic catalysis. Angew Chem Int Ed Engl 1977;16:449–59.
[5] Bardi JS, Luque I, Freire E. Structure-based thermodynamic analysis of HIV-1 protease inhibitors. Biochemistry 1997;36:6588–96.
[6] Velazquez-Campoy A, Todd MJ, Freire E. HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. Biochemistry 2000;39:2201–7.
[7] Searle MS, Williams DH. The cost of conformational order: entropy changes in molecular associations. J Am Chem Soc 1992;114:10690–7.
[8] Williams DH, Westwell MS. Aspects of weak interactions. Chem Soc Rev 1998;27:57–64.
[9] Williams DH, Stephens E, O'Brien DP, Zhou M. Understanding noncovalent interactions: Ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. Angew Chem Int Ed 2004;43:6596–616.
[10] Böhm H-J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. J Comput Aided Mol Des 1994;8:243–56.
[11] Dunitz JD. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. Chem Biol 1995;2:709–12.
[12] Gilli P, Ferretti V, Gilli G, Borea PA. Enthalpy-entropy compensation in drug-receptor binding. J Phys Chem 1994;98:1515–18.
[13] Velazquez-Campoy A, Luque I, Freire E. The application of thermodynamic methods in drug design. Thermochim Acta 2001;380:217–27.
[14] Weiland GA, Minneman KP, Molinoff PB. Fundamental difference between the molecular interactions of agonists and antagonists with the beta-adrenergic receptor. Nature 1979;281:114–17.
[15] Miklavc A, Kocjan D, Mavri J, Koller J, Hadzi D. On the fundamental difference in the thermodynamics of agonist and antagonist interactions with β-adrenergic receptors and the mechanism of entropy-driven binding. Biochem Pharmacol 1990;40:663–9.
[16] Borea PA, Dalpiaz A, Varani K, Guerra L, Gilli G. Binding thermodynamics of adenosine A2a receptor ligands. Biochem Pharmacol 1995;49:461–9.
[17] Davis AM, Teague SJ. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. Angew Chem Int Ed 1999;38:736–49.
[18] Hruby VJ, Li G, Haskell-Luevano C, Shenderovich M. Design of peptides, proteins, and peptidomimetics in chi space. Peptide Sci 1997;43:219–66.
[19] Gardiner J, Abell AD. Synthesis and solid-state conformation of phenylalanine mimetics constrained in a proline-like conformation. Org Biomol Chem 2004;2:2365–70.
[20] Gibson SE, Jones JO, McCague R, Tozer MJ, Whitcombe NJ. A bromoarene based approach to phenylalanine analogues Hic and Nic. Synlett 1999;954–6.
[21] Hanessian S, Bernstein N, Yang R-Y, Maguire R. Asymmetric synthesis of l-azetidine-2-carboxylic acid and 3-substituted congener–conformationally constrained analogs of phenylalanine, naphthylalanine, and leucine. Bioorg Med Chem Lett 1999;9:1437–42.
[22] Cativiela C, Lasa M, López P. Synthesis of enantiomerically pure cis- and trans-cyclopentane analogues of phenylalanine. Tetrahedron Asymmetry 2005;16:2613–23.
[23] Pujadas G, Vaque M, Ardevol A, Blade C, Salvado MJ, Blay M, et al. Protein–ligand docking: a review of recent advances and future perspectives. Curr Pharm Anal 2008;4:1–19.
[24] Kim R, Skolnick J. Assessment of programs for ligand binding affinity prediction. J Comput Chem 2008;29:1316–31.
[25] Gao C, Park M-S, Stern HA. Accounting for ligand conformational restriction in calculations of protein-ligand binding affinities. Biophys J 2010;98:901–10.
[26] Nicklaus MC, Wang S, Driscoll JS, Milne GWA. Conformational changes of small molecules binding to proteins. Bioorg Med Chem 1995;3:411–28.
[27] Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45:2615–23.
[28] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23:3–25.

[29] Varma MVS, Obach RS, Rotter C, Miller HR, Chang G, Steyn SJ, et al. Physicochemical space for optimum oral bioavailability: Contribution of human intestinal absorption and first-pass elimination. J Med Chem 2010;53:1098–108.
[30] Zimmerman DM, Leander JD, Cantrell BE, Reel JK, Snoddy J, Mendelsohn LG, et al. Structure-activity relationships of trans-3,4-dimethyl-4-(3-hydroxyphenyl)piperidine antagonists for μ- and κ-opioid receptors. J Med Chem 1993;36:2833–41.
[31] Zimmerman DM, Nickander R, Horng JS, Wong DT. New structural concepts for narcotic antagonists defined in a 4-phenylpiperidine series. Nature 1978;275:332–4.
[32] Le Bourdonnec B, Goodman AJ, Michaut M, Ye H-F, Graczyk TM, Belanger S, et al. Elucidation of the bioactive conformation of the n-substituted trans-3,4-dimethyl-4-(3-hydroxyphenyl)piperidine class of μ-opioid receptor antagonists. J Med Chem 2006;49:7278–89.
[33] Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, et al. Crystal structure of the μ-opioid receptor bound to a morphinan antagonist. Nature 2012;485:321–6.
[34] Ōki M. Recent advances in atropisomerism N.p. In: Allinger NL, Eliel EL, Wilen SH, editors. Topics in stereochemistry, vol. 14. John Wiley & Sons, Inc; 1983. p. 1–81
[35] Christie GH, Kenner J. LXXI: the molecular configurations of polynuclear aromatic compounds. Part I. The resolution of γ-6 : 6′-dinitro- and 4 : 6 : 4′ : 6′-tetranitro-diphenic acids into optically active components. J Chem Soc Trans 1922;121:614–20.
[36] LaPlante SR, Edwards PJ, Fader LD, Jakalian A, Hucke O. Revealing atropisomer axial chirality in drug discovery. Chem Med Chem 2011;6:505–13.
[37] LaPlante SR, Fader LD, Fandrick KR, Fandrick DR, Hucke O, Kemper R, et al. Assessing atropisomer axial chirality in drug discovery and development: natural products and drugs. Chirality 2011;25:265–74.
[38] Zask A, Murphy J, Ellestad GA. Biological stereoselectivity of atropisomeric natural products and drugs. Chirality 2013;25:265–74.
[39] Food and Drug Administration. Policy statement on the development of new stereoisomeric drugs. Available from: <http://www.fda.gov/drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm122883.htm>; 2011 [accessed 13.10.20].
[40] Olshansky SJ, Passaro DJ, Hershow RC, Layden J, Carnes BA, Brody J, et al. A potential decline in life expectancy in the United States in the 21st century. N Engl J Med 2005;352:1138–45.
[41] Ohki-Hamazaki H, Watase K, Yamamoto K, Ogura H, Yamano M, Yamada K, et al. Mice lacking bombesin receptor subtype-3 develop metabolic defects and obesity. Nature 1997;390:165–9.
[42] Ladenheim EE, Hamilton NL, Behles RR, Bi S, Hampton LL, Battey JF, et al. Factors contributing to obesity in bombesin receptor subtype-3-deficient mice. Endocrinology 2008;149:971–8.
[43] Guan X-M, Chen H, Dobbelaar PH, Dong Y, Fong TM, Gagen K, et al. Regulation of energy homeostasis by bombesin receptor subtype-3: selective receptor agonists for the treatment of obesity. Cell Metab 2010;11:101–12.
[44] Welch CJ, Gong X, Schafer W, Chobanian H, Lin L, Biba M, et al. Factors influencing the interconversion of a new class of dibenzodiazepine sulfonamide atropisomers. Chirality 2009;21:E105–9.
[45] Liu P, Lanza TJ, Chioda M, Jones C, Chobanian HR, Guo Y, et al. Discovery of benzodiazepine sulfonamide-based bombesin receptor subtype 3 agonists their unusual chirality. ACS Med Chem Lett 2011;2:933–7.
[46] Chobanian HR, Guo Y, Liu P, Chioda M, Lanza TJ, Chang L, et al. Discovery of MK-7725, a potent, selective bombesin receptor subtype-3 agonist for the treatment of obesity. ACS Med Chem Lett 2012;3:252–6.
[47] Sugane T, Tobe T, Hamaguchi W, Shimada I, Maeno K, Miyata J, et al. Atropisomeric 4-phenyl-4H-1,2,4-triazoles as selective glycine transporter 1 inhibitors. J Med Chem 2013;56:5744–56.
[48] Bröer S, Cavanaugh JA, Rasko JEJ. Neutral amino acid transport in epithelial cells and its malfunction in Hartnup disorder. Biochem Soc Trans 2005;33:233–6.
[49] Sur C, Kinney GG. Glycine transporter 1 inhibitors and modulation of NMDA receptor-mediated excitatory neurotransmission. Curr Drug Targets 2007;8:643–9.
[50] Danysz W, Parsons CG. Glycine and N-methyl-D-aspartate receptors: physiological significance and possible therapeutic applications. Pharmacol Rev 1998;50:597–664.
[51] Sugane T, Tobe T, Hamaguchi W, Shimada I, Maeno K, Miyata J, et al. Synthesis and biological evaluation of 3-biphenyl-4-yl-4-phenyl-4H-1,2,4-triazoles as novel glycine transporter 1 inhibitors. J Med Chem 2011;54:387–91.
[52] Gause GF, Brazhnikova MG. Gramicidin S. and its use in the treatment of infected wounds. Nature 1944;154:703.
[53] Stern A, Gibbons WA, Craig LC. A conformational analysis of gramicidin S-A by nuclear magnetic resonance. Proc Natl Acad Sci 1968;61:734–41.
[54] Yamada K, Unno M, Kobayashi K, Oku H, Yamamura H, Araki S, et al. Stereochemistry of protected ornithine side chains of gramicidin S derivatives: X-ray crystal structure of the bis-Boc-tetra-N-methyl derivative of gramicidin S. J Am Chem Soc 2002;124:12684–8.
[55] Xiao J, Weisblum B, Wipf P. Electrostatic versus steric effects in peptidomimicry: synthesis and secondary structure analysis of gramicidin S analogues with (E)-alkene peptide isosteres. J Am Chem Soc 2005;127:5742–3.
[56] Grotenbreg GM, Kronemeijer M, Timmer MSM, El Oualid F, van Well RM, Verdoes M, et al. A practical synthesis of gramicidin S and sugar amino acid containing analogues. J Org Chem 2004;69:7851–9.
[57] Wipf P, Xiao J, Stephenson CRJ. Peptide-like molecules (plms): a journey from peptide bond isosteres to gramicidin S mimetics and mitochondrial targeting agents. Chimia 2009;63:764–75.
[58] Skoda EM, Davis GC, Wipf P. Allylic amines as key building blocks in the synthesis of (E)-alkene peptide isosteres. Org Process Res Develop 2012;16:26–34.
[59] Frantz M-C, Pierce JG, Pierce JM, Kangying L, Qingwei W, Johnson M, et al. Large-scale asymmetric synthesis of the bioprotective agent JP4-039 and analogs. Org Lett 2011;13:2318–21.
[60] Xiao J, Weisblum B, Wipf P. Trisubstituted (E)-alkene dipeptide isosteres as β-turn promoters in the gramicidin S cyclodecapeptide scaffold. Org Lett 2006;8:4731–4.
[61] Fink MP, Macias CA, Xiao J, Tyurina YY, Jiang J, Belikova N, et al. Hemigramicidin-TEMPO conjugates: novel mitochondria-targeted anti-oxidants. Biochem Pharmacol 2007;74:801–9.

[62] Tyurina YY, Tyurin VA, Kaynar AM, Kapralova VI, Wasserloos K, Li J, et al. Oxidative lipidomics of hyperoxic acute lung injury: mass spectrometric characterization of cardiolipin and phosphatidylserine peroxidation. Am J Physiol 2010;299:L73–85.
[63] Ji J, Kline AE, Amoscato A, Samhan-Arias AK, Sparvero LJ, Tyurin VA, et al. Lipidomics identifies cardiolipin oxidation as a mitochondrial target for redox therapy of brain injury. Nat Neurosci 2012;15:1407–13.
[64] Nesbitt KM, Jaquins-Gerstl A, Skoda EM, Wipf P, Michael AC. Pharmacological mitigation of tissue damage during brain microdialysis. Analyt Chem 2013;85:8173–9.
[65] Xun Z, Rivera-Sánchez S, Ayala-Peña S, Lim J, Budworth H, Skoda EM, et al. Targeting of XJB-5-131 to mitochondria suppresses oxidative DNA damage and motor decline in a mouse model of Huntington's disease. Cell Rep 2012;2:1137–42.
[66] Frantz M-C, Skoda EM, Sacher JR, Epperly MW, Goff JP, Greenberger JS, et al. Synthesis of analogs of the radiation mitigator JP4-039 and visualization of BODIPY derivatives in mitochondria. Org Biomol Chem 2013;11:4147–53.
[67] Barreiro EJ, Kümmerle AE, Fraga CAM. The methylation effect in medicinal chemistry. Chem Rev 2011;111:5215–46.
[68] Leung CS, Leung SSF, Tirado-Rives J, Jorgensen WL. Methyl effects on protein–ligand binding. J Med Chem 2012;55:4489–500.
[69] Némethy G. Hydrophobe wechselwirkungen [Hydrophobic interactions]. Angew Chem 1967;79:260–71 [in German].
[70] Southall NT, Dill KA, Haymet ADJ. A view of the hydrophobic effect. J Phys Chem B 2001;106:521–33.
[71] Schönherr H, Cernak T. Profound methyl effects in drug discovery and a call for new C-H methylation reactions. Angew Chem Int Ed 2013;52:12256–67.
[72] Quancard J, Bollbuck B, Janser P, Angst D, Berst F, Buehlmayer P, et al. A potent and selective S1P1 antagonist with efficacy in experimental autoimmune encephalomyelitis. Chem Biol 2012;19:1142–51.
[73] Goodman AJ, Le Bourdonnec B, Dolle RE. Mu opioid receptor antagonists: recent developments. Chem Med Chem 2007;2:1552–70.
[74] Jenkins GM, Frohman MA. Phospholipase D: a lipid centric review. Cell Mol Life Sci 2005;62:2305–16.
[75] Elvers M, Stegner D, Hagedorn I, Kleinschnitz C, Braun A, Kuijpers MEJ, et al. Impaired αIIbβ3 integrin activation and shear-dependent thrombus formation in mice lacking phospholipase D1. Sci Signal 2010;3:1–10.
[76] Oliveira TG, Chan RB, Tian H, Laredo M, Shui G, Staniszewski A, et al. Phospholipase D2 ablation ameliorates Alzheimer's disease-linked synaptic dysfunction and cognitive deficits. J Neurosci 2010;30:16419–28.
[77] O'Reilly MC, Scott SA, Brown KA, Oguin TH, Thomas PG, Daniels JS, et al. Development of dual PLD1/2 and PLD 2 selective inhibitors from a common 1,3,8-triazaspiro[4.5]decane core: discovery of ML298 and ML299 that decrease invasive migration in U87-MG glioblastoma cells. J Med Chem 2013;56:2695–9.
[78] Czekaj M, Klein SI, Guertin KR, Gardner CJ, Zulli AL, Pauls HW, et al. Optimization of the β-aminoester class of factor Xa inhibitors. Part 1: P4 and side-chain modifications for improved in vitro potency. Bioorg Med Chem Lett 2002;12:1667–70.
[79] Schärfer C, Schulz-Gasch T, Ehrlich H-C, Guba W, Rarey M, Stahl M. Torsion angle preferences in druglike chemical space: a comprehensive guide. J Med Chem 2013;56:2016–28.
[80] Engelman JA, Luo J, Cantley LC. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. Nat Rev Genet 2006;7:606–19.
[81] Fontana L, Partridge L, Longo VD. Extending healthy life span-from yeast to humans. Science 2010;328:321–6.
[82] Liu KKC, Huang X, Bagrodia S, Chen JH, Greasley S, Cheng H, et al. Quinazolines with intra-molecular hydrogen bonding scaffold (iMHBS) as PI3K/mTOR dual inhibitors. Bioorg Med Chem Lett 2011;21:1270–4.
[83] LaPlante SR, Gillard JR, Jakalian A, Aubry N, Coulombe R, Brochu C, et al. Importance of ligand bioactive conformation in the discovery of potent indole-diamide inhibitors of the Hepatitis C virus NS5B. J Am Chem Soc 2010;132:15204–12.
[84] Spring DR. Diversity-oriented synthesis: a challenge for synthetic chemists. Org Biomol Chem 2003;1:3867–70.
[85] Virshup AM, Contreras-García J, Wipf P, Yang W, Beratan DN. Stochastic voyages into uncharted chemical space produce a representative library of all possible drug-like compounds. J Am Chem Soc 2013;135:7296–303.
[86] Fink T, Bruggesser H, Reymond J-L. Virtual exploration of the small-molecule chemical universe below 160 daltons. Angew Chem Int Ed 2005;44:1504–8.
[87] Cnop M, Welsh N, Jonas J-C, Jörns A, Lenzen S, Eizirik DL. Mechanisms of pancreatic β-cell death in type 1 and type 2 diabetes: many differences, few similarities. Diabetes 2005;54:S97–107.
[88] Chou DH-C, Bodycombe NE, Carrinski HA, Lewis TA, Clemons PA, Schreiber SL, et al. Small-molecule suppressors of cytokine-induced β-cell apoptosis. ACS Chem Biol 2010;5:729–34.
[89] Chou DH-C, Duvall JR, Gerard B, Liu H, Pandya BA, Suh B-C, et al. Synthesis of a novel suppressor of β-cell apoptosis via diversity-oriented synthesis. ACS Med Chem Lett 2011;2:698–702.
[90] Scully SS, Tang AJ, Lundh M, Mosher CM, Perkins KM, Wagner BK. Small-molecule inhibitors of cytokine-mediated STAT1 signal transduction in β-cells with improved aqueous solubility. J Med Chem 2013;56:4125–9.
[91] Hartl FU. Molecular chaperones in cellular protein folding. Nature 1996;381:571–80.
[92] Fewell SW, Smith CM, Lyon MA, Dumitrescu TP, Wipf P, Day BW, et al. Small molecule modulators of endogenous and co-chaperone-stimulated Hsp70 ATPase activity. J Biol Chem 2004;279:51131–40.
[93] Fewell SW, Day BW, Brodsky JL. Identification of an inhibitor of Hsc70-mediated protein translocation and ATP hydrolysis. J Biol Chem 2001;276:910–14.
[94] Werner S, Turner DM, Lyon MA, Huryn DM, Wipf P. A focused library of tetrahydropyrimidinone amides via a tandem Biginelli-Ugi multi-component process. Synlett 2006;14:2334–8.
[95] Huryn DM, Brodsky JL, Brummond KM, Chambers PG, Eyer B, Ireland AW, et al. Chemical methodology as a source of small-molecule checkpoint inhibitors and heat shock protein 70 (Hsp70) modulators. Proc Natl Acad Sci 2011;108:6757–62.

CHAPTER

12

Application Strategies for the Primary Structure–Activity Relationship Exploration

Camille G. Wermuth[1], *Serge Grisoni*[2], *Bruno Villoutreix*[3] *and Jean-Philippe Rocher*[4]

[1]Founder of Prestwick Chemical, Strasbourg, France; [2]Pierre Fabre Research Institute, Toulouse, Cédex, France; [3]Université Paris Diderot, Sorbonne Paris Cité, Inserm, Paris, France; [4]Pierre Fabre Research Institute, Castres, France

OUTLINE

Le bon sens est la chose du monde la mieux partagée:
car chacun pense en être si bien pourvu, que ceux même qui
sont les plus difficiles à contenter en tout autre chose,
n'ont point coutume d'en désirer plus qu'ils n'en ont.

Common sense is the worldly thing which is the best shared:
as each of us thinks to be so well provided with,
that even those who are the most difficult to satisfy in any other thing,
don't want to desire more of it than they already have.

René Descartes (1596–1650) [1]

I. INTRODUCTION

When confronted with a new lead structure or when needing to enlarge the protection perimeter around newly discovered structures for patent reasons, the medicinal chemist may be daunted by the immensity of the task. The possibilities of molecular variations around the lead structure are immense, and *a priori* the synthesis of several thousand potential analogs can be envisaged. The aim of this chapter is to provide some guidelines and strategies rendering easier and more efficacious the decision on which compounds to prepare and which ones to reject. The proposed guidelines derive essentially from common-sense reasons, a feature that may explain why they are often forgotten. In addition to his personal experience, the author was inspired by the articles of Messer [2], Cavalla [3], Craig [4], and Austel [5]. The guidelines will also refer to the multiparametric principles of hit optimization [6].

II. PRELIMINARY CONSIDERATIONS

Before considering the different possibilities of molecular variation presented in the previous chapters (e.g., homology, bioisosterism, conformational restriction, optical isomerism, ring system modifications, and synthesis of twin drugs), one has to decide what kind of general strategy should be applied. Depending on the lead structure's size and its degree of complexity, the strategy may involve a simplification (disjunctive approach), conservation of the same level of complexity (analogical approach), or enlargement through additional elements (conjunctive approach).

Simplification of the original lead compound is especially appropriate for natural substances. This approach, known as the "disjunctive approach" [7] consists of a molecular dissection that deletes functions, structural elements, or cycles. Classical examples of disjunctive approaches are found in the pruning of the acetylcholinesterase inhibitor physostigmine to yield neostigmine and, later on, rivastigmine (Figure 12.1), or the change from somatostatin to a simplified hexapeptide [8].

Other examples of disjunctive approaches are collected in Table 12.1. The main result of the method is the identification of the portions of the molecule that are essential for the expected biological activity and those that are not.

Conservation of the lead compound's degree of complexity proceeds usually through isosteric exchanges or functional inversions and can be considered the "analogical approach" [9] (Table 12.2).

Finally, when additional moieties are grafted to the molecule, one speaks of "conjunctive approaches." [7] They can also consist in the attachment of additional structural elements, as in the association of two separate drugs (associative synthesis; nonsymmetrical twin drugs), the symbiotic approach [10], or the duplication of the parent drug (symmetrical twin drugs).

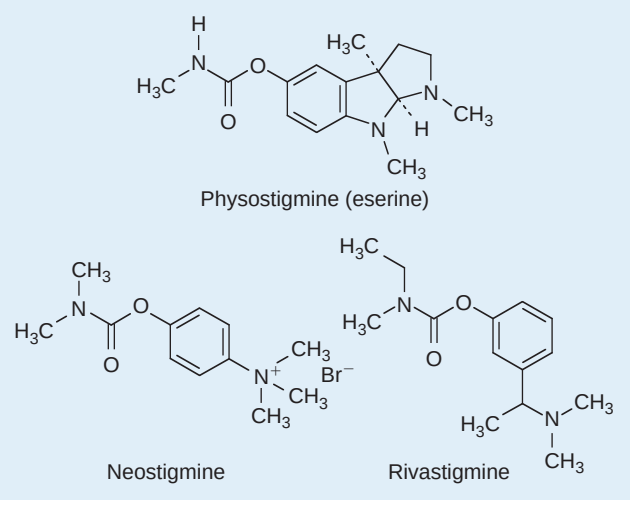


FIGURE 12.1 Neostigmine and rivastigmine are the result of disjunctive approach applied to physostigmine.

The change of the γ-aminobutyric acid (GABA)B receptor agonist CGP 27 492 to the GABAB receptor antagonist CGP 54 062 (Figure 12.2) represents a typical example of a conjunctive approach resulting from the attachment of additional structural elements [11]. A similar case is provided by the design of the H2 receptor agonist impromidine [12].

Examples of drugs resulting from associative synthesis (nonsymmetrical twin drugs) and of duplication of the parent drug (symmetrical twin drugs) are listed in Tables 12.3 and 12.4.

III. HIT OPTIMIZATION STRATEGIES

The strategy of hit optimization will heavily depend on the amount of information available at the start of the study. Particularly, if some knowledge of the 3D structure of the target is available, the synthesis program can take the corresponding information into account immediately. This will also be the case if some earlier structure−activity relationship (SAR) studies are available. However, in the most frequent cases, the target is new and original. In such a situation, all the initial SAR exploration rests in the hands of the medicinal chemist. The purpose of this chapter is to provide them with insight from more than thirty years' experience and to help them in the choice of the most appropriate strategy.

TABLE 12.1 Drugs Resulting from Disjunctive Manipulations

Lead	Derivative
Cocaine	Procaine
Tubocurarine	Decamethonium
Morphine	Morphinanes
	Benzomorphanes
	Phenylpiperidines
Atebrine	Chloroquine
Asperlicine	Benzodiazepine analog
Phylloquinone (vitamin K_1)	Menadione (vitamin K_3)
Triampterene	Amiloride
Cimetidine	Roxatidine
Somatostatine	Simplified peptide
Bothrops jaracaVenin	Teprotide Captopril

TABLE 12.2 Drug Analogs Possessing a Similar Size than the Model Compound

Initial drug	Analog
Chlorpromazine	Thioridazine
Imipramine	Amitryptyline
Propranolol	Pindolol
Furosemide	Bumetamide
Enalapril	Perindopril
Cimetidine	Ranitidine
Pravastatine	Fluindostatine

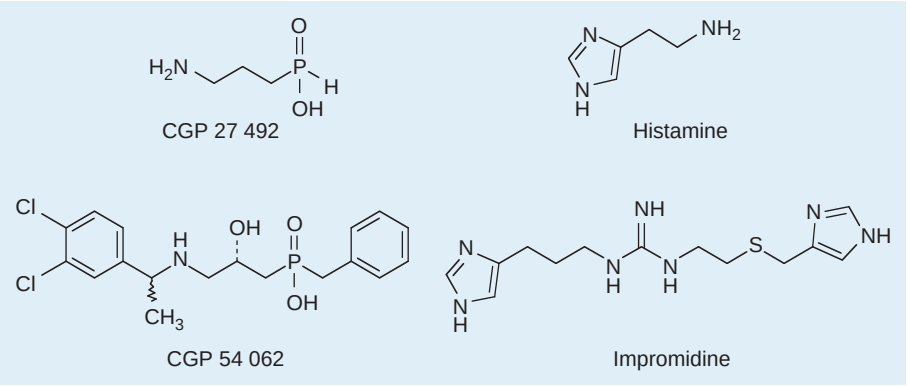


FIGURE 12.2 Conjunctive approach in drug design. The attachment of two benzylic groups and a hydroxyl in S-configuration changes a $GABA_B$ receptor agonist into a $GABA_B$ antagonist [11]. Similarly, the H_2-histaminergic agonist impromidine is the result of a conjunctive approach applied to histamine [12].

TABLE 12.3 Nonsymmetrical Twin Drugs

Drug No. 1	Drug No. 2	Twin drug
Cafeine	Amphetamine	Fenethylline
Aspirine	Paracetamol	Benorylate
Clofibric acid	Nicotinic acid	Etofibrate
Hydrazinopyridazine	β-blocker	Prizidilol
Pindolol	Captopril	BW-B385C [13]

TABLE 12.4 Symmetrical Twin Drugs

Drug	Activity
Bialamicol	Anti-amebic
Ethambucol	Tuberculostatic
Probucol	Antihyperlipoproteinemic
Thiamine disulfide	Vitamin
Dicumarol	Anticoagulant
Netropsin	DNA binding agent
Succinylcholine	Skeletal muscle relaxant

A. Some Information About the Target is Available

The first point to consider is to ascertain if the hit (or the lead) that has to be optimized is relevant to a computer-aided design. If the X-ray 3D structure of the target protein is already described, it becomes possible to match the different candidate molecules with the target structure and eliminate those that evidently are too bulky or possess an inadequate geometry. A similar situation occurs if a non-experimental 3D model of the target is available. For more details on computer-assisted drug design (CADD) see Chapters 6, 21 and 22. An important point that is always to be kept in mind is that CADD has limitations. Thus the two corticotropin-releasing factor (CRF) antagonists of Figure 12.3 are structurally very close positional isomers. However, one of them shows nanomolar affinity for the CRF receptor whereas the other one is practically inactive [14]. No evident CADD explanation is available for this behavior.

The support of experimental data on the ligand–protein interactions, in particular X-ray, is essential for molecular modelling studies, which can propose a unique picture of the protein but can hardly predict any change in the protein consecutive to its interaction with a new ligand. The example of kinase inhibitor models is representative. The first marketed inhibitor, Imatinib (Glivec), interacts with the ATP pocket, which is structurally highly conserved among the kinases and efficiently modelized. However, structural biology reveals another action

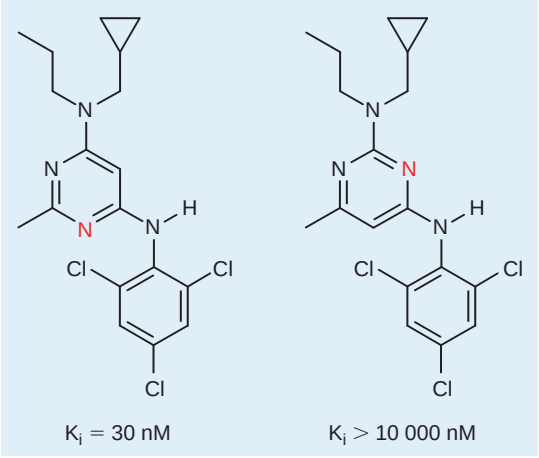


FIGURE 12.3 In a rather complex molecule, the simple shift of a nitrogen atom to another position of the diazine ring atom almost completely abolishes the affinity for the CRF receptor.

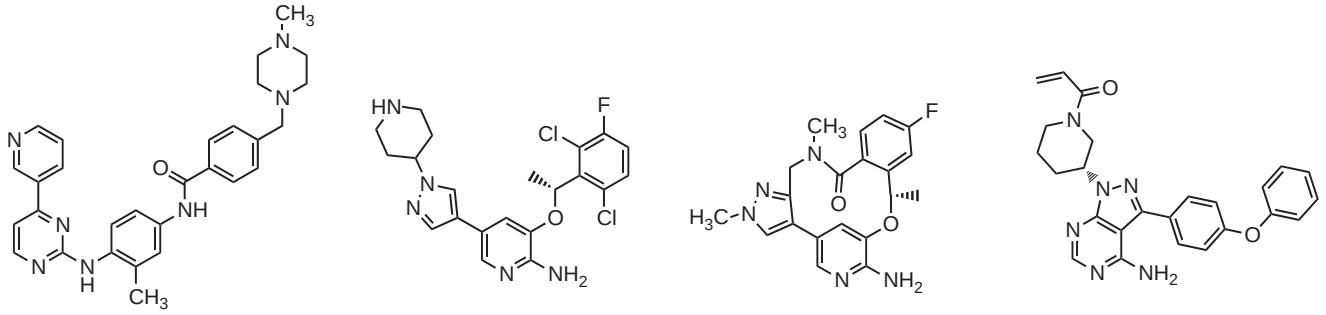


FIGURE 12.4 Molecular modelling and structural biology have helped in the optimizations and design of ATP binding site kinase inhibitors.

mode that corresponds to the inactive form of the kinase [15]. Three types of interactions towards protein kinases have been described that allow the design of inhibitors addressing the selectivity [16], the resistance [17] over mutated kinases, and the design of covalent drugs (Figure 12.4) [18,19]. The contribution of Structure Based Drug Design (SBDD) is highlighted in the discovery of the Crizotinib resistant ALK inhibitor PF-06463922 by using a chemical macrocyclization approach [20].

In the absence of any target structure, some information can be gathered in comparing the hit structure with the endogenous ligand (if known) or with other structures showing affinities for the same target (if known). If several compounds exist that are recognized by the target, it becomes possible to practice the active analog approach with a reference.

B. No Information About the Target is Available

When one is faced with a truly novel target, the only way to procede is to design analogs to identify the molecular features that are favorable and those that are detrimental to the activity. Such molecular variation programs can be performed in a number of ways. Activity against the target is essential, but other qualities of the future drug molecules are relevant for drug optimization. One can mention selectivity, satisfactory ADME (absorption, distribution, metabolism, and excretion) and toxicity profiles, and optimal physicochemical properties (such as chemical stability, water solubility, and the absence of polymorphs). Finally, the compounds must be patentable. Table 12.5, from Baxter et al [21], summarizes the different criteria practiced at the Astra-Zeneca company.

TABLE 12.5 Astra-Zeneca "Generic Lead Target Profile" for Progressing HTS Hits to Leads

Parameter	Values
Potency	100 nM
Molecular weight	<450
C log *P*	<3
Log *D*	<3
Solubility	>10 μg/mL
Clear SAR around the potential lead	
Structure must provide patent opportunities	
Selectivity – Use PanLabs/Cerep batteries	
Need *in vivo* biological validation	
HERG screening	
Early toxicity *in vivo* screening	
Metabolism	
P450 inhibition IC_{50}	10 μM for 5 major isoenzymes
Rat hepatocyte intrinsic clearance	<14 μL/min/mg
Human microsome intrinsic clearance	<23 μL/min/mg
Rat IV clearance	<35 mL/min/mg
Volume	>0.5 L/kg
$t_{1/2}$	>0.5 h
Pharmacokinetics	
Rat PO bioavailability	>10%
Plasma protein binding	<99.5

C. The Predominant Objective is Potency

One of the most fruitful strategies revealing the features associated with high potency consists of what we call the "topological exploration of the lead compound." In this approach, the possible modifications of the molecule are considered from the four cardinal points (the south, the north, the west, and the east) and from the center of the molecule (Figure 12.5).

As an illustration, let us assume that our starting lead compound is the muscarinic antagonist pirenzepine. We can then consider successively the different modification sites.

1. East Side Modifications

In the present example, the east side modifications mainly concern changes at the level of the pyridine ring, and the questions that should be answered are the following ones (Figure 12.6):

- Is the pyridine nitrogen necessary? Can the pyridine ring be replaced by a phenyl ring (a)?
- Can the pyridine nitrogen be displaced to other positions (b, c, d)?
- What is the influence of substituents on the pyridine ring (e)? Can it be substituted by various functions (associated with typical electronic, steric, or lipophilic changes)?
- Can the pyridine ring be changed to other aromatic heterocycles, such as pyridazine (f), pyrimidine (g), pyrazine (h), triazines (i), and thiazoles (j)?

2. North Side Modifications

On the north side of pirenzepine, one can consider the NH and the C = O groups—separately or together—as an amide function (Figure 12.7).

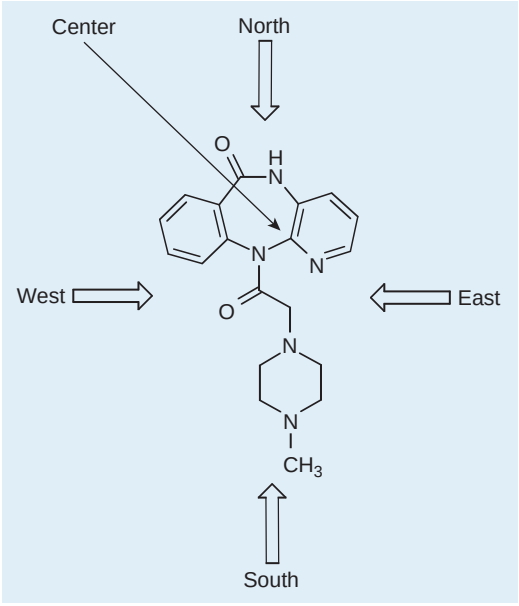


FIGURE 12.5 Scheme of the topological exploration of the pirenzepine molecule.

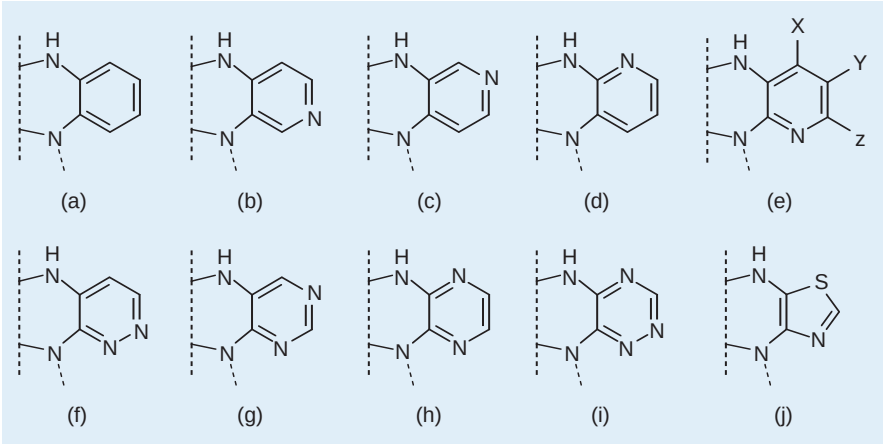


FIGURE 12.6 East side modifications on the pirenzepine molecule.

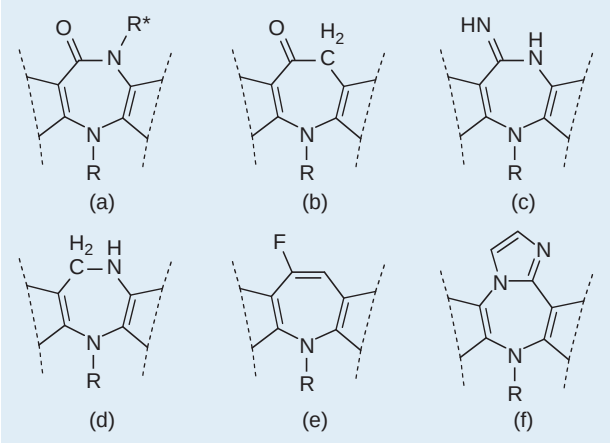


FIGURE 12.7 North side modifications on the pirenzepine molecule.

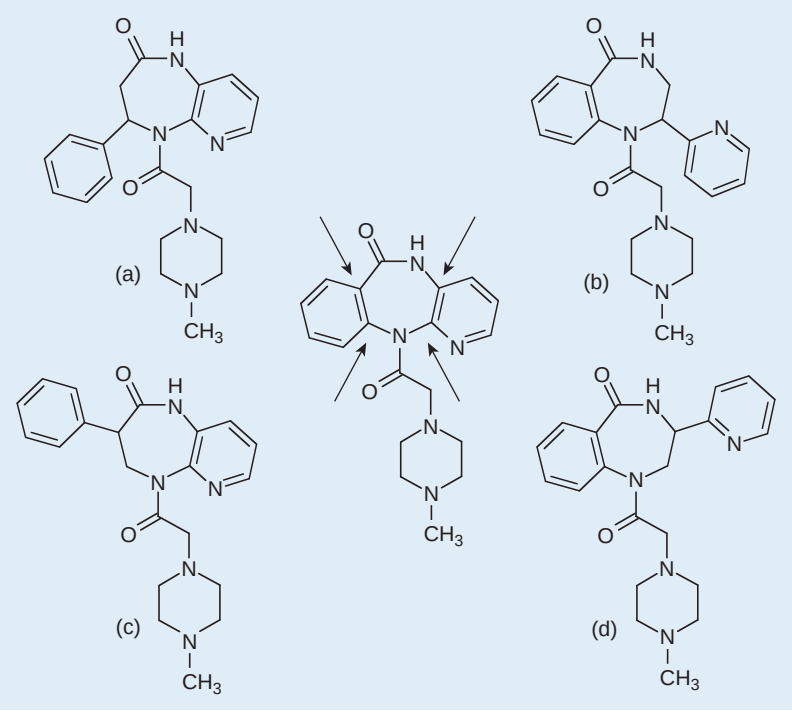


FIGURE 12.8 "Benzo-splitting" applied to the east and the west side of pirenzepine.

- Can the NH group be substituted (a)? (Possible role as hydrogen bond donor).
- Can the NH group be replaced by a CH_2 group (b) or any other possible bioisosteric group?
- Is the carbonyl group necessary? Can it be changed to a CH_2 group (d)?
- Can the amide be replaced by a bioisostere reference?

3. *West Side Modifications*

Due to the almost symmetrical structure of pirenzepine, the same kind of modifications can be applied as those suggested for the east side.

For both east and west sides, benzo-splitting can be considered (Figure 12.8).

4. *South Side Modifications*

These concern the changes made on the basic side-chain. A large number of possibilities exist on different levels:

On the piperazine ring (Figure 12.9):

– Can the *N*-methyl group be replaced by higher alkyl, aralkyl, or aryl groups (a)?
– Can it be replaced by homopiperazine (b), by piperidine (c and c), or by ring-opened diamines (e)?
– Can it be substituted (g and h) or bridged (i)?
– Can it be replaced by some vague bioisosteric equivalent such as a guanidino group (*N*-methyl piperazine was used as guanidine substitute in the design of thrombine inhibitors)?

On the carboxamido function (Figure 12.10):

– Can the carbonyl group be reduced to a CH_2 group (a)?
– Can the carboxamido function be replaced by a carbon–carbon double bond (b)?
– Can it be included in a bioisosteric and constraint ring system (c)?

5. *Center Modifications*

The literature available for the tricyclic psychotropic drugs proposes a number of possible variations that are potentially applicable to the present central diazepinone ring (Figure 12.11). They include ring contraction (a), ring extension (b), ring bridging (c), and changes in the nature and number of the heteroatoms.

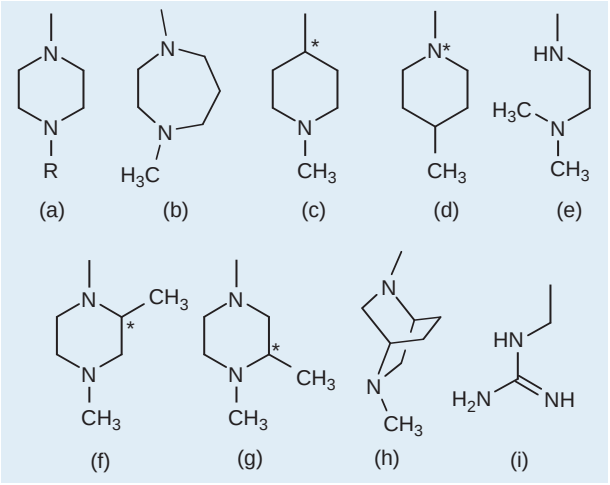


FIGURE 12.9 Variations on the N-methylpiperazine ring.

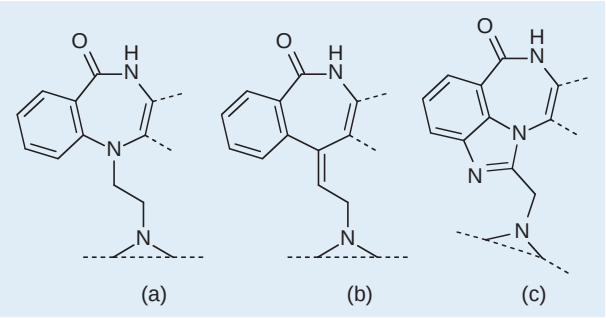


FIGURE 12.10 Variations on the side-chain carboxamido group of pirenzepine.

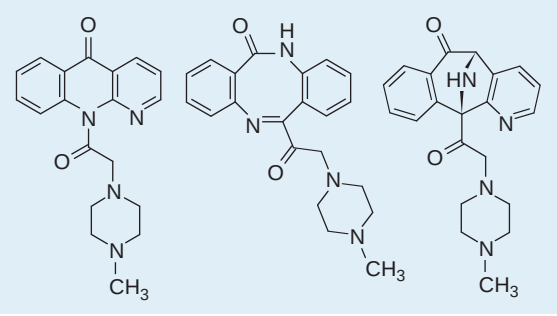


FIGURE 12.11 Modifications of the central ring of pirenzepine.

Taken together, these modifications suggest the synthesis of an impressive number of analogs of a given hit structure and hopefully open the way to some more potent original analogs.

D. The Predominant Objective is the Establishment of SAR

When one deals with a hitherto unknown active compound, he or she wants to acquire information about the atoms and the functionalities involved in the interaction with the target and about the nature of the chemical bonds they achieve in the drug–target complex. In the examples below, the hit compound is imaginary. First, a topological attribution of the functional groups similar to that of the previous paragraph is achieved (Figure 12.12). Second, a case-by-case discussion of each functional feature is undertaken.

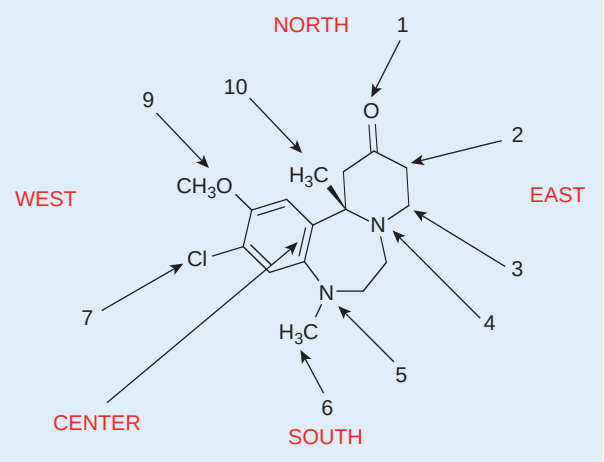


FIGURE 12.12 An imaginary hit compound serves as demonstration molecule. The different sites (1–10) of the molecule are potentially able to influence the drug–receptor interaction.

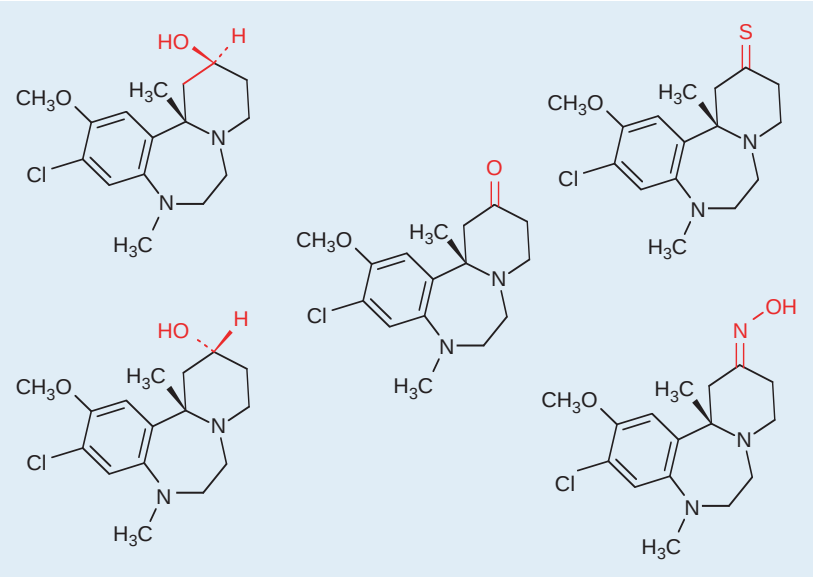


FIGURE 12.13 The ketonic carbonyl function of the hit compound can play an electronic role in forming hydrogen bonds or an architectural role in maintaining the oxygen atom coplanar with nearest three carbon atoms.

If the carbonyl group of the molecule is considered, one can assume that either it interacts by exchanging hydrogen bonds with its target protein or it plays an architectural role. The latter is due to the planar structure imposed by the sp^2 character of its carbon atom. In order to decide between the two alternatives, some key substances have to be prepared and tested (Figure 12.13).

The synthesis of the thiocarbonyl analog leads to a compound with potency similar to that of the carbonyl derivative but which is unable to make hydrogen bonds. If it is still active, it means that hydrogen bonding at this region of the molecule is not necessary for activity. If it is no longer active, the ketonic carbonyl can be kept. The carbonyl group can possibly be replaced by the corresponding ketoxime, a better hydrogen bond former. The reduction of the carbonyl group to the two enantiomeric alcohols transforms the planar sp^2 carbon atom into pyramidal sp^3 carbons. These compounds may be active if the carbonyl group plays a conformational role.

The presence of an α-methylene group permits the enolization of the hit molecule. This tautomery is blocked through the insertion of a spiro cyclopropyl ring (Figure 12.14). The result of formal oxidation yields the secondary alcohol. Double-bond possibilities are α-methylene ($X = CH_2$), the diketone ($X = O$), and the corresponding ketoxime ($X = N\text{-}OH$).

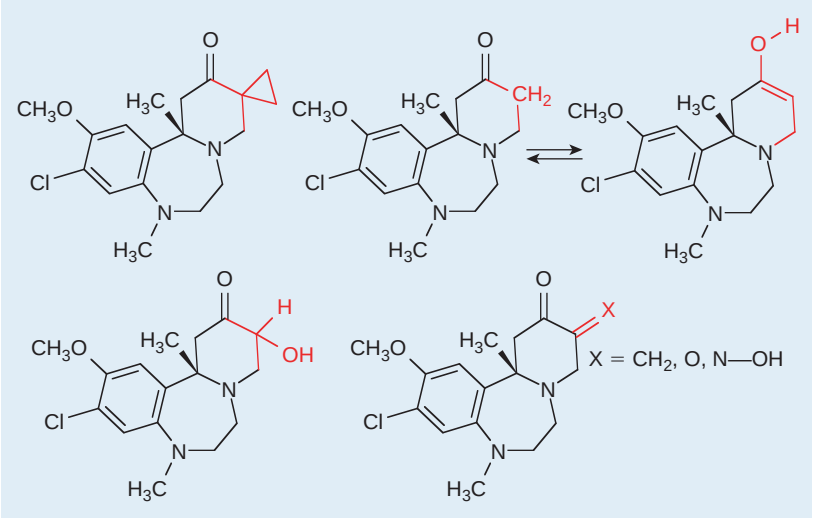


FIGURE 12.14 The α-methylene group allows enolization of the hit compound. This is no longer possible with the α-cyclopropyl analog. Various oxidative functionalities are possible.

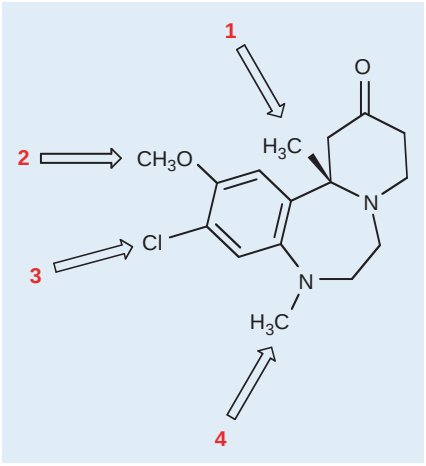


FIGURE 12.15 Investigations about the roles played by the various substituents of the hit molecule.

Further investigations concern the different substituents borne by the hit molecule (Figure 12.15):

1. Can the methyl group be replaced by a hydrogen? By a larger group (e.g., isopropyl)?
2. Can the methoxy substituent be replaced by an isoelectronic substituent or by another hydrogen bond acceptor?
3. What is the role of the chlorine: electronic, lipophilic, or metabolic?
4. What is the nitrogen substituent: desmethyl, higher alkyl, acyl, carbamate, or urea?

E. The Predominant Objective Consists of Analog Design

In the Losartan molecule, the substituted imidazole moiety is attached to the typical tetrazolyl-diphenyl unit (Figure 12.16). In practicing analog synthesis, the Novartis scientists conserved unchanged this latter part of the molecule but tried to prepare a bioisosteric equivalent of the substituted imidazole possessing similar interaction possibilities. The lipophilic *n*-butyl chain was maintained, the CN dipole was replaced by a CO dipole, and the chlorine substituent plus two imidazolic carbon atoms were replaced by an isolipophilic isopropyl group. Finally, because it was known that *in vivo* the primary alcoholic group of Losartan was oxidized into the corresponding carboxylic group, the carboxylic function was directly introduced in Valsartan.

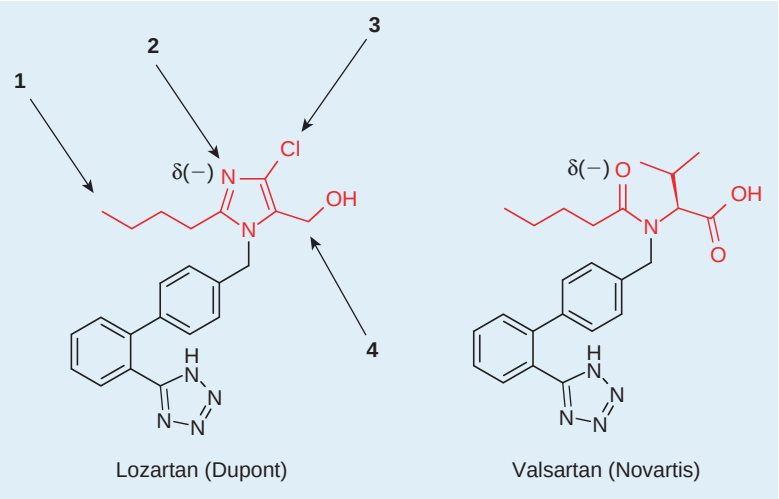


FIGURE 12.16 The angiotensin agent Valsartan results from a rational analog design.

IV. APPLICATION RULES

It would certainly be tiresome and beyond the possibilities of a medicinal chemistry team to be forced to prepare all the compounds imagined by means of "paper" chemistry. The following rules aim to codify precisely the use of all the strategies and to increase their efficacy in establishing priorities and selection rules.

A. Rule Number One: The Minor Modifications Rule

This rule [8] can be defined as the priority given to the design of analogs that are close to the lead structure and that result from only minor changes. Minor changes are achieved by very simple organic reactions such as hydrogenations, hydroxylations, methylations, acetylations, racemate resolutions, changes in substituents, and isosteric replacements. The modification can produce either an increase in potency or an increase in selectivity, or even sometimes the suppression of unwanted toxic or side effects (Table 12.6).

A simple change in the aliphatic chain length can abolish mutagenic properties. Thus, in a series of muscarinic M_1 partial agonists (Figure 12.17), the compound with a dimethylene side-chain is potent (IC_{50} [^{3}H]-pirenzepine = 3 nM) but mutagenic. The higher trimethylenic homolog is less potent (IC_{50} [^{3}H]-pirenzepine = 15 nM) but safe in terms of mutagenicity [22].

In some instances, very slight changes such as hydrogenation or dehydrogenation can induce dramatic changes in the activity profile of drug molecules. Examples are found in the imidazoline I_3 receptor ligands, which act on insulin secretion. The imidazolinic compound Efaroxan acts as an agonist [23], but the corresponding imidazole acts as an antagonist [24] (Figure 12.18). A similar passage is found for the passage from the agonistic benzofuranic compound 2-BFI [25] to its antagonistic dihydro derivative [26].

Despite being supported by prestigious results, the minor modifications rule is largely unrecognized. Making use of ordinary chemistry, it is not always accepted with enthusiasm by organic chemists. The very simple reactions that are involved do not add much to their fame, and they are more fascinated by the challenge of a total synthesis, especially of a natural substance bearing many chiral centers. Seen from a practical point of view, priority has nevertheless to be given to this principle. The simplicity of implementation, especially given the spectacular results that it brings, strongly advocate for the minor modification rule.

B. Rule Number Two: The Biological Logic Rule

The second rule of application rests on the earliest possible utilization of biochemical data. In particular, biological activity may be rationalized if it stems from the chemical or physicochemical properties of the series.

Very general properties can be foreseen insofar as functions or moieties present in the structure can suggest interference with a biological system—for example, hydrazines or hydroxylamines and pyridoxal-containing coenzymes, complexing agents and metallic coenzymes, electron donors or acceptors and oxido-reduction

TABLE 12.6 Minor Modifications

Original compound	Modified compound	Result[a]
Ergotamine	Dihydroergotamine	Increase in potency as α-adrenergic antagonist decrease in toxicity
Chlorothiazide	Hydrochlorothiazide	20-fold increase in potency
Chloroquine	Hydroxychloroquine	Decrease in toxicity
Morphine	Codeine	Change in activity profile (analgesic→antitussive)
Carbachol	Bethanechol	Increase in selectivity (exclusively muscarinic)
Imipramine	Desmethylimipramine	Change in activity profile (noradrenergic→serotonergic)
Tolbutamide	Chlorpropamide	Longer duration of action (5–7 h→24–48 h)
Racemic amphetamine	Dexamphetamine	Less cardiovascular side effects

[a]*Taken from Goodman and Gilman [16].*

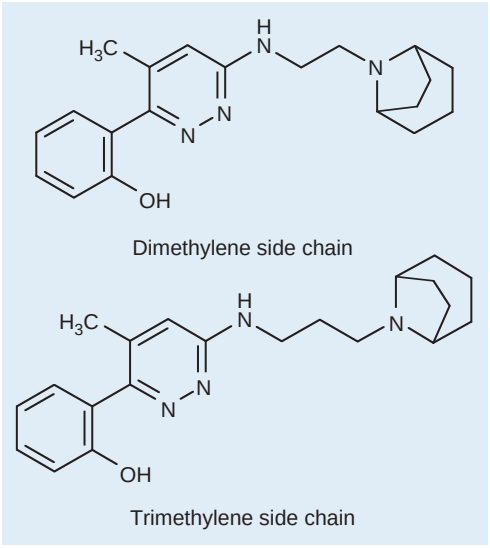


FIGURE 12.17 The mutagenicity of the original dimethylenic muscarinic M_1 partial agonist could be abolished in changing the dimethylene to a trimethylene side-chain.

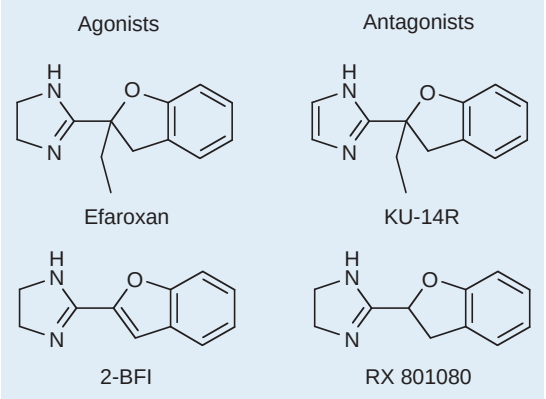


FIGURE 12.18 Hydrogenations and dehydrogenations can induce switches from agonist to antagonist profiles.

coenzymes, tensioactive amphiphilic substances, and production of hemolysis in erythrocytes. In a more precise manner, the alkylating properties of compounds such as the nitrogen mustards, the nitrosoureas and the mitomycins, or the cross-linking properties of *cis*-platinum derivatives relate to their anticancer activity. The activity of anthracyclines, ellipticine, and the anthracenediones has been shown to be due to intercalation in the double helix of DNA. The anticoccidial polyether ionophores are potent complexing agents for mono- and divalent cations.

The biological action of a compound is also readily explicable if it mimics a natural substrate or mediator. This is the case for enzymes with inhibitors (angiotensin I and captopril), suicide substrates (GABA and vigabatrin), antimetabolites (*p*-aminobenzoic acid and *p*-aminobenzene-sulfonamide), and receptors with agonists (acetylcholine and muscarine), antagonists (GABA and gabazine), and uptake inhibitors (GABA and nipecotic acid). The analogy with the endogenous substance can be tenuous as exemplified by the quaternary ammonium compounds, which all present with similar affinity for the cholinergic receptors.

The pathways of drug metabolism follow some general rules (see Chapter 25), and the metabolites of a given substance can, at least qualitatively, be imagined in advance. As a consequence, various measures can be taken to favor or—conversely—to slow down the biodegradation. One possibility recently reviewed is the introduction of deuterium on metabolic soft spots [27]. Some chemical groupings are more prone than others to yield unwanted toxic metabolites (see Chapter 16). Among the best known are the aromatic nitro, nitroso, azo, and amino compounds; the bromoarenes, the hydrazines and the hydroxylamines; and the polyhalogenated aliphatic or aromatic compounds. A proposed explanation of the arylamine toxicities (the so-called "para effect") is their facile oxidation to an electrophilic quinonic system followed by addition of thiol nucleophiles. A process that models well-known hapten formation reactions [28]. Besides biological activity and disposition of the drug, the safety and exploratory toxicology are addressed along the drug discovery process, and several guidelines inspired by marketed drugs have been proposed, in particular those adressing the exploration of the physicochemical properties of lead series. For example, the strategies leading to an attenuation of hERG cardiovascular safety are compulsory step towards high quality drug candidates [6,29–31]. Computational methods can help to establish QSAR models for hERG (or phospholipidosis or CYP binding) and can be combined with approaches that use docking and machine learning strategies [32]. Finally, if the active principle is an acid or a base, the choice of the salifying counterion has also to follow some selection criteria. Oxalates and nitrates, for example, are not very popular, whereas hydrochlorides represent a satisfactory compromise.

C. Rule Number Three: The Structural Logic Rule

This rule implies that as soon as some structural data are available (e.g., intercharge distances, E or Z conformations, axial or equatorial substituent orientations, misoriented substituents), they have to be fed back into drug design. When dealing with enzymes or receptors of unknown structure, one route to such information consists of comparing already known active compounds recognized by the same molecular target and deducing the important stereoelectronic features associated with potency and selectivity. This approach is called "pharmacophore identification" or "receptor mapping" (see Chapters 21 and 22). Initially presented by Marshall et al [33], it has some predictive merit [34,35] and at least avoids unnecessary syntheses of *a priori* inactive compounds. In practice the most efficient method consists of steady comings and goings between synthetic and computer chemistry in order to achieve the ideal interplay between intuition and computer assistance.

A structural guide is also available for drugs designed to bind to the neurotransmitters of the central nervous system (CNS). On the basis of a comparison of the crystal structure of recognized representative compounds from each of eight major CNS active drug classes, Andrews and Lloyd [36,37] identified a common structural basis, essentially characterized by an aromatic plane distant by about 5 Å from a nitrogen moiety. The explanation for this finding resides in the biochemical origin of the neurotransmitters [35]. Most of them are of the arylethylamine type, as a result of the decarboxylation of aromatic amino acids such as DOPA or histidine.

Another example is given by acetylcholine, which can be considered a bioisosteric with GABA (Figure 12.19). This property explains the observation that a compound such as the GABAA receptor antagonist bicuculline is recognized by both the GABAA and the nicotinic receptors [35].

D. Rule Number Four: The Right Substituent Choice

Half of all the existing drugs contain easily substituted aromatic rings. The replacement in such rings of a hydrogen by a substituent (e.g., alkyl, halogen, hydroxyl, nitro, cyano, alkoxy, amino, carboxylate) can

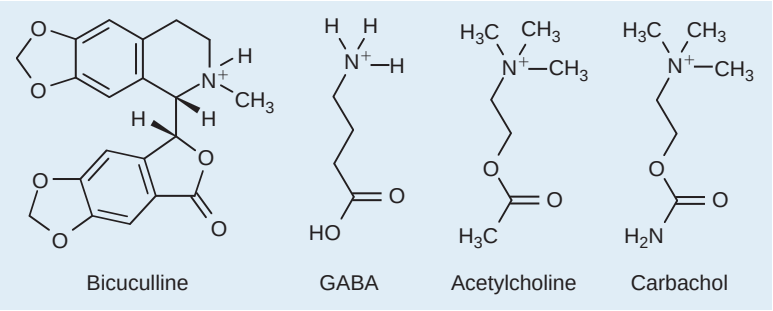


FIGURE 12.19 Similar intercharge distances between the protonated nitrogen and the carbonyl dipole exist in bicuculline, GABA, acetylcholine, and carbachol.

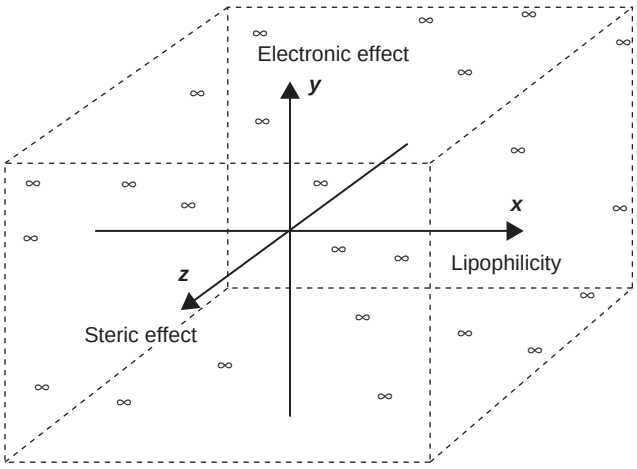


FIGURE 12.20 3D space formed by lipophilic, electronic, and steric coordinates.

dramatically modify the intensity, the duration, and perhaps even the nature of the pharmacological effect (see Chapters 13 and 15). It becomes therefore of prime importance to proceed to the optimal choice of substituents so as to explore with the smallest set possible the 3D space formed by lipophilic, electronic, and steric parameter coordinates (Figure 12.20).

The right substituent choice minimalizes the number of test compounds that have to be synthesized to ensure a significant space volume. This point represents a 3D extension of the Craig plot discussed by Craig [4] and by Austel [5]. In this context, the decision tree proposed by Topliss [38] allows a quick identification of the substituents associated with the highest potency. Application examples of the Topliss scheme are discussed by Martin and Dunn [39].

Multiple methods recently appeared as useful in quantifying the optimization of hits, with Ligand Efficiency (LE) [40] being one of the most commonly used metrics. It comes from Fragment Based Drug Discovery (FBDD), which consists of adding small substituents piece by piece on the hit in order to improve the interaction with the target protein binding site without disturbing the general structure of the hit (see Chapter 7). It can be applied on hits coming from screening that are usually not designed for the target with which they interact. In this case it is important to identify the minimal scaffold required to keep the activity. The first step consists in removing part of the molecule—in most cases losing activity in order to identify the most efficient core template. This first step of exploration allows the identification of pharmacokinetically promising leads. Usually, the optimization phase lead to an increase in the MW; consequently, starting from a low MW (<300) hit will keep this parameter in a favorable developability range.

E. Rule Number Five: The Easy Organic Synthesis (EOS) Rule

Synthesis of new compounds is a costly and lengthy process. Therefore, any measure able to render it more efficacious is welcome. Thus, for example, when the decision is taken to prepare a given set of compounds, why

not first prepare those whose synthesis is the easiest? In the same line of thought, why not prepare first compounds for which intermediates are commercially available?

A particular recommendation is to synthesize heterocycles. In statistics established in 1982 on 1,522 drug molecules and that can still be considered as valid, Kleemann and Engel [41] highlighted the fact that, among the synthetic drugs, 62 percent contained at least one heterocyclic ring, the percentage within natural compounds being even higher (77 percent). Indeed, heterocycles present many advantages. (1) They allow the insertion of elements capable of giving interactions that the carbocycles do not give. (2) They allow a greater number of combinations. It becomes therefore easier to be original. (3) They represent rigid analogs of endogenous substances that themselves are often nitrogenous metabolites of amino acids. (4) Often their facile synthesis permits the preparation of large series [42]. One of the major problems when dealing with isosteric or bioisosteric replacements in heterocyclic systems is the selection of the *a priori* most promising candidate among several dozens of possible rings. A simple clue, which reflects the dipolar moment, can be given by knowledge and comparison of the *boiling points* of the basic heterocycles (see Chapter 8).

F. Rule Number Six: Eliminate the Chiral Centers

Although optical isomerism is discussed in Chapter 18, some practical considerations on chiral molecules are appropriate here. Nowadays, it is well accepted that racemates and both enantiomers are usually three different pharmacological entities and that it requires extensive pharmacological, toxicological, and clinical pharmacological research before it can be decided whether it is advantageous to use racemates or enantiomers in clinical practice. According to Soudijn [42], these research efforts could be reduced to about one-third when drugs without centers or planes of asymmetry could be developed with the same or higher affinity.

Effectively, asymmetry is far from being an absolute requisite for activity. The alkaloid morphine possesses five chiral centers; on the other hand, its synthetic derivative fentanyl is devoid of any asymmetric center but nonetheless one of the most potent analgesics known. Usually, chiral centers are eliminated in creating symmetry (see Chapter 18, Figures 26.9 and 26.10). A typical example of this process is the design of non-chiral immunosuppressive 2-aminopropane-1,3-diols starting from the natural compound myriocin (Figure 12.21) [43].

In some instances, the chiral centers can be at least partially eliminated. This is the case for the synthetic analogs of the hydroxymethyl-glutaryl coenzyme A (HMG-CoA) reductase inhibitor mevinolin. Mevinolin itself has seven asymmetric centers, but the five chiral centers contained in the hexahydronaphtalene ring system are unnecessary for HMG-CoA inhibition. The second generation of mevinolin analogs retains only two of the initial seven chiral centers [44].

When one nevertheless has to deal with chiral centers, why not first prepare the racemic compound and start with an enantioselective synthesis only if an interesting activity is found? This latter point is complicated by the possibility that the two enantiomers might happen to antagonize each other. There are numerous examples published in the literature [42,45–47]. In reality, two optical isomers are never antagonists at comparable dosages. This comes from the space-relationship required for the interaction with the receptor site that is only slightly altered by passing from S to R forms or vice versa. If one of the enantiomers achieves the optimal fit to the

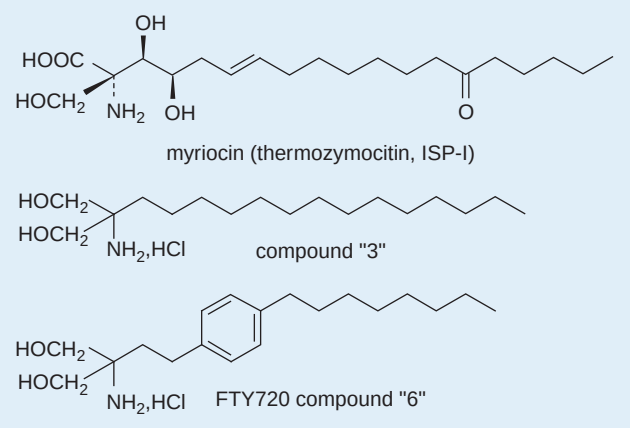


FIGURE 12.21 Suppression of chiral centers through introduction of symmetry [43].

receptor site in exchanging the highest number of noncovalent linkages, its antipode gives rise only to weaker interactions, even under the most favorable conditions (see Chapter 17).

From a practical point of view, this absence of stoichiometric antagonism entails two consequences: (a) if a racemic mixture does not show any activity, it is useless to carry out the separation of the two antipodes; and (b) a racemic mixture usually has the average potency of both constituents, so the maximal benefit one can achieve in resolving racemic mixtures is to increase the potency to twice that of the racemate.

G. Rule Number Seven: The Pharmacological Logic Rule

In Chapter 2, we already insisted on the fact that a correctly performed pharmacological study must satisfy certain criteria (relationship between dose and effect, presentation of the confidence limits, comparison with a reference compound, and determination of the time of the peak action). On the chemical side, it is also extremely important to provide the pharmacologists with reference compounds published by the competitors laboratories. Even if it is tedious and time-consuming to resynthesize an already described compound, the operation is always worthwhile and sometimes surprising. Often a good looking published molecule—for which attractive activities are claimed—loses much of its charm once it is reinvestigated by one's own team. Another point that may contribute to increasing the credibility of your work is the so-called *a contrario* probe. In other words, when your own SAR studies allow identification of the molecular features associated with high activity, proceed, of course, to the synthesis of the most interesting representatives, but also prepare at least one compound that, according to your results, should be inactive. The medicinal chemist should put all things in the pharmacological and therapeutic context: mechanistic studies, target engagement studies, and PK/PD interpretation of pharmacological data will drive the progression of the well-balanced (activity and property) lead compound towards a valuable clinical candidate [48].

References

[1] Descartes R. Discours de la méthode, 1ére partie [Discourse on method. Part I.]. Paris: Librairie Larousse; 1972. p. 27, [in French].
[2] Messer M. Traditional or pragmatic research. In: Jollès G, Wooldridge KRH, editors. Drug design: fact or fantasy? London: Academic Press; 1984. p. 217–24.
[3] Cavalla JF. Drug design valuable for refining an active drug. In: Gross F, editor. Decision making in drug research. New York: Raven Press; 1983. p. 165–72.
[4] Craig PN. Guidelines for drug and analog design. In: Wolff ME, editor. The basis of medicinal chemistry, vol. 1. New York: Wiley-Interscience; 1980. p. 331–48.
[5] Austel V. Features and problems in practical drug design. In: Charton M, Motoc I, editors. Steric effects in drug design, topics in current chemistry, vol. 114. Berlin: Lange & Springer; 1984. p. 8–19.
[6] Meanwell NA. Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. Chem Res Toxicol 2011;24:1420–56.
[7] Schueller FW. Chemobiodynamics and drug design. New York: McGraw-Hill; 1960.
[8] Freidinger RM, Veber DF. Design of novel cyclic hexapeptide somatostatin analogs from a model of the bioactive conformation. In: Vida JA, Gordon M, editors. Conformationally directed drug design. Washington, DC: American Chemical Society; 1984. p. 169–87.
[9] Wermuth CG. Modifications chimiques des médicaments en vue de l'amélioration de leur action [Chemical modification of drugs to improve their action]. Agressologie 1966;7:213–19 [in French].
[10] Nicolaus BJR. Symbiotic approach to drug design. In: Gross F, editor. Decision making in drug research. New York: Raven Press; 1983. p. 173–86.
[11] Froestl W, Furet P, Hall RG, Mickel SJ, Strub D, Sprecher GV, et al. $GABA_B$ antagonists: novel CNS-active compounds. In: Testa B, Kyburz E, Fuhrer W, Giger R, editors. Perspectives in medicinal chemistry. Weinheim: VHC; 1993. p. 259–72.
[12] Durant GJ, Duncan WAM, Ganellin CR, Parsons ME, Blakemore RC, Rasmussen AC. Impromidine (SK&F 92 676) is a very potent and specific agonist for histamine H_2 receptors. Nature 1978;276:403–5.
[13] Hardy GW, Allan G. BW-B3895C. Drugs Future 1988;13:204–6.
[14] Chen S, Dagnino R, De Souza EB, Grigoriadis DE, Huang CQ, Kim KI, et al. Design and synthesis of a series of non-peptide high-affinity human corticotropin-releasing factor 1 receptor antagonists. J Med Chem 1996;39:4358–60.
[15] Simard JR, Grütter C, Pawar V, Aust B, Wolf A, Rabiller M, et al. High-throughput screening to identify inhibitors which stabilize inactive kinase conformations in p38r. J Am Chem Soc 2009;131:18478–88.
[16] Cui JJ, Tran-Dub M, Shen H, Nambu M, Kung PP, Pairish M, et al. Structure based drug design of Crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal_epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). J Med Chem 2011;54:6342–63.
[17] Huang Q, Johnson TW, Bailey S, Brooun A, Bunker KD, Burke BJ, et al. Design of potent and selective inhibitors to overcome clinical anaplastic lymphoma kinase mutations resistant to Crizotinib. J Med Chem 2014;57:1170–87.
[18] Leproult E, Barluenga S, Moras D, Wurtz JM, Winssinger N. Cysteine mapping in conformationally distinct kinase nucleotide binding sites: application to the design of selective covalent inhibitors. J Med Chem 2011;54:1347–55.

[19] Pan Z, Scheerens H, Li SJ, Schultz BE, Sprengeler PA, Burrill LC, et al. Discovery of selective irreversible inhibitors for Bruton's tyrosine kinase. Chem Med Chem 2007;2:58–61.
[20] Johnson TW, Richardson PF, Bailey S, Brooun A, Burke BJ, Collins MR, et al. Discovery of (10R)-7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(metheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and crosOncogene1 (ROS1) with preclinical brain exposure and broad-spectrum potency against ALK-resistant mutations. J Med Chem 2014;57:4720–44.
[21] Baxter A, Bennion C, Bent J, Boden K, Brough S, Cooper A, et al. Hit-to-lead studies: the discovery of potent, orally bioavailable triazolethiol CXCR2 receptor antagonists. J Med Chem 2003;13:2625–8.
[22] Kan J-P. Sanofi-synthelabo recherche [Sanofi-Synthelabo Research]. Montpellier, France. Personal communication, January 17, 1997.
[23] Chan SL, Pallett AL, Morgan NG. Clotrimazole and efaroxan stimulate insulin secretion by different mechanisms in rat pancreatic islets. Naunyn-Schmiedebergs Arch Pharmacol 1997;356:763–8.
[24] Chan SL, Atlas D, James RF, Morgan NG. The effect of the putative endogenous imidazoline receptor ligand, clonidine-displacing substance, on insulin secretion from rat and human islets of Langerhans. Br J Pharmacol 1997;120:926–32.
[25] Morgan NG, Chan SL, Mourtada M, Monks LK, Ramsden CA. Imidazolines and pancreatic hormone secretion. Ann NY Acad Sci 1999;881:217–28.
[26] Brown CA, Chan SL, Stillings MR, Smith SA, Morgan NG. Antagonism of the stimulatory effects of efaroxan and glibenclamide in rat pancreatic islets by the imidazoline, RX801080. Br J Pharmacol 1993;110:1017–22.
[27] Gant TG. Using deuterium in drug-discovery: leaving the label in the drug. J Med Chem 2014;57:3595–611.
[28] Sanner MA, Higgins TJ. Chemical basis for immune mediated idiosyncratic drug hypersensitivity. In: Bristol JA, editor. Ann Rep Med Chem, vol. 26. San Diego, CA: Academic Press; 1991. p. 181–90.
[29] Young RJ. Physical properties in drug design. Top Med Chem 2015;9:1–68. Available from: http://dx.doi.org/10.1007/7355_2013_35.
[30] Hornberg JJ, Laursen M, Brenden N, Persson M, Thouggaard V, Toft DB, et al. Exploratory toxicology as an integrated part of drug discovery. Part I. Why and how. Drug Discov Today 2014;19:1131–6.
[31] Leishman DJ, Rankoviv Z. Drug discovery vs hERG. Top Med Chem 2015;9:225–59. Available from: http://dx.doi.org/10.1007/7355_2014_38.
[32] Moroy G, Martiny VY, Vayer P, Villoutreix BO, Miteva MA. Toward in silico structure-based ADMET prediction in drug discovery. Drug Discov Today 2012;17:44–55.
[33] Marshall GR, Barry CD, Bosshard HE, Dammkoehler RA, Dunn DA. The conformational parameter in drug design: the active analog approach. In: Olson EC, Christoffersen RE, editors. Computer-assisted drug design. Washington, DC: American Chemical Society; 1979. p. 205–26.
[34] Marshall GR, Cramer III DR. Three-dimensional structure–activity relationships. Trends Pharmacol Sci 1988;9:285–9.
[35] Wermuth CG, Langer T. Pharmacophore identification. In: Kubinyi H, editor. 3D QSAR in drug design: theory, methods, and applications. Leiden: ESCOM; 1993. p. 117–36.
[36] Andrews PR, Lloyd EJ. A common structural basis for CNS drug action. J Pharm Pharmacol 1983;35:516–18.
[37] Lloyd EJ, Andrews PR. A common structural model for central nervous system drugs and their receptors. J Med Chem 1986;29:453–62.
[38] Topliss JG. Utilization of operational schemes for analog synthesis in drug design. J Med Chem 1972;15:1006–11.
[39] Martin YC, Dunn WJ. Examination of the utility of the Topliss schemes for analog synthesis. J Med Chem 1973;16:578–9.
[40] Hopkins FAL, Keserü GM, Leeson PD, Rees DC, Reynolds CH. The role of ligand efficiency metrics in drug discovery. Nat Rev Drug Discov 2014;13:105–21.
[41] Kleemann A, Engel J. Pharmazeutische wirkstoffe: synthese, patente, anwendungen (Preface) [Pharmaceutical ingredients: synthesis, patents, applications]. Stuttgart: Georg Thieme Verlag; 1982 [in German].
[42] Soudijn W. Advantages and disadvantages in the application of bioactive racemates or specific isomers as drugs. In: Ariëns EJ, Soudijn W, Timmermans PBMWM, editors. Stereochemistry and biological activity of drugs. Oxford: Blackwell Scientific Publications; 1983. p. 89–102.
[43] Kiuchi M, Adachi K, Kohara T, Minoguchi M, Hanano T, Aoki Y, et al. Synthesis and immunosuppressive activity of 2-substituted 2-aminopropane-1,3-diols and 2-aminoethanols. J Med Chem 2000;43:2946–61.
[44] Baader E, Bartmann W, Beck G, Bergmann A, Granzer E, Jendrella H, et al. Rational approaches to enzyme inhibitors: new HMG-CoA reductase inhibitors. In: Claassen V, editor. Trends in drug research. Amsterdam: Elsevier; 1990. p. 49–71.
[45] Towart R, Wehninger E, Meyer H. Effects of unsymmetrical ester substituted 1,4-dihydropyridine derivatives and their optical isomers on contraction of smooth muscle. Naunyn Schmiedeberg's Arch Pharmacol 1981;317:183–5.
[46] Lotti VJ, Taylor DA. α2-Adrenergic agonist and antagonist activity of the respective (−)- and (1)-enantiomers of 6-ethyl-9-oxaergoline. Eur J Pharmacol 1982;85:211–15.
[47] Hof RP, Rüegg UT, Hof A, Vogel A. Stereoselectivity at the calcium channel: opposite action of the enantiomers of a 1,4-dihydropyridine. J Cardiovasc Pharmacol 1985;7:689–93.
[48] Rossi T, Braggio S. Quality by design in lead optimization: a new strategy to address productivity in drug discovery. Curr Opin Pharmacol 2011;11:515–20.

CHAPTER

13

Substituent Groups

Patrick Bazzini[1] *and Camille G. Wermuth*[2]

[1]PC SAS, Bd Gonthier d'Andernach, Illkirch, France; [2]Founder of Prestwick Chemical, Strasbourg, France;

OUTLINE

Fifty percent of the currently used drugs contain at least one aromatic ring that can be a matter of substitution ***John Taylor*** *[1]*

I. INTRODUCTION

In an active molecule, the replacement of a hydrogen atom by a substituent (e.g., alkyl, halogen, hydroxyl, nitro, cyano, alkoxy, amino, carboxylate) or a functional group can deeply modify the potency, the duration, and perhaps even the nature of the pharmacological effect. Structure–activity relationship studies implying substituent modifications therefore represent a common practice in medicinal chemistry, all the more since half of the existing drugs contain easy-to-substitute aromatic rings. The perturbations brought by the substituent can affect

various parameters of a drug molecule, such as its partition coefficient, its electronic density, its steric environment, its bioavailability and pharmacokinetics, and its capacity to establish direct interactions between the substituent and the receptor or the enzyme.

In reality, it is impossible to modify only one of these five parameters. Thus, for example, the replacement of a hydrogen atom by a methyl group is going to play on the five parameters listed above simultaneously. Nevertheless, through a careful selection of the adequate substituent, it is possible to vary one of the considered parameters in a dominant manner.

To illustrate the repercussions on the biological activity resulting from substituent effects, we will successively study the effects of methyl groups, of unsaturated groups, and of halogen substitution. Hydroxy groups, thiols, and acidic or basic functions will be discussed more briefly. A final section deals with the attachment of large lipophilic additional binding moieties.

II. METHYL GROUPS

In this section, we show how methyl groups—so often considered chemically inert—are able to alter the pharmacological properties of a molecule deeply. We will successively envisage effects on the solubility, conformational effects, electronic effects, and effects on the bioavailability and the pharmacokinetics. In the last paragraph, we will present some replacement possibilities of the methyl groups by related groups and extend the study to some larger alkyl groups.

A. Effects on Solubility

As a rule, the grafting of one or several methyl groups on an active molecule renders the former more lipophilic and therefore less soluble in water. However, in some particular cases, grafting one or several methyl groups to a molecule results in an increase of the water solubility by mechanisms such as the increase of hydrophobic bonding possibilities or diminution of the crystal lattice energy.

1. Increase in Lipophilicity

Normally, one expects methyl groups to increase the lipophilicity. Indeed, the log P (logarithm of the partition coefficient P between n-octanol and water) is 2.69 for toluene, compared to log $P = 2.13$ for benzene [2]. More generally, the passage of (M)-H to (M)-CH_3 gives place to a positive increment of 0.52 in Hansch constants calculations (see Chapter 12).

The increase in lipophilicity due to methylation can drastically modify the bioavailability of the drug and thus its efficacy. This can be illustrated by the morphine derivatives and their CNS activities as outlined in Box 13.1. Moreover, tyrosine dimethylation of the synthetic opioid DPDPE, for example (Figure 13.1), produces an increase in affinity both for the δ-opioid and the μ-opioid receptors. The *in vivo* analgesic activity also increases in comparison to the nonmethylated DPDPE, notably as a result of enhanced bioavailability [3].

Another example of correlation between lipophilicity and activity is found in a series of imidazolinediones (Table 13.1). The introduction of a single methyl group (compound 1 and 2) increases the lipophilic character of the molecule and its ability to displace rimonabant (SR-141716A), the specific antagonist of CB_1 cannabinoid receptors [4].

2. Hydrophobic Interactions

As stated above, the usual result of methyl group addition to a given molecule is the augmentation of its lipophilicity. There are, however, exceptions to this rule, especially when the grafting of one or several methyl groups can render the molecule more compact (more "globular"). A good illustration of this effect is provided by aliphatic alcohols [6]. As expected, one observes that the increase in lipophilicity when passing from n-butanol to n-pentanol is accompanied by a decrease in water solubility. However, 2-pentanol and even more neopentanol—although possessing one methyl more than n-butanol—are less lipophilic, which means they are more soluble in water (Table 13.2).

Similarly, one can observe in Table 13.1 (compound 3) that the imidazolinedione is less lipophilic when R2 is an isopropyl than when R2 is an n-propyl.

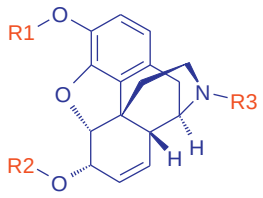
How to explain this anomaly? It has to be simply attributed to an entropic effect [7]. In aqueous solution, the particle is imprisoned in a three-dimensional network (a cluster) of structured water molecules. On the other

BOX 13.1

THE MORPHINE FAMILY [105]

Morphine was first isolated in 1804 by both Seguin and Courtois, but it was the German pharmacist Sertürner who first published its results in 1805. He called this white powder "morphium," because its effects point out the Greek god of dreams, Morpheus. Structure–activity relationships of morphine analogs illustrate well how simple methyl group can change the pharmacological profile.

- The replacement of the methyl by a hydrogen reduces the potency. This can be explained by the fact that the secondary NH obtained in compound 2 is more polar and therefore crosses the BBB less easily.
- Codeine (compound 3), which differs from morphine by a single methyl, is used for treating moderate pain, coughs, and diarrhea. The *in vitro* assays on isolated receptors indicate that codeine should be 1,000 times less active than morphine. But when codeine is given orally to patients, it is only five times less active. This difference between *in vitro* and *in vitro* is due to a demethylation of codeine in the liver, the removal of the R1 methyl group leading to morphine.
- In compound 4, heterocodeine exhibits the highest potency in this table. One could conclude that a substitution on R2 increases activity, but we must remember that these tests have generally been done *in vitro*. In reality, the improvement in activity is more due to the modification of the pharmacokinetic properties of the compound than to its intrinsic affinity for the receptor. Indeed, the methylation of the OH gives a more lipophilic drug, with better BBB crossing and thus higher morphine concentration in the CNS.
- For the acetylated analogs of morphine, 3-acetylmorphine has a weaker activity (proof of the importance of the free phenolic OH) than morphine, while 6-acetylmorphine and diamorphine (heroin) have an increased activity. For both, it is due to the fact that they are less polar than morphine and enter the brain more quickly and in higher concentrations. But whereas 6-acetylmorphine acts immediately on the receptor, the R1 acetyl group of diamorphine has to be removed first. This explains the difference of activity between compounds 6 and 7.



Compound	Name	R1	R2	R3	*Analgesic activity compared to morphine*
1	Morphine	H	H	Me	100
2	Normorphine	H	H	H	25
3	Codeine	Me	H	Me	20
4	Heterocodeine	H	Me	Me	500
5	3-Acetylmorphine	Ac	H	Me	<10
6	6-Acetylmorphine	H	Ac	Me	400
7	Diamorphine	Ac	Ac	Me	200

hand, a smaller amount of structured water molecules is needed to create a cluster around a compact molecule than around an extended one (Figure 13.2). This new structural arrangement is energetically favorable.

For the same reason, some aromatic rings bearing protonated basic side-chains are more soluble than anticipated. In these derivatives, the chain folds in such a way that the cationic head becomes placed under the aromatic ring and can establish a typical donor–acceptor interaction with the π cloud of the aromatic ring (folding effect) [8].

Compound	R	Receptor binding K_i (nM)	
		μ	δ
DPDPE	H	2018	17.7
DMT-DPDPE	Me	58.3	1.8

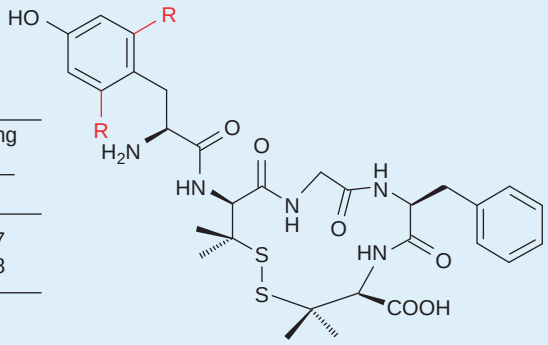
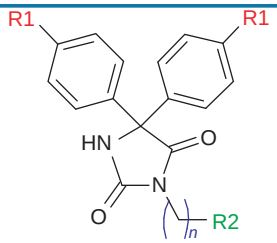


FIGURE 13.1 Structure–activity of DPDPE and its dimethyl-analog (Tyr-D-Pen-Gly-Phe-D-Pen).

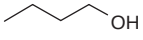
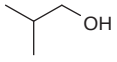
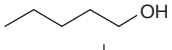
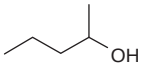
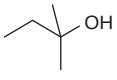
TABLE 13.1 Structure, Affinity to CB_1 Cannabinoids Receptors and Lipophilicity of Imidazolinediones [4]



Compound	R1	*n*	R2	Percentage of displacement at 10 μM	Lipophilicity (log *P*)[a]
1	H	2	*N*-morpholine	<5	2.16
	CH_3	2	*N*-morpholine	24	3.49
2	H	5	—CH_3	36	5.37
	CH_3	5	—CH_3	47	6.48
3	H	0	—$(CH_2)_2$—CH_3	<20	3.84
	H	0	—$CH(CH_3)_2$	<5	2.89

[a]*The lipophilicity was calculated using the CLIP method [5].*

TABLE 13.2 Solubility in Water at 20°C of *n*-Butanol **1**, Isobutanol **2**, *tert*-Butanol **3**, *n*-Pentanol **4**, 2-Pentanol **5** and Neopentanol **6** [6]

No.	Compound	Solubility (g/100 g H_2O)
1		8.2
2		5
3		Miscible
4		2.4
5		4.9
6		12.2

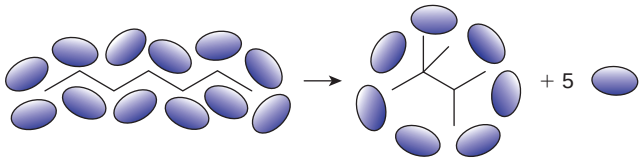
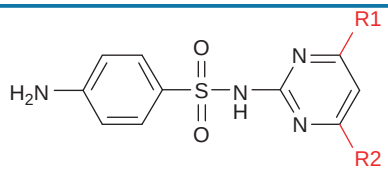


FIGURE 13.2 A lesser amount of structured water molecules is needed to wrap a compact molecule (2,2,3-trimethylbutane) than to wrap an extended one (*n*-heptane).

TABLE 13.3 Increased Water Solubility Caused by Insertion of Methyl Groups [9]



R1	R2	Drug	p*K* (acidic)	Percent ionized at pH 5.2	Solubility pH 5.2–37°C (M)
H	H	Sulfadiazine	6.5	3.9	0.0005
CH_3	H	Sulfamerazine	7.1	1.4	0.0013
CH_3	CH_3	Sulfamidine	7.4	0.7	0.0024

3. *Crystal Lattice Cohesion*

Greater water solubility can also result from a decrease of the crystal lattice energy, with the methyl groups hindering the various intermolecular interactions (e.g., hydrogen bonds, dipole–dipole bonds). In the antibacterial sulfonamide series, the substitution of the pyrimidine ring of sulfadiazine by one, then two, methyl groups causes an increase in solubility (Table 13.3) [9]. *A priori*, one would expect the methylated analogues to be less soluble, in particular because they show increased lipophilicity and they are less dissociated than the parent molecule.

Indeed, the inductive character of the methyl groups disfavors ionization, and the non-ionized form of a molecule is always less soluble than the corresponding ionized form. Despite this unfavorable electronic effect, sulfamidine is approximately five times more soluble than sulfadiazine. Similarly, the grafting of only one methyl group to the herbicide simazine provides atrazine that is fourteen times more soluble in water [10].

B. Conformational Effects

The steric hindrance generated by a methyl group can create constraints and impose particular conformations that may be favorable or unfavorable for ligand–receptor interactions.

Harms and Nauta have studied the effects of methyl substitution on the aromatic ring of the spasmolytic diphenhydramine [11]. The presence of a methyl in *para* position corresponds to a 3.7-fold increase in antihistaminic activity compared to the nonsubstituted derivative (Figure 13.3). Conversely, the presence of a methyl in *ortho* position inactivates the molecule (one fifth of the activity of the nonsubstituted derivative).

The explanation proposed by the authors is as follows: the methyl group in *ortho* position would prevent the side-chain from adopting the usual "antihistaminic" conformation, such as found for phenindamine. Curiously the *ortho*–*ortho*′-disubstituted analog of diphenhydramine shows local anesthetic properties (forty times those of diphenhydramine).

In the same way, the activity of histamine on the H_1 receptor is directly correlated with its conformation. That is why the presence of a methyl group in position 4, by modifying the orientation of the imidazole ring relative to the side-chain, decreases dramatically the potency of 4-methylhistamine (Table 13.4) on the H_1 receptor (400-fold less than histamine), whereas its potency on the H_2 receptor is almost unchanged [13]. Indeed, a different chemical property is needed at the H_2 receptor, namely the tautomeric property of the imidazole ring, to act as a proton-transfer agent (see C. Electronics Effects.).

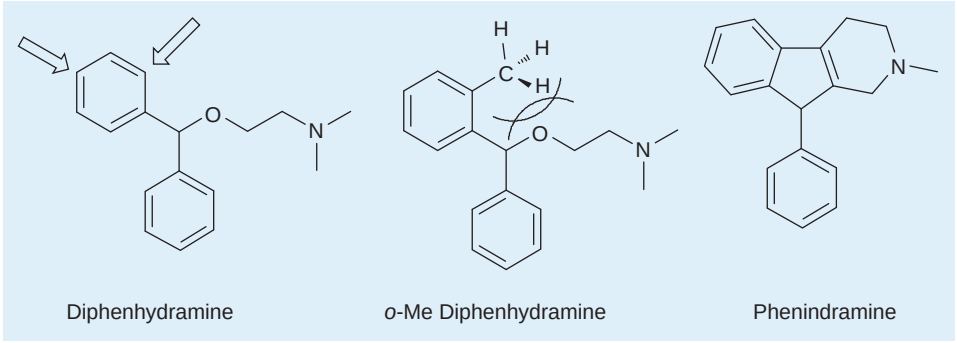
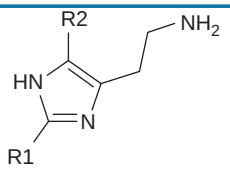


FIGURE 13.3 The presence on o-Me diphenhydramine of an *ortho*-methyl group prevents the side-chain from adopting the favorable coplanar conformation as found in phenindamine [11].

TABLE 13.4 H_1 and H_2 Receptor Agonist Activity of Histamine Derivatives [12]



R1	R2	Derivative	H_1 receptor[a]	H_2 receptor[a]
H	Me	4-Methylhistamine	0.23	39.0
Me	H	2-Methylhistamine	16.5	2.0

[a]*Agonist activity relative to histamine (=100).*

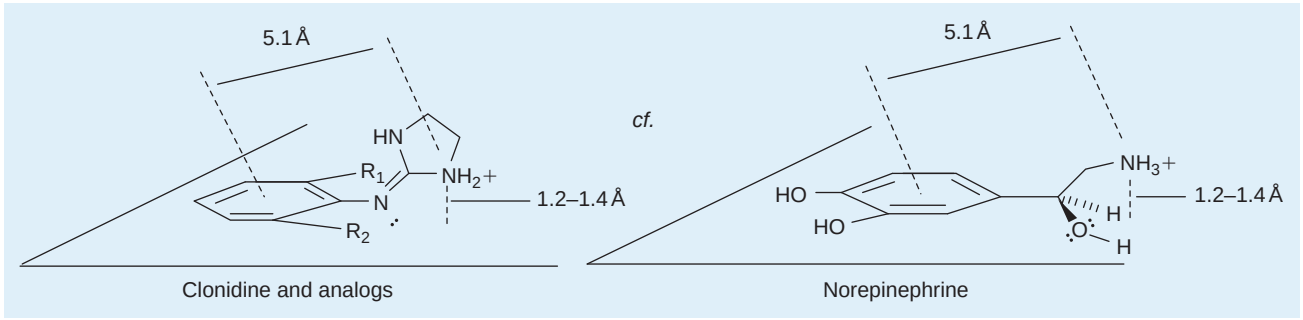


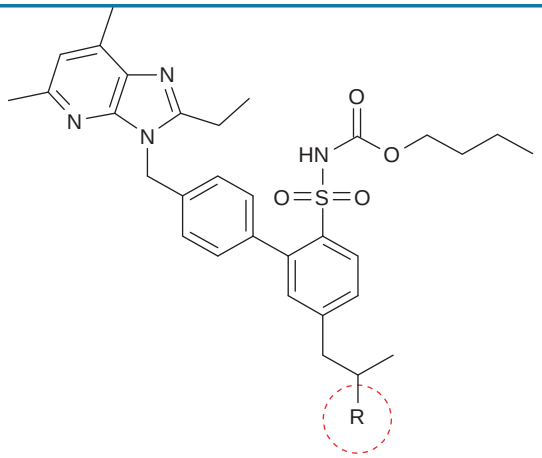
FIGURE 13.4 The restricted rotation resulting from *o*- and *o*′-substitution imposes a quasi-perpendicular orientation of the imidazolinic ring toward the phenyl ring [15].

In steroids, the two angular methyl groups in position 18 and 19 stand on the surface and form a screen above the β face. This entails selective attacks on the rear face (α face) of the molecule [14]. The presence of the methyl in position 18 imposes a preferential conformation to the methylketone chain placed in position 17 [14].

The antihypertensive imidazoline clonidine (Figure 13.4; R1 = R2 = Cl) and its analogs activate norepinephrine as well as specific receptors of the central nervous system (CNS). The maximal activity in this series is always observed when both the *ortho* positions are substituted (R1 = R2 = methyl, chlorine, ethyl, etc.). This situation implies a restrained rotation of the atropisomery type and the impossibility for the two cycles to lie in a coplanar situation. Correspondingly, the geometry of the molecule becomes close to that of the norepinephrine [15].

An interesting example of the effects induced by the presence or the absence of a simple methyl is found in two nonpeptide angiotensin AT_1 ligands (Table 13.5). Both bind with high affinity to the AT_1 and stimulate phosphatidylinositol turnover. However, L-162,782 acts as a powerful partial agonist with a bell-shaped dose–response of 64 percent of the maximal level reached with angiotensin II. Compound L-162,389 only gave a response of 6 percent, which characterized it an antagonist. The authors speculated that the receptor/ligand complex of these compounds after binding is able to change between inactive and active conformations [16].

TABLE 13.5 Dual Agonist and Antagonist Nonpeptide Ligands of the Angiotensine AT1 Receptor [16]



Compound	R	IC_{50} human AT1 wild-type (nM)	Maximal response obtained during stimulation
L-162,389	H	3.97 ± 0.69	5.8 ± 1.3%
L-162,782	Me	24.6 ± 2.3	64 ± 3%

TABLE 13.6 Some Common Aromatic Substituent Constants [17]

Group	π_{para}	σ	MR
H	0.00	0.00	1.03
CH_3	0.56	− 0.17	5.65
CF_3	0.88	0.54	5.02
Cl	0.71	0.23	6.03
OH	− 0.67	− 0.37	2.85
OCH_3	− 0.02	− 0.27	7.87
NH_2	− 1.23	− 0.66	5.42
NH_3^+	–	0.60	–
NO_2	− 0.28	0.78	7.36
CN	− 0.57	0.66	6.33
CO_2H	− 0.32	0.45	6.93
$COCH_3$	− 0.55	0.50	11.18

C. Electronics Effects

The methyl group—and more generally all alkyl groups—are the only substituents acting by an inductive electron-donating effect. All the other groups are electron donors by mesomeric effects. This means that the methyl and the alkyls are electron donors in any environment, while a basic group, dimethylaminoethyl, for example, will be a mesomeric donor in a basic or neutral medium but will become strongly electron attracting by protonation in gastric medium (pH ≈ 2). Table 13.6, taken from Chu [17], presents some numerical values for substituents commonly met in medicinal chemistry [18]. Hansch's π constant accounts for the contribution of lipophilicity, Hammett's σ constants reflect the electronic effects, and the molar refraction (MR) is related to the volume of the substituent. The table illustrates clearly the dramatic change in Hammett's σ parameter when passing from a free amino group ($\sigma = -0.66$) to a protonated one ($\sigma = +0.60$).

A practical consequence is that it is always judicious to include a methyl (or an alkyl) group in a structure–activity relationships (SAR) study. Thus, in any series of R-substituted molecules where one wants to vary R,

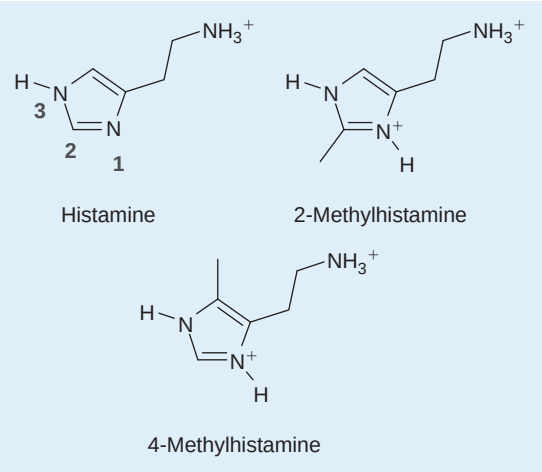


FIGURE 13.5 Prevalent species in aqueous solution at pH 7.4 [13].

the methyl group is generally chosen as a representative of an electron-donating group, the second substituent being chosen from among the electron attractors (e.g., Cl, CN, NO_2, CF_3).

As we have seen previously (Table 13.4), the activity of histamine on H_1 receptor is profoundly altered by methylation. This is also the case concerning the H_2 receptor. To be active on this receptor, the aqueous solution of histamine derivatives must be in a monocationic form, which is considered most likely to be the physiologically active form. We now know that in aqueous solution at pH 7.4, the prevalent tautomer of histamine is the N_3H (Figure 13.5) [19,20]. Indeed, the imidazole ring must not be protonated to act as a proton-transfer agent [13]. However, the presence of a methyl group on the imidazole ring, by its electron-donating effect, has a noticeable effect on the degree of protonation of the imine nitrogen N1, giving it a higher proton affinity. Thus, at pH 7.4, there are more diprotonated species (the non-active form) in solution for 2-methylhistamine and 4-methylhistamine than for histamine (99.85 percent, 87.7 percent, and 3.4 percent, respectively). This explains their relative weak activity compared to histamine (Table 13.4) [12].

D. Effects on Metabolism

Seen from the metabolic point of view, the methyl group plays a particularly important role. Three possibilities are currently met: (a) the methyl group is oxidized; (b) the methyl group is shifted; and (c) the methyl group is not (or only slightly) attacked and can then serve as blocking group.

1. *Oxidation of the Methyl Group*

The oxidation of the methyl group generally begins with the formation of the hydroxymethyl analog and continues usually until the carboxyl step. This is observed for simple compounds like camphor or 2-methyl-pyridine but also for drugs like tolbutamide and alpidem, explaining the relatively short half-life of these latter compounds.

Sometimes the oxidation of the methyl group gives rise to an active metabolite, contributing thus to a reasonable half-life to the drug (Figure 13.6) [21].

The grafting of a methyl group—especially on aromatic rings— often represents a good means of detoxification. It is rapidly oxidized to an inactive and easy-to-eliminate carboxylic group. When the grafted chains are longer than methyl, the attack takes place rather at the benzylic position, at position ω-1, or on ramifications (Figure 13.7) [22].

Angular methyl groups of steroids are usually resistant to metabolic oxidation, probably in relation with a local steric hindrance.

2. *The Methyl Group is Shifted*

When grafted on a nitrogen or sulfur atom, a methyl group can transform this latter in an "onium," which is able to act as methyl donor. In living organisms, the usual suppliers of methyl rests are choline and methionine. Methionine is first activated *in vivo* by combination with adenosine to yield *S*-adenosyl-methionine (SAM; Figure 13.8).

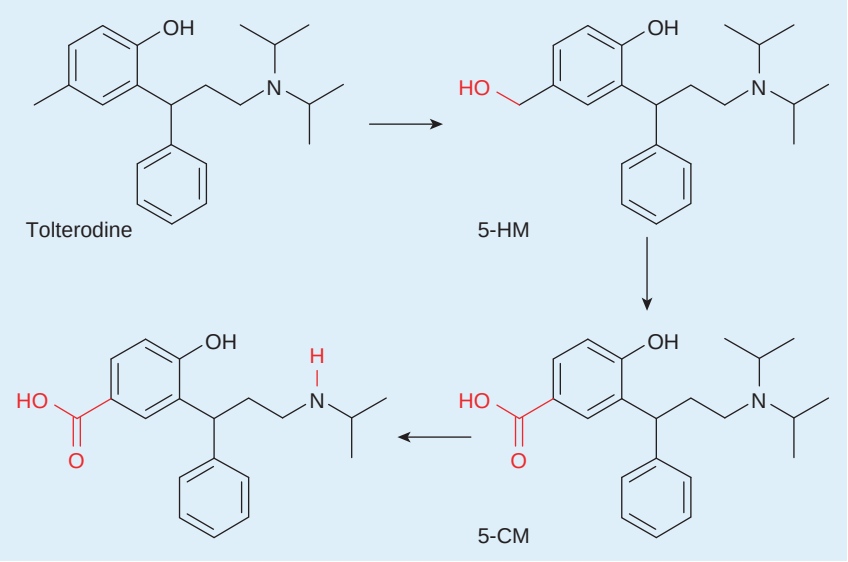


FIGURE 13.6 Metabolic pathway of tolderodine in mice and dogs, leading to its active metabolite 5-HM [21].

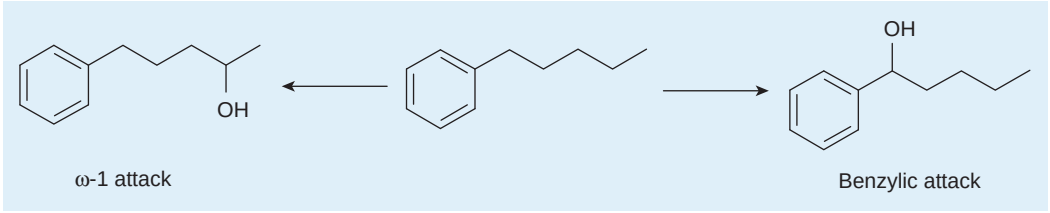


FIGURE 13.7 Privileged oxidative attacks of long chains.

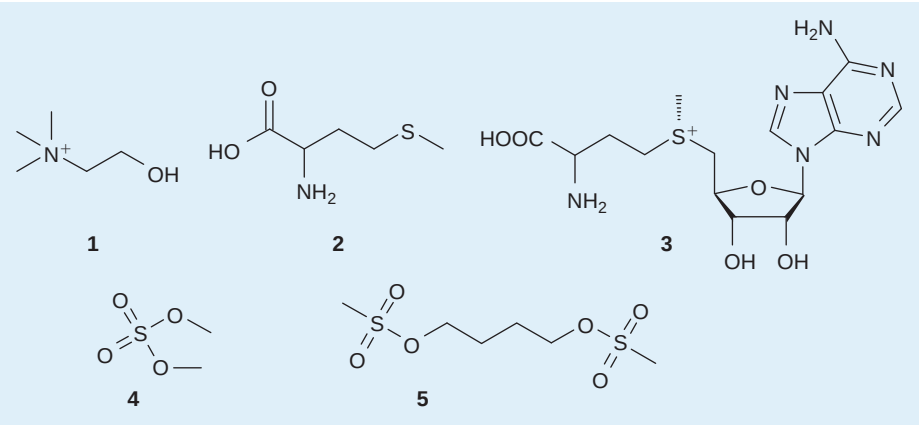


FIGURE 13.8 Methyl donors and alkylating compounds: choline **1**, methionine **2**, *S*-adenosyl-methionine **3**, dimethyl sulfate **4**, and butane-1,4-diol bis-methane sulfonate **5** (busulfan).

More generally, any *S*- or *N*-methylated drug can *a priori* constitute a methyl donor. On the other hand, when the methyl (or alkyl) rest is linked to a good leaving group, as found for alkyl sulfates or sulfonates such as methyl sulfate or busulfan, alkylating reagents are produced, and a huge risk of carcinogenicity exists.

3. *The Methyl Serves to Block a Reactive Function*

A reactive function, such as an active hydrogen belonging to a hydroxyl, thiol, or amino, can be masked by methylation. Methyl groups can thus serve to protect sensitive functionalities from metabolic hydroxylation.

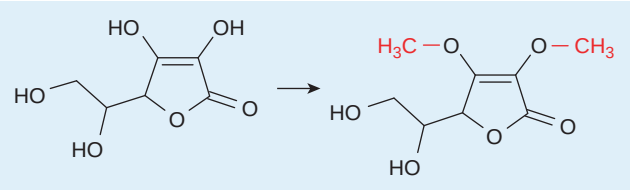


FIGURE 13.9 The methylation of the ene-diol function of ascorbic acid leads to a chemically stable, but pharmacologically inactive compound.

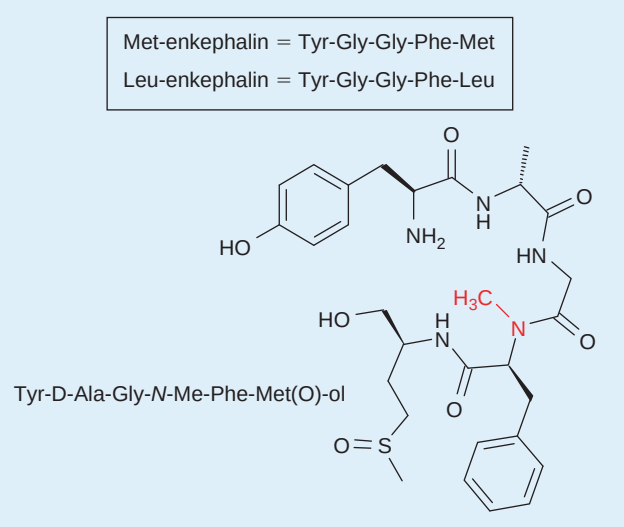


FIGURE 13.10 Less vulnerable analog of Met-enkephalin [23].

The ene-diol function is essential to the antioxidant properties of vitamin C. It is therefore not surprising that its methylation leads to an inactive compound (Figure 13.9).

As such, the endogenous peptides methionine- and leucine enkephalin are inactive by the oral route. Starting from Met-enkephalin, Roemer et al prepared a less vulnerable analog of methionine enkephalin (Tyr-D-Ala-Gly-*N*-Me-Phe-Met(O)-ol), with prolonged parenteral and oral analgesic activity (Figure 13.10) [23]. Several modifications were needed: replacement of glycine by the unnatural D-alanine, *N*-methylation of the Gly-Phe amide bond, oxidation of methionine to the sulfoxide, and reduction of the C-terminus to the corresponding alcohol. For other examples starting from peptide leads, see the excellent reviews of Plattner and Norbeck [24] and of Fauchère [25].

In steroids, the 6α-position (e.g., prednisolone; Figure 13.11) is a position that is normally hydroxylated. Grafting a methyl in this place prevents its hydroxylation. Halogens (particularly fluorine) are even better suited because they are not sensitive at all to oxidative attacks.

Allopregnanolone (3α-hydroxy-5α-pregnan-20-one) has anticonvulsant properties. However, this steroid is not orally active, and its first metabolite is the hormonally active 3-keto-5α-pregnan-20-one. Ganaxolone is a 3β-methylated analog of allopregnanolone. Its 3β-methylation has two effects: (1) it prevents the metabolic attack of the 3α-hydroxy function; and (2) it enhances the bioavailability of the pregnane steroids [26]. Ganaxolone is presently in phase IIb study as antiepileptic (Figure 13.11) [27].

E. Extensions to Other Small Alkyl Groups

The methyl group is the prototype of a saturated aliphatic substituent with lipophilic and electron-donor inductive effects. In some instances, it can advantageously be replaced by related groups bringing symmetry, more lipophilicity, or an increased inductive effect. We remind below of some possibilities.

Keep in mind, however, that modifications, even by alkyl groups, can dramatically affect the properties of a substance, as found some 1,5-benzodiazepine where hydrogen and methyl substituents yield weak antagonists of cholecystokinin, whereas ethyl, propyl, and butyl substituents yield agonists (Figure 13.12) [28].

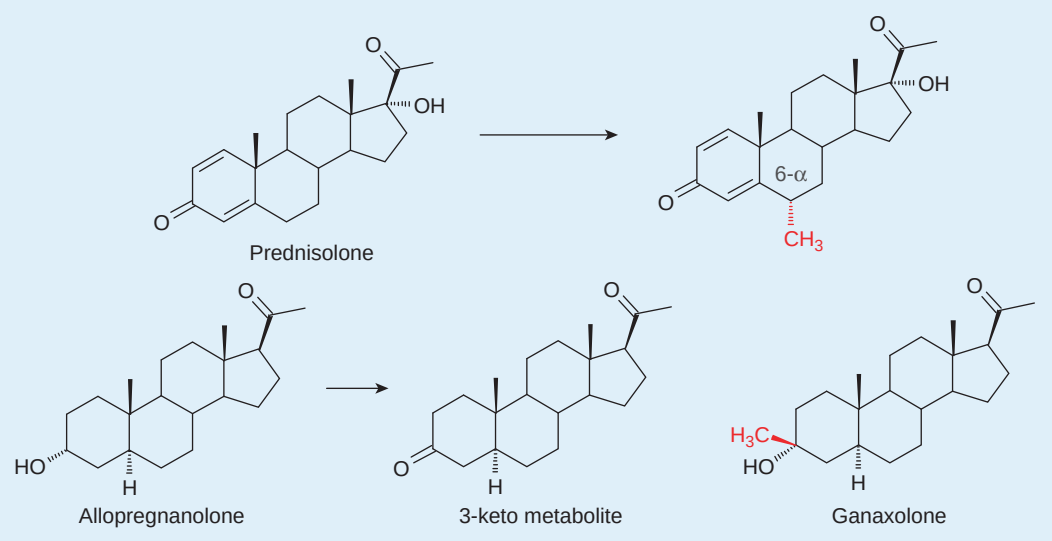


FIGURE 13.11 Protection of steroids against metabolic attacks.

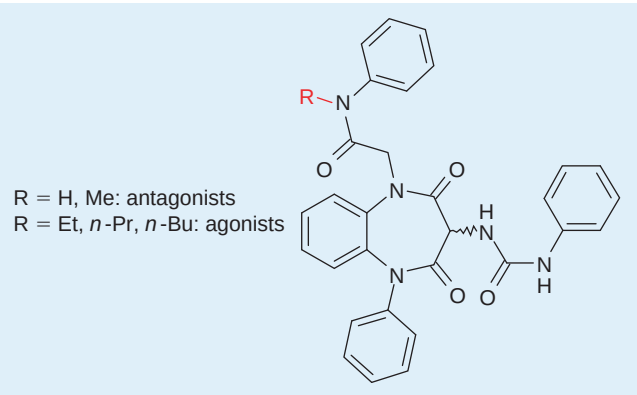


FIGURE 13.12 Structure of 1,5-benzodiazepine active on cholecystokinin receptors [28].

1. *Numerical Values*

The values reported in Table 13.7 [29] allow the comparison of some characteristic alkyl groups in aromatic substitution. One will note the comparable bulkiness (E_s) of the isopropyl and cyclopentyl groups, while the tertiobutyl group is far more voluminous. Furthermore, it is remarkable to observe that the electron-donor effect of the cyclopentyl group is superior to that of the cyclohexyl group.

2. *Gem-Dimethyl and Spiro-Cyclopropyl*

Gem-dimethyl and spiro-cyclopropyl are useful to render a carbon atom quaternary and therefore resistant to metabolic attacks.

Starting from retinoic acid (RA), the active vitamin A metabolite, one of the first constrained analogs, TTNPB, was synthesized (Figure 13.13). The incorporation of a second *gem*-dimethyl group into the tetrahydronaphthalene ring aimed to block its oxidation of the ring. The resulting compound is 10-fold more potent in the tracheal organ culture (TOC) than all-*trans*-RA [32] but also 10,000 times more toxic [33]. However, other less toxic retinoids are used in chemoprevention. Bexarotene, in combination with chemotherapeutic agents, has demonstrated interesting results with advanced nonsmall-cell lung cancer. Two phase III trials are currently underway to characterize the role of bexarotene in the treatment of these cancers [34].

Gem-dimethyl can also constitute solutions to introduce symmetry into a chiral center or to protect a close and sensitive function, as in the case of gemeprost, an analog of prostaglandin E_1 used in medical abortion, where the *gem*-dimethyl groups at C-16 protect the alcohol moiety at C-15 from rapid metabolic oxidation [35].

TABLE 13.7 Aromatic Lipophilic, Electronic and Steric Descriptors for Some Current Aliphatic Rests [29]

Group	π (Hansch)	σ (Hammett)	E_s (Taft)
Methyl	0.50	− 0.17	0.00
Isopropyl	1.30	− 0.19	− 1.08
Cyclopropyl	1.20[a]	− 0.30[b]	–
Cyclobutyl	1.80	− 0.20	− 0.67
Tertiobutyl	1.98	− 0.30	− 2.46
Cyclopentyl	2.14	− 0.20	− 1.12
Cyclohexyl	2.51	− 0.15	− 1.40

[a]*Taken from Ref. [30].*
[b]*Taken from Ref. [31].*

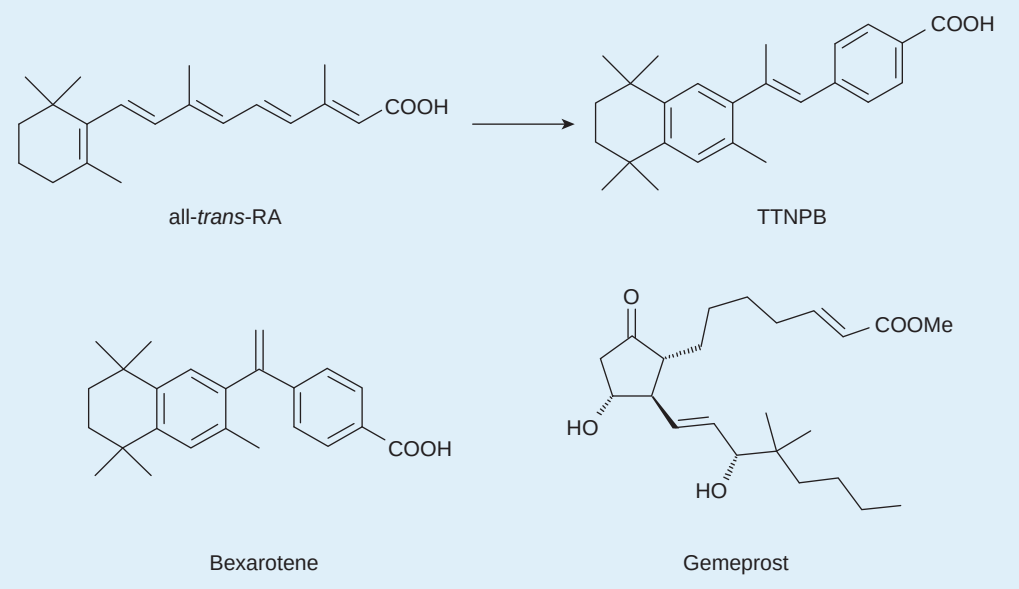


FIGURE 13.13 *Gem*-dimethyl derivatives.

3. *Isopropyl and Cyclopropyl*

The cyclopropyl rest is less bulky than the isopropyl group for a maximal electron-donor effect. This electronic effect is involved when the cyclopropyl group of efavirenz, a nonnucleoside reverse transcriptase inhibitor, interacts with the aromatic ring of tyr181 via a π-aryl interaction that is presumably favorable to binding [36].

The lipophilic effect of cyclopropyl explains why abacavir, a nucleoside reverse transcriptase inhibitor, has an improved absorption in the CNS compared to diaminopurine dioxolane (DAPD) (Figure 13.14) [38].

For a review on cyclopropane derivatives in medicinal chemistry, see Cussac et al [39].

4. *The Cyclopentyl Groups*

The cyclopentyl group creates the maximal inductive effect for a relatively reasonable bulkiness. It is often a good filling of a hydrophobic pocket, as illustrated for the cAMP-phosphodiesterase inhibitor rolipram (Figure 13.15). The inhibitory activity toward type IV cAMP-phosphodiesterase is increased ten times when the *meta*-methoxy group is replaced by a *meta*-cyclopentyl group (rolipram) [40].

Presumably, the cyclopentyl rest fills—in an optimal manner—a hydrophobic pocket of the active site of the enzyme. The cyclopentyl group has also proven advantageous in replacing a *gem*-dimethyl in a series of inhibitors of acyl-CoA-cholesterol acyltransferase, which is an enzyme implied in the absorption of the alimentary cholesterol [41].

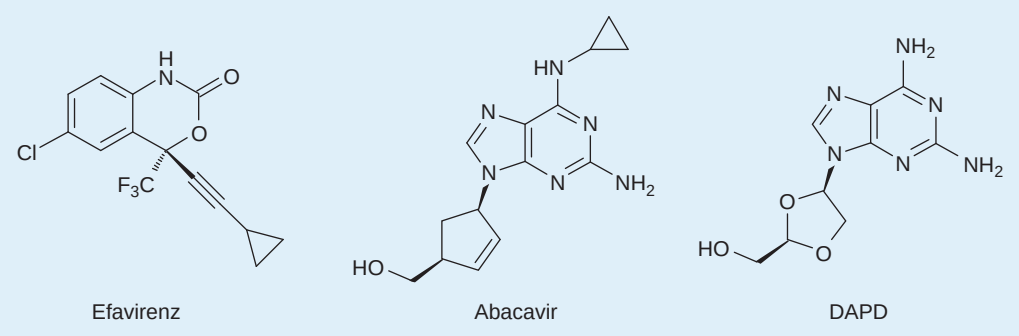


FIGURE 13.14 Reverse transcriptase inhibitors [37].

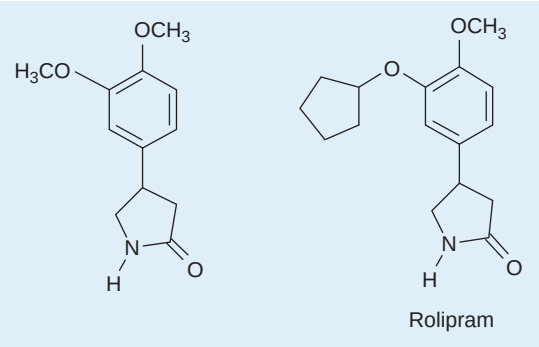


FIGURE 13.15 Structures of rolipram and of its dimethoxy analog [40].

III. EFFECTS OF UNSATURATED GROUPS

The introduction of an unsaturated group (e.g., vinyl, ethynyl, allyl) in a drug molecule generally entails one or several of the following consequences [42,43].

1. *Existence of electronic effects*: the unsaturated rests behave as electron attractors through inductive effects. Furthermore, direct interactions of donor–acceptor type are possible thanks to the π electron cloud surrounds present in multiple bonds.
2. *Possibility of existence of a geometrical isomery* (e.g., *cis–trans* geometric isomery).
3. *Possibility of activation through conjugation*: the association of several unsaturated functions in conjugated position (dienes, enynes, enones, enolides, polyunsaturated derivatives) renders the corresponding molecules very reactive. It especially facilitates the addition of biological nucleophiles and notably of thiols.
4. *Facilitation of the metabolism*: The unsaturated element often constitutes a vulnerable site on the molecule that will be attacked first (e.g., by formation of an epoxide that evolves into a diol that, on its turn, can undergo oxidative cleaving), but this is not always the case. Therefore, one should pay attention to the problems posed by the formation of these metabolites (e.g., aldehydes, carboxylic acids), as they can also be biologically active.
5. *Increase of the narcotic power and the toxicity* in comparison with the corresponding saturated compound. Ethylene, acetylene, trichlorethylene, divinyl oxide and, by extension, cyclopropane are examples of unsaturated narcotics.

With regard to their classification, we will distinguish four series of unsaturated derivatives: the vinyl series, the allyl series, the acetylenic series, and the ring-unsaturated derivatives that are bioisosteric to aromatic rings.

A. Vinyl Series

Beside active substances containing actual vinyl rests, this series comprises molecules containing substituted vinyl groups as well as cyclopropyl groups.

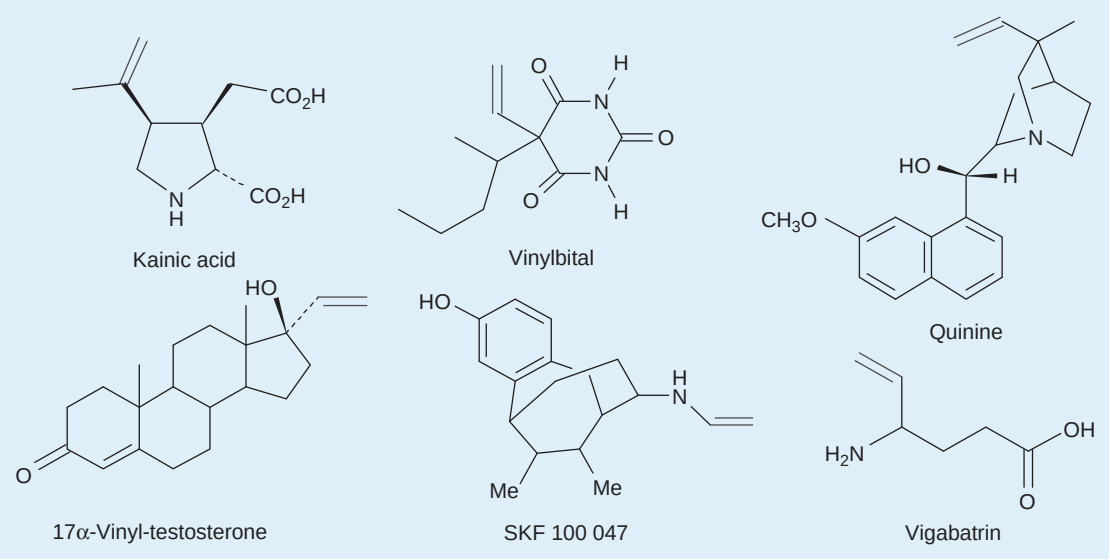
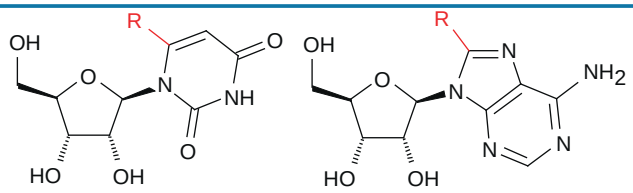


FIGURE 13.16 Medicines containing a vinyl group.

TABLE 13.8 Inhibitory Effects of Nucleoside Derivatives on Human T-lymphoblast Molt/4F and MT4 Cells [45]



Nucleoside	R	Molt/4F IC_{50} (μM)	MT4 IC_{50} (μM)
Uridine	Vinyl	14 ± 0.12	11
	Ethynyl	>372	286 ± 78
Adenosine	Vinyl	6.5 ± 0.1	15 ± 5
	Ethynyl	25 ± 0.1	22 ± 7

1. *Vinyl Groups*

These groups are not excessively used in medicinal chemistry. Divinyl oxide is an excellent general anesthetic, but it polymerizes easily and forms peroxides. Stabilization of the compound is usually achieved by addition of 0.01 percent of *N*-phenyl α-naphtylamine. On the other hand, compounds such as kainic acid, vinylbital, quinine, 17α-vinyl-testosterone, compound SKF 100 047, and vigabatrin (Figure 13.16) are perfectly stable vinyl derivatives.

Nowadays, several articles report the interest in vinyl groups. Indeed, the introduction of alkyl, akenyl, and alkynyl groups into purine and pyrimidine nucleosides is of great interest for their potential activity [44]. As an example, the introduction of vinyl and ethynyl on uridine and adenosine shows an interesting inhibitory effect toward human tumor cell lines (Table 13.8). According to the authors, these groups induce the opposite conformation of the glycoside bond if compared to the natural nucleoside, which made them more cytostatic than the natural ones [45].

Rancourt et al [46] have shown some interesting results with aminocyclopropane carboxylic acid (ACCA) derivatives as inhibitors of the hepatitis C virus NS3 protease. The replacement of hydrogen by an ethyl group affords a modest 3-fold improvement of potency (from $IC_{50} = 14$ to 4.8 μM). But the modification of the ethyl

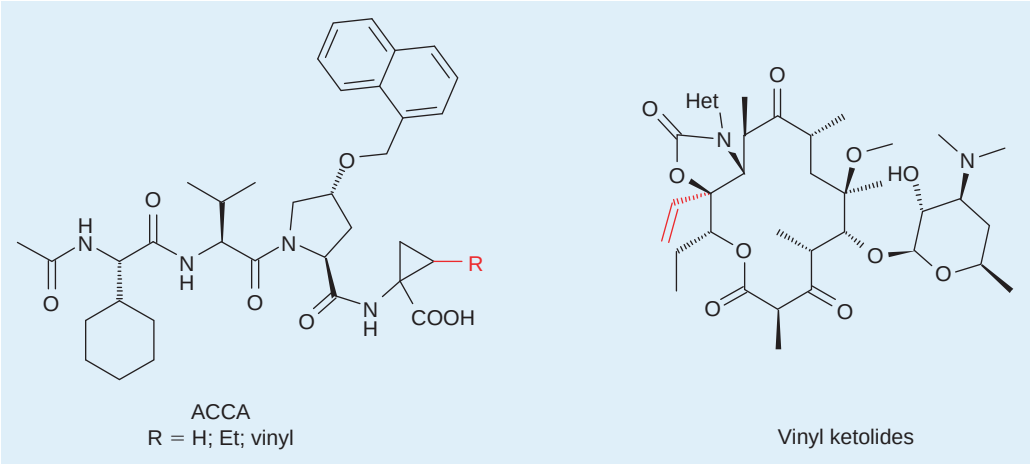


FIGURE 13.17 Recent studies on vinyl-containing compounds [46,47].

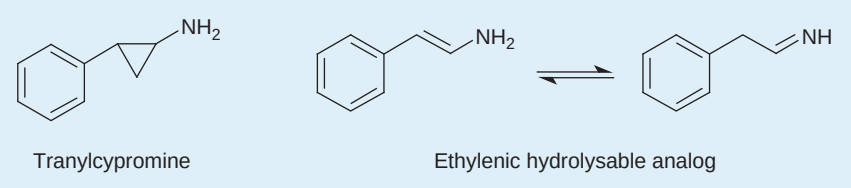


FIGURE 13.18 Tranylcypromine represents a stable substitute of the enamine aminostyrene.

group to a vinyl group allows once again a 7-fold increase (from $IC_{50} = 4.8$ to 0.63 μM) (Figure 13.17). This difference may be due in a part to a beneficial electronic interaction between the π-electrons of the vinyl group and those of the phenyl group of the near Phe154.

The incorporation of vinyl groups into antibacterial C12 ketolides has a favorable impact on the pharmacokinetic and pharmacodynamic properties, increasing the lung-to-plasma AUC ratio, the bioavailability, and the half-life in plasma and lung. These properties directly impact on the *in vivo* potent efficacy in rat lung infection models [47].

2. *Cyclopropyl Groups*

Cyclopropyl rings can constitute interesting substitutes for vinyl rests when the latter are too fragile or give place to unwanted isomeries or tautomeries. Thus tranylcypromine, an antidepressant acting by inhibition of the monoamine-oxidases (MAO), is a stable compound, while its ethylenic analog no longer is (Figure 13.18). A supplementary advantage in the use of cyclopropanic analogs comes from their fixed stereochemistry; there is no spontaneous conversion from *cis* to *trans* isomer as frequently observed with ethylenic derivatives.

B. Allylic Series

All allylic derivatives are relatively hepatotoxic and irritant. Allylic alcohol itself serves to create experimental hepatic lesions that allow testing hepatoprotecting drugs. We will consider three categories of allylic derivatives: *C*-allyl derivatives, *N*-allyl derivatives, and *O*-and *S*-allyl derivatives, which often possess alkylating properties.

1. C-*allyl Derivatives*

C-allyl derivatives present the double advantage of being lipophilic (rapid onset) and giving place to fast biodegradation (short duration of action). However, they often conserve the intrinsic hepatotoxicity of the allyl group [48]. Allobarbital is a sedative-hypnotic that is no longer used; allylestrenol acts as a pure progestative hormone, and alprenolol is a β-blocker (Figure 13.19).

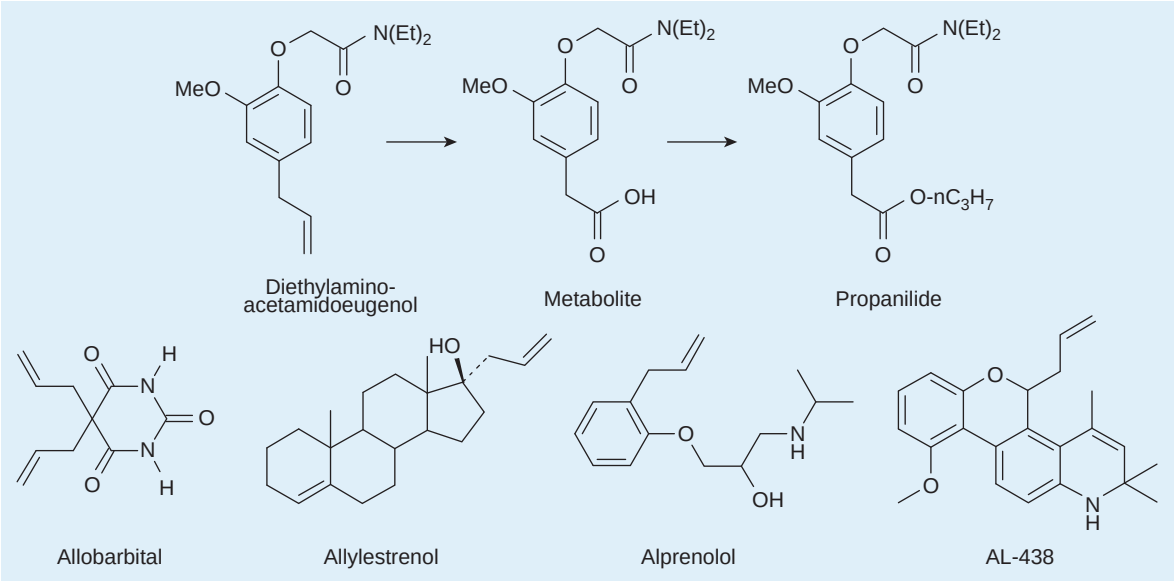


FIGURE 13.19 *C*-allyl derivatives.

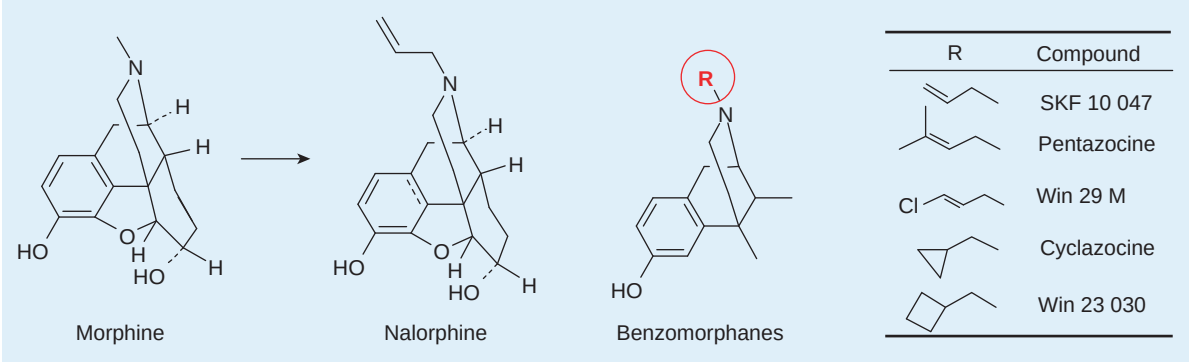


FIGURE 13.20 Nalorphine and cognate derivatives [51].

Acetamidoeugenol is an intravenous anesthetic of ultra-short duration of action, it has been withdrawn because it provokes irritations and lesions of the vascular wall. Acetamidoeugenol is oxidized very rapidly *in vivo* into the corresponding aryl-acetic acid. This observation was the basis of the synthesis of another intravenous short-acting anesthetic, propanilide [49].

Compound AL-438, a nonsteroidal selective glucocorticoid modulator, is a potent anti-inflammatory agent (in the same range as prednisolone). This 5-allyl compound has low side effects compared to other glucocorticoid treatment with no increase of glucose level and no effects over bone formation rates. Compound AL-438 was considered a lead compound, and [50] clinical trials were made before the end of its development.

2. N-*allyl Derivatives*

The replacement in morphine and in some of its simplified analogs of the *N*-methyl group by a *N*-allyl group (and, later on, by some related groups) [51] has constituted a decisive step in the study of opiate analgesics. Indeed, this modification had for the first time achieved the passage of morphinic receptor agonists to the corresponding antagonists (Figure 13.20).

Aloxidone and albutoïne are anticonvulsivant *N*-allyl derivatives of hydantoin and thiohydantoin (Figure 13.21). The dibenzazepine azapetine is an α-adrenergic blocking agent used as peripheral vasodilator.

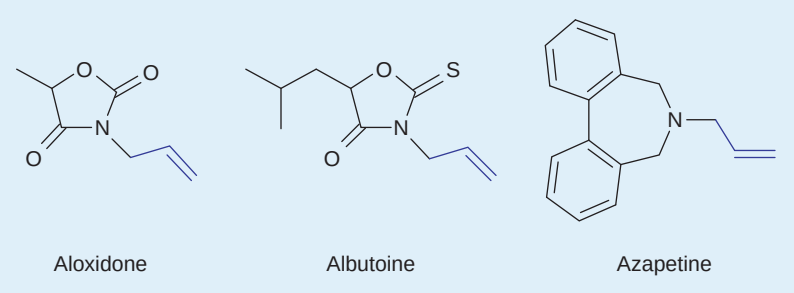


FIGURE 13.21 *N*-allyl derivatives of hydantoin, thiohydantoin, and dibenzazepine.

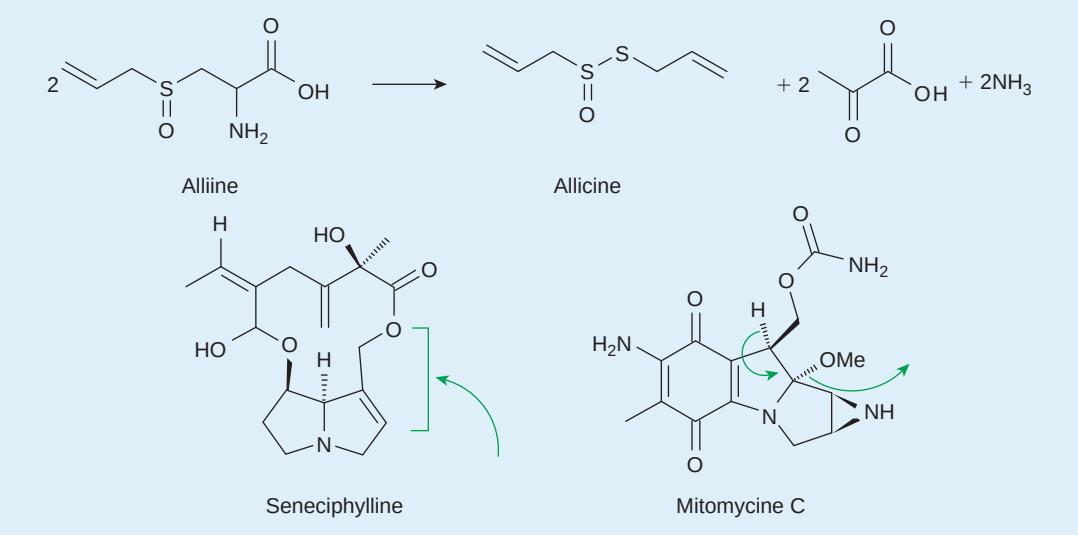


FIGURE 13.22 Alkylating allylic derivatives.

3. O- *and* S-*allyl Derivatives*

Several of these compounds are listed in the Merck Index. The β-blocking oxprenolol, the arylacetic analgesic-anti-inflammatory drug aclofenac, and the fungicide enilconazole are *O*-allyl derivatives. Penicillin O and penicillin S are both *S*-allyl derivatives.

Alkylating allyl derivatives: When the allyl rest bears a good leaving group, it easily generates the allylic cation. This cation is stabilized by mesomery and is an excellent electrophile. Many natural compounds can release allylic alcohols. A first example is found in allicine, the antibacterial principle of garlic, which results from the action of alliinase on alliine (Figure 13.22). Several varieties of senecio (e.g., *Senecio vulgaris* L., Compositae, *Senecio platyphyllus*) contain alkaloids derived from pyrrolizidine (e.g., senecionine, seneciphylline). These substances provoke hepatic cancers, notably in cattle. Here too, the alkylating properties are due to the allylic structure (Figure 13.22) [52] It is reasonable to mention the formation of allylic carbamates also in mitomycines A, B, and C, which are antimitotic drugs used in cancerology. Presumably, the allyl function is created by an elimination reaction involving the departure of an acetalic hydroxy or methoxy group.

C. Acetylenic Series

Acetylenic groups are used for their electronic effects, as equivalents of aromatic rings, and for imposing structural constraints.

1. *Electronic Effects*

The acetylene function exerts an electron-attracting effect. This effect can be reinforced by substitution of the acetylenic hydrogen. Ethynyl compounds are essentially found among the light sedative-hypnotic drugs

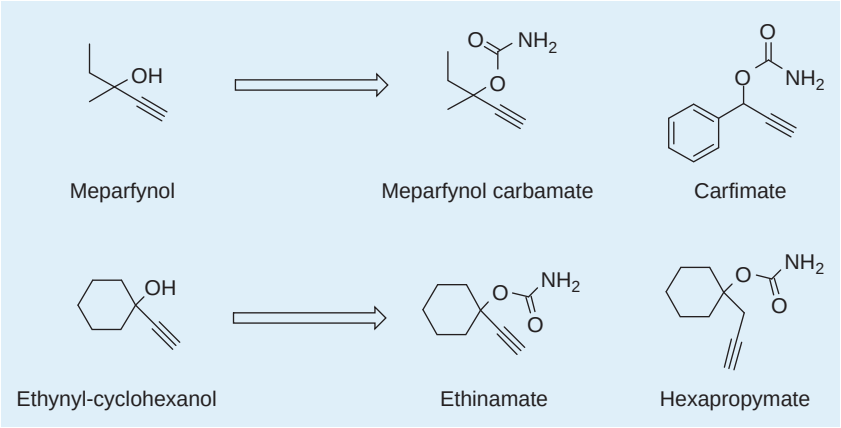


FIGURE 13.23 Ethynylated sedatives.

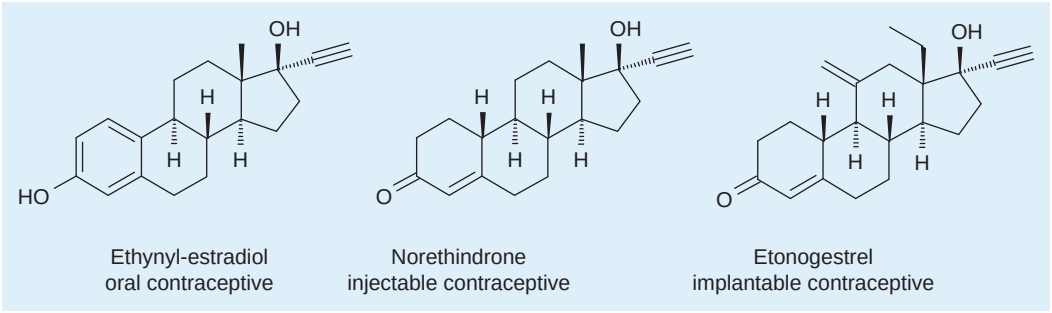


FIGURE 13.24 Ethynylated steroids.

(CNS-depressing effect of the unsaturated derivatives), where most of the acetylenic alcohols are used as carbamic esters (e.g., meparfynol, ethinamate; Figure 13.23).

The acetylenic CH can act as an hydrogen bond donor (like for the antiparkinsonian rasagiline) and thus allow an additional interaction with the receptor [53].

In steroid series, the presence of ethynyl groups in position 17α provides orally active compounds that are metabolized to 17α-methylketones by hydration of the triple bond (Figure 13.24).

Because of the potential biologic activity of ethynyl groups, they should not be systematically avoided during SAR exploration, even if mono-substituted ethynyl derivatives can sometimes be highly cytotoxic [54].

2. *Aromatic Ring Equivalents*

Thanks to their π electron clouds and their small volume, ethynyl groups can sometimes function as bioisosteres of aromatic rings and give similar donor−acceptor interactions.

An X-ray structure of ABT-279 (Figure 13.25), an inhibitor of dipeptidyl peptidase-IV (DPP-IV) bound in the human active site, shows an interesting feature: Tyr547 and Phe357 of DPP-IV formed a narrow tunnel between their phenyl rings that accommodates the ethynyl group. This suggests a π−π stacking effect between the triple bond and the aromatic systems of the receptor. Further investigations showed good pharmacokinetics and an excellent preclinical safety profile of the ethynyl compound, allowing ABT-279 to be selected as a candidate for clinical evaluation in humans for the treatment of type 2 diabetes [55].

Lee et al also supposed a stacking effect for a 6-ethynyl quinazoline. This additional interaction with the receptor increased the potency of this compound five times when compared to the 6-chloro analog. However, this acetylenic analog is reactive and rapidly metabolized, giving, for example, Michael addition in a biological system, and thus led the authors not to test this compound further [56].

Another example of the rapid metabolization of ethynyl groups with sodium methohexital, which is used as an injectable barbituric for very short anaesthesias.

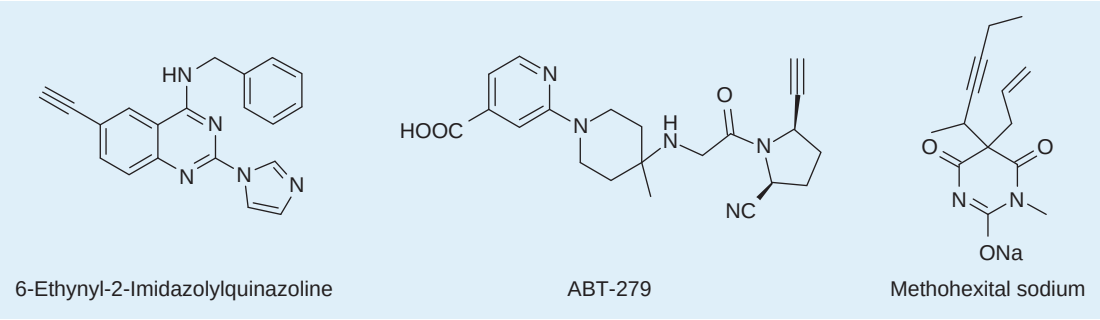


FIGURE 13.25 Acetylenic groups as aromatic ring equivalents.

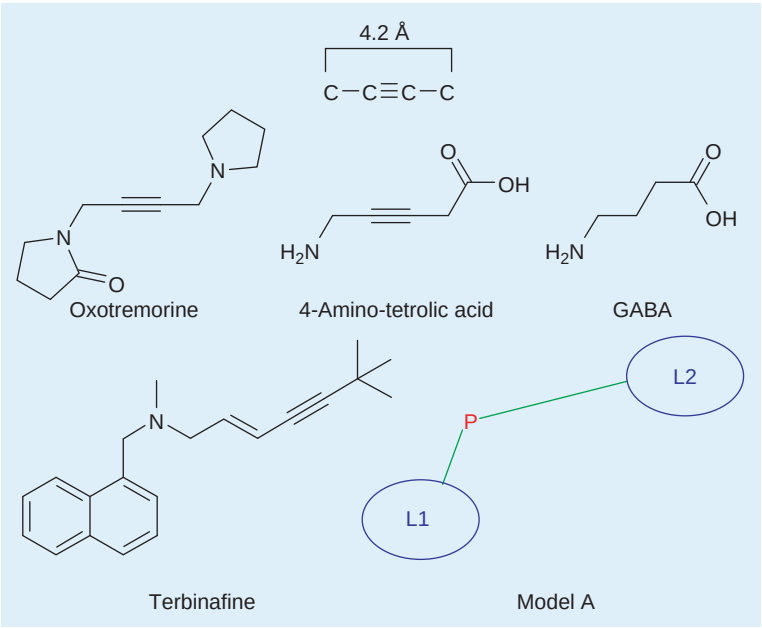


FIGURE 13.26 Rigidity and extension imposed by a triple bond.

3. *Structural Constraints*

In inserting an acetylenic function between two carbon atoms, one achieves a structure with four "on-line" atoms representing a rigid entity with a distance of 4.2 Å between the two extreme atoms (Figure 13.26). This kind of arrangement is found in the cholinergic agonist oxotremorine and in the γ-aminobutyric acid (GABA) analog 4-amino-tetrolic acid. This compound is recognized, like GABA itself, by the enzyme GABA-transaminase, for which it acts as an inhibitor [57].

A structural constraint is also present in the antimycotic terbinafine [58]. SAR exploration led to a simple pharmacophore model A in which the *trans*-enyne group achieved the best length spacing between the lipophilic domain (L2) and the polar group (P) [59].

D. Cyclenic Equivalents of the Phenyl Ring

The cyclohexenyl ring and, to a lesser extent the cyclopentenyl and cycloheptenyl rings, can possibly replace a phenyl ring. This is the case for the barbiturics cyclobarbital and heptabarbital, which are entirely comparable to phenobarbital. From the metabolic point of view, the cyclohexenyl ring is oxidized in position α to the double bond to produce the corresponding cyclohexenone (Figure 13.27).

Another example comes from the benzodiazepine series, where tetrazepam can be compared to diazepam. However, in this case a slight difference in the activity profile exists. Tetrazepam is less sedative, hypnotic, and anticonvulsant than diazepam. On the other hand, it has more muscle-relaxant and analgesic effects, hence its indication

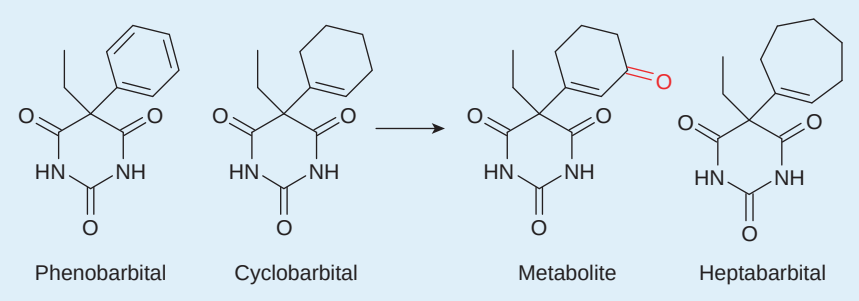


FIGURE 13.27 Cyclenic equivalents of the phenyl ring.

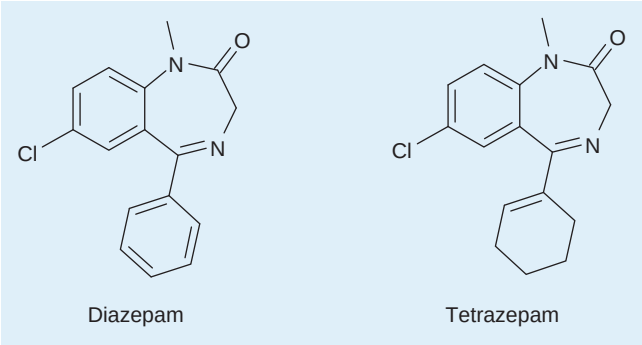


FIGURE 13.28 Cyclenic equivalents in the benzodiazepine series.

in visceral and particular pain. From the chemical point of view, the cyclohexenic double bond is introduced in a rather unexpected manner by means of a radicalar rearrangement of a *N*-chloroamide (Figure 13.28) [60].

IV. EFFECTS OF HALOGENATION

Presently, one drug in three is a halogenated derivative, and halogens are found in drugs belonging to practically all therapeutic classes. That has not always been the case. Indeed, in the past medicines were mostly of natural origin, and natural substance chemistry is relatively deprived of halogenated substances. Halogen-containing drugs have entered usage only after 1820. The first organic halogenated drugs were mainly used for their depressive action on the CNS (e.g., production of a general anesthesia with chloroform, sedation or hypnosis with chloral and bromural). From 20th century onward, a regular growth in the number of halogenated drugs had been observed, a growth that exploded at the end of World War II. Even halogenated drugs from natural origins became available, such as chlortetracycline or chloramphenicol, as well as substances from marine origin or from fermentation broths.

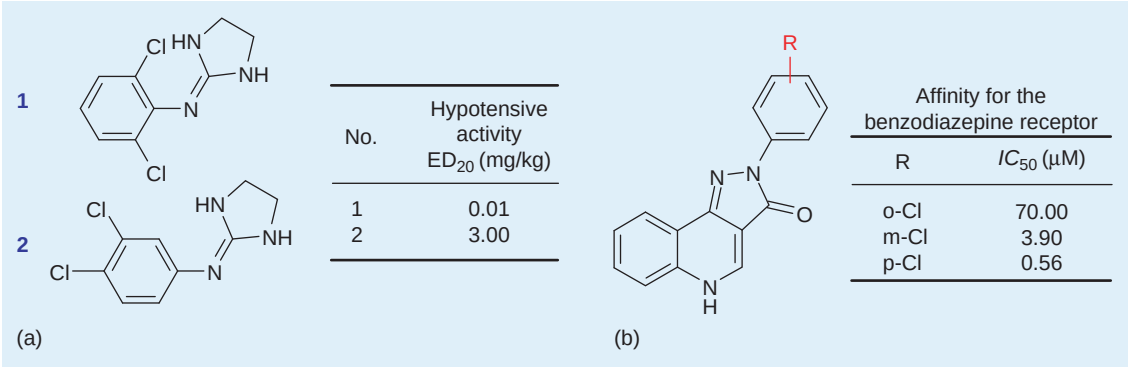
A. The Importance of the Halogens in the Structure–Activity Relationship

1. Steric Effects

The obstruction of a molecule by means of halogen substitution can impose certain conformations or mask certain functions. In the case of clonidine, the bulky halogen atoms prevent the free rotation and maintain the planes of the aromatic rings in a perpendicular position to each other (Figure 13.29a) [15].

In a series of benzodiazepine receptor ligands derived from CGS 9896, strong steric effects are described by Fryer et al (Figure 13.29b) [61]. Indeed, the *ortho* and the *para* isomers can be considered as having the same lipophilicity and very similar electronic effects. Thus, the reduction of binding affinity of the *ortho*-chloro compound is attributed to the steric effect.

The halogen van der Waals radii give a good idea of the size of each atom or groupment. In Table 13.9, the values proposed by Bondi [62] show that the size of fluorine is close to the size of oxygen. Chlorine, bromine, sulfur, and methyl groups are close to one another; due to the length of the C—F bond (1.39 Å), trifluoromethyl is equal to isopropyl.



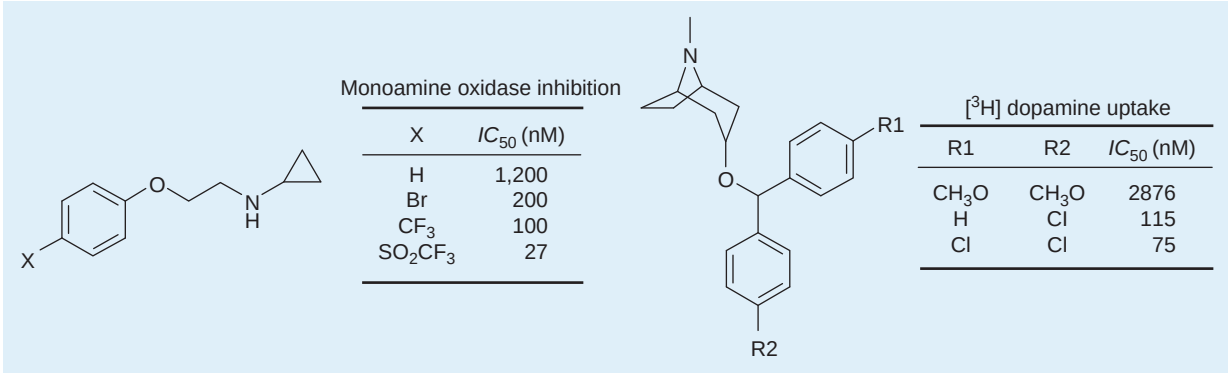
No.	Hypotensive activity ED_{20} (mg/kg)
1	0.01
2	3.00

Affinity for the benzodiazepine receptor	
R	IC_{50} (μM)
o-Cl	70.00
m-Cl	3.90
p-Cl	0.56

FIGURE 13.29 (a) The *ortho–ortho′* substitution in clonidine maintains the planes of the aromatic rings in a perpendicular position to each other; (b) The *ortho*–chloro isomer of the benzodiazepine ligand CGS 9896 has a 125-fold lower affinity than the parent molecule.

TABLE 13.9 van der Waals Radii [62]

	Radii in Å		**Radii in Å**
H	1.20	O	1.52
F	1.47	S	1.80
Cl	1.75	CH_3	1.80
Br	1.85	CF_3	2.20
I	1.98	$CH(CH_3)_2$	2.20



Monoamine oxidase inhibition

X	IC_{50} (nM)
H	1,200
Br	200
CF_3	100
SO_2CF_3	27

[^{3}H] dopamine uptake

R1	R2	IC_{50} (nM)
CH_3O	CH_3O	2876
H	Cl	115
Cl	Cl	75

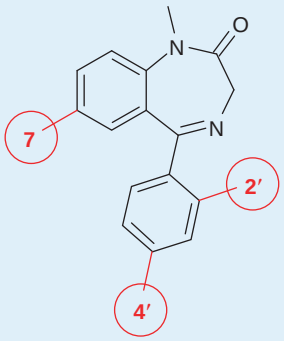
FIGURE 13.30 Influence of halogenated substituents on MAO inhibition potency *in vitro*.

A difference exists between the steric size of an atom or a groupment (the absolute size) and its steric effect (the volume at which its electronic effects are sensible).

2. *Electronic Effects*

The electronic effects of the halogens are ascribed to their inductive electron-attracting properties. These are maximal for chlorine and bromine, less marked for iodine, and very weak for fluorine. The mesomeric donor effect of the halogen atoms is usually not involved in biological media. The influence of halogens on the potency of MAO inhibition and of dopamine uptake blockade *in vitro* are shown in Figure 13.30. The choice of the optimal substituent allows noticeable gains in potency compared to the parent molecule [63,64].

Progressive mono- and di-substitution of diazepam-related benzodiazepinones enhances the affinity for the mitochondrial benzodiazepine receptor (MBR) by a factor of 233 (Figure 13.31) [65].



Compound	Substitution pattern 2′	4′	7	IC_{50} nM MBR
Ro 5-3464	H	H	H	700
Ro 5-5115	H	Cl	H	54
Diazepam	H	H	Cl	72
Ro 5-6900	Cl	H	Cl	11
Ro 5-4864	H	Cl	Cl	3

FIGURE 13.31 Chlorine effects in the benzodiazepine series [65].

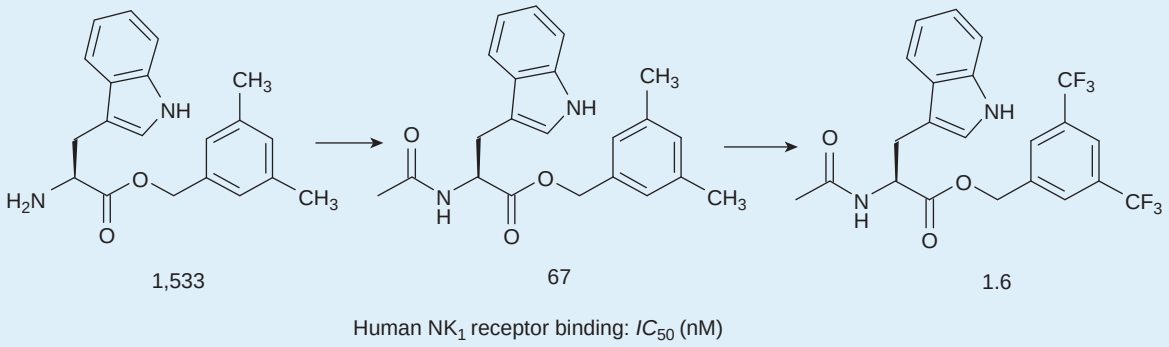


FIGURE 13.32 Successive *N*-acetylation and CH3→CF3 replacement achieve a 958-fold increase in affinity [65].

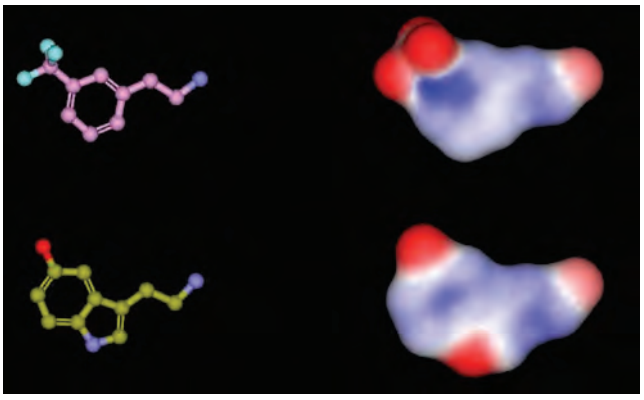


FIGURE 13.33 Electrostatic potential maps of *m*-trifluoromethyl-phenylethylamine and 5-hydroxy-tryptamine (Generated with DS Viewer Pro 6.0).

N-acylation of L-tryptophan benzyl esters, followed by replacement of the 3,5-dimethyl groups by their trifluoromethyl analogs, achieved an almost 1,000-fold increase in potency in a series of substance P receptor antagonists (Figure 13.32) [65].

3. *Electrostatic Similitude*

In certain active molecules, the role of the fluorine or chlorine atoms is not apparent at first glance. For example, two compounds such as *m*-trifluoromethyl-phenylethylamine and 5-hydroxy-tryptamine, which are chemically different, show many pharmacological similarities. In this case, the explanation lies in the similitude of the electrostatic potential maps (Figure 13.33).

Compound	log *P*	Compound	log *P*
$C_2H_5—O—C_2H_5$	0.89	$CHBr_2Cl$	2.16
$CHCl_3$	1.97	$CHBr_3$	2.40
$CHBrCl_2$	2.00	$CF_3CHBrCl$	2.30

The anaesthetic activity increases in the same order

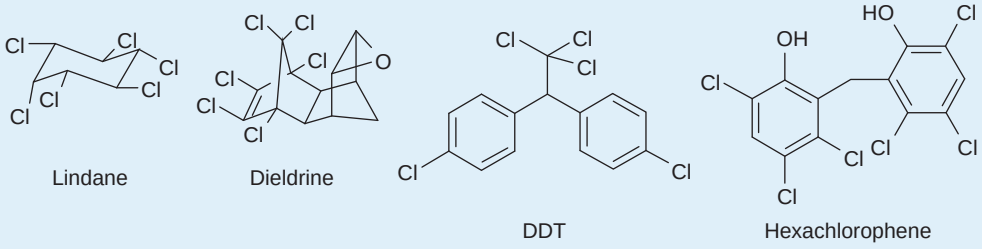


FIGURE 13.34 Compounds in which the halogens play essentially a lipophilic role.

TABLE 13.10 Selected Values of Hansch's π Parameter [69]

X	π	X	π
H	0.00	OH	− 0.67
F	0.14	OCH_3	− 0.02
Cl	0.71	OC_2H_5	0.38
Br	0.86	$COCH_3$	− 0.55
CH_3	0.56	CF_3	0.88
C_2H_5	1.02	NH_2	− 1.23

Conversely, two closely related pyrazoloquinolines, compounds CGS 8216 and its *para*-chloro analog CGS 9896, present totally opposed activity profiles on the same benzodiazepine receptor [66]. A dramatic effect resulting from chlorine substitution is also found in the change from β-phenyl-GABA to β-(*p*-chlorophenyl)-GABA [67].

4. Hydrophobic Effects

The predominantly lipophilic influence of halogen substitution is seen in the classical cases of the halocarbon anaesthetics [68], the halogenophenol antiseptics, and the halogenated insecticides (Figure 13.34).

These compounds show a direct correlation between biological activity and certain physicochemical parameters, such as partition coefficient, surface tension, or vapor pressure. The accumulation of halogen atoms favors the passage of the biomembranes and access to the CNS. Hansch has defined a parameter, π, representing the hydrophobic contribution of a substituent. It measures the individual participation of each group to the partition coefficient of the molecule (see Chapter 12). Table 13.10 shows π values for several substituent groups on bis-substituted phenyl rings.

As is evident, even if the participation of fluorine is modest compared to hydrogen, introducing halogens like chlorine (Figure 13.35) [68], bromine, or a trifluoromethyl group induce an important lipophilic contribution (larger than the contribution of a methyl group) [69].

5. Reactivity of the Halogens

In terms of bond strength, all C-halogen bonds are weaker than the C—H bond (Table 13.11), except for the C—F bond, due to the high electronegativity of the fluorine and an orbital size similar to that of carbon [71].

B. Usefulness of the Halogens and of Cognate Functions

Depending on their physical properties and their reactivity, the derivatives of fluorine, chlorine, bromine, and iodine present various degrees of usefulness (Table 13.12).

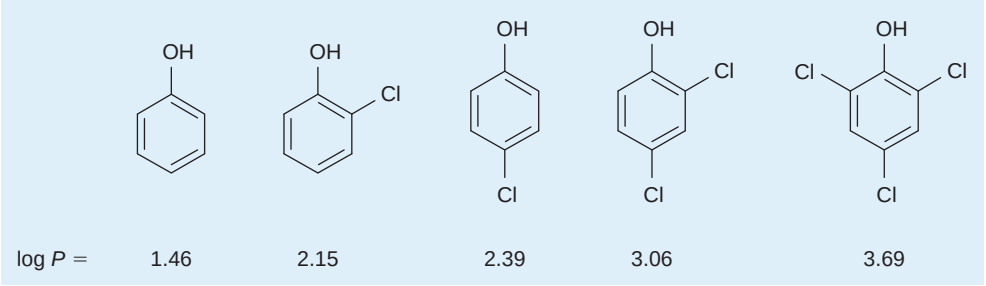


FIGURE 13.35 Chlorophenols experimental log *P* values [68].

TABLE 13.11 Atomic Radii and Characteristics of Carbon-Halogen Bonds [70]

Atomic radius (Å)	Bond[a]	Interatomic distance (Å)	Bond strength (kcal/mol)
H: 0.29	C—H	1.14	93
F: 0.64	C—F	1.45	114
Cl: 0.99	C—Cl	1.74	72
Br: 1.14	C—Br	1.90	59
I: 1.33	C—I	2.12	45

[a]*In aliphatic series.*

TABLE 13.12 Some Aromatic Substituent Constants for Halogens and Equivalent Functions [17]

Group	π	π	MR
H	0.00	0.00	1.03
F	0.14	0.06	0.92
Cl	0.71	0.23	6.03
Br	0.86	0.23	8.88
I	1.12	0.18	13.94
CF_3	0.88	0.54	5.02
CH_3	0.56	− 0.17	5.65
CN	− 0.57	0.66	6.33
SO_2CF_3	0.55	0.93	12.86
SCF_3	1.44	0.50	13.81
SCN	0.41	0.52	13.40

1. *The Case of Fluorine*

Along with chlorine, fluorine is one of the most used halogens in medicinal chemistry (around 20 percent of used medicines and 28 percent of the agrochemicals include fluorine) when they are attached to a non-activated carbon atom.

For a long time, fluorine was considered a bioisostere of hydrogen, an idea which is no longer accepted. Indeed, fluorine is very different from hydrogen, with a van der Waals radius comparable to that of oxygen. It induces an increase in lipophilicity, and its electronegativity is the highest in the periodic classification.

The difference in electronegativity between fluorine and carbon creates a large dipole moment in this bond. This dipole may contribute to the molecule's ability to be engaged in intermolecular interactions. Fluorine is able to participate in hydrogen bonds with the hydrogen of water. These bonds are weaker than those obtained with

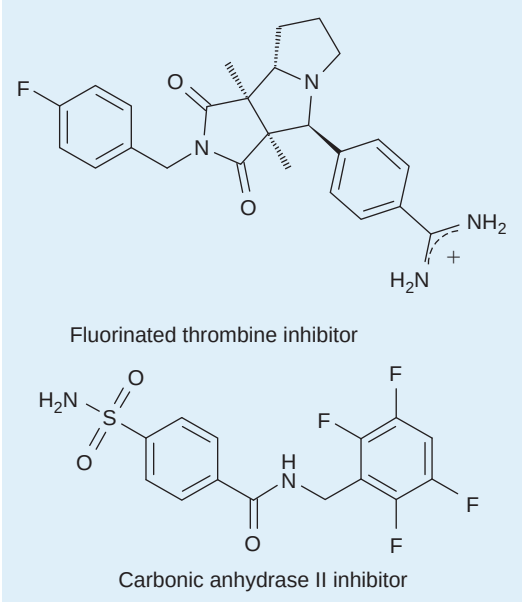


FIGURE 13.36 Example of fluoroaromatic compounds in which an intermolecular participation of the fluorine has been proposed [72].

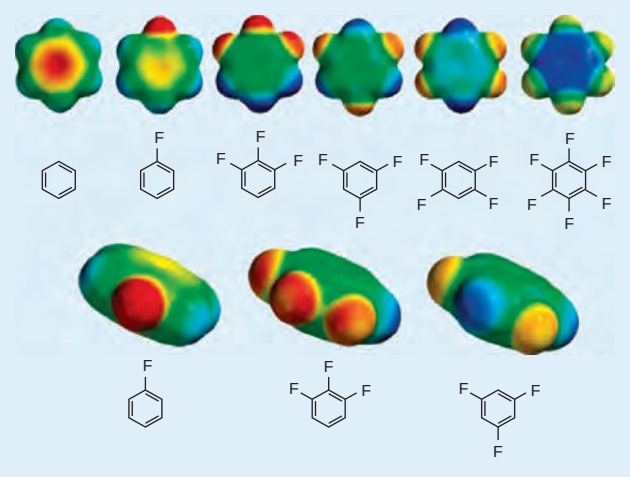


FIGURE 13.37 Electrostatic potential surface of fluorophenyls, view from the front and the edge. Source: *Taken from Ref. [72].*

oxygen, but they are still strong enough to contribute to the binding of fluoroaromatic compounds to active sites and receptors (Figure 13.36) [72].

The importance of electrostatics in the interaction of aromatics fluorine with cations and hydrogen bond donors can be visualized using electrostatic potential surfaces (Figure 13.37) [72]. In monofluorobenzene, the potential of the fluorine is concentrated on the unique fluorine present, whereas in polyfluorobenzene, the negative charge is spread over several fluorine atoms. For this reason, monofluorobenzene may give stronger interactions.

Another major use of fluorine is to block metabolically sensitive positions of a molecule (for example the para position of a phenyl group). The CF3 group is more or less comparable to chlorine over various parameters (van der Waals Radii, Hansch's π parameter, molecular refraction) and can advantageously replace it, particularly when halogen properties should be available on an activated position (e.g., R—CO—Cl → R—CO—CF_3), like for the antineoplastic valrubicin.

2. *The Case of Chlorine*

A chlorine substituent simultaneously produces an increase in lipophilicity, an electron-attracting effect, and a metabolic obstruction. The discovery of bicalutamide, a nonsteroidal selective androgen receptor modulators

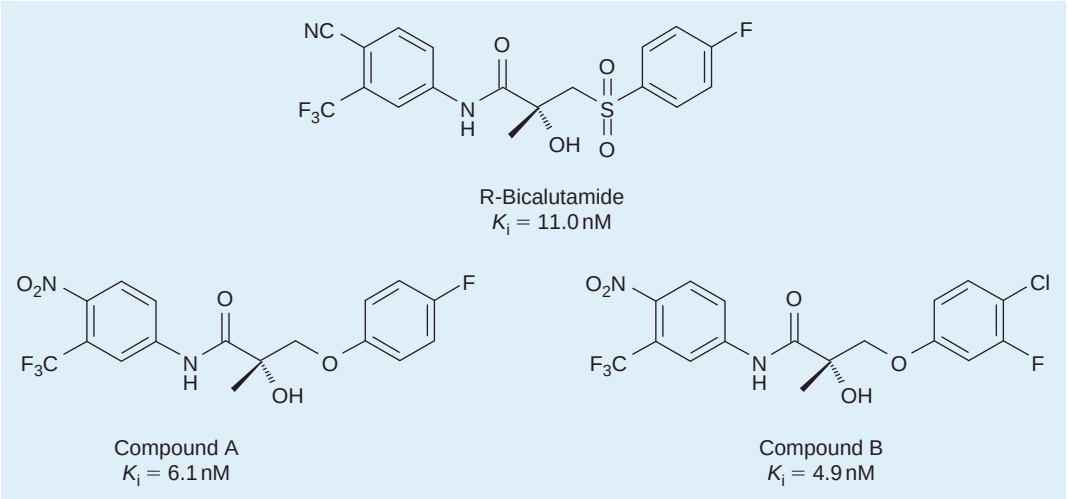


FIGURE 13.38 Structure and binding affinity to androgen receptor of nonsteroidal SARMs [73].

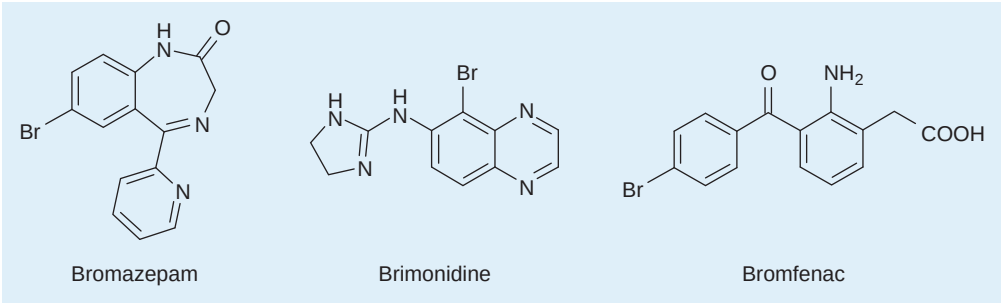


FIGURE 13.39 Bromoaryl compounds.

(SARM) gave birth to a series of analogs such as compounds A and B (Figure 13.38). Chen et al have found that compared to compound A, compound B has an higher affinity, an improved half-life, a smaller volume of distribution, and a lower clearance [73]. The improved half-life and lower clearance might be explained by the ability of the two halogens to prevent metabolism [74].

3. *The Case of Bromine*

Bromine is the least used halogen. When it serves, it is usually incorporated as a bromo-aryl. The reproach against bromine is to generate reactive alkylating intermediates more easily than chlorine or fluorine. Therefore, in long-term treatment it can confer toxic potentialities to the molecule that bears it. This was the case for the anti-inflammatory-analgesic drug bromfenac sodium, withdrawn from the US market due to reports of hepato-toxicity [75].

Other bromoaromatic compounds have been marketed, such as the anxiolytic bromazepam and brimonidine, which used to lower pressure in the eyes of patients who have glaucoma and ocular hypertension (Figure 13.39).

4. *The Case of Iodine*

Although tolerated even less well than bromine, iodine is indispensable to the treatment of certain thyroidal deficiencies. Administrated by internal route, iodine and iodine derivatives trigger either acute hypersensitivity reactions (e.g., larynx oedema, cutaneous hemorrhages, fever, arthralgies) or chronic reactions (iodism). In addition to its use in certain dysfunctions of the thyroid gland, iodine presents two specific uses: (1) covalent iodine derivatives serve as radiological contrast substances; and (2) 131iodine (half-life: 8 days) is used as radioactive tracing agent.

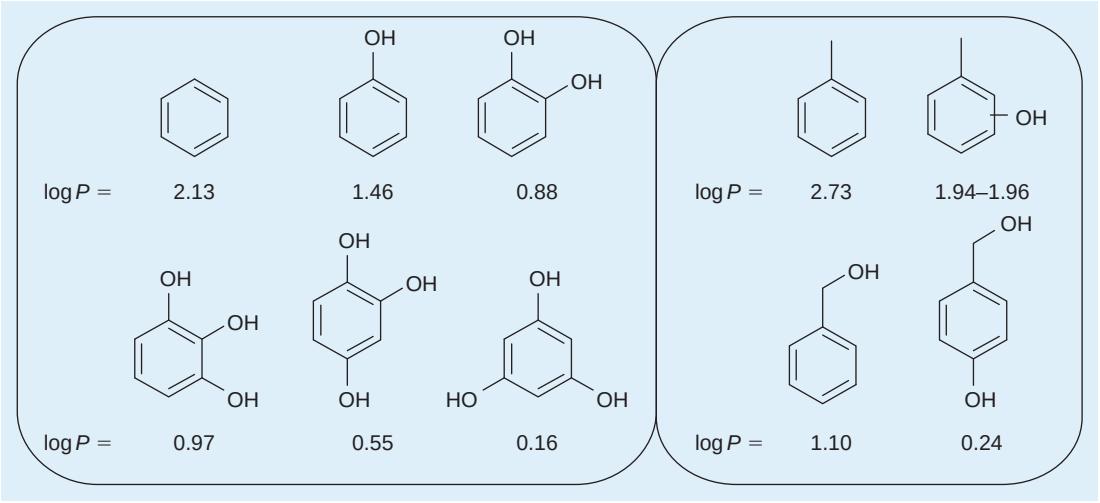


FIGURE 13.40 Experimental log *P* for various hydroxyaryl compounds [68].

5. *Extensions-Cognate Groups*

Chlorine, trifluoromethyl, cyano or azido groups are more or less bioisosteres. Other possible candidates are SCN, SCF_3, SO_2CF_3, and $CH{=}CF_2$ (see Chapter 8).

V. EFFECTS OF HYDROXYLATION

The substitution of OH for H affects biological activity profoundly, as in the conversion of ethane to ethanol or benzene to phenol. Simple alcohols have narcotic effects, and simple phenols have bacteriostatic properties. Polyfunctional compounds can act as chelating or complexing agents.

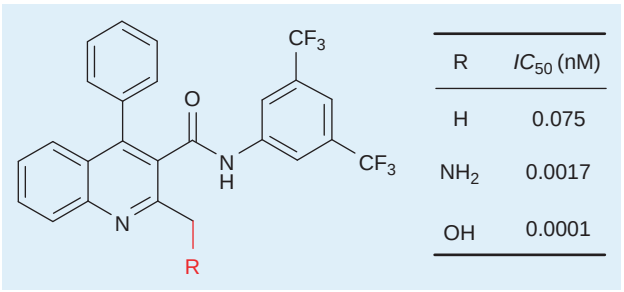
A. Effects on Solubility

The introduction of an alcoholic or a phenolic hydroxy group into an active molecule changes the partition coefficient toward more hydrophilicity and renders the molecule more water soluble (see Figure 13.40). The value of the Hansch π constant for an aromatic hydroxy group is -0.67. That means that the introduction of an hydroxy group modifies the physical properties of a compound, and particularly the hydrophobic contribution (Hansch parameter). If one wants to keep this parameter unchanged, for example to avoid the loss in lipophilicity, it is necessary to add either an atom like a chlorine ($\pi = -0.71$) or a methyl or a CF3 group ($\pi = 0.56$ and 0.88 respectively) to compensate the presence of the hydroxy.

An interesting parameter has to be noted about the hydrophilicity of trihydroxybenzenes and—by extention—of other polyhydroxylated compounds. The fact that the hydroxyl groups of 1,2,3-trihydroxybenzene, and to a lesser extent those of 1,2,4-trihydroxybenzene, are able to form intramolecular hydrogen bonds leaves these hydroxyls less available to generate hydrogen bonds with water, making them more hydrophobic than 1,3,5-trihydroxybenzene.

B. Effects on the Ligand–Receptor Interaction

For some hydroxylated drugs like morphine, dopamine, haloperidol, γ-hydroxybutyrate, serotonine, or most of the steroids, the hydroxy group is an essential element for hydrogen bonding with the receptor. For others, the attachment of a hydroxy group can result in potency changes. Examples are found in hycanthone, which is ten times more active against schistosomes than lucanthone [76], or in hydroxylated minaprine analogs, which show a ten-fold better affinity for M_1 muscarinic receptors than the parent drugs [77]. One spectacular example of hydrogen bonding with hydroxyl (or amino) group has been shown by Cappelli et al (Figure 13.41) [78]. By the simple replacement of an hydrogen by an hydroxyl group, they increased 750-fold the binding affinity of the compound, allowing it to show subpicomolar affinity.



R	IC_{50} (nM)
H	0.075
NH_2	0.0017
OH	0.0001

FIGURE 13.41 Structure and binding affinity of 3-quinolinecarboxamides to human endogenous NK1 receptor in UC11MG cells [78].

C. Hydroxylation and Metabolism

As a rule, metabolic hydroxylation of an active compound represents a detoxification (phase I) mechanism. It generally results from a first-pass effect and can be followed or not by a conjugation reaction (see Chapter 32). Classical examples of drugs detoxified through hydroxylation are paracetamol, oxyphenbutazone, and hydroxychloroquine. Other important reactions of hydroxy compounds, whether alcoholic or phenolic, are based on their capacity to accept activated groups through the action of group-transferring enzymes (e.g., methylation, sulfation, phosphorylation, glycosylation).

VI. EFFECTS OF THIOLS AND OTHER SULFUR-CONTAINING GROUPS

Thiol and disulfide groups occur widely in natural products. They are found in small molecules such as lipoic acid, glutathione, and thiamine, as well as in cysteine-containing peptides and proteins (hormones, enzymes, antibiotics). In all these substances, thiol and disulfide groups are clearly associated either with high chemical reactivity or with the consolidation of peptide and protein architecture. Being too reactive, the thiol and the disulfide groups are normally not used in medicinal chemistry as substituents in quantitative structure–activity relationships (QSAR) studies. Occasionally, methylthio substitution on aromatic rings is practiced, but even then the obtained thioethers are very reactive. They are easily converted to sulfoxides and *vice versa* (see Chapter 25).

A. Drugs Containing Thiol

Drugs containing thiol groups are mainly used for the strong affinity that the thiolate anion presents toward heavy metals. This is the case for thiol-containing angiotensin-converting enzyme inhibitors, which bind to a zinc-containing enzyme (see Chapter 5, and Ganellin and Roberts [79]). Captopril, a marketed drug used in the treatment of arterial hypertension and congestive heart failure, was obtained after manipulation of succinylproline. However, some adverse reactions can occur due to the presence of the mercapto group which led to the replacement of the thiol group by a carboxyl group. Finally, another potent angiotensin-converting enzyme was obtained, enalapril.

Presently, other captopril analogs, like zofenopril (a prodrug in which the cleavage of the thioester led the active compound zofenoprilat), are still marketed. They are able to coordinate Zn(II) and act as free radical scavengers. Their lipophilicity allows them to attain high levels in heart tissue, and they can therefore be used as cardioprotective drugs (Figure 13.42) [80].

The heavy-metal chelating properties of thiols were taken advantage of in the design of dimercaprol ("British Anti-Lewisite," BAL) as counter poison of the arsenical war gas lewisite (Figure 13.43). Today dimercaprol is used to treat poisoning by compounds of gold, mercury, antimony, and arsenic. The toxic nature of the heavy metals is masked, and chelate is stable enough to be excreted as such in the urine.

Penicillamine (D-β,β-dimethylcysteine) is an effective chelator of copper, mercury, zinc, and lead that promotes the excretion of these metals in the urine. It is clinically used in patients with Wilson's disease, with rheumatoid arthritis, and with heavy-metal intoxications [81]. Ziram and ferbam are the zinc and the iron salts respectively of dimethyl-dithiocarbamic acid (Figure 13.44). They are widely used as selective fungicides in agriculture. Pyrithione (1-hydroxy-2(1*H*)-pyridinethione), as its zinc salt, is used in dermatology as antiseborrheic.

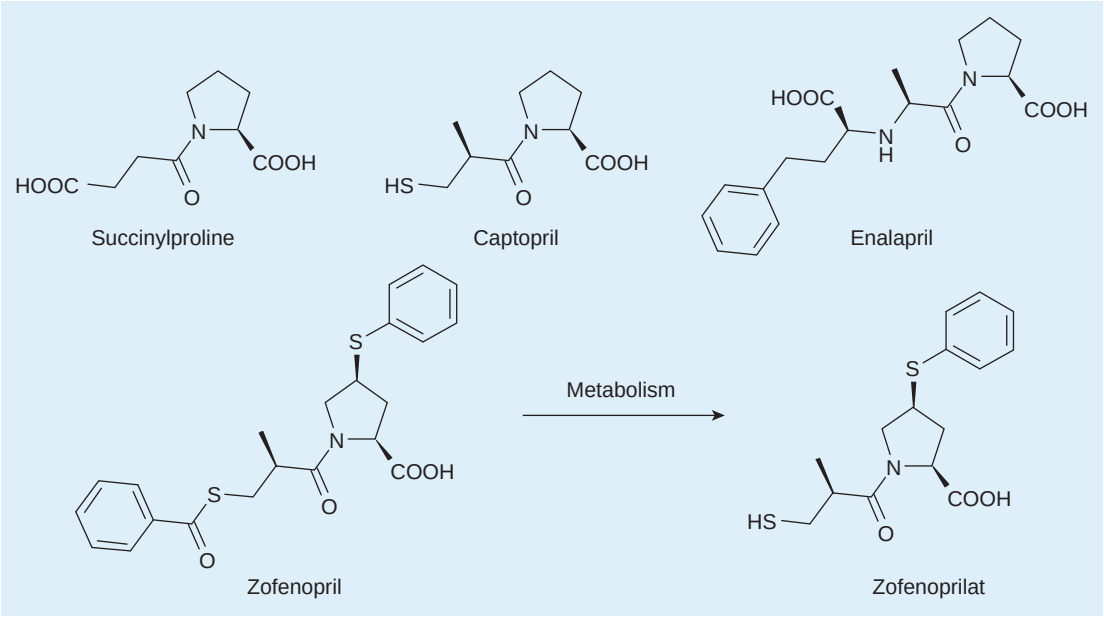


FIGURE 13.42 Antihypertensive drugs.

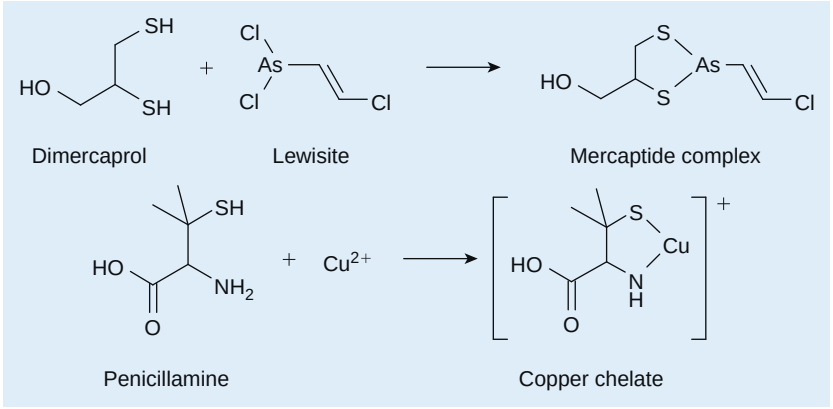


FIGURE 13.43 Chelating properties of thiol derivatives.

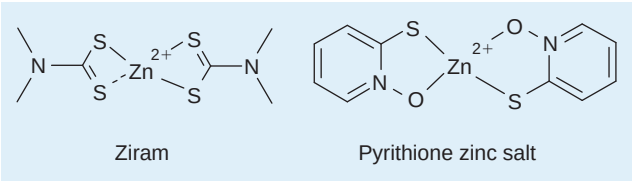


FIGURE 13.44 Ziram and pyrithione-zinc are chelated salts.

B. Drugs Containing Oxidized Sulfides

The sulfoxide (S═O) and the sulfone (O═S═O) functions are very polar and usually confer mediocre CNS bioavailability. But this characteristic should not be a barrier to the use of this type of derivatives. As an example, the triptan class of compounds does generally have poor brain penetration data, but these characteristics are not a good guide to central activity, especially with very potent drugs such as the triptans (in the nanomolar range on 5-$HT_{1B/1D}$ receptors). Indeed, in contrast to most other CNS agents that are antagonists, triptans are agonist and so will require only low fractional receptor occupancy to exert central effects [82]. As another

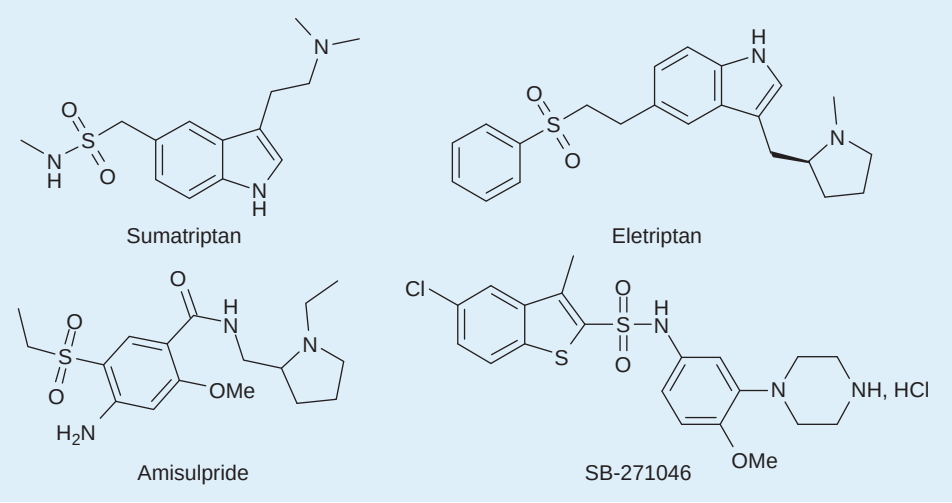


FIGURE 13.45 Compounds acting on CNS receptors.

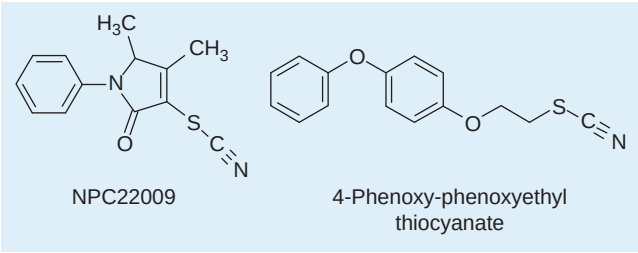


FIGURE 13.46 Covalently bound thiocyanates [84,85].

example, sulfonamides presenting moderate brain penetration (10 percent) are known, like the 5-HT_6 receptor antagonist SB-271046 (Figure 13.45) [83].

Many other sulfone containing drugs are marketed, most for antibacterial use but also as cardioprotective agents, antihypertensives, analgesics...

C. Drugs Containing Thiocyanate or Thiourea

Covalently bound thiocyanates are rather unusual. Recent examples are the corticotropin-releasing factor (CRF) antagonist NPC 22009 [84] and 4-phenoxy-phenoxyethyl thiocyanate (Figure 13.46), which acts as inhibitor of the sterol biosynthesis in *Trypanosoma cruzi* [85], the protozoon agent responsible for Chagas disease. Sometime thiocyanates are used as counteranions for the preparation of pharmaceutically acceptable salts [86].

Heterocyclic thioureas such as 6-propylthiouracil, methimidazole and carbimidazole (Figure 13.47) are used as antithyroid drugs. They inhibit the formation of thyroid hormones. One of the presumed mechanisms is the inhibition of the iodine incorporation into the tyrosyl residues of thyroglobulin. It was proposed that the iodine atom is bound to a protein as a sulfenyl iodide. The thioureas may act in establishing covalent —S—S— bonds and in displacing iodine as HI [87].

Some thioureas are also used in short-acting anesthetic compounds, like thialbarbital, thiamylal, and thiopental. These compounds are used to induce anesthesia before the injection of anesthetic products and are now slowly being replaced by propofol (Figure 13.48). More generally, thiourea and its simpler aliphatic derivatives, as well as thioamides, produce goiter and have to be avoided in drug design. A natural compound, L-5-vinyl-2-thiooxazolidone (goitrin) is responsible for the goiter of cattle eating turnips or cruciferous plants [88].

Besides their affinity for metallic ions, thiol groups have other characteristics such as the ability to interconvert to disulfides through redox reactions, to add to conjugated double bonds, and to form complexes with the pyridine nucleotides by nucleophilic attack at the 4-position of the pyridine ring.

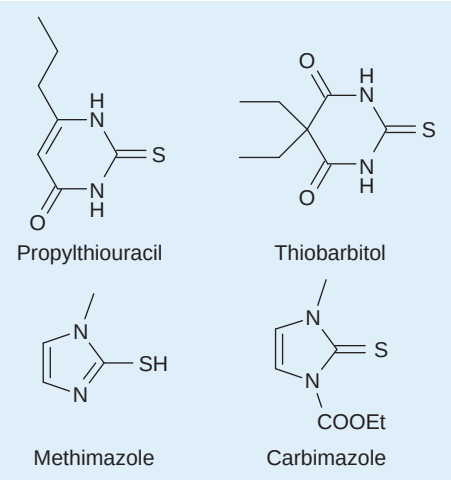


FIGURE 13.47 Heterocyclic thioureas with antithyroid effects.

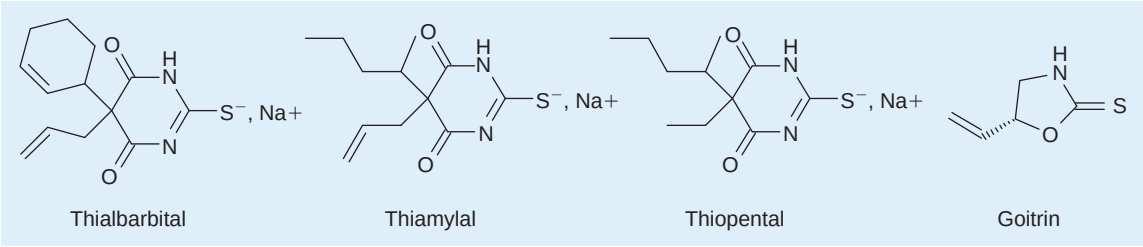


FIGURE 13.48 Heterocyclic thioureas with anesthetic effects, and goitrin.

VII. ACIDIC FUNCTIONS

The prototypical representatives of the group are the carboxylic acids. However, a huge number of bioisosteres, such as sulfonic or phosphonic acids and tetrazoles or 3-hydroxyisoxazoles, are available (see Chapter 8). In addition, functions like esters, amides, peptides, aldehydes, primary alcohols, and related functions can work as prodrugs or bioprecursors (see Chapter 28).

A. Effects on Solubility

The introduction of an acidic group into a biologically active compound that contains no such group has as its consequence essentially a solubilizing effect. This effect can even be enhanced through salt formation. Carboxylic acids are often highly ionized at the physiological pH values, and this is even more the case for sulfonic acids. As a rule, strong and highly ionized acids cannot cross the biological membranes, which are permeable only to nondissociated molecules or in some case to molecules whose charge is very distributed; they are then subject to a rapid clearance from the animal body. However, once absorbed, they can establish strong ionic interactions with the basic amino acids, especially with lysine, contained in the blood serum albumine, or the enzyme and receptor proteins.

A visible effect of the solubility modifications happening after introduction of a carboxylic group is found in the history of antihistaminic compounds (Figure 13.49). The first generation of antihistaminic drugs, such as hydroxyzine, were lipophilic compounds. They were able to cross the blood-brain barrier (BBB) and had a sedating action because they were not P-glycoprotein (P-gp) substrate (which means they were not considered as xenobiotics and pumped out of the brain). Nowadays, hydroxyzine is still used as anxiolytic.

The second generation of antihistaminic compounds are less lipophilic, thanks to the replacement of the hydroxy group by a carboxylic acid. They are also P-gp substrate, which limits CNS exposure [89,90].

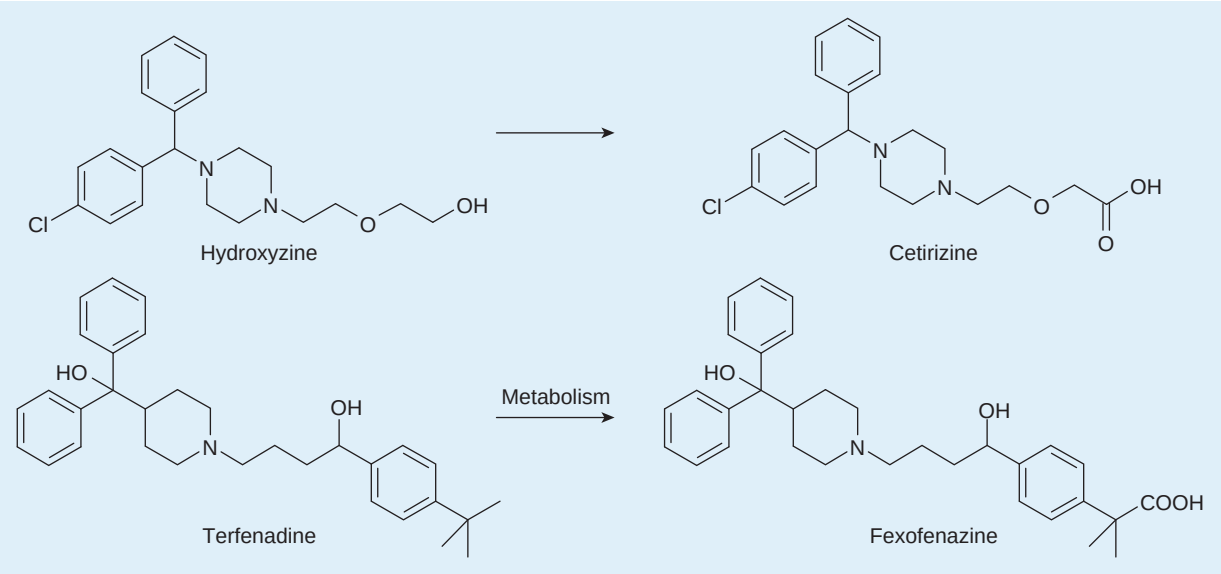


FIGURE 13.49 Antihistaminic compounds.

B. Effects on Biological Activity

Changes in biological activity distinguish the sulfonic from the carboxylic acids. Broadly speaking, the sulfonic acids as a class are generally not biologically active. Exceptions are certain complex dyes or trypanocides (e.g., Trypan blue, suramin) and sulfonic amino acids such as taurine and hypotaurine, for which an active transport mechanism exists. For carboxylic acids the situation depends on whether the carboxylic function is introduced in small or large molecules.

1. In Small Molecules

The introduction of a carboxylic group fundamentally changes the biological activity. Very often the initial biological activity is destroyed and the toxicity of the parent compound is reduced. In a series of cyproheptadine analogs, the replacement of a chlorine substituent on a benzo ring by a carboxylic group resulted in a 4,000-fold loss in affinity for the spiperone-labeled dopamine receptor [91]. Conversely, the presence of the carboxylic group can sometimes create the conditions necessary for activity (Figure 13.50).

2. In Large Molecules

High pharmacological activity is maintained despite the presence of the carboxylic group. Examples are the anti-inflammatory arylacetic acids, the prostaglandins, cromolin and related anti-asthmatics, and finally the β-lactam antibiotics. In these drugs, the relative weight of the carboxyl is notably smaller. An illustrative example is found in a series of acids and esters derived from coformycin and acting as AMP deaminase inhibitors (Figure 13.51). In this series, the free acid is clearly more potent than the corresponding ethyl ester [92].

The discovery of losartan (Figure 13.52) and the numerous analogs obtained after (e.g., candesartan, eprosartan, irbesartan, olmesartan, tasosartan, telmisartan, valsartan) [93,94] is an example of manipulation of acidic groups.

Starting with the benzyl-imidazole (compound 1), which binds to angiotensin II (AII) receptor but exhibits a very poor potency, Duncia et al [93] supposed that a second acidic group (compound 2) would provide, at physiological pH, a negative charge in the area of space where Tyr-4 and Asp-1 residues reside. This substitution increased 10-fold the potency of the compound. Another 10-fold increase has been obtained by the introduction of a second phenyl ring between the two first rings (compound 3). The replacement of the carboxylic acid group by a more lipophilic acid isostere (compound 6) gave some oral activity but much less active than by intravenous injection. Finally, a greater oral activity was obtained with the tetrazole substituent, due in part to the greater oral bioavailability, and rendered possible the development of the compound losartan (compound 7) [94].

Antitumor sulfonamides targeting G1 phase provide an *a contrario* example. Despite the relatively large size of the molecule, the presence of a carboxylic function completely abolishes the *in vitro* antiproliferative activity [95], whereas the corresponding carboxamide and sulfonamide are highly active (Figure 13.53).

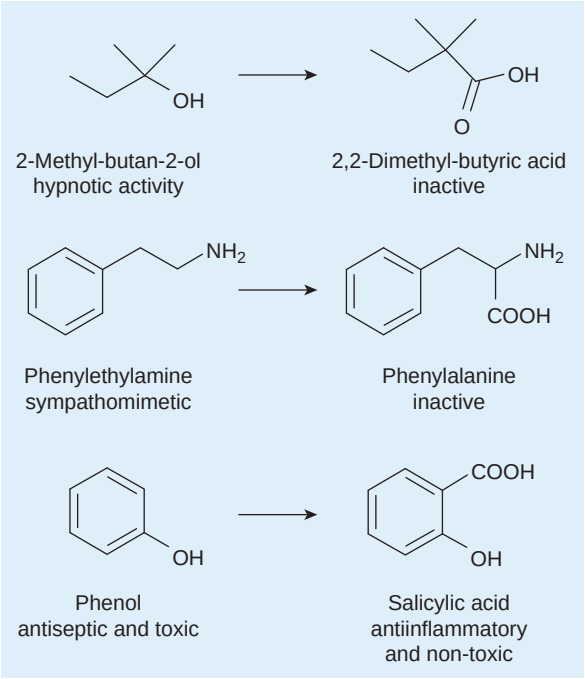
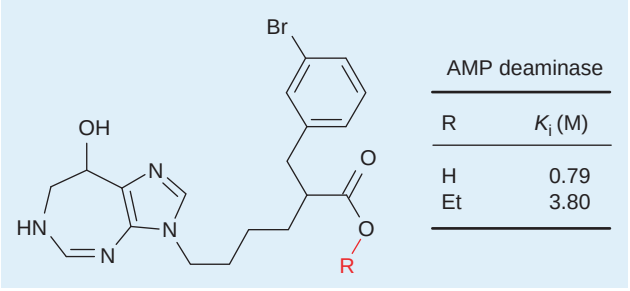


FIGURE 13.50 Introduction of a carboxylic group in a small molecule can destroy the activity originally present or, conversely, create the conditions necessary for activity.



AMP deaminase	
R	K_i (M)
H	0.79
Et	3.80

FIGURE 13.51 The free carboxylic acid is more active than the corresponding ethyl ester.

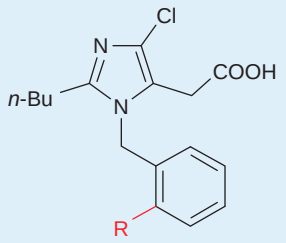
3. *Other Carboxylic Functions*

With carboxyl-derived functions such as esters and amides, the initial activity of the drug, lost in introducing the carboxyl group, is often regained. Amides, ureides, hydantoins, and barbiturates share CNS-depressing properties and are frequently indispensable elements of sedative, tranquillizing, and anticonvulsant drugs. Nitriles as substituents are often comparable to chlorine atoms but are sometimes more toxic.

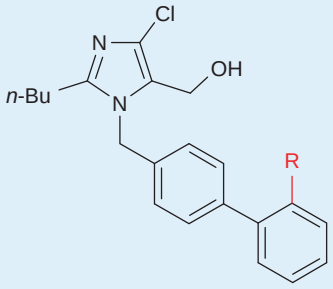
VIII. BASIC GROUPS

The basic groups met in medicinal chemistry are the amines, the amidines, the guanidines, and practically all nitrogen-containing heterocycles. Basic groups are polar, and one would expect that highly ionized bases (especially quaternary ammonium salts) would resemble the sulfonic acids and show limited activity due to their mediocre membrane permeability. In practice bases with pK_a values superior to ten have very limited chance to reach the CNS.

As seen for the acidic groups, the introduction of a basic group into a biologically active compound that contains no such group already has a solubilizing effect. This effect can also be enhanced through salt formation. In drug–protein interaction,s the classical counteranions of organic bases are the aspartic and the glutamic carboxylates.

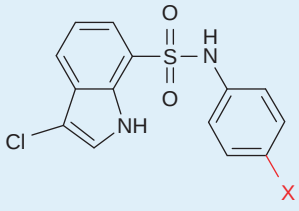


Compound	R	IC_{50} (nM)
1	H	150
2	COOH	1.6



Compound	R	Est. pK_a	IC_{50} (nM)
3	COOH	5	0.23
4	$COONHSO_2Ph$	8.44	0.14
5	$NHCOCF_3$	9.5	6.3
6	$NHSO_2CF_3$	4.5	0.083
7	Tetrazole	5–6	0.019

FIGURE 13.52 Genesis of losartan [93,94].



X	Antiproliferative activity IC_{50} (g/mL) Colon 38 murine adenocarcinoma	P388 murine leukemia
COOH	>100	>100
$CONH_2$	0.19	0.87
SO_2NH_2	0.10	0.45

FIGURE 13.53 The antitumor activity of the carboxamide and of the sulfonamide is completely abolished for the corresponding free carboxylic acid [92].

The biological activity of amines and basic heterocycles is immense and justifies the adage "no biological activity without nitrogen." Steroids, prostaglandins, and nonsteroidal anti-inflammatory drugs are, of course, exceptions. Primary amines often demonstrate less specific effects than secondary or tertiary amines. Acylation deactivates the amines strongly, as does the introduction in some other places of the molecule of a carboxylic or sulfonic group (formation of zwitterions: bipolar ions). Diamines and polyamines are usually more active than monoamines. Aromatic amines are always more hazardous than aliphatic amines and form toxic metabolites. Examples are 2-naphtylamine, benzidine, aniline (see Chapter 25). They are easy to detoxicate by introducing a carboxyl group, as evidenced by the change from aniline to the nontoxic *p*-aminobenzoic acid.

IX. ATTACHMENT OF ADDITIONAL BINDING SITES

A. To Increase Lipophilicity

Many polar active molecules selected through *in vitro* screening tests are unable to cross the biomembranes, and their bioavailability is particularly low. Attaching a very lipophilic moiety can sometimes help to overcome this drawback.

A typical example is given by the development of the anticonvulsant drug thiagabine [90] starting from nipecotic acid (Figure 13.54). Cyclic amino acids such as nipecotic acid and guvacine have been shown to inhibit GABA uptake. However, these small amino acids do not readily cross the BBB and thus limit their potential clinical usefulness. A considerable improvement has been the discovery by Yunger et al [96,97] of compound SKF 89976A, a *N*-(4,4-diphenyl-3-butenyl) substituted nipecotic acid. *In vitro*, this compound is approximately twenty times more potent than the parent amino acid as inhibitor of [3H] GABA uptake. Moreover, it is active orally as an anticonvulsant in mice and rats with ED_{50}s around 8 mg/kg.

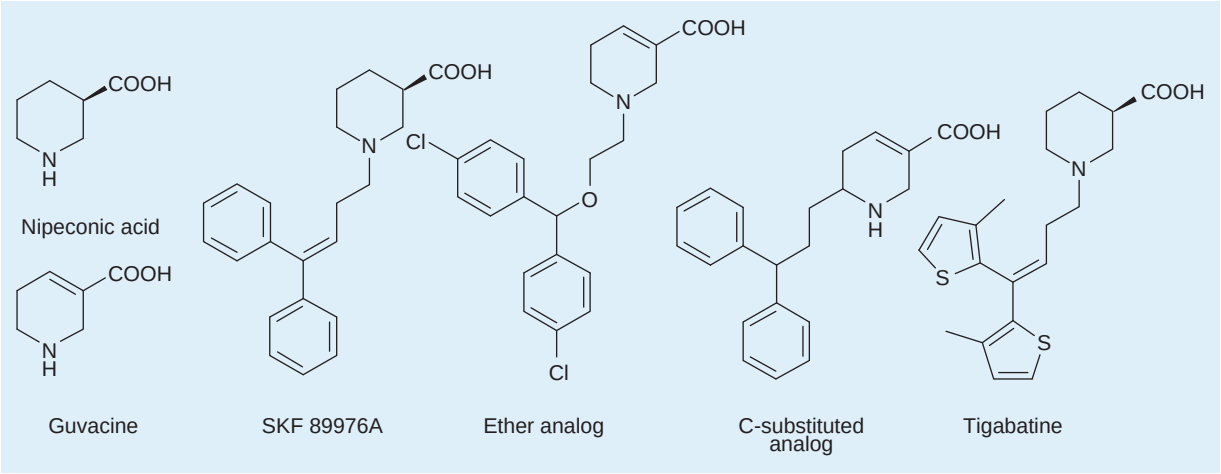


FIGURE 13.54 Lipophilic derivatives of nipecotic acid and of guvacine.

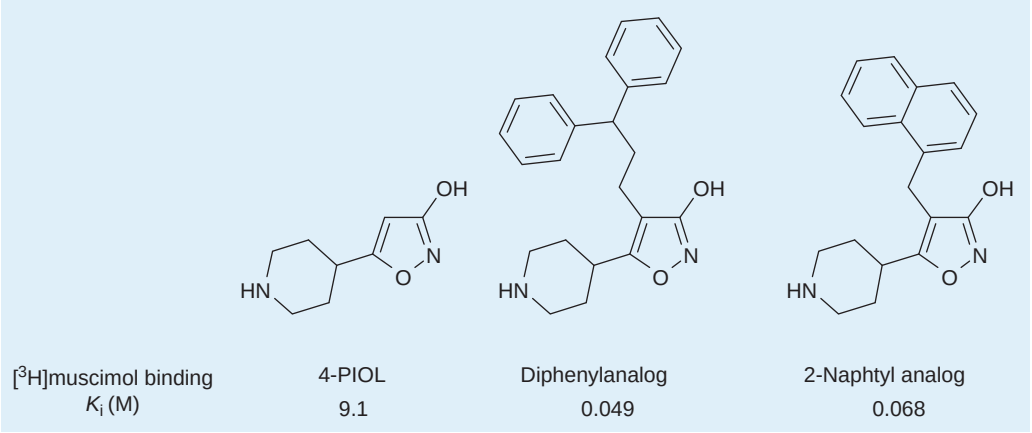


FIGURE 13.55 Large aromatic substituents attached to 5-(4-piperidyl)-3-isoxazolol.

Further developments demonstrated that the 4,4-diphenyl-3-butenyl moiety could be replaced by ether-type analogs [98,99] and that the attachment of a lipophilic 3,3-diphenylpropyl side-chain can take place at the carbon atom at the 6-position of the amino acid [100]. In this latter case, however, despite reasonable *in vitro* activity (IC_{50} = 100 nM), no *in vivo* activity is observed. Final optimization of SKF 89976A led to the bis-thiophene tiagabine [101]. In a similar way, ω,ω-diphenyl-alkyl chains (butterflies) were also attached to L-glutamic acid to yield glutamate metabotropic m3 receptor-selective agonists [102].

B. To Achieve Additional Interactions

The fixation of large aromatic substituents such as 2-naphtyl and 3,3-diphenylpropyl to the low-efficacy partial agonist 4-PIOL transforms this series into powerful $GABA_A$ receptor antagonists. This improvement is due to the substantial size of the cavity and the strong interactions between the aromatics and the binding cavity of the $GABA_A$ receptor (Figure 13.55) [103].

Histaprodifen, a 3,3-diphenylpropyl substituted analog of histamine, is a potent histamine H_1 receptor agonist. The increase in potency between histamine and histaprodifen may be due to a new orientation of the imidazole ring caused by the space filling substitution at the 2-position. This new orientation allows an additional interaction with the H_1 receptor [104].

Instead of ensuring high lipophilicity, these aralkyl groups serve to achieve additional interactions with the target macromolecule. This is typically the case for the angiotensin-converting enzyme inhibitor enalaprilat. The exchange

in captopril of the thiol function for a carboxylic group as ligand for the enzyme zinc atom entails an important decrease in activity (Chapter 4). This decrease could be compensated by the attachment of a phenethyl moiety.

Several marketed compounds are using the "butterfly" chains with different spacer length, with substituents on the phenyl rings (most often fluorine) and/or substituents onto the carbon bearing the two phenyls (alkyl, amide, hydroxyl, nitrile ...).

References

[1] Taylor J. The topliss approach-opinion of an industrial scientist on QSAR methods. Stasbourg: New York: Lecture given at the Louis Pasteur University; 1984.
[2] Rekker RF, Mannhold R. Calculation of drug lipophilicity. Weinheim: Verlag VCH; 1992.
[3] Witt KA, Davis TP. CNS drug delivery: opioid peptides and the blood-brain barrier. AAPS J 2006;8(1):E76–88.
[4] Ooms F, Wouters J, Oscari O, Happaerts T, Bouchard G, Carrupt PA, et al. Exploration of the pharmacophore of 3-alkyl-5-arylimidazolidinediones as new CB(1) cannabinoid receptor ligands and potential antagonists: synthesis, lipophilicity, affinity, and molecular modelling. J Med Chem 2002;45(9):1748–56.
[5] Gaillard P, Carrupt PA, Testa B, Boudon A. Molecular lipophilicity potential, a tool in 3D QSAR: method and applications. J Comput Aided Mol Des 1994;8(2):83–96.
[6] Ginnings P, Baum R. Aqueous solubilities of the isomeric pentanols. J Am Chem Soc 1937;59:1111–13.
[7] Nemethy G. Hydrophobic interactions. Angew Chem Int Ed Engl 1967;6(3):195–206.
[8] Hansch C, Anderson SM. The effect of intramolecular hydrophobic bonding on partition coefficients. J Org Chem 1967;32:2583–6.
[9] Gilligan D, Plummer N. Comparative solubilities of sulfadiazine, sulfamerazine, and sulfamethazine and their N4-acetyl derivatives at varying pH levels. Proc Soc Exp Biol Med 1943;53:142–5.
[10] Albert A. Selective toxicity. London: Chapman & Hall; 1979. p. 41.
[11] Harms AF, Nauta WT. The effects of alkyl substitution in drugs: I. Substituted dimethylaminoethyl benzhydryl ethers. J Med Pharm Chem 1960;2:57–77.
[12] Reggio P, Topiol S, Weinstein H. Molecular determinants for the agonist activity of 2-methylhistamine and 4-methylhistamine at H2-receptors. J Med Chem 1986;29(11):2412–15.
[13] Durant GJ, Ganellin CR, Parsons ME. Chemical differentiation of histamine H1- and H2-receptor agonists. J Med Chem 1975;18(9):905–9.
[14] Velluz L. La stucture nor-stéroïde [The structure of nor steroids].. In: Hazard R, Cheymol J, editors. Actualités pharmacologiques. Paris: Masson & Cie; 1961. p. 221–43. [in French].
[15] Wermuth CG, Schwartz J, Leclerc G, Garnier JP, Rouot B. Communication préliminaire: conformation de la clonidine et hypothèses sur son interaction avec un récepteur α-adrénergique [Preliminary communication: conformation of clonidine and assumptions about its interaction with an α-adrenergic receptor]. Eur J Med Chem 1973;8:115–16. [in French].
[16] Perlman S, Costa-Neto CM, Miyakawa AA, Schambye HT, Hjorth SA, Paiva AC, et al. Dual agonistic and antagonistic property of non-peptide angiotensin AT1 ligands: susceptibility to receptor mutations. Mol Pharmacol 1997;51(2):301–11.
[17] Chu KC. The quantitative analysis of structure–activity relationships. In: 4th ed. Wolf ME, editor. The basis of medicinal chemistry/Burger's medicinal chemistry, vol. I. New York: John Wiley & Sons; 1980. p. 393–418.
[18] van de Waterbeemd H, el Tayar N, Carrupt PA, Testa B. Pattern recognition study of QSAR substituent descriptors. J Comput Aided Mol Des 1989;3(2):111–32.
[19] Ganellin CR. Conformation of histamine derivatives. 3. A relationship between conformation and pharmacological activity. J Med Chem 1973;16(6):620–3.
[20] Paiva TB, Tominaga M, Paiva AC. Ionization of histamine, N-acetylhistamine, and their iodinated derivatives. J Med Chem 1970;13(4):689–92.
[21] Andersson SH, Lindgren A, Postlind H. Biotransformation of tolterodine, a new muscarinic receptor antagonist, in mice, rats, and dogs. Drug Metab Dispos 1998;26(6):528–35.
[22] Herz JE, Swaminathan S, Pinkerton FD, Wilson WK, Schroepfer Jr GJ. Inhibitors of sterol synthesis: a highly efficient and specific side-chain oxidation of 3 beta-acetoxy-5 alpha-cholest-8(14)-en-15-one for construction of metabolites and analogs of the 15-ketosterol. J Lipid Res 1992;33(4):579–98.
[23] Roemer D, Beuscher HH, Hill RC, Pless J, Bauer W, Cardinaux F, et al. A synthetic enkephalin analogue with prolonged parenteral and oral activity. Nature 1977;268:547–9.
[24] Plattner JJ, Norbeck DW. Obstacles to drug development from peptide leads. In: Clark CR, Moos WH, editors. Drug discovery technologies. Chichester: Ellis Horwood Limited; 1990. p. 92–126.
[25] Fauchère JL. Elements for the rational design of peptide drugs. In: Testa B, editor. Adv Drug Res, vol. 1. London: Academic Press; 1986. p. 29–69.
[26] Gee KW, McCauley LD, Lan NC. A putative receptor for neurosteroids on the GABAA receptor complex: the pharmacological properties and therapeutic potential of epalons. Crit Rev Neurobiol 1995;9(2–3):207–27.
[27] Carter RB, Wood PL, Wieland S, Hawkinson JE, Belelli D, Lambert JJ, et al. Characterization of the anticonvulsant properties of ganaxolone (CCD 1042; 3alpha-hydroxy-3beta-methyl-5alpha-pregnan-20-one), a selective, high-affinity, steroid modulator of the gamma-aminobutyric acid(A) receptor. J Pharmacol Exp Ther 1997;280(3):1284–95.
[28] Aquino CJ, Armour DR, Berman JM, Birkemo LS, Carr RA, Croom DK, et al. Discovery of 1,5-benzodiazepines with peripheral cholecystokinin (CCK-A) receptor agonist activity. 1. Optimization of the agonist "trigger". J Med Chem 1996;39(2):562–9.
[29] Tute MS. Principles and practice of Hansch analysis: a guide to structure–activity correlation for the medicinal chemist. In: Harper NJ, Simmonds AB, editors. Advan Drug Res, vol. 6. London: Academic Press; 1971. p. 1–77.

[30] Hansch C, Anderson SM. The structure–activity relationship in barbiturates and its similarity to that in other narcotics. J Med Chem 1967;10:745–53.
[31] Martin YC. Quantitative drug design: a critical introduction. New York: Marcel Dekker, Inc.; 1978.
[32] Dawson MI, Hobbs PD, Derdzinski K, Chan RL, Gruber J, Chao WR, et al. Conformationally restricted retinoids. J Med Chem 1984;27(11):1516–31.
[33] Lindamood C, Cope FO, Dillehay DL, Everson MP, Giles HD, Lamon EW, et al. Pharmacological and toxicological properties of arotinoids SMR-2 and SMR-6 in mice. Fundam Appl Toxicol 1990;14(1):15–29.
[34] Rigas JR, Dragnev KH. Emerging role of rexinoids in non-small cell lung cancer: focus on bexarotene. Oncologist 2005;10(1):22–33.
[35] Lednicer D, Mitscher LA, Georg GI. The organic chemistry of drug synthesis, vol. 4. New York: John Wiley & Sons; 1990. p. 11.
[36] Rizzo RC, Wang D, Tirado-Rives J, Jorgensen WL. Validation of a model for the complex of HIV-1 reverse transcriptase with sustiva through computational resistance profiles. J Am Chem Soc 2000;122(51):12898–900.
[37] Bai DL, Tang XC, He XC. Huperzine A, a potential therapeutic agent for treatment of Alzheimer's disease. Curr Med Chem 2000;7(3):355–74.
[38] Ray AS, Yang Z, Chu CK, Anderson KS. Novel use of a guanosine prodrug approach to convert 2′,3′-didehydro-2′,3′-dideoxyguanosine into a viable antiviral agent. Antimicrob Agents Chemother 2002;46(3):887–91.
[39] Cussac M, Pierre JL, Boucherle A, Favier F. Intérêt des dérivés du cyclopropane en chimie thérapeutique Cyclopropane derivatives, interest in medicinal chemistry. Ann Pharm Fr 1975;33:513–29 [in French].
[40] Marivet MC, Bourguignon JJ, Lugnier C, Mann A, Stoclet JC, Wermuth CG. Inhibition of cyclic adenosine-3′,5′-monophosphate phosphodiesterase from vascular smooth muscle by rolipram analogues. J Med Chem 1989;32:1450–7.
[41] Trivedi BK, Holmes A, Stoeber TL, Blankley CJ, Roark HW, Picard JA, et al. Inhibitors of acyl-CoA:cholesterol acyltransferase. 4. A novel series of urea ACAT inhibitors as potential hypocholesterolemic agents. J Med Chem 1993;36:3300–7.
[42] Sexton WA. Chemical constitution and biological activity. 3rd ed. London: E & F N Spon, Ltd; 1963. p. 103.
[43] Craig PN. Guidelines for drug and analog design. In: 4th ed. Wolf ME, editor. The basis of medicinal chemistry/Burger's medicinal chemistry, vol. I. New York: John Wiley & Sons; 1980. p. 331–48.
[44] Hobbs JB. Purine and pirimidine targets. In: Sammes PG, editor. Comprehensive medicinal chemistry. Oxford: Pergamon Press; 1990. p. 299–331.
[45] Manfredini S, Baraldi PG, Bazzanini R, Marangoni M, Simoni D, Balzarini J, et al. Synthesis and cytotoxic activity of 6-vinyl- and 6-ethynyluridine and 8-vinyl-and 8-ethynyladenosine. J Med Chem 1995;38(1):199–203.
[46] Rancourt J, Cameron DR, Gorys V, Lamarre D, Poirier M, Thibeault D, et al. Peptide-based inhibitors of the hepatitis C virus NS3 protease: structure–activity relationship at the C-terminal position. J Med Chem 2004;47(10):2511–22.
[47] Burger MT, Lin X, Chu DT, Hiebert C, Rico AC, Seid M, et al. Synthesis and antibacterial activity of novel C12 vinyl ketolides. J Med Chem 2006;49(5):1730–43.
[48] Browning E. Toxicity and metabolism of industrial solvents. New York: Elsevier; 1965. pp. 377–381.
[49] Scholtan W, Sy L. Kolloid-chemische Eigenschaften eines neuen Kurznarkotikums: teilgewicht der Mizelle und Verteilungsgleichgewicht des Wirkstoffes [Colloidal chemical properties of a new short-acting anesthetic: particle weight of the micelle and distribution equilibrium of the active compound]. Arzneimit-Forsch 1966;16:679–91. [in German].
[50] Rosen J, Miner JN. The search for safer glucocorticoid receptor ligands. Endocr Rev 2005;26(3):452–64.
[51] Milne Jr GM, Johnson MR. Narcotic antagonists and analgesics. In: Clarke FH, editor. Annual reports in medicinal chemistry, vol. 11. New York: Academic Press; 1967. p. 23–32.
[52] Culvenor CCJ, Dann AT, Dick AT. Alkylation as the mechanism by which the hepatotoxic pyrrolizidine alkaloids act on cell nuclei. Nature 1962;195:570–3.
[53] Sterling J, Herzig Y, Goren T, Finkelstein N, Lerner D, Goldenberg W, et al. Novel dual inhibitors of AChE and MAO derived from hydroxy aminoindan and phenethylamine as potential treatment for Alzheimer's disease. J Med Chem 2002;45(24):5260–79.
[54] Cristofoli WA, Wiebe LI, De Clercq E, Andrei G, Snoeck R, Balzarini J, et al. 5-Alkynyl analogs of arabinouridine and 29-deoxyuridine: cytostatic activity against herpes simplex virus and varicella-zoster thymidine kinase gene-transfected cells. J Med Chem 2007;50(12):2851–7.
[55] Madar DJ, Kopecka H, Pireh D, Yong H, Pei Z, Li X, et al. Discovery of 2-[4-{{2-(2S,5R)-2-cyano-5-ethynyl-1-pyrrolidinyl]-2-oxoethyl]amino]-4-methyl-1-piperidinyl]-4-pyridinecarboxylic acid (ABT-279): a very potent, selective, effective, and well-tolerated inhibitor of dipeptidyl peptidase-IV, useful for the treatment of diabetes. J Med Chem 2006;49(21):6416–20.
[56] Lee SJ, Konishi Y, Yu DT, Miskowski TA, Riviello CM, Macina OT, et al. Discovery of potent cyclic GMP phosphodiesterase inhibitors. 2-Pyridyl- and 2-imidazolylquinazolines possessing cyclic GMP phosphodiesterase and thromboxane synthesis inhibitory activities. J Med Chem 1995;38(18):3547–57.
[57] Beart PM, Uhr ML. Johnston GAR. Inhibition of GABA transaminase activity by 4-aminotetrolic acid. J Neurochem 1972;19:1855–61.
[58] Stutz A, Petranyi G. Synthesis and antifungal activity of (E)-N-(6,6-dimethyl-2-hepten-4-ynyl)-N-methyl-1-naphtha lenemethanamine (SF 86-327) and related allylamine derivatives with enhanced oral activity. J Med Chem 1984;27(12):1539–43.
[59] Nussbaumer P, Leitner I, Mraz K, Stutz A. Synthesis and structure–activity relationships of side-chain-substituted analogs of the allylamine antimycotic terbinafine lacking the central amino function. J Med Chem 1995;38(10):1831–6.
[60] Schmitt J. Sur un nouveau myorelaxant de la classe des benzodiazépines: le tétrazepam [On a new muscle relaxant in the benzodiazepine class: tetrazepam]. Eur J Med Chem 1967;2:254–9. [in French].
[61] Fryer RI, Zhang P, Rios R, Gu Z-Q, Basile AS, Skolnick P. Structure–activity relationship studies at the benzodiazepine receptor (BZR): a comparison of the substituent effects of pyrazoloquinolinone analogs. J Med Chem 1993;36:1669–73.
[62] Bondi A. van der Waals volumes and radii. J Phys Chem 1964;68(3):441–51.
[63] Taylor JB, Kennewell PD. Introductory medicinal chemistry. Chichester: Ellis Horwood, Ltd; 1981. p. 89.
[64] Newman AH, Allen AC, Izenwasser S, Katz J. Novel 3α-(diphenylmethoxy)tropane analogs: potent dopamine uptake inhibitors without cocaine-like behavioral profiles. J Med Chem 1994;37:2258–61.

[65] MacLeod AM, Merchant KJ, Cascieri MA, Sadowski S, Ber E, Swain CJ, et al. N-acyl-l-tryptophan benzyl esters: potent substance P receptor antagonists. J Med Chem 1994;36:2044–5.
[66] Gee KW, Yamamura HI. A novel pyrazoloquinoline that interacts with brain benzodiazepine receptors: characterization of some in vitro and in vivo properties of CGS 9896. Life Sci 1982;30:2245–52.
[67] Bowery NG, Bittiger H, Olpe H-R. $GABA_B$ receptors in mammalian function. Chichester: John Wiley & Sons; 1990.
[68] Katritzky AR, Dobchev DA, Fara DC, Hur E, Tamm K, Kurunczi L, et al. Skin permeation rate as a function of chemical structure. J Med Chem 2006;49(11):3305–14.
[69] Hansch C, Leo A, Hoekman D. Exploring QSAR: hydrophobic, electronic and steric constants. Editor American chemical society, Washington, DC; A.C.S. Professionnal Reference Books; 1995.
[70] Buu-Hoï NP. Les dérivés organiques du fluor d'intérêt pharmacologique [Organic fluorine derivatives of pharmacological interest]. In: Jucker E, editor. Progress in drug research, vol. 3. Basel: Birkhäuser Verlag; 1961. p. 9–74. [in French].
[71] Lide DR. CRC handbook of chemistry and physics. 86th ed. Boca-Raton, FL: CRC Press; 2005.
[72] Razgulin AV, Mecozzi S. Binding properties of aromatic carbon-bound fluorine. J Med Chem 2006;49(26):7902–6.
[73] Chen J, Hwang DJ, Bohl CE, Miller DD, Dalton JT. A selective androgen receptor modulator for hormonal male contraception. J Pharmacol Exp Ther 2005;312(2):546–53.
[74] Wu D, Wu Z, Yang J, Nair VA, Miller DD, Dalton JT. Pharmacokinetics and metabolism of a selective androgen receptor modulator in rats: implication of molecular properties and intensive metabolic profile to investigate ideal pharmacokinetic characteristics of a propanamide in preclinical study. Drug Metab Dispos 2006;34(3):483–94.
[75] Bromfenac Sodium, Duract^R. Drugs Fut 1998;23:1234–1235.
[76] Rosi D, Peruzzotti G, Dennis EW, Berberian DA, Freele H, Archer S. A new, active metabolite of "miracil D". Nature 1965;208:1005–6.
[77] Wermuth CG. Aminopyridazines: an alternative route to potent muscarinic agonists with no cholinergic syndrome. Farmaco 1993;48:253–74.
[78] Cappelli A, Giuliani G, Pericot Mohr GI G, Gallelli A, Anzini M, Vomero S, et al. A non-peptide NK1 receptor agonist showing subpicomolar affinity. J Med Chem 2004;47(6):1315–18.
[79] Ganellin CR, Roberts SM. Medicinal chemistry: the role of organic chemistry in drug research. 2nd ed. London: Academic Press; 1993.
[80] Avendano Lopez MC. Drug optimization by molecular manipulation. Meth Find Exp Clin Pharmacol 2002;24(Suppl. A):9–11.
[81] Klaassen CD. Heavy metals and heavy-metal antagonists. In: Goodman and Gilman's the pharmacological basis of therapeutics, Editor: McGraw-ill; 8th ed. New York: Pergamon Press; 1990. p. 1592–614.
[82] Evans DC, O'Connor D, Lake BG, Evers R, Allen C, Hargreaves R. Eletriptan metabolism by human hepatic CYP450 enzymes and transport by human P-glycoprotein. Drug Metab Dispos 2003;31(7):861–9.
[83] Bromidge SM, Brown AM, Clarke SE, Dogson K, Gager T, Grassam HL, et al. 5-Chloro-N-(4-methoxy-3-piperazin-1-yl-phenyl)-3-methyl-2-benzothiophenesulfonamide (SB-271046): a potent, selective, and orally bioavailable 5-HT_6 receptor antagonist. J Med Chem 1999;42:202–5.
[84] Abreu ME, Rzeszotarski W, Kyle DJ, Hiner R. Preparation of oxopyrazolinyl thiocyanate and bis (oxopyrazolinyl) disulfides as corticotropin-releasing factor antagonist. US Patent 5063245 A. Nova Pharmaceutical Corporation; 1991.
[85] Szajnman SH, Yan W, Bailey BN, Docampo R, Elhalem E, Rodriguez JB. Design and synthesis of aryloxyethyl thiocyanate derivatives as potent inhibitors of Trypanosoma cruzi proliferation. J Med Chem 2000;43:1826–40.
[86] Stahl PH, Wermuth CG. Pharmaceutical salts. Helvetica chimica acta. Zürich: John Wiley & Sons; 2002.
[87] Jirousek L, Pritchard E. On the chemical iodination of tyrosine with protein sulfenyl iodide and sulfenyl periodide derivatives. Biochim Biophys Acta 1971;243:230–8.
[88] Haynes Jr RC. Thyroid and antithyroid drugs. In: Goodman Gilman A, Rall TW, Nies AS, Taylor P, editors. Goodman and Gilman's the pharmacological basis of therapeutics. 8th ed. New York: Pergamon Press; 1990. p. 1361–83.
[89] Montoro J, Sastre J, Bartra J, del Cuvillo A, Davila I, Jauregui I, et al. Effect of H1 antihistamines upon the central nervous system. J Invest Allergol Clin Immunol 2006;16(Sup 1):24–8.
[90] Chen C, Hanson E, Watson JW, Lee JS. P-glycoprotein limits the brain penetration of nonsedating but not sedating H1-antagonists. Drug Metab Dispos 2003;31(3):312–18.
[91] Remy DC, Britcher SF, King SW, Anderson PS, Hunt CA, Randall WC, et al. Synthesis and receptor binding dtudies relevant to the neuroleptic activities of some 1-methyl-4-piperidylene-9-substituted-pyrrolo[2,1-b][3]benzazepine derivatives. J Med Chem 1983;26:974–80.
[92] Bookser BC, Kasibhatla SR, Erion MD. AMP deaminase inhibitors. 4. Further N3-substituted coformycin aglycon analogues: N3-alkylmalonates as ribose 59-monophosphate mimetics. J Med Chem 2000;43(8):1519–24.
[93] Duncia JV, Chiu AT, Carini DJ, Gregory GB, Johnson AL, Price WA, et al. The discovery of potent nonpeptide angiotensin II receptor antagonists: a new class of potent antihypertensives. J Med Chem 1990;33(5):1312–29.
[94] Carini DJ, Duncia JV, Aldrich PE, Chiu AT, Johnson AL, Pierce ME, et al. Nonpeptide angiotensin II receptor antagonists: the discovery of a series of N-(biphenylylmethyl)imidazoles as potent, orally active antihypertensives. J Med Chem 1991;34(8):2525–47.
[95] Owa T, Yoshino H, Okauchi T, Yoshimatsu K, Ozawa Y, Sugi NH, et al. Discovery of novel antitumor sulfonamides targeting G1 phase of the cell cycle. J Med Chem 1999;42(19):3789–99.
[96] Yunger LM, Fowler PJ, Zarevics P, Setler PE. Novel inhibitors of γ-aminobutyric acid (GABA) uptake: anticonvulsant actions in rats and mice. J Pharmacol Exp Ther 1984;228:109–15.
[97] Ali FE, Bondinell WE, Dandridge PA, Frazee JS, Garvey E, Girard GR, et al. Orally active and potent inhibitors of γ-aminobutyric acid uptake. J Med Chem 1985;28:653–60.
[98] Falch E, Krogsgaard-Larsen P. GABA uptake inhibitors: syntheses and structure–activity studies on GABA analogues containing diarylbutenyl and diarylmethoxyalkyl N-substituents. Eur J Med Chem 1991;26:69–78.
[99] Pavia MR, Lobbestael SJ, Nugiel D, Mayhugh DR, Gregor VE, Taylor CP, et al. Structure–activity studies on benzhydrol-containing nipecotic acid and guvacine derivatives as potent, orally-active inhibitors of GABA uptake. J Med Chem 1992;35:4238–48.

[100] N'Goka V, Schlewer G, Linget JM, Chambon JP, Wermuth CG. GABA-uptake inhibitors: construction of a general pharmacophore model and successful prediction of a new representative. J Med Chem 1991;34(8):2547–57.
[101] Andersen KE, Braestrup C, Gronwald FC, Jorgensen AS, Nielsen EB, Sonnewald U, et al. The synthesis of novel GABA uptake inhibitors. 1. Elucidation of the structure–activity studies leading to the choice of (R)-1-[4,4-bis(3-methyl-2-thienyl)-3-butenyl]-3-piperidinecarboxylic acid (tiagabine) as an anticonvulsant drug candidate. J Med Chem 1993;36(12):1716–25.
[102] Wermuth CG, Mann A, Schoenfelder A, Wright RA, Johnson BG, Burnett JP, et al. (2S,4S)-2-amino-4-(4,4-diphenylbut-1-yl)- pentane-1,5-dioic acid: A potent and selective antagonist for metabotropic glutamate receptors negatively linked to adenylate cyclase. J Med Chem 1996;39(4):814–16.
[103] Frolund B, Jorgensen AT, Tagmose L, Stensbol TB, Vestergaard HT, Engblom C, et al. Novel class of potent 4-arylalkyl substituted 3-isoxazolol GABA(A) antagonists: synthesis, pharma-cology, and molecular modelling. J Med Chem 2002;45(12):2454–68.
[104] Elz S, Kramer K, Pertz HH, Detert H, ter Laak AM, Kühne R, et al. Histaprodifens: synthesis, pharmacological in vitro evaluation and molecular modeling of a new class of highly active and selective histamine H_1 receptor agonists. J Med Chem 2000;43:1071–84.
[105] Patrick GL. The opium analgesics. An introduction to medicinal chemistry. 1st ed. Oxford: Oxford University Press; 1995. pp. 250–254.

CHAPTER

14

The Role of Functional Groups in Drug–Receptor Interactions

Laurent Schaeffer

Prestwick Chemical Inc., Boulevard Gonthier, Illkirch, France

OUTLINE

Alice remained looking thoughtfully at the mushroom for a minute, trying to make out which were the two sides of it; and, as it was perfectly round, she found this a very difficult question. ***(Alice's Adventures in Wonderland, Lewis Carroll)***

I. INTRODUCTION

An understanding of noncovalent interactions in ligand–receptor complexes is essential for the appreciation of drug action mechanisms as well as for rational drug design. The purpose of this chapter is to provide an overview of the physical and chemical factors that contribute most significantly to the strength of drug–receptor interactions [1–3]. The first part consists of a physical description of the influence of electrostatic and steric match on the various types of nonbonded drug–receptor interactions. The second part provides a chemical interpretation, concentrating on the intrinsic strengths of individual functional group contributions to the affinity of drugs for their receptors. Some practical applications for the medicinal chemist will then be proposed. The concept of cooperative binding will be exposed in the conclusion in order to underline the limitations of the pure additive calculation method based on average binding energies.

II. GENERAL PRINCIPLES

When a ligand and a receptor are sufficiently close, the ligand can diffuse up to and dock into its binding site on the receptor. This may first be mediated by the long-range electrostatic interactions between the ligand and the receptor, and then strengthened by short-range hydrogen bonds and van der Waals' interactions. Binding is accompanied by conformational changes ranging from modest shifts of a few atoms to movements of whole macromolecule domains. When the ligand is a drug, its biological activity is directly related to its affinity for the receptor, that is, the stability of the drug–receptor complex. The strength of this interaction is measured by its K_d, the dissociation constant for the complex at equilibrium:

$$K_d = \frac{[\text{drug}][\text{receptor}]}{[\text{complex}]} \tag{14.1}$$

The smaller the K_d is the greater is the affinity of the drug for the receptor. This dissociation constant is related to the corresponding Gibbs free energy change, which itself is composed of an enthalpic (ΔH) and entropic contribution ($T\Delta S$) [1,4]:

$$\Delta G = -2.303\ RT \log K_d = \Delta H - T\Delta S \tag{14.2}$$

Under physiological conditions (T = 310 K) this is approximated (in kJ/mol) by

$$\Delta G = -5.85 \log K_d \tag{14.3}$$

The experimental measurement of the equilibrium constant thus provides a direct calculation of ΔG. Typically, K_d is in the range of 10^{-2} and 10^{-12} M, meaning that the affinity of a ligand toward its receptor falls into an energy interval between -10 and -70 kJ/mol in aqueous solution [5]. A drug binding with a K_d of 10^{-9} M requires, for example, $(-5.85) \times (-9) = 52.6$ kJ/mol to dissociate from the receptor.

According to Eq. 14.2, ligand–receptor interactions are characterized by enthalpy–entropy compensation in which one term favors and the other disfavors binding. While enthalpic contributions include electrostatic, hydrogen bond, and van der Waals' interactions, entropic contributions arise from several sources. On the one hand, the loss of flexibility upon binding has an important entropic cost, which is counterbalanced on the other hand by the displacement of ordered water molecules. This will be discussed in the next section along with the various types of drug–receptor interactions.

III. THE IMPORTANCE OF THE ELECTROSTATIC AND STERIC MATCH BETWEEN DRUG AND RECEPTOR

What determines K_d? In other words, how does the binding affinity relate to structural properties of a complex and its formation from separate, individually solvated species? What are the prerequisites that allow a receptor to bind a ligand tightly and selectively? On first glance, the most important requirement appears to be a good steric and electronic complementarity between receptor and ligand, usually described by van der Waals and Coulomb interactions. Although most noncovalent interactions depend to some degree on both types of complementarity, we will separate them in the following discussion into those which are primarily electrostatic and those which are primarily steric.

A. Electrostatic Interactions

Electrostatic interactions are the net result of the attractive forces between the positively charged nuclei and the negatively charged electrons of the two molecules. The attractive force between these opposite charges leads to three main bond types: charge–charge, charge–dipole, and dipole–dipole interactions. The strength of any electrostatic interaction can be calculated with Eq. 14.4, where q_i and q_j are two charges separated by a distance r_{ij} in a medium of dielectric constant ε. This equation applies equally to ionic interactions, where the charges q_i and q_j are integer values, and to polar interactions, in which the total energy is summed over the contributions calculated from the partial charges on all the individual atoms.

$$E = \frac{q_i \cdot q_j}{\varepsilon \cdot r_{ij}} \tag{14.4}$$

TABLE 14.1 Main Ionizable Groups in Proteins and Nucleic Acids [6,7]

Chemical function	Charge	pK_a
FULLY OR ALMOST FULLY IONIZED GROUPS AT pH 7.4		
Carboxyl α (terminal COOH)	−	1.8–2.4
Carboxyl β (Asp)	−	3.7
Carboxyl γ (Glu)	−	4.3
Primary phosphoryl	−	0.7–1.0
Secondary phosphoryl	−	5.9–6.0
α-Ammonium (Lys)	+	10.5
Guanidinium	+	12.5
PARTIALLY IONIZED GROUPS AT pH 7.4		
Sulfhydryl (Cys)	−	8.2
Imidazolium (His)	+	6.0
α-Ammonium (terminal peptide NH_2)	+	7.5–10.3
N Amidic (Glu)	+	0.1
N Amidic (Asp)	+	8.8
NONIONIZED GROUPS AT pH 7.4		
Phenolic hydroxyl (Tyr)	−	10.0
Heteroaromatic hydroxyl (uracyl, thymine, guanine)	−	9.2–9.8
Osidic hydroxyl	−	12.3–12.6
Amino residue (adenine, guanine, cytosine)	+	3.3–4.6

1. *Charge–Charge Interactions or Ionic Bonds*

According to Eq. 14.4, the strength between two charges is inversely proportional to the distance separating them. Since the strengths of other noncovalent bonds are even more sharply dependent on distance than that of ionic bonds, ionic attraction frequently dominates the initial long-range interactions between drugs and receptors. A simple ionic interaction basically provides a ΔG of -20 kJ/mol. It also follows from Eq. 14.4 that the strengths of ionic interactions are crucially dependent on the dielectric constant ε of the surrounding medium. Indeed, in biological systems, charges are often separated by water or other molecules and dielectric micro-environments are variable, with less shielding of charges in regions of hydrocarbon side-chains and greater shielding in regions of polar side-chains. For instance, in hydrophobic pockets, like the interior of a protein molecule, the dielectric constant is around 4, whereas in bulk-phase water the corresponding value is 80. In other environments, intermediate values are appropriate, for example, for interactions occurring near the surface of a protein, an ε value of 28 is commonly used. Charge–charge interactions between biological systems and drugs are possible insofar as ionic species are strongly present in biomacromolecules and drugs at physiological pH. Cationic environments are provided by protonation of basic groups such as the amino acid side-chains lysine, arginine, and—to a much lesser extent—histidine (Table 14.1). On the other hand, acidic groups, such as the carboxylic acid side-chains of glutamic acid and aspartic acid, are deprotonated to give anionic groups.

Concerning drugs, both cationic and anionic compounds are commonly used (protonated basic side-chain, protonated aza heterocycles, deprotonated carboxylic acids, enolic species, and acidic sulfonamides; Figure 14.1).

The antihypertensive drug captopril is an example of a molecule that participates in an ionic bond with the Lys1087 residue of the angiotensin-converting enzyme (ACE) receptor (Figure 14.2).

2. *Charge–Dipole and Dipole–Dipole Interactions*

Molecules composed of atoms of different electronegativities usually have an asymmetric distribution of electrons, which produces electronic dipoles. These dipoles within a cell or in aqueous medium can be attracted by a

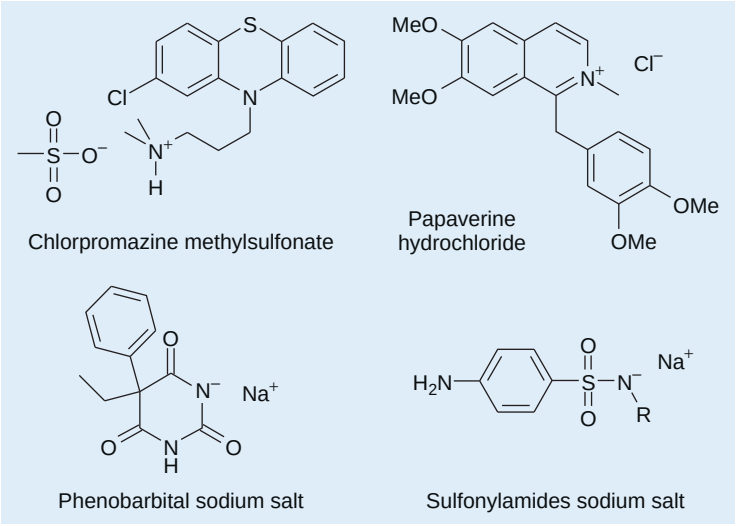


FIGURE 14.1 Cationic and anionic drugs.

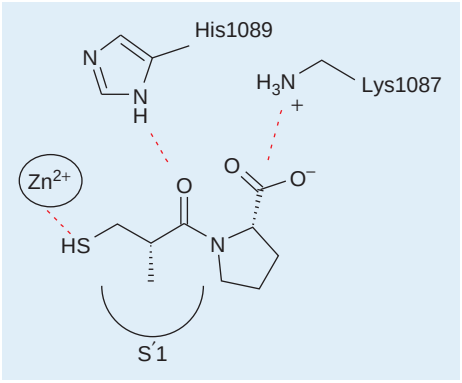


FIGURE 14.2 Captopril as example of a charge–charge interaction with the ACE receptor.

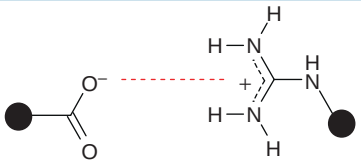
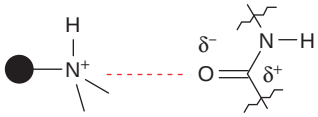
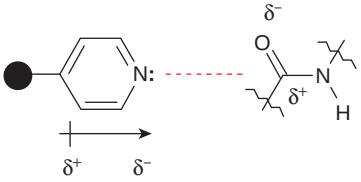
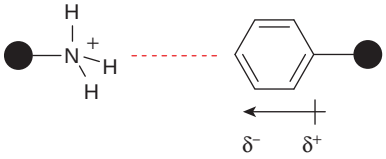
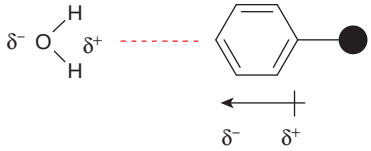
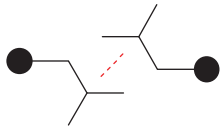
close-by ion, establishing so-called charge–dipole interactions. A permanent dipole can also interact with another permanent dipole, leading to a dipole–dipole interaction. In a recent review, Diederich et al listed these orthogonal multipolar interactions, among which are C–X···C=O (X=halogen), C=O···CO, C≡N···C=O, S=O···C=O, C–OH···C=O, and H_2O···C=O [8]. Because the charge of a dipole is less than that of an ion, charge–dipole and dipole–dipole interactions are weaker than ionic bonds. They are nevertheless key contributors to the overall strengths of drug–receptor interactions, since they occur in any molecule in which electronegativity differences between atoms result in significant bond, group, or molecular dipole moments. The key differences between ionic and dipolar interactions relate to their dependence on distance and orientation (Table 14.2).

Indeed, while steric effects are of little importance in ionic interactions, stricter geometric requirements apply to dipolar interactions, which may be either attractive or repulsive, depending on the orientation of the dipole moments. In most cases, these interactions provide a ΔG of -5 to -30 kJ/mol. The insomnia drug zafirlukast gives an illustration of dipolar interactions (Figure 14.3).

3. *Inductive Interactions*

The electric field generated by a charged molecule or a molecule with a permanent dipole can induce a dipole in a second molecule that is located nearby in space. The strength of the interaction depends on the dipole moment of the first molecule and the polarizability of the second. That means when an electron-donating molecule (or group) comes into contact with an electron-withdrawing molecule (or group), the donor may transfer some of its charge to the acceptor, forming a charge–transfer complex. In the case of an intramolecular

TABLE 14.2 Types of Noncovalent Interactions

Type of interaction	Energy and dependency on distance	Example
Charge–charge: Longest-range force, nondirectional	ΔG of −20 to −40 kJ/mol1/r	
Charge–dipole: Depends on orientation of dipole	ΔG of −12 to −20 kJ/mol1/r^2	
Dipole–dipole: Depends on mutual orientation of dipoles	ΔG of −4 to −12 kJ/mol1/r^3	
Charge-induced dipole: Depends on polarizability of molecule in which dipole is induced	ΔG of −2 to −10 kJ/mol1/r^4	
Dipole-induced dipole: Depends on polarizability of molecule in which dipole is induced	ΔG of −2 kJ/mol1/r^6	
Dispersion: Involves mutual synchronization of fluctuating charges	ΔG of −2 to −4 kJ/mol1/r^6	
Hydrogen bond: Charge attraction 1 partial covalent bond	ΔG of −4 to −30 kJ/mol	

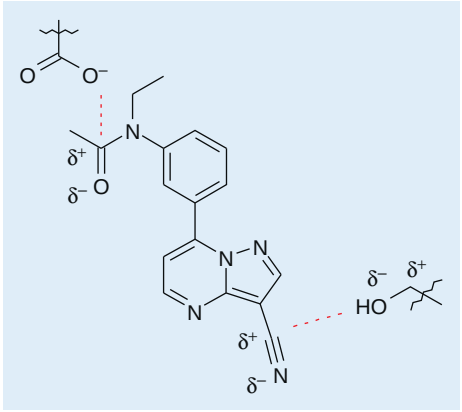


FIGURE 14.3 Zaleplon as example of ion–dipole and dipole–dipole interactions.

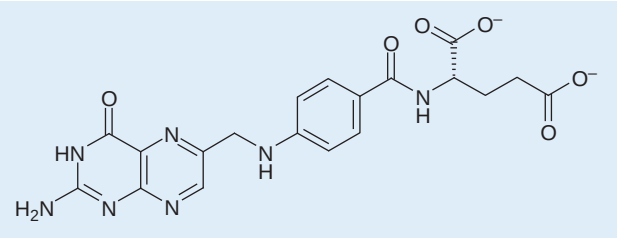


FIGURE 14.4 Folate structure.

TABLE 14.3 Potential Hydrogen Bond Donor and Acceptor Groups Classified According to Their Strength of Interaction [12]

	Donor[a]	Acceptor
Very strong	N^+H_3, X^+–H, F–H	CO_2^-, O^-, N^-, F^-
Strong	O–H, N–H, Hal–H	O = C, O–H, N, S = C, F–H, Hal^-
Weak	C–H, S–H, P–H, M–H	C = C, Hal–C, π, S–H, M, Hal–M, Hal–H, Se

[a]*X is any atom, Hal is any of the lighter halogens, and M is a transition metal.*

redistribution of charge, this will be referred to as an induced polarization, whereas a redistribution of charge between two molecules is described as a charge–transfer interaction. In either case, the resulting interactions are always attractive and strongly dependent on the distance separating the two molecules, as well as on the difference between the ionization potential of the donor and the electron affinity of the acceptor. Donor groups contain π-electrons, such as alkenes, alkynes, and aromatic moieties with electron-donating substituents, or groups presenting nonbonded electrons pairs (O, N, S). Acceptor groups contain electron-deficient π-orbitals, such as alkenes, alkynes, and aromatic moieties having electron-withdrawing substituents, and weakly acidic protons. Groups on receptors acting as electron donors are, for instance, the aromatic ring of tyrosine or the carboxylate group of aspartate. Cystein is an example of an electron-withdrawing group, whereas histidine, tryptophan, and asparagine are both electron donors and acceptors. In general, this type of interaction can contribute as much as 2–3 kJ/mol.

An interesting example of the importance of inductive interactions is the calculation by Bajorath et al on the binding of folate (Figure 14.4) and dihydrofolate to dihydrofolate reductase. This revealed a shift in net charge equivalent to half an electron from the pteridine ring to the glutamate moiety on binding to the enzyme, with the major change in density being focused on the bonds that are catalytically reduced [9].

4. *Hydrogen Bonds*

Hydrogen bonds are specific, short-range, and directional nonbonded interactions. They occur between a hydrogen atom bound covalently to an electronegative atom (usually N, S, or O) and an additional electronegative atom (Table 14.3). Distances of 2.5–3.2 Å between hydrogen-bond donor X and Y and X–H···Y angles of 130–180° are typically found [10]. Their strength is optimal when the three concerned atoms are aligned and when the H donor tends to point directly at the acceptor electron pair. As a result of its electrostatic nature, the strength of a hydrogen bond depends also on its microscopic environment and on the local dielectric constant ε of the surrounding medium (Coulombic interaction energy is proportional to ε^{-1}). Therefore, buried hydrogen bonds are regarded as more important for protein–ligand interactions than those formed in solvent-exposed regions [11]. The free energy for hydrogen bonding can be between −4 and −30 kJ/mol, but usually is in the range of −12 to −20 kJ/mol. Binding affinities increase by about one order of magnitude per hydrogen bond.

Although their strength is weaker than ionic or covalent bonds, they are in general the predominant contribution to the specificity of molecular recognition [13,14]. They help also to determine the conformation and folding ways of numerous macromolecules.

The double helical structure of DNA, for example, is due largely to hydrogen bonding between the base pairs, which link one complementary strand to the other and enable replication (Figure 14.5). Hydrogen bonds are also

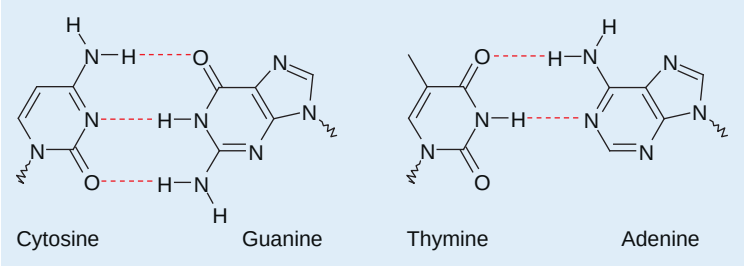


FIGURE 14.5 Hydrogen bonds between DNA base pairs.

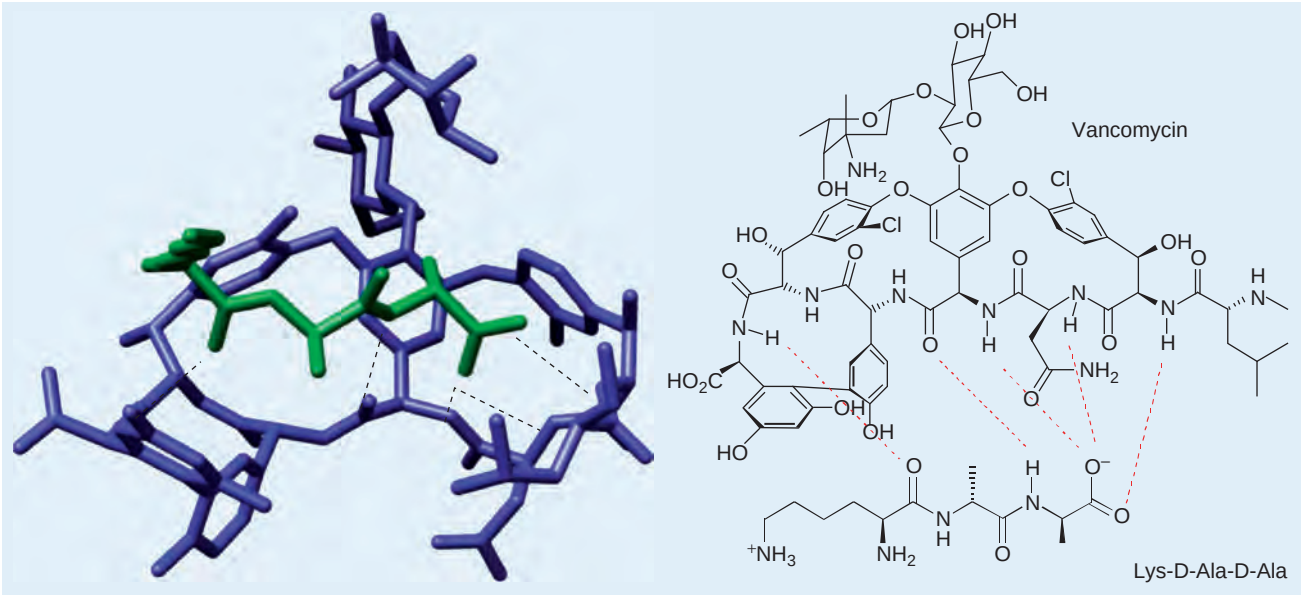


FIGURE 14.6 Crystal structure of a short peptide L-Lys-D-Ala-D-Ala (bacterial cell wall precursor [in green]) bound to the antibiotic vancomycin (in blue) through five hydrogen bonds [15].

essential to maintaining the structural integrity of α-helix and β-sheet conformations of peptides and proteins. By causing the macromolecules to fold into a specific shape, the hydrogen bonds contribute to the apparition of their biochemical functions.

In drug design, hydrogen bonds are exploited to obtain specificity, which is achieved through favorable, short-range, directionally specific interactions and the fact that ligand-receptor arrangements that leave bonding capacity unsatisfied are disfavored [15a]. The number of hydrogen bonds in a drug molecule may be limited by requirements on polarity for absorption and permeation. The Lipinski rule-of-five, for example, suggests that compounds with more than five hydrogen-bond donors or more than ten hydrogen bonds acceptors are likely to have poor absorption or permeation characteristics.

Among the numerous examples of drug−receptor interactions through hydrogen bonds, the antibiotic vancomycin is especially interesting because it binds selectively with peptides having a terminal D-Ala-D-Ala moiety in a bacterial cell through five hydrogen bonds (Figure 14.6). Vancomycin is lethal to the bacteria, since once it has bound to these particular peptides they are unable to be used to construct the bacteria's cell wall.

5. *Cation−π Interactions*

Such interaction occurs between a cation and the large, permanent quadrupole moment of an aromatic ring. The interaction energy depends on both the nature of the π-system and the nature of the cation. The importance of cation-π interactions were first recognized by Ma and Dougherty [16]. In fact, cation-π interactions play a key role in molecular recognition in biological receptors. They have been considered in such diverse systems as acetylcholine receptors (nicotinic, muscarinic, and Ach esterase), K^+ channels, the cyclize enzymes of steroid biosynthesis, and enzymes that catalyze methylation reactions involving S-adenosylmethionine [17]. A remarkable

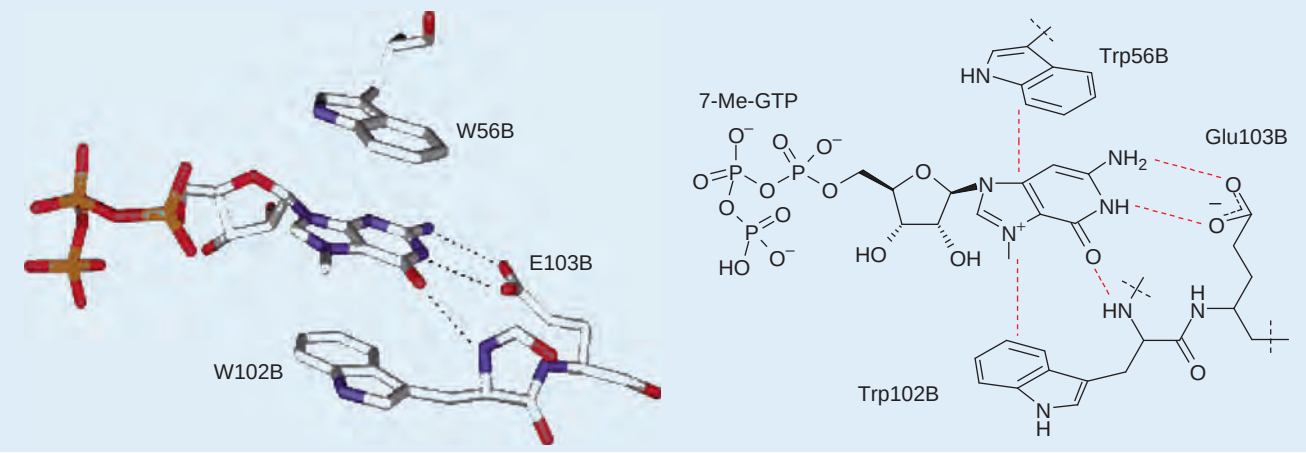


FIGURE 14.7 Partial view of the X-ray crystal structure (PDB code: 1L8B) of the messenger RNA 5′-cap-binding protein eIF4E bound to 7-methyl-GTP, which shows the sandwiching of the cationic nucleobase between the side chains of Trp102B and Trp56B.

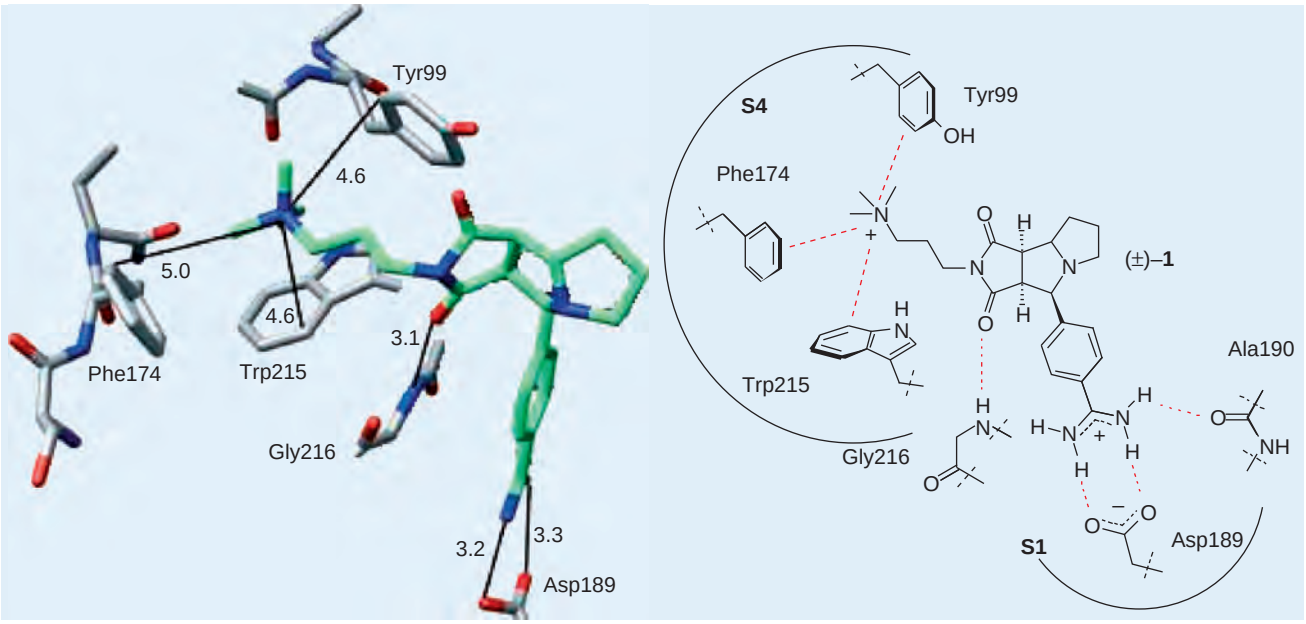


FIGURE 14.8 Schematic representation of the complex between factor Xa and the tricyclic inhibitor (±) − 1.

example is given by the strong complexation of 7-methyl-GTP (dissociation constant $K_d \approx 1.1 \times 10^8\ M^{-1}$, $\Delta G = -45\ kJ/mol)^{-1}$ by a messenger RNA 5′-cap-binding protein, the eukaryotic translation initiation factor eIF4E. The cationic nucleobase in this complex is sandwiched at van der Waals distance (c. 3.5 Å) between two tryptophan side-chains (Figure 14.7) [18].

The study of the factor Xa, a serine protease from the blood coagulation cascade, has led Diederich et al. [19] to observe that the aromatic box formed by the side chains of Phe174, Tyr99, and Trp215 in the S4-pocket is a very effective onium binding site. By comparing the affinity of the quaternary ammonium ion (±)-**1** to its tert-butyl analog, the free enthalpy increment for cation-π interactions in this box was determined as $\Delta G = -2.8 \times 4.18\ kJ/mol$ (Figure 14.8).

6. *Arene–Arene Interactions [17,20]*

Despite their weak and poorly directional character, π−π interactions have been recognized to play an important role in molecular recognition. Burley and Petsko demonstrated in a study involving thirty-four proteins that,

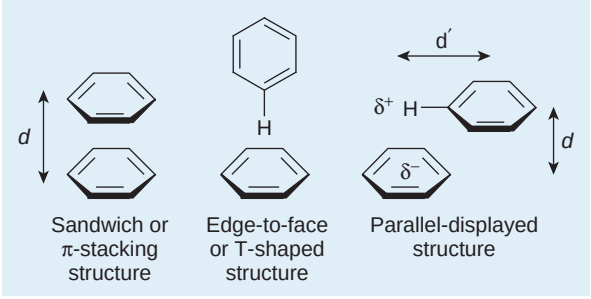


FIGURE 14.9 Proposed lowest energy structures of the benzene dimer. *d*: distance between planes, d′: lateral offset.

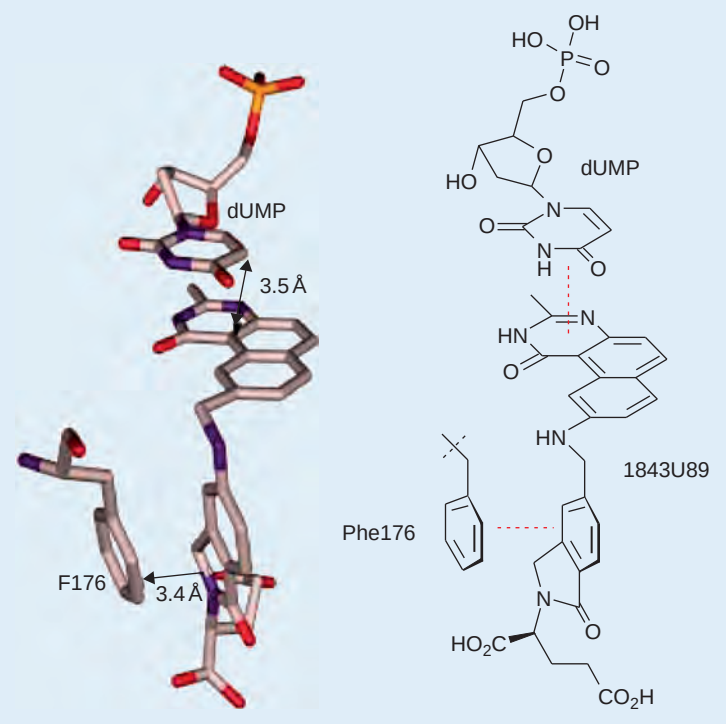


FIGURE 14.10 Heterocyclic π stacking between dUMP and the anticancer drug 1843U89 bound at the active site of thymidylate synthase (PDB code: 1TSD).

on average, 60 percent of aromatic side-chains (Phe, Trp, Tyr) are involved in aromatic–aromatic interactions [21]. It is now recognized that London dispersion interactions are the major source of stabilization energy between two aromatic molecules; however, the electrostatic component associated with the quadrupole moment of the aromatic ring is an influential factor in determining the geometry of the interaction. Assessment of binding interactions in aqueous solution is complicated by additional hydrophobic effects leading to apolar complexation or to intramolecular hydrophobic collapse.

Three lowest energy arrangements are commonly involved in π–π interactions (Figure 14.9), among which the T-shaped structure is the predominant one. The highest edge-to-face attraction is observed when an electron-attracting substituent renders the interacting H atom more acidic (higher positive partial charge) and when an electron-donating substituent increases the basicity (π-electron density) of the interacting π system. The other important arrangement of aromatic rings besides the edge-to-face contact is the parallel alignment. In this case two aromatic partners—one bearing strong electron-donor and the other strong electron-acceptor groups—form parallel stacking complexes in solution, with the geometry largely determined by molecular orbital interactions (charge–transfer complexes). The term "polar/π" interaction was introduced to emphasize the importance of the electrostatic term in π-stacking. Since attractive electrostatic interactions between atoms of opposite partial charges often overcome the repulsion between close-shell π clouds, π–π stacking interactions are also abundant between heterocyclic π systems. A nice example is provided by the ternary complex of the anticancer drug 1843U89 and dUMP formed at the active site of thymidylate synthase (Figure 14.10) [22].

B. Steric Interactions

As the lock-and-key model suggests, shape complementarity is very important for ligand–receptor binding and specificity. While electrostatic interactions are the dominant interactions involving polar molecules, there are also strong interactions between nonpolar molecules, particularly at short intermolecular distances.

1. *Dispersion Forces*

Van der Waals or London dispersion forces are the universal forces responsible for attractive interactions between nonpolar molecules. The occurrence of these short-range interactions is due to the fact that any atom will, at any given instant, be likely to possess a finite dipole moment as a result of the movement of electrons around the nuclei. When molecules are approaching each other, the temporary dipoles of one molecule induce opposite dipoles in the other approaching molecules, thus resulting in a net attractive force. Although the individual interactions between pairs of atoms are relatively weak (about 2 kJ/mol), the total contribution to binding from dispersion forces can be very significant if there is a close fit between drug and receptor. The quality of the steric match is thus the dominant factor in nonpolar interactions.

2. *Short-Range Repulsive Forces*

The short-range repulsive forces resulting from the overlap of the electron clouds of any two molecules increase exponentially with decreasing internuclear separation. The balance between these repulsive interactions and the dispersion forces thus determines both the minimum and the most favorable nonbonded separation between any pair of atoms. The equilibrium distance can be determined from crystal data, and is equivalent to the sum of the van der Waals radii of the two interacting atoms. For nonpolar molecules, this balance between the attractive dispersion forces and the short-range repulsive forces is generally defined in terms of the Buckingham (6-exp) potential given in Eq. 14.5 or the alternative Lennard–Jones 6–12 potential given in Eq. 14.6.

$$E = \frac{Ae^{-Br}}{r^d} - \frac{C}{r^6} \tag{14.5}$$

$$E = \frac{A^r}{r^{12}} - \frac{C}{r^6} \tag{14.6}$$

3. *Conformational Energy*

While intramolecular interactions within the drug molecule are the primary factor in determining the lowest energy conformation of the unbound drug, intermolecular interactions with the receptor also have a significant effect on conformation. If the bound conformation of a flexible molecule is also its lowest energy conformation, there is no conformational energy cost involved in binding. If, on the other hand, the optimal interaction between drug and receptor requires a higher-energy conformation, this energy difference will reduce the apparent strength of the interaction between the two molecules.

C. Enthalpy–Entropy Compensation

Consider the formation of a specific noncovalent bond (e.g., $\mathbf{A}\cdots\mathbf{B}$ for the transformation $\mathbf{A}+\mathbf{B}\rightarrow\mathbf{A}\cdots\mathbf{B}$). An increase in its strength (which corresponds to an increasing negative contribution to ΔH and a more favorable binding process) will be accompanied by an increasing restriction in the relative motion of **A** and **B** in $\mathbf{A}\cdots\mathbf{B}$ (which corresponds to a negative contribution to ΔS, and so is unfavorable to binding). This opposing interplay between enthalpy and entropy is known as "enthalpy–entropy compensation" and is a fundamental property of noncovalent interactions [23]. An enhancement of intermolecular binding is accompanied by a loss in degrees of freedom of mobility and vice versa. The two effects can be traded off against each other because the strength of noncovalent bonds at room temperature is comparable to the thermal energies that oppose them. The enthalpy–entropy compensation is of particular importance for the prediction of receptor–ligand interactions. Whereas the individual enthalpic and entropic contributions can vary over large ranges, the total change in free enthalpy is frequently close to zero. As a consequence, small relative errors in the prediction of ΔH and ΔS can have significant influence on ΔG. This concept is less important in the study of covalent bonds, which are typically too strong to be effectively opposed by thermal motions at room temperature.

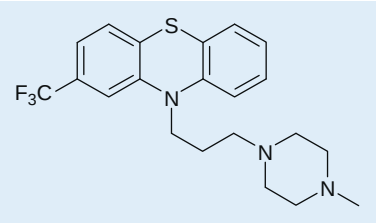


FIGURE 14.11 Trifluoperazine, a ligand of Ca^{2+}-calmodulin that induces an adaptation of the protein binding pocket.

1. *Hydrophobic Interactions*

The hydrophobic effect is that nonpolar molecules tend to self-associate in the presence of aqueous solution. This short-range attractive interaction is due to both enthalpic and entropic effects. It describes the energetic preference of nonpolar molecular surfaces to interact with other nonpolar molecular surfaces, and thereby to displace water molecules from the interacting surfaces. When a nonpolar molecule is surrounded by water, stronger than normal water−water interactions are formed around the solute molecule to compensate for the weaker interactions between solute and water [24]. This results in an increasingly ordered arrangement of water molecules around the solute and thus a negative entropy of dissolution. The decrease in entropy is roughly proportional to the nonpolar surface area of the molecule. The association of two such nonpolar molecules in water reduces the total nonpolar surface area exposed to the solvent, thus reducing the amount of structured water, and therefore providing a favorable entropy of association. The enthalpic contribution to hydrophobic interactions is due to the water molecules occupying lipophilic binding sites, which are consequently unable to form hydrogen bonds with the receptor. Their release from the hydrophobic pocket lets them form strong hydrogen bonds with the bulk water. As for van der Waals forces, hydrophobic interactions are individually weak (0.1 to 0.2 kJ/mol for every square angstrom of solvent-accessible hydrocarbon surface) [25], but the total contribution of hydrophobic bonds to drug−receptor interactions is substantial. Similarly, the overall strength of the hydrophobic interaction between two molecules is highly dependent on the quality of the steric match between the two molecules. If this is not sufficiently close to squeeze all of the solvent from the interface, a substantial entropy penalty must be paid for each of the trapped water molecules. Hydrophobic interactions are also regarded as the main driving force for conformational changes of the receptor upon ligand binding. This induced fit can be viewed as a "collapse" of the receptor about the ligand [26]. As an extreme case, the binding of trifluoroperazine (Figure 14.11) to Ca^{2+} calmodulin induces a conformational change of the protein from an extended to a compact form [27].

2. *Translational and Rotational Entropy*

The transformation of two mobile molecules into one mobile complex results in the loss of translational and rotational entropy. Indeed, by binding to its receptor, a drug molecule is losing three translational and three rotational degrees of freedom, which are replaced by six vibrational degrees of freedom in the complex [28]. The resulting entropy change is dependent on the relative tightness of the complex that is formed. For a typical ligand−protein interaction, the estimated change in free energy resulting from the loss of entropy on binding (at 310 K) ranges from 12 kJ/mol for a very weak interaction to 60 kJ/mol for a tightly bound complex [29].

3. *Conformational Entropy*

In the case of flexible drug molecules, a further entropy loss is due to the conformational restriction that accompanies binding. Based on the observed entropy changes accompanying cyclization reactions, the extent of this entropy loss is estimated [30] at 5−6 kJ/mol per internal rotation, although the actual figure again depends on the overall strength of the interaction between the drug and the receptor. In the case of rigid analogs, no such loss of conformational entropy occurs on binding. Provided that they offer a good steric and electrostatic match to the receptor, rigid analogs should therefore have a free energy advantage relative to more flexible drugs. To optimize entropic contributions, compounds are usually designed to be relatively rigid with few rotatable bonds. Conformational flexibility is, however, important in biomolecular binding processes. A recent conformational analysis of drug-like ligands binding to proteins shows that many ligands do not bind in a minimum energy conformation. Energetically unfavorable conformational rearrangements can be tolerated in some cases without penalizing the tightness of binding [31]. On the other hand, small-scale motions including bond stretching, bond angle bending, and dihedral angle variations are able to reduce slightly the receptor affinity for its ligand. The

timescale of these motions is around 10^{-12} seconds and the amplitude is less than 1 Å. For enzyme catalysis, for example, movement of less than 1 Å can alter catalytic rates by several orders of magnitude.

IV. THE STRENGTHS OF FUNCTIONAL GROUP CONTRIBUTIONS TO DRUG–RECEPTOR INTERACTIONS

The total free energy of interaction between a drug and its receptor provides a measure of the strength of the association between the two molecules but tells us little or nothing about the overall quality of their match. Does the observed binding reflect a composite of interactions between every part of the drug and its receptor, or is it a case of one or two strong interactions contributing sufficient energy to disguise an otherwise mediocre fit? Is the observed increase in interaction energy resulting from the addition of a new functional group consistent with what might have been anticipated? To answer these questions, we need some means of estimating the individual functional group contributions to drug–receptor interactions.

A. Measuring Functional Group Contributions

When cooperativity is ignored, contributions of ΔG values to the total free energies of binding may be added together. Approaches based on functional group additives (Eq. 14.7) or the additivity of free enthalpy components (Eq. 14.8) have frequently been applied to understand and predict protein–ligand interactions [32]. Pioneering studies in this field were performed by Andrews et al. [31] and Lau and Pettitt [33].

$$\Delta G = \Delta G_{\mathrm{Me}} + \Delta G_{\mathrm{OH}} + \Delta G_{\mathrm{Ph}}(\mathrm{Ph}) + \cdots \tag{14.7}$$

$$\Delta G = \Delta G_{H\text{-bridge}} + \Delta G_{\mathrm{solvation}} + \Delta G_{\mathrm{conformation}} + \cdots T\Delta S \tag{14.8}$$

In order to have a brief overview of the methods used to predict the free energy of binding of a ligand to its receptor, we will describe and discuss some of them here.

As a first approximation, the free energy of binding can be defined in terms of the binding energies for the individual functional groups that make up the drug molecule according to Eq. 14.9.

$$\Delta G = T\Delta S_{t,r} + n_r E_r = \Sigma n_{\mathrm{x}} E_{\mathrm{x}} \tag{14.9}$$

where $T\Delta S_{t,r}$ is the loss of overall translational and rotational entropy associated with binding of the drug molecule, n_r is the number of internal degrees of conformational freedom lost on binding the drug molecule, and E_r is the energy equivalent of the entropy loss associated with the loss of each degree of conformational freedom on receptor binding.

1. *Intrinsic Binding Energy*

The final term in Eq. 14.9 is the sum of the binding energies E_{x} associated with each functional group X, of which there are n_{x} present in the drug. In the ideal case, when the specified functional group is aligned optimally and without strain with the corresponding functional group in the receptor, E_{x} is called the "intrinsic binding energy" [34]. In other cases, the term "apparent binding energy" is used.

It should be noted that each binding energy E_{x} is actually a combination of the various enthalpic and entropic interactions outlined above. These include the enthalpy of interaction between the functional group and its corresponding binding site on the receptor, the enthalpy changes associated with the removal of water of hydration from the functional group and its target site and the subsequent formation of bonds between the displaced water molecules, and the corresponding entropy terms associated with the displacement and subsequent bonding of water molecules (Figure 14.12).

It is apparent that these intrinsic binding energies may be regarded, at least approximately, as properties of the functional group that should be relatively independent of the groups to which the particular functional group is attached. Such intrinsic binding potentials might thus reasonably be used in an additive manner to provide an overall estimate of the drug–receptor interaction.

2. *Anchor Principle*

It follows from Eq. 14.9 that the binding energy E_{x}, due to the interaction between the receptor and a specific functional group X, can be estimated by comparing the binding energies for pairs of compounds that differ only

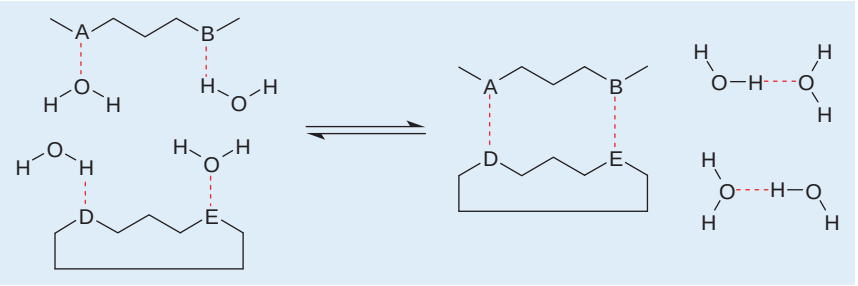


FIGURE 14.12 Complex formation between a ligand (containing polar functionalities **A** and **B**) and a receptor (containing polar functionalities **D** and **E**) with exchange of four water molecules to the bulk solvent.

in the presence or absence of the specific functional group. This approach was first applied by Page [29] who referred to it as the "anchor principle." It is based on the premise that the difference in binding of a drug molecule with or without the particular functional group is due to factors associated solely with that group, that is, the binding energy E_x plus any degrees of conformational freedom lost specifically as a result of binding of group X. Other degrees of conformational freedom lost on binding and the loss of overall rotational and translational entropy associated with the remainder of the drug molecule (the anchor) are assumed to be unaffected by the presence or absence of X.

Similarly, the impact of a single amino acid substitution in the active site of an enzyme on transition-state stabilization, as determined by the change in either catalytic efficiency or inhibitor binding, provides a measure of the relative binding energy of the two side-chains.

Clearly, the magnitude of the binding energies obtained using the anchor principle will vary widely with the quality of the interaction. If the functional groups are not properly aligned, as might reasonably be expected in many mutant proteins, a small or even repulsive interaction may result. Alternatively, the strength of the additional bond may be offset by a reduction in the strengths of the existing bonds. Under these circumstances, the anchor principle will lead to an underestimation of the true bond strength.

3. *Average Binding Energy*

An alternative to the pair-by-pair approach inherent in the anchor principle was developed by Andrews et al. [31], who sought to average the contributions of individual functional groups to the observed binding energies of 200 ligand–protein interactions in aqueous solution. For this purpose, the average loss of overall rotational and translational entropy accompanying drug–receptor binding, $T\Delta S_{t,r}$ in Eq. 14.9, was estimated at 58.5 kJ/mol at 310 K. Regression analysis against n_r (obtained by counting the number of degrees of conformational freedom in each of the 200 ligand structures) and n_x (the number of occurrences of each functional group X in each of the 200 ligand structures) as the independent variables was then used to obtain average values of the binding energies associated with each functional group and for the loss of entropy associated with each degree of conformational freedom.

The results of this analysis showed that the loss of entropy associated with each internal rotation ΔG_r on receptor binding is equivalent to a reduction in the free energy of binding by average of 3 kJ/mol.

The corresponding binding energies obtained by the averaging process were: C (sp^2 or sp^3), 3 kJ/mol; O, S, N, or halogen, 5 kJ/mol; OH and C=O, 10 and 14 kJ/mol, respectively; and CO_2^-, OPO_3^{2-}, and N^+, 34, 42, and 48 kJ/mol, respectively. Once again, it should be stressed that these are not intrinsic binding energies in the sense defined above. This would be the case only if each functional group in each drug in the series was optimally aligned with a corresponding functional group in the receptor. In fact, since every functional group of every drug was included in the analysis, the calculated values are averages of apparent binding energies, including those for some groups that may not interact with the receptor at all [35]. The calculated averages are thus almost certainly smaller than the corresponding intrinsic binding energies, although they follow expected trends in that charged groups lead to stronger interactions than polar groups, which in turn are stronger than nonpolar groups such as sp^2 or sp^3 carbons.

The apparent contributions of some functional groups and/or bond types to overall binding energies derived from the various studies reviewed above are summarized in Table 14.4. Also included are corresponding values used or suggested for the overall loss of rotational and translational entropy, $T\Delta S_{t,r}$, and the loss of conformational entropy resulting from restriction of free rotation, ΔG_r. These first attempts at a semiquantification of drug–receptor interactions in terms of costs and benefits were later refined by Williams et al. [36,37], who applied an additional term to account for the hydrophobic effect. They suggested that the magnitude of this effect

TABLE 14.4 Functional Group Contributions to Drug–Receptor Interactions (kJ/mol)

	Technique employed to determine interaction energy		
Functional group type	**Anchor principle**	**Site-directed mutagenesis**	**Average energy**
Nonpolar (per carbon atom)	12–14	1–3	3–6
H–bonding (uncharged)	16	2–6	5–14
H–bonding (charge-assisted)	20–42	15–19	
Charged (carboxyl, amine)	18–28	12–25	34–48
$T\Delta S_{t,r}$	12–60		58.5
ΔG_r (internal rotation)	5–6		3

TABLE 14.5 Average Values for the Parameters of Eq. (14.10) [25]

Parameter	**Physical process**	**Value (kJ/mol)**
$\Delta G_{t,r}$	Energy cost of bimolecular association	+5.4
ΔG_r	Energy cost of restriction of an internal rotor	+1.4
ΔG_h	Benefit of the hydrophobic effect (per $Å^2$ of buried hydrocarbon)	−0.17 ($Å^2$)
ΔG_p	Benefit of making a neutral hydrogen bond of ideal geometry	−4.7
ΔG_{ionic}	Benefit of making an ionic hydrogen bond of ideal geometry	−8.3

is proportional to the surface area of the hydrocarbon that is removed from exposure to water upon formation of the complex. It can be estimated in terms of ΔG_h per unit of area A of hydrocarbon buried, which can readily be measured with the aid of computer graphics. It leads to the following equation:

$$\Delta G = \Delta G_{t,r} + n\Delta G_r + A\Delta G_h + \Sigma\Delta G_p \tag{14.10}$$

ΔG is the observed free energy of binding, $\sum\Delta G_p$ is the sum of the free energies of binding for all the polar interactions made in the binding site, and the other terms are as defined above. Böhm "trained" a variant of Eq. 14.10 with a set of forty-five interactions of experimentally known binding constants from the association of ligands of small molecular weight with proteins through sets of known interactions [38]. He divided the original ΔG_p values into two groups: those involving ionic interactions (ΔG_{ionic}) and those involving hydrogen bonds formed between neutral entities (the term ΔG_p was retained). Since the modified form of the equation has only five types of ΔG contributions and the forty-five binding sites involve different combinations of these five types of ΔG contributions, average values for them can be obtained. These values (Table 14.5) have proven very useful in the pharmaceutical industry.

It is interesting to note that the average value $\Delta G_{t,r} = +5.4$ kJ/mol is remarkably small and represents only about one tenth of the maximum theoretical entropy loss corresponding to complete immobilization of the ligand. This small value presumably reflects, at least in part, the large residual motion that the drugs can exercise relative to the receptor to which they are bound. The average cost of restricting the rotation of an internal bond in the drugs (+1.4 kJ/mol) is slightly less than that found for the formation of crystals from neat liquids that contain internal rotors (2–3 kJ/mol). This finding probably reflects the fact that rotations are somewhat less restricted in these binding sites than they are in crystals. Most importantly, the application of the equation gives useful approximate binding constants in many cases.

B. The Methyl Group and Other Nonpolar Substituents

The initial application of the anchor principle described by Page [29] related to data on the selectivity of amino acid-tRNA synthetases, from which he estimated intrinsic binding energies for the methylene group in the range 12–14 kJ/mol. For example, the calculated binding energies of Eq. 14.3 for isoleucine **2** (Figure 14.13) and its

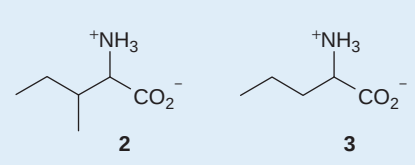


FIGURE 14.13 Isoleucine **2** and desmethyl-isoleucine **3**.

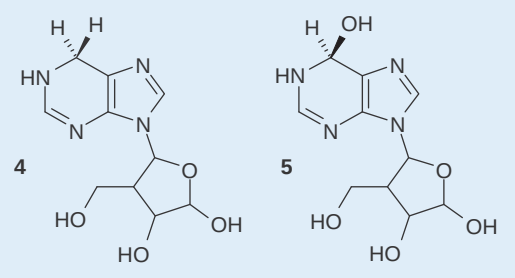


FIGURE 14.14 1,6-Dihydropurine ribonucleoside **4** and its 6-hydroxy analog **5**.

desmethyl analog **3** to isoleucyl-tRNA synthetase are 29.7 and 15.9 kJ/mol, respectively, indicating that the methyl group contributes a total of 13.8 kJ/mol to the overall interaction.

This estimate, having been derived from observations on a highly selective enzyme–substrate interaction, is probably also approaching the intrinsic limit for the binding contribution of a methyl group. Indeed, according to the calculation of Williams [36], the burial of a methyl group ($\approx$25 Å^2) contributes only -4 kJ/mol. For longer hydrocarbon side-chains the positive contribution due to dispersion forces and hydrophobic interactions tends to be offset by the loss of conformational entropy on binding. Thus, the "average" binding energy of 3 kJ/mol obtained by Andrews et al for sp^2 and sp^3 carbon groups is identical to the "average" reduction in free energy of binding estimated for the loss of conformational freedom around a single bond [31]. Clearly, this effect will be greater in saturated hydrocarbon chains than in their more conformationally constrained unsaturated or cyclic analogs.

C. The Hydroxyl Group and Other Hydrogen-bond Forming Substituents

The most extensive studies of hydroxyl group contributions to drug–receptor interactions are those of Wolfenden et al. on the contribution of hydrogen bonds formed by hydroxyl groups in transition-state analogs. In a series of thirteen examples of paired ligands with and without hydroxyl groups, they used [39] the anchor principle to determine apparent binding energies for single hydroxyl groups ranging from 20–42 kJ/mol.

Thus, in comparing the binding of 1,6-dihydropurine ribonucleoside **4** (Figure 14.14) and its 6-hydroxy derivative **5** to adenosine deaminase, they observed a difference in binding energy of 41 kJ/mol [40]. The authors suggested that the 6-hydroxyl group, which has very limited freedom of movement, is likely to be in almost ideal alignment with the active site for forming a hydrogen bond. This conjecture has been verified by the determination of the crystal structure of the inhibitory complex between adenosine deaminase and 6-hydroxy-1,6-dihydropurine ribonucleoside, which showed that the 6-hydroxyl group interacts with a zinc atom, with a protonated histidyl residue, and with an aspartic acid residue at the enzyme's active site [41].

Once again, the data from active-site mutagenesis studies are less striking, but nevertheless reveal some very substantial hydrogen-bonding interactions. In Fersht's studies [42] on tyrosyl-tRNA synthetase, for example, hydrogen bonds between this enzyme and uncharged substrate groups contributed between 2 and 6 kJ/mol toward specificity, while hydrogen bonds to charged groups contributed between 15 and 19 kJ/mol, corresponding to a factor of 1000 in specificity. These numbers are, however, higher than the average contributions determined by Böhm [38] (4.7 and 8.3 kJ/mol, respectively).

D. Acidic and Basic Substituents

Simple observations on the interactions of individual charged groups with appropriate enzymes may lead to an indication of their binding energies. The phosphate ion, for example, binds alkaline phosphatase [43] with a dissociation constant of 2.3×10^{-6} M, equivalent to a ΔG value of approximately 33 kJ/mol. Taking the most conservative estimate for the loss of rotational and translational entropy associated with this interaction (12 kJ/mol for a loosely bound complex), Eq. 14.9 then gives a lower estimate for binding of the phosphate ion of 45 kJ/mol. If the same value of $T\Delta S_{t,r}$ is applied to the binding of oxalate ion to transcarboxylase [44], for which the dissociation constant is 1.8×10^{-10} M (57 kJ/mol), Eq. 14.9 gives an apparent binding energy of 24 kJ/mol per carboxylate group after allowance for a minimal conformational entropy loss of 3 kJ/mol.

These figures are broadly consistent with the average values of Andrews et al. [31], which were in the range 34–48 kJ/mol for charged phosphate [45], amine, and carboxyl groups.

E. Practical Applications for the Medicinal Chemist

1. Assessing a Lead Compound

Summation of the average contributions of individual binding groups, including allowance for conformational, rotational, and translational entropy terms as shown in Eq. 14.9, provides a simple back-of-the-envelope calculation of the strength of binding that might be expected for a drug forming a typical interaction with a receptor. This figure, when compared to the observed affinity of the drug for the target receptor, then gives a direct indication of the actual quality of the electrostatic and steric match between the drug and the receptor.

A. BINDING IS TIGHTER THAN EXPECTED

If the observed binding of a drug to its receptor turns out to be substantially stronger than that calculated from Eq. 14.9, it is reasonable to expect that the drug structure offers a good fit to the receptor in a reasonably low-energy conformation. The structure should therefore provide an excellent starting point for the development of even more bioactive compounds.

A good example of this is biotin (Figure 14.15), which was the most extreme case of a positive deviation from the calculated "average" in the original set of 200 ligand-protein interactions studied by Andrews et al [31].

$$\begin{aligned}\Delta G_{av} &= T\Delta S_{rt} + 5E_r + 8E_{\mathrm{Csp^3}} + 2E_{\mathrm{N}} \\ &\quad + E_2 + E_{\mathrm{C{=}O}} + E_{\mathrm{COOH}} \\ &= -58.5 + 5(-3) + 8(3) + 2(5) \\ &\quad + 5 + 14 + 34 = 13.5\ \mathrm{KJ/mol}\end{aligned} \tag{14.11}$$

$$\begin{aligned}\Delta G_{\mathrm{obs}} &= -5.85 \log K_d \\ &= -5.85(-15) = 87.7\ \mathrm{KJ/mol}\end{aligned} \tag{14.12}$$

Application of Eq. 14.9 to biotin (see above) gives an average binding energy of 13.5 kJ/mol, whereas substitution into Eq. 14.3 of the experimentally observed binding constant to the protein avidin (10^{-15} mol^{-1}) gives a binding energy of 87.7 kJ/mol. The difference of almost 74 kJ/mol implies an exceptionally good fit between biotin and the structure of the protein. It has since been established that this is indeed the case, with polarization of the biotin molecule by the protein actually leading to an ionic interaction where a neutral hydrogen-bonding interaction had been assumed.

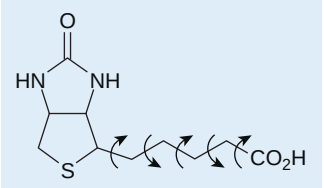


FIGURE 14.15 Structure of biotin.

B. BINDING IS LOOSER THAN EXPECTED

If the observed binding is significantly weaker than anticipated on the basis of an "average" energy calculation, the fit between the drug and the receptor is less than perfect. In some cases, this will be because the match between drug and receptor is less a matter of "hand and glove" than of "square peg and round hole," and the only realistic option for the drug designer is to start again.

In other cases, simpler remedies may be followed:

1. The fit may be unsatisfactory because only part of the drug is interacting with the receptor. This situation applies particularly to large drug molecules (e.g., peptide hormones), for which selective pruning of unused parts of the structure may produce simpler compounds without loss of affinity;
2. The drug may be binding to the receptor in a comparatively high energy conformation. In this case, the design of more rigid structures that are already fixed in the desired conformation will give an increase in binding energy equivalent to the conformational energy cost of binding the more flexible analog.

2. *Assessing the Effectiveness of Substituents*

Equally simple back-of-the-envelope calculations based on Eq. 14.9 can be used to predict the increase in binding energy that might be expected upon the addition of a functional group that is optimally aligned with a corresponding group in the receptor. This figure, when compared to the observed increase in affinity, gives direct feedback on whether or not the new group is actually performing the function anticipated in the design strategy.

An interesting example of how this approach can be used to assess the validity of a drug design hypothesis is provided by the receptor-based design of sialidase inhibitors as potential anti-influenza drugs. Starting from the knowledge [46] of the structurally invariant active site of influenza A and sialidases, von Itzstein et al. [47] postulated that substitution of the 4-hydroxyl group of the nonselective sialidase inhibitor 2-deoxy-2,3-didehydro-D-acetylneuraminic acid **6** (Figure 14.16) with a positively charged substituent would fill an occupied pocket lined with anionic residues. Synthesis and testing of the 4-guanidino analog **7** revealed a reduction in K_i from 10^{-6} to 10^{-10} mole/l, equivalent to an additional binding energy of 23 kJ/mol. Although not at the upper limit of the increments in binding energy anticipated for well-aligned ionic interactions on the basis of the data in Table 14.4, this figure is certainly consistent with the design hypothesis, as is borne out by the crystal structure of the complex [47]. This shows that the guanidino lies between two target carboxyl groups in the active site of the enzyme, although only one appears to be optimally placed for a strong interaction.

F. Ligand Efficiency

The idea of ligand efficiency (LE) has recently emerged as a useful guide to optimize fragment and lead selection in the discovery process [48,49]. Preliminary work has been published by Kuntz et al. [50]. This key contribution, where affinities were examined for a variety of ligands against many different targets, showed that ΔG tends to increase little with molecular mass when the ligand contains more than c. 15 heavy atoms (HA). Later, Hopkins et al. [48] proposed to define LE as the binding free energy for a ligand divided by its number of HA:

$$\mathrm{LE} = -\Delta G/\mathrm{HA} \tag{14.13}$$

According to them, comparison of lead compounds on the basis of LE rather than the potency alone could be useful in deciding the potential for further optimization for particular hits and chemical scaffold. LE has since

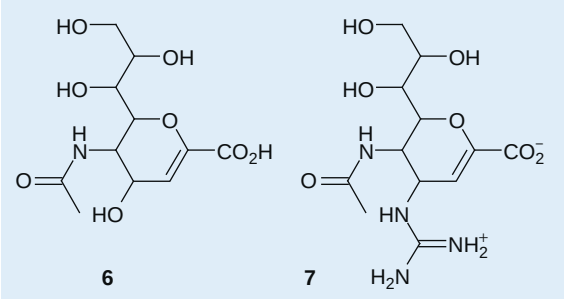


FIGURE 14.16 Sialidase inhibitors: 2-deoxy-2,3-didehydro-D-*N*-acetyl neuraminic acid **7** and its **4**-guanidino analog **8**.

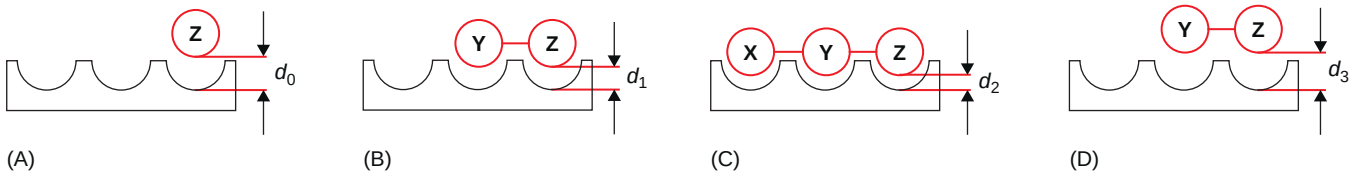


FIGURE 14.17 Schematic representation of a receptor that binds ligands **X**, **Y**, and **Z** with affinities ΔG_X, ΔG_Y, and ΔG_Z, respectively. (a) Binding of **Z** results in a structure with an intermolecular distance d_0. (b) When **Y** and **Z** are connected by a rigid, strain-free linker (**Y**–**Z**), they bind to the receptor with positive cooperativity (ΔG_{Y-Z} more negative than $\Delta G_Y + \Delta G_Z$) and there is structural tightening ($d_1 < d_0$). (c) If **X** is connected to **Y**–**Z** by a rigid, strain-free linker to form **X**–**Y**–**Z**, then further structural tightening will occur ($d_2 < d_1$), leading to a further cooperative enhancement. (d) The shorter linker between **Y** and **Z** does not allow both these binding interactions to occur with optimal geometry. **Y**–**Z** binds the receptor with negative cooperativity (ΔG_{Y-Z} more positive than $\Delta G_Y + \Delta G_Z$) and there is structural loosening ($d_3 > d_0$).

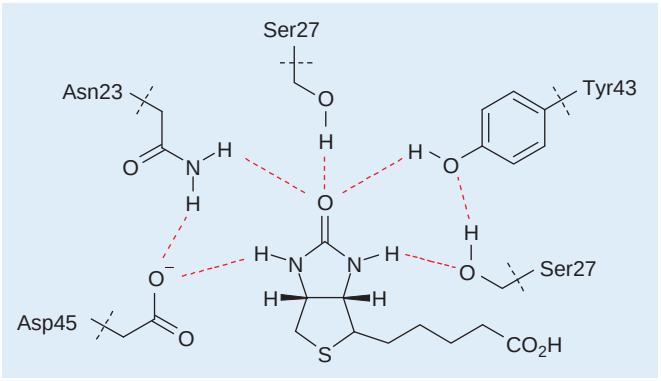


FIGURE 14.18 Schematic representation of streptavidin–biotin interactions.

become an important concept in drug discovery, partly due to the realization that large ligands have a decided disadvantage in terms of the molecular properties necessary for bioavailability [51,52]. LE is notably used when prioritizing the output from HTS or other screening strategies. It is helpful, particularly when trying to assess the relative value of fragments for follow-up in fragment-based drug design [53]. To obtain a final compound with MW $<$500 and $\approx$10 nM potency, LE needs to stay above 1.25 kJ/mol per HA [48].

V. COOPERATIVE BINDING

Equations 14.9 and 14.10 are based on the assumption that contributions to binding energies can be partitioned in term of individual interactions, and that these individual binding energies are additive and independent of each other. However, in general, it is impossible to study one binding interaction in isolation from the others at an interface. In practice, cooperativity between noncovalent interactions is observed.

Noncovalent interactions are said to interact with each other in a positively (or negatively) cooperative manner when the binding energy that is derived from their acting together is greater (or smaller) than would be derived from the sum of their acting separately. This concept, developed by Williams et al. [25,54,55], is illustrated in Figure 14.17.

The consequence of a positive cooperativity is that a structural tightening occurs in the bound state with a benefit in enthalpy and a smaller cost in entropy. By studying the unusually strong reversible binding of biotin by avidin ($K_a \approx 10^{15}$ mol/l) and streptavidin ($K_a \approx 10^{13}$ mol/l), Houk et al. [56] observed that the five hydrogen bonds of the ligand–receptor complex act cooperatively (Figure 14.18), leading to stabilization that is larger than the sum of individual hydrogen-bonding energies. The charged aspartate is the key residue that provides the driving force for cooperativity by greatly polarizing the urea of biotin. If the residue is removed, the network is disrupted.

Following the same principle, a negative cooperativity produces a decrease of the ligand–receptor interaction strength. This is illustrated by Figure 14.17d, where **Y** and **Z** are rigidly held in a conformation that does not

allow both binding interactions to occur with the preferred geometry. This situation could be induced, for example, by introducing a linker between **Y** and **Z** that is too short. The "pull" of **Y** toward its preferred binding geometry will adversely affect the binding of **Z** by forcing it away from its preferred binding geometry, and vice versa (ΔG_{Y-Z} is less negative than $\Delta G_Y + \Delta G_Z$).

References

[1] Klebe G, Dullweber F, Böhm H-J. Thermodynamic models of drug-receptor interactions: a general introduction. In: Raffa RB, editor. Drug-receptor thermodynamics: introduction and applications. Chichester: Wiley; 2001. p. 83–104.
[2] Silverman RB. Drug-receptor interactions: the organic chemistry of drug design and drug action. Amsterdam: Elsevier Academic Press; 2004. p. 123–31.
[3] Motiejunas D, Wade RC. Structural, energetic, and dynamic aspects of ligand-receptor interactions N.p. In: Taylor JB, Triggle DJ, editors. Comprehensive medicinal chemistry, vol. 4. Amsterdam: Elsevier; 2007. p. 193–213.
[4] Di Cera E. Thermodynamic theory of site-specific binding processes in biological macromolecules. Cambridge: Cambridge University Press; 1995.
[5] Böhm H-J, Klebe G. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs? Angew Chem Int Ed Engl 1996;35:2588–614.
[6] Goldstein A, Aronow L, Kalman MS. Principles of drug action: the basis of pharmacology. New York: Hoeber Medical Division; 1968. p. 884.
[7] Dawson RMC, Elliot DC, Elliot WH, Jones KM. Data for biochemical research. Oxford: Oxford University Press; 1969.
[8] Paulini R, Muller K, Diederich F. Orthogonal multipolar interactions in structural chemistry and biology. Angew Chem Int Ed Engl 2005;44:1788–805.
[9] Bajorath J, Kraut J, Li ZQ, Kitson DH, Hagler AT. Theoretical studies on the dihydrofolate reductase mechanism: electronic polarization of bound substrates. Proc Natl Acad Sci USA 1991;88:6423–6.
[10] Jeffrey GA. An introduction to hydrogen bonding. New York: Oxford University Press; 1997.
[11] Stahl M, Bohm HJ. Development of filter functions for protein-ligand docking. J Mol Graph Model 1998;16:121–32.
[12] Williams MA, Ladbury JE. Hydrogen bonds in protein-ligand complexes. In: Böhm H-J, Schneider G, editors. Protein–ligand interactions: from molecular recognition to drug design. Weinheim: Wiley-VCH; 2003. p. 139.
[13] Fersht A. Enzyme structure and mechanism. New York: Freeman; 1985.
[14] Fersht A. The hydrogen bond in molecular recognition. Trends Biochem Sci 1987;12:301–4.
[15] Knox JR, Pratt RF. Different modes of vancomycin and D-alanyl-D-alanine peptidase binding to cell wall peptide and a possible role for the vancomycin resistance protein. Antimicrob Agents Chemother 1990;34:1342–7.
[15a] Horbert R, Pinchuk B, Johannes E, Schlosser J, Schmidt D, Cappel D, et al. Optimization of potent *DFG-in* inhibitors of platelet derived growth factor receptorβ (PDGF-Rβ) guided by water thermodynamics. J Med Chem 2015;58(1):170–82.
[16] Ma JC, Dougherty DA. The cation-π interaction. Chem Rev 1997;97:1303–24.
[17] Meyer EA, Castellano RK, Diederich F. Interactions with aromatic rings in chemical and biological recognition. Angew Chem Int Ed Engl 2003;42:1210–50.
[18] Niedzwiecka A, Marcotrigiano J, Stepinski J, Jankowska-Anyszka M, Wyslouch-Cieszynska A, Dadlez M, et al. Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 59 cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. J Mol Biol 2002;319:615–35.
[19] Schärer K, Morgenthaler M, Paulini R, Obst-Sander U, Banner DW, Schlatter D, et al. Quantification of cation-pi interactions in protein–ligand complexes: crystal-structure analysis of factor Xa bound to a quaternary ammonium ion ligand. Angew Chem Int Ed Engl 2005;44:4400–4.
[20] Di Fenza A, Heine A, Koert U, Klebe G. Understanding binding selectivity toward trypsin and factor Xa: the role of aromatic interactions. Chem Med Chem 2007;2:297–308.
[21] Burley SK, Petsko GA. Weakly polar interactions in proteins. Adv Protein Chem 1988;39:125–89.
[22] Weichsel A, Montfort WR. Ligand-induced distortion of an active site in thymidylate synthase upon binding anticancer drug 1843U89. Nat Struct Biol 1995;2:1095–101.
[23] Dunitz JD. Win some, lose some: enthalpy–entropy compensation in weak intermolecular interactions. Chem Biol 1995;2:709–12.
[24] Silverstein KAT, Haymet ADJ, Dill KA. The strength of hydrogen bonds in liquid water and around nonpolar solutes. J Am Chem Soc 2000;122:8037–41.
[25] Williams DH, Stephens E, O'Brien DP, Zhou M. Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. Angew Chem Int Ed Engl 2004;43:6596–616.
[26] Davis AM, Teague SJ. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. Angew Chem Int Ed Engl 1999;38:736–49.
[27] Vandonselaar M, Hickie RA, Quail JW, Delbaere LT. Trifluoperazine-induced conformational change in Ca(2+)-calmodulin. Nat Struct Biol 1994;1:795–801.
[28] Yu YB, Privalov PL, Hodges RS. Contribution of translational and rotational motions to molecular association in aqueous solution. Biophys J 2001;81:1632–42.
[29] Page MI. Entropy, binding energy, and enzymic catalysis. Angew Chem Int Ed Engl 1977;16:449–59.
[30] Page MI, Jencks WP. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. Proc Natl Acad Sci USA 1971;68:1678–83.
[31] Andrews PR, Craik DJ, Martin JL. Functional group contributions to drug-receptor interactions. J Med Chem 1984;27:1648–57.

[32] Gohlke H, Klebe G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. Angew Chem Int Ed Engl 2002;41:2644–76.
[33] Lau WF, Pettitt BM. Selective elimination of interactions: a method for assessing thermodynamic contributions to ligand binding with application to rhinovirus antivirals. J Med Chem 1989;32:2542–7.
[34] Jencks WP. On the attribution and additivity of binding energies. Proc Natl Acad Sci USA 1981;78:4046–50.
[35] Andrews PR, Kubinyi H, editors. 3D QSAR in drug design: theory, methods, and applications. Leiden: ESCOM Science Publishers BV; 1993.
[36] Williams DH, Cox JPL, Doig AJ, Gardner M, Gerhard U, Kaye PT, et al. Toward the semiquantitative estimation of binding constants: guides for peptide–peptide binding in aqueous solution. J Am Chem Soc 1991;113:7020–30.
[37] Searle MS, Williams DH. The cost of conformational order: entropy changes in molecular associations. J Am Chem Soc 1992;114:10690–7.
[38] Böhm HJ. The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. J Comput Aided Mol Des 1994;8:243–56.
[39] Wolfenden R, Kati WM. Testing the limits of protein–ligand binding discrimination with transition-state analogue inhibitors. Acc Chem Res 1991;24:209–15.
[40] Kati WM, Wolfenden R. Contribution of a single hydroxyl group to transition-state discrimination by adenosine deaminase: evidence for an "entropy trap" mechanism. Biochemistry 1989;28:7919–27.
[41] Wilson DK, Rudolph FB, Quiocho FA. Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations. Science 1991;252:1278–84.
[42] Fersht AR, Shi JP, Knill-Jones J, Lowe DM, Wilkinson AJ, Blow DM, et al. Hydrogen bonding and biological specificity analysed by protein engineering. Nature 1985;314:235–8.
[43] Levine D, Reid TW, Wilson IB. The free energy of hydrolysis of the phosphoryl-enzyme intermediate in alkaline phosphatase catalyzed reactions. Biochemistry 1969;8:2374–480.
[44] Northrop DB, Wood HG. Transcarboxylase. VII. Exchange reactions and kinetics of oxalate inhibition. J Biol Chem 1969;244:5820–7.
[45] Hirsch AK, Fischer FR, Diederich F. Phosphate recognition in structural biology. Angew Chem Int Ed Engl 2007;46:338–52.
[46] Colman PM, Krug RM, editors. The influenza virus. New York: Plenum Press; 1989.
[47] von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dyason JC, Jin B, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 1993;363:418–23.
[48] Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today 2004;9:430–1.
[49] Abad-Zapatero C, Metz JT. Ligand efficiency indices as guideposts for drug discovery. Drug Discov Today 2005;10:464–9.
[50] Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. Proc Natl Acad Sci USA 1999;96:9997–10002.
[51] Abad-Zapatero C. Ligand efficiency indices for effective drug discovery. Exp Opin Drug Discov 2007;2:469–88.
[52] Reynolds CH, Bembenek SD, Tounge BA. The role of molecular size in ligand efficiency. Bioorg Med Chem Lett 2007;17:4258–61.
[53] Leach AR, Hann MM, Burrows JN, Griffen EJ. Fragment screening: an introduction. Mol Biosyst 2006;2:430–46.
[54] Calderone CT, Williams DH. An enthalpic component in cooperativity: the relationship between enthalpy, entropy, and noncovalent structure in weak associations. J Am Chem Soc 2001;123:6262–7.
[55] Williams DH, Davies NL, Zerella R, Bardsley B. Noncovalent interactions: defining cooperativity. Ligand binding aided by reduced dynamic behavior of receptors. Binding of bacterial cell wall analogues to ristocetin A. J Am Chem Soc 2004;126:2042–9.
[56] DeChancie J, Houk KN. The origins of femtomolar protein–ligand binding: hydrogen-bond cooperativity and desolvation energetics in the biotin-(strept)avidin binding site. J Am Chem Soc 2007;129:5419–29.

CHAPTER

15

Compound Properties and their Influence on Drug Quality

Hongming Chen, Ola Engkvist and Thierry Kogej

Chemistry Innovation Centre, Discovery Sciences, AstraZeneca R&D, Mölndal, Sweden

OUTLINE

I. INTRODUCTION

Drug discovery and development is generally a long and expensive process coupled with a high risk of failure. Typically, it takes around fifteen to twenty years to develop a drug from the initial stages of target identification to the introduction of the drug into the market. The cost of introducing a drug molecule into the market has been approximated to be almost $1.2 billion [1]. The US Food and Drug Administration (FDA) estimates that eventually only 8 percent of the compounds that enter Phase 1 clinical trials will reach the market [2]. Though the investment in pharmaceutical research has increased through the years, associated developmental costs of drugs also seems to be increasing [3] and the number of approved drugs has been flat or even reduced. Improving drug quality, including drug efficacy, pharmacokinetic and safet is important for resolving this challeng and a lot of effort has been expended for reducing attrition and speeding up the process of drug development [4,5].

Through the years, various studies have indicated that a major portion of drug attrition is attributable to poor absorption, distribution, metabolism, excretion, and toxicology (ADMET) properties [6,7]. This called for a move toward the "fail early, fail cheap' strategy by investigating drug quality in early phases of drug discovery, since the associated costs increase as a compound progresses through the drug development cascade [8]. Compounds with "off-target" activity (effects on various targets unrelated to the therapeutic target) carry adverse drug reaction (ADR) liabilities and could severely restrict the use of the drug or prevent its entry into the clinical phases. Therefore, compound promiscuity (selectivity) is another property that is related to drug quality. A drug molecule's potency with its pharmacological target is also an important factor for drug quality, since it has direct effect on the drug's efficacy.

This chapter critically reviews the recent advances in assessing drug quality and is arranged to cover the following aspects: compound drug-likeness, target promiscuity, ADMET properties, and binding efficiency metrics. Important physico-chemical properties that influence drug qualities have also been highlighted and discussed.

II. COMPOUND DRUG-LIKENESS ANALYSIS

Starting in the 1980s with the molecular biology revolution and the introduction of combinatorial chemistry in the 1990s, the whole drug-discovery paradigm changed from optimizing molecules with already known effects *in vivo* to screening large sets of molecules *in vitro*. Pioneering efforts in analyzing compound properties in the new era of molecular biology and combinatorial chemistry were published by Lipinski et al [9]. The paper has been cited close to 4,000 times, highlighting the large influence of the analysis. Altogether, 2,245 drugs from the WDI (World Drug Index) [10] were analyzed in order to find out the limit of the 90th percentile for different physico-chemical descriptors. They found that most oral drugs showing passive absorption are in the range of MW (molecular weight) <500, cLogP <5, HBD (hydrogen bond donors) ≤ 5 and HBA (hydrogen bond acceptors) ≤ 10. These guidelines form what is called Lipinski's "rule-of-five." The property distributions of drug-like compounds have also been investigated by others. For instance, it was found that most drugs have between zero and two HBDs and between two and nine HBAs [11]. Recently, a comprehensive review of oral drugs and clinical candidates with properties outside the "rule-of-five" has been published [12]. It has also been shown that oral drugs and clinical candidates have different physico-chemical properties. Drugs on the market have significantly lower lipophilicity and are smaller than clinical candidates [13]. The conclusions are still valid when taking into account which targets the drugs are modulating [14]. The potential advantages and disadvantages of basic and acidic drugs have very recently been discussed [15].

Analysis of the drug-discovery process highlighted that there is a difference in properties between the lead (i.e., the starting point of drug-discovery projects) and the molecule nominated for clinical trials (drug). Several studies have been comparing the identified leads and the final drugs. An initial study compared eighteen lead/drug pairs [16]. It was later extended to include ninety-six lead/drug pairs [17]. The conclusion in both studies was that leads are smaller and more hydrophilic than drugs. The conclusions were later strengthened in a larger study [18]. The importance of lead-likeness has also been reinforced by Leeson et al [19]. Also the concept of "metabolite-likeness" has been discussed in literature and it has been found that there is a general similarity between drugs and human metabolites [20]. The differences between different types of molecules are illustrated in Figures 15.1 and 15.2. In Figure 15.1 the evolution from a fragment to a lead-like molecule and the final drug Vemurafenib [21] is shown. Figure 15.2 illustrates the similarity between a human metabolite ATP and the drug Ticagrelor.

Besides focusing on physico-chemical properties to better understand drug-likeness, significant efforts have been made in statistical analysis and machine learning approaches to understand and predict compound drug-likeness. Schematically, a drug-likeness prediction consists of three important parts: a dataset, a set of descriptors that could be structure or property based, and a statistical analysis or machine learning algorithm. The establishment of cheminformatics as a scientific discipline in the 1990s together with availability of drug databases made it possible to do a statistical analysis of the properties of drugs. The first analysis of the statistical properties of drugs based on structural descriptors was done by Bemis et al in 1996 [22]. The authors introduced a scheme for dividing a molecule into frameworks, ring systems, linkers, and side-chains (Figure 15.3). The molecular frameworks were further divided into atomic and graph frameworks. This division of molecules has been very

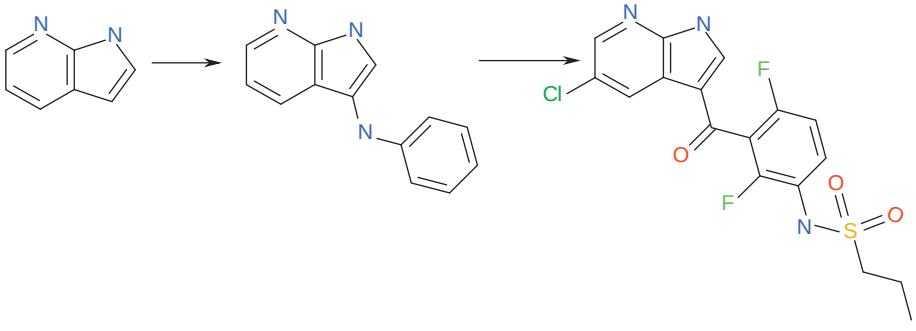


FIGURE 15.1 An illustration of the differences in size between the starting fragment, the following lead-like molecule, and the final drug, Vemurafenib.

influential in the cheminformatics community and is still heavily in use. The authors used the CMC (Comprehensive Medicinal Chemistry) database [24] which contains approximately 5,000 drugs. The results reflected the known druggable target space at that time. Besides very simple molecules that only contain a benzene ring when the side-chains were stripped off, the most common frameworks were from steroids and compounds binding to bio-aminergic GPCRs (G protein coupled receptor). The study clearly showed that certain structural frameworks are overrepresented among drugs. In a follow-up study, the frequency of various side-chains was investigated [25]. On average, there are four side-chains per molecule and each consists, on average, of two heavy atoms. For the compounds in CMC, there are in total approximately 15,000 side-chains, of which approximately 12,000 are represented by the top twenty side-chains. One further analysis comparing the topology of drugs, clinical compounds, and bioactive molecules has also been published [26].

Besides analyzing the molecular structures of drugs to identify distinct features, machine-learning approaches have also been utilized to build predictive models to distinguish drugs from ordinary organic molecules. Several different models were published in the late 1990s and early 2000s. In a study, Gillet et al [27] used substructure analysis to predict bioactivity profiles. The WDI was used as the bioactive training set, and the SPRESI [28] database as the inactive set. The weights for the different substructures were optimized through a genetic algorithm (GA). Two different studies used neural networks (NNs) to distinguish drugs and non-drugs. Ajay et al [29] used Bayesian NNs to distinguish the drugs in the CMC database and the non-drugs in ACD (Available Chemical Directory) [30] database. Over 90 percent of the compounds in CMC were correctly classified as drugs. The drugs in MDDR database were used as an external test set and over 80 percent of the compounds in the database were correctly classified. A similar study based on NN was published by Sadowski et al [31]. It was found that their model could clearly separate the drugs in WDI and the non-drugs in ACD. Other machine-learning techniques

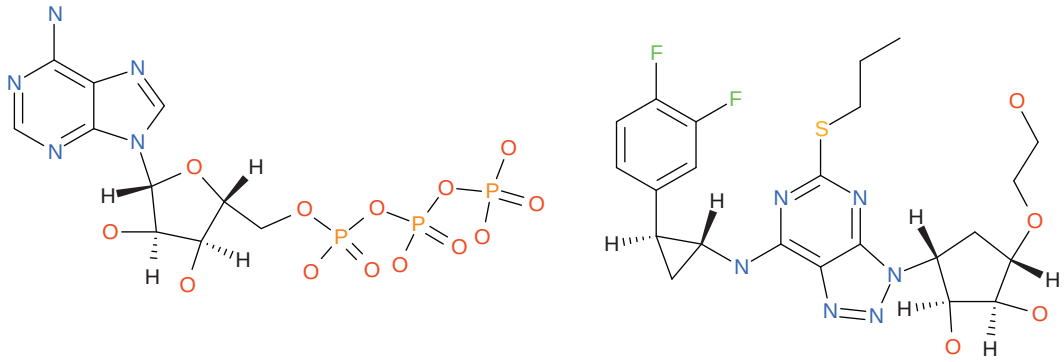


FIGURE 15.2 An illustration of "metabolite-likeness." To the left is the human metabolite ATP (http://en.wikipedia.org/wiki/Adenosine_triphosphate/), and to the right is the drug Ticagrelor (http://en.wikipedia.org/wiki/Ticagrelor/).

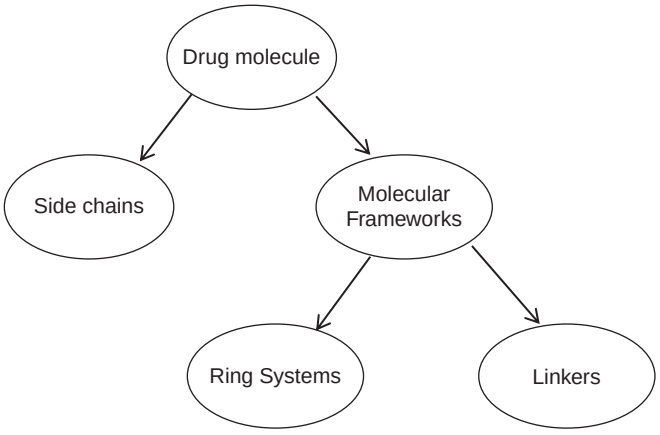


FIGURE 15.3 The partitioning of molecules for property analysis as described in Beamis [23].

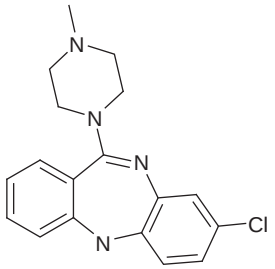
have also been successfully utilized to separate drugs from non-drugs. Byvatov et al [32] used Support Vector Machines (SVM) to separate drugs and non-drugs, with similar results to that of NN models. Many of the early drug-likeness studies were of a "black-box" type (i.e., it was not possible to interpret the models). Thus it was not possible to understand which descriptors were important in discriminating drugs from non-drugs. An exception was a study utilizing decision trees (DT). Wagener et al [33] used DT to separate drugs in WDI from assumed non-drugs in ACD. The classification accuracy was similar to other machine-learning algorithms. The authors found that drugs have more functional groups like hydroxyl, tertiary, or secondary amino and carbonyl groups, while non-drugs are enriched in aromatic rings and halogens. There were also many more studies over the years describing drug-likeness with different types of datasets and machine-learning algorithms [34–40]. Models have also been built to differentiate different types of drugs, such as CNS drugs versus other types of drugs [41]. A comprehensive review of different type of drug-likeness models has been published covering developed models until 2010 [42]. A novel way to predict drug-likeness has recently been proposed by Bickerton et al [43]. The method is called QED (quantitative estimate of drug-likeness). Eight relevant physico-chemical properties were selected (molecular weight, ALogP, number of HBAs, number of HBDs, polar surface area, number of rotatable bonds, number of aromatic rings, and number of structural alerts). Based on a set of 771 small molecule oral drugs, individual desirability functions were defined for each molecular property, mimicking the distribution of the 771 drugs. The QED was created by calculating the geometric mean of the individual desirability functions.

While the above mentioned studies have clearly shown that it is possible to separate drugs from other types of organic molecules, there is a lack of evidence that these models have had any real influence on the drug-discovery process. ADMET models, ligand efficiency metrics, and rule-of-five like rules-of-thumb seem to have been much more influential in the drug-discovery community. There are probably many different reasons for this, which might be related the fact that drug-likeness is not a single property but consists of many independent properties relating to ADMET and binding affinity. Difficulties in defining what a non-drug is have contributed as well. In contrast, models built on individual properties give a medicinal chemist better guidance on how to improve a molecule. Drug-likeness concepts might be useful in library design. However, it might be more advantageous to design a library according to lead-likeness criteria rather than drug-likeness criteria, since the identified hits from an HTS (high-throughput screening) tend to grow in the lead generation and lead optimization phases. An indication that the concept has been less impactful than other concepts like ligand efficiency metrics is that the number of drug-likeness studies has decreased during the last decade after a peak in publications around the year 2000.

III. COMPOUND PROMISCUITY

One of the main goals of a drug-discovery project is to develop highly selective compounds for a therapeutically relevant target while avoiding side effects or ADRs. Pharmacological promiscuity is especially undesirable for this "one drug, one target" paradigm [44–46]. However, there are cases (e.g., complex diseases such as Alzheimer's disease [47] and central nervous system disorders [48]) for which the design of new medicines that address multiple targets simultaneously may be desirable. There is also evidence for promiscuity as a key contributor to the clinical efficacy of many newer anticancer drugs [49]. An understanding of the molecular and structural basis for compound promiscuity could help to design, optimize, and prioritize suitable lead structures at the earliest possible stage of a drug discovery project and therefore significantly increase research productivity.

A compound's off-target activities (secondary pharmacology) are usually inferred from *in vitro* testing of the compound against a panel of proteins [50]. A compound's probability to be active in the target panel is closely related to its ability to interact with different target proteins. This is often called "molecular promiscuity." For example, Clozapine (shown in Figure 15.4), an antipsychotic drug, is a highly promiscuous compound. It is a serotonin antagonist, with strong binding to the 5-HT 2A/2C receptor subtype, and also displays strong affinity to several dopaminergic receptors [51]. The role of physical properties in influencing drug promiscuity has been studied. Leeson et al [19] found that the overall promiscuity is predominantly controlled by lipophilicity and ionization state after they examined the Bioprint database [52]. When lipophilicity is increased, there is an increased likelihood for the compound binding to multiple targets (i.e., higher promiscuity). Bases and quaternary bases are notably more promiscuous than acids, neutral compounds, or zwitterions, while the influence of molecular weight on promiscuity is complex and varies among different ionization states. Bases and neutrals show optimal promiscuity in the 350–500 Da range. Acids, however, show a positive correlation between promiscuity and



Receptor/Transporter protein	Clozapine binding affinity Ki (nM)
5-HT_{1A} (human)	123.7
5-HT_{1B} (human)	519
5-HT_{2A} (human)	5.35
5-HT_{2C} (human)	9.44
D_1 (human)	266.3
D_2 (human)	157
SERT (human)	1624
NET (human)	3168
α_{1A} (human)	1.62
α_{1B} (human)	7

FIGURE 15.4 The structure and some selected biological activity data for Clozapine.

molecular weight. An analysis of Novartis safety pharmacology profiling data [50] showed that molecules with higher MW are more promiscuous. However, researchers from Pfizer analyzed their HTS data and concluded that there is an inverse correlation between promiscuity and MW [53]. These different conclusions indicate that a potential correlation between size and promiscuity may be highly context dependent.

Recently, Yang et al [54] have investigated the relationship between promiscuity and a molecular topological descriptor, the fraction of molecular framework (f_{MF}). They first divided the molecule into a molecular framework (MF) and side-chains as defined by Bemis and Murcko, as exemplified in Figure 15.5 [22,25]. The descriptor f_{MF} is defined as the number of heavy atoms (Nheavy) in the MF divided by the total number of heavy atoms in the molecule, as shown in Equation 15.1, where $f_{MF} \in [0, 1]$.

$$f_{MF} = \frac{Nheavy_{MF}}{Nheavy_{total}} \tag{15.1}$$

Molecules with low f_{MF} value tend to have small frameworks and are decorated with many large side-chain atoms, while a high f_{MF} value corresponds to a molecule having a large framework and fewer or smaller side-chain atoms. A positive correlation has been found between the f_{MF} descriptor and promiscuity for the Bioprint dataset. It was found that molecules with f_{MF} above 0.65 are generally more promiscuous than other molecules. Accordingly, molecules with a large molecular framework and only a few side-chain atoms exhibit higher promiscuity than other types of molecules. More importantly this relationship seems to be independent of lipophilicity, since after dividing the compound set into several cLogP [55] bins, the trend is still valid in each cLogP bin.

Peters et al [56] looked into molecular promiscuity for the Bioprint dataset from a molecular structural and target class perspective. Their analysis demonstrated that molecules with a positive charge center show highest promiscuity, which is consistent with Leeson's conclusion [19]. When they divided the Bioprint dataset according to target class and calculated the average hit rate for each target class, the aminergic GPCR set seemed to be the most promiscuous target class with an average hit rate of 5.6 percent. A closer inspection revealed that some individual aminergic GPCR targets have a surprisingly high hit rate. For example, 5-HT_{2b} GPCR interacts with

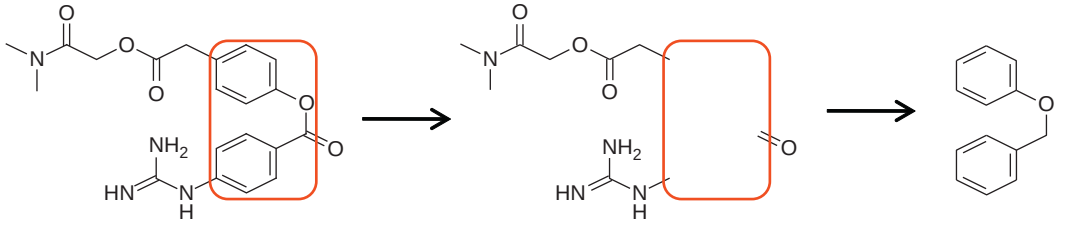


FIGURE 15.5 Disconnecting side-chains from the original molecule leads to its molecular framework [54]. *Reprinted with permission from the Journal of Medicinal Chemistry, Copyright 2009, American Chemical Society.*

almost one-third of positively charged compounds. In terms of promiscuity of structural motifs, their analysis has shown that tricyclic scaffolds and "ergoline-like" motifs form the most promiscuous substructures. Indole, piperazine motifs are also overrepresented in a lot of promiscuous aminergic ligands.

IV. COMPOUND ADMET PROPERTIES

Compound ADMET properties can be divided into two classes based on the way of measuring them, namely, the *in vitro* and *in vivo* classes. For example, some common *in vitro* measured ADMET properties in drug-discovery projects are aqueous solubility, logarithm of octanol−water partition coefficient (logP/logD), pKa, Caco-2/MDCK permeability, hERG (human *Ether-à-go-go*-Related Gene) inhibition, various CYPs (Cytochromes P450 enzyme) inhibition, plasma protein binding, and liver microsome stability. The *in vivo* ADMET properties are measured in animal models and include various pharmacokinetic properties such as oral bioavailability (F), human intestinal absorption (HIA), urinary excretion, area under the plasma concentration−time curve (AUC), total body clearance (Cl), volume of distribution, and elimination half-time (t1/2).

In 1991, about 40 percent of drug candidate attrition was caused by adverse pharmacokinetics and a lack of bioavailability, while by 2000 this amount had been dramatically reduced to 10 percent, and the primary reason for attrition was the lack of efficacy, which accounted for 30 percent of all failures in drug development [5]. As to the attrition caused by drug toxicity, the percentages were about 15 percent and 25 percent in 1991 and 2000, respectively. Clearly, the ADMET related attrition significantly dropped when the pharmaceutical industry began to remedy the biggest cause of attrition. To reduce late-stage attrition, it is critical to identify compounds that are unlikely to succeed and to terminate their development as early as possible. Only compounds exhibiting a good ADMET profile should be advanced into clinical trials, thus saving development resources. Now, HTS of ADMET properties such as inhibition of human CYP enzymes and human hERG ion channels are routinely conducted in pharmaceutical companies. However, the *in vitro* and *in vivo* assays are very time consuming and costly. Numerous studies were carried out on *in silico* prediction of ADMET properties, and they can be loosely classified into two categories. One category of research is to build predictive QSAR models using advanced statistical algorithms along with large numbers of molecular descriptors. During the years, many *in silico* models have been developed for almost every ADMET end point and are extensively reviewed elsewhere [57−62]. These models are often complicated and not interpretable. Another kind of research focuses on building simple interpretable ADMET rules-of-thumb to provide general guidance for medicinal chemists in the optimization of a lead compound's ADMET properties. In this case, only a few interpretable molecular descriptors are used in the study, and the qualitative relationship between descriptor and end point property are identified. Here we mainly summarize research works in the latter category.

Since Lipinski et al [9] profiled a range of physico-chemical properties of drugs and derived the well-known "rule of 5" (see above), an increasing body of literature [19,63−66] suggests that poor ADMET outcomes are predominantly correlated with increasing lipophilicity and molecular size. Gleeson [63] did an analysis of a number of key ADMET assays run within GSK: solubility, permeability, bioavailability, volume of distribution, plasma protein binding, CNS penetration, brain tissue binding, P-gp efflux, hERG inhibition, and CYP 1A2/2C9/2C19/2D6/3A4 inhibition. A set of rules has been formulated using molecular properties that chemists intuitively know how to alter in a molecule, namely, molecular weight, cLogP, and ionization state. Gleeson's analysis shows that almost all ADMET parameters deteriorate with increasing molecular weight, cLogP, or both. The

ionization state plays either a beneficial or detrimental role depending on the parameter in question. In general, compounds with a cLogP >4 and MW >400 Da have a less favorable ADMET profile. Leeson and Springthorpe's [19] analysis on patent data from several pharmaceutical companies indicate that recent medicinal chemistry efforts are still often producing compounds with much higher molecular mass and cLogP values than historical drugs. They have linked this trend to the likelihood of compounds failing in development as a result of poor ADMET characteristics.

Yan et al [67] proposed a descriptor (Fsp3) indicating the degree of aliphatic, the fraction of sp3 hybridized carbon in total carbon count, to represent the carbon-bond saturation. The Fsp3 descriptor can also be regarded as a shape descriptor by nature: a higher Fsp3 value corresponds to less flatness or more "three dimensionality" of a compound. Lovering et al [68] found that this simple descriptor significantly correlates with the success rate of a clinical candidate to become a drug. Their analysis shows that more saturated compounds, corresponding to a higher Fsp3 value, are more likely to succeed in the transitions. Also, compounds with higher solubility and lower melting points tend to have higher values of Fsp3. Inspired by this finding, Yang et al [69] studied the influence of Fsp3 on several ADMET properties based on in-house historical data and confirmed the influence of Fsp3 on solubility, with aqueous solubility positively correlating the values of Fsp3 for over 32,000 in-house compounds. The effect of Fsp3 was also found to be independent from lipophilicity (cLogP) and ionization state in this data set. Increasing Fsp3 was shown to reduce plasma protein binding but had lesser effects on Caco-2 permeability, hERG binding, and CYP 3A4 inhibition. Richie and Macdonald [70] have looked into the impact of aromatic ring count (the number of aromatic rings and hetero aromatic rings in the molecule) on in-house ADMET benchmark data: aqueous solubility, lipophilicity, serum albumin binding, hERG inhibition, and CYP inhibition. In their analysis, it was found that as the number of aromatic rings increases, there is a concomitant decrease in aqueous solubility and an increase in human serum albumin binding, CYP450 3A4 inhibition, and hERG binding. They concluded that the fewer aromatic rings are contained in an oral drug candidate, the more developable the candidate is likely to be. More than three aromatic rings in a compound generally correlate with poor developability and thus an increased risk of attrition in drug development. The effect of aromatic ring count was further investigated by Ritchie et al [71] in 2011, where carboaromatic (i.e., benzene) ring count, heteroaromatic ring count, and carboaliphatic and heteroaliphatic ring count were analyzed separately on a number of ADMET related readouts. It was demonstrated that the decreasing order of detrimental effect on compound developability for these four ring types is: carboaromatics >> heteroaromatics > carboaliphatics > heteroaliphatics. This result highlights the particular detrimental effect of aromatic ring count on ADMET properties. Another aromaticity related descriptor was proposed by Yan [67], and it was defined as the numbers of aromatic atoms divided by the total number of atoms (excluding hydrogen atoms) in the molecule, often abbreviated to Ar/HA (heavy atoms) or AP (aromatic proportion). It is actually inversely related to Fsp3 and has been used to build predictive models for solubility. The rationale of using an aromaticity descriptor in a solubility model is that aromatic compounds are more likely to form dense crystals with higher melting points and are less flexible, leading to lower solubility [72]. In contrast to the fraction based descriptors like Fsp3 and Ar/HA, Leeson et al [73] introduced an additional metric for the aromatic/aliphatic balance (Ar-sp3), being the aromatic atom count minus the number of sp3 carbon atoms. The advantage of this parameter is that compounds with zero aromatic atoms or zero sp3 carbons are also parameterized. Their analysis shows that Ar-sp3, lipophilicity and the number of hydrogen bond donors of oral drugs remain constant over time, and there is a higher degree of aromaticity among pharmaceutical company patent compounds (63 percent compounds having more than two aromatic rings) than oral drugs (14 percent drugs having more than two aromatic rings). This conclusion implies that the aromaticity/aliphaticity balance probably has impact on drug ADMET properties. Roughley and Jordan [74] examined the number and type of aromatic rings in the medicinal chemistry literature data. They found that carboaromatic rings were more prevalent than heteroaromatic rings and thus introduced the benzenoid index (BI), being equal to the number of carboaromatic (i.e., benzene) rings/total number of aromatic rings. A higher value of BI, indicating a higher proportion of benzenoid character, is more likely to result in poorer compound developability.

Recently, Hill and Young studied GSK solubility data [75] and proposed the so-called "solubility forecast index" (SFI), which is a simple summation of the computed or chromatographically measured logD7.4 and the aromatic ring count. This parameter was found to have more predictive power than the individual parameters, which had low correlation with each other, implying that aromatic ring count influences solubility by both lipophilicity dependent and independent mechanisms. The same parameter was then extended to other ADMET screening data, such as permeability, CYP450 inhibition, intrinsic clearance, hERG binding, and promiscuity [76]. It was shown to be a generally applicable index to predict ADMET behavior and was then renamed to "property forecast index" (PFI). As the value of PFI increases, the number of compounds that could meet defined target

values with respect to solubility, permeability, CYP450 inhibition, clearance, hERG inhibition, and promiscuity decrease. It was suggested that a PFI of 7 or less is desirable to ensure that compounds exhibit good levels of developability.

Most of the abovementioned studies were carried out by binning continuous data and making statistical trend analyses based on the average values of bins. However, Kenny et al [77] have called for caution in doing this kind of trend analysis by arguing that averaging groups of data points prior to analysis could create correlation inflation and therefore exaggerate the observed trend.

V. LIGAND BINDING EFFICIENCY METRICS

Over the last three decades, an overall increase in MW of synthesized molecules in drug discovery has been observed. This increase of 100 Da in the mean of MW is 50 percent higher than in the approved drugs set [43]. Furthermore, studies have demonstrated that clinical candidates [13] or pre-clinically interesting molecules published in the literature [78] or patents [79] are generally more lipophilic, larger, and less three-dimensional [68,73] than oral drugs. This trend has been attributed to an increasing usage of HTS [80] for finding chemical starting points in the lead generation phase. As larger and more lipophilic compounds may potentially present higher binding affinities, they have emerged more easily as hits in HTS than smaller molecules. Subsequently, a tendency of inflating the molecular size and lipophilicity during the follow-up lead optimization process has been generally noticed [17,81]. This is supported by thermodynamic observations revealing the necessity of increasing the size of ligands to achieve higher potency [82,83]. In addition, there is a growing interest in modulating targets that require bigger ligands [83,84]. However, it is well established that molecules with high MW and/or lipophilicity have a higher probability to fail at the different stages of drug development due to ADMET issues [13,63,85,86] (see above).

A. Ligand Efficiency

Ligand efficiency (LE) has been proposed as a useful parameter to mitigate the influence of MW, lipophilicity, and other physico-chemical properties when increasing the potency of hit and lead compounds. The LE concept has been derived from the observation that the maximum affinity achievable by ligands is -1.5 kcal/mol per heavy atom (e.g., nonhydrogen atom) [84,87]—ignoring cations and anions [88]—and from functional group binding energy studies [89]. Basically, if we assign all the intermolecular interaction to the ligand alone, LE corresponds to the Gibbs free energy of binding per heavy atom: [88]

$$LE = -\frac{2.303\ RT}{HA}\log(K_d/C) \tag{15.2}$$

where HA denotes the number of heavy atoms, R is the ideal gas constant ($1.987\ 10^{-3}$ kcal.mol/K), T is the absolute temperature (Kelvin), K_d is the dissociation constant of the ligand−protein complex, and C is the ligand concentration. Thus, LE allows comparing ligands according to their average binding energy per atom [87,88]. It assumes that the contribution of all atoms of the ligand are additive and play an equal role in the molecular recognition with the target protein. This clearly represents a limitation, since different types of atoms like carbon, nitrogen, oxygen, sulfur, and halogen have different size and binding capabilities [84,90]. Moreover, some groups having the same number of heavy atoms may have completely differently charged species (e.g., CH_3 and NH_2). Inversely, in cases where some groups present similar contributions to binding, the preference might be given to the one having the lower number of heavy atoms (higher LE), while the larger group could present a better starting point for an optimization campaign by improving the overall physico-chemical profile (e.g., lower lipophilicity) [91]. It was then recommended to consider other properties in parallel—such as potency, lipophilicity, solubility, and pharmacokinetics—instead of LE alone [92].

Assuming within standard conditions of temperature (T = 300K), neutral pH, and a concentration of 1M, Equation 15.2 becomes:

$$LE = -\frac{1.37\log K_d}{HA} = 1.37\ pK_d/HA \tag{15.3}$$

It was pointed out that the choice of the concentration (here, 1M) can dramatically influence how the LE value evolves within a given chemical series [93].

By using the negative logarithmic values of the half maximal inhibitory concentration (pIC_{50}) or the negative logarithmic value of the inhibition constant (pK_i), Equation 15.3 leads to:

$$LE = 1.37\ pIC_{50}/HA \tag{15.4}$$

and

$$LE = 1.37\ pK_i/HA \tag{15.5}$$

As IC50 values in competitive binding experiments are dependent on the concentration of the competing ligand, Equation 15.4 should preferably be used to compare inhibitors under the same assay condition.

Thus, for a given value of LE, HA is linearly correlated to potency (pK_d, pIC_{50}, or pK_i), with a slope of 1.37/LE. Obtaining a 10-fold increase in affinity for a constant value of LE requires a decrease of 1.37/LE of the number of heavy atoms. Therefore, for ligands having different LE values, the change in HA values to increase the affinity of one order of magnitude may differ [90], as the change of the number of heavy atoms in a small molecule will have more effect on ΔLE (i.e., the difference of the LE of the ligand before and after structural modification) than a larger molecule [84,88,92,94,95]. The LE concept is summarized in Figure 15.6, where we plot the p(Activity) (e.g., pIC_{50}, pK_i, pK_D) with respect to HA for a set of fifty-nine compounds presenting different LE values, while another parameter, LLE (see below), is used to guide the structural modification. These data demonstrate that the optimization of these compounds is concurrent with an increase of the LE values.

The concept of LE has been extended to take into account other important physico-chemical properties in drug design, such as lipophilicity [19] (see below), molecular mass [23,96] (Binding Efficiency Index (BEI)), polar surface area [23,96] (Surface Efficiency Index (SEI)), combinations of physico-chemical properties [80,97], and functional group contributions [98]. While it can be used at different stages of the drug-discovery process, the LE concept has been mainly focused on *in vitro* binding affinity data and not on *in vivo* data [84].

B. Thermodynamic Aspects of Binding

While a detailed analysis of the thermodynamic terms related to LE is beyond the scope of this chapter, it is worth highlighting several major attributes. An increase of ligand affinity corresponds to a decrease of the free energy of binding (ΔG), which combines the binding enthalpy (ΔH) and the binding entropy (TΔS):

$$\Delta G = \Delta H - T\Delta S \tag{15.6}$$

ΔH can be optimized by increasing the number of specific polar interaction, such as hydrogen bonding between the ligand and its protein binding site [99,100], while TΔS can be increased through nonspecific interaction by having a larger ligand size and higher ligand lipophilicity [101]. The tendency to favor the optimization of the entropic term of the binding free energy over the enthalpy term has probably contributed to the inflation of the physico-chemical properties used in lead optimization [17,80,81,84,102]. Several studies show that increasing ligand size does not necessarily result in a more favorable entropic term [95,103], as many large molecules have more conformational constraints than one might expect based on the number of heavy atoms or rotatable bonds. Similarly, ligand enthalpy and entropy efficiencies per heavy atom [104] can also be calculated from Equation 15.5 and 15.6. Isothermal titration calorimetry (ITC) experiments on a broad range of ligands revealed that the average entropy efficiency does not evolve significantly with increasing molecular size, while enthalpy efficiency values do, and the variation of overall LE values is in line with enthalpy efficiency [105–107]. This suggests that the changes in LE are mainly due to the enthalpy efficiency [103,107], and the contribution of entropy is not as important as it is usually believed [84]. It has been reported that the chance to increase the enthalpy driven affinity contribution diminishes as a molecule becomes larger than a lead-size compound (~20–30 heavy atoms), so that further affinity improvements could only be achieved via entropy driven optimization [82]. In other words, this stresses the necessity to optimize specific binding interactions (increase enthalpy driven affinity) as early as possible [82].

C. Lipophilic Ligand Efficiency

Lipophilic ligand efficiency (LLE), also referred as LipE (=lipophilic efficiency) [91], has been proposed as a better alternative than LE or lipophilic ligand efficiency (LELP = cLogP/LE) to capture the enthalpic component of ligand binding, especially in series of analogs presenting similar lipophilic binding interaction [91]. Recent

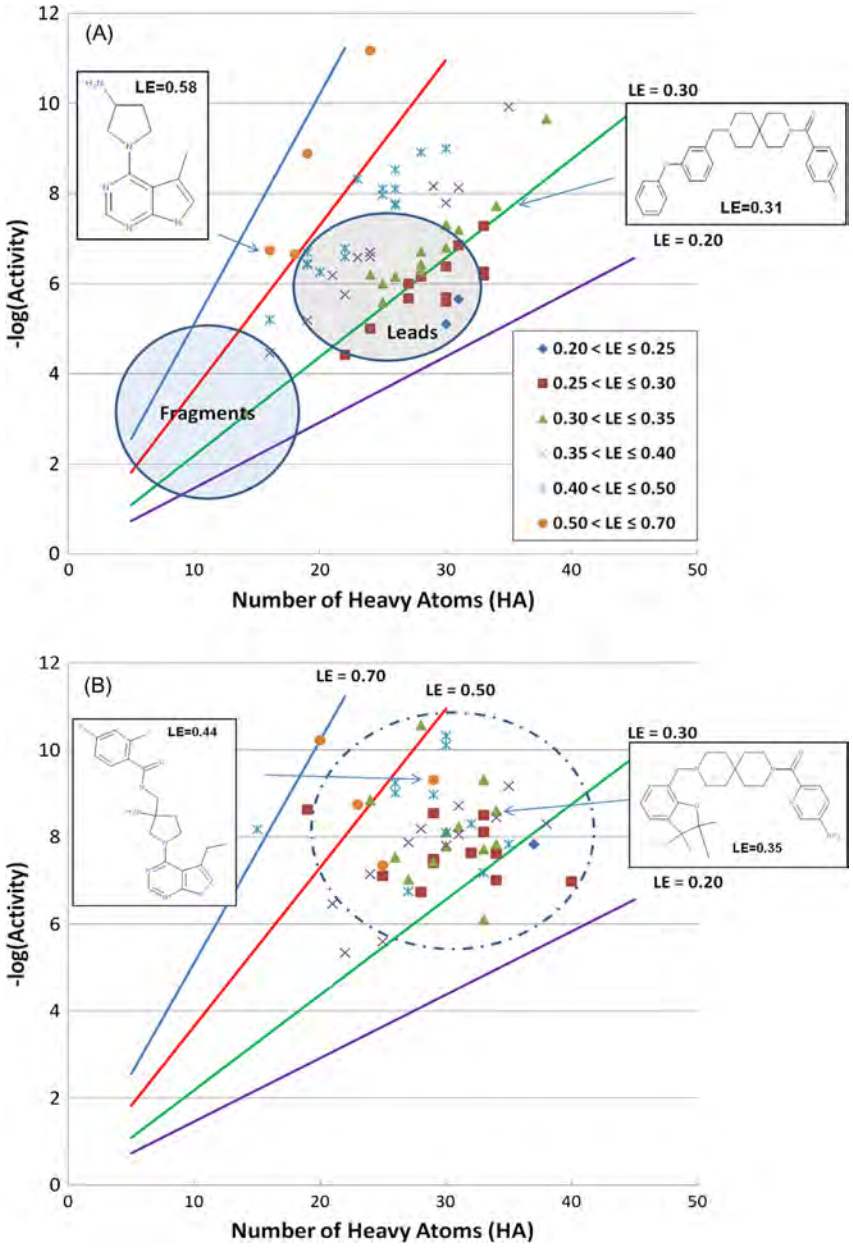


FIGURE 15.6 Representation of the LE concept. The straight lines show the theoretical evolution of the "pActivity" (e.g., pIC_{50}, pK_i, pK_D) with respect to the number of heavy atoms (HA) for different values of LE. On plot (a), we display two hypothetical regions where fragment and lead-like ligands are usually located. The two plots (a and b) correspond to the same series of fifty-nine active compounds to forty-seven different proteins (additional data on those compounds can be found in the supplementary materials of reference [75]) before (a) and after (b) optimization. We observe that the values of LE tend to increase upon optimization, as the dashed circle covering most of the area where the compounds are localized moves up on plot (b) compared to plot (a). The presented structures are chosen to specifically point out that the optimization does not necessarily imply an increase of the LE value. Note also that these compounds were optimized using the lipophilic ligand efficiency (LLE, see below).

analyses demonstrate that LLE is a more robust metric over LE in supporting decision-making in selection and optimization programs [91,108].

LLE can be derived from the following equation:

$$LLE = pIC_{50} - cLogP \tag{15.7}$$

Namely, it corresponds to the difference between the pIC_{50} (or pK_i, pK_d) and the lipophilicity expressed by cLogP (which refers here to a calculated partition coefficient) or, similarly, the LogD (distribution coefficient). LLE is an estimate of the preference of a ligand to bind to the target protein rather than partition into 1-octanol [19] when the partition factor is determined in a water/1-octanol experiment. A ligand with a LLE equal to zero based on cLogP can be seen as having the same affinity for its target as for 1-octanol partition. Similarly, a ligand with a LLE of 6 has a one-million-fold higher affinity for its target than for 1-octanol. From Equation 15.7, we can assume that in order to achieve 1−10 nM range or higher affinity for a drug-like molecule having a cLogP in the range of 2.5-3.0, LLE must be in the range of 5−7 or higher [19], while, on the other hand, negative LLE values

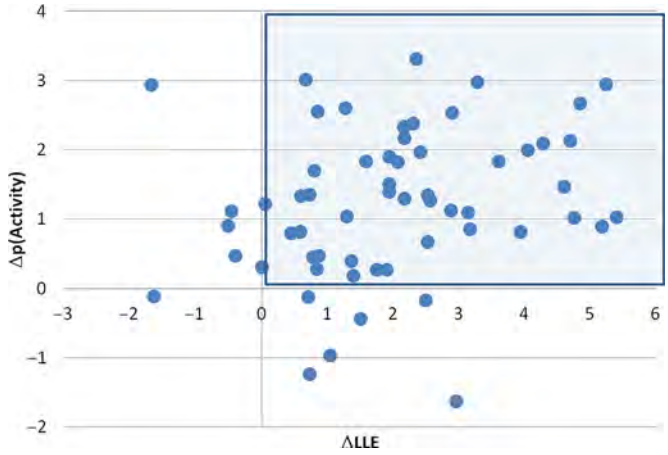


FIGURE 15.7 Examples from the literature in which lipophilic ligand efficiency was explicitly used in the optimization process of fifty-nine compounds targeting forty-seven different proteins. Changes in p(Activity) between the starting point and the optimized compound are plotted against the corresponding changes in the lipophilic ligand efficiency (LLE). The upper right quadrant (highlighted area) contains molecules where potency and LLE are increased (forty-eight out of fifty-nine compounds).

should clearly be avoided. HTS usually reveals hits with LLE <2, as common affinity is <1 μm and cLogP for a lead-like compound is around 4 [80]. Finally, it is important to note that the LLE values depend on the experimental condition of the LogD measurements or the software and the normalization procedure of chemical structures when using cLogP.

Recent publications demonstrate a growing number of successful examples for applying LLE in drug discovery [84]. A list of fifty-nine examples covering forty-seven different target proteins is summarized in Figure 15.7 (generated using the data found as supplementary materials of [84]).

Despite the success of LLE in supporting hit selections and optimizations [84], it is worth mentioning some of its limitations here. First, it can be challenging to use LLE for comparing molecules with very different size and properties [84]. Secondly, LLE is regarded as less efficient when the target protein favors binding highly polar compounds. An example has been reported in the case of influenza neuraminidase, where a more lipophilic compound, oseltamivir (cLogP = 2.1), led to a more successful orally available compounds compared to the first non-orally available drug zanamivir (cLogP = −5.6), which then present a better (but confusing) LLE value [84]. At a molecular recognition level, any modification occurring at distal position, such as one that is solvent exposed, may lead to a misinterpretation of the binding contribution to the LLE value. This may occur when solvent-exposed polar groups are incorporated to the ligand, resulting in a decrease of lipophilicity and an increase of LLE, while no new interaction between the ligand and the protein have been created (and no change of potency has been measured) [91].

D. Ligand Efficiency and Fragment-Based Lead Generation

Fragment hits usually have low potency ranging from 1 mM to 10 μM due to their small size (100 Da to 250 Da). The use of LE can facilitate the identification of such hits, since their affinity is normalized in such a way that fragments can have similar or even higher LE values than more potent but bigger molecules [88,95]. However, adding 200−250 Da. molecular mass (i.e., roughly 15−20 heavy atoms) to the initial active fragment to achieve nanomolar potency is not a trivial task. Several examples of successfully employing ligand efficiencies metrics to improve drug-like properties in fragment based lead generation (FBLG) have been reported [109−113]. In the case of a heat shock protein 90 (HSP90) FBLG program [109−111], six orders of increased affinity have been achieved by efficiency adding only six heavy atoms (potency increased from 0.79 μM to 0.70 nM). In this study, a group efficiency (GE) has been introduced as a useful guide to improve affinity while controlling chemical properties [98]. This was done via a matched-pair procedure, which led to a GE value for each group in the final molecule. The authors stressed that the use of the X-ray crystal structure and careful monitoring the chemical properties was crucial to achieve these results.

E. Size-independent Measures of Ligand Efficiency

It was noticed that the average or optimal LE values are systematically higher for small molecules than for large ones [94,95] and the relationship between LE and the number of HA is nonlinear. Attempts were therefore made to transform LE into metrics that are more consistent within a broader range of molecular size. Generally, LLE evolves linearly with respect to cLogP, and the gradient of the slope depends on the influence of ligand lipophilicity on affinity (the slope is equal to -1 if there is no relationship between potency and ligand lipophilicity; see supplementary material of [84]). Efforts to normalize LLE with a heavy atom count led to the suggestion of lipophilicity-corrected ligand efficiency (LELP) and LLE adjusted for heavy atom count (LLE_{AT}): [80]

$$LELP = cLogP/LE \tag{15.8}$$

$$LLE_{AT} = 0.111 + 1.37(LLE/HA) \tag{15.9}$$

Optimum values for LELP are in the range of 0−7.5 and correspond to a LE range of 0.40−0.45 and a cLogP interval from 0 to 3 [80]. LLE_{AT} was derived from an FBLG effort [97]. LELP and LLE_{AT} are also valid for lead optimization [84]. Studies have shown that optimizing LLE and LELP helped in improving the ADMET properties of the active molecules [114].

Two size-independent modification of LE using only HA have been developed: fit quality (FQ) [95] and the size-independent ligand efficiency (SILE) [115]. FQ is obtained by binning the LE values for a large number of ligand−protein complexes and applying a scaling factor derived from the most potent compounds in each bin: [95]

$$FQ = [(pIC_{50}\ or\ pK_i)/HA]/[0.0715 + (7.5328/HA) + (25.7079/HA^2) - (361.4722/HA^3)] \tag{15.10}$$

SILE is similar to FQ, where a different fitting function was employed: [112]

$$SILE = (pIC_{50}\ or\ pK_i)/HA^{0.3} \tag{15.11}$$

The reasons for replacing HA count alone as an estimate of the molecular size are mainly based on three factors. First, the HA count is being used to surrogate for molecular surface, which seems to be a more relevant parameter to molecular recognition and binding affinity. Second, computational works show that the effect of increasing the molecular surface by adding heavy atoms decreases with the molecular size [95]. As a result, the larger molecules tend to fold at a larger extent and present more buried (internal) molecular surface than smaller molecules [84]. Third, it was also observed that the possible adjustments in bond distance and angle within a molecular structure without introducing extra intramolecular energy strain decreases as the molecule size increases [95].

VI. CONCLUSIONS AND FUTURE OUTLOOK

Cheminformatics and data analysis are now key parts of the drug-discovery process. The first analysis was published in the 1990s, and from then onward a steady stream of new models and articles have been published. The models have varied tremendously in complexity and in impact. A main observation is that the simplest models might also have been the most successful, which might be due to how easy they are to understand and interpret. The Lipinski "rule-of-five" is still used as a drug-likeness filter, despite the development of much more complex models that can better separate between drugs and other compounds.

Drug selectivity is an important aspect for evaluating the ADRs of drugs, and there is evidence that compound target promiscuity is largely correlated with its lipophilicity and ionization states. It is widely agreed that drug attrition is closely related with its ADMET properties. Both ADMET *in silico* models and rules-of-thumb are now well established and routinely used in the drug-discovery process. The LE concept has had tremendous impact in FBLG as well as in lead generation and lead optimization. Despite the criticism that it is an oversimplified measure, it has been used pragmatically as a parameter in optimizing the compounds in drug discovery. While complex models of predicting drug-likeness aroused a lot of excitement when they were developed, over time they have been less impactful. Efficiency measures other than ligand efficiency and ligand lipophilic efficiency have been developed, but so far have had less impact than these. We believe that for normal small molecules existing in current screening collections, drug-likeness analysis, ADMET modeling, and LE metrics are now mature tools that are used as a normal part of the drug-discovery process. We expect that a major future

challenge will be to extend these concepts to new modalities, such as macrocycles, peptides, and microRNA. It is believed that for modulating relevant targets in many disease areas, modalities other than small molecules are needed. An example of work in this direction is an analysis of the properties of macrocycles that has recently been published [116].

References

[1] Michielan L, Moro S. Pharmaceutical perspectives of nonlinear QSAR strategies. J Chem Inf Model 2010;50(6):961–78.
[2] Peck RW. Driving earlier clinical attrition: if you want to find the needle, burn down the haystack. Considerations for biomarker development. Drug Discov Today 2007;2(7–8):289–94.
[3] Rawlins MD. Cutting the cost of drug development? Nat Rev Drug Discov 2004;3(4):360–4.
[4] Schuster D, Laggner C, Langer T. Why drugs fail: a study on side effects in new chemical entities. Curr Pharm Des 2005;11(27):3545–59.
[5] Kola I, Landis J. Can the pharmaceutical industry reduces attrition rates? Nat Rev Drug Discov 2004;3(8):711–16.
[6] Hou TJ, Xu XJ. ADME evaluation in drug discovery. 3. Modelling blood-brain barrier partitioning using simple molecular descriptors. J Chem Inf Comput Sci 2003;43(6):2137–52.
[7] Clark DE. In silico prediction of blood-brain barrier permeation. Drug Discov Today 2003;8(20):927–33.
[8] Dickson M, Gagnon JP. Key factors in the rising cost of new drug discovery and development. Nat Rev Drug Discov 2004;3(5):417–29.
[9] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. Adv Drug Deliv Rev 1997;23(1–3):3–25.
[10] World Drug Index database. Mission Viejo, CA: Daylight Chemical Information Systems, Inc.
[11] Oprea TI. Property distribution of drug-related chemical databases. J Comp Aid Mol Des 2000;14(3):251–64.
[12] Over B, Giordanetto F, Kihlberg J. Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. Chemi Biol 2014;21(9):115–1142.
[13] Wenlock MC, Austin RP, Barton P, Davis AM, Leeson PD. A comparison of physiochemical property profiles of development and marketed oral drugs. J Med Chem 2003;46(7):1250–6.
[14] Tyrchan C, Blomberg N, Engkvist O, Kogej T, Muresan S. Physicochemical property profiles of marketed drugs, clinical candidates, and bioactive compounds. Bioorg Med Chem Lett 2009;19(24):6943–7.
[15] Charifson PS, Walters WP. Acidic and basic drugs in medicinal chemistry: a perspective. J Med Chem 2014;57(23):9701–17.
[16] Teague SJ, Davis AM, Leeson PD, Oprea T. The design of leadlike combinatorial libraries. Angew Chem Int Ed 1999;38(24):3743–8.
[17] Oprea TI, Davis AM, Teague SJ, Leeson PD. Is there a difference between leads and drugs? A historical perspective. J Chem Inf Comput Sci 2000;41(5):1308–15.
[18] Allu TK, Oprea TI. Rapid evaluation of synthetic and molecular complexity for in silico chemistry. J Chem Inf Comput Sci 2005;45 (5):1237–43.
[19] Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 2007; 6(11):881–90.
[20] Chen H, Engkvist O, Blomberg N, Li J. A comparative analysis of the molecular topologies for drugs, clinical candidates, natural products, human metabolites, and general bioactive compounds. Med Chem Comm 2012;3(3):312–21.
[21] de Kloe GE, Bailey D, Leurs R, de Esch IJP. Transforming fragments into candidates: small becomes big in medicinal chemistry. Drug Disc Today 2009;14(13/14):630–46.
[22] Bemis GW, Murcko MA. The properties of known drugs. 1. Molecular frameworks. J Med Chem 2000;39(15):2887–93.
[23] Abad-Zapatero C, Metz JT. Ligand efficiency indices as guideposts for drug discovery. Drug Discov Today 2005;10(7):464–9.
[24] Comprehensive Medicinal Chemistry Database, Accelrys, Incorporation, San Diego, CA.
[25] Bemis GW, Murcko MA. The properties of known drugs. 2 Side chains. J Med Chem 2000;42(25):5095–9.
[26] Chen H, Yang Y, Engkvist O. Molecular topology analysis of the differences between drugs, clinical candidate compounds, and bioactive molecules. J Chem Info Model 2010;50(2):2141–50.
[27] Gillet VJ, Willet P, Bradshaw J. Identification of biological activity profiles using substructural analysis and genetic algorithms. J Chem Inf Comput Sci 1998;38(2):165–79.
[28] SPRESI database. Mission Viejo, C.A.: Daylight Chemical Information Systems, Inc.
[29] Ajay A, Walters WP, Murcko MA. Can we learn to distinguish between “drug-like” and “nondrug-like” molecules? J Med Chem 1998;41 (18):3314–24.
[30] Available chemistry directory. San Diego, C.A.: Accelrys, Inc.
[31] Sadowski J, Kubinyi H. A scoring scheme for discriminating between drugs and nondrugs. J Med Chem 1998;41(18):3325–8.
[32] Byvatov E, Fechner U, Sadowski J, Schneider G. Comparison of support vector machine and artificial neural network systems for drug/ nondrug classification. J Chem Inf Comput Sci 2003;43(6):1882–9.
[33] Wagener M, van Geerestein VJ. Potential drugs and nondrugs: prediction and identification of important structural features. J Chem Inf Comput Sci 2000;40(2):280–92.
[34] Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. J Comb Chem 1999;1(1):55–68.
[35] Frimurer T, Bywater R, Naerum L, Nörskov-Lauritsen L, Brunak S. Improving the odds in discriminating “drug-like” from “non drug-like” compounds. J Chem Inf Comput Sci 2000;40(6):1315–24.
[36] Xu J, Stevenson J. Drug-like index: a new approach to measure drug-like compounds and their diversity. J Chem Inf Comput Sci 2000;40 (5):1177–87.
[37] Good AC, Hermsmeier MA. Measuring CAMD technique performance. 2. How “druglike” are drugs? Implications of random test set selection exemplified using druglikeness classification models. J Chem Inf Model 2007;47(1):110–14.

[38] Hutter MC. Separating drugs from nondrugs: a statistical approach using atom pair distributions. J Chem Inf Model 2007;47(1):186–94.
[39] Ursu O, Oprea TI. Model-free drug-likeness from fragments. J Chem Inf Model 2010;50(8):1387–94.
[40] García-Sosa AT, Oja M, Hetényi C, Maran U. DrugLogit: logistic discrimination between drugs and nondrugs including disease-specificity by assigning probabilities based on molecular properties. J Chem Inf Model 2012;52(8):2165–280.
[41] Ajay A, Murcko A. Designing libraries with CNS activity. J Med Chem 1999;42(24):4942–51.
[42] Ursu O, Rayan A, Goldblum A, Oprea TI. Understanding drug-likeness. Comp Mol Sci 2011;1(5):760–81.
[43] Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. Nat Chem 2012;4(2):90–8.
[44] Druker BJ. STI571 (Gleevec) as a paradigm for cancer therapy. Trends Mol Med 2002;8(4 Sup):S14–18.
[45] Kalgutkar AS, Crews BC, Rowlinson SW, Marnett AB, Kozak KR, Remmel RP, et al. Biochemically based design of cyclooxygenase-2 (COX-2) inhibitors: facile conversion of nonsteroidal antiinflammatory drugs to potent and highly selective COX-2 inhibitors. Proc Natl Acad Sci USA 2000;97(2):925–30.
[46] Mencher SK, Wang L. Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). BMC Clin Pharmacol 2005;5:3.
[47] Stephenson VC, Heyding RA, Weaver DF. The "promiscuous drug concept" with applications to Alzheimer's disease. FEBS Lett 2005;579 (6):1338–42.
[48] Roth BL, Sheffler DJ, Kroeze WK. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat Rev Drug Discov 2004;3(4):353–9.
[49] Hampton T. "Promiscuous" anticancer drugs that hit multiple targets may thwart resistance. J Am Med Assoc 2004;292(4):419–22.
[50] Azzaoui K, Hamon J, Faller B, Whitebread S, Jacoby E, Bender A, et al. Modeling promiscuity based on in vitro safety pharmacology profiling data. Chem Med Chem 2007;2(6):874–80.
[51] Roth BL, Driscol J. PDSP Ki database. University of North Carolina at Chapel Hill and the United States National Institute of Mental Health, <http://pdsp.med.unc.edu/pdsp.php>.
[52] Krejsa CM, Horvath D, Rogalski SL, Penzotti JE, Mao B, Barbosa F, et al. Predicting ADME properties and side effects: the BioPrint approach. Curr Opin Drug Discov Devel 2003;6(4):470–80.
[53] Hopkins AL, Mason JS, Overington JP. Can we rationally design promiscuous drugs? Curr Opin Struct Biol 2006;16(1):127–36.
[54] Yang Y, Chen H, Nilsson I, Muresan S, Engkvist O. Investigation of the relationship between topology and selectivity for druglike molecules. J Med Chem 2010;53(21):7709–14.
[55] cLogP. Version 4.3. Claremont, CA: BioByte Corp.
[56] Peters JU, Hert J, Bissantz C, Hillebrecht A, Gerebtzoff G, Bendels S, et al. Can we discover pharmacological promiscuity early in the drug discovery process? Drug Discov Today 2012;17(7–8):325–35.
[57] Norinder U, Bergstrom CA. Prediction of ADMET properties. Chem Med Chem 2006;1(9):920–37.
[58] Hou TJ, Wang J, Zhang W, Wang W, Xu X. Recent advances in computational prediction of drug absorption and permeability in drug discovery. Curr Med Chem 2006;13(22):2653–67.
[59] Chohan KK, Paine SW, Waters NJ. Quantitative structure activity relationships in drug metabolism. Curr Top Med Chem 2006;6 (15):1569–78.
[60] de Groot M. Designing better drugs: predicting cytochrome P450 metabolism. Drug Discov Today 2006;11(13–14):601–6.
[61] Hansch D, Leo A, Mekapati SB, Kurup A. QSAR and ADME. Bioorg Med Chem 2004;12(12):3391–400.
[62] Lombardo F, Gifford F, Shalaeva MY. In-silico ADME prediction: data, models, facts, and myths. Mini Rev Med Chem 2003;3(8):861–75.
[63] Gleeson MP. Generation of a set of simple, interpretable ADMET rules of thumb. J Med Chem 2008;51(4):817–34.
[64] Waring MJ. Defining optimum lipophilicity and molecular weight ranges for drug candidates molecular weight dependent lower log D limits based on permeability. Bioorg Med Chem Lett 2009;19(10):2844–51.
[65] Hughes JD, Blagg J, Price DA, Bailey S, DeCrescenzo GA, Devraj RV, et al. Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorg Med Chem Lett 2008;18(17):4872–5.
[66] Waring MJ, Johnstone C. A quantitative assessment of hERG liability as a function of lipophilicity. Bioorg Med Chem Lett 2007;17 (6):1759–64.
[67] Yan A, Gasteiger J. Prediction of aqueous solubility of organic compounds based on a 3D structure representation. J Chem Inf Comput Sci 2003;43(2):429–34.
[68] Lovering F, Bikker J, Humblet C. Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 2009;52(21):6752–6.
[69] Yang Y, Engkvist O, Llinas A, Chen H. Beyond size, ionization state, and lipophilicity: influence of molecular topology on absorption, distribution, metabolism, excretion, and toxicity for druglike compounds. J Med Chem 2012;55(8):3667–77.
[70] Ritchie TJ, Macdonald SJ. The impact of aromatic ring count on compound developability: are too many aromatic rings a liability in drug design? Drug Discov Today 2009;14(21–22):1011–20.
[71] Ritchie TJ, Macdonald SJ, Young RJ, Pickett SD. The impact of aromatic ring count on compound developability: further insights by examining carbo- and hetero-aromatic and -aliphatic ring types. Drug Discov Today 2011;16(3–4):164–71.
[72] Delaney JS. ESOL: estimating aqueous solubility directly from molecular structure. J Chem Inf Comput Sci 2004;44(3):1000–5.
[73] Leeson PD, St Gallay SA, Wenlock MC. Impact of ion class and time on oral drug molecular properties. Med Chem Commun 2011;2:91–105.
[74] Roughley SD, Jordan AM. The medicinal chemist's toolbox: an analysis of reactions used in the pursuit of drug candidates. J Med Chem 2011;54(10):3451–79.
[75] Hill AP, Young RJ. Getting physical in drug discovery: a contemporary perspective on solubility and hydrophobicity. Drug Discov Today 2010;15(15–16):648–55.
[76] Young RJ, Green DVS, Luscombe CN, Hill AP. Getting physical in drug discovery II: the impact of chromatographic hydrophobicity measurements and aromaticity. Drug Discov Today 2011;16(17–18):822–30.
[77] Kenny PW, Montanari CA. Inflation of correlation in the pursuit of drug-likeness. J Comput Aided Mol Des 2013;27(1):1–13.

[78] Walters WP, Green J, Weiss JR, Murcko MA. What do medicinal chemists actually make? A 50 year retrospective. J Med Chem 2011;54(19):6405–16.
[79] Leeson PD, St Gallay SA. The influence of the "organizational factor" on compound quality in drug discovery. Nat Rev Drug Discov 2011;10(10):749–65.
[80] Keserü GM, Makara GM. The influence of lead discovery strategies on the properties of drug candidates. Nat Rev Drug Discov 2009;8(3):203–12.
[81] Hann MM, Leach AR, Harper G. Molecular complexity and its impact on the probability of finding leads for drug discovery. J Chem Inf Comput Sci 2001;41(3):856–64.
[82] Ferenczy GG, Keserü GM. Enthalpic efficiency of ligand binding. J Chem Inf Model 2010;50(9):1536–41.
[83] Carlson HA, Smith RD, Khazanov NA, Kirchhoff PD, Dunbar Jr JB, Benson ML. Differences between high- and low-affinity complexes of enzymes and nonenzymes. J Med Chem 2008;51(20):6432–41.
[84] Hopkins AL, Keserü GK, Leeson PD, Rees DC, Reynolds CH. The role of ligand efficiency metrics in drug discovery. Nat Rev Drug Discov 2014;13(2):105–21.
[85] Paolini GV, Shapland RH, van Hoorn WP, Mason JS, Hopkins AL. Global mapping of pharmacological space. Nat Biotechnol 2006;24(7):805–15.
[86] Blake JF. Examination of the computed molecular properties of compounds selected for clinical development. Biotechniques Suppl 2003;34(6):16–20.
[87] Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today 2004;9(10):430–1.
[88] Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. Proc Natl Acad Sci USA 1999;96(18):9997–10002.
[89] Andrews PR, Craik DJ, Martin JL. Functional group contributions to drug–receptor interactions. J Med Chem 1984;27(12):1648–57.
[90] Shultz MD. Setting expectations in molecular optimizations: strengths and limitations of commonly used composite parameters. Bioorg Med Chem Lett 2013;23(21):5980–91.
[91] Shultz MD. The thermodynamic basis for the use of lipophilic efficiency (LipE) in enthalpic optimizations. Bioorg Med Chem Lett 2013;23(21):5992–6000.
[92] Murray CW, Erlanson DA, Hopkins AL, Keserü GM, Leeson PD, Rees DC, et al. Validity of ligand efficiency metrics. ACS Med Chem Lett 2014;5(6):616–18.
[93] Kenny PW, Leitão A, Montanari A. Ligand efficiency metrics considered harmful. J Comput Aided Mol Des 2014;28(7):699–710.
[94] Reynolds CH, Bembenek SD, Tounge BA. The role of molecular size in ligand efficiency. Bioorg Med Chem Lett 2007;17(15):4258–61.
[95] Reynolds CH, Tounge BA, Bembenek SD. Ligand binding efficiency: trends, physical basis, and implications. J Med Chem 2008;51(8):2432–8.
[96] Abad-Zapatero C. Ligand efficiency indices for effective drug discovery. Expert Opin Drug Discov 2007;2:469–88.
[97] Mortenson PN, Murray CW. Assessing the lipophilicity of fragments and early hits. J Comput Aided Mol Des 2011;25(7):663–7.
[98] Verdonk ML, Rees DC. Group efficiency: a guideline for hits to leads chemistry. Chem Med Chem 2008;3(8):1179–80.
[99] Freire E. Do enthalpy and entropy distinguish first in class from best in class? Drug Discov Today 2008;13(19–20):869–74.
[100] Ferenczy GG, Keserü GM. Thermodynamics of ligand binding N.p. In: Luque J, Barril X, editors. Physico-chemical and computational approaches to drug discovery. Royal Society of Chemistry; 2012. p. 23–79.
[101] Olsson TSG, Williams MA, Pitt WR, Ladbury JE. The thermodynamics of protein–ligand interaction and solvation: insights for ligand design. J Mol Biol 2008;384(4):1002–17.
[102] Hann MM. Molecular obesity, potency, and other addictions in drug discovery. Med Chem Comm 2011;2:349–55.
[103] Loving K, Alberts I, Sherman W. Computational approaches for fragment-based and de novo design. Curr Top Med Chem 2010;10(1):14–32.
[104] Ladbury JE, Klebe G, Freire E. Adding calorimetric data to decision making in lead discovery: a hot tip. Nat Rev Drug Discov 2010;9(1):23–7.
[105] Freire E. A thermodynamic approach to the affinity optimization of drug candidates. Chem Biol Drug Des 2009;74(5):468–72.
[106] Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. Nucleic Acids Res 2007;35:D198–201.
[107] Reynolds CH, Holloway MK. Thermodynamics of ligand binding and efficiency. ACS Med Chem Lett 2011;2(6):433–7.
[108] Freeman-Cook KD, Hoffman RL, Johnson TW. Lipophilic efficiency: the most important efficiency metric in medicinal chemistry. Future Med Chem 2013;5(2):113–15.
[109] Drysdale MJ, Brough PA. Medicinal chemistry of Hsp90 inhibitors. Curr Top Med Chem 2008;8(10):859–68.
[110] Murray CW, Carr MG, Callaghan O, Chessari G, Congreve M, Cowan S, et al. Fragment-based drug discovery applied to Hsp90: discovery of two lead series with high ligand efficiency. J Med Chem 2010;53(16):5942–55.
[111] Woodhead AJ, Angove H, Carr MG, Chessari G, Congreve M, Coyle JE, et al. Discovery of (2,4 dihydroxy 5 isopropylphenyl)-[5-(4 methylpiperazin 1 ylmethyl)- 1,3 dihydrois oindol 2 yl]methanone (AT13387), a novel inhibitor of the molecular chaperone Hsp90 by fragment based drug design. J Med Chem 2010;53(16):5956–69.
[112] Ferenczy GG, Keserü GM. How are fragments optimized? A retrospective analysis of 145 fragment optimizations. J Med Chem 2013;56(6):2478–86.
[113] Jhoti H, Williams G, Rees DC, Murray CW. The "rule of three" for fragment-based drug discovery: where are we now? Nat Rev Drug Discov 2013;12(8):644–5.
[114] Tarcsay A, Nyiri K, Keserü GM. Impact of lipophilic efficiency on compound quality. J Med Chem 2012;55(3):1252–60.
[115] Nissink JWM. Simple size-independent measure of ligand efficiency. J Chem Inf Model 2009;49(6):1617–22.
[116] Giordanetto F, Kihlberg J. Macrocyclic drugs and clinical candidates: what can medicinal chemists learn from their properties? J Med Chem 2013;57(2):278–95.

CHAPTER

16

Pharmacological Space

Andrew L. Hopkins

College of Life Sciences, University of Dundee, Dundee, Scotland, UK

OUTLINE

When you know you've come up with something completely novel, the medical imperative is to come up with a good use for it. That takes imagination. And the final step is to persuade the patents people that something has actually been discovered. ***Paul Janssen***

I. WHAT IS PHARMACOLOGICAL SPACE?

The medicinal chemist is faced with an apparently endless series of choices in undertaking a drug-discovery project, such as which disease to attempt to treat, which proteins to target, which assays to employ, and which compounds to synthesize and in what order. Sir James Black elegantly outlined six critical requirements for decision-making in a drug-discovery project (Black's Rules, Box 16.1). The concept of "pharmacological space" provides a theoretical framework for navigating the apparently infinite number of choices in drug discovery. Pharmacological space attempts to chart the limits of chemical space, targets space, and disease space in order to reduce and systematize the search for new drugs in these spaces. Thus, charting pharmacological space is an attempt to outline the areas where opportunities for new drugs may lie based on an extrapolation of the knowledge we have gained from our experiments to date. In this chapter, we shall outline some of the recent theoretical arguments and empirical evidence for navigating target space and chemical space for drug discovery.

II. CHEMICAL SPACE

At the invention of modern drug discovery, Paul Ehrlich screened just over 600 synthetic compounds to discover arsphenamine (Salvarsan) [2], a novel treatment for syphilis. With advances in screening technology, researchers can now routinely test millions of compounds in protein-based bioassays. Yet even the compound files of the largest pharmaceutical companies—which typically contain 10 [3] compounds—offer only a cursory examination of all the possible organic compounds that comprise "chemical space," which even up to a limit of 500 Da molecular weight (MW) per molecules is, for all practical purposes, infinite and limited only by the

BOX 16.1

BLACK'S RULES

Sir James Black's (Nobel Laureate in Physiology and Medicine, 1988) requirements for a drug-discovery project:[1]

1. Is the project purged of wishful thinking?
2. Is a chemical starting point identified?
3. Are relevant bioassays available?
4. Will it be possible to confirm laboratory-specificity in humans?
5. Is a clinical condition relevant to the specificity mentioned in point four?
6. Does the project have a champion—someone with the necessary passion, conviction, and energy?

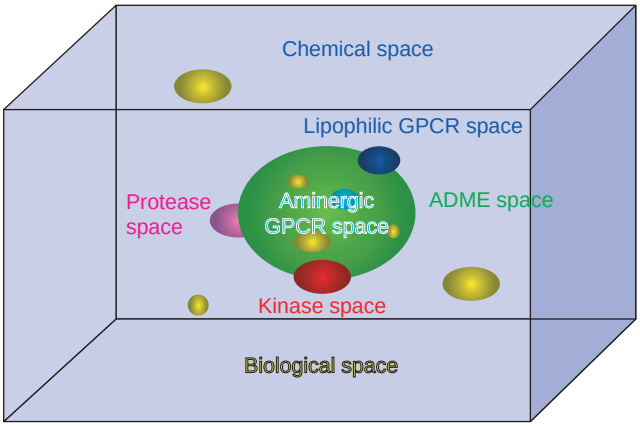


FIGURE 16.1 Graphical representation of property space have been observed for compounds active against individual genes families [7–9] and for bio-availability [3]. The independent intersection of compounds with drug-like properties or "ADME space" (absorption, distribution, metabolism, and excretion) is shown in green.

chemist's imagination. Yet the medicinal chemist's goal is not to explore chemical space for its own sake but to identify the small, discrete islands of compounds that interact with biological systems from the vast ocean of possible chemical structures. Measured in terms of physicochemical properties and topological descriptors, therapeutically useful compounds—that is, drugs—appear to form clusters in chemical space rather than be evenly distributed. The primary explanation for this is that historically about eight out of ten approved drugs are incremental improvements on existing classes of drugs and therefore often possess similar chemical structures. On average, of the thirty drugs approved each year over the past two decades only five new proteins are targeted by new drugs each year—of which fewer than two targets belong to previously undrugged target classes or protein families [4]. A similar discrete phenomenon is also observed in the chemical scaffolds of known drugs. While chemical space may be apparently limitless, the repertoire of molecular framework—the scaffolds of atom connectivity in a molecule—is extremely low. Only thirty-two molecular frameworks represent the chemical structures of half of all drugs [5], while 73 percent of all side-chains in drugs attached to the molecular frameworks are represented by only twenty side-chain groups [6].

Our experience from a century of medicinal chemistry and thousands of high-throughput screening (HTS) programs suggests that compounds that bind to certain "target classes"—that is, proteins from the same protein family, such as G-protein-coupled receptors (GPCRs)—are clustered together in discrete regions of a chemical space that can be defined by particular chemical descriptors. Figure 16.1 depicts a cartoon representation of the relationship between the continuum of chemical space (light blue) and the discrete areas of chemical space that are occupied by compounds with specific affinity for biological molecules, such as those from major gene families (shown in yellow, with specific gene families color-coded as proteases (purple), lipophilic GPCRs (blue) and kinases (red)), in terms of molecular property descriptors.

A. Drug-Like Space

A further restriction on chemical space is that a compound not only has to be biologically active but also contain the desired physicochemical properties to be administered as a drug. A biologically active compound may be too lipophilic to be orally absorbed, be too polar to cross the gastrointestinal wall, or have too much vulnerable chemistry functionality that is attacked by liver metabolizing systems and so does not exist intact long enough to generate a useful *in vivo* biological effect. Observations of the characteristics of compounds that are far more likely to yield safe, orally bioavailable medicines has led to the concept of "drug-likeness" to describe compounds that have the potential to be developed into orally administered drugs, which are generally favored owing to their ease of use by patients. Essential to the design of a drug are the physicochemical characteristics of the lead compound. A balance of solubility and polar/hydrophobic properties is crucial for specific routes of absorption, membrane permeabilities, and other biological barriers that a drug needs to penetrate to reach the desired site of action, in order to affect the biological equilibrium of a whole organism. The presence of such biological barriers limits the range of molecular properties and thus the chemical space within which medicinal chemists can design.

The distribution of molecular properties of small-molecule, launched drugs has changed little in the past twenty years, despite changes in the range of indications and targets [7] and despite changes in the underlying properties of investigational compounds (see Box 16.2). Lipinski's seminal analysis of the Derwent World Drug Index introduced the concept of drug-likeness: orally administered drugs are far more likely to reside in areas of chemical space defined by a limited range of molecular properties. These properties have been encapsulated in Lipinski's "rule-of-five." [3] The analysis by Lipinski et al shows that, historically, 90 percent of orally absorbed drugs have had fewer than five hydrogen-bond donors, fewer than ten hydrogen-bond acceptors, MW of less than 500 Da, and log *P* values (a measure of lipophilicity) of less than five [3]. Since this work, various definitions of and methods to predict drug-likeness have been proposed in the literature [17–28], but the consensus is that drug-likeness is defined by a range of molecular properties and descriptors that can discriminate between drugs and non-drugs for such characteristics as oral absorption, aqueous solubility, and permeability—what could be considered ADME space. Computational property filters can be used to assess rapidly the drug-likeness of chemical libraries *in silico* before purchase or synthesis.

In recent times, toxicity has replaced poor drug metabolism properties as the major cause of failure in the early clinical phase of drug development. Toxicity may derive from modulation of the biological target or from the compound's structure or physicochemical properties. Extrapolating from large-scale gene knock-out studies in model organisms, approximately a fifth of human proteins are thought to be essential, at least in embryo development. Despite this, only a handful of targets are considered routinely as general "anti-targets," such as the

BOX 16.2

CHANGES IN MEDICINAL COMPOUND PROPERTIES

Over the past twenty-five years, there has been a steady, inexorable rise in the median molecular weight of reported medicinal chemistry compounds [7,10] (Figure 16.2a). Comparing five-year averages from 1986–1990 to those of 1999–2003, the median MW of all reported medicinal chemistry compounds in the literature rose 68 Da (20 percent), from 354 to 422 Da, respectively. Interestingly, this growth is also reflected in the increase of the median MW of disclosed ligands for several gene families. For example, compounds binding to aminergic GPCRs have increased in MW by around 56 Da, from 337 to 393 Da between the two five-year periods. No significant increase in mean or median potency is observed in the data to explain the increase in MW, although it is suspected that with the rise of the availability of recombinant proteins assay, there is been an increase in selectivity driving this increase. The changing target portfolio of the industry is also a significant driver in the increase in MW of new lead compounds and investigational drugs. Even so, this rise in MW contrasts with the much slower rise in the mean MW of approved drugs of only 2.2 Da per launch year over the same period. In contrast, there is a steady decline in MW through each subsequent stage of clinical development and increase in the proportion of compounds that are "rule-of-five" compliant [7,11,12,14–16] (Figure 16.2b).

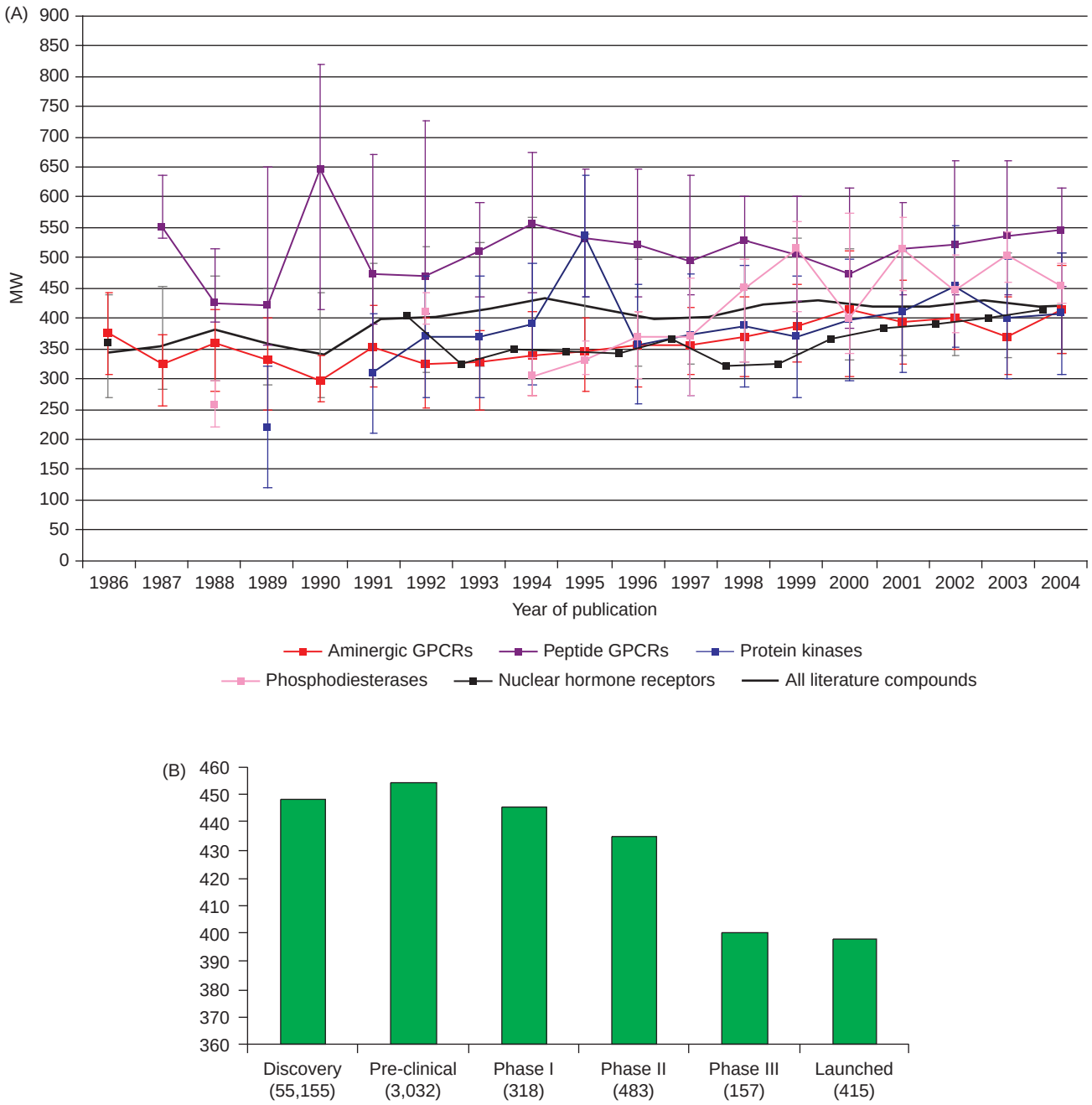


FIGURE 16.2 (a) Increase in MW over time of published medicinal compounds. (b) Mean MW through clinical development (Source: Investigational Drugs Database). Similar results found by Blake [11] and Wenlock et al [12]. The increase in MW between discovery compounds (leads) and preclinical compounds (candidates) has been studied in detail by Oprea et al [13].

HERG ion channel, the binding of which should be avoided. In some circumstance, inhibition of a vital metabolic process may be beneficial in specific tissue types (such as statins predominately targeting liver HMG-CoA reductase) or in specific circumstances where disease-specific differentiation is possible (such as bortezomib inhibition of the proteasome). The most common tactic medicinal chemists apply to decrease the chance of isosyncratic toxicity is avoiding compounds that contain specific chemical groups (toxicophores) that have been associated with toxic effects, and compounds that interact covalently with protein targets, which suffer from problems such as lack of specificity and unsuitability for optimization by medicinal chemistry techniques [29–31]. Although several well-known drugs, such as omeprazole and β-lactamase inhibitors, are known to act *via* irreversible mechanisms, medicinal chemists and toxicologists are becoming more wary of incorporating reactive groups

within tools or drugs that can form covalent bonds to the target or other proteins. Recent studies have also highlighted the relationship between high lipophilicity and the increased chance of toxicity *via* the increased target promiscuity of lipophilic molecules [10]. To avoid promiscuous off-target effects, Leeson et al [10] recommend medicinal chemists should be wary of synthesizing compounds with *c* Log $P > 3.5$ [10].

III. TARGET SPACE

Drugs act by binding to and modifying the function of biological macromolecules, predominately proteins but also DNA, RNA, carbohydrates, and phospholipids membranes. Drug target space, unlike chemical space, is limited and discrete. With the availability of genome sequences, drug target space is capable of being fully defined [4,32–37]. Of the 1,357 unique drugs approved by the US Food and Drug Administration (FDA), 1,204 are classed as small-molecule drugs and 166 are classed as biological drugs. Of the 1,204 small-molecule drugs, 803 can be administered orally, 421 can be dosed parenterally, and 275 can be used as topical agents (including buccal, rectal, inhalational, and other such routes of administration for "topical" agents) [4]. At least 192 (16 percent) of small-molecule drugs are administered as prodrugs. Of the 1,204 small-molecule drugs, 885 pass the "rule-of-five" test; of these, 619 (70 percent) are dosed orally, whereas 159 (20 percent) of orally dosed drugs fail at least one of the rule-of-five parameters [4]. For all the current drugs for which we can determine the mode of action, 324 distinct molecular drug targets have been identified. Of these, 266 are human proteins, and the remainder are present in bacterial, viral, fungal, or other pathogenic organism targets. Small-molecule drugs modulate 248 proteins, of which 207 are targets encoded by the human genome. Oral small-molecule drugs target 227 molecular targets, of which 186 are human targets [4] (Table 16.1).

The pharmacopoeia of molecular drug targets represents approximately only 1 percent of the total number of predicted genes in the human genome. Medicinal chemistry has, however, explored a greater number of targets for which chemical tools, lead compounds, and investigational drugs have been discovered. Paolini et al have used large-scale integration of proprietary and published screening data to attempt to identify the number of unique molecular targets for which chemical tools, leads, or drugs have been discovered [7]. The global survey of the data from Pfizer, Warner-Lambert, and Pharmacia, integrated with a large body of medicinal chemistry SAR results published in the literature, identified over 1,300 proteins from fifty-five organisms with biologically active chemical matter. These include a nonredundant list of 836 genes in the human genome for which small-molecule chemical tools have been discovered, of which 727 human targets have at least one compound with binding affinity below 10 μM compliant with Lipinski's "rule-of-five" criteria for oral drug absorption [3], and 529 human targets have at least one "rule-of-five" compound below 100 nM (Table 16.2).

A. Druggability

Development of the ideas of drug-likeness of compounds has led to the parallel concept of druggability of proteins. The druggability hypothesis proposes that the likelihood of discovering a drug-like ligand for a protein can be assessed *a priori* derives from the biophysical basis of molecular recognition [34,38–40]. In short, the binding sites on biological molecules must be complementary in terms of volume, topology, and physicochemical

TABLE 16.1 Molecular Targets of FDA Approved Drugs (2006)

Class of drug target	Species	Number of molecular targets
Targets of approved drugs	Pathogen and human	324
Human genome targets of approved drugs	Human	266
Targets of approved small-molecule drugs	Pathogen and human	248
Targets of approved small-molecule drugs	Human	207
Targets of approved oral small-molecule drugs	Pathogen and human	227
Targets of approved oral small-molecule drugs	Human	186
Targets of approved therapeutic antibodies	Human	15
Targets of approved biologicals	Pathogen and human	76

TABLE 16.2 Human Proteins with Identified Chemical Tools, by Gene Family, Lipinski Rule-of-Five Properties and Potency [7]

Gene taxonomy	All targets at, 10 μM[a]	Human targets at, 10 μM	Human targets at, 10 μM	Human targets at, 10 μM Ro5[b] $n>1$	Human targets at, 10 μM	Human targets at, 100 nM Ro5 $n>1$
Protein kinases	131	105	99	98	86	83
Peptide GPCRs	10	63	59	59	55	42
Transferases	75	49	42	36	33	24
Aminergic GPCRs	72	35	35	35	35	35
GPCRs class A – others	68	44	44	40	38	32
Oxidoreductases	68	40	36	38	29	25
Metalloproteases	63	44	41	41	36	35
Hydrolases	56	36	29	30	25	21
Ion channels – ligand gated	55	29	28	24	25	22
Nuclear hormone receptors	47	24	24	22	23	19
Serine proteases	37	30	30	28	29	21
Ion channels – others	24	18	16	16	13	11
PDEs	23	19	19	19	18	18
Cysteine proteases	20	16	16	14	14	13
GPCRs class C	20	10	10	10	6	6
Kinases – others	16	12	9	11	6	5
GPCRs class B	14	7	7	4	7	3
Aspartyl proteases	10	7	7	4	6	4
Others	241	139	119	108	83	63
Enzymes – others	156	109	97	90	69	47
Total	**1,306**	**836**	**767**	**727**	**639**	**529**

[a]*Molecular targets for 1,357 unique FDA approved drugs (including New Chemical Entities and New Biological Entities) as derived by Overington et al [4] from a normalized database analysis of the FDA Orange Book. Following a comprehensive analysis of the literature unique molecular targets could be assigned to 1,065 FDA drugs as the mechanisms of action.*

[b]*Human proteins with known small-molecule chemical tools, lead or drugs as determined by Paolini et al. [7] survey of the medicinal chemistry literature and corporate databases. Biologically active compounds are defined as this with a binding affinity below 10mM against the molecular target. Lipinski's "rule-of-five" criteria [3] of fewer than 5 H-bond donors, fewer than 10 H-bond acceptors, MW below 0.5 kDa and clog P below 5. Compounds that fail Lipinski's criteria are more likely to show poor absorption or permeation because such compounds are unlikely to show good oral bioavailability. Over 761 proteins have more than one compound reported active.*

properties with their ligands. Then only certain binding sites on putative drug targets are compatible with the high affinity binding of compounds with drug-like properties [34]. The binding energy (ΔG) of a ligand to a molecular target such as a protein, RNA, DNA, or carbohydrate is defined as $\Delta G = -RT \ln K_d$. The van der Waals and entropy components are considered to be the predominate contributors to the binding energy by the burying of hydrophobic surfaces and the liberation of ordered waters. A low affinity ligand of $K_i = 1$ μM affinity equates to a binding energy of $\Delta G = -8.4$ kcal/mol. A high affinity drug molecule binding with an affinity of $K_i = 10$ nM requires a binding energy (ΔG) of -11 kcal/mol. Thus, 1.36 kcal/mol of binding energy is equivalent to a 10-fold increase in potency. The binding energy potential of a ligand is approximately proportional to the available surface area and its properties, assuming there are no strong covalent or ionic interactions between the ligand and the protein. For small-molecule (less than 500 Da) drug-like molecules, a near linear correlation between molecular surface area and MW has been observed. The van der Waals attractions between atoms and the hydrophobic effect from the displacement of water contributes approximately 0.03 kcal/mol/Å^2 [41]. Thus, assuming there are no strong ionic interactions between the protein and the ligand, a ligand with a 10 nM dissociation constant would be required to bury 370Å [4] of the hydrophobic surface area. The contribution of the hydrophobic surface to binding energy is demonstrated by the medicinal chemistry phenomenon of the "magic methyl," where a single methyl group placed in the correct position can increase ligand affinity by 10-fold. The

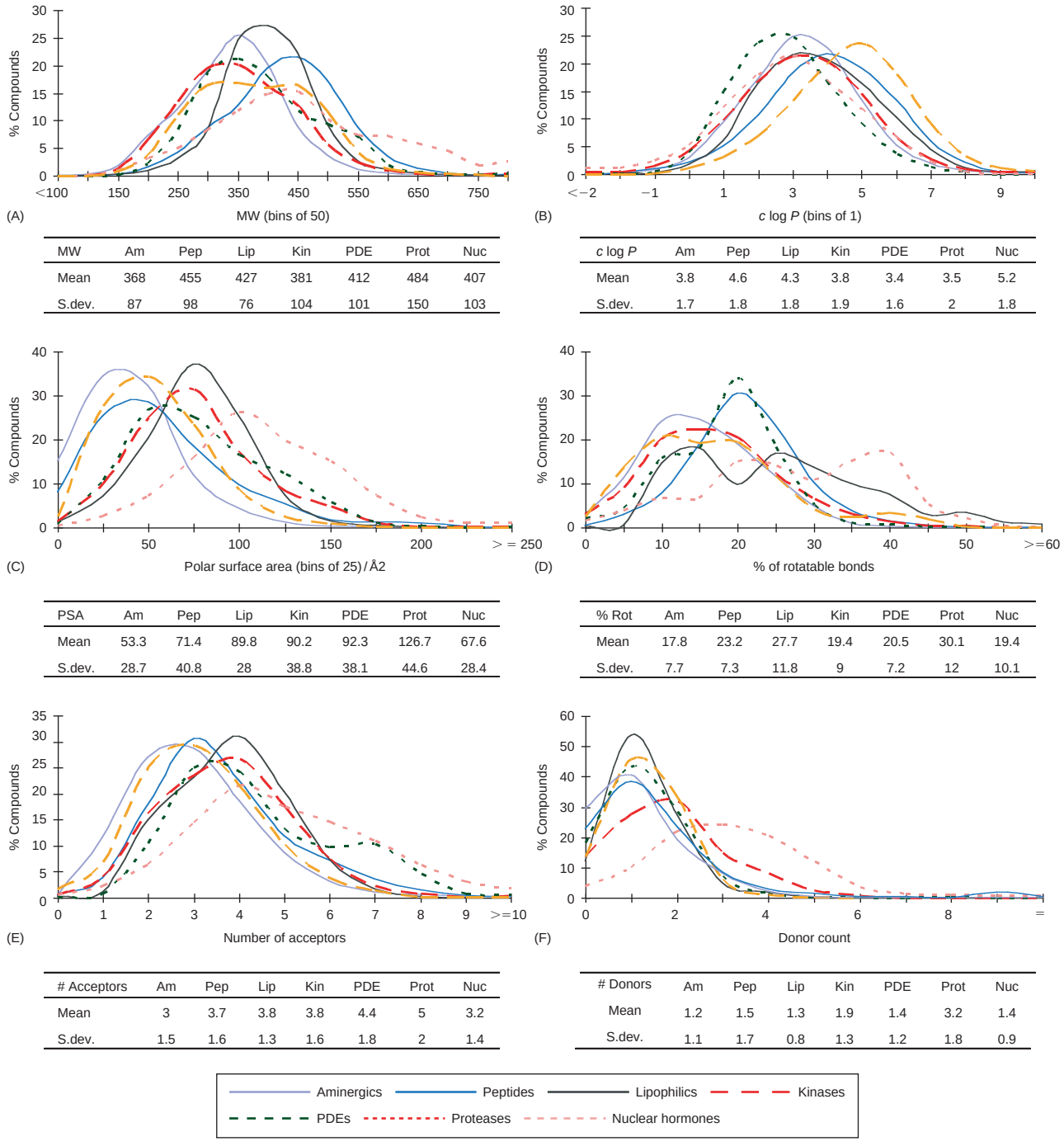
accessible hydrophobic surface area of a methyl group is approximately 46 $Å^2$ (if one assumes all of the hydrophobic surface area is encapsulated by the protein binding site and thus makes full contact with the protein) with a hydrophobic effect of 0.03 kcal/mol/$Å^2$, approximately equal to 1.36 kcal/mol, which is equivalent to the observed 10-fold affinity increase (approximately the maximal affinity per non-hydrogen atom) [42]. In addition to the predominantly hydrophobic contribution to the binding of many drugs, ionic interactions, such as those found in zinc proteases (such as angiotensin converting enzyme (ACE) inhibitors) contribute to the binding energy. The attraction of complementary polar groups contributes up to 0.1 kcal/mol/$Å^2$, with ionic salt bridge approximately three times greater, thereby allowing low MW compounds to bind strongly. Unlike hydrophobic interactions, complementary polar interactions are dependent on the correct geometry.

B. Structure-Based Druggability

The physicochemical and energetic constraints of molecular recognition leads to the conclusion that a drug target needs a "pocket," whether the pocket is predefined by the protein's architecture or formed on binding by allosteric mechanisms. Druggable cavities on proteins that are complementary with the high affinity binding of noncovalent, small-molecule, "rule-of-five" compliant ligands (whose binding energy is predominantly driven the entropic, hydrophobic, and van der Waals contributions) are predominately apolar cavities of 400−1,000 $Å^3$, where over 65 percent of the pocket is buried or encapsulated, with an accessible hydrophobic surface area of at least 350 $Å^2$ [38]. Encapsulated cavities maximize the ratio of the surface area to the volume and are thus capable of binding low MW compounds with high affinities.

The hypothesis that the physicochemical properties of cavities on protein structures can be analyzed *a priori* to predict the druggability of a protein has been developed further into automatic algorithms to assess the protein structures in the Protein Data Bank (PDB) and the stream of novel structures determined by the structural genomics initiative. Empirical druggability predictions have been explored experimentally using heteronuclear NMR to identify and characterize the binding surfaces on protein at Abbott by screening ~10,000 low MW molecules (average MW 220, average *c* Log *P* 1.5) [39]. In a small sample of thirty-three proteins, the screening results reveal that about 90 percent of the ligands bind to sites known to be small-molecule ligand binding sites. Only three out of the twenty-three proteins were distinct uncompetitive new binding sites. In the relatively small sample of proteins studied, Hajduk et al note a high correlation between experimental NMR hit rates and the ability to find high affinity ligands. From the experimental screening hit rates, Hajduk et al constructed a simple model that included physicochemical property descriptors such as cavity dimensions, surface complexity, and polar and apolar surface area that accurately predicts the experimental screening hit rates with an R^2 of 0.72, an adjusted R^2 of 0.65.

A decision tree approach to assessing the druggability of protein structure has been developed by Al-Lazikani and Overington [43]. A range of physicochemical properties of the identified binding sites and cavities were calculated from the protein structures including volume, depth, curvature, accessibility, hydrophobic surface area and polar surface area. The algorithm was trained set against a test set of 400 protein complexes binding small-molecule, "rule-of-five" compliant ligands. From this analysis, a decision tree was derived to predict the druggability of a binding site or cavity from calculated physicochemical properties. The decision tree predicts whether a cavity is druggable within the statistical confidence of the tree. A success rate of 91 percent when predicting druggability on the protein drug targets has been claimed for this approach [43]. The method requires either an experimentally derived structure or a high-quality homology model. Ideally, because of the inherent flexibility of many protein−ligand binding sites, a sample of multiple conformations is preferred. The decision tree method was applied to the entire PDB (December 2004 release). Following a cleanup process, 27,409 files were suitable for analysis, further classified into 76,322 structural domains using SCOP [44], of which 28 percent (21,522) were found to have at least one site predicated to have some degree of druggability. From this analysis, a nonredundant set of 427 human proteins were predicted to contain a druggable binding site, including 281 proteins having no prior known compounds or drugs developed against them. In a similar analysis, Hudjuk et al calculated the druggability of 1,000 nonredundant human proteins derived from the PDB, of which 35 percent of entries contain at least one site predicted to be highly druggable, which is slightly higher but comparable to Al-Lazikani's prediction. A third approach to structure-based druggability has been proposed by Cheng et al [40], who developed the topical and surface area arguments of druggability, as outlined above, into an automated algorithm for analyzing binding sites on protein structures. Assuming a hydrophobic binding energy of de-solvated flat surface inside a binding site, where a drug-like compound of 500 Da MW ligand is assumed



MW	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	368	455	427	381	412	484	407
S.dev.	87	98	76	104	101	150	103

c log *P*	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	3.8	4.6	4.3	3.8	3.4	3.5	5.2
S.dev.	1.7	1.8	1.8	1.9	1.6	2	1.8

PSA	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	53.3	71.4	89.8	90.2	92.3	126.7	67.6
S.dev.	28.7	40.8	28	38.8	38.1	44.6	28.4

% Rot	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	17.8	23.2	27.7	19.4	20.5	30.1	19.4
S.dev.	7.7	7.3	11.8	9	7.2	12	10.1

# Acceptors	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	3	3.7	3.8	3.8	4.4	5	3.2
S.dev.	1.5	1.6	1.3	1.6	1.8	2	1.4

# Donors	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	1.2	1.5	1.3	1.9	1.4	3.2	1.4
S.dev.	1.1	1.7	0.8	1.3	1.2	1.8	0.9

FIGURE 16.3 Physicochemical property distributions of approximately 75,000 biologically active ligands found in Pfizer's screening data as a function of target type. *Data courtesy of Barker, Snarey, Groom, and Hopkins.*

from observation to correlate 300 $Å^2$ solvent accessible surface area. A limitation of this method is that drugs that exploit strong ionic interactions (e.g. targeting ACE or HIV-1 integrase) are predicted to have a low druggability, due to the lack of hydrophobic surface area in the binding site.

C. Degrees of Druggability

Distinct differences are observed in the distribution of molecular properties between sets of compounds active against different gene families (Figure 16.3). The relationship between target class and the physicochemical

properties of ligands has been explored by calculating a set of physicochemical descriptors of hundreds of thousands of biologically active compounds across more than a thousand proteins where the protein sequences assigned to each of the pharmacological targets were classified into gene families. For example, ligands for the nuclear hormone receptors are significantly most lipophilic, as measured by *c* Log *P*, mirroring the properties of steroids. By comparison, the mean MW of ligands binding to aminergic GPCRs is close to the mean MW of approved drugs (383 Da, St. Dev. = 155 Da), while the mean MW of peptide GPCR ligands is greater and close to Lipinski's "rule-of-five" limit of 500 Da.

By linking predicted druggable targets to orthologs and homologs with known chemical matter in a structure–activity database, the likely physicochemical properties of potential ligands can be assessed. This premise is based on the assumption that protein targets that are closely related in sequence space are closely related in chemical space. Analysis of the diversity of the physicochemical properties of ligands for a protein family supports this general assumption. Prediction of the likely physicochemical properties of ligands for a novel drug targets can be used to assess their drug-like properties and attractiveness for drug-discovery programs. Lipinski's rule-of-five is commonly used as a metric to assess drug-like properties, but Lipinski's parameters (N + O < 10, log *P* <5, MW <500, H-bond donors <5) do not allow for a continuum of probabilities of compound properties to be assessed.

The ligand efficiency or binding energy per atom (Δg) of a compound can be calculated by converting the K_d into the free energy of binding (Equation 16.1) at 300 K and dividing by the number of "heavy" (i.e., nonhydrogen atoms) atoms (Equation 16.2):

Free energy of ligand binding:

$$\Delta G = -RT \ln K_d \tag{16.1}$$

Binding energy per atom (ligand efficiency):

$$\Delta g = \frac{\Delta G}{N_{\text{non-hydrogen atoms}}} \tag{16.2}$$

The logarithmic relationship between free energy of binding and dissociation constant potency means that every ΔG change of −1.36 kcal/mol results in a 10-fold change in potency. Kuntz et al surveyed the dissociation or IC_{50} values of ~150 ligand complexes and concluded that the maximum affinity per atom for organic compounds is −1.5 kcal/mol/nonhydrogen atom. The medicinal chemistry phenomenon of "magic methyls"—the addition of a single methyl group increasing potency by 10-fold—is explained in terms of the maximum achievable free energy for burying the surface area of a single "heavy" atom. The vast majority of medicinal chemistry compounds have efficiencies far below the observed maximal affinity per atom. A simple calculation can define the lowest limit of acceptable ligand efficiency in a typical pharmaceutical project where we wish to obtain a compound with a potency of 10 nM and an upper MW of 500:

- 500 MW compounds contain on average thirty-eight nonhydrogen atoms.
- 10 nM binding constant = 10.99 kcal/mol.
- Therefore, a 500 MW compound with a binding constant of 10 nM compound possesses a ligand efficiency of 0.29 kcal/mol/non-H atom.

Small differences in ligand efficiency (Δg) may have large consequences for the type of compounds that may be possible in a chemical series or against a particular target. For example:

- A compound with a $\Delta g = -0.27$ kcal/mol/non-H atom requires forty-one atoms (541 MW) to bind with $K_d = 10$ nM.
- A compound with a $\Delta g = -0.36$ kcal/mol/non-H atom requires only thirty atoms (405 MW) to bind with $K_d = 10$ nM.

Potency is an important criteria for assessing leads (or "hits" discovered in HTS). Potency alone, however, is often a false prophet. Indeed, the screening parameters, reagent concentrations, and false positive filters used make the detection of weak, low MW leads unlikely in many HTS. The bias of the HTS toward high MW compounds has often confounded further optimization as increase in potency often track increase in MW and result in compounds falling outside of the "rule-of-five" profile for acceptable absorption and permeability properties [5,20]. Scaffold and lead series selection could be aided by considering a parameter that "normalizes" the potency of a lead with respect to MW to allow comparisons between different series and scaffolds. Indeed, small compounds with low molecular complexity are predicted to have an improved probability of binding to the targets of interest [45].

TABLE 16.3 Distribution of Ligand Efficiencies Against GPCR Sub-Families

Gene family (ligand sub-family)	Mean ligand efficiency[a]	Mean efficiency of top 10%[b]	Mean MW of top 10%	Mean IC_{50} (nM) of top 10%
GPCR (metabotropic class)	0.40	0.48	281	1,384
GPCR (aminergic)	0.36	0.44	364	191
GPCR (proteins)	0.29	0.34	449	643
GPCR (small peptides)	0.28	0.33	486	99
GPCR (peptide)	0.28	0.32	471	424
GPCR (secretin class)	0.28	0.32	468	279
GPCR (amidated peptides)	0.28	0.31	490	291
GPCR (cyclic peptides)	0.28	0.29	467	365
GPCR (lipophilics)	0.27	0.33	437	323
GPCR (nucleotide)	0.25	0.32	536	10

[a]*kcal/mol/non-hydrogen atom.*
[b]*Mean of upper decile of compounds ranked by potency.*
Note: These data are derived from over 31,000 compounds. Each sub-family is represented by between 3 and 29 targets.

Some chemical series may exploit a specific binding site more successfully than others, yet the fundamental physiochemical nature of the target site defines the upper limits of the binder energy available. Thus a quantifiable value of the "druggability" of a particular target or gene family can be derived from the mean ligand efficiency of active compounds. An example of the use of ligand efficiency as a measure of the "degree of druggability" of a drug target is illustrated by analyzing ligands of the various subfamilies of GPCRs. Over 31,000 compounds with recorded activity and measurements against members of the GPCR class of targets were analyzed in Pfizer's chemogenomics database [7]. Within this set of compounds, we observe significant differences in the MW and ligand efficiencies against the distinct GPCR subfamilies (Table 16.3). Aminergic GPCR compounds tend to be significantly smaller and have higher ligand efficiencies than compounds that are active against peptide-binding GPCRs. GPCR subfamilies differ in terms of medicinal chemistry tractability. Peptide-binding GPCR targets with maximum ligand efficiencies are around 0.29 kcal/mol/non-H atom. This implies that one would need a compound with a MW of around 500 to achieve potency in the nanomolar range. Analysis across a range of targets suggests ligand efficiency is a useful metric for assessing target druggability [7] and attractive, efficient lead compounds.

D. Druggable Genome

The knowledge of the proteins against which medicinal chemistry has developed drugs, leads, and tool compounds can be used to infer the subset of the proteins expressed by the human genome that have a high probability of being "druggable," that is capable of binding drug-like small molecules with high affinity. The first systematic estimate of the number of druggable proteins—the "druggable genome"—following the publication of the draft human genome [46,47] was based on a search for membership of an extensive list of druggable gene families [34]. Gene family-based analysis assumes that the sequence and functional similarities underlie a conservation of binding site architecture between protein family members. Thus, the explicit assumption is that if one member of a gene family is modulated by a drug molecule, other members of the druggable protein domain family are likely to also be able to bind a compound with similar physicochemical properties. Thus, analysis based on druggable protein families or domains are likely to overestimate the number of druggable targets. Following the construction of a drug target sequence database of 399 targets of approved and experimental drugs and leads, 376 sequences could be assigned to 130 drug-binding domains as captured by their InterPro domain annotation. Of these, 130 InterPro domains have orthologs present in the human proteome. At the time of the initial draft of the human genome [46,47], 3,051 genes were identified as belonging to the 130 druggable protein domains and thus predicted to encoded proteins that are inferred to bind to drug-like molecules. Further refinements of the initial druggable

genome analysis have been published [48,49], reflecting how the number of predicted protein-expressing genes in the human genome has been modified since the initial draft. Orth et al estimate that there are 3,080 genes belonging the druggable genome, with over 2,950 druggable gene sequences in public databases in 2004 based on an estimate of the InterPro domain assignments of druggable gene families [49]. Russ and Lampel conducted an analysis of the 130 druggable protein domains using InterPro and PFAM on the final assembly of the human genome. [48,50] Overall, the PFAM protein domain annotation predicted fewer false positives than the InterPro classification used. When corrected for the overestimate of olfactory and taste GPCRs, the author again identify 3,050 druggable genes from the previously defined set of druggable protein domains [34], but with some significant changes within individual gene families. Using more stringent predictions for enzyme proteases and other subfamilies, a conservative estimate of approximately 2,200 druggable genes are identified [48].

In order to expand the homology analysis method for identifying which targets the survey include that most recent catalogs of known biological targets of drugs and lead compounds. Al-Lazikani and Overington have conducted the most extensive analysis to date on identifying the druggable genome, based on the homology to chemically tractable drug targets [43]. Using the BLAST sequence alignment algorithm to search each of the sequences against the human genome, over 900 distinct genes are identified that show close homology to 200 human proteins of approved small molecular drugs [4] at a cut-off of 30 percent sequence identity and *E*-value less than or equal to 10 [7]. Expanding the analysis to include human proteins from the small-molecule chemical leads as published in the medicinal chemistry journal data (i.e., J Med Chem 1980−2004, Bioorg Med Chem Letts 1990−2004), a set of 1,155 protein targets known with at least one drug or lead compound with a binding affinity below 10 kM, 707 of which are human molecular targets. BLAST sequence analysis of this database of medicinal chemistry literature [7,43] identified 2,921 protein sequences (within the same sequence identify cut-offs) that are predicted to be druggable proteins expressed by the human genome.

The distribution of the population size of gene families in the druggable genome follows a power law (Figure 16.4), with the top five target classes of GPCRs, protein kinases, proteases, transporters, and ion channels accounting for 54 percent of identified sequences in the druggable genome. The remaining 46 percent consists of a long tail of over 130 small-gene families and singleton enzymes. Thus, many of the new leads that appear in the literature are targeting new members of existing large gene families, such as protein kinases, and GPCRs do not necessarily increase the size of the predicted druggable genome. At present, novel potent chemical tools and lead compounds are reported in the peer-reviewed literature for 80−100 new molecular targets each year [7]. Undoubtedly, many more are disclosed only in patents. The increase in the rate of discovery of chemical tools for new targets doubled from an average of thirty new targets with leads being disclosed in the 1980s to an average of sixty new targets per year in the 1990s. By comparison, an average of four first-in-class drugs targeting novel targets have reached the market each year during the 1990s [7]. The new leads that target new families and proteins only increase the size of the druggable genome in an incremental fashion, as most gene families are small. For example, the discovery that imiquimod acts *via* its agonist binding to toll-like receptor 7 [51] suggests that the other members of the TLR gene family may be druggable with small-molecule compounds, yet this new, potentially druggable gene family introduced only ten new members to the druggable genome.

IV. CONCLUSIONS

Sir James Black once remarked that "the most fruitful basis for the discovery of a new drug is to start with an old drug." Indeed, approximately 80 percent of new drugs approved for marketing over the past decade by the FDA have been incremental improvements upon existing chemical designs [4]. While incremental innovations have been challenged by critics of the pharmaceutical industry, it must be remembered that no two drugs have exactly the same biological profile. A range of drugs in the same class expands the treatment options available to clinicians, as patients often respond differently to different drugs. Incremental innovations, however, starkly illustrate how important *a priori* knowledge of drug targets and chemical structure are to the medicinal chemist in designing new drugs. However, *a priori* information can be used in the hunt for innovative new medicines and new classes that can be used in all stages of medicinal chemistry design, from the selection of targets to the identification of lead chemical matter, and in the choice of chemical design-modification and substituent replacement [52,53]. Analysis of the properties of drugs reveals, despite changes in chemical structures and the targets of drugs, the overall distribution of their physical properties has changed little over the past twenty-five years, leading to concepts such as drug-likeness and the "rule-of-five." [7,10,26] Knowledge of the shared experience of

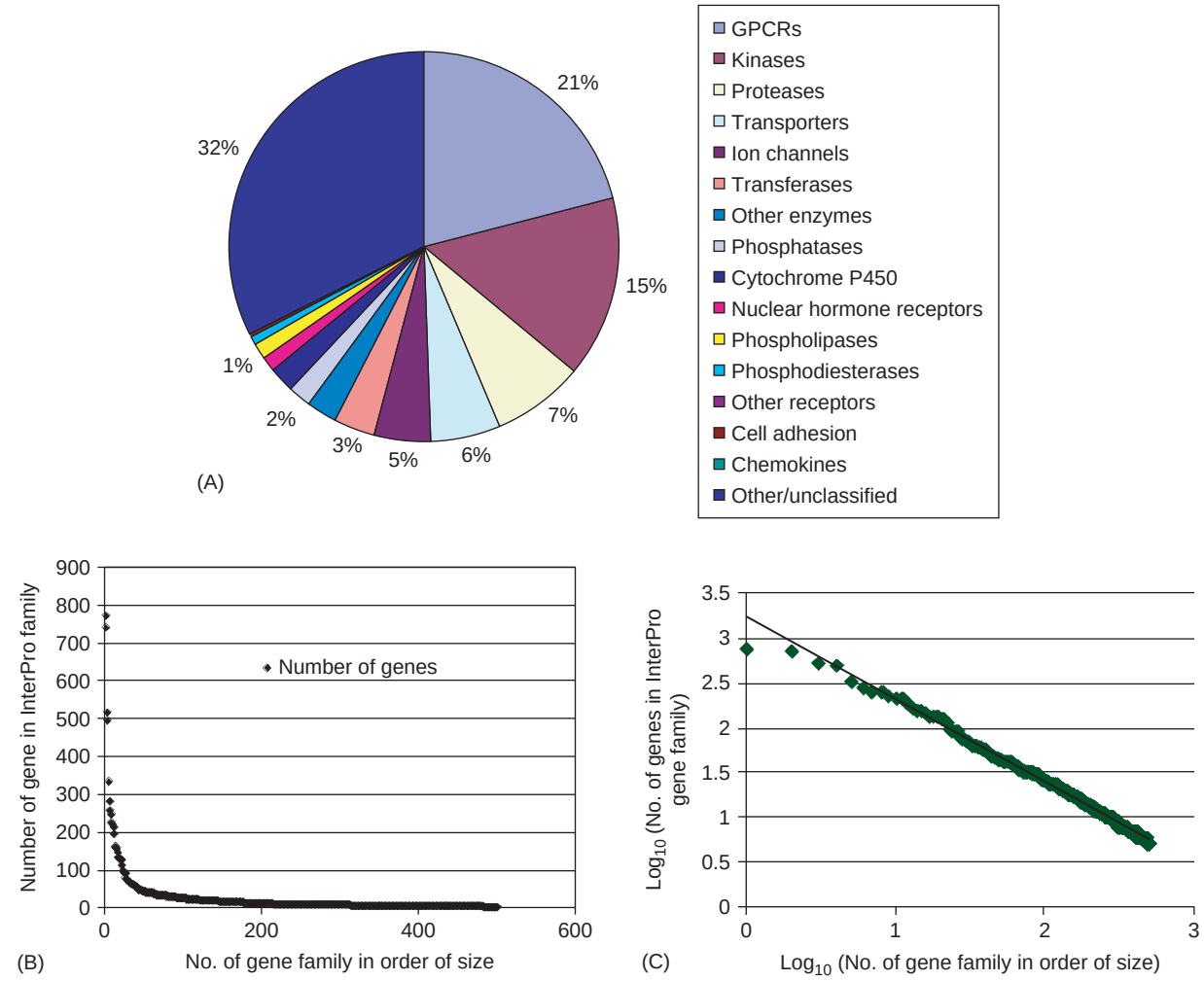


FIGURE 16.4 (a) Gene family distributions of human druggable genome [43]. (b) Population distribution of gene families size in the human genome as represented by a plot of top 500 InterPro families (http://www.ensembl.org/Homo_sapiens/interpro/IPtop500.html). (c) Distribution of gene families size in the human genome as represented as a logarithm scale to illustrate the power of law distribution of gene family population size.

compound library screening and the elucidation of X-ray structures of protein–ligand complexes led to the hypothesis that the druggability of a protein is a function of the physicochemical properties and topology of a binding site. While concepts like drug-likeness and druggability are often criticized for limiting the options of the drug hunter, these concepts have been derived from empirical observations and fundamental physical principles of molecular recognition. Given the fact that chemistry space is practically infinite, drug-discovery resources are finite, and cost and time to market equate to prolonged patient suffering, medicinal chemistry is wise to learn from its previous endeavors to chart areas of chemical space that may be rich in pharmacological compounds.

We started this chapter with a quote from one of the most successful drug discoverers, Dr. Paul Janssen. In many ways, Janssen's drug-discovery strategy demonstrated the power of using knowledge of pharmacological space in the search from drugs. Janssen produced a large number of new medicines by the careful exploration of the pharmacological space and medicinal applicability around a narrow seam of chemical space centered on piperidine derivatives [54]. Within this relatively narrow area of chemical space, Janssen et al explored the wider range of molecular targets and disease indications. This successful strategy of exploring pharmacological space in its widest sense is in contrast to drug discovery as it is usually conducted today, which focuses on the narrow spectrum of a single disease or molecular target but screens a wide sample of chemical space.

Acknowledgments

I would like to thank Gaia Paolini, Colin Groom, Carolyn Barker, Mike Snarey, Bissan Al-Lazikani, and John Overington for all their efforts to understand pharmacological space.

References

[1] Black JW. Receptors as pharmaceutical targets. In: Johanson JFT, editor. Textbook of receptor pharmacology. Boca Raton, FL: CRC Press; 2003. p. 271–9.
[2] Sneader W. Drug prototypes and their exploitation. London: John Wiley & Sons; 1996.
[3] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23:3–25.
[4] Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006;5(12):993–6.
[5] Bemis GW, Murcko MA. The properties of known drugs. 1. Molecular frameworks. J Med Chem 1996;39(15):2887–93.
[6] Bemis GW, Murcko MA. Properties of known drugs. 2. Side chains. J Med Chem 1999;42(25):5095–9.
[7] Paolini GV, Shapland RHB, van Hoorn WP, Mason JS, Hopkins AL. Global mapping of pharmacological space. Nat Biotechnol 2006;24(7): 806–15.
[8] Morphy R. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. J Med Chem 2006;49(10):2969–78.
[9] Vieth M, Sutherland JJ. Dependence of molecular properties on proteomic family for marketed oral drugs. J Med Chem 2006;49 (12):3451–3.
[10] Leeson PD, Springthorpe B. The influence of drug-like concepts on decision-making in medicinal chemistry. Nat Rev Drug Discov 2007;6 (11):881–90.
[11] Blake JF. Examination of the computed molecular properties of compounds selected for clinical development. BioTechniques 2003;34: S16–34.
[12] Wenlock M, Austin RP, Barton P, Davis AM, Leeson PD. A comparison of physiochemical property profiles of development and marketed oral drugs. J Med Chem 2003;46(7):1250–6.
[13] Oprea TI, Davis AM, Teague SJ, Leeson PD. Is there a difference between leads and drugs? A historical perspective. J Chem Inf Comput Sci 2001;41(5):1308–15.
[14] Blake JF. Identification and evaluation of molecular properties related to preclinical optimization and clinical fate. Med Chem 2005;1 (6):649–55.
[15] Proudfoot JR. The evolution of synthetic oral drug properties. Bioorg Med Chem Lett 2005;15(4):1087–90.
[16] Leeson PD, Davis AM. Time-related differences in the physical property profiles of oral drugs. J Med Chem 2004;47(25):6338–48.
[17] Ajay A, Walters WP, Murcko MA. Can we learn to distinguish between "drug-like" and "nondrug-like" molecules? J Med Chem 1998;41 (18):3314–24.
[18] Wang J, Ramnarayan K. Towards designing drug-like libraries: a novel computational approach for prediction of drug feasibility of compounds. J Comb Chem 1999;1(6):524–33.
[19] Walters WP, Ajay, Murcko MA. Recognizing molecules with drug-like properties. Curr Opin Chem Biol 1999;3(4):384–7.
[20] Walters WP, Murcko MA. Prediction of "drug-likeness". Adv Drug Deliv Rev 2002;54(3):255–71.
[21] Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 2000;44 (1):3–25.
[22] Podlogar BL, Muegge I, Brice LJ. Computational methods to estimate drug development parameters. Curr Opin Drug Discov Dev 2001;4 (1):102–9.
[23] Muegge I. Selection criteria for drug-like compounds. Med Res Rev 2003;23(3):302–21.
[24] Muegge I, Heald SL, Brittelli D. Simple selection criteria for drug-like chemical matter. J Med Chem 2001;44(12):1841–6.
[25] Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple K. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45(12):2615–23.
[26] Proudfoot JR. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. Bioorg Med Chem Lett 2002;12(12):1647–50.
[27] Egan WJ, Walters WP, Murcko MA. Guiding molecules towards drug-likeness. Curr Opin Drug Discov Dev 2002;5(4):540–9.
[28] Lajiness MS, Vieth M, Erickson J. Molecular properties that influence oral drug-like behavior. Curr Opin Drug Discov Dev 2004;7 (4):470–7.
[29] Williams DP, Naisbitt DJ. Toxicophores: groups and metabolic routes associated with increased safety risk. Curr Opin Drug Discov Dev 2002;5:104–15.
[30] Willams DP, Park K. Idiosyncratic toxicity: the role of toxicophores and bioactivation. Drug Discov Today 2003;8(22):1044–50.
[31] Hakimelahi GH, Khodarahmi GA. The identifications of toxicophores for the prediction of mutagenicity, hepatotoxicity, and cardiotoxicity. J Iran Chem Soc 2005;2(4):244–67.
[32] Drews J, Ryser S. Classic drug targets. Nat Biotechnol 1997;15:1318–19.
[33] Overington J. Prioritizing the proteome: identifying pharmaceutically relevant targets. Drug Discov Today 2002;7(9):516–21.
[34] Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov 2002;1:727–30.
[35] Golden J. Towards a tractable genome: knowledge management in drug discovery. Curr Drug Discov 2003;3:17–20.
[36] Golden JB. Prioritizing the human genome: knowledge management for drug discovery. Curr Opin Drug Discov Dev 2003;6(3):310–16.
[37] Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 2006;34(1):D668–72.

[38] Hopkins AL, Groom CR. Target analysis: a priori assessment of druggability. Ernst Schering Res Found Workshop 2003;42:11–17.
[39] Hajduk PJ, Huth JR, Fesik SW. Druggability indices for protein targets derived from NMR-based screening data. J Med Chem 2005;48:2518–25.
[40] Cheng AC, Coleman RG, Smyth KT, Cao Q, Soulard P, Caffrey DR, et al. Structure-based maximal affinity model predicts small-molecule druggability. Nat Biotechnol 2007;25(1):71–5.
[41] Chothia C. Hydrophobic bonding and accessible surface area in proteins. Nature 1974;248:338–9.
[42] Kuntz ID, Chen K, Sharp KA, Kollman PA. The maximal affinity of ligands. Proc Natl Acad Sci USA 1999;96(18):9997–10002.
[43] Al-Lazikani B, Gaulton A, Paolini G, Lanfear J, Overington J, Hopkins A. The molecular basis of predicting druggability. In: Wess G, Schreiber S, editors. Chemical biology. London: John Wiley & Sons; 2007.
[44] Hubbard TJ, Ailey B, Brenner SE, Murzin AG, Chothia C. SCOP: a structural classification of proteins database. Nucleic Acids Res 1999;27(1):254–6.
[45] Hann MM, Leach AR, Harper G. Molecular complexity and its impact on the probability of finding leads for drug discovery. J Chem Inf Comput Sci 2001;41(3):856–64.
[46] Lander E, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.
[47] Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001;1304–51.
[48] Russ AP, Lampel S. The druggable genome: an update. Drug Discov Today 2005;10(23–24):607–10.
[49] Orth AP, Batalov S, Perrone M, Chanda SK. The promise of genomics to identify novel therapeutic targets. Expert Opin Ther Targets 2004;8(6):587–96.
[50] Consortium article. Finishing the euchromatic sequence of the human genome. Nature 2004;431(7001):931–45.
[51] Hemmi H, Kaisho T, Takeuchi O, Sato S, Sanjo H, Hoshino K, et al. Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signaling pathway. Nat Immunol 2002;3(2):196–200.
[52] Hopkins AL, Polinsky A. Knowledge and intelligence in drug design. Annu Rep Med Chem 2006;41:425–37.
[53] Stewart KD, Shiroda M, James CA. Drug Guru: a computer software program for drug design using medicinal chemistry rules. Bioorg Med Chem 2006;14(20):7011–22.
[54] van Gestel S, Schuermans V. Thirty-three years of drug discovery and research with Dr Paul Janssen. Drug Dev Res 1986;8(1–4):1–13.

CHAPTER

17

Systems Biology: A New Paradigm for Drug Discovery

Hiba Abi Hussein[1], *Alexandre Borrel*[1,2], *Leslie Regad*[1], *Delphine Flatters*[1], *Anne Badel*[1], *Colette Geneix*[1], *Michel Petitjean*[1], *Anne-Claude Camproux*[1] *and Olivier Taboureau*[1]

[1]UMR-S973, MTi, University Paris Diderot, Paris, France;
[2]Faculty of Pharmacy, University of Helsinki, Helsinki, Finland

OUTLINE

Everything should be as simple as it is, but not simpler —**Albert Einstein**

I. INTRODUCTION

For fifty years, the "one target, one drug" paradigm has been the driving force for developments in pharmaceutical research. Although this strategy allowed bringing new drugs to the market, a significant decrease in the rate of new drug candidates has been observed. The number of failures of drug candidates in advanced stages of clinical trials has increased, and the number of submissions for the US Food and Drug Administration's (FDA) approval has decreased in the last decade [1]. The reasons for this attrition are essentially the result of a lack of efficacy and clinical safety or toxicology. Therefore, to enrich our knowledge of the mechanisms of disease and drug effect that subsequently leads to greater success for the pharmaceutical industry, alternative approaches are needed [2].

Recent advances in chemical biology and systems biology have shown that most drugs interact with multiple targets, and the pharmacological profile of a drug is not as reductionist as we believed. For example, celecoxib, often considered as a selective cyclooxygenase-2 nonsteroidal anti-inflammatory drug (NSAID) [3], has been documented to be active on at least two additional targets, namely carbonic anhydrase II and 5-lipoxygenase [4]. Rosiglitazone, which has been used for the treatment of type II diabetes mellitus, not only stimulates the peroxisome proliferator-activated receptor gamma, but also blocks interferon gamma-induced chemokine expression in Graves' disease or ophthalmopathy [5–6]. Cyclobenzaprine, a skeletal muscle relaxant [7], has been reported recently to block the serotonin and norepinephrine transporters and to bind to five other serotonin receptors in addition to its already known binding profile for muscarinic and histaminic receptors [8]. Mometasone furoate, an anti-inflammatory corticosteroid that is active through its agonist effect on glucocorticoid receptor, has shown some inhibitor activity to the ATP-binding cassette B1 (ABCB1) efflux pump [9], in addition to showing progesterone receptor agonism and androgen receptor antagonism [10]. Imatinib, a drug used in the treatment of chronic myelogenous leukemia (CML) [11] is reported to block the activity of several nonspecific tyrosine kinases [12].

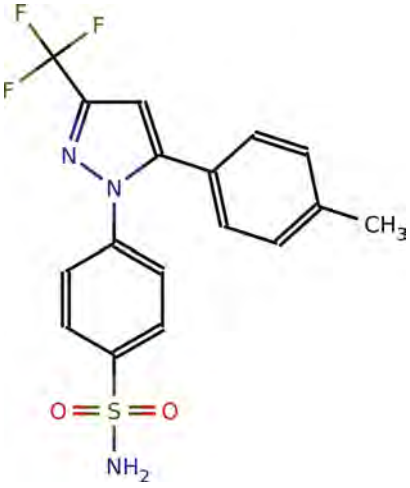
Although multiple activities against several targets might be beneficial, the drawback is that it can also lead to dramatic side effects and toxicity. Cisapride, which acts as a serotonergic 5-HT4 receptor agonist [13], as well as astemizole, which blocks histamine H1 receptors (H1R), have both been withdrawn from all markets due to the risk of fatal cardiac arrhythmia associated with their blockade of the hERG potassium ion channel, an unanticipated and undesirable "anti-target" associated to QT prolongation and torsades de pointes [14]. Fenfluramine and phentermine were also withdrawn from the market after the proliferation of valvular heart disease on numerous patients [15]. An illustration of few successful ligand-based target-fishing experiments is shown in Table 17.1.

In addition to the boost of biochemical assays or cell based high-throughput screening lead discovery, the recent advances in systems biology and genomic are expected to pave the way for major improvements in drug efficacy and safety. Since proteins rarely operate in isolation inside and outside cells but rather function in highly interconnected cellular pathways, systems biology approaches attempt to consider a more global physiological environment of protein targets and biological processes. It aims to identify components and networks that are either participating in the generation of a phenotype or diseases or are deregulated by a treatment.

So far, two major issues have limited our understanding how small molecules perturb biological systems:

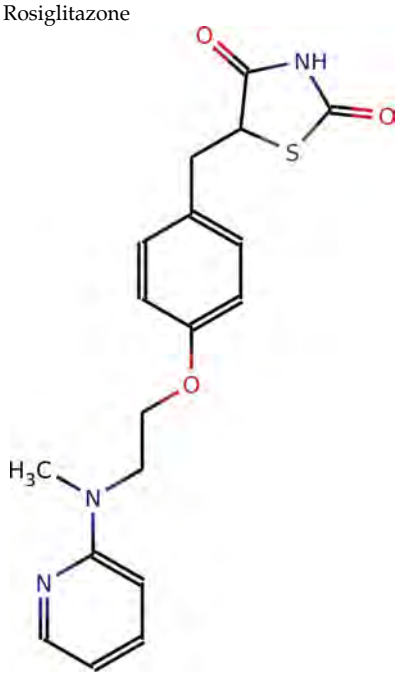
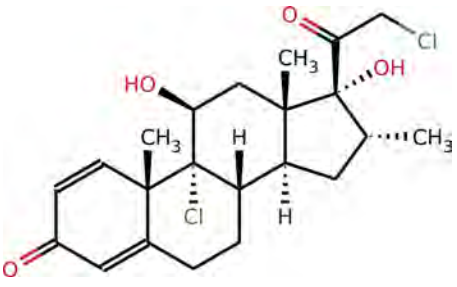
- The generation, gathering, and availability of experimental data are sparse.
- The integration, organization, and visualization of these data are complex.

TABLE 17.1 An Illustration of a Few Drugs with the Primary Target Associated to the Main Effect and the Off-Target Associated to a Second Effect. The Second Effect Might be Beneficial (and so can be considered for drug repurposing) or Associated with Strong Side Effects (which can cause the withdraw of the drug)

Compound	Main effect	Second effect	Primary target	Off-target	References
Celecoxib 	Anti-inflammatory drug	Anticancer properties reduce the number of adenomatous colorectal polyps	COX-2 inhibitors	– CA2 inhibitor – ALOX5 inhibitor	[3,4]

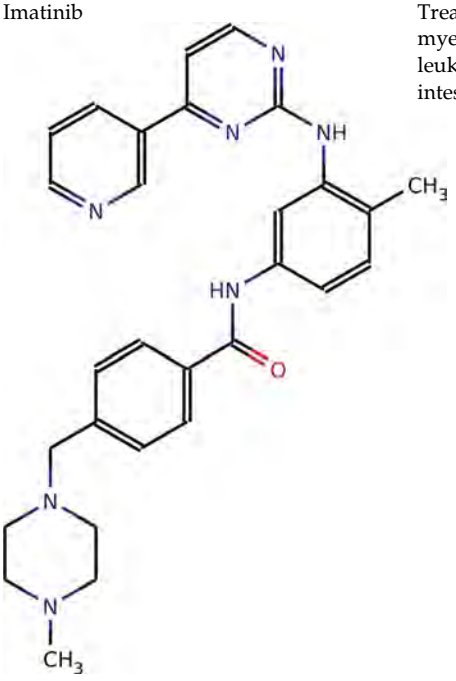
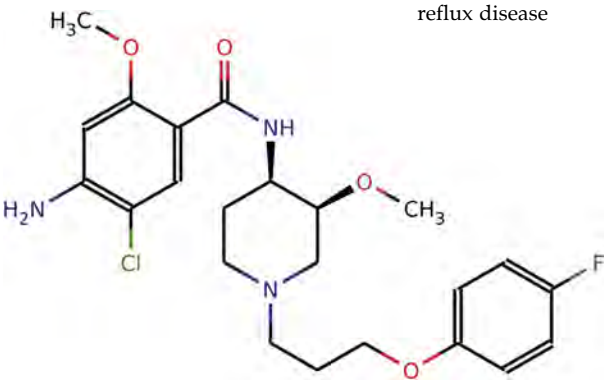
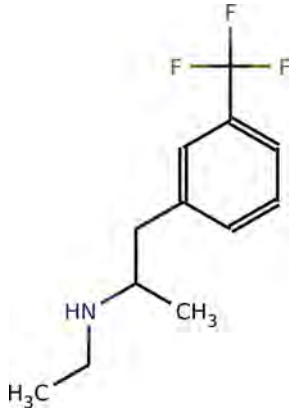
(*Continued*)

TABLE 17.1 (Continued)

Compound	Main effect	Second effect	Primary target	Off-target	References
Rosiglitazone 	Antidiabetic drug for type II diabetes mellitus	Anti-inflammatory effect	PPARγ agonists	IFNγ inhibitor	[4,5]
Cyclobenzaprine 	Skeletal muscle relaxant	Antidepressant	5-HT2 antagonist	– CHRM1 blocker – CHRM3 blocker – H1R antagonist	[6–8]
Mometasone furoate 	Anti-inflammatory corticosteroid	Tumoricidal effects in leukemia	NR3C1 agonist	– ABCB1 antagonist – AR antagonist	[7,8]

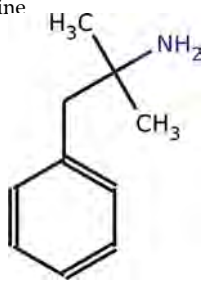
(*Continued*)

TABLE 17.1 (Continued)

Compound	Main effect	Second effect	Primary target	Off-target	References
Imatinib 	Treatment of chronic myelogenous leukemia and gastro intestinal cancer	Left ventricular contractile dysfunction (cardiotoxic)	BCR-ABL tyrosine kinase inhibitor	no specific tyrosine kinases inhibitor	[9]
Cisapride 	gastroesophageal reflux disease	– torsades de pointes – cardiac arrest	– HTR4 agonist – HTR3 agonist – HTR2A agonist	hERG inhibitor	[13]
Fenfluramine 	The management of exogenous obesity	valvular heart disease	HTT inhibitor	– 5-HTR2B agonist – 5 HTR2C agonist	[15]

(*Continued*)

TABLE 17.1 (Continued)

Compound	Main effect	Second effect	Primary target	Off-target	References
Phentermine 	The management of exogenous obesity	valvular heart disease	– HTT inhibitor – NET inhibitor – DAT inhibitor	– 5-HTR2B agonist – 5-HTR2C agonist	[15]

Fortunately, massive amounts of data are generated and accumulated by new experimental technologies such as toxicogenomics, proteomics, metabolomics, and genomics (through next-generation sequencing). Furthermore, centralized systems that facilitate the integration and standardization of diverse federated resources are in development. So, drug action can be explored across multiple scale of complexity, from molecular and cellular to tissue and organism levels over time and dose. For example, it has become apparent that many common diseases such as cancer, cardiovascular diseases, and mental disorders are much more complex than initially anticipated, as they are caused by multiple molecular and cellular dysfunctions rather than being the result of a single defect. As a result, network-centric therapeutic approaches that consider entire pathways rather than single proteins must be investigated [16].

In this chapter, we will describe the different biological and clinical sources of information that can be integrated with chemistry in the aim to enrich our knowledge on drug effects and how they can be explored using systems biology. We will start with a description of molecular interactions (drug-target space). We will then explain how the integration of protein−protein interactions, pathways, tissues and organs, and genetic variations to such data can provide new insight into the mechanism of action of a drug (systems biology space). We will finally highlight how this information is of interest to explain patients' reactions to drugs (phenotype space; Figure 17.1) and for the design of new drugs in drug discovery.

II. DRUG-TARGET SPACE (OFF-TARGET)

Drugs exert their therapeutic and adverse effects by interacting with molecular targets. Following the lock and key model proposed by Ehrlich more than a century ago [17], drug discovery efforts have focused on identifying single selective drugs that target a single mechanism; that is, identifying ligands (keys) that fit into specific targets (locks). This paradigm has guided the pharmaceutical industry throughout the last three decades [18]. Nowadays, it has become obvious that the problem is much more complicated than Ehrlich's magic bullet, which is likely the exception, not the rule, and would explain why the attrition rate at the late stage of clinical trials is typically as high as 90 percent, which is a result of the lack of both efficacy and clinical safety [19]. These failures are extremely costly to pharmaceutical companies. A single new drug typically costs more than US$800 million and may take fifteen to twenty years to be developed [20]. Contrary to the hypothesis that a drug has only one target and one indication, one drug or a combination of different drugs can hit one or more targets, modulating multiple pathways with potential benefit for multiple disease conditions. In recent years, mounting evidence makes it increasingly obvious that many drugs elicit their therapeutic activities by modulating multiple targets [21−24]. By analogy with the lock and key model, the goal of lead identification is being expanded accordingly to identify master key compounds that favorably interact with multiple targets (i.e., operate a set of desired locks to gain access to the expected clinical effects) [25].

Multi-target pharmacology exploration increases with readily available information, linking the relationships between the entire chemical and target spaces. Thanks to homogenization of data archives and mining, knowledge on protein − ligand binding data is increasing at an amazing gait [26−27]. For example, ChEMBL [28] is a repository of more than 12 million bioactivity data gathered from literature and addressing more than 1.3 million ligands (with their molecular structure) and 9,300 molecular targets (data obtained in Dec 2013). PubChem [29],

ChemBank [30], ChemProt [31], and OPS [32] are other large chemical biology resources publicly available. Protein structure data are also of high relevance. The Protein Data Bank [33] stores 89,000 3D structures of proteins and protein − ligand complexes. With the increase facility for academics to get access to this data, chemogenomic, proteochemometric, and polypharmacology approaches have started to be developed [34]. Such methods help to mine this vast amount of protein−ligand interactions [34] and allow predicting a single ligand against a set of heterogeneous targets [35].

Globally, *in silico* prediction of novel protein−ligand interactions can be classified into three groups of methods: (1) ligand-centric profiling methods, based on the concept that similar ligands bind to similar targets [36]; (2) a second group of methods relies on the concept that similar ligands bind to similar binding sites; and (3) approaches focusing on protein − ligand complexes that can be described either as simple 1D fingerprints [37], protein − ligand-derived pharmacophores, [38−39], or protein − ligand docking poses [40]. Ligand-centric profiling methods are, however, restricted to targets for which sufficient ligand information is available. Comparison of binding sites, can be carried out with the aim of exploring the relationship with the corresponding ligands requiring or not the 3D structures of the targets [41−45]. These methods have inherited limitations, as it is confined to targets for which a 3D structure is available (structure-based approach) or to a target subfamily in order to avoid binding site-based misalignments (sequence-based approach). With the targets promiscuity concept, proteins similar in terms of sequences are expected to interact with similar drugs. Combining the two approaches (ligand-based, target-based similarity) proteochemometric modeling allowed to extrapolate the activities of known ligands for known targets to novel targets and, conversely, to screen for selective compounds that are solely active on a single member of a subfamily of targets [46−48] and to investigate the so-called "ligand−target interaction space" [49−51].

This step of drug-target interaction annotation is the foundation of a pharmacological analysis, as the assessment of clinical efficacy and adverse effects of drugs begins at the molecular level, involves complex biological networks, and is ultimately measured by clinical outcomes at the level of the whole organism. Beyond determining putative on/off-targets, there is a need to decipher the impact of such binding on the complete biological system, with the ultimate goal of being able to predict the phenotypic outcome. The effect of the drug response at the systems level is particularly important for understanding polypharmacology, since there are many reported examples where the therapeutic efficacy is enhanced through synergistic relationships among multiple targets [52−55].

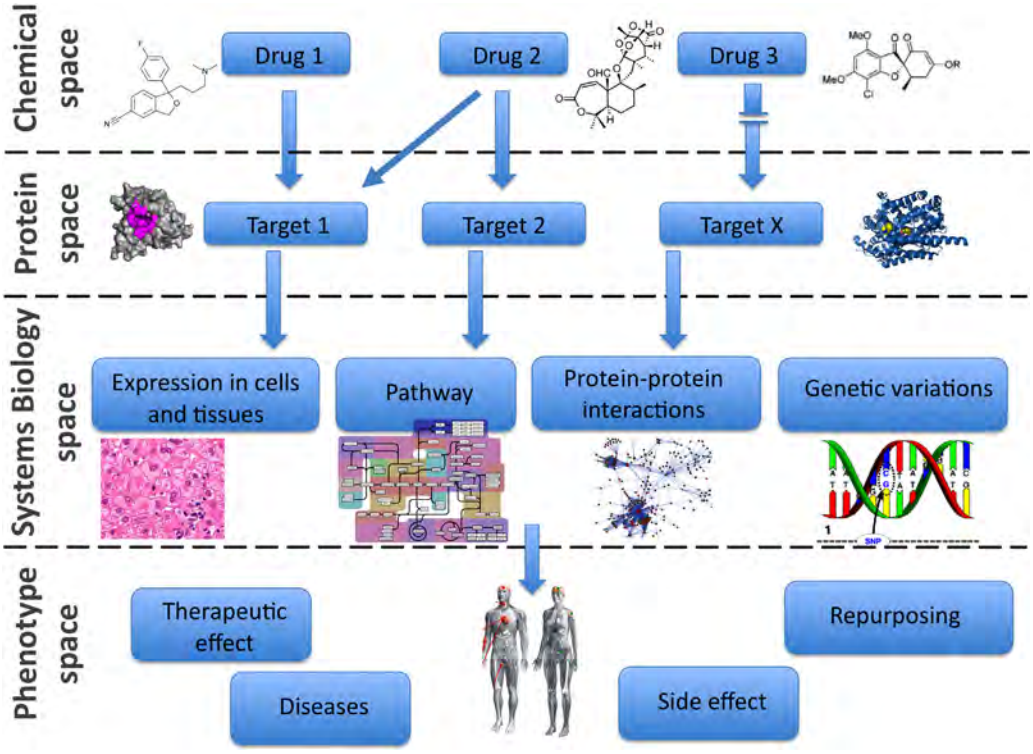


FIGURE 17.1 Interaction schema between the chemical space, the protein space, the systems biology space, and the phenotype space that will be discussed in this review.

III. SYSTEMS BIOLOGY SPACE

A. From Target to Protein-Protein Interactions and Pathways Using Network Biology Approaches

The objective of systems biology is to study the biological mechanism associated with a phenotype, a disease, or a treatment with the integration of several components at different layers of complexity (from molecular and cellular to tissue, organs, and systems). In this concept, it is assumed that the functions of molecular components in a human cell are closely connected, and thus a disease is not only a consequence of a genetic variation but also a consequence of perturbations of complex intracellular and extracellular networks linking tissue and organ systems [56]. So, the perception that human diseases are associated with a few dominant factors (reductionist) is replaced by the view of diseases as the outcome of many weak contributors (holistic).

Network biology is an interesting approach that has been used in many studies to gain insight into disease and phenotype mechanisms, study comorbidities, and analyze therapeutic drugs and their targets. The biological network should not be random but instead should follow an organized interaction between the components. Features such as node (i.e., the entity on which the network is building up), edges (which can represent physical interactions, functional interactions, or simply connection between multi-scale data), degree of distribution around a node and the forming hubs, distance, degree of clustering or betweenness centrality are usually considered to assess the properties of the biological network (Figure 17.2). More information about network biology can be found in these reviews [57,58].

Several data types can be used and integrated in a network-centric aspect, and we will focus on protein–protein interactions and pathways, transcriptomics data, and genomics variation.

B. Protein–Protein Interactions

Protein–protein interactions (PPI) have become a valuable resource of data implemented in network biology. PPI are derived from high-throughput approaches, including yeast two-hybrid screens, immunoprecipitation studies followed by mass spectrometry analysis, and small-scale experiments [59]. Based on theses PPI information a lot of studies considering phenotype as a result of a deregulation of an ensemble of proteins interconnected we reported. For example, network-based analyses applied to diverse phenotypes demonstrated that products of genes that are implicated in similar diseases are clustered together into highly connected subnetworks (or modules) in PPI networks [60,61]. Following this trend, network biology was applied to pharmacological data. For example, Yildirim et al provided systematic information about drugs and their targets in the context of cellular and disease networks by generating a bipartite graph of all known FDA approved drugs and their targets [62]. In addition, the authors overlaid the drug-protein network with a network of physical PPI. They demonstrated that the drug-protein networks have 42 percent more interacting proteins than the average protein in the PPI network, a feature that has been correlated with the essentiality of the proteins. Finally, they examined etiological and

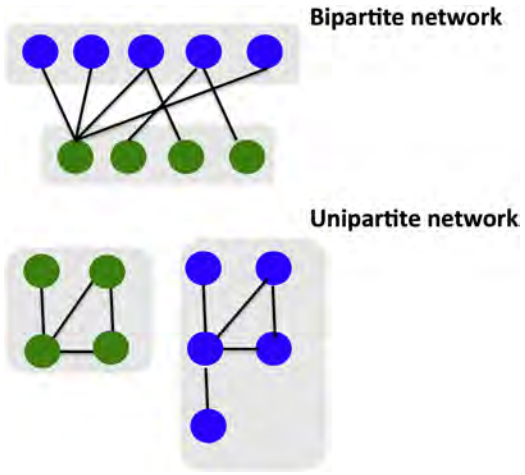


FIGURE 17.2 Example of a network biology graph between two sets of data. The bipartite network represents the interconnection between the two sets. The unipartie network shows the connection inside each set.

palliative drugs in the network by estimating the number of molecular steps between a drug-target and the corresponding disease cause. Based on comparison with randomized expectation, drugs for cancer, endocrine, psychiatric, and respiratory disease classes appear to be more rationally designed than developmental, muscular, and ophthalmologic disease classes.

Similarly, Hopkins expanded drug-target interactions with protein−protein interactions in order to develop a network pharmacology that could better explain the drug-phenotype relationship [63]. Berger et al and Taboureau et al illustrated the usefulness of such PPI networks to predict and identify drugs likely to cause Long QT Syndrome (LQTS), not only because of a direct drug-target interaction but also because gene products involved in LQTS formed a distinct neighborhood [64,65]. Lage et al [66] went further and showed a functional interaction between the development of specific anatomical structures of the heart and protein network. So, in contrast to the single-target approach, network pharmacology identifies a complex (or combination) of proteins whose perturbation (even indirect) results in the clinical outcome observed. An example with arrhythmia is presented in Figure 17.3.

C. Transcriptomics Data

Complementary to protein−protein interactions, transcriptomic data and gene expression differentiation, triggered by drugs, have been increasingly investigated, thanks to technological advances, initially with microarray and now with next-generation sequencing [67]. Such experiments allow the exploration of drug effects on a cellular or tissue level, *in vitro* and *in vivo*, and at different time and dose conditions. Interestingly, several of these data have been made available to the public through large compendium such as the Gene Expression Omnibus (GEO) [68], ArrayExpress [69], the connectivity maps (CMAP) [70], DrugMatrix [71], and The Japanese Toxicogenomics Project (TG-GATEs) [72]. Based on these data, several initiatives have been developed. For example, Iskar et al combined drug-targets information and genes expression profiles after drug treatment to identify the deregulation of new drug-targets induced by drugs that could explain the repurposing of drugs or potential side effects associated to them [73]. Emig et al took as input a disease gene expression signature and a high quality interaction network to prioritize drug-targets associated with a particular disease [74]. Baron et al gathered all the public microarray data related to skeletal and heart muscle from GEO, including seven different animal species from invertebrates and vertebrates [75]. Subsequently, they developed a tool for meta-analysis of the muscle transcriptome data allowing the user to identify robust signatures of pathology or drug treatment among different studies. On those grounds they generated lists of co-expressed genes, and they identified gene lists that overlap between human and animal models or/and are similarly modulated between disease states or after drug treatment.

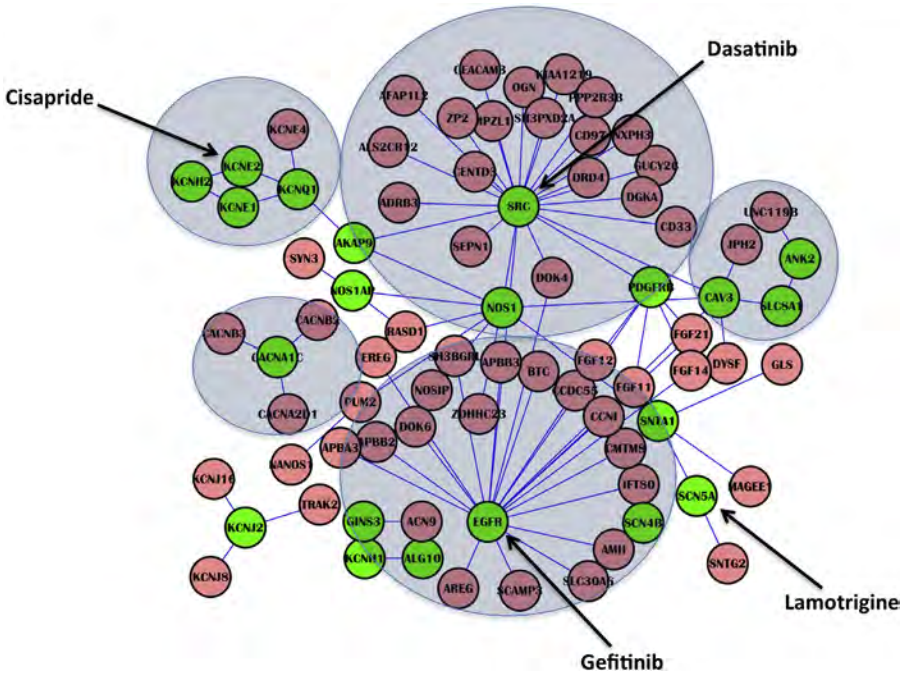


FIGURE 17.3 Example of drugs (Cisapride, Gefitinib, Lamotrigine, Dasatinib). Targets susceptible to being associated with arrhythmia (in green) and their direct protein−protein interactions associated (pink). Clusters (blue circles) represent hubs of proteins. The cumulative effect of several drugs might be assumed as they interact with proteins involved directly or indirectly in a protein−protein interaction.

The CMAP consortium developed pattern-matching tools to detect similarities in the gene expression signatures and discover unexpected connections between genes (new therapeutics), disease states (signature conservation across diverse cell types and settings), and drug molecules (recognize drugs with common mechanisms of action). A Gene-Set Enrichment Analysis (GSEA) is also implemented in CMAP. With the GSEA, genes are scored based on their differential expression between two biological states (such as healthy versus diseased). The use of GSEA allowed understanding of the chemical mechanism of action in risk assessment and with toxicogenomics data. A number of gene sets whose expression levels are closely associated with certain toxicological endpoints have been reported [76]. An example of how to detect similar gene expression signature between compounds is depicted in Figure 17.4.

D. Biological Enrichment

Once a target list is generated from these experiments, additional information can be integrated to enrich the biological outcomes associated with the list. For example, proteins are preferably expressed in some tissues over others, and such data are available through database like the Human Protein Atlas [77], the GNF tissues atlas [78], or Pagenbase [79]. For example, a combination of gene expression in human protein complexes revealed tissue specificity and pathology [80]. Similarly, McCall et al [81] leveraged data from the GEO and ArrayExpress public repositories to build statistical models for the most annotated genes for 131 human and 89 mouse tissue types to address which genes are expressed in a given cell type. Such information is useful in understanding the chemical mechanism of action linked to the drug tolerance, drug efficacy, side effects, and risk assessment.

Other biological sources of information (e.g., Gene Ontology (GO) annotations [82], metabolic or signaling pathway (KEGG [83], Reactome [84], and inherent diseases such as OMIM [85]) can also be integrated and might help in the interpretation of drug effects due to promiscuity (proteins with similar functions) or pathways perturbation.

E. Pharmacogenomics

With the increasing amount of human high-throughput genotyping and sequencing, the study of genetic factors that contribute to the interindividual variability in drug efficacy and safety (i.e., pharmacogenomics) is a

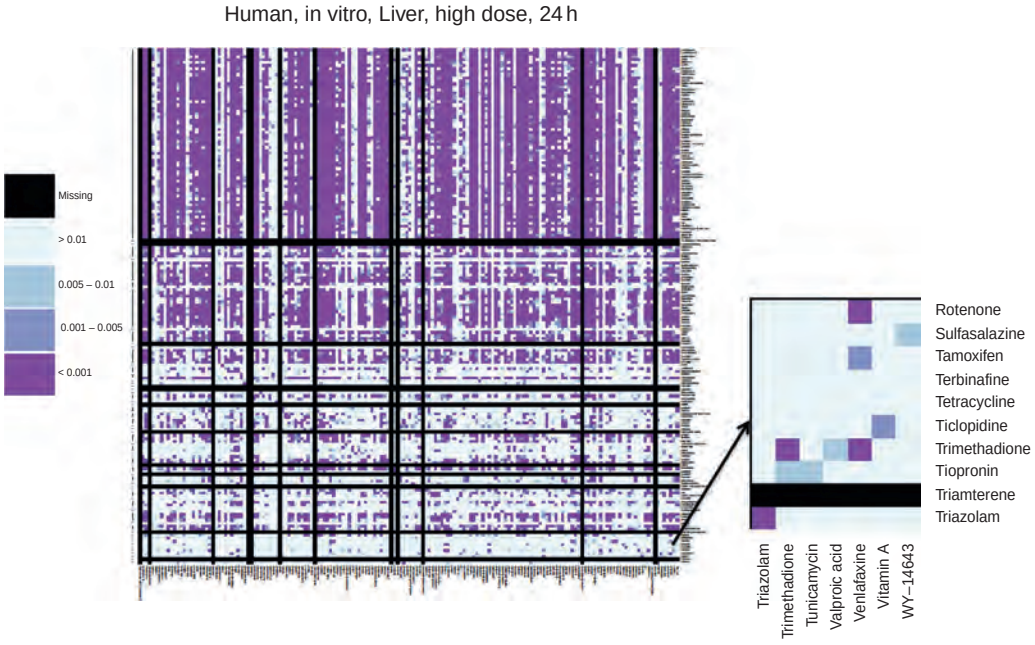


FIGURE 17.4 The heatmap represents the similarity of gene expression profiles for compounds from TG-GATEs, tested on the same conditions in human hepatocytes cells. When a cell is magenta, it means compounds share highly similar gene profiles. For example, venlafaxine causes a similar deregulation of genes as rotenone or trimethadione. When the cell is white, the cell response to the two compounds is different. A black cell means the compounds have not been tested in the same condition.

tremendous challenge in drug development. During the last few years, the number of known single nucleotide polymorphisms (SNPs) has increased exponentially [86]. A major database, dbSNP, contributes to the era of personal genomics with more than 40 million human SNPs (as of Build 138, April 2013) and is the largest repository. dbSNP enables biomedical researchers to retrieves SNPs reported for genes and to connect their potential influence in the protein function and diseases associated to the genes [87]. It is estimated that the genetic component accounts for 20–95 percent of variability in drug response [88]. Many functionally relevant polymorphisms have been identified in drug metabolizing enzymes (DMEs) and transporters [89]. Still, most of the genetic variability in drug response is likely to be associated with complex traits involving multiple genes, including relevant pathways. Differences in age, sex, race, organ dysfunction, disease characteristics, comorbidities, co-medications, drug–drug interactions, life-style, and environmental factors can contribute to the drug response variability. There is evidence that SNPs are the most frequent type of sequence variation between individuals (0.1 percent of sequence variation in a diploid genome [90]), and identifying the functional effect of predisposed SNPs to a drug response is of particular interest in the context of system biology. To do such analysis, integration of pharmacological and clinical data is needed.

IV. PHENOTYPE SPACE

There are many online sources that contain clinical information and allow the collection of data from clinicians, patients, and pharmaceutical companies. Among others, there is AERSs (created by the US FDA, World Health Organization, and Health Canada [91]), EudraVigilance (a system designed for collecting reports of suspected side effects created by the European medicines agency [92]), and JAPIC (which manages all package insert information of pharmaceutical products in Japan, under the approval of Health and Welfare Minister of Japan [93]). However, there is no common terminology for therapeutic effects and side effects in each organization. Thus, to facilitate the coding of "regulatory data" in biopharmaceutical development and clinical trials, and the reporting of therapeutic and side effects, several dictionaries (terminologies) have been developed. The most used are MeSH (Medical Subject Headings [94]), MedDRA (Medical Dictionary for Regulatory Activities [95]), ICD-10 (International classification of diseases [96]), SNOMED CT (Systematized Nomenclature of Medicine-Clinical Terms [97]), the ATC Classification System (Anatomical Therapeutic Chemical Classification System [98]), UMLS (Unified Medical Language System [99]), and J-ART (Japanese Adverse Reaction Terminology). Many databases have been built based on these data sources (e.g., DrugBank [100], SIDER [101], OFFSIDES [102], and METAADEDB [103]).

Such clinical information, in association with the drug-targets and the omics data, can provide a holistic picture of biological functions for a better understanding of the drug effects (therapeutic and side effects). It can also contribute to the field of drug repurposing and drug–drug interactions. An example is shown in Figure 17.5.

V. EXAMPLES

The following sections, drug-diseases and drug-side effects, will describe several examples where the integration of systems biology into chemical biology can be applied.

A. Drug-Diseases

Classically, successful drugs have been analyzed through their binding and alteration of a protein activity [104]. But gene expression data profiles after drug treatment reveal new mechanisms of action and new potential drug-targets. For example, Iskar et al [73] used the CMAP to study the mechanisms of positive and negative feedback loops that adjust the expression level of drug-targets. The authors developed a state-of-the-art normalization and scoring procedure for assessing the drug-induced differential regulation of drug-targets in three cell lines. Using the normalized data of 1,290 drug-target relations, they found that 8 percent of drug-induced drug-targets are regulated at the mRNA level. They also identified unknown drug-induced target expression changes that were linked to the development of drug tolerance. Similarly, Ioro et al considered the variability in the transcriptional responses to a small molecule due to differences in cell lines or dosages. By removing unspecific effects

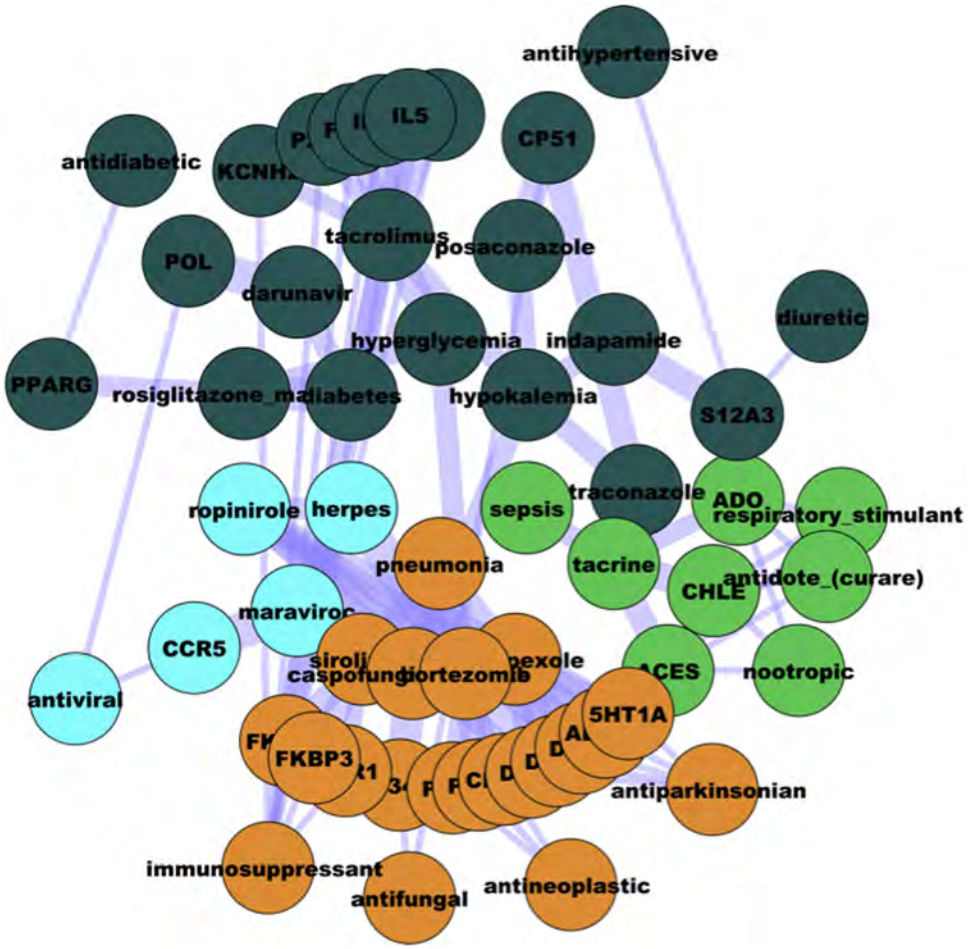


FIGURE 17.5 Integration of drug-target, therapeutic, and side effect into biological network. Each color corresponds to the system and organ where the side effect with the highest frequency is occurring for a set of drugs described in DrugBank. In this figure, cyan is for the nervous system, orange for the respiratory, dark green for metabolic, and lime green for systemic side effects.

and computing a consensus transcriptional response, they constructed drug networks that were further divided to interconnected modules. These modules were analyzed for similarities and differences in pharmacological effects, and they were applied for predicting previously unreported modes of action for anticancer drugs [105]. Instead of considering genes in isolation, a global distance biology network can be used to consider both direct and indirect paths in the network [106].

The application of systems biology with drug treatment into different tissues and cell lines is also investigated for drug repurposing and drug–drug interactions. Drug repurposing is of high interest for pharmaceutical companies, as this approach has the potential to increase success rates, decrease time, and reduce development costs, resulting in the acceleration of drug development compared with conventional drug design. In addition, a repositioned drug has valid toxicology and safety assessments to reduce the risk of failure [107].

A recent publication integrating genes expression, protein–protein interactions, and pathways has allowed the identification of new activities from some drugs. For example, vinburnine (a vasodilator) and sulconazole (a topical antifungal) blocked the cell cycle during G2/M transition and could potentially be used in anticancer therapy. Zaprinast, an experimental phosphodiesterase inhibitor was confirmed to be a modulator of PPARy (peroxisome proliferator-activated receptor y) and could be investigated in the context of diabetes. Hexetidine, an antibacterial showed some adrenergic activity, could have some new therapeutic effects in vasoconstriction and attention deficit-hyperactivity disorder [108].

Similarly, systems biology approaches started to be used for drug–drug interactions. Although DDI is usually associated with adverse effects, it might be of interest for diseases that requires a combination of several drugs, such as cancer or AIDS, not simply by targeting the same protein but by deregulating proteins involved in a

same metabolic pathway and PPI network, and looking at the similarity of the whole-expression profiles [109–113]. For example, Tatonetti et al [114] demonstrated recently that pravastatin (used to lower cholesterol) and paroxetine (an antidepressant) interact with biological pathways that are critical for glucose metabolism (Figure 17.6).

B. Drug-Side Effect

Another challenge in drug discovery with which systems biology can help is side effects and adverse drug reactions. It is expected to guide scientists in both characterizing the molecular mechanism of drugs and in early evaluation of the risk assessment of drug candidate compounds [115,116].

In general, *in silico* chemical-based approaches are used to analyze the SE-target associations [18,37,117–121]. As side effecst can also be the result of downstream pathway perturbation, kinetic and dosage effects, cell types, and tissue and species specificity [122], assessment of chemical safety using chemical genomics approaches across multiple scales of complexity from molecular to systems level is increasingly applied [123,124].

Studies demonstrate that drugs are more likely to cause side effects in the organ where it is most likely to be accumulated. For example, Scheiber et al [117] analyzed side effects according to the system organ class (SOC) extracted from MedDRA terminology and showed that the common chemistry leads to side effects only in the same organ. By integrating drug-targets, side, and therapeutic effects into a neural network, Oprea et al [125] concluded that a part of side-effect occurrence could be explained by drug compartmentalization into different systems.

Based on the assumption that drugs able to modulate the same pathway may induce similar SE, Wallach et al studied the relation between drugs activated pathways and side effects [126]. They identified 185 significant SE-pathways associations, of which thirty-two (among forty-five selected pathways) are supported by scientific literature. Some studies considered signaling network and PPIs. Wang et al [27] showed through a signaling network approach that drugs with a distance less than three or higher than four in network biology had significantly more side effects than others drugs. Similarly, Brouwers et al [127] showed that in PPI networks, two-thirds of the side-effect similarities were related to shared targets, while 5.8 percent of the side-effect similarities were due to drugs targeting proteins close in the human interactome.

The method developed by Xie et al [128] is also one of the approaches using side effect–target associations. It consists of the characterization of the ligand binding site of the primary target using the 3D experimental structure or homology model, the identification of the off-target proteins having similar ligand binding site (identified by sequence alignments), the docking of these putative off-targets on the drug using protein–ligand docking methods, and the ranking of these off-targets in order to identify the ones having the highest affinities to the drug. Finally, the selected off-targets are clustered and incorporated into a network that includes multiple metabolic, signal transduction, and gene regulation pathways. It has been shown that using this approach, side effects can be early identified, and the interactions with off-targets can be greatly reduced.

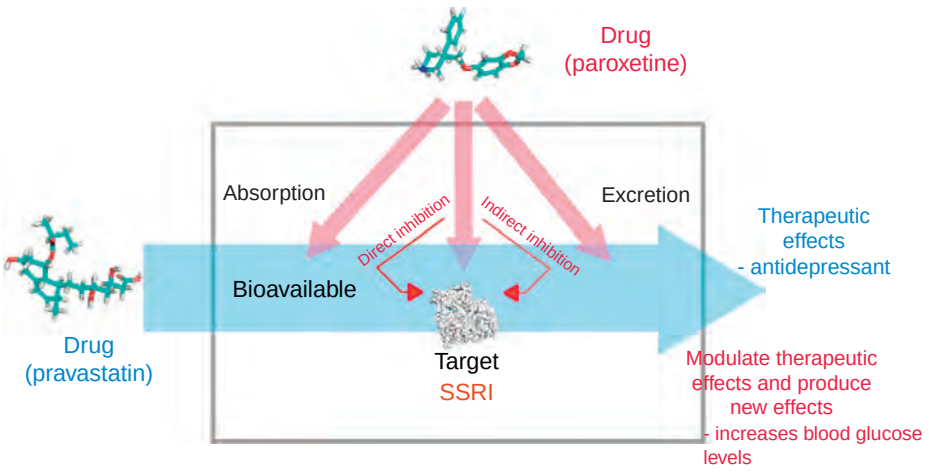


FIGURE 17.6 Drug–drug interactions between paroxetine and pravastatin susceptible to increase blood glucose levels.

VI. CONCLUSION

To understand the impact of a ligand in targeting multiple proteins in a biological system, we have to answer several questions. What are the common features of the protein–ligand interaction that persists across gene families [129]? Hence, which specific proteins are inhibited/activated by the ligand? How are these proteins connected in biological pathways? Given these pathways, how does inhibition/activation of these proteins affect the overall physiological process? Further questions arise when considering an individual response to a given therapeutic or in the treatment of a fast evolving pathogen. Namely, how do single or multi-amino acid mutations alter ligand binding and consequently the physiological effect? In general, the model performance increases when the model combines different parameters. In their study, Bresso et al developed a method to predict the side effect–side effect association [130]. They showed that in addition to classical parameters (target, chemical structure, and biological process annotation), information about pathways, protein–protein interactions, and protein domain also plays an important role in side-effect characterization. More and more drug-target interactions models are based on the genomic expression profiles. With the omics revolution, a massive amount of data is now available and can be used in drug discovery. The mechanisms of action and the safety of drugs start to be explored not only at the molecular level but also at the level of the whole biological system. Network biology seems to be an essential component of the systems approach, which in turn can generate testable hypothesis based on analyzing and modeling large amounts of data. It allows the assessment of the chemical effect at the biological systems (i.e., systems pharmacology and systems toxicology) more comprehensively.

At the population level, whereas genetic disorders have been well studied, understanding the individual differences in drug response in the context of biological networks is the new challenge in pharmacogenomics and personalized medicine. To do so, methods that integrate drug-target, clinical-outcome, and genetic factors using network biology have started to be reported [131–133]. Such analysis would definitively contribute to a better understanding of the variability in drug response and a more personalized approach to therapy.

References

[1] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004;3:711–15.
[2] Schadt EE, Zhang B, Zhu J. Advances in systems biology are enhancing our understanding of disease and moving us closer to novel disease treatments. Genetica 2009;136:259–69.
[3] Sud'ina GF, Pushkareva MA, Shephard P, Klein T. Cyclooxygenase (COX) and 5-lipoxygenase (5-LOX) selectivity of COX inhibitors. Prostaglandins Leukot Essent Fatty Acids 2008;78:99–108.
[4] Weber A, Casini A, Heine A, Kuhn D, Supuran CT, Scozzafava A, et al. Unexpected nanomolar inhibition of carbonic anhydrase by COX-2-selective celecoxib: new pharmacological opportunities due to related binding site recognition. J Med Chem 2004;47:550–7.
[5] Antonelli A, Ferrari SM, Fallahi P, Piaggi S, Paolicchi A, Franceschini SS, et al. Cytokines (interferon-gamma and tumor necrosis factor-alpha)-induced nuclear factor-kappaB activation and chemokine (C-X-C motif) ligand 10 release in Graves' disease and ophthalmopathy are modulated by pioglitazone. Metabolism 2011;60:277–83.
[6] Antonelli A, Ferrari SM, Frascerra S, Ruffilli I, Pupilli C, Bernini G, et al. β (CCL2) and α (CXCL10) chemokine modulations by cytokines and peroxisome proliferator-activated receptor-α agonists in Graves' ophthalmopathy. J Endocrinol 2012;213:183–91.
[7] Mestres J, Seifert SA, Oprea TI. Linking pharmacology to clinical reports: cyclobenzaprine and its possible association with serotonin syndrome. Clin Pharmacol Ther 2011;90:662–5.
[8] Gautam D, Heard TS, Cui Y, Miller G, Bloodworth L, Wess J. Cholinergic stimulation of salivary secretion studied with M_1 and M_3 muscarinic receptor single- and double-knockout mice. Mol Pharmacol 2004;66:260–7.
[9] Winter SS, Lovato DM, Khawaja HM, Edwards BS, Steele ID, Young SM, et al. High-throughput screening for daunorubicin-mediated drug resistance identifies mometasone furoate as a novel ABCB1-reversal agent. J Biomol Screen 2008;13:185–93.
[10] Madauss KP, Deng SJ, Austin RJ, Lambert MH, McLay I, Pritchard J, et al. Progesterone receptor ligand binding pocket flexibility: crystal structures of the norethindrone and mometasone furoate complexes. J Med Chem 2004;47:3381–7.
[11] Nadal E, Olavarria E. Imatinib mesylate (Gleevec/Glivec): a molecular-targeted therapy for chronic myeloid leukemia and other malignancies. Int J Clin Pract 2004;58:511–16.
[12] Kerkelä R, Grazette L, Yacobi R, Iliescu C, Patten R, Beahm C, et al. Cardiotoxicity of the cancer therapeutic agent imatinib mesylate. Nat Med 2006;12:908–16.
[13] Crema F, Modini C, Croci T, Langlois M, de Ponti F. Intestinal prokinesia by two esters of 4-amino-5-chloro-2- methoxybenzoic acid: involvement of 5-hydroxytryptamine-4 receptors and dissociation from cardiac effects in vivo. J Pharmacol Exp Ther 1999;288:1045–52.
[14] Walker BD, Singleton CB, Bursill JA, Wyse KR, Valenzuela SM, Qiu MR, et al. Inhibition of the human ether-a-go-go-related gene (HERG) potassium channel by cisapride: affinity for open and inactivated states. Br J Pharmacol 1999;128:444–50.
[15] Rothman RB, Baumann MH, Savage JE, Rauser L, McBride A, Hufeisen SJ, et al. Evidence for possible involvement of 5-HT2b receptors in the cardiac valvulopathy associated with fenfluramine and other serotonergic medications. Circulation 2000;102:2836–41.
[16] Hansen NT, Brunak S, Altman RB. Generating genome-scale candidate gene lists for pharmacogenomics. Clin Pharmacol Ther 2009;86:183–9.
[17] Winau F, Westphal O, Winau R. Paul Ehrlich: in search of the magic bullet. Microbes Infect 2004;6:786–9.

[18] Koutsoukas A, Simms B, Kirchmair J, Bond PJ, Whitmore AV, Zimmer S, et al. From in silico target prediction to multi-target drug design: current databases, mehods, and applications. J Proteomics 2011;74:2554–74.
[19] Nwaka S, Hudson A. Innovative lead discovery strategies for tropical diseases. Nat Rev Drug Discov 2006;5:941–55.
[20] O'Connor KA, Roth BL. Screening the receptorome for plant-based psychoactive compounds. Life Sci 2005;78:506–11.
[21] Paolini GV, Shapland RH, van Hoorn WP, Mason JS, Hopkins AL. Global mapping of pharmacological space. Nat Biotechnol 2006;24:805–15.
[22] Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. Nat Biotechnol 2007;25:197–206.
[23] Morphy R, Kay C, Rankovic Z. From magic bullets to designed multiple ligands. Drug Discov Today 2004;9:641–51.
[24] Azzaoui K, Hamon J, Faller B, Whitebread S, Jacoby E, Bender A, et al. Modeling promiscuity based on in vitro safety pharmacology profiling data. Chem Med Chem 2007;2:874–80.
[25] Medina MA. Systems biology for molecular life sciences and its impact in biomedicine. Cell Mol Life Sci 2013;70:1035–53.
[26] Ekins S, Williams AJ. When pharmaceutical companies publish large datasets: an abundance of riches or fool's gold? Drug Discov Today 2010;15:812–15.
[27] Wang J, Li Z, Qiu C, Wang D, Cui Q. The relationship between rational drug design and drug side effects. Brief Bioinform 2012;13:377–82.
[28] Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, et al. ChEMBL: a large-scale bioactivity database for drug discovery. Nucleic Acids Res 2011;40(Database issue):D1100–7. Available from: http://dx.doi.org/10.1093/nar/gkr777.
[29] Chen B, Wild D, Guha R. PubChem as a source of polypharmacology. J Chem Inf Model 2009;49:2044–55.
[30] Seiler KP, George GA, Happ MP, Bodycombe NE, Carrinski HA, Norton S, et al. Chembank: a small-molecule screening and chemoinformatics resource database. Nucleic Acids Res 2008;36:D351–9.
[31] Taboureau O, Nielsen SK, Audouze K, Weinhold N, Edsgärd D, Roque FS, et al. ChemProt: a disease chemical biology database. Nucleic Acids Res 2011;39:D367–72.
[32] Williams AJ, Harland L, Groth P, Pettifer S, Chichester C. Open PHACTS: semantic interoperability for drug discovery. Drug Discov Today 2012;17:1188–98.
[33] Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, et al. The Protein Data Bank and the challenge of structural genomics. Nat Struct Biol 2000;7:957–9.
[34] Rognan D. Chemogenomic approaches to rational drug design. Br J Pharmacol 2007;152:38–52.
[35] Schneider G. Virtual screening: an endless staircase? Nat Rev Drug Discov 2010;9:273–6.
[36] Keiser MJ, Setola V, Irwin JJ, Laggner C, Abbas AI, Hufeisen SJ, et al. Predicting new molecular targets for known drugs. Nature 2009;462:175–81.
[37] Van Westen GJ, Wegner JK, Geluykens P, Kwanten L, Vereycken I, Peeters A, et al. Which compound to select in lead optimization? Prospectively validated proteochemometric models guide preclinical development. PLoS One 2011;6:e27518.
[38] Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 2005;45:160–9.
[39] Meslamani J, Li J, Sutter J, Stevens A, Bertrand HO, Rognan D. Protein-ligand-based pharmacophores: generation and utility assessment in computational ligand profiling. J Chem Inf Model 2012;23:943–55.
[40] Yang L, Luo H, Chen J, Xing Q, He L. SePreSA: a server for the prediction of populations susceptible to serious adverse drug reactions implementing the methodology of a chemical-protein interactome. Nucleic Acids Res 2009;37:W406–12.
[41] Morris RJ, Najmanovich RJ, Kahraman A, Thornton JM. Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons. Bioinformatics 2005;21:2347–55.
[42] Najmanovich R, Kurbatova N, Thornton J. Detection of 3D atomic similarities and their use in the discrimination of small molecule protein-binding sites. Bioinformatics 2008;24:105–11.
[43] Hoffmann B, Zaslavskiy M, Vert JP, Stoven V. A new protein binding pocket similarity measure based on comparison of clouds of atoms in 3D: application to ligand prediction. BMC Bioinformatics 2010;11:99.
[44] Yamanishi Y, Pauwels E, Saigo H, Stoven V. Extracting sets of chemical substructures and protein domains governing drug-target interactions. J Chem Inf Model 2011;51:1183–94.
[45] Pérot S, Sperandio O, Miteva MA, Camproux AC, Villoutreix BO. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. Drug Discov Today 2010;15:656–67.
[46] Junaid M, Lapins M, Eklund M, Spjuth O, Wikbeg JE. Proteochemometric modeling of the susceptibility of mutated variants of the HIV-1 virus to reverse transcriptase inhibitors. PLoS One 2010;5:e14353.
[47] Dakshanamurthy S, Issa NT, Assefnia S, Seshasayee A, Peters OJ, Madhavan S, et al. Predicting new indications for approved drugs using a proteochemometric method. J Med Chem 2012;55:6832–48.
[48] Van Westen GJ, Overington JP. A ligand's-eye view of protein similarity. Nat Methods 2013;10:116–17.
[49] Lapinsh M, Prusis P, Uhlén S, Wikberg JE. Improved approach for proteochemometrics modeling: application to organic compound—amine G protein-coupled receptor interactions. Bioinformatics 2005;21:4289–96.
[50] Weill D, Rognan D. Development and validation of a novel protein-ligand fingerprint to mine chemogenomic space: application to G protein-coupled receptors and their ligands. J Chem Inf Model 2009;49:1049–62.
[51] Pérot S, Regad L, Reynès C, Spérandio O, Miteva MA, Villoutreix BO, et al. Insights into an original pocket-ligand pair classification: a promising tool for ligand profile prediction. PLoS One 2013;8:e63730.
[52] Keith CT, Borisy AA, Stockwell BR. Multicomponent therapeutics for networked systems. Nat Rev Drug Discov 2005;4:71–8.
[53] Rasmussen K. Creating more effective antidepressants: clues from the clinic. Drug Discov Today 2006;11:623–31.
[54] Zimmerman GR, Lehár J, Keith CT. Multi-target therapeutics: when the whole is greater than the sum of the parts. Drug Discov Today 2007;12:34–42.

[55] Fitzgerald JB, Schoeberl B, Nielsen UB, Sorger PK. Systems biology and combination therapy in the quest for clinical efficacy. Nat Chem Biol 2006;2:458–66.
[56] Barabási AL, Oltvai Z. Network Biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13.
[57] Csermely P, Korcsmáros T, Kiss HJ, London G, Nussinov R. Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review. Pharmacol Ther 2013;138:333–408.
[58] Furlong LI. Human diseases through the lens of network biology. Trends Genet 2013;29:150–9.
[59] Sardiu ME, Washburn MP. Building protein–protein interaction networks with proteomics and informatics tools. J Biol Chem 2011;286:23645–51.
[60] Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. Proc Natl Acad Sci USA 2007;104:8685–90.
[61] Dalgaard MD, Weinhold N, Edsgärd D, Silver JD, Pers TH, Nielsen JE, et al. A genome-wide association study of men with symptoms of testicular dysgenesis syndrome and its network biology interpretation. J Med Genet 2012;49:58–65.
[62] Yildirim MA, Goh KI, Cusick ME, Barabási AL, Vidal M. Drug-target network. Nat Biotechnol 2007;25:1119–26.
[63] Hopkins AL. Network pharmacology: the next paradigm in drug discovery. Nat Chem Biol 2008;4:682–90.
[64] Berger SI, Iyengar R. Network analyses in systems pharmacology. Bioinformatics 2009;25:2466–72.
[65] Taboureau O, Jørgensen F. In silico predictions of hERG channel blockers in drug discovery: from ligand-based and target-based approaches to systems chemical biology. Comb Chem High Throughput Screen 2011;14:375–87.
[66] Lage K, Greenway SC, Rosenfeld JA, Wakimoto H, Gorham JM, Segré AV, et al. Genetic and environmental risk factors in congenital heart disease functionally converge in protein networks driving heart development. Proc Natl Acad Sci USA 2012;109:14035–40.
[67] Yilmaz S, Jonveaux P, Bicep C. Gene-disease relationship discovery based on model-driven data integration and database view definition. Bioinformatics 2009;25:230–6.
[68] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 2013;41:D991–5.
[69] Rustici G, Kolesnikov N, Brandizi M, Burdett T, Dylag M, Emam I, et al. ArrayExpress update—trends in database growth and links to data analysis tools. Nucleic Acids Res 2013;41:D987–90.
[70] Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, et al. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science 2006;313:1929–35.
[71] Ganter B, Snyder RD, Halbert DN, Lee MD. Toxicogenomics in drug discovery and development: mechanistic analysis of compound/class-dependent effects using the DrugMatrix database. Pharmacogenomics 2006;7:1025–44.
[72] Uehara T, Ono A, Maruyama T, Kato I, Yamada H, Ohno Y, et al. The Japanese toxicogenomics project: application of toxicogenomics. Mol Nutr Food Res 2010;54:218–27.
[73] Iskar M, Campillos M, Kuhn M, Jensen LJ, Van Noort V, Bork P. Drug-induced regulation of target expression. PLoS Comput Biol 2010;6:e1000925.
[74] Emig D, Ivliev A, Pustovalova O, Lancashire L, Bureeva S, Nikolsky Y, et al. Drug target prediction and repositioning using an integrated network-based approach. PLoS One 2013;8:e60618.
[75] Baron D, Dubois E, Bihouee A, Teusan R, Steenman M, Jourdon P, et al. Meta-analysis of muscle transcriptome data using the MADMuscle database reveals biologically relevant gene patterns. BMC Genomics 2011;12:113.
[76] Kiyosawa N, Ando Y, Manabe S, Yamoto T. Toxicogenomics biomarkers for liver toxicity. J Toxicol Pathol 2009;22:35–52.
[77] Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonasson K, Zwahlen M, et al. Towards a knowledge-based human protein atlas. Nat Biotechnol 2010;28:1248–50.
[78] Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, et al. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci USA 2004;101:6062–7.
[79] Pan J, Hu S, Shi D, Cai M, Li Y, Zou Q, et al. PaGenBase: a pattern gene database for the global and dynamic understanding of gene function. PLoS One 2013;8:e80747.
[80] Börnigen D, Pers TH, Thorrez L, Huttenhower C, Moreau Y, Brunak S. Concordance of gene expression in human protein complexes reveals tissue specificity and pathology. Nucleic Acids Res 2013;41:e171.
[81] McCall MN, Uppal K, Jaffee HA, Zilliox MJ, Irizarry RA. The gene expression barcode: leveraging public data repositories to begin cataloging the human and murine transcriptomes. Nucleic Acids Res 2011;39:D1011–15.
[82] Camon E, Magrane M, Barrell D, Lee V, Dimmer E, Maslen J, et al. The gene ontology annotation (GOA) database: sharing knowledge in Uniprot with gene ontology. Nucleic Acids Res 2004;32:D262–6.
[83] Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, et al. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res 2006;34:D354–7.
[84] Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, et al. Reactome: a knowledgebase of biological pathways. Nucleic Acids Res 2005;33:D428–32.
[85] Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM): a knowledgebase of human genes and genetic disorders. Nucleic Acids Res 2005;33:D514–17.
[86] Fernald GH, Capriotti E, Daneshjou R, Karczewski KJ, Altman RB. Bioinformatics challenges for personalized medicine. Bioinformatics 2011;27:1741–8.
[87] Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2011;39:38–51.
[88] Kalow W, Tang BK, Endrenyi L. Hypothesis: comparisons of inter- and intra-individual variations can substitute for twin studies in drug research. Pharmacogenetics 1998;8:283–9.
[89] Kirk RJ, Hung JL, Horner SR, Perez JT. Implications of pharmacogenomics for drug development. Exp Biol Med 2008;233:1484–97.
[90] Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, et al. The diploid genome sequence of an individual human. PLoS Biol 2007;5:e254.

[91] Szarfman A, Tonning JM, Doraiswamy PM. Pharmacovigilance in the 21st century: new systematic tools for an old problem. Pharmacotherapy 2004;24:1099–104.
[92] Vermeer NS, Straus SM, Mantel-Teeuwisse AK, Domergue F, Egberts TC, Leufkens HG, et al. Traceability of biopharmaceuticals in spontaneous reporting systems: a cross-sectional study in the FDA Adverse Event Reporting System (FAERS) and EudraVigilance. Drug Saf 2013;36:617–25.
[93] Shiokawa T. Background, introduction, and activity of the Japan Primary Registries Network. J Evid Based Med 2009;2:41–3.
[94] US National Library of Medicine, National Institutes of Health. Medical subject headings, <http://www.nlm.nih.gov/mesh/>.
[95] Wang L, Jiang G, Li D, Liu H. Standardizing drug adverse event reporting data. Stud Health Technol Inform 2013;192:1101.
[96] Hohl CM, Karpov A, Reddekopp L, Stausberg J. ICD-10 codes used to identify adverse drug events in administrative data: a systematic review. J Am Med Inform Assoc 2014;21:547–57. Available from: http://dx.doi.org/10.1136/amiajnl-2013-002116.
[97] Wasserman H, Wang J. An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list. AMIA Annu Symp Proc 2003;2003:699–703.
[98] WHO Collaborating Centre for Drug Statistic Methodology. ATC/DDD Index, <http://www.whocc.no/atc_ddd_index/>; 2015.
[99] Gabetta M, Larizza C, Bellazzi R. A unified medical language system (UMLS) based system for literature-based discovery in medicine. Stud Health Technol Inform 2013;192:412–16.
[100] Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, et al. Drugbank 4.0: shedding new light on drug metabolism. Nucleic Acids Res 2014;42(1):D1091–7.
[101] Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol 2010;6:343.
[102] Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. Sci Transl Med 2012;4:125ra31.
[103] Cheng F, Li W, Wang X, Zhou Y, Wu Z, Shen J, et al. Adverse drug events: database construction and in silico prediction. J Chem Inf Model 2013;53:744–52.
[104] Hopkins AL, Groom CR. The drugable genome. Nat Rev Drug Discov 2002;1:720–30.
[105] Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaokar P, Ferriero R, et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci USA 2010;107:14621–6.
[106] Köhler S, Bauer S, Horn D, Robinson PN. Walking the interactome for prioritization of candidate disease genes. Am J Hum Genet 2008;82:949–58.
[107] Von Eichborn J, Murgueitio MS, Dunkel M, Koerner S, Bourne PE, Preissner R. PROMISCUOUS: a database for network-based drug-repositioning. Nucleic Acids Res 2011;39:D1060–6.
[108] Iskar M, Zeller G, Blattmann P, Campillos M, Kuhn M, Kaminska KH, et al. Characterization of drug-induced transcriptional modules: towards drug repositioning and functional understanding. Mol Syst Biol 2013;9:662.
[109] Piscitelli SC, Gallicano KD. Interactions among drugs for HIV and opportunistic infections. N Engl J Med 2001;344:984–96.
[110] Ioro F, Tagliaferri R, di Bernardo D. Identifying network of drug mode of action by gene expression profiling. J Comput Biol 2009;16:241–51.
[111] Perlman L, Gottlieb A, Atias N, Ruppin E, Sharan R. Combining drug and gene similarity measures for drug-target elucidation. J Comput Biol 2011;18:133–45.
[112] Hennessy S, Flockhart DA. The need for translational research on drug-drug interactions. Clin Pharmacol Ther 2012;91:771–3.
[113] Fichtenbaum CJ, Gerber JG. Interactions between antiretroviral drugs and drugs used for the therapy of the metabolic complications encountered during HIV infection. Clin Pharmacokinet 2002;41:1195–211.
[114] Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, Castro V, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther 2011;90:132–42.
[115] Boverhof DR, Zacharewski TR. Toxicogenomics in risk assessment: applications and needs. Toxicol Sci 2006;89:352–60.
[116] Wilson VS, Keshava N, Hester S, Segal D, Chiu W, Thompson CM, et al. Utilizing toxicogenomics data to understand chemical mechanism of action in risk assessment. Toxicol Appl Pharmacol 2013;271:299–308.
[117] Schreiber G, Keating AE. Protein binding specificity versus promiscuity. Curr Opin Struct Biol 2011;21:50–61.
[118] Scheiber J, Jenkins JL, Sukuru SC, Bender A, Mikhailov D, Milik M, et al. Mapping adverse drug reactions in chemical space. J Med Chem 2009;52:3103–7.
[119] Bender A, Scheiber J, Glick M, Davies JW, Azzaoui K, Hamon J, et al. Analysis of pharmacology data and the prediction of adverse drug reactions and off-target effects from chemical structure. Chem Med Chem 2007;2:861–73.
[120] Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins J, et al. Large scale prediction and testing of drug activity on side effect targets. Nature 2012;486:361–7.
[121] Cheng F, Yu Y, Zhou Y, Shen Z, Xiao W, Liu G, et al. Insights into molecular basis of cytochrome p450 inhibitory promiscuity of compounds. J Chem Inf Model 2011;51:2482–95.
[122] Campillos M, Kuhn M, Gavin AC, Jensen LJ, Bork P. Drug target identification using side-effect similarity. Science 2008;321:263–6.
[123] Gautier L, Taboureau O, Audouze K. The effect of network biology on drug toxicology. Expert Opin Drug Metab Toxicol 2013;9:1409–18.
[124] Bai JPF, Abernethy DR. Systems pharmacology to predict drug toxicity: integration across levels of biological organization. Annu Rev Pharmacol Toxicol 2013;53:451–73.
[125] Oprea TI, Nielsen SK, Ursu O, Yang JJ, Taboureau O, Mathias SL, et al. Associating drugs, targets, and clinical outcomes into an integrated network affords a new platform for computer-aided drug repurposing. Mol Inform 2011;30:100–1.
[126] Wallach I, Jaitly N, Lilien R. A structure-based approach for mapping adverse drug reactions to the perturbation of underlying biological pathways. PLoS One 2010;5:e12063.
[127] Brouwers L, Iskar M, Zeller G, van Noort V, Bork P. Network neighbors of drug targets contribute to drug side-effect similarity. PLoS One 2011;6:e22187.

[128] Xie L, Li J, Xie L, Bourne PE. Drug discovery using chemical systems biology: identification of the protein-ligand binding network to explain the side effects of CETP inhibitors. PLoS Comput Biol 2009;5:e1000387.
[129] Nobeli I, Favia AD, Thornton JM. Protein promiscuity and its implications for biotechnology. Nat Biotechnol 2009;27:157–67.
[130] Bresso E, Grisoni R, Marchetti G, Karaboga AS, Souchet M, Devignes MD, et al. Integrative relational machine-learning approach for understanding drug side-effect profiles. BMC Bioinformatics 2013;14:207.
[131] Chen L, Zhang L, Zhao Y, Xu L, Shang Y, Wang Q, et al. Prioritizing risk pathways: a novel association approach to searching for disease pathways fusing SNPs and pathways. Bioinformatics 2009;25:237–42.
[132] O'Dushlaine C, Kenny E, Heron EA, Segurado R, Gill M, Morris DW, et al. The SNP ratio test: pathway analysis of genome-wide association datasets. Bioinformatics 2009;25:2762–3.
[133] Schlosberg CE, Schwantes-An TH, Duan W, Saccone NL. Application of Bayesian network structure learning to identify causal variant SNPs from resequencing data. BMC Proc 2011;5:S109.

SECTION FOUR

Substituents and Functions

CHAPTER

18

Optical Isomerism in Drugs

Camille G. Wermuth

Founder of Prestwick Chemical, Strasbourg, France

OUTLINE

Most natural organic compounds, the essential products of life, are asymmetric and possess such asymmetry that they are not superposable on their images ... This establishes perhaps the only well marked line of demarcation that at present can be drawn between the chemistry of dead matter and the chemistry of living matter.

(Pasteur, Van t'Hoff, Le Bel, and Wislicenus, Memoirs, 1901)

I. INTRODUCTION

This chapter is concerned with bioactive compounds bearing on their skeleton one or more asymmetric carbon atom(s). For such compounds the term "configuration" defines the implantation mode of the four covalent linkages on the central, asymmetric, carbon atom. The terms "optical isomers," "optical antipodes," "enantiomorphs," or "enantiomers" are synonyms and relate to molecules that are mirror images of each other and are not, therefore, superimposable. Owing to their non-identical 3D structure, enantiomers may elicit differentiated biological responses and thus provide useful information on drug–receptor interactions and on receptor characteristics. A great number of books and publications deal with chirality and drug design [1–9].

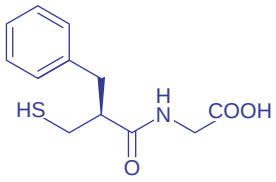
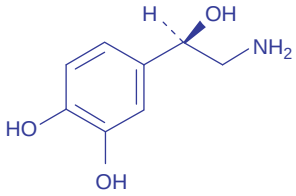
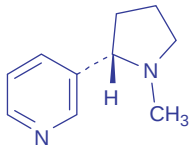
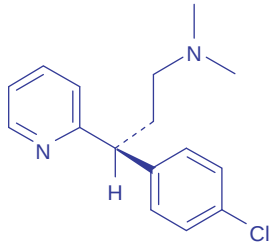
II. EXPERIMENTAL FACTS AND THEIR INTERPRETATION

A. Stereoselectivity in Biologically Active Compounds

Toward a biological target, the potency of two enantiomers can sometimes differ considerably and sometimes be very similar (Table 18.1). Often the activity is concentrated in only one enantiomer. When such a high stereoselectivity arises, it is admitted that the mechanism of action at the molecular level involves a highly specific interaction between the ligand, a chiral molecule, and the recognition site, a chiral environment. It is to be expected that the most active isomer, in terms of affinity, achieves a better steric complementarity to the receptor than the less active one.

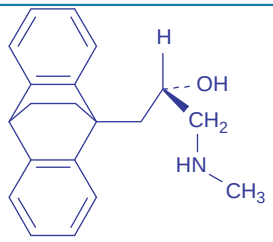
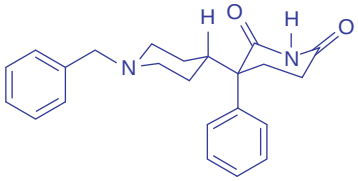
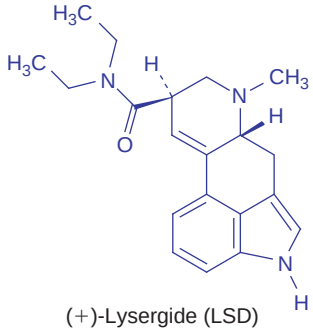
When considering *in vivo* activities, the difference in activity observed for the two enantiomers is neither always nor exclusively the result of the quality of the ligand–receptor fit. It must be kept in mind that *in vivo*, the pharmacokinetic processes (ADME) may account for the observed difference in activity. The interpretation of pharmacological data obtained from *in vivo* assays should thus be questioned and does not allow anticipating the quality of the ligand–receptor interaction.

TABLE 18.1 Differences in Activity or in Affinity (eudismic index) between Couples of Enantiomers

SUBSTANCE		EUDISMIC INDEX
 S-Thiorphan	K_i values for inhibition of neutral endopeptidase 24.11 (former enkephalinase)	S/R = 1,2 [10]
 R(-)-Noradrenaline [R(-)-Norepinephrine] natural enantiomer	Adrenergic activity at rat aorta α_1 sites	R(–)/S(+) = 33 [11]
 S(-)-Nicotine (Natural nicotine)	Affinity constants for rat brain thalamus sites labeled by [3H]-(–)-nicotine	S(–)/R(+) = 35 [12]
 S(+)-Chlorpheniramine (Polaramine)	K_i values for human brain frontal lobe sites labeled by [3H]-mepyramine	S(+)/R(–) = 83 [13]

(*Continued*)

TABLE 18.1 (Continued)

SUBSTANCE		EUDISMIC INDEX
 S(+)-Oxaprotiline	IC_{50} values for noradrenaline uptake into rat brain synaptosomes	S(+)/R(−) = 1000 [14]
 S(+)-Dexetimide	[^{3}H]-dexetimide binding to brain muscarinic receptors	S(+)/R(−) = 2000 [15]
 (+)-Lysergide (LSD)	(+)-[3H] lysergic acid diethylamide binding on rat forebrain suspension	R(+)/S(−) = 24000 [16]

B. The Three-Point Contact Model

When only one asymmetric center is present in the molecule in a compound exhibiting stereoselectivity, it is thought that the substituents on the chiral carbon atom make a three-point contact with the receptor. Such a fit ensures a very specific molecular orientation that can only be obtained for one of the two isomers. A three-point fit of this type was first suggested by Easson and Stedman [17], and the corresponding model was proposed by Beckett [18] in the case of R(−)-adrenaline (=R(−)-epinephrine). The more active natural R(−)-adrenaline establishes contacts with its receptor through the following three interactions (Figure 18.1):

1. acceptor−donor or hydrophobic interaction between the aromatic ring of adrenaline and an aromatic ring of the receptor protein;
2. a hydrogen bond at the alcoholic hydroxyl; and
3. an ionic bond between the protonated amino group and an aspartic or glutamic carboxylic group of the receptor.

The combination of these interactions can generate binding energies in the order of 12−17 kcal/M, corresponding to binding constants in the order of 10^{-9}−10^{-12} M [19]. The biologically weak optical isomer, S(+)-adrenaline, can make contact through only two groups. According to this hypothesis, it would be anticipated that deoxyadrenaline (epinine) should have much the same activity as S(+)-adrenaline. This has been found to be basically true [17,20,21].

Computer-generated receptor models for protein-G linked receptors are now available [22,23], and Figure 18.2 illustrates the fit of R(−)-adrenaline into the active site of the β_2-adrenergic receptor. It clearly appears that the docking involves more interactions than only the above-mentioned three points:

1. The two phenolic hydroxyl groups exchange hydrogen bonds with Ser^{505} and Ser^{508} respectively.
2. The aromatic ring of adrenaline is stabilized by means of π-π interactions with Phe^{509} and Phe^{617}.

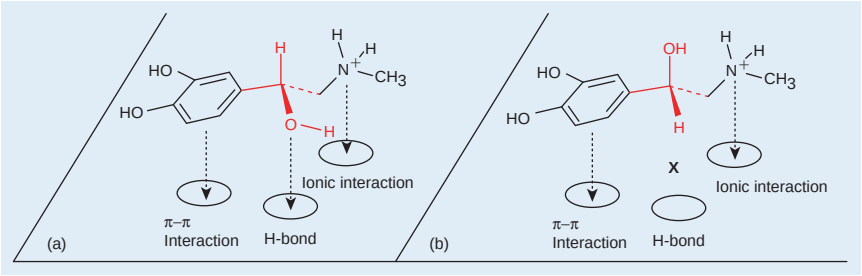


FIGURE 18.1 Interaction capacities of the natural R(−)-epinephrine and its S(+) antipode. In simply assuming that the natural R(−)epinephrine establishes a three-point interaction with its receptor (a), the combination of the donor−acceptor interaction, the hydrogen bond, and the ionic interaction will be able to generate energies in the order of 12−17 kcal/mol, which corresponds to binding constants of 10^{-9} to 10^{-12} M[18] The less active isomer, S(+) epinephrine, may establish only a two-point contact (b). The loss of the hydrogen bond interaction equals approximately 3 kcal/mol. This isomer should therefore possess an approximately 100-fold lesser affinity. Experience confirms this estimate. If we consider less abstract models, it becomes apparent that the less potent enantiomer is also able to develop three intermolecular bonds to the receptor, provided that it approaches the receptor in a different manner. However, the probability of this alternate binding mode triggering the same biological response is close to null.

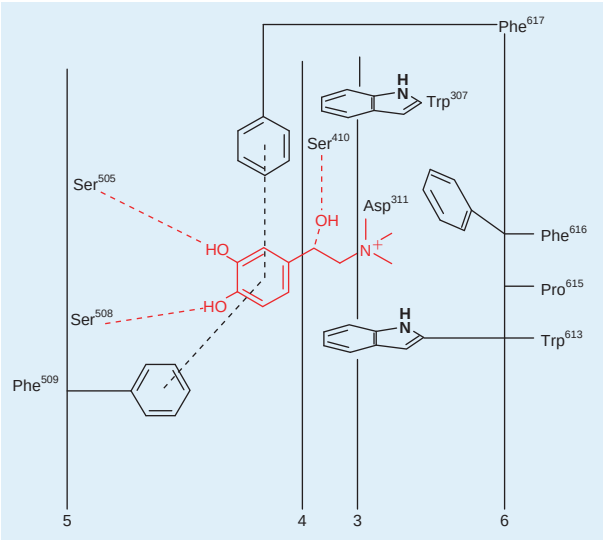


FIGURE 18.2 Interaction capacities of the natural R(−) epinephrine with a model of its receptor (after references [22,23]).

3. The cationic head exerts a coulombic interaction with the Asp^{311} carboxylate and is located in a hydrophobic pocket made of Trp^{307}, Phe^{616}, and Trp^{613}.
4. Finally, the secondary benzylic hydroxyl exchanges a hydrogen bond with Ser^{410}.

Even taking into account these newer findings, it can be speculated that the Easson−Stedman hypothesis still holds. The nonnatural S(+)-adrenaline, having the wrong orientation of its benzylic hydroxyl, is unable to exchange a hydrogen bond with Ser^{410}, and achieves therefore a weaker interaction with the receptor.

An alternative model of the adrenergic receptor-active site shows that natural R-(−) epinephrine (adrenaline) can establish a hydrogen bond with the Ser^{410} alcoholic group, whereas this interaction is not possible with the nonnatural S-(+) epinephrine (Figure 18.3).

C. Diastereoisomers

When more than one asymmetric center is involved, the complexity of the problem increases rapidly (Figure 18.5). For the four isomers of ephedrine, which represent a set of diastereoisomers, the R configuration of the β-carbon (as found for adrenaline, noradrenaline, nordefrin, phenylephrin, and octopamine; see Patil et al [24]) is not automatically associated with the highest alpha-agonistic activity. Both (−)-ephedrine and

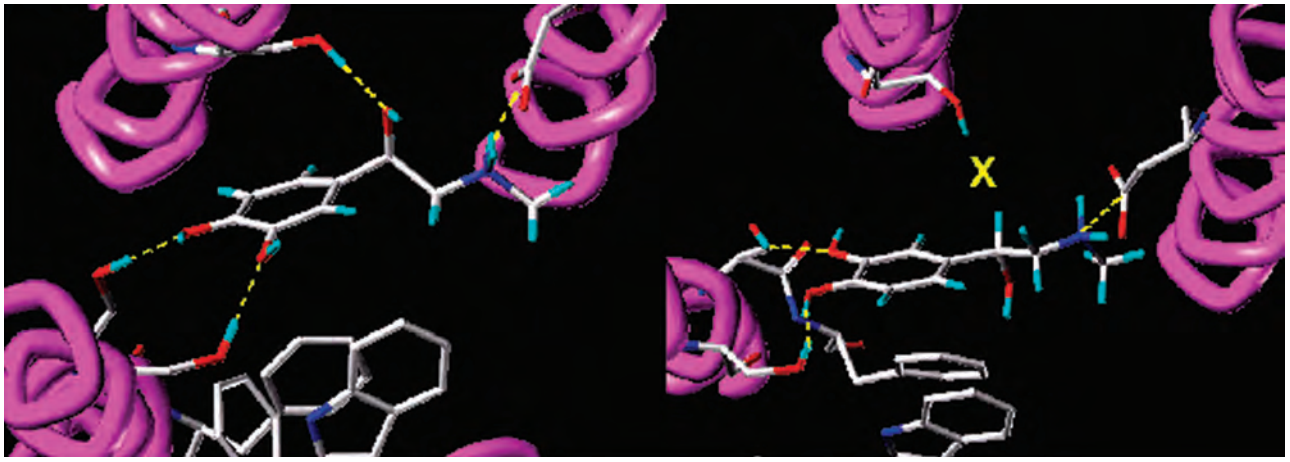


FIGURE 18.3 In an adrenergic receptor model the interaction of the natural R-(−) epinephrine (left) implies an hydrogen bond with the Ser ^{410} alcoholic group.

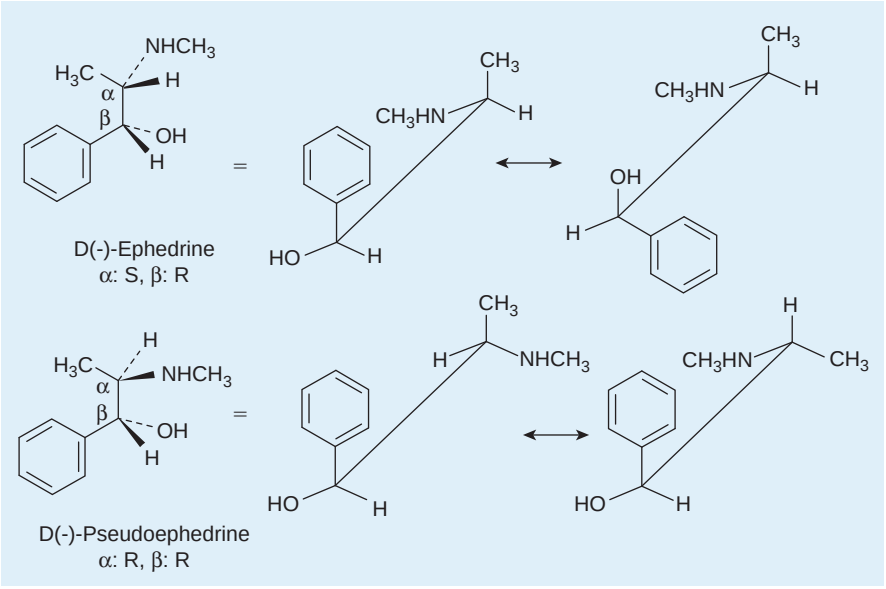


FIGURE 18.4 Preferred conformations of D-(−)-ephedrine and of D-(−)-pseudo-ephedrine.

(−)-pseudo-ephedrine possess the β-(R) configuration, yet only (−)- ephedrine acts as an agonist. This anomaly can be explained if one takes into account the preferred conformations of these two compounds, calculated by using the ETH (Extended Hückel Theory) [25,26]. In the (−)-ephedrine molecule, the methyl group attached to the carbon in alpha-position to the amino function is projected above the plane of the phenyl-ethyl-amino group, whereas in (−)-pseudo-ephedrine the methyl group is oriented below the plane and thus prevents an efficient interaction of the drug with the receptor (Figure 18.4).

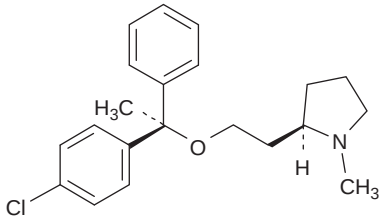
The antibacterial activity of chloramphenicol isomers represents a similar example. Significant activity is only found for the (−)-threo-chloramphenicol [27]. The clinical formulation of the adrenergic receptor-blocking agent labetalol consists of a mixture of equal proportions of the four optical isomers (RR, SS, RS, and SR). Each possesses different pharmacological properties. The most active RR enantiomer was developed some years ago as dilevalol [28], but had to be withdrawn after some months due to a slightly higher than average degree of hepatic toxicity [29].

Despite the fact that it contains two chiral centers, the antihistaminic drug clemastine (Tavegyl) provides one of the few examples of chiral antihistamines employed clinically in the form of a single isomer. Data on the antihistaminic activity of clemastine and its isomers [30] are summarized on Table 18.2.

(a) (b)

FIGURE 18.5 An artistic illustration of the RR/SS relationship in diastereoisomers. Auguste Rodin's famous sculpture "The Cathedural" (image of (a)) represents two right hands. Its mirror image (image of the (b)) represents two left hands. The images symbolize the RR versus SS relationship as found in diastereoisomers.

TABLE 18.2 Antihistamine Activities of Clemastine and its Isomers [30]



R, R-Clemastine

Isomer	Prevention of histamine toxicity ED_{50} (mg/kg s)	Prevention of histamine spasm	pA_2
RR (clemastine)	0.04	~ +7	9.45
SS	5.1	~ −1.5	7.99
SR	11.0	~ −6	8.57
RS	0.28	~ +5	9.40

The stereoisomers of some oxotremorine analogs containing two chiral centers and acting as oxotremorine antagonists show *in vivo* (tremorolytic activity) stereoselectivity ratios as high as 1:200 [31].

D. Stereoselectivity Ratios

Stereoselectivity was defined by Rauws [32] as follows: "Stereoselectivity is the extent to which an enzyme or other macromolecule, or macromolecular structure (antibody or receptor) exhibits affinity towards one molecule of a pair of isomers in comparison with and in contrast to the other isomer." Lehmann [33,34] has expressed this in mathematical form: the ratio of activity of the better fitting enantiomer (eutomer; Greek, "eu" = good) to that of the less fitting enantiomer (distomer; Greek, "dys" = bad) is defined as the eudismic ratio. From this, a eudismic affinity quotient can be derived (Table 18.3).

In a series of agonists or antagonists one can write Equation 18.1:

$$\mathrm{EI} = a + b\mathrm{Log}\ \ \mathrm{Affin}.Eu \tag{18.1}$$

in which a is a constant, b is the quotient of eudismic affinity (QEA) that precisely accounts for the stereoselectivity.

TABLE 18.3 Nomenclature and Definitions in Drug Stereoselectivity

Eutomer (*Eu*): Enantiomer presenting the *highest* affinity (or activity)
Distomer (*Dis*): Enantiomer presenting the *lowest* affinity (or activity)
Eudismic quotient: $\frac{\text{Affin.}Eu}{\text{Affin.}Dis}$
Eudismic index (*EI*): Log Affin. $_{Eu}$-Log Affin.$_{Dis}$

When the activity of the eutomer "Eu" is compared to that of the racemic mixture "Rac," four possibilities can arise: [35,36]

1. The activity ratio is equal to 2: Eu/Rac = 2/1. In this case the activity is only concentrated in the eutomer and the distomer does not contribute significantly to the observed activity. The chiral compound shows stereoselectivity.
2. The activity ratio is higher than 2: Eu/Rac >2 (e.g., Eu/Rac = 2/0.3). This means that the distomer represents a competitive antagonist of the eutomer. In practice, such a situation is rather exceptionally encountered. See section "V. Practical Considerations; B. The distomer counteracts the eutomer."
3. If the activity ratio is lower than 2: Eu/Rac <2 (e.g., Eu/Rac = 2/1.6), we are in the presence of two active isomers. The distomer reinforces the activity of the eutomer. Such a situation indicates a decrease of the receptor selectivity.
4. The activity ratio is Eu/Rac = 1. In this case both isomers are equipotent and no stereoselectivity is observed. This can be explained by the assumptions: (a) that the compounds act through a nonspecific mechanism; (b) that the active compound and the receptor make only a two-point contact with the chiral center; and (c) that the chiral center is not involved in the contact (is located in a "silent region").

E. Pfeiffer's Rule

One usually admits that the discriminative effect between the two enantiomers increases with the proximity of the chiral center to the site of interaction with the receptor. An empirical rule published by Pfeiffer in 1956 [37] states that the isomeric activity ratio (eudismic quotient) of a highly active couple of isomers is always superior to that of a less active couple. In other words, "the greater the difference between the pharmacological activity of the R and the S isomers, the greater is the potency of the active isomer." However, there are some exceptions to Pfeiffer's rule. Some of the reasons are conformational flexibility of the ligands [38]. Others reside in an improper selection of "homologous" sets of compounds as illustrated with muscarinic agonists and antagonists [39]. Quantitative analyses of the correlations between biological activity and the structure of stereoisomeric compounds are difficult [40,41].

III. OPTICAL ISOMERISM AND PHARMACODYNAMIC ASPECTS

The biological response induced by a pair of enantiomers can differ in potency (quantitative difference) or in nature (qualitative difference). In the latter case, it is assumed that one enantiomer acts at one receptor site, whereas its antipode is recognized by other sites and possesses a different activity and toxicity profile.

A. Differences in Potency and Antagonism between Two Enantiomers

Two optical isomers are never antagonists, at least at comparable dosages. This comes from the space relationship required for the interaction with the receptor site, which is only slightly altered by passing from S to R forms or vice-versa. If one of the enantiomers achieves the optimal fit to the receptor site in exchanging the highest number of noncovalent linkages, its antipode can only give rise to a weaker interaction, even in the most favorable conditions (Figure 18.1). From a practical point of view, this absence of stoechiometric antagonism entails two consequences:

1. If a racemic mixture does not show any activity, it is useless to carry out the separation of the two antipodes.

TABLE 18.4 Differences in Pharmacological Profile of Couples of Enantiomers

Racemate	Levorotatory enantiomer	Dextrorotatory enantiomer	References
Quinine/quinidine (racemate not in use)	Quinine: antipyretic, antimalarial	Quinidine: antiarrythmic antimalarial	White et al, [44] White et al, [45] Alexander et al [46]
Sotalol	(−)-Sotalol β-adrenoceptor blocker	(+)-Sotalol antiarrythmic agent	Drayer [28]
Racemorphane	(−)-*N*-Methyl-3-methoxy- morphinane antitussive	(+)-*N*-Methyl-3-methoxy-morphinane analgesic	Benson et al [47]
Indacrinone	R(−)-Indacrinone diuretic	S(+)-Indacrinone uricosuric	Drayer [28]
Propoxyfene	α-Levopropoxyfene (Novrad) antitussive	α-Dextropropoxyfene (Darvon) analgesic	Drayer [28]
Tetramisole	S(−)-Levamisole: nematocidal, immunostimulant	R(+)-Dexamisole: antidepressant	Bullock et al, [48] Schnieden [49]
3-Amino-1-hydroxypyrrolid- 2-one (HA-966)	3R-(+)-HA-966 partial agonist at the glycine site of the NMDA receptor	3S-(−)-HA-966 γ-butyrolactone-like sedative	Singh et al [50]
3-Methoxy-cyproheptadine	(−)-3-Methoxy-cyproheptadine anticholinergic activity	(+)-3-Methoxy-cyproheptadine antiserotonin activity	Remy et al [51]

2. A racemic mixture usually has the average potency of both constituents. Thus, the maximal benefit in resolving racemic mixtures is achieved by increasing the activity of one of the antipodes to twice of that of the racemate.

B. Differences in the Pharmacological Profile of Two Enantiomers

Besides the difference in potency, it often happens that two enantiomers show differences in their pharmacological profile [42,43]. In such a case, resolving the racemic mixture can generate two pharmacologically different and useful compounds, and can also separate the more active compound from its less well tolerated or more toxic isomer. In the quinine−quinidine couple (Table 18.4), both isomers share antimalarial, antipyretic, and oxytoxic, as well as skeletal and cardiac muscle depressant activities. However, whereas antipyresis and treatment of malaria represent the main use of quinine, quinidine is more effective on the cardiac muscle and is used in the therapy of atrial fibrillation and in certain other arrhythmias [52]. In the *N*-methyl-3-methoxy morphinane racemate (racemorphane), most of the analgesic and addictive properties are concentrated in the (+)-isomer. The corresponding (−)-isomer is non-addictive and retains only antitussive properties [47]. The same kind of discrimination is found for the antitussive levopropoxyphene and its well-known analgesic enantiomer dextropropoxyphene [28].

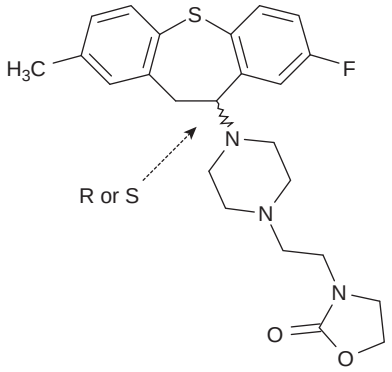
The substituted imidazo-thiazole, dexamisole, has antidepressant properties, and its isomer, levamisole, possesses anthelmintic and immunostimulant properties [48,49]. Enantiomers of HA-966 (3-amino-1-hydroxypyrrolid-2-one) exhibit distinct central nervous system effects: (+)-HA-966 is a selective glycine/*N*-methyl-D-aspartate receptor antagonist, but (−)-HA-966 is a potent γ-butyrolactone-like sedative [50]. A comparison of (+) and (−)-3-methoxycyproheptadine shows that all of the anticholinergic activity of the (±)-3-methoxycyproheptadine resides solely in the dextrorotatory enantiomer, while the antiserotonin activity resides in the levorotatory enantiomer [51].

Table 18.5 shows some experimental data for the active isomers of the fluoro analogs of the tricyclic neuroleptic clotepin as compared to the corresponding racemate [53]. In this example, it appears that the neuroleptic activity is concentrated in the dextrorotatory compound (+)**2**, whereas the toxicity resides in the (−)**3** antipode.

In the present case, the 1:10 therapeutic index of the racemate— unsatisfactory for a clinical outlook— was raised to the much more acceptable 1:50 ratio for the isolated S(+) antipode [53].

Primaquine enantiomers show differences in activity and toxicity in different species, suggesting that (−)-(*R*)-Primaquine may have a better safety margin than the racemate in human; [54] the lesser toxic effect of the (−)-enantiomer of primaquine may be due to its faster transformation to (−)-carboxyprimaquine [55].

TABLE 18.5 Comparison of the Racemate and the Two Enantiomers of a Fluoro Analogs of the Tricyclic Neuroleptic Clotepin [53]



Test	Measurements	RS(±) 1	S(+) 2	R(−)3
Increase in brain homovanillic acid	At 100 mg/Kg	256%	316%	128%
Adenylate-cyclase inhibition	$c = 10^{-6}$ M	48%	72%	27%
Inhibition of conditioned flight reflexes in rats	ED_{50} (mg/Kg) per os	14	10	>100
Inhibition of apomorphine- induced emesis in dogs	ED_{50} (mg/Kg) per os	20	12	>30
Acute toxicity in mice	LD_{50} (mg/Kg) per os	200	515	68

IV. OPTICAL ISOMERISM AND PHARMACOKINETIC ASPECTS

After administration and before it arrives in the vicinity of its receptor site, a drug is subjected to a variety of physiological processes: absorption, distribution, metabolism, uptake at storage sites, and excretion. Many of these processes are stereoselective. For reviews, see Jamali et al [56] and Kroemer et al [57].

A. Isomer Effects on Absorption and Distribution

The higher narcotic potency *in vivo* of the S(+)-isomer of hexobarbital was shown to be related to higher central nervous system levels than for the R(−) form. This seems to be due to a better crossing of the blood−brain barrier [58]. In a distribution study of [^{14}C] (+) and (−)-alpha-methyl-DOPA in rats after intravenous injection [59], the (−)-isomer attained higher concentrations than the (+)-form in most organs, which is reflective of the fact that of the two isomers, only the (−) isomer has hypotensive activity [60].

B. Isomer Effects on Metabolism

Since all enzymes are chiral in nature and therefore probably possess some degree of asymmetry at the reactive center, it is not surprising that most metabolic reactions of isomers lead to qualitative and quantitative differences in the metabolites formed. For review articles, see Testa [1,2], Vermeulen [61], Kroemer [57] and Campo [62].

1. *Differential Metabolism of Two Antipodes*

The levo isomers of 3-hydroxy-*N*-methyl-morphinan and of methadone are demethylated by rat liver 2−3 times more rapidly than the corresponding dextro antipodes [63,64]. The S(+)-enantiomer of hexobarbital (Figure 18.6) is metabolized almost twice as rapidly as the R(−)-enantiomer by allylic hydroxylation [65], and in dogs, the dextrorotatory isomer of 5-ethyl-5-phenyl-hydantoin affords ten times more para-hydroxy-metabolite than the levorotatory isomer [66]. Hydroxylation takes place alpha to a carbonyl in the dextrorotary enantiomer of glutethimide, whereas the levorotamer is hydroxylated on the methylene group of the ethyl side-chain [67]. Numerous other examples are found in the literature [68,69].

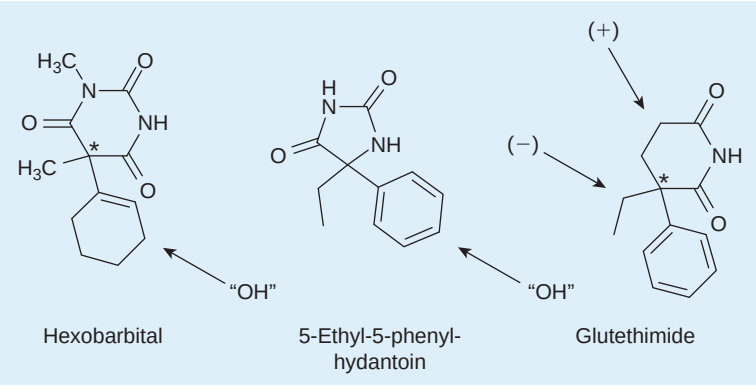


FIGURE 18.6 The structures of hexobarbital, 5-ethyl-5-phenyl-hydantoin, and glutethimide.

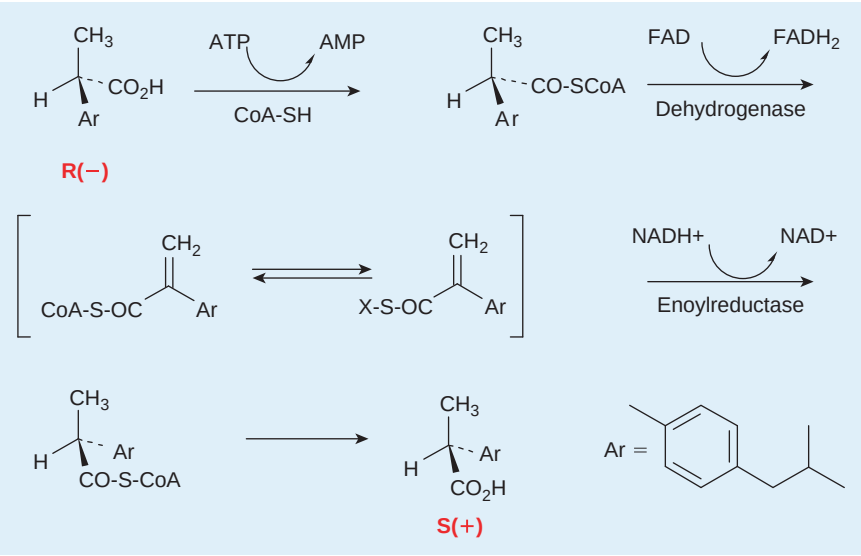


FIGURE 18.7 Mechanism of the enzymatic inversion of R(−)-ibuprofen [73].

2. Enzymatic Inversion

The energy requirements necessary for the conversion of a given sp^3 configuration into its optical antipode imply the formation of an intermediary carbenium ion, carbanion, or free radical, and are unlikely to arise in biological systems. Thus, racemization and epimerization involving non-oxygenated sp^3 carbon atoms are generally not encountered in mammals. They are usually restricted to microorganisms (e.g., alanine-racemase). One case of this unusual phenomenon is described in mammals for arylpropionic acids. More precisely, for the nonsteroidal anti-inflammatory agent ibuprofen (R,S-para-isobutyl-hydratropic acid), it has been demonstrated that only the S (+)-isomer is active *in vitro* as an inhibitor of the prostaglandin-synthesizing enzyme cyclo-oxygenase. Surprisingly, no significant differences could be observed *in vivo* between the S(+) or the R(−)-enantiomers and the racemate (=ibuprofen) [70]. It was therefore concluded that *in vivo* there must be an almost complete inversion of the poorly active R(−) form to the much more active S(+)-isomer. In humans, the main metabolites isolated after administration of (racemic) ibuprofen were dextrorotatory [71], the R(−)-enantiomer being converted to the S(+) −isomer [72]. A biochemical investigation using deuterium-labeled R(−)-isomer led to the hypothesis of the existence of an R-arylpropionic acid isomerase (R-APAI) enzyme system proceeding via the enzymes of lipid catabolism and anabolism, as outlined in Figure 18.7.

It is assumed that the coenzyme A ester (CoA-ester) of the R(−)-enantiomer acts as a substrate for the fatty acid deshydrogenase, thus eliminating the chiral center. The next step may or may not take place, depending on whether or not the CoA-ester must be transferred to an acyl-carrier protein or another site in the fatty acid synthetase system, so that a stereoselective reduction by an enoylreductase can take place. Thus, the nature of X is unknown [73]. Similar epimerization reactions were also described for some other arylpropionic acids such as benoxaprofen [74], carprofen [75], and isopropyl-indanyl-propionic acid [76]. It was demonstrated that the

configural inversion does not take place in the liver and that the responsible enzyme, R-(−)-arylpropionic acid isomerase, is located in the gut wall [73,77].

C. Isomer Effects on Uptake

As drugs are usually absorbed by passive diffusion and since enantiomers do not differ in their aqueous and lipid solubilities, absorption is not usually considered to be a stereoselective process. However, stereoselectivity has been described for drugs that are transported by a carrier-mediated process [78,79]. Typical uptake selectivity is observed for neurotransmitter reuptake inhibitors such as nipecotic acid, oxaprotiline, fluoxetine, and venlafaxine. Uptake of drugs by various organs can also be enantioselective. For example, the liver/plasma concentration ratios of S(−) and R(+)-phenprocoumon in rats were found to be different (6.9 and 5.2, respectively), indicating a preferential uptake of the more potent isomer [80]. Enantioselective skin permeation can also be observed [81].

D. Isomer Effects on Excretion

The kinetics of excretion are a direct consequence of the kinetics of metabolic transformations. The faster a drug is metabolized, the faster its elimination can be expected. In accordance with this assertion, rats given R,S (±), S(+), and R(−)-amphetamine were found to excrete less (+)-*p*-hydroxy-amphetamine than its (−)-isomer. This may be the basic explanation of the more pronounced pharmacological properties of the dextro- compared to the levoamphetamine [82]. For the hypnotic agent hexobarbital, the elimination half-life in humans is about three times longer for the more active (+)-isomer than for the less active (−)-isomer. This was attributed to a difference in hepatic metabolic clearance and not in volumes of distribution or plasma binding between the enantiomers [83].

V. PRACTICAL CONSIDERATIONS

A. Racemates or Enantiomers?

Many drugs having a center of asymmetry are still used in clinical practice as racemates. Racemic mixtures were estimated to represent 10−15 percent of all marketed drugs [3]. For certain types of therapeutics, such as the β-adrenergic agents, β-adrenergic blockers, antiepileptics, and oral anticoagulants, up to 90 percent of the compounds are in fact racemic mixtures according to Ariëns [3]. For antihistaminics and local anesthetics, this holds true for about 50 percent of the drugs currently used [3]. Often, racemic drugs were introduced in clinical practice because the animal and clinical pharmacology, the toxicology, and the teratology were performed with the racemates. The reasons for that is that, at the time of the discovery of the drug, the resolution (or the chiral synthesis) appeared to be too difficult, too costly, or even impossible.

The question now is when and why to use rather racemic mixtures or pure enantiomers. Although it seems good sense to use pure eutomers and to consider the distomer as an unwanted load of xenobiotic (a kind of pollution, or even an impurity), there are instances where it is recommended to use racemates rather than eutomers.

Thus, racemates may be more stable, more active, or less toxic, or present a favorable combination of the properties of each separate isomer (see below). Finally, one can ask if it would not be wise to design effective drugs without centers of asymmetry.

B. The Distomer Counteracts the Eutomer

Contrary to a well-established belief, there are no examples of inactive racemates in which the distomer antagonizes the activity of the eutomer in a stoichiometric manner.

Thus, a dihydropyridine-derived calcium inhibitor, the R(−) enantiomer of compound Sandoz 202-791, inhibits the uptake of [$^{45}Ca^{++}$] with an IC_{50} of 4.3×10^{-8}, whereas its S(+) enantiomer increases the uptake with an IC_{50} of approximately 10^{-6}−10^{-7} M [84]. The corresponding racemic mixture inhibits the uptake with an IC_{50} of 1.7×10^{-7} M. Some other examples are reported in Table 18.6.

As shown in Table 18.6, a more or less important residual activity is always present in the racemate but resolution would generally be beneficial. Picenadol (LY 150720) seems to be an exception to the rule that pure eutomers

TABLE 18.6 Antagonism in Couples of Enantiomers

Compound	Eutomer	Distomer	Racemate	Reference
N-Isopropyl- norepinephrine	(−) α-Adrenergic agonist	(+) Inactive competitive antagonist	(±) Partial agonist	Page 15 of [3]
5-Ethyl-5(1,3- dimethylbutyl) barbituric acid	(+) Convulsant	(−) Depressant	(±) Convulsant	Hof and Adron-Harris [85]
Ozolinone (metabolite of etazoline)	(−) Diuretic	(+) Inhibits low doses of (−) or of furoxemide	(±) Diuretic	Greven et al [86]
Picenadol	(+) Morphinomimetic	(−) Narcotic antagonist	(±) Partial agonist	Zimmerman and Gesellchen [87]
Alpha-(2,4,5)-trichloro-phenoxy-propionic acid	(+) Auxin-like plant growth regulator	(−) Decreases activity of (+)	(±) Auxin-like plant growth regulator	Smith et al [88]
6-Ethyl-9-oxaergoline (EOE)	(−) Dopamine agonist	(+) Dopamine antagonist	(±) Dopamine agonist	Lotti and Taylor [89]

should be used when the distomer shows antagonistic properties. For clinical trials as a narcotic analgesic, the racemate was the preferred preparation owing to its partial agonist profile.

C. Racemic Switches

Presently, a general trend in the pharmaceutical industry is to switch from racemates to single enantiomers [90,91]. Examples are given by (R)-(−)-verapamil, (S)-fluoxetin, (S)-ketoprofen, (R)-albuterol, levofloxacin, esopra-zole, cetirizine, cisapride, and many others [92,93]. In addition to the quality improvement of the drug, this switch represents a way to prolong its life insofar as the isolated eutomer is legally considered a new drug entity. As a consequence, drug companies are increasingly adopting racemic switches as a management strategy. The company first develops a chiral drug as a racemate, and later on patents and develops the single isomer [93]. This strategy does not always work successfully [94]. This is illustrated by the S(−)-eutomer of propranolol. This compound shows reduced β-blocking activity when administered as single isomer, compared with its bioavail-ability when administered as a racemate, suggesting that the presence of R(+)-propranolol had a beneficial effect on the availability of S(−)-propranolol [95]. The same phenomenon happened when the racemate of fluoxetin was compared with its eutomer. The consequence was that the management at Eli Lilly decided not to practice the racemic switch for this compound.

D. The Distomer is Metabolized to Unwanted or Toxic Products

Racemic deprenyl, a monoamine−oxidase inhibitor used in the treatment of depression, is metabolized to (+)- and (−)-metamphetamine [96], the former being much more active than its (−)-isomer as central stimulant, lead-ing to drug abuse (Table 18.7).

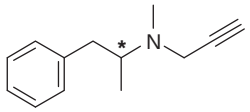
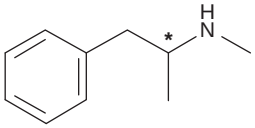
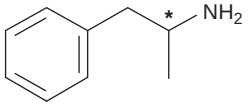
On the other hand, the (−)-isomer of deprenyl is a much more potent MAO-B inhibitor than the (+)-isomer. For these reasons, racemic deprenyl has been replaced by (−)-deprenyl in clinical practice. In the racemic local anesthetic prilocaïne (Figure 18.8), only the R-(−)-isomer is metabolized to an aniline derivative (ortho-toluidine) and to the corresponding para- and ortho-aminophenols that are highly toxic and responsible for methemoglobinemia [97].

The S-(+) enantiomer is not a substrate for the metabolizing enzyme and would probably be chemically safe.

Many of the side effects encountered with racemic DOPA (e.g., granulocytopenia) were not seen with levo-DOPA and therefore can be attributed to the (+)-enantiomer [98]. For this reason, the racemate is no longer given. Post-anesthesia reactions to the anesthetic and analgesic agent ketamine are overwhelmingly associated with the R(−) antipode [99].

In vitro studies suggest that the beneficial antiarrhythmic properties of disopyramide are concentrated in the S (+)-isomer, whereas the negative inotropic effect predominates in the R(−)-isomer [100]. In addition, the pharma-cokinetics (clearance and protein binding) differ [101,102]. For these reasons, selection of the S(+) isomer may have led to the development of a very effective drug with significantly fewer therapeutic problems.

TABLE 18.7 Activities of Deprenyl Enantiomers and their Metabolites

Formula	Activity	Ratio
 Deprenyl	MAO-B inhibition amphetamine effects	(–) ≫ (+) (–) < (±)
 Metamphetamine	Amphetamine effects	(–) ≪ (+)
 Amphetamine		

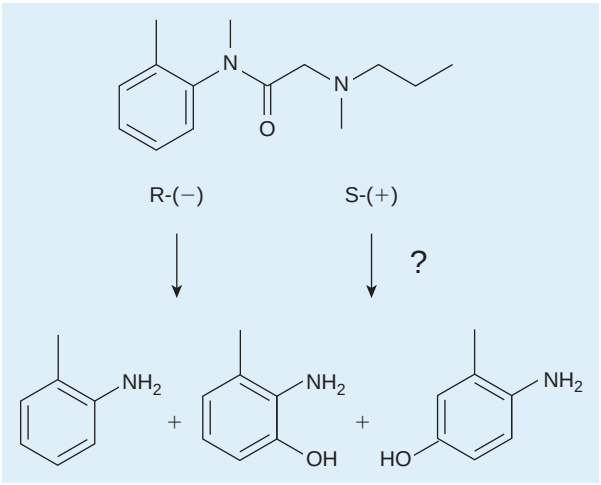


FIGURE 18.8 Stereoselective metabolic attack yielding toxic metabolites [97].

E. Deletion of the Chiral Center

Nowadays it is widely accepted that racemates and both enantiomers are usually three different pharmacological entities, and that it requires extensive pharmacological, toxicological, and clinical pharmacological research before it can be decided whether it is advantageous to use racemates or enantiomers in clinical practice. According to Soudijn [103], these research efforts could be reduced to about one-third when drugs without centers or planes of asymmetry could be developed with the same or higher affinity.

Effectively, asymmetry is far from being an absolute requisite for activity. The alkaloid morphine possesses five chiral centers. On the other hand, its synthetic derivative fentanyl is devoid of any asymmetric center but nonetheless is one of the most potent analgesics known. In some instances, the chiral centers can at least partially be eliminated. This is the case for the synthetic analogs of the HMG-CoA reductase inhibitor mevinolin. Mevinolin itself (Figure 18.9) has eight asymmetric centers, but structure–activity relationship (SAR) studies rapidly revealed that the six chiral centers contained in the hexahydronaphtalene unit are unnecessary for HMG-CoA inhibition. The second generation of mevinolin analogs, illustrated in Figure 18.9 by the compound HR 780, retains only two of the initial eight chiral centers [104].

Usually, chiral centers are eliminated in creating symmetry. Thus, in a series of muscarinic agonists derived from 3-aminopyridazines, one of the most favorable side-chains was the racemic 2-*N*-ethylpyrrolidinyl-methyl chain, that is, the side-chain of sulpiride (Figure 18.10). The 5-methyl-6-phenylpyridazine bearing this basic chain at its 3-amino function presented a 0.26 micromolar affinity for M1 muscarinic receptor preparations [105].

After resolution of the racemate, the corresponding enantiomers show only a six-fold difference in M1 affinity. It was therefore decided to eliminate the chiral center by introducing symmetry either by ring opening or ring closure, or even by replacing the 2-*N*-ethylpyrrolidinyl-methyl unit with the nonchiral tropane ring. The modified structures show affinities similar to those of the corresponding chiral molecule [105].

F. Usefulness of Racemic Mixtures

In practice, if both optical isomers are of similar potency and do have similar pharmacokinetic profiles, it may be useless to proceed to the resolution of the racemic mixture. Such situations are infrequent but may occur. An example is given by the antithrombotic acids 21-X and 21-Y (Figure 18.11) [106]. The corresponding pure enantiomers were first compared to the corresponding racemates for their *in vitro* activities.

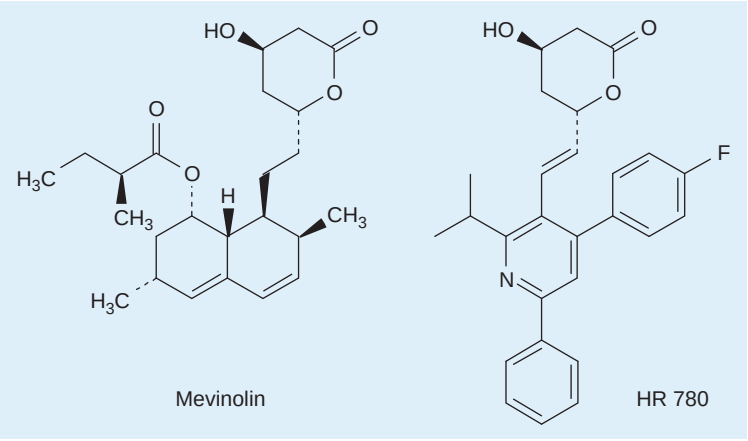


FIGURE 18.9 Deletion of six out of eight chiral centers yields still highly potent mevinolin analogs [104].

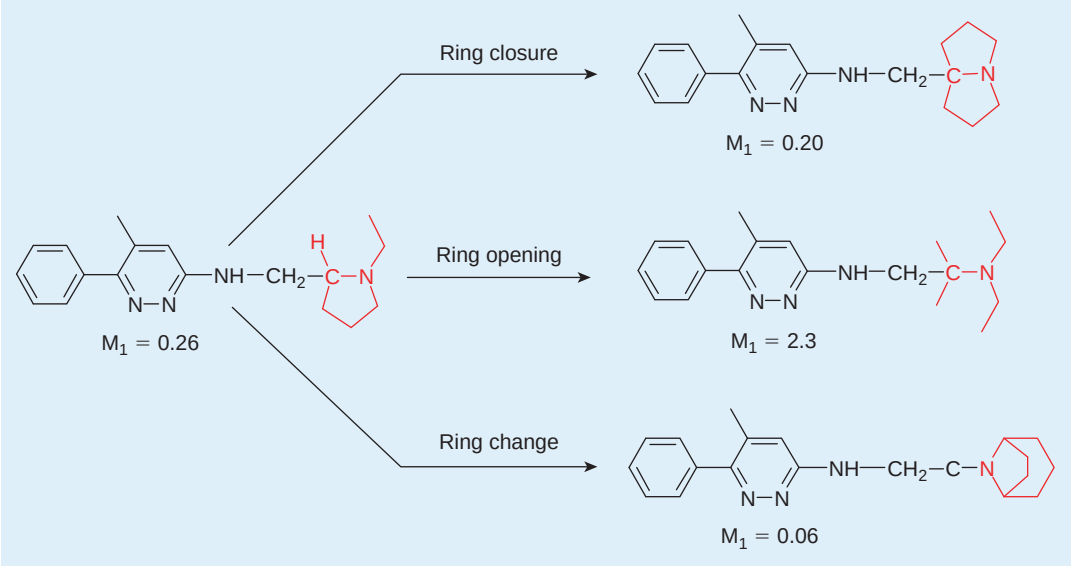


FIGURE 18.10 Introducing symmetry and abolishing thus a chiral center (affinity values for M_1 receptor preparations expressed as micromoles).

In both series, almost equipotent activities were observed for thromboxane receptor antagonism and thromboxane synthase inhibition ($IC_{50} = 2-30$ nM). Upon oral administration to guinea pigs, the enantiomers inhibited the *ex vivo* U-46619-induced platelet aggregation with potencies similar to those of the corresponding racemates. This indicates that the enantiomers have pharmacologic profile and bioavailability similar to that of the corresponding racemic compound (Figure 18.11).

The racemates can even be more potent than either of the enantiomers used separately. This is observed with the antihistaminic drug isothipendyl [107]. In other cases, it may be of interest to racemize a natural optically active molecule. Thus, to warrant a constant pharmacological activity of ergotamine, which is racemized in solution producing inactive ergotaminine, the commercial solution is produced as an equilibrium mixture of the two antipodes [108]. Another example of the utility of a racemic mixture is given by the lysine salts of aspirine (Box 18.1). The acetylsalicylate prepared from (R,S)-lysine is a stable, crystalline white powder that is freely soluble in water, giving a tasteless, odorless, and colorless solution, suitable for parenteral injections. Surprisingly, the corresponding salts of pure (R)-lysine or pure (S)-lysine do not crystallize [109]. Finally, when the distomer is

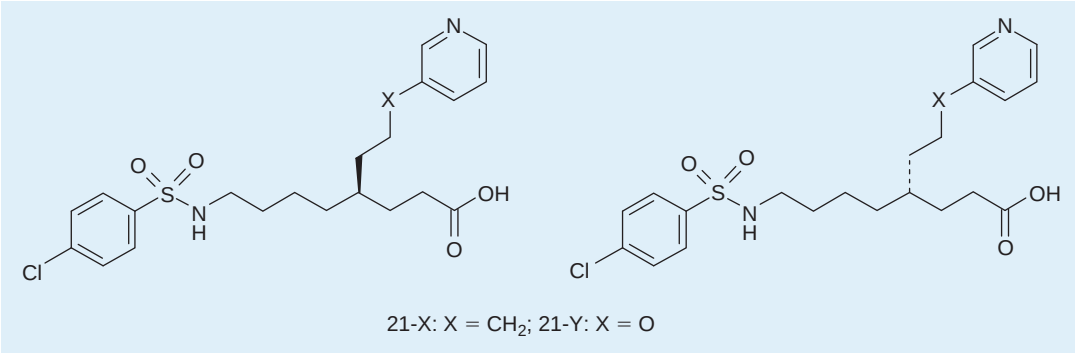


FIGURE 18.11 Isoactive antithrombotic enantiomers [106].

BOX 18.1

THE WATER-SOLUBLE D, L-LYSINE SALT OF ASPIRINE: A SUCCESS STORY BASED ON LUCK AND SERENDIPITY

Before World War II, in the little city of Chef-Boutonne in France's Poitou-Charente region, Gaston Baetz prepared in the back shop of his pharmacy some in-house remedies able to be sold over the counter. His start-up business, created in 1933 and named ATP (Association Technique Pharmaceutique), flourished reasonably well.

At the same period, about 160 miles eastward in the Auvergne city of Commentry, a chemical plant named Alimentation Equilibrée Commentry (AEC) produced some synthetic food additives such as vitamins, methionine, and lysine for the intensive farming of chicken and pigs. A spin-off company of AEC named l'Equilibre Biologique was then created with the objective to develop some amino-acid derived drugs for human use. The two companies, ATP and l'Equilibre Biologique, merged in 1945 under the control of l'Alimentation Equilibrée.

In 1953, Gaston Baetz, his oldest son Jacques, and some friends had the opportunity to take over the control of l'Equilibre Biologique and to develop their pharmaceutical company under the name Egic. Egic was specialized in hospital formulations of sterile injectable nutriments such as lipid emulsions, amino-acid mixtures, and glucose solutions.

With the objective of preparing water-soluble salts of aspirin (which is known to be poorly soluble in water and rapidly hydrolyzed), the chemists of l'Equilibre Biologique prepared and patented some aspirin salts of the basic amino-acids lysine and arginine, which were provided by AEC (French patent 1 295 304, May 7, 1956). The salts proved to be water-soluble, but the laboratory procedure used for their preparation yielded slightly hygroscopic salts that partly decomposed to acetic acid and free salicylic acid. The industrial development was therefore precluded for some years.

BOX 18.1 (*cont'd*)

In 1967, under the scientific direction of Dr. Pierre Baetz (the younger brother of Jacques), the question arose if an adequate pharmaceutical formulation could not rescue the compound. However, a preliminary physicochemical re-investigation had to be undertaken first. Indeed, the original batches of lysine aspirinate were prepared from the available AEC lysine. As this lysine was produced by synthesis, it was racemic, [(D,L)-lysine]. For many scientists of the company, notably Pierre Baetz, Francis Rosé, and Abkar Vartanian, it appeared that the natural L-lysine salt would be preferable. As a consequence, attempts were made to prepare the aspirin salt of the natural L-lysine. However, despite good will and obstinate efforts, the salt did not crystallize, nor did the aspirin salt of D-lysine. The decision was then taken to develop the D,L-salt corresponding to the synthetic racemic lysine. *A posteriori* it is interesting to note that fortunately, at the time of the preparation of the first batches of lysine aspirinate, the AEC chemists were probably the only ones in the world using racemic lysine.

The combination of a modified preparation procedure yielding a very anhydrous lyophilized salt (French Demande 2 115 060, August 11, 1972) on one hand, and working under controlled atmosphere on the other hand, allowed the industrial production of the racemic lysine salt. This compound is a white crystalline powder which is freely soluble in water, odorless and tasteless. It allowed injectable preparations of aspirin for the first time.

A possible explanation of the higher propensity of the racemate to crystallize can reside in a closer fitting of the molecules in the crystal grid. As some textbooks mention, "It is easier to put a pair of shoes into a box than two left shoes (or two right ones)."

converted to the eutomer *in vivo*—as seen above for ibuprofen and its analogs—it also becomes preferable to commercialize the racemate.

The recommendations of the European Community Working Party on drug quality, safety, and efficacy take into account two situations [110]. For already well-established racemates, the clinical use can continue as such. No specific study of the isolated enantiomers is required. For newly introduced chiral drugs, both enantiomers have to be prepared and studied separately with regard to their activity as well as their disposition *in vivo*. However, the final decision to introduce the drug on the market as enantiomer or as racemate belongs to the producer.

References

[1] Testa B. Chiral aspects of drug metabolism. Tr Pharmacol Sci 1986;7:60–4.
[2] Testa B, Mayer JM. Stereoselective drug metabolism and its significance in drug research. In: Jucker E, editor. Progress in drug research. Basel: Birkhäuser; 1988. p. 249–303.
[3] Ariëns EJ. Stereoselectivity of bioactive agents: general aspects. In: Ariëns EJ, Soudjin W, Timmermans PBMWM, editors. Stereochemistry and biological activity of drugs. Oxford: Blackwell Scientific Publications; 1983. p. 11–32.
[4] Ariëns EJ. Stereospecificity of bioactive agents. In: Ariëns EJ, van Rensen JJS, Welling W, editors. Stereoselectivity of pesticides. Amsterdam: Elsevier; 1988.
[5] Simonyi M. Problems and wonders of chiral molecules. Budapest: Akadémia Kiado; 1990. p. 400.
[6] Brown C. Chirality in drug design and synthesis. London: Academic Press; 1990. p. 243.
[7] Casy AF, Dewar GH. The steric factor in medicinal chemistry dissymmetric probes of pharmacological receptors. New York: Plenum Press; 1993.
[8] Federsel H-J. Chiral drug discovery and development-from concept stage to market launch. In: Taylor JB, Triggle DJ, editors. Comprehensive medicinal chemistry, vol. 2. Amsterdam: Elsevier; 2007. p. 713–34.
[9] Brooks H, Guida CW, Daniel GK. The significance of chirality in drug design and development. Curr Top Med Chem 2011;11(7):760–70.
[10] Fournié-Zaluski MC, Lucas-Soroca E, Devin J, Roques BP. ^{1}H NMR configural correlation for retro-inverso dipeptides: application to the determination of the absolute configuration of "enkephalinase" inhibitors. J Med Chem 1986;29:751–7.
[11] Jordan R, Midgeley JM, Thonoor CM, Williams CM. Beta-adrenergic activities of octopamine and synephrine stereoisomers on guinea-pig atria and trachea. J Pharm Pharmacol 1987;39:752–4.
[12] Martino-Barrows AM, Kellar KJ. [3H]Acetylcholine and [3H](-)nicotine label the same region in brain. Mol Pharmacol 1987;31:169–74.
[13] Chang RSL, Tran VT, Snyder SH. Heterogenicity of histamine H1-receptors: species variations in [3H]mepyramine binding of brain membranes. J Neurochem 1979;32:1653–63.
[14] Waldmeier PC, Baumann PA, Hauser K, Maitre L, Storni A. Oxaprotiline, a noradrenaline uptake inhibitor with an active and an inactive enantiomer. Biochem Pharmacol 1982;31:2169–76.
[15] Laduron PM, Vervimp M, Leysen JE. Stereospecific *in vitro* binding of [^{3}H]-dexetimide to brain muscarinic receptors. J Neurochem 1979;32:421–7.

[16] Lovell RA, Freedman DX. Stereospecific receptor sites for D-lysergic acid diethylamide in rat brain: effects of neurotransmitters, amine antagonists, and other psychotropic drugs. Mol Pharmacol 1976;12:620–30.
[17] Easson LH, Stedman E. Studies on the relationship between chemical constitution and physiological action. V. Molecular dissymmetry and physiological activity. Biochem J 1933;27:1257–66.
[18] Beckett AH. Stereochemical factors in biological activity. In: Jucker E, editor. Fortschritte der arzneimittel forschung, vol. I. Basel: Birkhäuser Verlag; 1959. p. 455–530.
[19] Farmer PS, Ariëns EJ. Speculations on the design of nonpeptide peptidomimetics. Tr Pharmacol Sci 1982;3:362–5.
[20] Patil PN, LaPidus JB, Campbell D, Tye A. Steric aspects of adrenergic drugs. II. Effects of dl isomers and deoxy derivatives on the reserpine-pretreated vas deferens. J Pharmacol Exp Ther 1967;155:13–23.
[21] Ruffolo Jr RR. Stereoselectivity in adrenergic agonists and adrenergic blocking agents. In: Ariëns EJ, Soudjin W, Timmermans PBMWM, editors. Stereochemistry and biological activity of drugs. Oxford: Blackwell Scientific Publications; 1983. p. 103–25.
[22] Hibert MF, Trumpp-Kallmeyer S, Bruinvels A, Hoflack J. Three-dimensional models of neurotransmitter G protein-coupled receptors. Mol Pharmacol 1992;40:8–15.
[23] Trumpp-Kallmeyer S, Hoflack J, Bruinvels A, Hibert M. Modelling of G protein-coupled receptors: application to dopamine, adrenaline, serotonin, acetylcholine and mammalian opsin receptors. J Med Chem 1992;35:3448–62.
[24] Patil PN. Steric aspects of adrenergic drugs. 8. Optical isomers of beta adrenergic receptor antagonists. J Pharmacol Exp Ter 1968;160:308–14.
[25] Kier LB. The preferred conformations of ephedrine isomers and the nature of the alpha adrenergic receptor. J Pharmacol Exp Ther 1968;164:75–81.
[26] Portoghese PS. Stereochemical studies on medicinal agents. IV. Conformational analysis of ephedrine isomers and related compounds. J Med Chem 1967;10:1057–63.
[27] Maxwell RE, Nickel VS. The antibacterial activity of the isomers of chloramphenicol. Antibiot Chemother 1954;4:289–95.
[28] Drayer DE. Pharmacodynamic and pharmacokinetic differences between drug enantiomers in humans: an overview. Clin Pharmacol Ther 1986;40(2):125–33.
[29] Fell AF. Current perspectives on chiral drug development. EUFEBS News Lett 1998;7:1–2.
[30] Ebnöther A, Weber H-P. Synthesis and absolute configuration of clemastine and its isomers. Helv Chim Acta 1976;59:2462–8.
[31] Ringdahl B, Resul B, Dahlbom R. Stereoselectivity of some oxotremorine antagonists containing two chiral centers. J Pharm Pharmacol 1979;31:837–9.
[32] Rauws AG. Origin and basis of stereoselectivity in biology. In: Ariëns EJ, Soudijn W, Timmermans PBMWM, editors. Stereochemistry and biological activity of drugs. Oxford: Blackwell Scientific Publications; 1983. p. 1–10.
[33] Lehman FPA, Ariëns EJ, Rodrigues de Miranda JF. Stereoselectivity and affinity in molecular pharmacology. In: Jucker E, editor. Progress in drug research, Vol. 20. Basel: Birkhäuser Verlag; 1976. p. 101–42.
[34] Lehman PA. Stereoisomerism and drug action. Tr Pharmacol Sci 1986;7:281–5.
[35] Casy AF. Stereochemistry and biological activity. In: Burger A, editor. Medicinal chemistry, vol. 1. New York: Wiley-Interscience; 1970. p. 81–107.
[36] Schröder E, Rufer C, Schmiechen R. Arzneimittelchemie [Pharmacological chemistry], vol. 1. Stuttgart: Georg Thieme Verlag; 1976. p. 48. [in German].
[37] Pfeiffer CC. Optical isomerism and pharmacological action: a generalization. Science 1956;124:29–31.
[38] Barlow RB. Enantiomers: how valid is Pfeiffer's rule? Trends Pharmacol Sci 1990;11:148–50.
[39] Gualtieri F. Pfeiffer's rule OK? Trends Pharmacol Sci 1990;2:315–16.
[40] Lien EJ, Rodrigues de Miranda JF, Ariëns EJ. Quantitative structure-activity correlation of optical isomers: a molecular basis for Pfeiffer's rule. Mol Pharmacol 1976;12:598–604.
[41] Portoghese PS, Williams DA. Stereochemical studies on medicinal agents. VIII. Absolute stereochemistries of isomethadol isomers. J Med Chem 1970;13:626–30.
[42] Wu F, Gaohua L, Zhao P, Jamei M, Huang S-M, Bashaw ED, et al. Predicting nonlinear pharmacokinetics of omeprazole enantiomers and racemic drug using physiologically based pharmacokinetic modeling and simulation: application to predict drug/genetic interactions. Pharm Res 2014;31(8):1919–29.
[43] Leach A-G, Pilling E-A, Rabow A-A, Tomasi S, Asaad N, Buurma N-J, et al. Enantiomeric pairs reveal that key medicinal chemistry parameters vary more than simple physical property based models can explain. Rev Article Med Chem Commun 2012;3:528–40.
[44] White NJ, Looareeswan S, Warrel DA. Quinidine in falciparum malaria. Lancet 1981;2:1069–71.
[45] White NJ, Looareeswan S, Warrel DA. Quinine and quinidine: a comparison of EKG effects during the treatment of malaria. J Cardiovasc Pharmacol 1983;5:173–5.
[46] Alexander F, Gold H, Katz LN. The relative value of synthetic quinidine, dihydroquinidine, commercial quinidine, and quinine in the control of cardiac arrhythmias. J Pharmacol Exptl Ther 1947;90:191–201.
[47] Benson WM, Stefko PL, Randall LO. Comparative pharmacology of levorphan, racemorphan, and dextrorphan and related methyl ethers. J Pharmacol Exptl Therap 1953;109:189–200.
[48] Bullock MW, Hand JJ, Waletzky E. Resolution and racemization of dl-tetramisole, dl–6-phenyl-2,3,5,6-tetrahydroimidazo-[2,1-b]thiazole. J Med Chem 1968;11:169–71.
[49] Schnieden H. Levamisole: a general pharmacological perspective. Int J Immunopharmacol 1981;3:9–13.
[50] Singh L, Donald AE, Foster AC. Enantiomers of HA-966 (3-amino-1-hydroxypyrrolid-2-one) exhibit distinct central nervous system effects: (+)-HA-966 is a selective glycine/N-methyl-D-aspartate receptor antagonist, but (–)-HA-966 is a potent γ-butyrolactone-like sedative. Proc Natl Acad Sci USA 1990;87:347–51.
[51] Remy DC, Rittle KE, Hunt CA, Anderson PS, Engelhardt EL, Clineschmidt BV, et al. (+) and (–)-3-Methoxycyproheptadine: a comparative evaluation of the antiserotonin, antihistaminic, anticholinergic, and orexigenic properties with cyproheptadine. J Med Chem 1977;20:1681–4.

[52] Roden DM. Antiarrhythmic drugs. In: Hardman JG, Limbird LE, Molinoff PB, Ruddon RW, Goodman Gilman A, editors. The pharmacological basis of therapeutics. 9th ed. New York: McGraw-Hill; 1995. p. 1905.
[53] Aschwanden W, Kyburz E, Schönholzer P. Stereospezifizität der neuroleptischen Wirkung und Chiralität von (+)-3-{2-[Fluor-2-methyl-10,11-dihydrodibenzo[b,f]thiepin-10-yl)-1-piperazinyl]-äthyl}-2-oxazolidinon (16) [Stereospecificity of the neuroleptic effect and chirality of (+)-3-{2-[Fluor-2-methyl-10,11-dihydrodibenzo[b,f]thiepin-10-yl)-1-piperazinyl]-äthyl}-2-oxazolidinon (16)]. Helv Chim Acta 1976;59:1245–52. [in German].
[54] Nanayakkara NP, Tekwani BL, Herath HM, Sahu R, Gettayacamin M, Tungtaeng A, et al. Scalable preparation and differential pharmacologic and toxicologic profiles of primaquine enantiomers. Antimicrob Agents Chemother 2014;58(8):4737–44.
[55] Brondz I. The technology for preparation of generic (monoenantiomeric) antimalarial drug primaquine by using supercritical fluid chromatography. Separation of primaquine from quinocide: simultaneous resolution of the enantiomers of primaquine and their separation from quinocide in one run. Am J Anal Chem 2012;3(12A). Available from: http://dx.doi.org/10.4236/ajac.2012.312a117.
[56] Jamali F, Mehvar R, Pasutto FM. Enantioselective aspects of drug action and disposition: therapeutic pitfalls. J Pharm Sci 1989;78:695–715.
[57] Kroemer HK, Gross AS, Eichelbaum M. Enantioselectivity in drug action and drug metabolism: influence on dynamics. In: Cutler NR, Sramek JJ, Narang PK, editors. Pharmacodynamics and drug development. Chichester: John Wiley & Sons; 1994. p. 103–14.
[58] Buch H, Rummel W, Brandenburger V. Versuche zur Aufklärung der Ursachen der unterschiedlichen narkotischen Wirksamkeit von (+)- und (−)-Evipan [Attempts to identify the causes of the different narcotic potency of (+)- and (−)-Evipan]. Arch Pharmakol Exptl Pathol 1967;257:270–1. [in German].
[59] Duhm B, Maul W, Medenwald H, Platzschke K, Wegner LA. Experimental animal studies with alpha-methyldopa-14C, with special attention to the optical isomers. II. Organ distribution. Z Naturforsch 1967;22b:70–84.
[60] Sjoerdsma J, Udenfriend S. Pharmacology and biochemistry of α-methyl-dopa in man and experimental animals. Biochem Pharmacol 1961;8:164.
[61] Vermeulen NPE. Stereoselective biotransformation: its role in drug disposition and drug action. In: Harms AF, editor. Innovative approaches in drug research. Amsterdam: Elsevier Science Publishers BV; 1986. p. 393–416.
[62] Campo V-L, Bernardes LSC, Carvalho I. Stereoselectivity in drug metabolism: molecular mechanisms and analytical methods. Curr Drug Metab 2009;10(2):188–205.
[63] Axelrod J. The enzymatic N-demethylation of narcotic drugs. J Pharmacol Expt Therap 1956;117:322–30.
[64] Elison C, Elliott HW, Look M, Rapoport H. Some aspects of the fate and relationship of the N-methyl group of morphine to its pharmacological activity. J Med Chem 1963;6:237–46.
[65] Degwitz E, Ullrich V, Staudinger H, Rummel W. Metabolism and cytochrome P-450 binding spectra of (+) and (−) hexobarbital in rat liver microsomes. Hoppe Seylers Z Physiol Chem 1969;350:547–53.
[66] Kupfer A, Bircher J, Preisig R. Stereoselective metabolism, pharmacokinetics, and biliary elimination of phenylethylhydantoin (nirvanol) in the dog. J Pharmacol Exptl Therap 1977;203:493–9.
[67] Keberle H, Riess W, Hoffman K. The stereospecific metabolism of the optical antipodes of β-phenyl-β-ethylglutarimide (Doriden). Arch Int Pharmacodyn 1963;142:117–24.
[68] Jenner P, Testa B. The influence of stereochemical factors on drug metabolism. Drug Metab Rev 1973;2:117–84.
[69] Testa B, Jenner P. Concepts in drug metabolism. Part A. New York: Marcel Dekker; 1980.
[70] Adams SS, Bresloff P, Mason CG. Pharmacological differences between the optical isomers of ibuprofen: evidence for metabolic inversion of the (−)-isomer. J Pharm Pharmacol 1976;28:256.
[71] Adams SS, Cliffe EE, Lessel B, Nicholson JS. Some biological properties of 2-(4-isobutylphenyl)-propionic acid. J Pharm Sci 1967;56:1686.
[72] Vane JR. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. Nature 1971;231:232–5.
[73] Wechter WJ, Loughhead DG, Reisher RJ, van Geissen GJ, Kaiser DG. Enzymatic inversion at saturated carbon: nature and mechanism of the inversion of R(−) p-*iso*-butyl hydratropic acid. Biochem Biophys Res Commun 1974;61:833–7.
[74] Bopp RJ, Nash JF, Ridolfo AS, Shepard ER. Stereoslective inversion of (R)-(−)-benoxaprofen to the (S)-(+)-enantiomer in humans. Drug Metab Dispos 1979;7:356–9.
[75] Kemmerer JM, Rubio FA, McClain RM, Koechlin BA. Stereospecific assay and stereospecific disposition of carprofen in rats. J Pharm Sci 1979;68:11274–80.
[76] Tanaka Y, Hayashi R. Stereospecific inversion of configuration of 2-(2-isopropylindan-5-yl)-propionic acid in rats. Chem Pharm Bull 1980;28:2542–5.
[77] Simmonds RG, Woodage TJ, Duff SM, Green JN. Stereospecific inversion of (R)-(−)-benoxafen in rat and in man. Eur J Drug Metab Pharmacokin 1980;5:169–72.
[78] Zhou Q, Yu L-S, Zeng S. Stereoselectivity of chiral drug transport: a focus on enantiomer–transporter interaction. Drug Metab Rev 2014;46(3):283–90.
[79] Sugane T, Tobe T, Hamaguchi W, Shimada I, Maeno K, Miyata J, et al. Atropisomeric 4–Phenyl–4H–1,2,4-triazoles as selective glycine transporter 1 inhibitors. J Med Chem 2013;56:5744–56.
[80] Schmidt W, Jahnchen E. Stereoselective drug distribution and anticoagulant potency of phenprocoumon in rats. J Pharm Pharmacol 1977;29:266–71.
[81] Zhang L, Yu W, Rong Y, Guo X, Ye J, Shen Z, et al. Enantiomeric separation of 2-arylpropionic acid nonsteroidal anti-inflammatory drugs and β-blockers by RP-HPLC using an amylose chiral stationary phase for the enantioselective skin permeation study. Anal Methods 2014;6:6058–65.
[82] Gunne LM, Galland L. Stereoselective metabolism of amphetamine. Biochem Pharmacol 1967;16:1374–7.
[83] Vermeulen NPE, Breimer DD. Stereoselectivity in drug and xenobiotic metabolism. In: Ariëns EJ, Soudjin W, Timmermans PBMWM, editors. Stereochemistry and biological activity of drugs. Oxford: Blackwell Scientific Publications; 1983.
[84] Hof RP, Rüegg UT, Hof A, Vogel A. Stereoselectivity at the calcium channel: opposite action of the enantiomers of a 1,4-dihydropyridine. J Cardiovasc Pharmacol 1985;7:689–93.

[85] Hof IK, Harris RA. Mechanism of action of barbiturates. Ann Rev Pharmacol Toxicol 1981;21:93–111.
[86] Greven J, Defrain W, Glaser G, Meywald K, Heidenreich O. Studies with the optically active isomers of the new diuretic drug ozolinone. Pflüger's Arch 1980;384:57–60.
[87] Zimmerman DM, Gesellchen PD. Analgesics (peripheral and central), endogenous opioids, and their receptors. Ann Rep Med Chem 1982;17:21–30.
[88] Smith SM, Wain RL, Wightman F. Studies on plant growth-regulating substances. V. Steric factors in relation to mode of action of certain aryloxyalkylcarboxylic acids. Ann Appl Biol 1952;39:295–307.
[89] Lotti VJ, Taylor DA. α2-Adrenergic agonist and antagonist activity of the respective (−) and (+)-enantiomers of 6-ethyl-9-oxaergoline. Eur J Pharmacol 1982;85:211–15.
[90] Gellad WF, Choi P, Mizah M, Good CB, Kesselheim AS. Assessing the chiral switch: approval and use of single-enantiomer drugs, 2001 to 2011. Am J Manag Care 2014;20(3):e90–7. <http://www.ajmc.com/publications/issue/2014/2014-vol20-n3/Assessing-the-Chiral-Switch-Approval-and-Use-of-Single-Enantiomer-Drugs-2001-to-2011/>.
[91] Agranat I, Silvya R, Wainschtein SR. The strategy of enantiomer patents of Drugs. Drug Discov Today 2010;15:163–70.
[92] Stinson ST. Chiral drugs. Chem Eng News 1995;44–74.
[93] Stinson SC. Chiral drug interactions. Chem Eng News 1999;101–20.
[94] Isvrael Agranat I, Wainschtein SR, Zusman EZ. The predicated demise of racemic new molecular entities is an exaggeration. Nat Rev Drug Discov 2012;11:972–3.
[95] Lindner W, Rath M, Stochitzky K, Semmelrock HJ. Pharmacokinetic data of propranolol enantiomers in a comparative human study with (S)- and (R,S)-propranolol. Chirality 1989;1:10–13.
[96] Reynolds GP, Elsworth JD, Blau K, Sandler M, Lees AJ, Stern GM. Deprenyl is metabolized to metamphetamine and amphetamine in man. Br J Clin Pharmacol 1978;6:542–4.
[97] Akerman B, Astrom A, Ross S, Telc A. Studies on the absorption, distribution, and metabolism of labeled prilocaine and lidocaine in some animal species. Acta Pharmacol Toxicol 1966;24:389–403.
[98] Cotzias GC, Papavasiliow PS, Gellene R. Modification of parkinsonism: chronic treatment with l-dopa. N Engl J Med 1969;280:337–45.
[99] White PF, Ham J, Way WL, Trevor AJ. Pharmacology of ketamine isomers in surgical patients. Anesthesiology 1980;52:231–9.
[100] Kidwell GA, Lima JJ, Schaal SF, Muir WM. Hemodynamic and electrophysiologic effects of disopyramide enantiomers in a canine blood superfusion model. J Cardiovasc Pharmacol 1989;13:644–55.
[101] Lima JJ, Boudoulas H, Shields BJ. Stereoselective phamacokinetics of disopyramide enantiomers in man. Drug Metab Dispos 1985;13:572–7.
[102] Giacomini KM, Nelson WL, Pershe RA, Valdevieso L, Turner-Tamayasu K, Blaschke TF. In vivo interactions of the enantiomers of disopyramide in human subjects. J Pharmacokin Biopharmacol 1986;14:335–56.
[103] Soudijn W. Advantages and disadvantages in the application of bioactive racemates or specific isomers as drugs. In: Ariëns EJ, Soudijn W, Timmermans PBMWM, editors. Stereochemistry and biological activity of drugs. Oxford: Blackwell Scientific Publications; 1983. p. 89–102.
[104] Baader E, Bartmann W, Beck G, Bergmann A, Granzer E, Jendrella H, et al. Rational approaches to enzyme inhibitors: new HMG-CoA reductase inhibitors. In: Claassen V, editor. Trends in drug research. Amsterdam: Elsevier; 1990. p. 49–71.
[105] Wermuth CG. Aminopyridazines: an alternative route to potent muscarinic agonists with no cholinergic syndrome. Il Farmaco 1993;48:253–74.
[106] Bhagwat SS, Gude C, Cohen DS, Dotson R, Mathis J, Lee W, et al. Thromboxane receptor antagonism combined with thromboxane synthase inhibition. 5. Synthesis and evaluation of enantiomers of 8-{[(4-chlorophenyl)sulfonyl]amino}-4-(3-pyridinylalkyl)octanoic acid. J Med Chem 1993;36:205–10.
[107] Yamamura S, Oda K, Mizoguchi T, Saito S, Iwasawa Y, Kiyomoto A. Study on the structure-activity relationships of adrenergic, beta-mimetic benzylamine derivatives. V. 9-aryl-1H-2,3,7,8,9,10-hexahydro-benzo[d,e]quinolines. Chem Pharm Bull 1979;27:858–69.
[108] Stoll A. Über Ergotamin [About ergotamine]. Helv Chim Acta 1945;28:1283–308.
[109] Baetz J. Scientific Director, Egic Laboratories, Paris. In Personal commmunication: Paris; 1982.
[110] Knabe J. Synthetische Enantiomere als Arzneistoffe [Synthetic enantiomers as drugs]. Pharmazie in unserer Zeit 1995;24:324–30. [in German].

CHAPTER

19

Multitarget Drugs: Strategies and Challenges for Medicinal Chemists

Richard Morphy[1] *and Zoran Rankovic*[2]

[1]Eli Lilly, Erl Wood Manor, Windlesham, Surrey, UK; [2]Eli Lilly, Indianapolis, IN, US

OUTLINE

When the whole is greater than the sum of the parts

I. INTRODUCTION

Historically, the compounds produced by medicinal chemists have been screened by *in vivo* pharmacologists in whole animal models of disease. This approach provided a means to identify, in a single test, compounds that exhibited a rare combination of desirable pharmacokinetic (PK) and pharmacodynamic (PD) properties. The downside of this approach was that the animal model was essentially a "black box," so when compounds were inactive, it was unclear whether this was because they no longer interacted with the PD target(s) or whether they had failed to reach the required site of action due to poor pharmacokinetics. Often, the molecular targets driving both the desired PD effect and any detrimental side effects were unclear so a rational, reductionist approach to drug discovery was impossible. In the latter decades of the twentieth century, the drug-discovery paradigm became cemented in a "one-target-one-disease" philosophy, increasingly dominated by *in vitro* high-throughput screening (HTS) technologies. Many successful drugs that are selective for a single target have emerged from this strategy, but despite the best efforts of drug discoverers, many diseases remain inadequately treated by such an

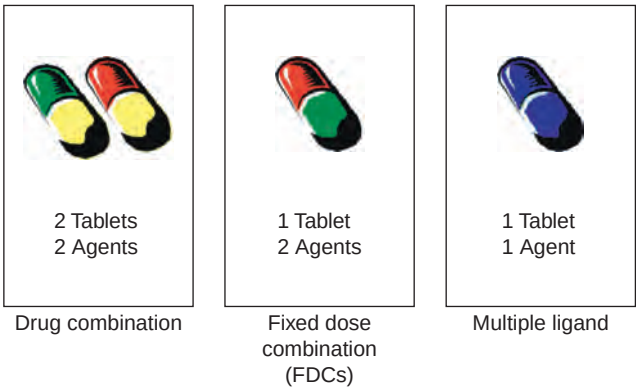


FIGURE 19.1 The three different approaches to multitarget therapy (polypharmacology).

approach. Evidence suggests that the main causes of failure of compounds in the clinic are now a lack of efficacy and poor safety [1]. Since agents that modulate multiple targets simultaneously (polypharmacology) have the potential to enhance efficacy or improve safety relative to drugs that address only a single target, it is not surprising that this area is attracting the attention of increasing number of drug discoverers [2–4].

There are three distinctly different approaches to multitarget therapy (Figure 19.1). Traditionally, clinicians have treated unresponsive patients by combining therapeutic mechanisms with cocktails of drugs. Most frequently the cocktail is administered in the form of two (or more) individual tablets (scenario A) [5,6]. However, the benefits of this approach are often compromised by poor patient compliance, particularly for treating asymptomatic diseases such as hypertension [7]. Recently, there has been a move toward fixed dose combination (FDC) drugs, whereby two (or more) agents are co-formulated in a single tablet to make dosing regimes simpler and thereby improve patient compliance (scenario B) [8,9]. An alternative strategy is to develop a multiple ligand which is a single-chemical entity that is able to modulate multiple targets simultaneously (scenario C) [2].

Across the pharmaceutical industry, the FDC approach increasingly provides an attractive opportunity for enhancing R&D output [10]. Several FDCs are very successful commercially. Vytorin, combining the cholesterol absorption inhibitor ezetimibe with the statin simvastatin for treating hypercholesterolemia, had sales of 2 billion dollars in 2006. Advair, combining a glucocorticoid, fluticasone, with a long-acting bronchodilator, salmeterol, for treating asthma, had sales of 6.5 billion dollars in 2006. However, significant risks can be involved in the development of FDCs. Commercial uncertainty arises from the risk that clinicians might still prefer prescribing combinations of existing monotherapies that may offer greater dose flexibility and lower cost treatment in the case of generic drugs. This is illustrated by the sales of the hypertension/hyperlipidemia FDC Caduet being 370 million dollars in 2006 compared with multi-billion dollar sales for the individual drugs atorvastatin and amlodipine [10]. Differences in the relative rates of metabolism between patients can produce highly complex PK/PD relationships for FDCs, leading to unpredictable variability between patients and necessitating extensive and expensive clinical studies.

Compared to FDCs, the multiple-ligand approach has a profoundly different risk–benefit profile (Table 19.1). A downside is that it is significantly more difficult to adjust the ratio of activities at the different targets. However, this increased complexity in the design and optimization of such ligands occurs at an earlier and therefore less expensive stage of the drug-discovery process. The risks and costs of developing multiple ligands is in principle no different from the development of any other single entity. Another advantage of a single-chemical entity is a lower risk of drug–drug interactions compared to cocktails or FDCs [11].

Several drugs currently on the market have been found to have activity at more than one target. In some cases, this nonselective activity serendipitously increases efficacy, whereas in others it is associated with side effects. Although these historical drugs were not designed for multiple activity, a recent trend has been to deliberately and rationally design ligands that act selectivity on multiple targets (selectively nonselective drugs). Numerous terms have been used to describe such ligands, with dual ligand, heterodimer, promiscuous drug, pan-agonist, and triple blocker being just a few of many examples. The complexity and inconsistency of this nomenclature has partly obscured overall developments in this field, so to improve communication and awareness, the common term “designed multiple ligands” (DMLs) has recently been introduced [2].

In a number of disease areas, drug discoverers have followed a three-stage evolutionary journey from a nonselective drug with undesirable side effects, to a target-selective ligand with a safer profile, and onward toward a

TABLE 19.1 Risk–Benefit Profile of FDCs and Multiple Ligands

Risks/benefits	FDCs	Multiple ligands
Patient compliance	Improved when compared to drug cocktails.	Improved when compared to drug cocktails.
PK/PD relationship	Often highly complex PK/PD correlation that requires sophisticated formulation solutions.	Single-chemical entity; generally no issues.
Drug–drug interactions	Increased risk of drug– drug interactions.	Risk similar to any other single-compound entity.
Titration of activities	Possible, but may be difficult and costly to develop; requires full clinical development, production, and marketing of a series of dose combinations.	Not possible.
R&D challenges	Potentially fast progress toward proof-of-concept; however, clinical development can be complicated by the requirement to demonstrate the superiority of combination versus individual agents, as well as potentially increased risk of drug– drug interactions and formulation issues.	Can be challenging to design a multiple ligand with the required ratio of activities and adequate selectivity at the discovery stage; however, the development program and regulatory approval process is the same as for a standard NCE.
Intellectual property	Patent life of old drugs can be prolonged when combined with a new drug.	Standard NCE position.

selectively nonselective DML that attempts to provide a more optimal balance of efficacy and safety. An example of a nonselective ligand is the atypical antipsychotic drug Clozapine, which shows antagonist activity at multiple aminergic G-protein-coupled receptors (GPCRs). To circumvent the side effects of Clozapine, a number of ligands that are selective for single receptors targeted by Clozapine were developed, such as dopamine D_4 and serotonin 5-HT_{2a} antagonists, but these lacked sufficient efficacy in the clinic [12]. Research then shifted toward DMLs, such as the dual D_2/5-HT_{2a} antagonists [13,14]. Nonselective tricyclic anti-depressants such as Amitryptyline were superseded by selective serotonin (5-HT) transporter inhibitors (SSRIs), which increased safety but had a slow onset of action and lacked efficacy in some patients. Dual serotonin and norepinephrine (NA) re-uptake inhibitors (SNRIs) were subsequently developed with the hope of addressing these deficiencies [15]. The same trend is observed in the area of nonsteroidal anti-inflammatory drugs (NSAIDs), starting from nonselective agents such as aspirin, to selective cyclooxygenase-2 (COX-2) inhibitors, and then to dual COX-2/5-lipoxygenase (5-LOX) inhibitors [16]. Similarly, for the treatment of asthma, nonselective adrenergic agonists (e.g., epinephrine) have been replaced by selective β_2-adrenoceptor agonists such as salbutamol, with a significantly improved therapeutic window. Most recently, dual M3 antagonist/β_2 agonist and D_2/β_2 agonist have been developed [17,18].

II. STRATEGIES FOR LEAD GENERATION

As with single-target projects, medicinal chemists have access to a number of different ways of generating the chemical matter with which to commence a DML project. Conceptually, there are two quite different methods of generating lead compounds: screening approaches, which rely largely upon serendipity, and knowledge-based approaches, which exploit information either from the general literature or proprietary information from within an organization (Figure 19.2).

The screening of either diverse or focused compound libraries can deliver a single molecule that has at least minimal activity at each of the targets of interest. To date, few examples of DMLs derived via the HTS approach have been reported. This could be due to the fact that HTS has only become the *de rigeur* method of lead generation in the last decade or so, and there is an inevitable time lag to publication. Other factors could be the logistical complications of screening against multiple targets in parallel or the inherently low probability of detecting a compound with a multiple profile of therapeutic interest from screening compounds at random. Due to the large number of compounds typically involved in diversity-based screening, they will usually be screened first at one target of interest, and any actives will then be filtered on the basis of activity at the other target(s). Even if activity is observed for the second target, usually the balance of affinities is non-optimal, so the activity ratio must be adjusted during optimization.

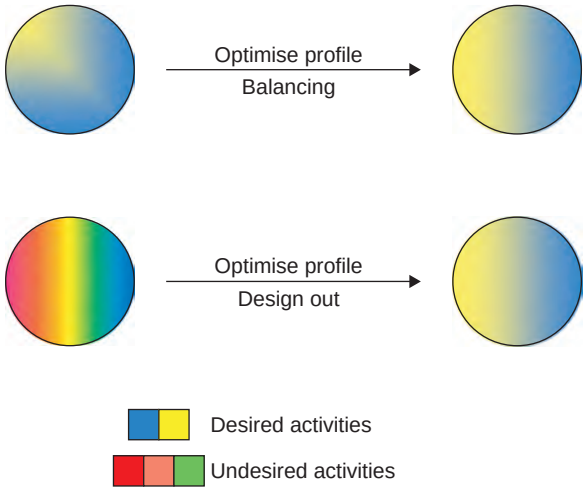


FIGURE 19.2 The screening of diverse or focused libraries can deliver a compound that has at least minimal activity at each target of interest. However, it is unlikely that the hit compound has the optimal affinity for all targets, so the profile must be balanced during optimization. Alternatively, screening might deliver a compound that—in addition to the desired activities—has undesired activities. These must be designed out during optimization.

Compared to HTS, the literature presents many more examples of the screening of focused libraries of compounds selected from single-target projects or using prior knowledge of the targets. In focused screening, compound classes that are already known to be active against one of the targets of interest are screened against another target. For example, DMLs for kinase targets are usually discovered serendipitously through the cross-screening of ligands from selective kinase programs against other kinases. In addition to the desired activities, screening frequently provides hit compounds that bind to other targets. To minimize the risk of side effects, the medicinal chemist will need to "design out" these undesired activities.

The second most common lead-generation strategy reported in the literature is a knowledge-based approach known as framework combination. It starts with two compounds, one of which binds with high selectively to one of the targets and the other with high selectively to the other target. In this case, the first goal is to "design in" both activities into a single lead molecule by combining the frameworks (and the underlying pharmacophores) of the two selective molecules (Figure 19.3). The intellectual elegance of the framework combination stems from the fact that often a wealth of structure−activity relationships (SAR) knowledge is on hand from previous selective-ligand projects that can be used to guide the optimization process.

DMLs arising from framework combination can be viewed as linked, fused or merged, depending on the degree to which the frameworks have been integrated (Figure 19.3). In linked DMLs (conjugates), the molecular frameworks are not at all integrated, and there is a distinct linker group between the two components that is not found in either of the selective ligands. This linker is usually intended to be metabolically stable so that the single compound is capable of interacting with both targets, albeit different ends of the molecule may be responsible for the activity at the different targets. Some linked DMLs contain a cleavable linker that is designed to be metabolized to release two ligands that interact independently with each target. This scenario represents a half-way point between a true DML and a FDC.

If the frameworks are essentially touching, so neither a discernable linker nor any framework overlap is present, the DML can be viewed as fused. In the most common and most sought-after type of DML, the frameworks are merged by taking advantage of commonalities in the structures of the starting compounds. Medicinal chemists will normally aspire to maximize the degree of overlap in order to produce smaller and simpler molecules. The degree of framework combination for the examples reported in the literature forms a continuum, with high molecular weight (MW) DMLs with lengthy linker groups at one extreme, and small DMLs with highly merged frameworks at the other.

The screening and framework-combination approaches to lead generation have various advantages and disadvantages that influence which one is best applied to a particular project (Table 19.2). Indeed, given the added challenges of multiple-ligand projects in general, it would make sense to employ both strategies if feasible to increase the overall chance of success. A major advantage of the screening approach is that you start from a compound that already has multiple activities built in, although these may be quite weak. Screening can add

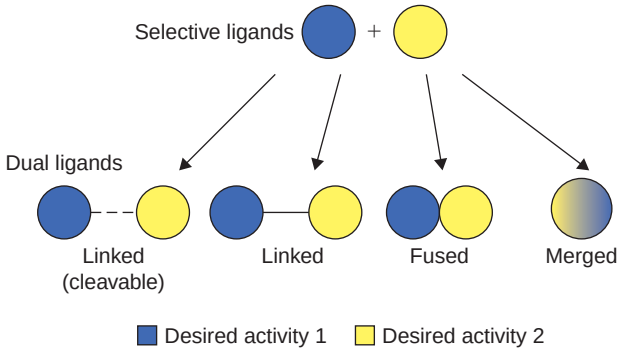


FIGURE 19.3 Framework combination is a knowledge-based approach to generating DMLs. There is a continuum in the degree of merger of the frameworks of the target-selective starting ligands. In linked DMLs, the frameworks are connected via a definable linker, which in some cases is designed to be cleaved *in vivo* to release two independently acting drugs. In fused DMLs, the frameworks are directly attached, and in the commonest form of DML, the frameworks are merged together.

TABLE 19.2 Features of the Screening and Framework Combination Approaches to DMLs

Screening approach	Framework combination approach
Can provide novel chemotypes not found in selective ligands.	Can be employed where selective ligands are known for each target.
Can provide ligands with complex profiles unavailable by framework combination.	Can employ existing SAR knowledge from selective ligand projects to assess feasibility and facilitate optimization.
Can provide ligands with improved physicochemical and pharmacokinetic properties compared to framework combination.	Can readily provide linked DMLs for use as IV drugs or biochemical tools, even for two targets with incompatible pharmacophores.
Chance of success is low for unrelated targets.	Can be difficult to incorporate a second activity while retaining the first activity and good physicochemical properties.

particular value if there is a lack of selective ligands for the targets of interest or little of the SAR information required for a knowledge-driven approach. Screening can deliver novel and unexpected chemotypes, sometimes providing hits for unusual target combinations that span unrelated receptor families. Since the framework-combination strategy almost invariably produces dual ligands, discovering ligands that bind to more than two targets usually demands that a screening approach be followed. Screening can also provide ligands with improved physiochemical and PK properties compared to framework combination (see Section V on physicochemical properties).

In the case of framework combination, incorporating a second activity into a compound that has no measurable affinity for that target while retaining affinity for the original target is by no means an easy task. However, many examples in the literature testify that it can often be achieved by effectively leveraging SAR knowledge from historical selective-ligand projects. Compared to screening, framework combination can provide rapid entry to conjugate molecules that can be used as intravenously (IV) administered drugs or biochemical tools, even for targets that are very different at the pharmacophore level (see Section V on physicochemical properties). The chance of success with a random screening approach would be expected to diminish rapidly as the targets in a combination become more dissimilar.

III. MAIN AREAS OF FOCUS IN DISCOVERY

Historically, the most common disease areas for DML projects have been psychiatry, neurodegeneration, and oncology, as well as metabolic, cardiovascular, and allergic diseases. A common theme is to focus on a primary target that has previously been well validated in the clinic for a given disease, and then add one or more

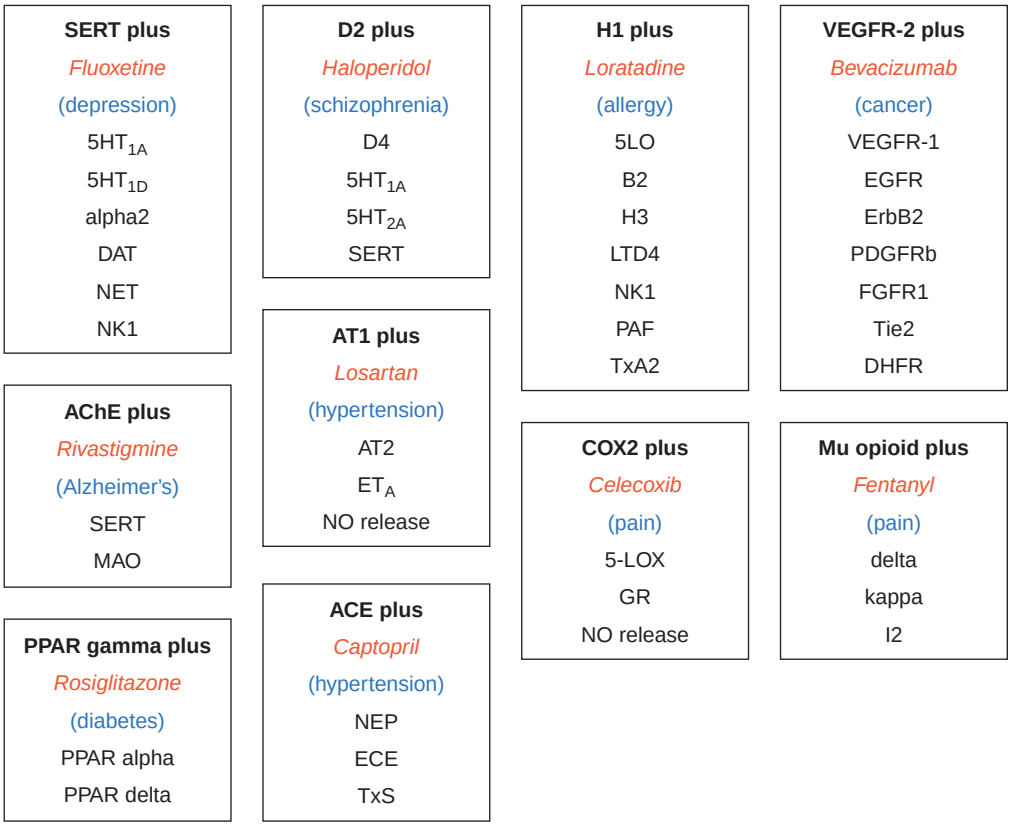


FIGURE 19.4 Secondary activities that have been added to a clinically validated primary target in an effort to enhance efficacy and reduce side effects. An example of a drug selective for the primary target is shown in red. Abbreviations: 5-HT_{1A} = 5-HT_{1A} receptor; 5-HT_{1D} = 5-HT_{1D} receptors; 5-HT_{2A} = 5-HT_{2A} receptor; 5-LOX = 5-lipoxygenase; alpha2 = alpha2 adrenergic receptor; ACE = angiotensin-converting enzyme; AChE = acetylcholinesterase; AT_1 = angiotensin-1 receptor; AT_2 = angiotensin-2 receptor; B2 = bradykinin-2 receptor; COX-2 = cyclooxygenase-2; delta = delta-opioid receptor; D_2 = dopamine-2 receptor; D_4 = dopamine-4 receptor; DAT = dopamine transporter; DHFR = dihydrofolate reductase; ECE = endothelin-converting enzyme; EGFR = epidermal growth factor receptor; ET_A = endothelin-A receptor; FGFR1 = fibroblast growth factor receptor 1; GR = glucocorticoid receptor; H_1 = histamine-1 receptor; H_3 = histamine-3 receptor; kappa = kappa-opioid receptor; I2 = Imidazoline-2 receptor; LTD_4 = leukotriene D4 receptor; mu = mu-opioid receptor; MAO = monoamine oxidase; NEP = neutral endopeptidase; NET = norepinephrine transporter; NK1 = neurokinin-1 receptor; NO = nitric oxide; PAF = platelet-activating factor receptor; PDGFRb = Platelet-derived growth factor receptor beta; PPAR = peroxisome proliferator-activated receptor; SERT = serotonin transporter; TxA2 = thromboxane-A2 receptor; TxS = thromboxane-A2 synthase; VEGFR-1 = vascular endothelial growth factor receptor-1; VEGFR-2 = vascular endothelial growth factor receptor-2.

secondary activities in an effort to enhance efficacy and reduce side effects. For example, there are a large number of reported combinations containing the 5-HT transporter (serotonin transporter (SERT)) for depression, the histamine H_1 receptor for allergy, and thevascular endothelial growth factor receptor-2 (VEGFR-2) kinase for cancer (Figure 19.4). A relatively small number of target combinations have predominated in terms of their percentage share of the total number of publications in the literature between 1990 and 2005. The six most commonly reported combinations were as follows:

1. Angiotensin-converting enzyme/neutral endopeptidase (ACE/NEP) for hypertension.
2. Cyclooxygenase-2/5-lipoxygenase (COX-2/5LOX) for inflammatory pain.
3. Thromboxane-A2 receptor/thromboxane-A2 synthase (TxA2/TxS) as antiplatelet agents.
4. Serotonin transporter/5-HT1A receptor (SERT : 5-HT1A) for depression.
5. Neurokinin-1 receptors (NK1/NK2) for asthma.
6. Peroxisome proliferator-activated receptors (PPARalpha/PPARgamma) for diabetes.

A. Sert-Plus DMLs for Depression

Depression is associated with reduced levels of 5-HT in the brain. Drugs that inhibit the re-uptake of 5-HT, such as fluoxetine, have been used clinically for many years. In an attempt to address the deficiencies of SSRIs as

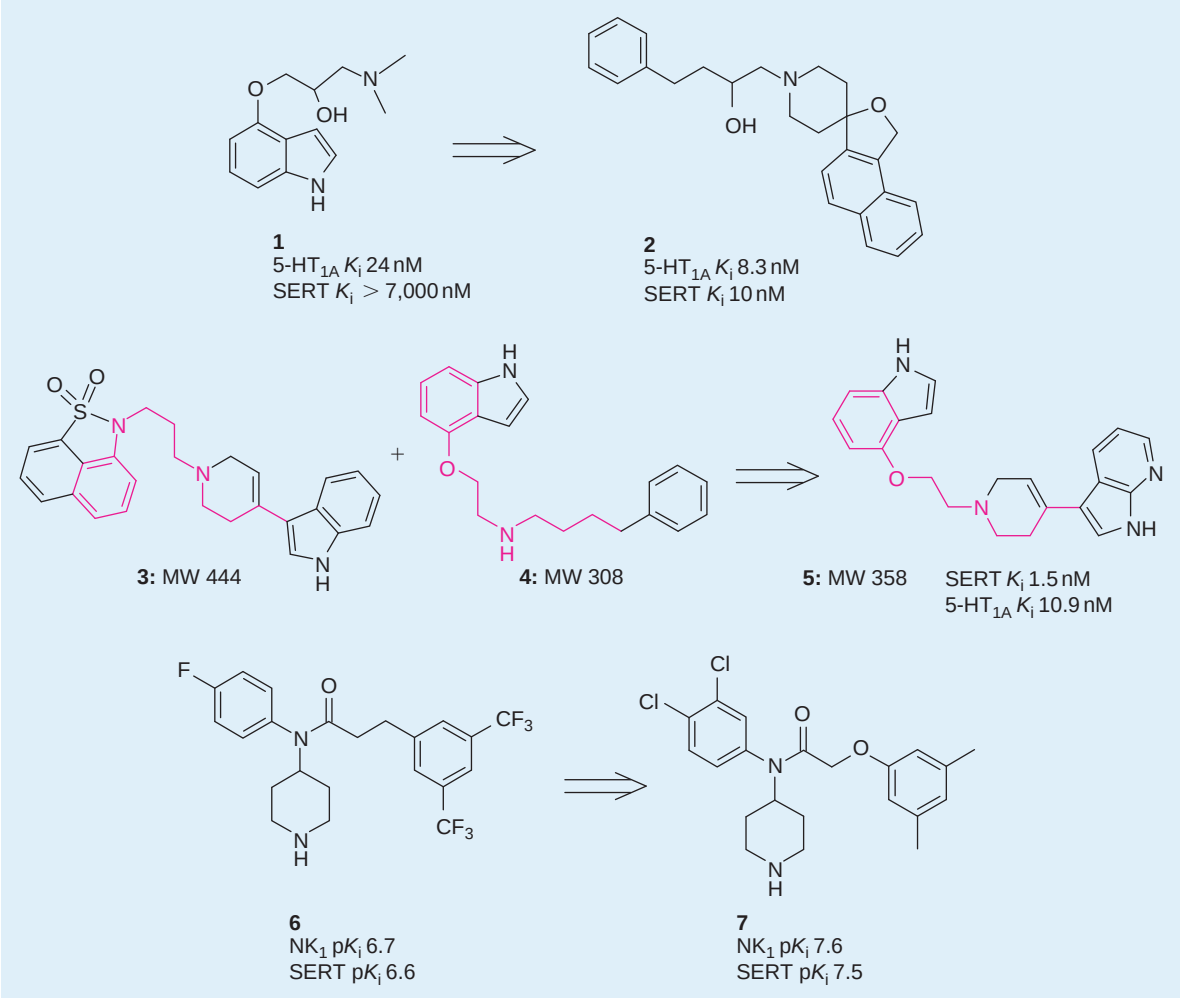


FIGURE 19.5 SERT-plus DMLs for depression.

anti-depressants in terms of efficacy or time of onset, SERT inhibition has been supplemented with activity at a secondary monoamine target, such as the 5-HT_{1A}, 5-HT_{1D}, alpha2, NET, or DAT (Figure 19.4). The delayed onset time for SSRIs has been attributed to the need for 5-HT_{1A} autoreceptors to become desensitized by sustained SERT blockade. By mimicking this desensitization with a 5-HT_{1A} antagonist, the onset time might be accelerated. The following three examples of dual 5-HT_{1A}/SERT blockers illustrate how the various lead-generation methods of screening and framework combination have been employed in this area.

Van Niel et al designed a focused screening library based on the 3-aryloxy-2-propanolamine scaffold found in the 5-HT_{1A} antagonist pinadol **1** (Figure 19.5) [19]. The variations at the amine and phenol positions included privileged structures, as well as fragments reported to have affinity for either 5-HT_{1A} or SERT. The SAR around the indole region was reasonably tolerant for both targets, but the only amine group that provided reasonable SERT inhibition was a spiro-piperidine **2**. This compound provided balanced inhibition as well as good oral exposure ($F = 65$ percent) and brain penetration in rats.

Using a framework-combination approach, compounds with dual 5-HT_{1A}/SERT activity were designed by Mewshaw et al, starting from a template known to possess robust SERT activity **3** and adding 5-HT_{1A} features in the form of the aryloxyethyl group found in **4** (Figure 19.5) [20]. The presence of a basic nitrogen was the common pharmacophoric feature that allowed the two frameworks to be merged to give **5**. The degree of framework overlap in this example (shown in magenta) is quite extensive, helping to produce a DML with a relatively low MW of 358 Da.

A high-throughput screen provided a multiple ligand **6** with a surprising combination of activities at a peptide GPCR, the neurokinin NK_1 receptor, and a monoamine transporter, SERT [21]. The two targets in this

combination have individually generated much interest for treating depression. While a NK_1-selective ligand, MK-869, gave disappointing results in clinical trials for depression, it is interesting that the target is now being pursued in combination with a clinically validated target, SERT. Although the hit **6** had only modest activity, systematic optimization of each aromatic moiety in turn provided a more potent compound with a balanced activity at both targets **7** (Figure 19.5). An aryl ether moiety was introduced to reduce lipophilicity.

B. Dopamine D_2-Plus DMLs for Schizophrenia

D_2-selective antagonists, such as haloperidol, are efficacious against the positive symptoms of schizophrenia (e.g., hallucinations and delusions). However, unlike the atypical antipsychotic drugs such as Clozapine, they do not address the negative symptoms such as social withdrawal, and they can cause extrapyramidal side effects (EPS) such as Parkinsonism. The aim of combining D_2 antagonism with activity at other targets (Figure 19.4) is to mimic the advantages of Clozapine in terms of efficacy without producing the disadvantages such as weight gain. One of a number of possible explanations for Clozapine's atypical profile is its higher antagonist affinity for the 5-HT_2 receptor than for the D_2 receptor. This observation lead to the so-called D_2/5-HT_2 ratio hypothesis whereby agents with >10-fold selectivity for 5-HT_2 over D_2 were sought.

Using a framework-combination approach, the structure of the endogenous agonist for the D_2 receptor, dopamine **8**, was fused with a large lipophilic group from the 5-HT ligand **9** (Figure 19.6) [22]. This transformed the D_2-agonist activity of the endogenous ligand into an antagonist. This DML is of the fused type, since there is only one nitrogen atom overlap between the frameworks of the starting compounds. Fused DMLs can have an undesirably high MW if the starting compounds are already quite large, but because the starting ligands are much smaller in this case, the resulting DML has a relatively low MW of 371 Da. Various heterocyclic groups were selected containing hydrogen bonding groups that might mimic the phenolic interaction, such as the oxindole found in **10**. Further optimization involved replacing the naphthyl group by a 1,2-benzisothiazole group **11**, which provided D_2 blockade comparable in potency to the typical antipsychotic haloperidol, together with a desirable D_2/5-HT_2 ratio of 11, comparable to the atypical agent, Clozapine [23]. The D_2/alpha1 ratio of 0.44 for **11** is substantially lower than that for Clozapine, suggesting the former should have a lower propensity to cause orthostatic hypotension. The ratio hypothesis was validated by clinical studies, and **11** (Ziprasidone) was launched in 2001 by Pfizer for the treatment of schizophrenia.

It has also been hypothesized that the unique profile of Clozapine in treating psychosis might be due to a precise ratio of D_2 and D_4 receptor affinities, with higher affinity required at D_4 than D_2. Zhao et al tried to reproduce this exact ratio with the goal of obtaining D_4 affinity of less than 10 nM and D_2 affinity of less than 200 nM [24]. They started from a nonselective D_2/D_4 compound **12** with undesired alpha-1 affinity, discovered via a screening approach (Figure 19.6). Introduction of a methyl group in the 2-position of the indoline ring gave an improvement in D_2 activity, **13**, and also good selectivity against a diverse range of other targets including alpha-1 [25]. It displayed activity in an *in vivo* test of psychosis, the inhibition of amphetamine-induced locomotor activity, and showed low activity in a catalepsy test, suggesting a low propensity to cause EPS. The behavioral data for this dual antagonist provided support for the D_2/D_4 ratio hypothesis, although the approach still needs clinical validation.

To maximize the efficacy and safety profile of an antipsychotic drug, much evidence now suggests that it is necessary to address more than two receptors. Using a screening approach, Garzya et al discovered a molecule **14** that had five activities regarded as being critical for an antipsychotic drug: blocking the D_2, D_3, 5-HT_{2A}, 5-HT_{2C}, and 5-HT_6 receptors [26]. Careful optimization produced a DML **15** with the optimal balance of affinities.

C. DMLs Targeting the Angiotensin System for Hypertension

The vasoconstricting peptide angiotensin II is a principal component of the renin-angiotensin system (RAS), a hormone system that helps regulate blood pressure and extracellular volume in the body. ACE inhibitors and angiotensin-1 receptor antagonists (AT_1)—such as Captopril and Losartan, respectively—have gained widespread acceptance for the treatment of hypertension and congestive heart failure (Figure 19.7). It has been postulated that DMLs such as dual ACE/NEP inhibitors or dual AT_1/endothelin receptor 1 (ET_A) antagonists may produce a beneficial synergistic effect in the management of hypertension and congestive heart failure (Figure 19.7).

A rational approach toward one of the earliest ACE/NEP dual inhibitors, dipeptide **18**, demonstrates how a good understanding of the pharmacophore requirements for the targets is highly desirable when designing

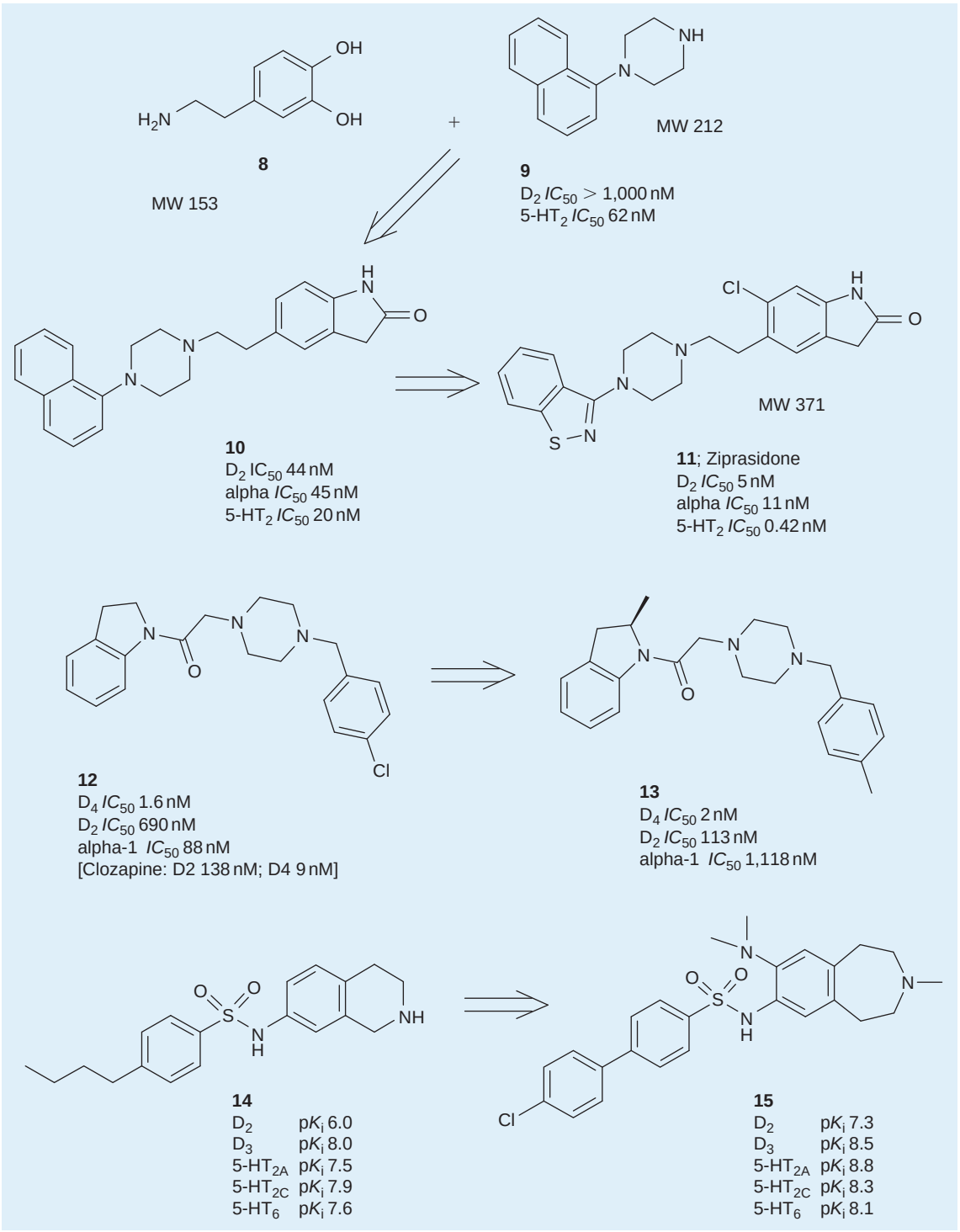


FIGURE 19.6 Dopamine D_2-plus DMLs for schizophrenia.

multiple ligands. For example, knowledge that NEP favors a hydrophobic substituent in the S1′ pocket, preferably a benzyl group such as the one present in the NEP selective inhibitor **16**, whereas ACE is more tolerant in this region but strongly favors a proline residue at P_2', as in ACE selective inhibitor captopril **17**, was instrumental in the design of **18** (Figure 19.7) [27]. In order to further improve the *in vitro* and *in vivo* potency of **18**, a range of diverse constrained analogs were designed, drawing extensively from the SAR generated around selective inhibitors. A particular challenge for this approach was a relatively tight SAR for NEP, which fortuitously was

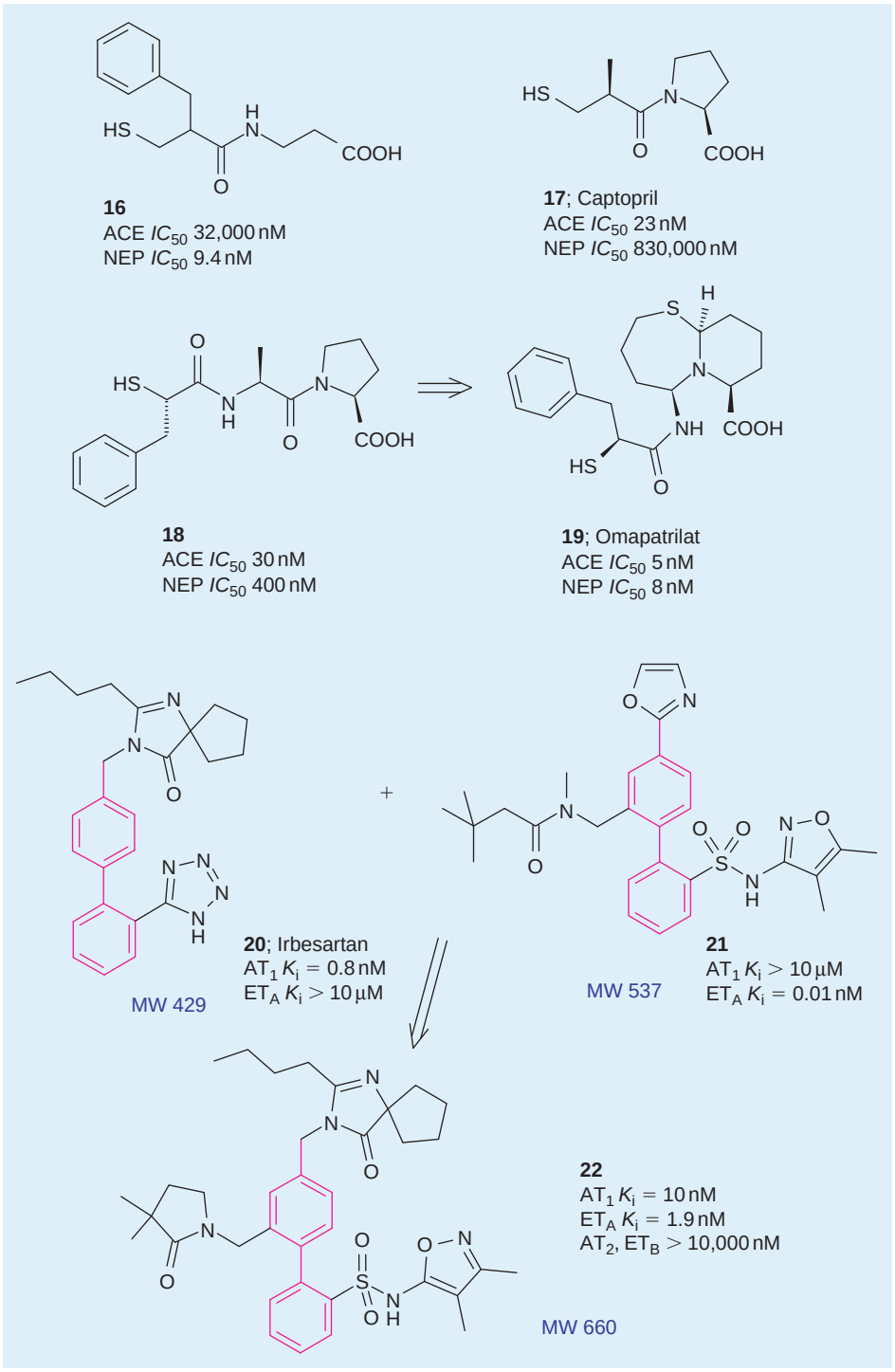


FIGURE 19.7 DMLs targeting the Angiotensin system for hypertension.

counterbalanced by a remarkably flexible SAR for ACE. The optimization efforts led to the discovery of a 7,6-bicyclic oxazepinone series, which produced omapatrilat **19**, a potent ACE/NEP inhibitor displaying a good PK profile and efficacy *in vivo* [28].

In another example of a merged DML, Murugesan et al were interested in simultaneous blockers of AT_1 and ET_A receptors, since a combination of the AT_1 selective antagonist Losartan and the ET_A/ET_B selective antagonist SB-290670 produced an additive reduction in blood pressure compared to either drug alone. Fortuitously, the selective AT_1 and ET_A antagonists **20** and **21** both contained a biaryl core (Figure 19.7) and the heterocycle in the 4′-position of the biaryl, required for AT_1 activity, was tolerated by ET_A, albeit with reduced affinity [29]. The

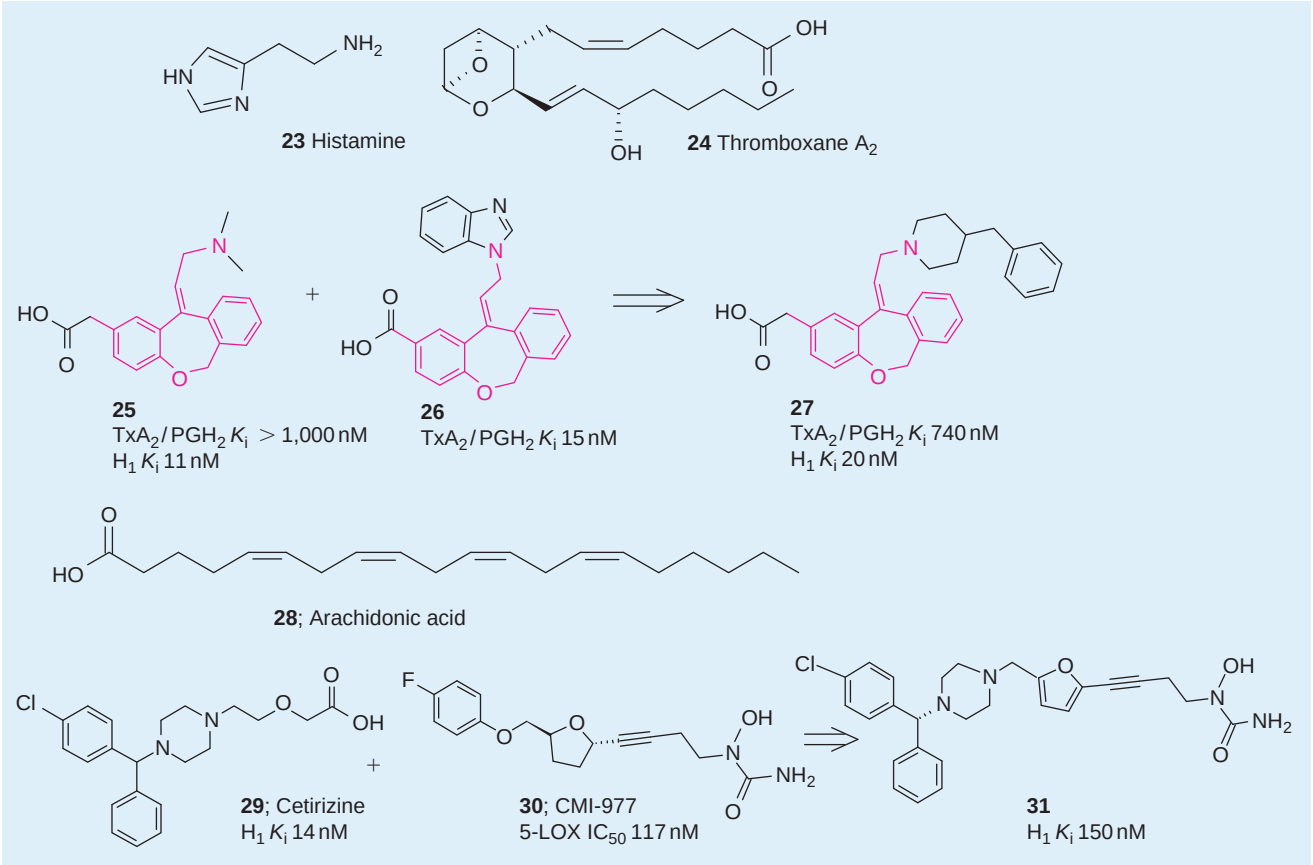


FIGURE 19.8 Histamine H_1-plus DMLs for allergies.

acylsulphonamide moiety was found to be a carboxylic acid bioisostere that was suitable for both receptors. By introducing a new substituent in the C2′-position of the biaryl **22**, a balanced dual activity at AT_1 and ET_A receptors was obtained. Compound **22** has a high MW of 660, reflecting the size of the starting compounds used for framework combination. Nonetheless, good oral bioavailability was observed for **22** in rats ($F = 38$ percent).

D. Histamine H_1-Plus DMLs for Allergies

Histamine is a primary mediator of the systemic inflammatory response to allergens in humans. H_1-antagonists such as loratidine have found widespread utility in the treatment of hay fever and other allergic reactions. However, selective H_1-antagonists have been largely ineffective for the treatment of asthma [30]. Almost all the H_1-antagonists that show some efficacy against asthma are reported to possess additional activities, suggesting that other chemical mediators are also involved in its pathogenesis. As a result, various groups have aimed to produce DMLs combining H_1-antagonism with a range of additional activities (Figure 19.8).

For example, the TxA_2 receptor has also been linked to allergic disease. Although both the H_1 and TxA_2 receptors are GPCRs, they might be expected to possess very different binding sites given that the endogenous ligand for the former is a small polar amine **23** and for the latter a lipophilic acid **24** (Figure 19.8). It might be anticipated that rationally "designing in" activity for targets with highly dissimilar endogenous ligands might be particularly difficult. However, a large number of recent examples show that this need not necessarily be a barrier to the discovery of a DML. It was observed by Ohshima et al that molecules with a common benzoxepine scaffold, the selective H_1-antagonist **25** and the TxA_2R antagonist **26** bound to both targets [31]. The tertiary amine group in **27** successfully mimicked the benzimidazole moiety that was known to be crucial for the TxA_2 activity of **26**. Compound **27** was active at both GPCRs, albeit with rather different binding affinities, as well as being selective over related GPCRs.

Perhaps an even more striking example is provided by the dual H_1-antagonist/5-lipoxygenase (5-LOX) inhibitor **31** that inhibits an enzyme that oxidizes highly lipophilic arachidonic acid **28**, while also antagonizing a GPCR that binds highly polar histamine **23**. The starting points for framework combination were the selective H_1-antagonist **29** and the 5-LOX inhibitor **30** (Figure 19.8) [32]. The strategy took advantage of the flat SAR around the basic nitrogen of the antihistamine to introduce a butynyl-hydroxyurea group into **31** required for 5-LOX inhibition.

E. AChE-Based DMLs for Alzheimer's Disease

Alzheimer's disease is associated with a progressive loss of cholinergic neurons in the brain that results in memory disturbances and cognitive dysfunction. One strategy for the treatment of Alzheimer's patients has been the use of acetylcholinesterase (AChE) inhibitors such as rivastigmine to enhance cholinergic activity in the central nervous system, although the result of such treatment is far from satisfactory in many patients. In an attempt to increase efficacy, AChE inhibition has been combined with SERT and monoamine oxidase (MAO) activity.

Kogen et al described work toward a dual AChE and 5-HT transporter (SERT) inhibitor, **35** [33]. This work also represents another example of a DML that crosses different proteomic families. A notable feature of this work is the elegant use of biostructural information to guide the combination of the frameworks of the starting compounds. A model of the active site of AChE showed that the AChE-selective inhibitor rivastigmine **32** possessed only three elements of the proposed AChE pharmacophore, lacking a fourth hydrophobic binding site (Figure 19.9). If the phenoxyethyl motif from the SERT blocker fluoxetine **33** could provide this hydrophobic interaction, potency should be improved relative to rivastigmine. Thus, hybridization of the two inhibitors, followed by optimization of the carbamate and phenoxy substituents, provided a dual inhibitor **34**. Conformational constraint using a seven-membered ring then gave a compound **35** with potent and balanced inhibition at the two diverse targets.

Youdim et al described dual AChE/MAO inhibitors as another approach to the treatment of Alzheimer's disease [34]. The structural framework of rivastigmine **32** was in this case combined with a selective MAO-B inhibitor rasagiline **36**, yielding a dual inhibitor ladostigil **37** (Figure 19.9). Reported SAR around this compound indicates that carbamate and propargylamine groups are key pharmacophoric elements responsible for the AChE and brain MAO inhibition, respectively. Ladostigil has shown efficacy in rhesus monkey cognition and neuroprotection models [35].

F. PPAR-Based DMLs for Metabolic Disease

The realization that the fibrate and glitazone classes of drugs used to treat dyslipidemia and type-2 diabetes respectively, exert their effects through activation of PPARalpha and PPARgamma, respectively, lead to the development of selective ligands for each of the PPAR receptor subtypes. However, findings suggesting that insulin resistance, dyslipidemia, and obesity can be seen as components of a complex mixture of abnormalities known as "metabolic syndrome" have stimulated interest in developing dual PPARalpha and PPARgamma agonists [36].

An interesting combination of screening and structure-based approaches was reported by Xu et al [37]. Their screening efforts resulted in identification of carboxylic acid **38** containing a bulky lipophilic group in the α-position as a moderate dual PPARα/γ agonist (Figure 19.10). The fact that **38** exhibited activity for both targets despite of lacking the lipophilic "tail" characteristic of PPAR ligands suggested that the α-benzyl group might improve the binding affinity of **39**, a well-balanced but weak dual agonist. The α-benzyl derivative **40** indeed showed improved activity at both PPARα and PPARγ. Shifting the oxygen adjacent to the quaternary stereogenic center in **40** to the alternative benzylic position provided a significantly more potent dual agonist **41**.

G. DMLs that Inhibit Multiple Kinases for Treating Cancer

Systematically targeting multiple kinases is currently of great interest in the fight against various forms of cancer. While there are a number of literature examples of the framework-combination approach being applied to GPCRs, transporters, nuclear receptors, proteases, and oxidases, we have so far identified no such examples for kinases. The absence of the framework-combination approach is probably due to the fact that obtaining selective ligands for kinases is still a major challenge, and this step precedes the rational "designing in" of multiple

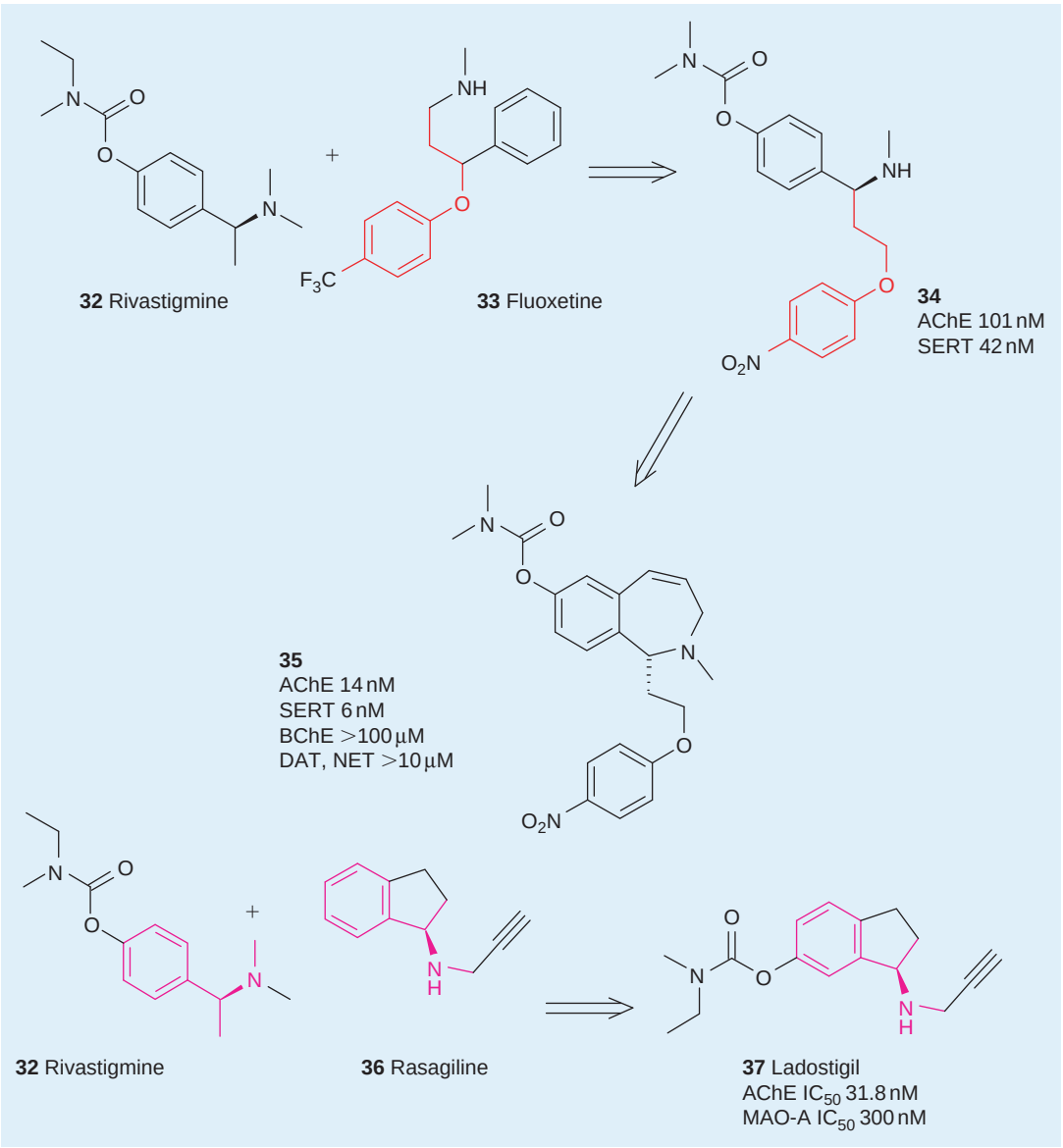


FIGURE 19.9 AChE-based DMLs for Alzheimer's disease.

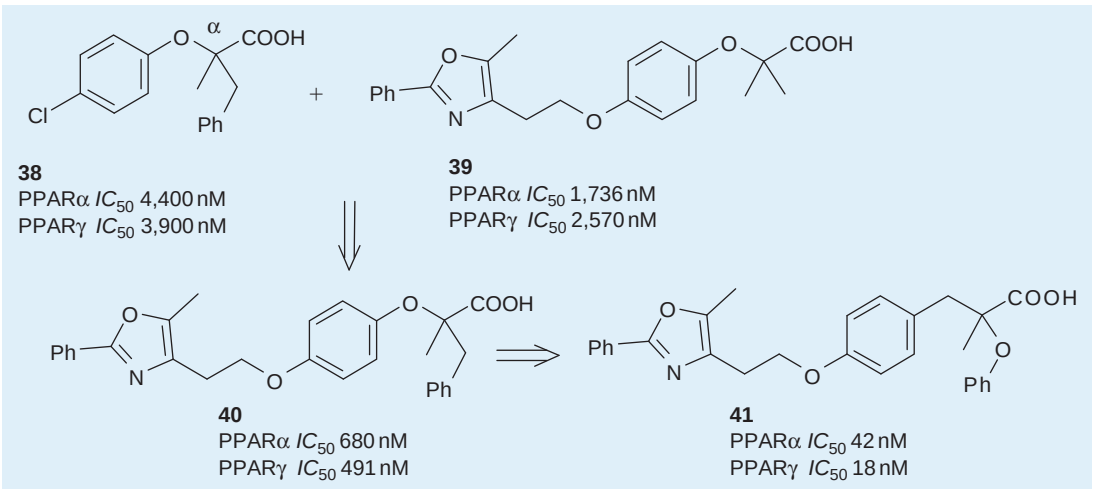


FIGURE 19.10 Dual PPARα/γ agonist for treating metabolic disease.

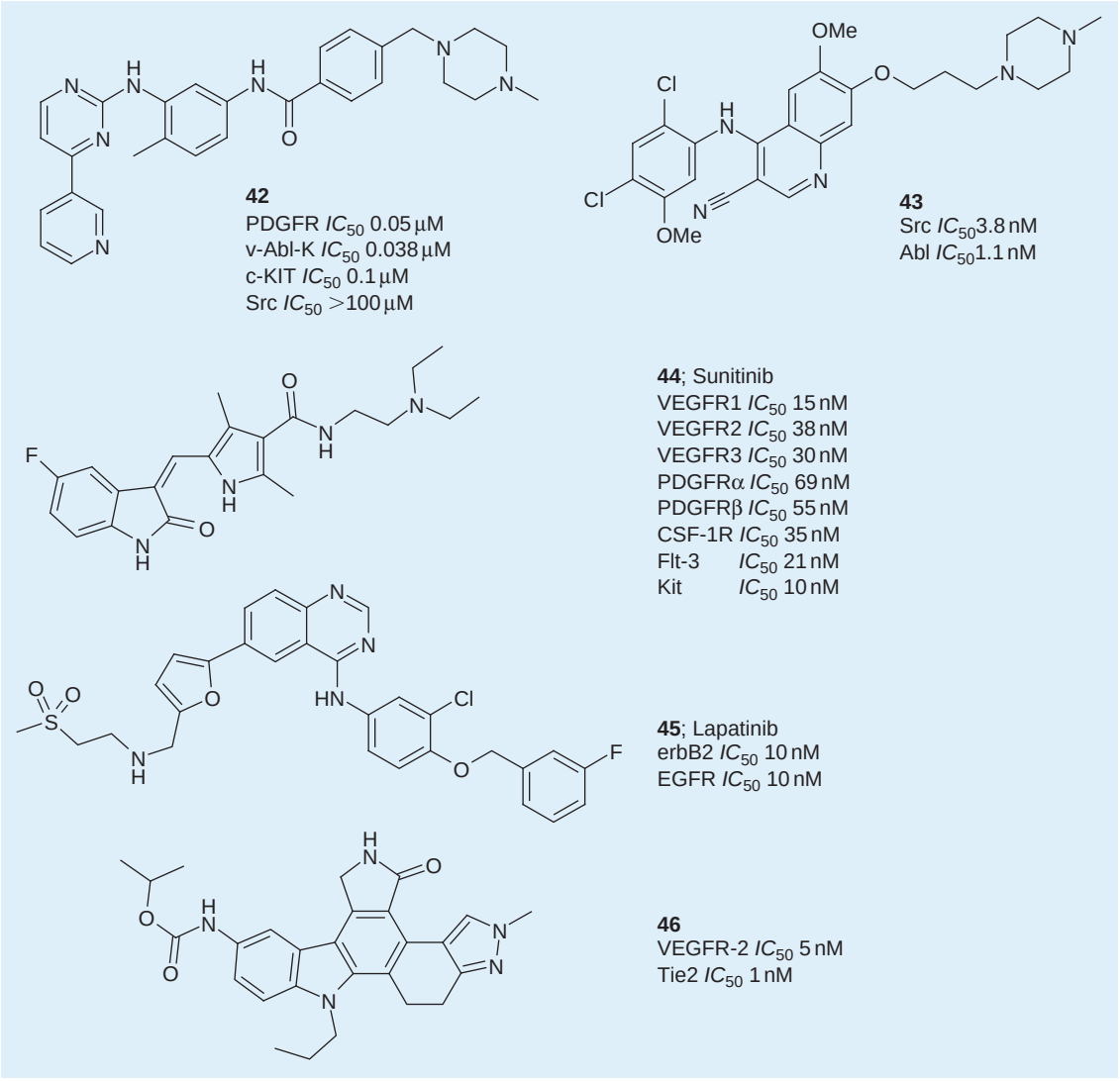


FIGURE 19.11 DMLs that inhibit multiple kinases for treating cancer.

activities, driven by knowledge of the selective ligand SARs. The most feasible strategy for designing multi-kinase inhibitors is focused screening to identify a nonselective inhibitor and then attempting to "design out" undesired kinase activities.

The first kinase inhibitor to be developed for clinical use was imatinib **42**, first marketed in 2001 for chronic myelogenous leukemia (CML). The clinical effectiveness of imatinib for the treatment of CML is now thought to be due to its multi-kinase activity, inhibiting PDGFR and c-KIT, in addition to its well-known activity as a Bcr-Abl kinase inhibitor. Resistance to imatinib can become a problem due to mutations in the Abl gene [38,39]. Dual Src/Abl inhibitors are currently of interest for the treatment of CML in patients who are resistant to imatinib. Whereas imatinib itself has no measurable activity against Src, Boschelli et al used a focused screening approach to identify an inhibitor **43** with dual Src/Abl activity (Figure 19.11) [40]. They found a very close correlation between the Src and Abl SARs, reflecting the close homology of these kinases.

Soon after the first generation of kinase inhibitors appeared on the market, the first multi-kinase inhibitors that were intentionally designed to have a particular profile were introduced for cancer treatment. For example, building on the success of the VEGFR-1/VEGFR-2 blocking antibody Avastin, a small molecule VEGFR-2 and PDGFRβ inhibitor sunitinib **44** was introduced in 2006 [41]. Similarly, the selective HER2 blocking antibody herceptin was followed to the market by the dual epidermal growth factor receptor (EGFR)/erbB2 inhibitor, lapatinib **45** in 2007.

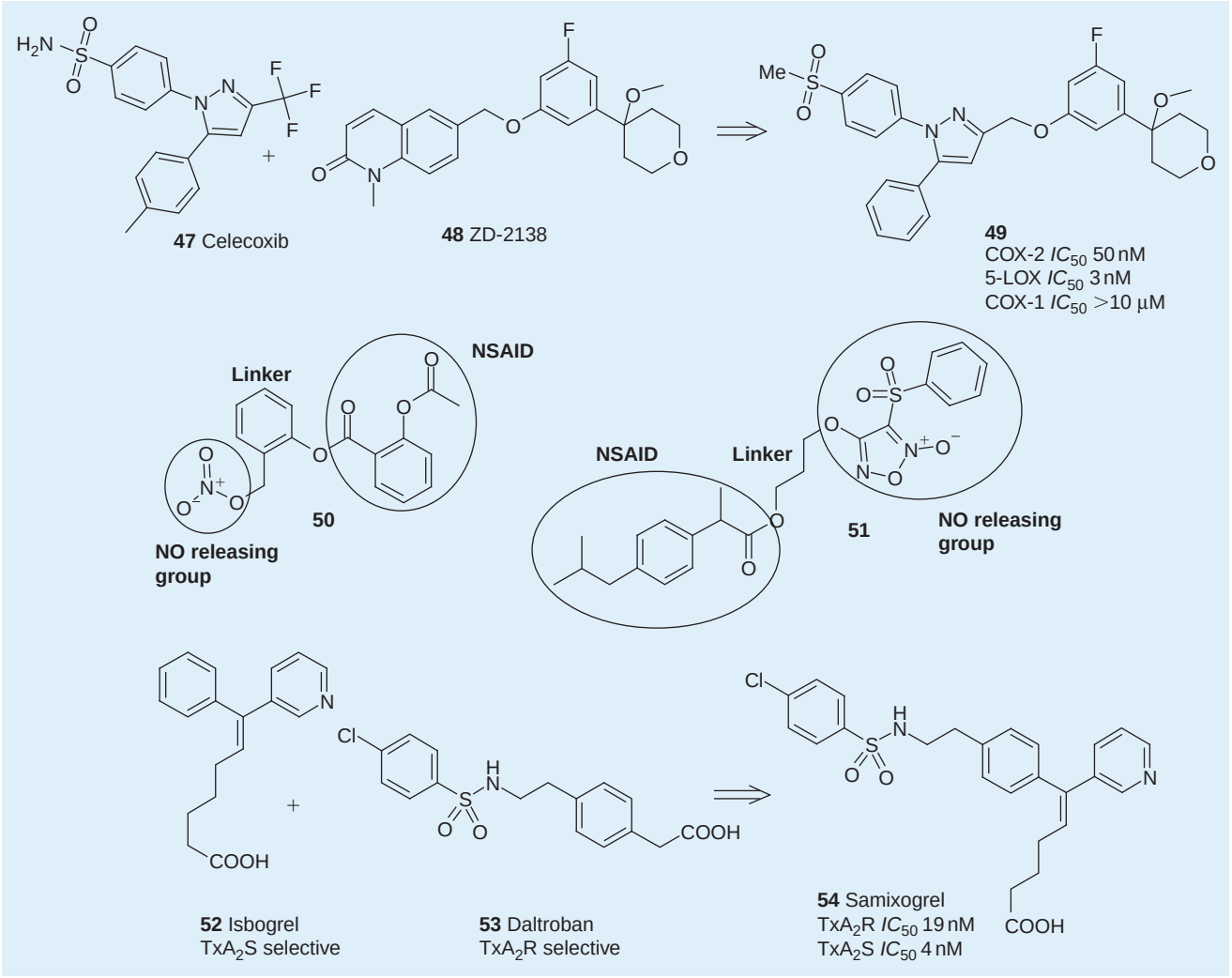


FIGURE 19.12 DMLs targeting the arachidonic acid cascade.

To enhance efficacy, various other kinase targets with a potential role in angiogenesis and tumor growth have been combined with VEGFR-2. Becknell et al developed a dual TIE-2/VEGFR-2 inhibitor **46** by cross-screening molecules from an earlier selective VEGFR-2 project (Figure 19.11) [42]. By inhibiting angiogenesis, such multi-kinase agents are showing promise for the treatment of solid tumors in the breast and kidney that were previously highly resistant to therapy.

H. DMLs Targeting the Arachidonic Acid Cascade

NSAIDs exert their anti-inflammatory effect by inhibiting cyclooxygenases-1 and -2 (COX-1 and COX-2), key enzymes in prostaglandin (PG) biosynthesis from arachidonic acid [43]. Side effects, in particular gastrointestinal ulcerogenic activity and renal toxicity, often limit their use [44]. A single-target strategy resulted in the development of COX-2 inhibitors like celecoxib **47** (Figure 19.12). So-called selective drugs have often been found to possess unexpected polypharmacological profiles and can therefore provide attractive starting points for a DML project. For example, the COX-2 inhibitor celecoxib **47** was reported to potently inhibit carbonic anhydrases hCA II and IX [45].

A number of DML approaches targeting multiple key proteins involved in the arachidonic acid biosynthesis have been reported including COX/5-lipoxygenase (5-LOX), 5-LOX/TxA_2, and TxA_2/TxA_2 synthase (TxS). In particular, the combination of 5-LOX with COX-2 inhibitory activity has attracted much attention in recent years [46]. Henichart et al reported a dual COX-2/5-LOX inhibitor designed by fusing the tricyclic moiety present in

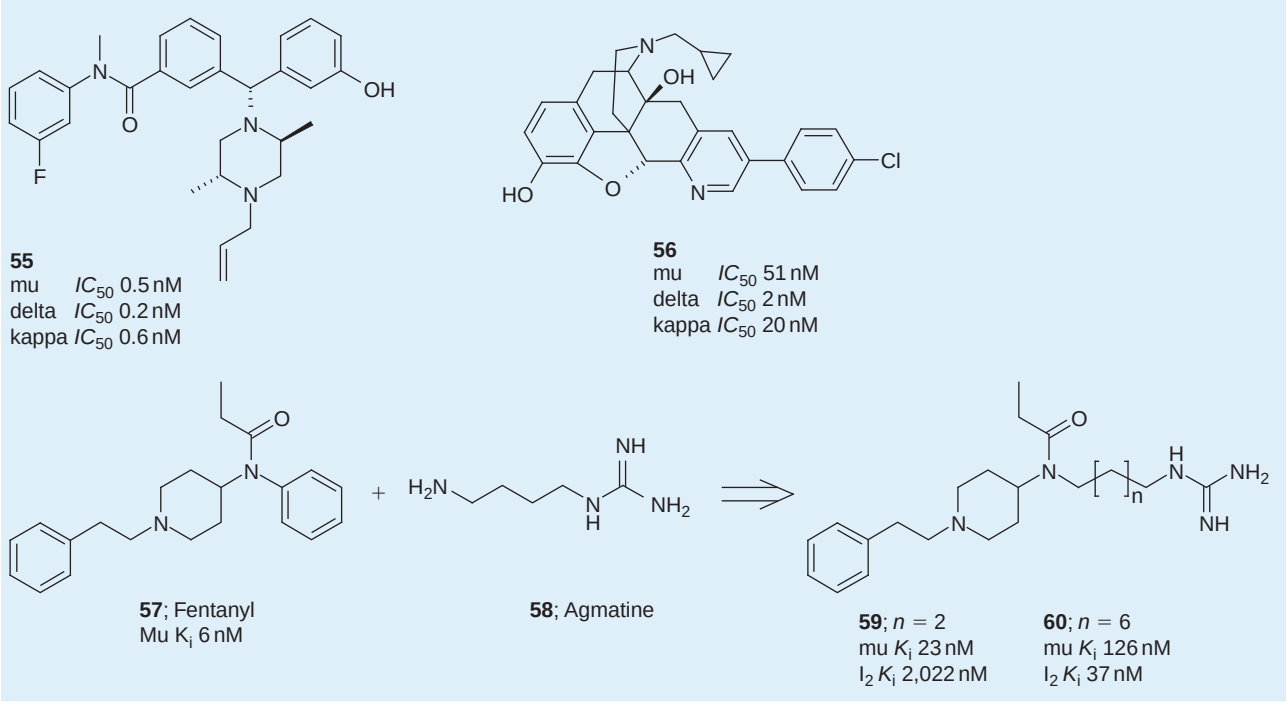


FIGURE 19.13 Mu-opioid-plus DMLs for treating pain.

celecoxib **47** with an aryltetrahydropyran moiety from the 5-LOX inhibitor ZD-23138 **48** (Figure 19.12) [47]. Both starting compounds were completely inactive at the second target, but the resulting DML **49** possessed nanomolar potencies for both enzymes.

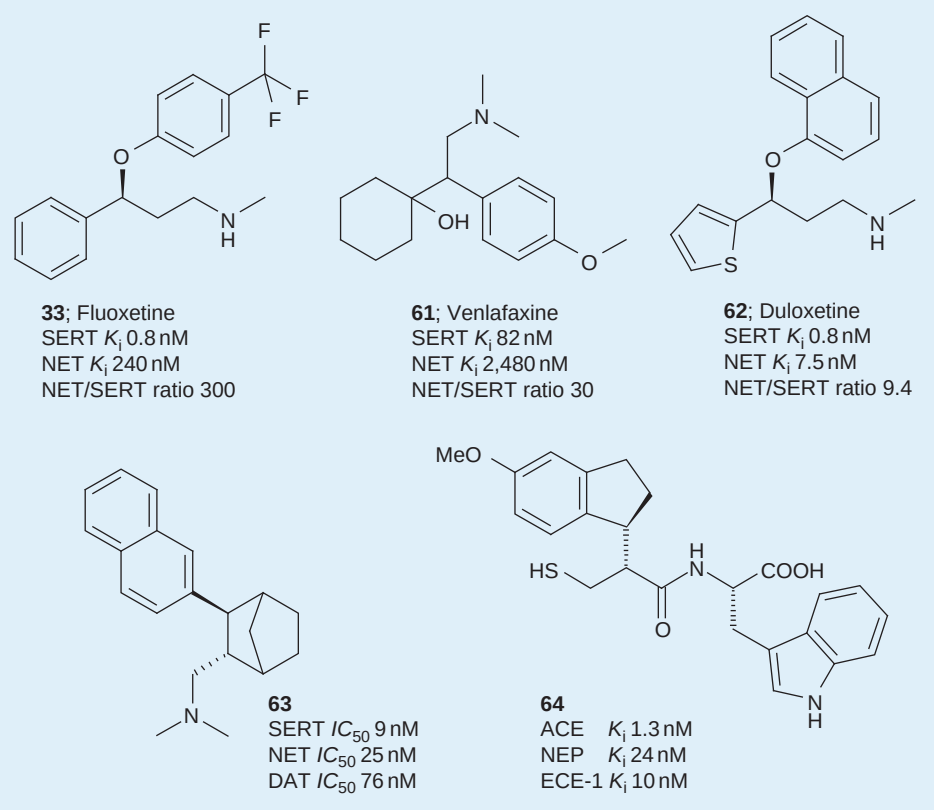
Nitric oxide-releasing NSAIDs, such as NO-aspirin **50** (NCX-4016) and the ibuprofen derivative **51**, contain a cleavable ester linker to a nitric oxide-releasing moiety (Figure 19.12) [48,49]. It was hoped that this dual activity would translate into a superior anti-inflammatory and antithrombotic profile in patients with cardiovascular diseases, while sparing the gastrointestinal tract.

It might seem that bridging two diverse proteomic superfamilies, the GPCRs and the oxidases, to design dual TxA_2/TxS inhibitors would be very challenging. However, a good understanding of target pharmacophore requirements proved to be very helpful in one example [50]. The essential structural features of TxS inhibitors like isbogrel **52** are a pyridine nitrogen and carboxylic group separated by between 8.5 and 10 Å (Figure 19.12). Since TxS is a cytochrome P-450 enzyme, it was postulated that the pyridine moiety forms a complex with the heme group of the enzyme catalytic site. A key feature of TxA_2 receptor antagonists like daltroban **53** is a carboxylic acid separated by a nonspecific spacer from a benzenesulphonamide group. Integration of the TxS and TxA_2 features produced compounds such as Samixogrel **54**, which showed low nanomolar activity at both targets [51].

I. Mu-Opioid-Plus DMLs for Treating Pain

Mu-opioid receptor agonists such as morphine and fentanyl remain the gold standard treatment for severe pain. The use of these agents is limited by mechanism-dependent side effects such as euphoria, respiratory depression, sedation, tolerance, and dependence. In an attempt to maintain efficacy while reducing these side-effect liabilities, mu-opioid activity has been combined with agonist activity at delta- and kappa-opioid receptors **55** [52]. One unusual example of a DML **56**, in terms of contrasting functional activity, combines mu-opioid agonism with delta-opioid antagonism again with the aim of circumventing mu-based side effects [53].

Montero et al combined agonism at the mu-opioid and I_2-imidazoline receptors in a single molecule (Figure 19.13) [54,55]. A guanidinium group from the I_2 ligand agmatine **57** was incorporated into the opioid fentanyl **58**. The lead compound **59** possessed activity at both receptors, but the activity was unbalanced, having significantly higher affinity for the opioid receptor. In this example, the frameworks of the starting compounds are slightly merged with the agmatine-derived alkyl chain replacing the aniline system in fentanyl. The identification of such a "tolerant region" for both receptors is a first key step in any DML program. The compound with an 8-carbon spacer **60** possessed the best balance of activities.



33; Fluoxetine
SERT K_i 0.8 nM
NET K_i 240 nM
NET/SERT ratio 300

61; Venlafaxine
SERT K_i 82 nM
NET K_i 2,480 nM
NET/SERT ratio 30

62; Duloxetine
SERT K_i 0.8 nM
NET K_i 7.5 nM
NET/SERT ratio 9.4

63
SERT IC_{50} 9 nM
NET IC_{50} 25 nM
DAT IC_{50} 76 nM

64
ACE K_i 1.3 nM
NEP K_i 24 nM
ECE-1 K_i 10 nM

FIGURE 19.14 Optimization of a DML profile to enhance efficacy and safety.

IV. OPTIMIZATION OF THE ACTIVITY PROFILE AND WIDER SELECTIVITY

Regardless of whether the lead compound is obtained by a screening and framework-combination approach, the compound will usually lack the optimal ratio of activities. Thus, a medicinal chemist working on a DML lead-optimization project is faced with the twin challenges of balancing the desired activities at an appropriate level while, if necessary, removing any undesired side activities. Establishing what the desired level of modulation for each target should be for optimal efficacy and safety is not a straightforward task. Moreover, understanding the relationship between *in vivo* target modulation and activity in a simple *in vitro* test, such as receptor affinity in a recombinant cell assay, is difficult. Factors such as the distribution of the compound, whether the targets are located in different tissues, the receptor/enzyme densities, and receptor reserve in different tissues will influence the optimal balance of *in vitro* activities. Ideally, knowledge from clinical studies will guide researchers toward the optimal ratio, though for novel mechanisms of action, this clearly will not be available. In the absence of this knowledge, the aim of most historical DML projects has been to obtain the same degree of *in vitro* activity for each target, with the assumption that this will also lead to similar levels of enzyme modulation or receptor occupancy *in vivo*. Assuming a validated animal model is available, the testing of a lead candidate *in vivo* may help to clarify the required ratio of *in vitro* activities.

In the antidepressant field, the historical trend has been toward developing agents with both potent and balanced activity at both the serotonin (SERT) and norepinephrine (NET) transporters, starting with fluoxetine **33**, moving to venlafaxine **61**, and most recently duloxetine **62** (Figure 19.14). Although classified as a dual SERT/NET blocker (SNRI), venlafaxine has a 30-fold difference in *in vitro* potency at the two transporters, meaning that it behaves as a multiple ligand *in vivo* only at high doses [56]. A newer drug, duloxetine, has a more potent and balanced *in vitro* profile [57].

A difference in the *in vitro* activities may sometimes be desirable where a different level of receptor occupancy for each target is associated with a desired pharmacological effect. The "atypical" profile of the antipsychotic drug Clozapine has been variously associated with its lower activity at the D_4 or 5-HT_{2A} receptors compared to the D_2 receptor. Neuroimaging studies have shown that an optimal D_2 receptor

occupancy of 60–70 percent is sufficient to produce an atypical antipsychotic effect, and if D_2 receptor occupancy is too high, the atypical profile can be lost even in the presence of high 5-HT_2 occupancy [58]. Several atypical antipsychotics with low D_2/5-HT_2 binding ratios have now been introduced onto the market, such as ziprasidone **11**.

As the number of targets to be balanced increases, the complexity of the task for a medicinal chemist can increase exponentially. It is therefore not surprising that the vast majority of reported DMLs are dual ligands. However, for targets which are closely related, such as combinations of monoamine transporters, monoamine GPCRs, proteases, or kinases, triple blockers are known. By concurrently blocking the re-uptake of dopamine as well as serotonin and NA, super mixed uptake blockers (SMUBs), such as **63**, may possess mood-elevating properties and deliver better control of depression than either SSRIs or SNRIs. One of potential limitations of the ACE/NEP dual inhibition approach for the management of hypertension and congestive heart failure is an increase in plasma levels of another vasoconstricting peptide, endothelin-I (ET-1). This might be overcome by additionally inhibiting a closely related zinc metallopeptidase, endothelin-converting enzyme-1 (ECE-1). Triple blockers of ACE, NEP, and ECE-1, such as **64**, may produce a beneficial synergistic effect.

In addition to adjusting the ratio of activities, optimizing wider selectivity against a broad panel of targets is often required. Many publications in the DML area do not even discuss the key issue of global selectivity, so it is frequently difficult to judge whether real selectivity for the disease-relevant targets has been achieved. Animal models and subsequent clinical studies can provide essential feedback on the level of cross-reactivity that can be tolerated. In cases where a large number of closely related receptor subtypes or isozymes exist and the compound possesses more than one undesired activity, the task of achieving wider selectivity will be particularly intricate. Two therapeutic areas where this is currently a critical issue are psychiatric drugs, which frequently hit multiple monoamine GPCRs and transporters, and oncology drugs, which often hit multiple kinases. In both these areas, it can be difficult, if not impossible, to obtain absolute selectivity for the desired targets with no affinity for any off-target. This current reality has led to a pragmatic approach whereby DMLs are developed that are deemed to be "selective enough" to be progressed into toxicity testing in animal studies.

DMLs for monoamine GPCRs and transporters frequently hit adrenergic GPCRs that are associated with cardiovascular side effects. Bonnert et al successfully "designed out" adrenergic α_1 receptor activity from a dual dopamine D_2/adrenergic beta-2 (β_2) agonist [59], and Atkinson et al removed adrenergic receptor β_2 activity from a 5-HT_{1A}/SERT ligand [60]. Atypical antipsychotics often have a complex multireceptor profile and off-target activities can be associated with undesirable side effects. In particular, cross-reactivity at the histamine H_1 receptor has been highlighted as the main cause of the weight gain caused by agents such as Clozapine and olanzepine [61]. During the optimization of the antipsychotic agent **15**, Garzya et al had to balance the five desired activities (D_2, D_3, 5-HT_{2A}, 5-HT_{2C}, and 5-HT_6), while avoiding undesired activity at the H_1 receptor, as well as at other monoamine targets, α_{1B}, M_{1-4}, and β_{1-3} [26].

Similarly, discovering multi-kinase inhibitors is complicated by the risk of inhibiting kinases that are critical to normal cellular function. At the present time, it is difficult to design an inhibitor that inhibits two or three kinases specifically while being inactive at all others. Unanticipated activities for even well-studied inhibitors such as imatinib **42** are still being found via panel screening [62]. This information can reveal which kinases are "safe" to inhibit and which are critical to normal cellular function and should be avoided. Although the recently launched drug Sunitinib was designed as a dual VEGFR-2 and PDGFR β inhibitor [41], it was later reported to inhibit no less than eight kinases with IC_{50} values of less than 100 nM and yet has an acceptable side-effect profile in humans (Figure 19.11) [63]. It remains to be seen whether such a pragmatic approach to kinase selectivity profiles can be extended beyond oncology to non-life-threatening disease areas such as inflammation, where side-effect liabilities will be particularly critical. Cardiotoxicity associated with multi-kinase inhibition is one area of possible concern [64]. Screening for multi-kinase inhibitors sometimes provides compounds with undesired off-target activity at non-kinase targets. In a recent example, activity at the hERG ion channel was successfully designed out of a multi-kinase inhibitor [65].

Several examples give encouragement to the medicinal chemist that surprising activity and selectivity profiles can sometimes be achieved. The dual AChE/SERT blocker **35** possesses high selectivity over several closely related targets, including butyrylcholinesterase and the NET/DAT [33]. Similarly the COX-2/5-LOX inhibitor **49** possesses surprising selectivity for COX-2 over COX-1 and the AT_1/ET_A antagonist **22** is inactive at AT_2 and ET_B receptors [47].

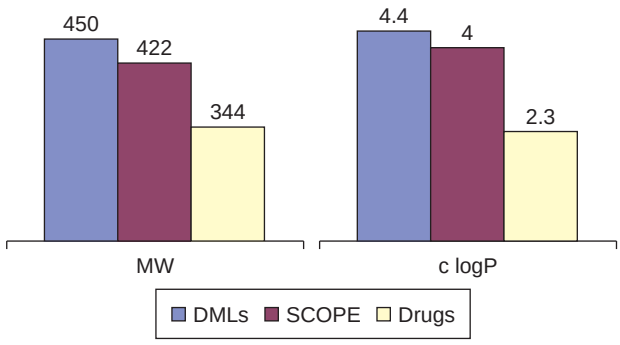


FIGURE 19.15 The median MW and *c* Log *P* values for DMLs are higher than those for oral drugs [67] or a general set of preclinical compounds from Organon's SCOPE database [68].

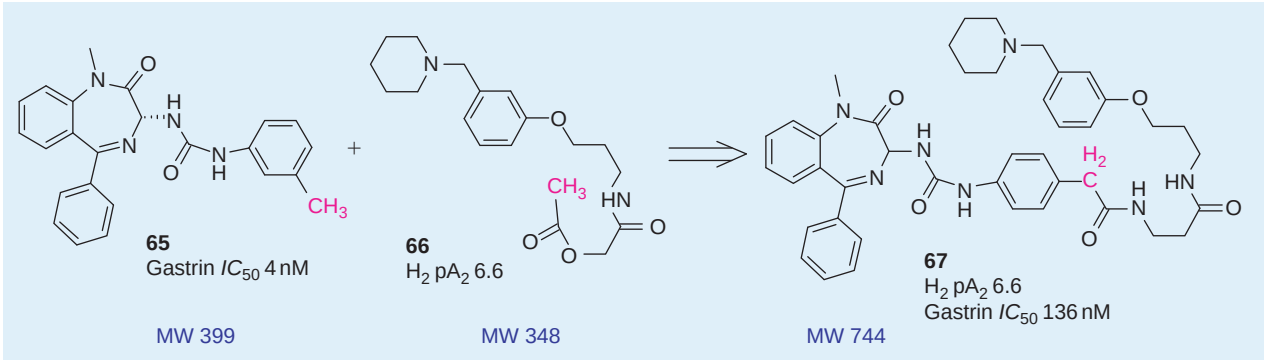


FIGURE 19.16 A example of a "fused" DML with a high MW and low oral absorption.

V. THE PHYSICOCHEMICAL CHALLENGE

Compared to optimizing the balance of affinities and the wider selectivity, an even greater challenge for medicinal chemists when designing multiple ligands is to obtain physicochemical and PK properties consistent with developing an oral drug [2]. The influence of physicochemical properties on the PK behavior of orally administered drugs has been the subject of intense interest over the past few years since the publication of Lipinski's seminal work on the "rule-of-five" in 1997 [66].

On average, the current generation of DMLs have been found to be larger and more lipophilic than marketed drugs [67] or preclinical compounds in general (Figure 19.15) [68,69]. Larger and more lipophilic molecules are often associated with poorer oral absorption profiles, and yet this route of administration is required for most DMLs [66,70]. Optimizing the pharmacokinetics—in addition to attaining a balanced profile—can easily become the most challenging aspect of working with DMLs. One explanation for this has been the popularity of the knowledge-based framework-combination strategy whereby the molecular frameworks from two selective ligands are combined. Given that the selective ligands used as the starting points are already drug-like in size and the extent to which the frameworks can be integrated is often low, this process can result in large property increases that compromise oral bioavailability.

This Achilles' heel of the framework-combination strategy is illustrated by the example in Figure 19.16, wherein the framework of a selective gastrin receptor antagonist **65** was combined with that of a histamine H_2 ligand **66** [71]. Compound **67** is a classic example of a "fused" DML, since the degree of overlap that was possible was just a single carbon atom. The incompatibility of the hydrophobic gastrin pharmacophore with the hydrophilic H_2 pharmacophore produces "tolerated regions" that are only relevant for binding at one of the targets, having the effect of increasing the size of the resulting molecule (MW 744) and compromising oral absorption.

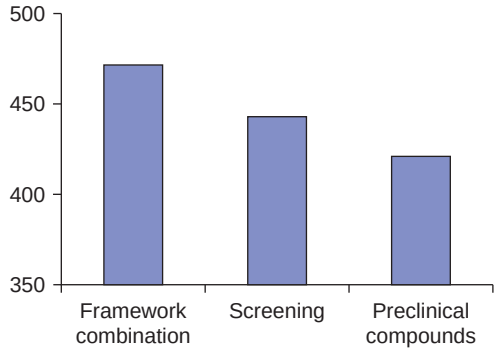


FIGURE 19.17 : Median MW of DMLs derived via framework combination and screening compared to a general set of preclinical compounds.

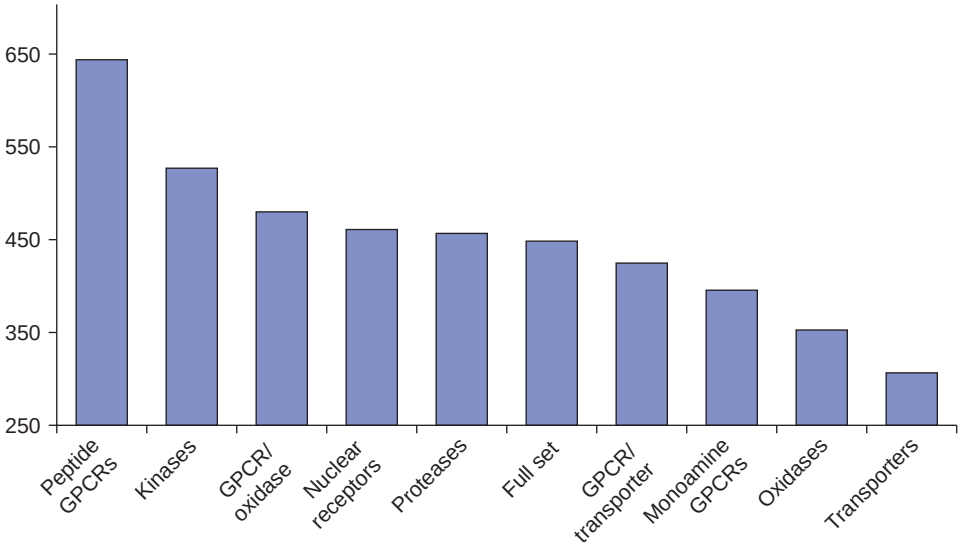


FIGURE 19.18 Median MW of DMLs classified according to proteomic target family.

Nonetheless, the framework-combination approach is a conceptually elegant knowledge-driven strategy that effectively uses SAR knowledge derived from selective-ligand projects. Furthermore, there are successful examples of oral drugs having been discovered by this strategy reaching the market, such as ziprasidone **11** [23]. To achieve an orally active DML, it is important that the degree of framework overlap is maximized and the size and complexity of the selective ligands is minimized. These goals will typically be more feasible for targets with simple endogenous ligands and conserved binding sites, such as monoamine GPCRs and transporters.

The MW for screening-derived DMLs is frequently lower than for the framework-combination strategy, suggesting that this approach may provide a route to smaller and less complex leads (Figure 19.17). A starting compound obtained via screening already possesses multitarget activity to some extent. During optimization, the activities are usually balanced by adding modestly sized groups or modifying the existing functionality. This typically has less of an effect on the overall size and physicochemical properties of the molecule than the combination of two frameworks.

Over recent years, an increasing amount of evidence shows that physicochemical properties are less favorable for the ligands from some proteomic target families of interest in drug discovery than for others, which makes the discovery of orally active drugs for those targets more challenging [68]. Similar trends amongst the target families have also been reported for DMLs [69]. The target family that has consistently given the highest property values for both preclinical compounds in general and DMLs is the peptide GPCRs (Figure 19.18). For example, DMLs for peptide GPCRs had a median MW of 636 and a median *c* Log *P* of 5.1,

figures in excess of those defined in the "rule-of-five" for drug-likeness [66]. At the other end of the spectrum, the ligands for transporters, monoamine GPCRs, and oxidases generally possess favorable physicochemical properties, and the feasibility of such targets for DML projects using a variety of lead-discovery strategies will be relatively high.

The analysis indicates that designing DMLs for peptide GPCRs will be a more difficult endeavor than for other types of GPCR or indeed for selective ligands for individual peptide GPCRs. However, with perseverance and skill, even difficult families such as peptide GPCRs can sometimes be addressed with a framework-combination approach such as the dual AT_1/ET_A antagonist program that delivered compounds, such as **22**, with good oral bioavailability [29]. In such cases, a strong emphasis is often required during lead optimization on simplifying the structure of the lead compound (Figure 19.7).

In a number of other literature examples, the combination of a desirable *in vitro* profile with the PK profile required for the development of an oral drug was not achievable. Where the pharmacophores are fundamentally different, it may not be possible to integrate the requirements of both binding sites into a small, compact molecule, and a higher MW compound may be unavoidable. Inevitably, this will mean that some combinations of targets will be more difficult—if not impossible—to address with a drug-like molecule, illustrated by example 67 in Figure 19.16.

While the framework-combination strategy tends to produce large molecules, this is less of an issue when the goal is the discovery of pharmacological tools for validating novel target combinations or the production of injectable drugs. An important goal for future research in this field—particularly in academic institutions—will be to develop high-quality pharmacological tools to explore the potential therapeutic value of novel target combinations. Here, less attention can be paid to oral exposure and overall developability criteria. More important will be the wider selectivity profile of these pharmacological tools. Portoghese et al reported a range of homo- and hetero-dimeric conjugates with varying linker length designed to investigate pharmacodynamic and organizational features of opioid receptors [72]. For example, recently reported heterodimeric conjugates containing delta-antagonist (naltrindole) and kappa-agonist (ICI-199,441) pharmacophores tethered by variable length oligoglycyl-based linkers **68** (Figure 19.19) were demonstrated to possess significantly greater potency and selectivity compared to their monomer congeners, providing further evidence for the opioid receptor hetero-oligomerization phenomenon [73].

The use of alternative routes of administration, such as intravenous and transdermal, is applicable for some DML applications. High MW DMLs (conjugates) containing a linker group separating the frameworks of the two selective ligands have been successfully employed as IV administered drugs. Van Boeckel et al designed compound **69** with a metabolically stable linker (Figure 19.19) as a dual inhibitor of thrombin, via NAPAP, and antithrombin III (ATIII)-mediated factor Xa, via a heparin-derived pentasaccharide fragment [74]. The polyethylene glycol linker in this antithrombotic compound confers good aqueous solubility, making it suitable for intravenous administration. Since the pentasaccharide demonstrated a much longer half-life (13−15 h in humans) than NAPAP (18 min), the authors postulated that a conjugate with NAPAP might possess improved duration of action. *In vivo* studies confirmed that **69** provided a stronger and longer-lasting antithrombotic effect than a mixture of free pentasaccharide and NAPAP.

VI. SUMMARY

Compounds that act at multiple targets (DMLs) can deliver superior efficacy against complex diseases compared to compounds with high specificity for a single target. The medicinal chemistry literature contains many elegant and increasingly rational approaches to the discovery of DMLs, a small cross-section of which has been described above. To address the "physicochemical challenge," new design strategies will certainly be needed, and some new approaches, such as fragment-based approaches, have been proposed [75]. Inevitably, medicinal chemists will face target combinations that are particularly compelling in terms of biological rationale but problematical from the perspective of combining appropriately balanced *in vitro* and *in vivo* activities with acceptable oral bioavailability, duration of action, and safety. In many cases, alternative formulations and routes of administration will need to be investigated. Without a doubt, the field of multiple ligands will present future generations of medicinal chemists with many challenges but also numerous opportunities to discover a range of new and superior medicines.

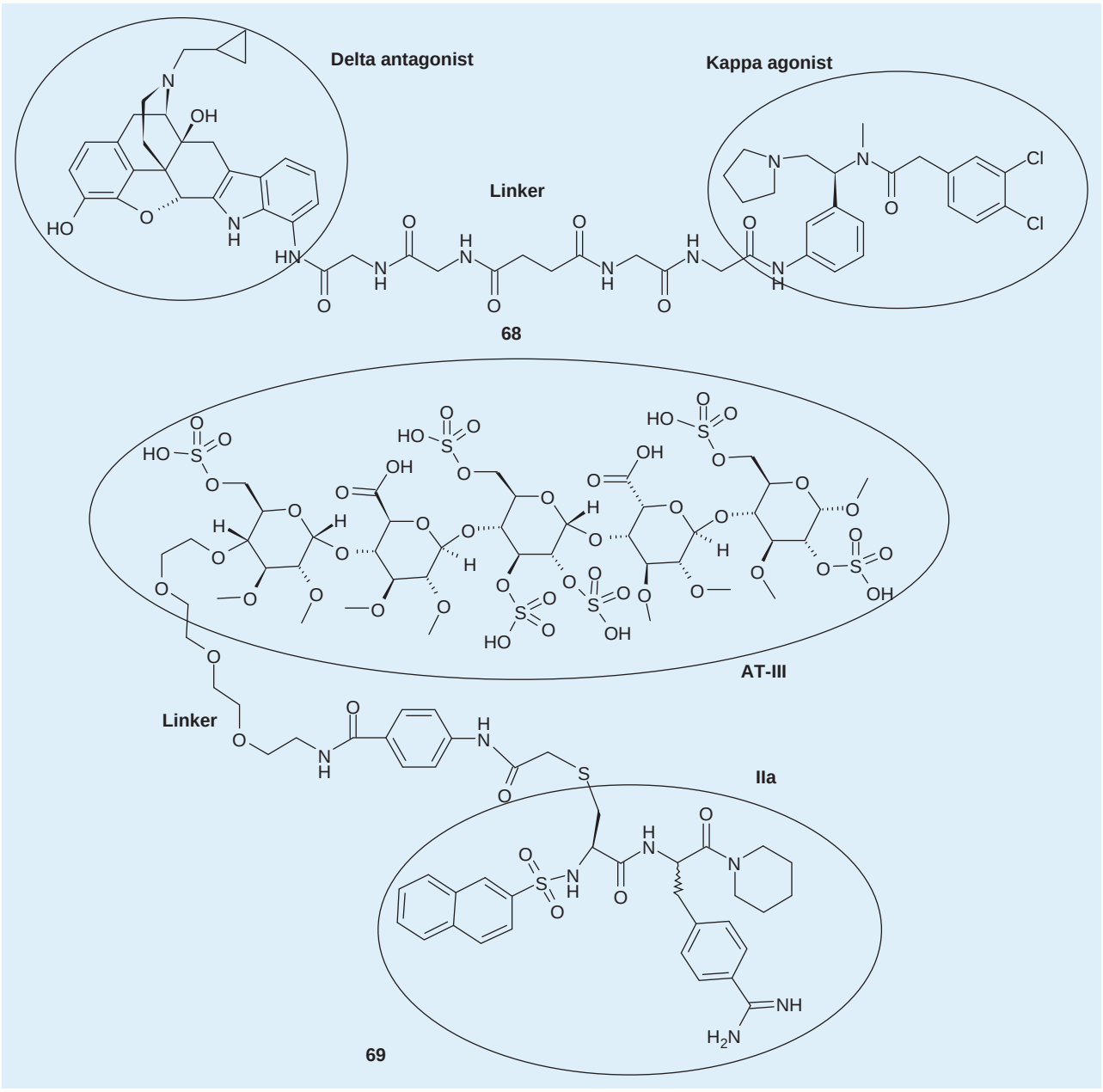


FIGURE 19.19 Linked DMLs derived via a framework-combination approach can make useful biochemical tools or IV administered drugs.

References

[1] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov 2004;3:711–16.
[2] Morphy R, Kay C, Rankovic Z. From magic bullets to designed multiple ligands. Drug Discov Today 2004;9:641–51.
[3] Anighoro A, Bajorath J, Rastelli G. Polypharmacology: challenges and opportunities in drug discovery. J Med Chem 2014;57:7874–87.
[4] Bornot A, Bauer U, Brown A, Firth M, Hellawell C, Engkvist O. Systematic exploration of dual-acting modulators from a combined medicinal chemistry and biology perspective. J Med Chem 2013;56:1197–210.
[5] Law MR, Wald J, Morris JK, Jordan RE. Value of low dose combination treatment with blood pressure lowering drugs: analysis of 354 randomised trials. Br Med J 2003;326:1427–31.
[6] Larder BA, Kemp SD, Harrigan PR. Potential mechanism for sustained antiretroviral efficacy of AZT-3TC combination therapy. Science 1995;269:696–9.
[7] Eisen SA, Miller DK, Woodward RS, Spitznagel E, Przybeck TR. The effect of prescribed daily dose frequency on patient medical compliance. Arch Intern Med 1990;150:1881–4.

[8] Skolnik NS, Beck JD, Clark M. Combination antihypertensive drugs: recommendations for use. Am Fam Physician 2000;61:3049–56.
[9] Glass G. Cardiovascular combinations. Nat Rev Drug Discov 2004;3:731–2.
[10] Herrick TM, Million RP. From the analyst's couch: tapping the potential of fixed-dose combinations. Nat Rev Drug Discov 2007;6:513–14.
[11] Edwards IR, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. Lancet 2000;356:1255–9.
[12] Kramer MS, Last B, Getson A, Reines SA. The effects of selective D_4 dopamine receptor antagonist (L-745, 870) in acutely psychotic inpatients with schizophrenia. Arch Gen Psychiatry 1997;54:567–72.
[13] Campiani G, Butini S, Fattorusso C, Catalanotti B, Gemma S, Nacci V, et al. Pyrrolo[1,3]benzo-thiazepine-based serotonin and dopamine receptor antagonists: molecular modeling, further structure–activity relationship studies, and identification of novel atypical antipsychotic agents. J Med Chem 2004;47:143–57.
[14] Jones HM, Pilowsky LS. New targets for antipsychotics. Expert Rev Neurother 2002;2:61–8.
[15] Stahl SM, Entsuah R, Rudolph RL. Comparative efficacy between venlafaxine and SSRIs: a pooled analysis of patients with depression. Biol Psychiatry 2002;202:1166–74.
[16] Charlier C, Michaux C. Dual inhibition of cyclooxygenase-2 (COX-2) and 5-lipoxygenase (5-LOX) as a new strategy to provide safer non-steroidal anti-inflammatory drugs. Eur J Med Chem 2003;38:645–59.
[17] Chao R, Rapta M, Colson PJ, Lee J. US Patent 0182092; 2005.
[18] Res Med 2003, 97, Suppl. A: S45–2.
[19] Van Niel M, Beer M, Castro J, Cheng S, Evans D, Heald A, et al. Parallel synthesis of 3-aryloxy-2-propanolamines and evaluation as dual affinity 5-HT_{1A} and 5-HT re-uptake ligands. Bioorg Med Chem Lett 1999;9:3243–8.
[20] Mewshaw R, Meagher K, Zhou P, Zhou D, Shi X, Scerni R, et al. Studies toward the discovery of the next generation of antidepressants. Part 2: incorporating a 5-HT1A antagonist component into a class of serotonin reuptake inhibitors. Bioorg Med Chem Lett 2002;12:307–10.
[21] Ryckmans T, Balançon L, Berton O, Genicot C, Lamberty Y, Lallemand B, et al. First dual NK_1 antagonists–serotonin reuptake inhibitors: synthesis and SAR of a new class of potential antidepressants. Bioorg Med Chem Lett 2002;12:261–4.
[22] Lowe J, Seeger T, Nagel A, Howard H, Seymour P, Heym J, et al. 1-Naphthylpiperazine derivatives as potential atypical antipsychotic agents. J Med Chem 1991;34:1860–6.
[23] Howard H, Lowe J, Seeger T, Seymour P, Zorn S, Maloney P, et al. 3-Benzisothiazolylpiperazine derivatives as potential atypical antipsychotic agents. J Med Chem 1996;39:143–8.
[24] Zhao H, Thurkauf A, He X, Hodgetts K, Zhang X, Rachwal S, et al. Indoline and piperazine containing derivatives as a novel class of mixed D_2/D_4 receptor antagonists. Part 1: identification and structure–activity relationships. Bioorg Med Chem Lett 2002;12:3105–9.
[25] Zhao H, He X, Thurkauf A, Hoffman D, Kieltyka A, Brodbeck R, et al. Indoline and piperazine containing derivatives as a novel class of mixed D_2/D_4 receptor antagonists. Part 2: asymmetric synthesis and biological evaluation. Bioorg Med Chem Lett 2002;12:3111–15.
[26] Vincenzo Garzya Ian TF, Gribble AD, Hadley MS, Lightfoot AP, Payne AH, Smith AB, et al. Studies towards the identification of a new generation of atypical antipsychotic agents. Bioorg Med Chem Lett 2007;17:400–5.
[27] Robl J, Sieber-McMaster E, Asaad MM, Bird JE, Delaney NG, Barrish JC, et al. Mercaptoacyl dipeptides as dual inhibitors of angioten-sin converting enzyme and neutral peptidase: preliminary structure–activity studies. Bioorg Med Chem 1994;4:1783–9.
[28] Robl JA, Sun C-Q, Stevenson J, Ryono DE, Simpkins LM, Cimarusti MP, et al. Dual metalloprotease inhibitors: mercaptoacetyl-based fused heterocyclic dipeptide mimetics as inhibitors of angiotensin-converting enzyme and neutral endopeptidase. J Med Chem 1997;40:1570–7.
[29] Murugesan N, Tellew J, Gu Z, Kunst B, Fadnis L, Cornelius L, et al. Discovery of N-isoxazolyl biphenylsulfonamides as potent dual angiotensin II and endothelin A receptor antagonists. J Med Chem 2002;45:3829–35.
[30] Eiser N. Histamine antagonists and asthma. Pharm Ther 1982;17:239–50.
[31] Ohshima E, Takami H, Harakawa H, Sato H, Obase H, Miki I, et al. Dibenz[b,e]oxepin derivatives: novel antiallergic agents possessing thromboxane A_2 and histamine H_1 dual antagonizing activity, 1. J Med Chem 1993;36:417–20.
[32] Lewis T, Bayless L, Eckman J, Ellis J, Grewal G, Libertine L, et al. 5-Lipoxygenase inhibitors with histamine H1 receptor antagonist activity. Bioorg Med Chem Lett 2004;14:2265–8.
[33] Toda N, Tago K, Marumoto S, Takami K, Ori M, Yamada N, et al. A conformational restriction approach to the development of dual inhibitors of acetylcholinesterase and serotonin transporter as potential agents for Alzheimer's disease. Bioorg Med Chem 2003;11:4389–415.
[34] Sterling J, Herzig Y, Goren T, Finkelstein N, Lerner D, Goldenberg W, et al. Novel dual inhibitors of AchE and MAO derived from hydroxy aminoindan and phenethylamine as potential treatment for Alzheimer's disease. J Med Chem 2002;45:5260–70.
[35] Sagi Y, Drigues N, Youdim MB. The neurochemical and behavioral effects of the novel cholinesterase-monoamine oxidase inhibitor, ladostigil, in response to l-dopa and l-tryptophan, in rats. Br J Pharmacol 2005;146:543–52.
[36] Henke B. Peroxisome proliferator-activated receptor α/γ dual agonists for the treatment of type 2 diabetes. J Med Chem 2004;47:4118–27.
[37] Xu Y, Rito CJ, Etgen GJ, Ardecky RJ, Bean SJ, Bensch WR, et al. Design and Synthesis of α-aryloxy-α-methylhydrocinnamic acids: a novel class of dual peroxisome proliferator-activated receptor α/γ agonists. J Med Chem 2004;47:2422–5.
[38] Lackey KE. Lessons from the drug discovery of lapatinib, a dual ErbB1/2 tyrosine kinase inhibitor. Curr Top Med Chem 2006;6:435–60.
[39] Daub H, Specht K, Ullrich A. Strategies to overcome resistance to targeted protein kinase inhibitors. Nat Rev Drug Discov 2004;12:1001–10.
[40] Boschelli D, Wang Y, Johnson S, Wu B, Ye F, Sosa A, et al. 7-Alkoxy-4-phenylamino-3-quinolinecarbonitriles as dual inhibitors of Src and Abl kinases. J Med Chem 2004;47:1599–601.
[41] Sun L, Liang C, Shirazian S, Zhou Y, Miller T, Cui J, et al. Discovery of 5-[5-fluoro-2-oxo-1,2-dihydroindol-(3Z)-ylidenemethyl]-2,4-dimethyl-1H-pyrrole-3-carboxylic acid (2-diethylaminoethyl)amide, a novel tyrosine kinase inhibitor targeting vascular endothelial and platelet-derived growth factor receptor tyrosine kinase. J Med Chem 2003;46:1116–19.

[42] Becknell NB, Zulli AL, Angeles TS, Yang S, Albom MS, Aimone LD, et al. Novel C-3 N-urea, amide, and carbamate dihydroindazolo[5,4-a]pyrrolo[3,4-c] carbazole analogs as potent TIE-2 and VEGF-R2 dual inhibitors. Bioorg Med Chem Lett 2006;16:5368–72.
[43] Vane JR. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. Nat New Biol 1971;231:232–5.Dannhard G, Kiefer W. Cyclooxigenase inhibitors-current status and future prospects. Eur J Med Chem 2001;36:109–26.
[44] Fosslien E. Adverse effects of nonsteroidal antiinflammatory drugs on the gastrointestinal system. Ann Clin Lab Sci 1998;28:67–81.
[45] Weber A, Casini A, Heine A, Kuhn D, Supuran CT, Scozzafava A, et al. Unexpected nanomolar inhibition of carbonic anhydrase by COX-2-selective celecoxib: new pharmacological opportunities due to related binding site recognition. J Med Chem 2004;46:550–7.
[46] Goossens L, Pommery N, Henichart JP. COX-2/5-LOX: dual acting anti-inflammatory drugs in cancer chemotherapy. Curr Top Med Chem 2007;7:283–96.
[47] Barbey S, Goossens L, Taverne T, Cornet J, Choesmel V, Rouaud C, et al. Synthesis and activity of a new methoxytetrahydropyran derivative as dual cyclooxygenase-2/5-lipoxygenase inhibitor. Bioorg Med Chem Lett 2002;12:779–82.
[48] Di Napoli MF, Papa F. NCX-4016 NicOX. Curr Opin Invest Drugs 2003;4:1126–39.
[49] Lolli ML, Cena C, Medana C, Lazzarato L, Morini G, Coruzzi G, et al. A new class of Ibuprofen derivatives with reduced gastrotoxicity. J Med Chem 2001;44:3463–8.
[50] Kato K, Ohkawa S, Terao S, Terashita ZI, Nishikava K. Thromboxane synthetase inhibitors (TXSI): design, synthesis, and evaluation of a novel series of omega-pyridylalkenoic acids. J Med Chem 1985;28:287–94.
[51] Soyka R, Heckel A, Nickl J, Eisert W, Muller TH, Weisenberger H. 6,6-Disubstituted hex-5-enoic acid derivatives as combined thromboxane A2 receptor and synthetase inhibitors. J Med Chem 1994;37:26–39.
[52] Gengo PJ, Pettit HO, O'Neill SJ, Wei K, McNutt R, Bishop MJ, et al. DPI-3290 [(+)-3-((alpha-R)-alpha-((2S,5R)-4-Allyl-2,5-dimethyl-1-piperazinyl)-3-hydroxybenzyl)-N-(3-fluorophenyl)-N-methylbenzamide]: mixed opioid agonist with potent antinociceptive activity. J Pharmacol Exp Ther 2003;307:1221–6.
[53] Ananthan S, Kezar III HS, Carter RL, Saini SK, Rice KC, Wells JL, et al. Synthesis, opioid receptor binding, and biological activities of naltrexone-derived pyrido- and pyrimidomorphinans. J Med Chem 1999;42:3527–38.
[54] Montero A, Goya P, Jagerovic N, Luis F, Callado L, Meana J, et al. Guanidinium and aminoimidazolinium derivatives of N-(4-piperidyl)propanamides as potential ligands for opioid and I_2-imidazoline receptors: synthesis and pharmacological screening. Bioorg Med Chem 2002;10:1009–18.
[55] Dardonville C, Jagerovic N, Callado L, Meana J. Fentanyl derivatives bearing aliphatic alkaneguanidinium moieties: a new series of hybrid molecules with significant binding affinity for μ-opioid receptors and I_2-imidazoline binding sites. Bioorg Med Chem Lett 2004;14:491–3.
[56] Stahl SM. Comparative efficacy between venlafaxine and SSRIs: a pooled analysis of patients with depression. Biol Psychiatry 2002;52:1166–74.
[57] Bymaster FP, Beedle EE, Findlay J, Gallagher PT, Krushinski JH, Mitchell S, et al. Duloxetine (Cymbalta TM), as dual inhibitor of serotonin and norepinephrine reuptake. Bioorg Med Chem Lett 2003;13:4477–80.
[58] Nyberg S, Eriksson B, Oxenstierna G, Halldin C, Farde L. Suggested minimal effective dose of risperidone based on PET-measured D2 and 5-HT(2A) receptor occupancy in schizophrenic patients. Am J Psychiatry 1999;156:869–75.
[59] Bonnert RV, Brown RC, Chapman D, Cheshire DR, Dixon J, Ince F, et al. Dual D_2-receptor and b_2-adrenoreceptor agonists for the treatment of airway diseases. 1. Discovery and biological evaluation of some 7-(2-aminoethyl)-4-hydroxybenzothiazol-2(3H)-one analogues. J Med Chem 1998;41:4915–17.
[60] Atkinson P, Bromidge S, Duxon M, Laramie M, Gaster L, Hadley M, et al. 3,4-Dihydro-2H-benzoxazinones are 5-HT1A receptor antagonists with potent 5-HT reuptake inhibitory activity. Bioorg Med Chem Lett 2005;15:737–41.
[61] Meltzer HY. Illuminating the molecular basis for some antipsychotic drug-induced metabolic burden. Proc Natl Acad Sci USA 2007;104:3019–20.
[62] Fabian MA, Biggs WH, Treiber DK, Atteridge CE, Azimioara MD, Benedetti MG, et al. A small molecule-kinase interaction map for clinical kinase inhibitors. Nat Biotechnol 2005;23:329–36.
[63] Roskoski R. Sunitinib: a VEGF and PDGF receptor protein kinase and angiogenesis inhibitor. Biochem Biophys Res Commun 2007;356:323–8.
[64] Force T, Krause DS, van Etten RA. Molecular mechanisms of cardiotoxicity of tyrosine kinase inhibition. Nat Rev Cancer 2007;7:332–44.
[65] Dinges J, Albert DH, Arnold LD, Ashworth KL, Akritopoulou-Zanze I, Bousquet PF, et al. 1,4-Dihydroindeno[1,2-c]pyrazoles with acetylenic side chains as novel and potent multitargeted receptor tyrosine kinase inhibitors with low affinity for the hERG Ion channel. J Med Chem 2007;50:2011–29.
[66] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23:3–25.
[67] Vieth M, Siegel MG, Higgs RE, Watson IA, Robertson DH, Savin KA, et al. Characteristic physical properties and structural fragments of marketed oral drugs. J Med Chem 2004;47:224–32.
[68] Morphy JR. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. J Med Chem 2006;49:2969–78.
[69] Morphy JR, Rankovic Z. The physicochemical challenges of designing multiple ligands. J Med Chem 2006;49:4961–70.
[70] Veber DF, Johnson SR, Cheng HY, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45:2615–23.
[71] Kawanishi Y, Ishihara S, Tsushim T, Seno K, Miyagoshi M, Hagishita S, et al. Synthesis and pharmacological evaluation of highly potent dual histamine H_2 and gastrin receptor antagonists. Bioorg Med Chem Lett 1996;6:1427–30.
[72] Portoghese PS. From models to molecules: opioid receptor dimers, bivalent ligands, and selective opioid receptor probes. J Med Chem 2001;44:2259–69.
[73] Daniels DJ, Kulkarni A, Xie Z, Bhushan RG, Portoghese PS. A bivalent ligand (KDAN-18) containing δ-antagonist and κ-agonist pharmacophores bridges δ2 andκ1 opioid receptor phenotypes. J Med Chem 2005;48:1713–16.
[74] Buijsman RC, Basten JE, van Dinther TG, van der Marel GA, van Boeckel CA, van Boom JH. Design and synthesis of a novel synthetic NAPAP-pentasaccharide conjugate displaying a dual antithrombotic action. Bioorg Med Chem Lett 1999;9:2013–18.
[75] Morphy R, Rankovic Z. Fragments, network biology, and designing multiple ligands. Drug Discov Today 2007;12:156–60.

CHAPTER

20

Selective Optimization of Side Activities (SOSA) in Drug Discovery

Tim Jonckers

Janssen Research & Development, Janssen Infectious Diseases – Diagnostics BVBA, Turnhoutseweg, Beerse, Belgium

OUTLINE

I. INTRODUCTION

If one were to ask a medicinal chemist whether he or she would deliberately design a molecule that acts on multiple targets, the most probable answer would be "no." At first glance, it seems to be a bit counter-intuitive to deliberately design molecules that are not selective toward one specific target but act upon different targets at the same time. The optimization and development of such "promiscuous" drugs appear to be more complicated because such dual activity requires a deep understanding of the role and functions not only of the targets of interest themselves but also of the underlying mechanisms that relates them to each other. Numerous examples of lead optimization programs have successfully optimized lead molecules from moderate- to high-affinity binders (obtaining single digit nM or even pM potencies) toward a single target. However, for a molecule to exert an effect on more than one target, high-affinity binders appear less attractive per se, since this would reduce the chance of interaction with another target. As a consequence, such an intrinsic "moderate-affinity profile" increases the importance of an optimized pharmacokinetic profile, which only adds to the complexity. Despite these apparent hurdles, several examples of such polypharmacology drugs are reported, especially in the field of oncology and central nervous system (CNS) diseases. A very recent review describes several examples [1].

A second—more literal—definition of multitarget drugs covers the strategy by which patients are given two or more individual drugs that each act via a different mode of action (MOA) to treat a specific disease or condition (Figure 20.1). This approach has found widespread application in the treatment of human immunodeficiency virus (HIV) infected patients, where combinations of nucleoside reverse transcriptase inhibitors (NRTIs), nonnucleoside reverse transcriptase inhibitors (NNRTIs), and protease inhibitors (PIs) were used [2]. Likewise, this approach has also found to be useful in treating patients infected with the hepatitis C virus (HCV) [3]. To reduce the bill burden of such combined therapy with the aim of improving patient adherence, so-called fixed dose combinations (FDCs) have been developed. Examples of this are Atripla® (used in HIV treatment), which combines an NNRTI and two NRTIs in a single tablet, and Vytorin® (used to treat hypercholesterolemia), which is a combination of the cholesterol absorption inhibitor ezetimibe and the statin simvastatin [4,5].

FIGURE 20.1 Definitions of multitarget drug strategies.

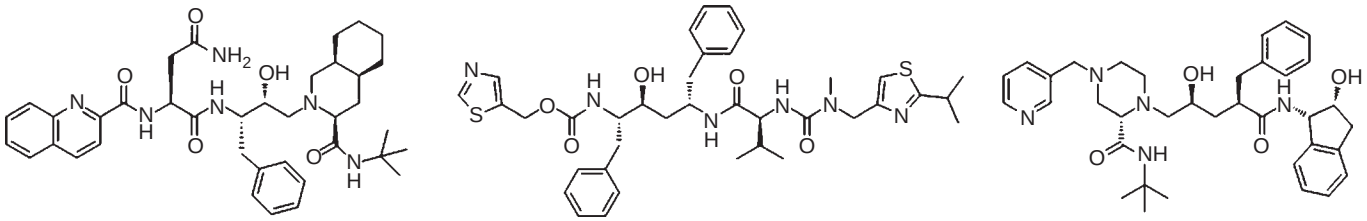
Finally, a third definition of multitarget drugs exists that has a more serendipitous meaning. Despite the extensive preclinical and clinical screening aimed at making molecules safer and more selective toward a chosen target, unforeseen secondary pharmacology is sometimes observed for molecules, and in some cases a novel application for a drug is discovered that is distinct from the one for which it was originally designed. There are a few remarkable examples of such accidental findings in drug discovery, and while these unintended discoveries make great stories, in general they are relatively rare. In contrast, a more rational approach has emerged in recent years, where chemists intentionally try to find a novel application for an existing drug that—for whatever reason—is no longer considered a viable option for its original application. This drug repurposing, often of compounds that already reached the clinical stage, can be the result of renewed screening on novel targets and pathogens. Also, new insights—which were unknown at the time of the initial drug discovery program—originating from a better understanding of certain disease-causing pathways and novel target elucidation strategies can also lead to the repositioning of certain drugs. Finally, *in silico*-based data mining strategies are used to build up novel models of diseases without singling out specific targets. Through combination of large biological data sets with in-depth knowledge about relevant biological targets, markers, and/or pathways, novel insights and possibilities for treatment of certain diseases can be discovered [6].

Several reasons contribute to the fact that "therapeutic switching" is considered an appropriate strategy to pursue. First, it may result in bringing new therapies to developing countries, as was the case for Miltefosine (see below). Secondly, the known safety profile of existing drugs would speed up their positioning for a novel application, since the necessary preclinical and clinical safety evaluations would already be completed. That obviously would positively impact the costs related to a novel commercialization. Current estimates range up to US $800 million for the discovery up to marketing of a single compound. If one were to find a "second life" for a drug, this would represent a significant reduction in costs, and the development time might be substantially reduced.

The purpose of this chapter is to highlight a few examples of drugs that *senso strictu* can be considered multitargeting drugs by any of the definitions given above and that have been repositioned for a novel application.

II. RITONAVIR: REJUVENATING A SUBOPTIMAL DRUG

In 1968, Spencer Silver working at 3M was attempting to create a novel super strong adhesive that would have potential application in the aerospace industry. While working diligently, he mistakenly obtained a weak adhesive substance that by no means had the properties he sought. It did, however, have two interesting features. First, the material could be peeled away from any material without leaving any residue, and—more



1 Saquinavir 2 Ritonavir 3 Indinavir

FIGURE 20.2 Structures of the first three approved HIV PIs.

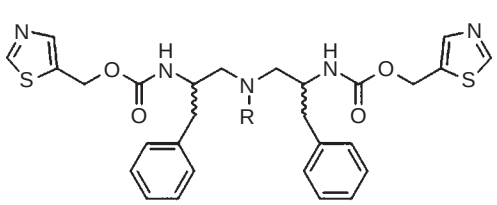
importantly—it seemed to be re-usable. When Silver spoke to his management about his finding, it was considered uninteresting and having no commercial viability. They therefore urged him to shelve his stuff and focus on his initial goal. At the time, without Silver himself could not come up with a marketable use for it, but he kept on advertising it to his colleagues without fully understanding the potential of his novel substance. Five years later, fate provided Silver a helping hand. Chemical engineer Art Fry realized the potential of the material when he looked for something that would prevent his song page markers from falling out of his hymnal. Fry suggested trying to find a way to attach the adhesive to a piece of paper that could then be stuck onto anything. The famous Post-it® notes were born, which since then have become familiar items in offices throughout the world [7].

The discovery and use of the drug Ritonavir has a striking resemblance to the story above. In both cases, a product was developed that was found to have suboptimal properties for its original application, but later a novel usage was found that could not have been predicted upfront.

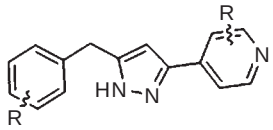
Few diseases have seen such tremendous progress in the development of novel drugs as there is for HIV. The pathogen itself was discovered and isolated in 1983. Soon thereafter, it was recognized as the causative agent of the acquired immunodeficiency syndrome (AIDS) [8]. This condition is characterized by a severe lowering of the CD4 + T cells, which are key immune response regulators. Reduction thereof makes a patient much more vulnerable for "opportunistic infections" to which a healthy person would not be sensitive. Ultimately, HIV patients would die as a result of these otherwise innocent infections. About thirty years ago, a patient diagnosed with HIV would effective hear his death sentence being announced, as without the availability of a proper treatment at that time, HIV would progress rapidly into AIDS. Thankfully, combined efforts from governmental sites, academic groups, and pharmaceutical companies has resulted in the gradual introduction of several new classes of drugs that have transformed the lives of millions of HIV patients from living with a lethal disease into getting older with a chronic one. To date, more than twenty different drugs representing six distinct MOAs have been approved for the treatment of HIV [9].

Until 1995, only NRTIs were available as chemotherapeutics for the treatment of HIV infection. While this class of compounds resulted in delaying the progression of HIV infection into AIDS, its success was limited by the rapid emergence of resistant HIV strains that were no longer sensitive to the NRTIs. Drugs acting via another MOA were urgently needed. In the two year span between 1995 and 1996, three PIs were approved [10]. This class of compounds targets the viral HIV-1 protease enzyme, which plays a critical role in the viral replication cycle. PIs bind into the active site of the HIV-protease by mimicking a transition state that resembles the one that is formed when the protease cleaves the Gag and Gag-Pol polyproteins, its natural substrates. The use of PIs has been a major breakthrough in the therapy for HIV-1 infection, substantially reducing morbidity and mortality in infected individuals. Early representatives of this class were Saquinavir (SQV, **1**) and Ritonavir (RTV, **2**), followed by Indinavir (IDV, **3**) (Figure 20.2).

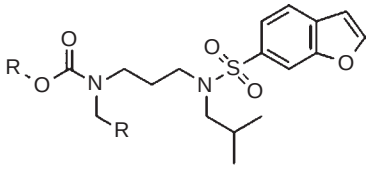
While PIs were originally studied in stand-alone therapy settings (not in combination with other MOA drugs), it became quickly apparent that the clinical efficacy of such a therapy was limited—similar to what was seen for NRTIs—due to the eventual appearance of viruses that contained PI resistant mutations. These findings, together with the introduction of several other MOA drugs like the NNRTI Nevirapine, paved the way for what is known as highly-active antiretroviral therapy (HAART), which combines at least three different drugs in one treatment regimen [11]. Later on, clever formulation work allowed the development of FDCs like Atripla, which combines several drugs from different classes in a single pill, making therapy adherence much easier and transforming HIV into a manageable disease with much-improved life expectancy.



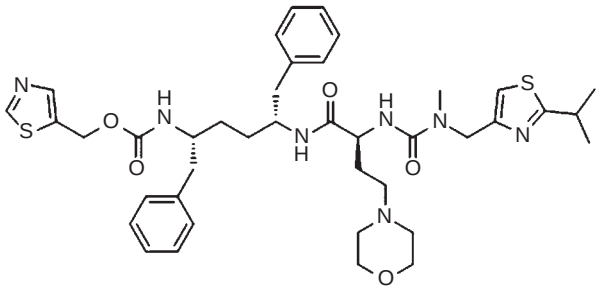
Abbott Laboratories



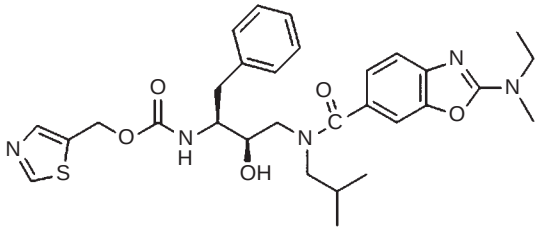
Pfizer



Sequoia Pharmaceuticals



4 Cobicistat / Gilead Sciences



5 TMC558445 /Janssen

FIGURE 20.3 Cobicistat and other CYP3A4 inhibiting scaffolds.

Being first generation PIs, relatively high doses of **1** and **2** were needed to ensure therapeutic efficacy. For **2**, a BID dosing of 600 mg was needed, and at that dose several unwanted side effects were reported, including gastrointestinal discomfort and headaches. More importantly, alterations in serum lipids were noted, including elevated cholesterol and triglyceride levels. These observations, together with the introduction of more potent PIs with improved tolerability profile, resulted the abandonment of RTV as PI in HIV treatment regimens.

When researchers studied the possible causes that led to the emergence of resistance for PIs, two important observations were made. First, a clear inverse relationship was found between resistance mutations popping up and the trough concentration of the inhibitor in the plasma [12]. It was concluded that maintaining high enough plasma concentrations of the active inhibitor was of critical importance for a successful therapy. Secondly, in a clinical study aimed to prevent or at least delay resistance emergence by combining **1** and **2** (which have a different resistance mutations selection), it was observed that co-dosing of **1** and **2** had a beneficial effect on the pharmacokinetic parameters of **1**, which in turn resulted in improved clinical efficacy [13]. This finding meant the start of Ritonavir's "second life" as a pharmacokinetic enhancer.

Although more potent and better tolerated PIs than **1** and **2** were developed over the years [14], they all suffered from a similar problem. Their primary metabolism occurs by the cytochrome P450 enzymes CYP3A4 and CYP3A5. In addition, most PIs are also substrates for efflux systems like P-glycoprotein (P-gp). These factors combined resulted in suboptimal PK profiles when PIs were dosed alone, which could only be compensated by increasing the dose. To avoid this, the majority of PI containing regimens involve the co-administration of a low dose of **2** together with the PI in order to ensure sufficient exposure to the latter agent. The mechanism by which **2** exerts its pharmacokinetic "boosting" is primarily driven by its potent CYP450-inhibiting activity [15]. Detailed spectroscopic studies showed that the N atom of the 5-thiazole group of **2** is able to interact directly with the iron atom present in the heme group of CYP3A4. This interaction results in an impaired functionality of the enzyme [16,17]. In addition to this, **2** also inhibits the P-glycoprotein (P-gp) transport system. As a result, the pharmacokinetic profile of concomitantly administered PIs is beneficially influenced (e.g., increased C_{max} and AUC-values, slower elimination) in RTV boosted regimens, which typically contain a daily dose of 100–200 mg of **2**.

As mentioned, the regular use of **2** is found to be associated with side effects such as gastrointestinal adverse events, changes in serum lipids, insulin resistance, lipoatrophy, and CYP induction. As a consequence, novel derivatives devoid of the drawbacks associated with the use of **2** have been pursued, and various groups have reported on their progress in this area (Figure 20.3) [18–21]. Noteworthy is the fact that in many cases the thiazolyl fragment is retained.

TABLE 20.1 Inhibitory Potency of **2** and **4** Against Human Hepatic Microsomal P450 Cytochromes

		Ritonavir (2)	Cobicistat (4)
Enzyme	**Activity**	IC_{50} (μM)	
CYP3A	midazolam 1′-hydroxylase	0.107	0.154
	testosterone 6β-hydroxylase	0.116	0.151
CYP1A2	phenacetin *O*-deethylase	>25	>25
CYP2B6	bupropion 4-hydroxylase	2.9	2.8
CYP2C8	paclitaxel-6α-hydroxylase	2.8	>25
CYP2C9	tolbutamide-4-hydroxylase	4.4	>25
CYP2C19	*S*-mephenytoin 4′-hydroxylase	>25	>25
CYP2D6	dextromethorphan *O*-demethylase	2.8	9.2

Cobicistat (**4**, formerly GS-9350), a compound structurally related to RTV, has emerged as a promising candidate and was recently approved as novel pharmacokinetic enhancer, an alternative to **2** [22]. The compound has a comparable CYP3A4 inhibition to **2** (Table 20.1), but more importantly it shows an improved tolerability and side effect profile. In addition, the superior physicochemical properties of **4** allow it to be co-formulated with several other drugs. Stribild®, a fixed dose combination of four different drugs (tenofovir DF/elvitegravir/emtricitabine and **4**), also known as the "quad-pill," was approved for treatment of HIV infected patients in 2012 [23]. Very recently, a fixed dose combination consisting of **4** with Prezista, an HIV PI, was also approved by the FDA under the name Prezcobix® [24].

While the concept of "boosting" originated from within the field of HIV, it has successfully been applied in treatments for other diseases. Recently, the FDA approved Viekira Pak™ for combatting HCV infection [25]. It combines three direct antivirals with **2** and has been the most recent introduction of a shorter, better tolerated, and much more efficacious treatment option for treatment of HCV patients. The above story of repositioning Ritonavir from PI to pharmacokinetic enhancer illustrates how an unexpected clinical observation can lead to the repositioning of a compound, giving it a second life and ultimately leading to novel treatment paradigms.

III. SILDENAFIL, SIDE EFFECTS ARE NOT ALWAYS BAD

Angina pectoris (AP)—better known as angina—is a term that refers to the unpleasant feeling of upper chest pains and an elevated sensation of pressure (choking sensation) on the chest that is the result of an insufficient blood flow to the heart muscle. Many different causes for AP are known, including anemia, cardiac arrhythmia, and heart failure. In many cases, however, AP is the result of coronary artery disease (CAD) in which plaque residues that are enriched with fatty acids and cholesterol thwart the blood flow to the heart. Differences in the occurrence of AP related symptoms lead to different classifications of the condition and in medicine; stable angina, unstable angina, and microvascular angina are distinguished. While the condition itself creates obvious discomfort, it can be the prelude to a more dangerous condition, namely a myocardial infarction, which is better known as a heart attack. It is believed that in the US on an annual basis, more than 500 thousand people suffer from AP, especially in the age group above 60. An unhealthy lifestyle (smoking, obesity, lack of exercise) appears to play a predominant role in the disease.

The obvious treatment of AP is aimed at restoring normal blood flow to the heart. In the most extreme cases, surgical interventions are needed in which permanent vasodilatation is achieved by inserting an arterial stent to dilate the artery, a procedure which is usually preceded by a balloon angioplasty [26]. As an alternative, a by-pass operation may be necessary. To provide immediate relief to patients suffering from AP, nitroglycerine (**6**) is often given. This drug can be administered sublingual and acts through the formation of nitric oxide (NO), which is a potent vasodilation agent. In many cases, AP sensitive patients will be on a drug regimen that may include β-blockers (e.g., propranolol (**7**)), calcium channel blockers (e.g., nifedipine (**8**)), ACE inhibitors (e.g., Captopril (**9**)), and statins (e.g., Lovastatin (**10**)), which contribute to keeping AP under control through different MOAs (Figure 20.4).

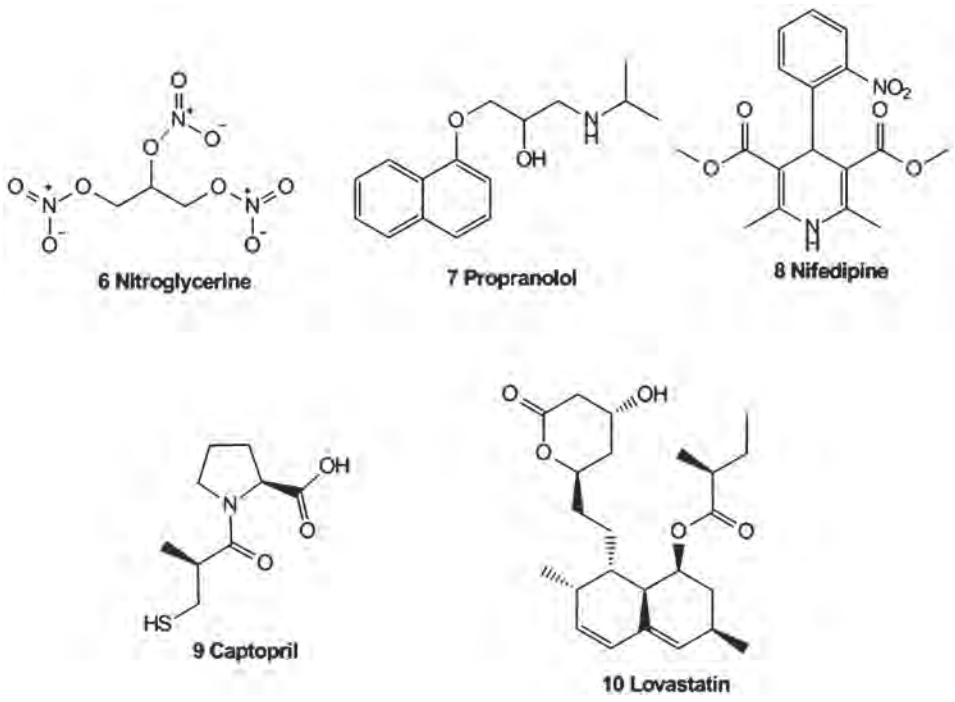


FIGURE 20.4 Potential drugs used in treatment regimens for AP patients.

In attempting to design a novel class of drugs that could help relieve the symptoms of AP, researchers at Pfizer studied the class of phosphodiesterase inhibitors (PDEIs), in particular PDE5 inhibitors (PDE5Is). PDEIs are a class of compounds that prevent the cleavage—and thereby inactivation—of so-called second messenger molecules like cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP), which play a critical role in activating various cellular processes like apoptosis and proliferation. PDE5Is in particular were studied for their potential application as pulmonary artery vasodilators, an effect which results in reduced pulmonary arterial pressure. This in turn alleviates the effort the heart needs to provide a sufficient amount of blood to the lungs.

When clinical test were performed with Sildenafil (**11**; Figure 20.5), its vasodilating properties were confirmed, although its effects in treating AP were found to be limited. But an unforeseen side effect was reported during the clinical trials by male subjects participating in the study, who experienced a stimulating effect of the compound on their erections. Based on this observation, Pfizer decided not to pursue marketing of the compound for AP treatment but to reposition it as a first in class treatment of erectile dysfunction [27]. The FDA approval for this application followed in 1998, and the drug turned out to be a huge commercial success ,with sales of Viagra® reaching more than US $1.9 billion at peak. Detailed studies later revealed that **11**—which is a selective inhibitor of cGMP-specific phosphodiesterase type 5—in particular prolongs the vasodilatory effect of cGMP in the corpus cavernosum of the penis. While the commercial success of **11** did not originate from its intended application, it is interesting that in 2005 the compound did get approval under the tradename Revatio™ for the treatment of pulmonary arterial hypertension [28].

Similar to the case of Ritonavir, the use of PDE5Is to treat erectile dysfunctions was seen by others as an opportunity to be exploited. It should therefore not come as a surprise to see that other companies also developed PDE5Is, as illustrated by the commercialization of Vardenafil (**12**), Tadalafil (**13**), and Avalafil (**14**) [29].

IV. NUCLEOTIDE PRODRUGS: CHEMICAL TROJAN HORSES

Discovered more than fifty years ago, Idoxuridine (**15**) was the first marketed antiviral nucleoside used for the treatment of the herpes simplex virus. Since then, nucleoside-based treatment regimens have started to emerge as real game-changers in therapeutic settings, especially in the field of oncology and virology. There are more than twenty-five nucleoside derivatives approved for therapeutic applications. Some representatives of the class are depicted in Figure 20.6.

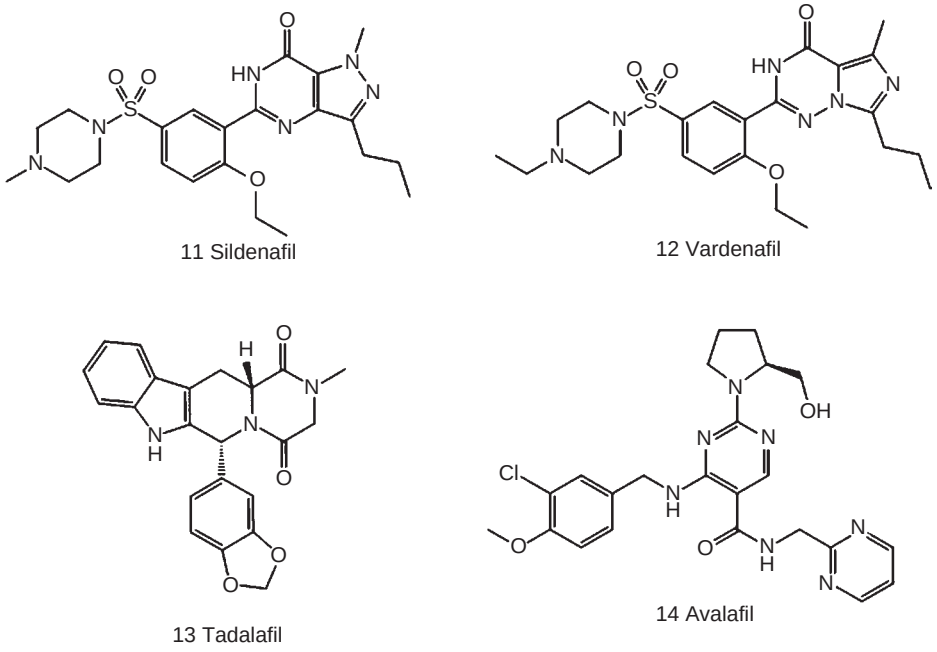


FIGURE 20.5 Structure of Sildenafil and related PDE5 inhibitors.

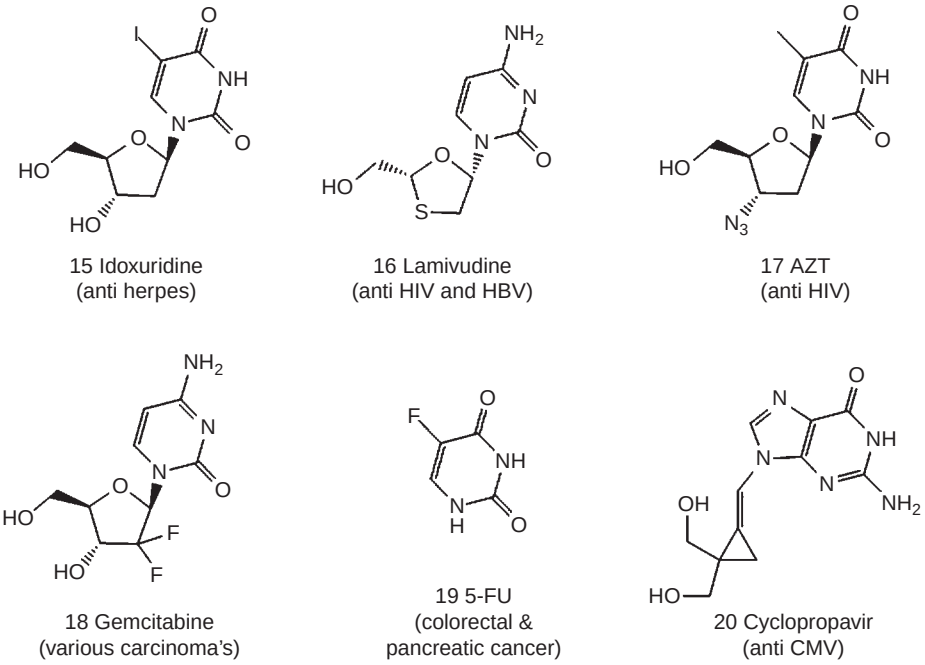


FIGURE 20.6 Nucleoside derivatives acting on different targets.

Interestingly, some synthetic nucleoside derivatives like lamivudine (**16**) display activity against multiple pathogens, which underlines the potential of this compound class. The MOA of antivirally active nucleosides is based on the competitive nature of the nucleoside 5′-triphosphate (NTP) analog of the antiviral, which antagonistically competes with natural NTPs for binding or recognition by a viral RNA or DNA polymerase, like the reverse transcriptase in the case of HIV. Incorporation of such an unnatural NTP into the nascent nucleic acid prevents the further growth of the elongating RNA or DNA chain, a process called "chain termination." MOAs of oncolytic nucleosides include apoptosis induction, DNA methyltransferase inhibition, and topoisomerase II inhibition [30].

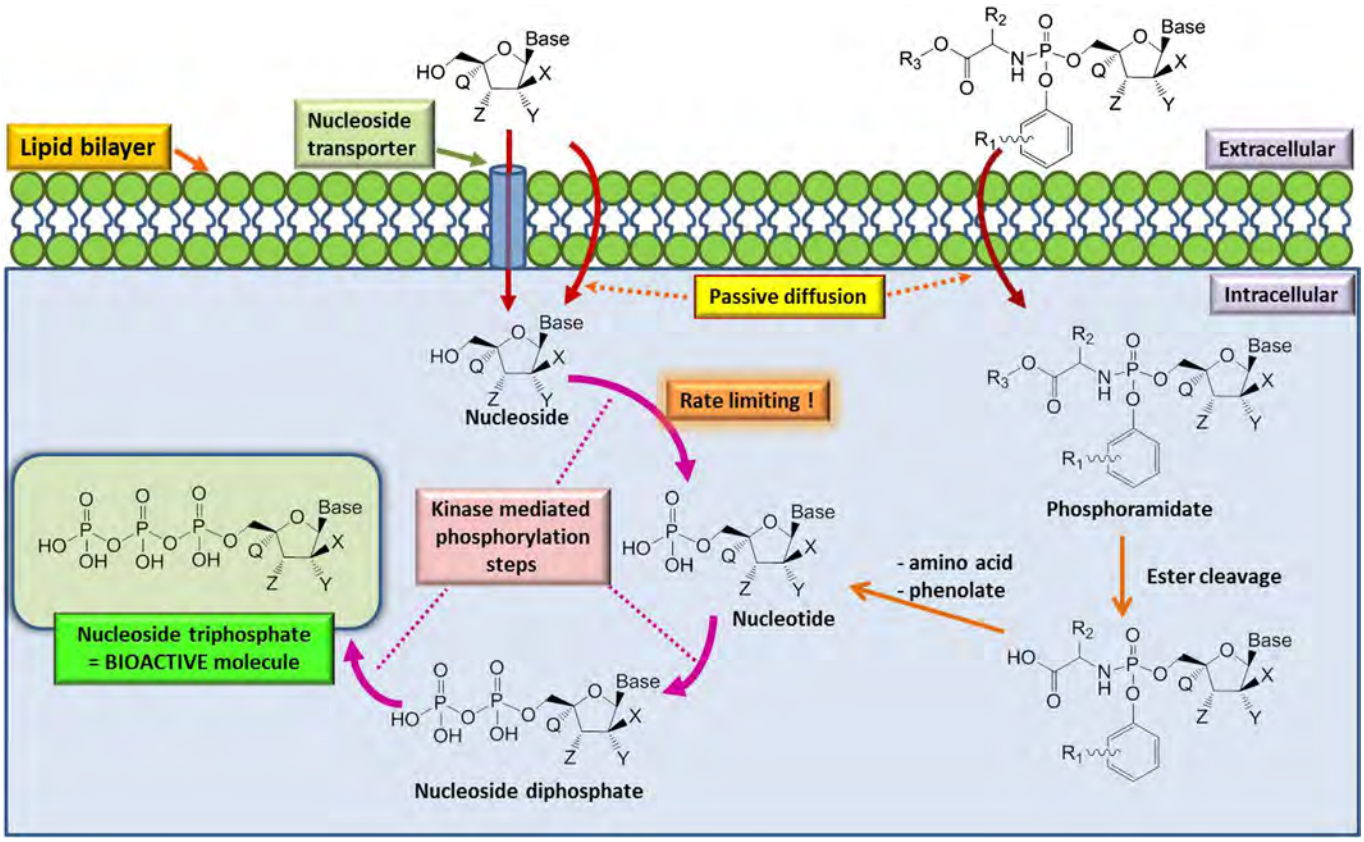


FIGURE 20.7 Intracellular generation of NTPs and phosphoramidate concept.

In principle, nucleoside analogs like **15** act as prodrugs whose efficacy is highly dependent on their conversion via cellular kinases into the corresponding NTPs, which are the active inhibitors. Limited formation of the NTP, resulting from one or more suboptimal phosphorylation steps, would lead to an apparently weaker or even inactive nucleoside inhibitor. In particular, the formation of the nucleoside monophosphate (=the nucleotide) is considered as the rate-limiting step toward NTP generation (Figure 20.7) [31]. Classically, cell-based screening assays are used to search for novel potential hits on a given target. But applying this approach to identify novel nucleoside derivatives could be unsuitable. Indeed, many false negative results might be obtained for two reasons. First, the cells used in the assay might not possess the necessary kinases to form the corresponding NTPs. Alternatively, the nucleosides might not be recognized by a kinase that is a prerequisite for its phosphorylation. As a result, neither nucleotide nor NTP are formed, and so no antiviral effect would be detected.

To avoid this, two possible solutions exist. The most obvious—but probably also most tedious—is to prepare the corresponding NTP derivative of a novel nucleoside and profile it in a biochemical experiment (e.g., a polymerase inhibition assay) to determine its inhibiting properties. This requires the availability of such a biochemical profiling assay, which is not a given for every pathogen. A second solution is to convert the novel nucleoside derivatives into a nucleotide prodrug or pronucleotide. This type of prodrug is designed to be metabolized intracellularly and to liberate the corresponding nucleotide analog of the nucleoside, effectively bypassing the rate-determining first phosphorylation step. Many different types of nucleotide prodrugs are reported in literature, showing the value of this method [32]. However, the class of aryloxy phosphoramidate (PA) derivatives as monophosphate prodrugs is probably the most extensively explored. Introduced in the 1990s by McGuigan et al, this pronucleotide (Protide) method has established itself as one of the preferred techniques to circumvent hampered nucleotide formation [33]. Especially in the field of HCV, the application of phosphoramidate derivatives has been studied extensively, which is not surprising since metabolic breakdown of phosphoramidates—similar to other drugs—is mostly a hepatic event. This is due to the increased lipophilicity of the prodrug compared with the parent nucleoside, which results in liberation of the nucleotide. This happens predominantly in the liver, the target organ for HCV.

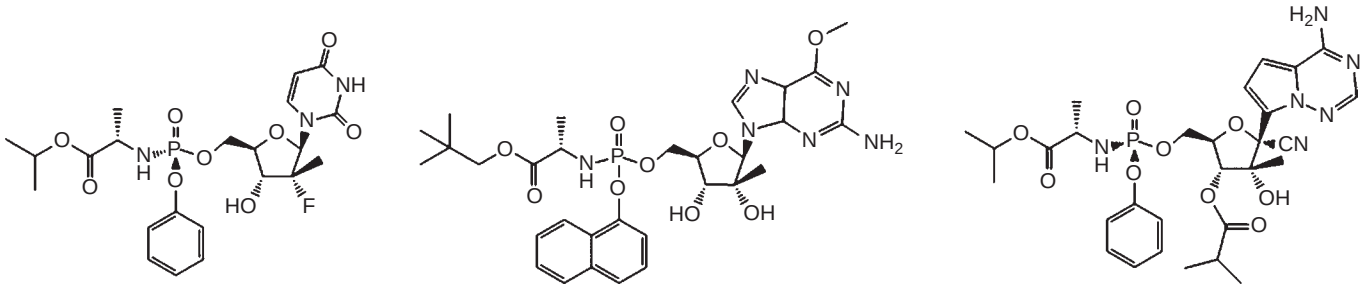


FIGURE 20.8 HCV-active phosphoramidate prodrugs reaching the clinical stage.

Various anti-HCV active nucleoside derivatives have been preclinically evaluated as phosphoramidates, and some have reached the clinical stage (Figure 20.8) [34,35]. Although only Sofosbuvir (**21**) has been approved, the compound has contributed to a complete shift in treatment paradigm for HCV patients, who can now benefit from a much safer, more efficacious, and much shorter treatment option than before [36].

As said, some nucleoside derivatives are known to be active inhibitors of more than one virus or target. However, some derivatives that were only tested as parent nucleoside could be repositioned thanks to application of the phosphoramidate method. A nice example illustrating this is acyclovir (**24**), a well-known nucleoside derivative active against herpes simplex virus type 2 (HSV-2). This virus manifests itself in many HIV-1 infected patients as a result of their reduced immune system response, and it is believed that infection with one of these viruses significantly increases the susceptibility to the other. When HSV-2/HIV-1 co-infected tissues were treated with **24**, it was shown not only to inhibit the HSV-2 virus but also to block HIV-1 virus growth. This was quite surprising, and detailed analysis revealed that the anti-HIV-1 activity of acyclovir was dependent on nucleotide formation by kinases of the co-infecting HSV-2 virus [37].

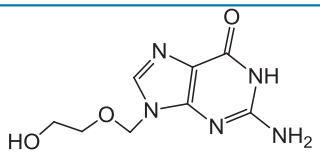
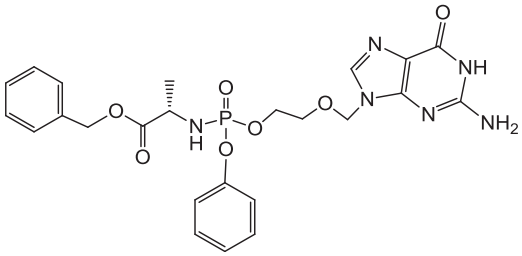
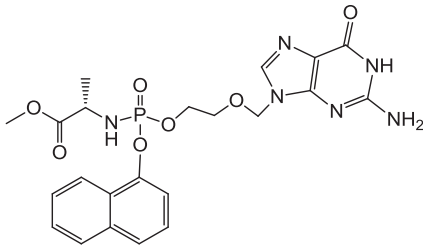
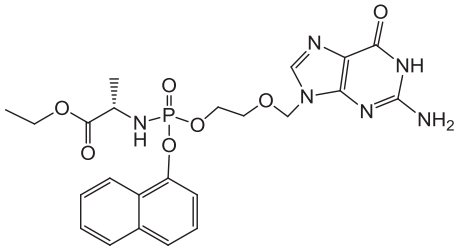
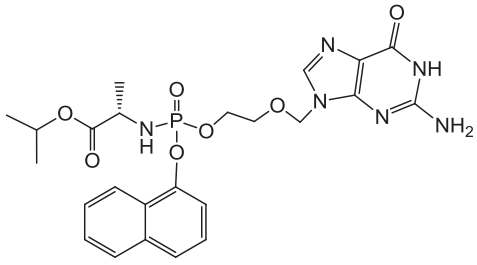
This finding exemplifies again that the activity of a nucleoside is heavily reliant on the availability of a suitable kinase to convert it into its nucleotide. As an extension of the above findings, researchers synthesized several phosphoramidate derivatives of acyclovir and demonstrated their antiviral activity against HIV reverse transcriptase, even in cell lines that were not co-infected with HSV-2 (Table 20.2) [38]. This finding effectively meant the first discovery of an acyclic nucleoside derivative blocking HIV's reverse transcriptase.

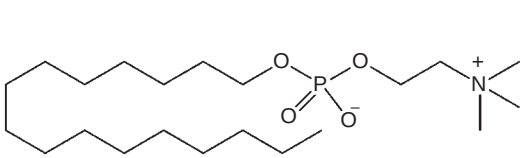
V. MILTEFOSINE

Miltefosine (**25**), discovered about thirty years ago, is a compound belonging to the group of alkylphosphocholines (APCs), which are the more metabolically stable analogs of lysophosphatidyl derivatives like lysophosphatidylcholine (**26**; Figure 20.9). This latter class of compounds structurally resembles lipids present in biological membranes, and some derivatives of this class possess oncolytic properties. The mechanism by which these molecules work is rather specific, since they do not target the cell DNA but insert themselves into the plasma membrane and influence signaling pathways, which in turn leads to apoptotic events. Due to this specific MOA, the proliferation state of the tumor has no influence on the activity of APCs, which make them an important class of compounds [39]. Historically, the general use of **25** as an oral anticancer drug was found to be limited because of its undesirable side effect profile, since at therapeutic relevant doses a high incidence of gastrointestinal toxicity and hematoxicity was observed. In contrast, the use of **25** for the topical treatment of cutaneous lymphomas and related skin metastasis of breast cancer was found to be efficacious and seen as a good complement to radiotherapy and surgical treatments. The compound has been marketed under the name Miltex® for this application [40].

Coincidentally with the discovery of the antineoplastic properties of **25**, another application was discovered for the compound. It was found to possess antiprotozoal activity. In particular, its encouragement of *in vitro* efficacy was observed against trypanosomatid parasites. The exact nature by which **25** exerts its antileismaniasis activity is not fully understood, but it believed to be a combination of several MOAs, of which impacting the phospholipid metabolism and induction of mitochondrial dysfunction are likely [41].

TABLE 20.2 Acyclovir Phosphoramidates are HIV-1 Inhibitors

Compound	Anti-HIV activity EC_{50} μM (MT-4 cells)
 **24 Acyclovir**	>250
	5.7 ± 1.6
	4.7 ± 2.1
	1.7 ± 0.8
	5.3


25 Miltefosine

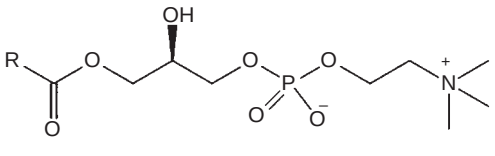
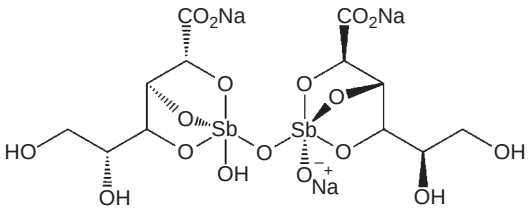
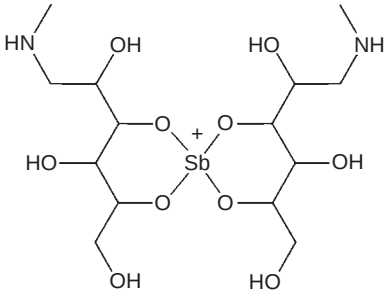

26 Lysophosphatidylcholine
(R = fatty acid chain)

FIGURE 20.9 Structure of alkylphosphocholines **25** and **26**.



27 Pentostam



28 Glucantime

FIGURE 20.10 Structure of antimony-based drugs to treat leishmaniasis.

Globally, three important human diseases caused by trypanosomatids are geographically spread:

a. *African trypanosomiasis*, also known as sleeping sickness, caused by Trypanosoma brucei and transmitted by Tsetse flies;
b. *South American trypanosomiasis*, also known as Chagas Disease caused by Trypanosoma cruzi and transmitted by triatomine bugs; and
c. *Leishmaniasis*, a set of trypanosomal diseases caused by various species of Leishmania and transmitted by sandflies.

In the latter case, one distinguishes between cutaneous leishmaniasis (CL) and the most debilitating form, which is called visceral leishmaniasis (VL). This variant consists of a visceral infection of the reticulo-endothelial system (liver, spleen, bone marrow) that is potentially fatal if not treated properly. Historically, treatment of both forms of the disease is based on the use of antimony based compounds like Pentostam® (**27**) and Glucantime® (**28**), which are dosed intravenously or intramuscular [42] (Figure 20.10). Despite the highly toxic nature of these compounds, they have long been considered as a first-line treatment option for patients suffering from leishmaniasis infections. Amphotericine B has also been used for the treatment of leishmaniasis infection, but as is the case for **27** and **28**, its side effect profile and the need of parenteral administration are limiting factors.

Especially in developing countries and regions were access to general medicine is not guaranteed, there used to be a clear need for an orally available, highly efficacious, and cost-effective cure for treatment of leishmaniasis infections that did not require hospitalization. This was finally achieved by the introduction of **25**. Based on the potent *in vitro* activity followed by promising outcomes in an *in vivo* mice model for VL, compound **25** was tested in a Phase II trial for the treatment of VL in India and demonstrated excellent activity. Further encouraging results from additional Phase II and Phase III trials followed, which ultimately led to the approval of **25** for the treatment of VL in 2002 under the trade name Impavido [43]. Since 2002 (2006 in the US), the compound has received "orphan drug" status, and it also has been added to the list of essential medicines. Despite the fact that the antileishmaniasis activity of **25** was discovered around the same time as its oncolytic properties, its repositioning from an anticancer drug to an best-in-class oral option to combat VL and CL still took many years. In view of this, the exemplary collaboration between industries, Indian governmental instances, and the WHO deserves to be mentioned. So far, **25** remains the only orally available drug for the treatment of leishmaniasis, and it has been approved in many countries, including the US and the whole of Europe.

VI. AZTREONAM

Cystic fibrosis (CF) is a genetic disease that is mainly characterized by the formation of thick and sticky mucus that accumulates in the lungs, severely hampering normal breathing. In addition, gastrointestinal problems are also seen in CF patients, with problems related to the pancreas, liver, and intestines most frequently observed. The disease is caused by a mutation in the gene that encodes for the so-called "cystic fibrosis transmembrane conductance regulator" (CFTR), a membrane protein/ion channel that is responsible for transporting ions like chloride and thiocyanate across epithelial cell membranes. Malfunctioning of this ion channel leads to a deregulated transport of epithelial fluids in the lung, pancreas, and other organs, which leads to CF. At present there is

no definite cure for CF, and patients usually die relatively early [44]. In general, treatment of CF is directed toward alleviation and prevention of symptoms characteristic of the disease, with the biggest focus directed toward preserving lung function. With chronic and/or acute lung infections being regular, many CF patients are on a constant regimen of antibiotics, even prophylactically. The reason for this is that the thickened mucus present in the lungs of CF patients appears to be the perfect habitat for bacteria to reside and grow, making it a difficult environment to reach for antibacterials and host immune cells.

One of the bacteria often causing pulmonary infection in CF patient is *Pseudomonas aeruginosa*, a Gram-negative bacteria that is also responsible for other pathologies like wound infections and urinary tract infections. In general, infections caused by *P. aeruginosa* can be treated with several classes of antibiotics, although its sensitivity to several of them is moderate. Some representative drugs are depicted below (Figure 20.11) [45].

Not all the antibiotics used for maintenance therapy in CF patients can be administered orally, and they require an intravenous or intramuscular injection. For prolonged administration, this severely limits their application, and alternative solutions were developed including the use of subcutaneous catheters and inhalation devices. The latter approach had also been considered for Aztreonam (**32**), a monobactam antibiotic approved in 1986 (Azactam®), which is active against Gram-negative bacteria and is often used for treatment of patients that are allergic to penicillins or aminoglycosides derivatives. When testing (**32**) in inhalation experiments in CF patients, it was found that the treatment resulted in increased inflammation of the airways and worsening of pulmonary symptoms, which was entirely opposite to what was anticipated. The reason for this finding was not related to **32** itself, but was found to be attributable to the formulation used for inhalation. In particular the presence of an arginine based buffer solution used in the formulation was a problem. It was discovered that the prolonged inhalation of arginine resulted in increased levels of nitric oxide being produced, which led to tissue injury. This problem was cleverly solved by changing from an arginine-based formulation buffer to a lysine-based one. This switch ultimately led to the repositioning of **32** for its use in the treatment of *P. aeruginosa* infections via inhalation in CF patients. In 2010, the FDA granted approval for the commercialization of Cayston®, which allows administration of **32** via an ultrasonic nebulizing device [46].

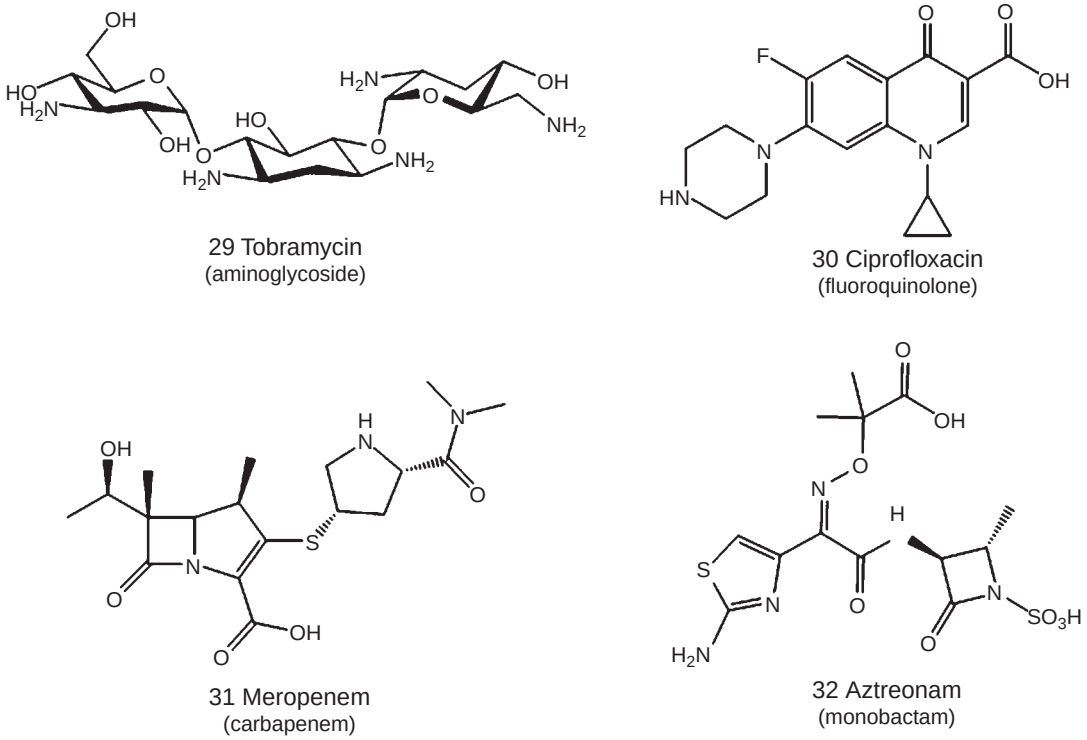


FIGURE 20.11 Structure of some antibiotics used for treatment of *P. aeruginosa*-based infections.

VII. CONCLUSIONS

Recent years have seen a growing interest in and increased effort to identify novel applications for existing drugs. The pharmaceutical industry has faced continuously growing budget pressure and a concomitant need to maintain productiveness over the years, which has surely contributed to this evolution. Importantly, none of the examples discussed above—irrespective of whether they were discovered serendipitously or not—would have been successful if there was not a convincing scientific rationale for their development. This chapter has highlighted a few examples of repositioned or repurposed drugs that have altered the lives of millions of patients. It is therefore realistic to believe that this approach will lead to more successful and surprising cases of "second lives" for drugs in the future, a development that can only be to the benefit of patients in need.

References

[1] Anighoro A, Bajorath J, Rastelli G. Polypharmacology: challenges and opportunities in drug discovery. J Med Chem 2014;57(19):7874–87.
[2] Bartlett JA, DeMasi R, Quinn J, Moxham C, Rousseau F. Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults. AIDS 2001;15(11):1369–77.
[3] Lawitz E, Sulkowski MS, Ghalib R, Rodriguez-Torres M, Younossi ZM, Corregidor A, et al. Simeprevir plus sofosbuvir, with or without ribavirin, to treat chronic infection with hepatitis C virus genotype 1 in non-responders to pegylated interferon and ribavirin and treatment-naive patients: The COSMOS randomised study. Lancet 2014;384(9956):1756–65.
[4] Deeks ED, Perry CM. Efavirenz/emtricitabine/tenofovir disoproxil fumarate single-tablet regimen (Atripla®): a review of its use in the management of HIV infection. Drugs 2010;70(17):2315–38.
[5] Ballantyne CM, Abate N, Yuan Z, King TR, Palmisano J. Dose-comparison study of the combination of ezetimibe and simvastatin (Vytorin) versus atorvastatin in patients with hypercholesterolemia: The Vytorin Versus Atorvastatin (VYVA) study. Am Heart J 2005;149(3):464–73.
[6] Lum PY, Derry JM, Schadt EE. Integrative genomics and drug development. Pharmacogenomics 2009;10(2):203–12.
[7] Petroski H. The evolution of useful things. New York: Knopf; 1992.
[8] Weiss R. How does HIV cause AIDS? Science 1993;260(5112):1273–9.
[9] Günthard HF, Aberg JA, Eron JJ, Hoy JF, Telenti A, Benson CA, et al. Antiretroviral treatment of adult HIV infection. 2014 recommendations of the international antiviral society–USA panel. JAMA 2014;312(4):410–25.
[10] Wensing AMJ, van Maarseveen NM, Nijhuis M. Fifteen years of HIV protease inhibitors: raising the barrier to resistance. Antiviral Res 2010;85(1):59–74.
[11] Palella FJ, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. N Engl J Med 1998;338:853–60.
[12] Molla A, Korneyeva M, Gao Q, Vasavanonda S, Schipper PJ, Mo HM, et al. Ordered accumulation of mutations in HIV protease confers resistance to ritonavir. Nat Med 1996;2(7):760–6.
[13] Merry C, Barry MG, Mulcahy F, Ryan M, Heavey J, Tjia JF, et al. Saquinavir pharmacokinetics alone and in combination with ritonavir in HIV-infected patients. AIDS 1997;11:F29–33.
[14] McKeage K, Perry CM, Keam SJ. Darunavir: a review of its use in the management of HIV infection in adults. Drugs 2009;69(4):477–503.
[15] Zeldin RK, Petruschke RA. Pharmacological and therapeutic properties of ritonavir-boosted protease inhibitor therapy in HIV-infected patients. J Antimicrob Chemot 2004;53:4–9.
[16] Kempf DJ, Marsh KC, Kumar GN, Rodrigues AD, Denissen JF, McDonald E, et al. Pharmacokinetic enhancement of inhibitors of the human immunodeficiency virus protease by coadministration with ritonavir. Antimicrob Agents Chemother 1997;41(3):654–60.
[17] Lin JH, Lu AY. Inhibition and induction of cytochrome P450 and the clinical implications. Clin Pharmacokinet 1998;35(5):361–90.
[18] Gulnik S, Eissenstat M, Afonina E, Ludtke D Erickson J, Dagger R, et al. Preclinical and early clinical evaluation of SPI-452, a new pharmacokinetic enhancer. 16th Conference on Retroviruses and Opportunistic Infections (CROI): Montreal. 2009; Abstract 41. Available from: <http://img.thebody.com/confs/croi2009/posters/41_Guttendorf_slides.pdf>.
[19] See: Patents WO2008022345, WO2007034312, and WO2008004100.
[20] Flentge CA, Randolph JT, Huang PP, Klein LL, Marsh KC, Harlan JE, et al. Synthesis and evaluation of inhibitors of cytochrome P450 3A (CYP3A) for pharmacokinetic enhancement of drugs. Bioorg Med Chem Lett 2009;19(18):5444–8.
[21] Jonckers THM, Rouan M-C, Haché G, Schepens W, Hallenberger S, Baumeister J, et al. Benzoxazole and benzothiazole amides as novel pharmacokinetic enhancers of HIV protease inhibitors. Bioorg Med Chem Lett 2012;22(15):4998–5002.
[22] Xu L, Liu H, Murray BP, Callebaut C, Lee MS, Hong A, et al. Cobicistat (GS-9350): A potent and selective inhibitor of human CYP3A as a novel pharmacoenhancer. ACS Med Chem Lett 2010;1(5):209–13.
[23] Cada DJ, Torres S, Levien TL, Baker DE. Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Disoproxil Fumarate tablets. Hosp Pharm 2013;48(1):48–56.
[24] The AIDS InfoNet. Darunavir/Cobicistat (Prezcobix). Available from: <http://www.aidsinfonet.org/fact_sheets/view/451/>.
[25] A 4-drug combination (Viekira Pak) for hepatitis C. The Medical Letter 2015;57(1461):15–17.
[26] Dotter CT, Judkins MP. Transluminal treatment of arteriosclerotic obstruction. Circulation 1964;30:654–70.
[27] Boolell M, Allen MJ, Ballard SA, Gepi-Attee S, Muirhead GJ, Naylor AM, et al. Sildenafil: an orally active type 5 cyclic GMP-specific phosphodiesterase inhibitor for the treatment of penile erectile dysfunction. Int J Impot Res 1996;8(2):47–52.
[28] Prasad S, Wilkinson J, Gatzoulis MA. Sildenafil in primary pulmonary hypertension. N Engl J Med 2000;343:1342.
[29] Raheem AA, Kell P. Patient preference and satisfaction in erectile dysfunction therapy: A comparison of the three phosphodiesterase-5 inhibitors sildenafil, vardenafil, and tadalafil. Patient Prefer Adherence 2009;3:99–104.

[30] Jordheim LP, Durantel D, Zoulim F, Dumontet C. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. Nat Rev Drug Discov 2013;12(6):447–64.
[31] Van Rompay AR, Johansson M, Karlsson A. Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. Pharmacol Ther 2000;87:189–98.
[32] Pradere U, Garnier-Amblard EC, Coats SJ, Amblard F, Schinazi RF. Synthesis of nucleoside phosphate and phosphonate prodrugs. Chem Rev 2014;114:9154–218.
[33] Mehellou Y, Balzarini J, McGuigan C. Aryloxy phosphoramidate triesters: a technology for delivering monophosphorylated nucleosides and sugars into cells. Chem Med Chem 2009;4:1779–91.
[34] Vernachio JH, Bleiman B, Dawn Bryant K, Chamberlain S, Hunley D, Hutchins J, et al. INX-08189, a phosphoramidate prodrug of 6-O-methyl-2-C-methyl guanosine, is a potent inhibitor of hepatitis C virus replication with excellent pharmacokinetic and pharmacodynamics properties. Antimicrob Agents Chemother 2011;55:1843–51.
[35] Cho A, Zhang L, Xu J, Lee R, Butler T, Metobo S, et al. Discovery of the first C-nucleoside HCV polymerase inhibitor (GS-6620) with demonstrated antiviral response in HCV infected patients. J Med Chem 2014;57(5):1812–25.
[36] Cholongitasa E, Papatheodoridisb GV. Sofosbuvir: a novel oral agent for chronic hepatitis C. Ann Gastroenterol 2014;27(4):331–7.
[37] Lisco A, Vanpouille C, Tchesnokov EP, Grivel JC, Biancotto A, Brichacek B, et al. Acyclovir is activated into an HIV-1 reverse transcriptase inhibitor in herpesvirus-infected human tissues. Cell Host Microbe 2008;4:260–70.
[38] Vanpouille C, Lisco A, Derudas M, Saba E, Grivel JC, Brichacek B, et al. A new class of dual-targeted antivirals: monophosphorylated acyclovir prodrug derivatives suppress both human immunodeficiency virus type 1 and herpes simplex virus type 2. J Infect Dis 2010;201(4):635–43.
[39] van Blitterswijk WJ, Verheij M. Anticancer alkylphospholipids: mechanisms of action, cellular sensitivity and resistance, and clinical prospects. Curr Pharm Des 2008;14(21):2061–74.
[40] Smorenburg CH, Seynaeve C, Bontenbal M, Planting AS, Sindermann H, Verweij J. Phase II study of miltefosine 6% solution as topical treatment of skin metastases in breast cancer patients. Anticancer Drugs 2000;11(10):825–8.
[41] Dorlo TP, Balasegaram M, Beijnen JH, de Vries PJ. Miltefosine: a review of its pharmacology and therapeutic efficacy in the treatment of leishmaniasis. J Antimicrob Chemother 2012;67(11):2576–97.
[42] Frézard F, Demicheli C, Ribeiro RR. Pentavalent antimonials: new perspectives for old drugs. Molecules 2009;14:2317–36.
[43] Sindermann H, Croft SL, Engel KR, Bommer W, Eibl HJ, Unger C, et al. Miltefosine (Impavido): the first oral treatment against leishmaniasis. Med Microbiol Immunol 2004;193(4):173–80.
[44] O'Sullivan BP, Freedman SD. Cystic fibrosis. Lancet 2009;373(9678):1891–904.
[45] Döring G, Conway SP, Heijerman HGM, Hodson ME, Høiby N, Smyth A, et al. Antibiotic therapy against Pseudomonas aeruginosa in cystic fibrosis: a European consensus. Eur Respir J 2000;16:749–67.
[46] Retsch-Bogart GZ, Quittner AL, Gibson RL, Oermann CM, McCoy KS, Montgomery AB, et al. Efficacy and safety of inhaled Aztreonam Lysine for airway pseudomonas in cystic fibrosis. Chest 2009;135(5):1223–32.

SECTION FIVE

Spatial Organization, Receptor Mapping and Molecular Modeling

CHAPTER

21

Pharmacophore Identification and Pseudo-Receptor Modeling

Gerhard Wolber[1] *and Wolfgang Sippl*[2]

[1]Computer-Aided Molecular Design, Institute of Pharmacy, Freie Universität Berlin, Königin-Luisestr., Berlin, Germany;
[2]Medicinal Chemistry, Institute of Pharmacy, Martin-Luther-Universität Halle-Wittenberg, Halle (Saale), Germany

OUTLINE

In theory, theory and practice are the same. In practice, they are not. ***Lawrence Peter Berra***

I. INTRODUCTION

In the large majority of cases, the basis for a pharmacodynamic effect is the interaction of a certain substance with a biomacromolecule of physiological importance. Proteins—like enzymes, receptors, and ion channels—but also nucleic acids serve as physiological binding partners for small organic ligands. In all cases, a highly specific 3D binding epitope must exist, which serves as complementary binding site for a drug molecule. Compounds exerting similar activities on the same enzyme or receptor therefore possess—in most cases—closely related binding properties. That is, these molecules present structural elements of identical chemical features in sterically consistent locations to the macromolecule. The highest common denominator of a group of ligands exhibiting a similar biological effect recognized by the same binding site is named a "pharmacophore." [1] In other words, a pharmacophore is an abstraction of the crucial molecular features responsible for the binding of a set of ligands to a macromolecular target.

As a practical matter, computer-aided molecular design is frequently split into disciplines that focus on either structure-based or ligand-based methods. When the 3D structure of a target protein and the binding site is available,

it is possible to invoke structure-based approaches. New candidate ligands may be docked into a particular binding site in order to study whether they can interact with the protein in an optimal way. If, however, knowledge about the structure of the macromolecular target is limited but a sufficient number of active analogues have already been discovered, then pharmacophore-based methods are applied to design novel active molecules. It may seem straightforward to develop new ligands for known proteins by applying structure-based approaches, but significant problems are involved. Induced fit mechanism, multiple binding modes, solvation, and entropic effects are some of the problems that must be overcome to end up with reliable models. Beside these problems, many target proteins of high pharmaceutical interest are membrane-bound receptors (e.g., G-protein coupled receptors [GPCRs]) and attempts to crystallize them have been only partially successful. Although twenty-four GPCR crystal structures from four different classes have been published at this point, structure-based 3D pharmacophore development with these structures remains a challenge. A larger amount of 3D structures for activated and inactivated structures would be necessary to be able to develop structure-based predictive models for agonists and inverse agonists, respectively.

In the absence of the 3D structure of a protein of interest or a biologically relevant conformation, ligand design may be performed by the use of a pharmacophore-based method. This is based on the assumption that several ligands bind to the same binding pocket of the protein. Thus, a flexible superposition can be identified, which represents the interaction pattern of the binding pocket from the view of the ligands.

A. Historical Background

The idea that bioactive substances interact with receptors began in 1878 with Langley, who introduced the term "receptive substance." [2] However, the term "receptor" was introduced several years later by Paul Ehrlich [3]. He also introduced the term "pharmacophore" to describe those parts of a molecule that are responsible for its activity. Together with the lock-and-key concept of Emil Fischer, it became clear that not all parts of a molecule—the "key"—are equally important for exerting its biological effect on the "lock." [4] Thus, sometimes small variations of distinct parts of a molecule can dramatically influence the activity, whereas variations of other parts only cause minor changes in the biological activity. The concepts of Langley, Ehrlich, and Fischer constitute the cornerstones of modern drug discovery and development up to this day. Half a decade later, their concepts were confirmed in an impressive manner by the first solved crystal structures of protein−ligand complexes [5].

Even before the advent of computers and modeling software, simple pharmacophores were described in the literature and considered tools for the discovery of novel molecules. Based on initial structure−activity relationship considerations, simple 2D models were introduced in the 1940s. With the advent of computers and modeling programs, the idea of displaying and manipulating 3D structures became possible [6]. Kier and Marshall pioneered the development of the pharmacophore concept and its application in structure−activity relationships [7,8]. In the 1970s, Peter Gund implemented the first *in silico* screening method with a program to screen a substance library for pharmacophoric patterns [9]. The active analogue approach developed by Garland Marshall's group was one of the first automated tools for pharmacophore generation. Marshall's approach was the basis for many following pharmacophore modeling programs in that area. Since these early days, a variety of automated pharmacophore discovery programs have been developed in academia and by software developing companies (for review, see [10,11]).

B. Definitions

The term pharmacophore is not always used by different groups of scientists in accordance with the official definition elaborated by the IUPAC working party, which states [1]: "A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interaction with a specific biological target structure and to trigger (or block) its biological response." Many scientists use the term "pharmacophore" or "pharmacophoric group" to define distinct functional groups or substance classes possessing biological activity (e.g., sulfonamides or dihydropyridines). In this context, the term pharmacophore is mixed with another concept of structure and activity, namely "privileged structures." The retrospective analysis of the chemical structures and scaffolds of drug molecules led to the detection of some structural motifs that are often associated with biological activity. Such motifs were called "privileged structures" by Evans et al to represent substructures that confer activity toward two or more different targets [12]. The idea behind this is that the privileged structure provides the scaffold and the substitution provides the reason for specificity. However, in terms of the IUPAC definition, the pharmacophore represents the common molecular interaction features of a set of molecules toward their receptor [13].

A pharmacophoric element (also called feature) is generally defined as an atom or a group of atoms (e.g., a hydrogen bond donor atom or an aromatic ring system) common to active compounds with respect to a target protein and essential for the activity. Thus, a pharmacophore model can also be regarded as the representation of a collection of pharmacophore features.

The above-described definition of a pharmacophore is based on a 3D point-of-view of molecules. It reflects the way medicinal chemists characterize the binding ability of molecules for a given target protein. However, depending on the different research areas, scientists have different views. Computational chemists often use the term pharmacophore in a more abstract way. Influenced by the structural representation of molecules, a set of topological connections is used to define the properties and dimension of a molecule in 2D. Here, the spatial and topological distribution of pharmacophoric features is converted to a lower dimensional representation (e.g., vectors). Such vectors, which represent pharmacophore descriptors, are called "fingerprints," "keys," "bitstrings," or "correlation vectors," depending on the type of information stored. The pharmacophore descriptors or fingerprints can be regarded as a transformed molecular representation instead of an explicit 3D structure. These fingerprints are often used to screen large compound libraries rapidly. In this chapter, we will only focus on 3D pharmacophore concepts.

Starting from a preliminary pharmacophore model, a hypothetical receptor consisting of individual amino acid residues can be constructed surrounding a set of superimposed ligands. Guided by permanent correlation of biological data and model-derived calculated free energies of binding, a complex system is generated, mimicking the interaction pattern of a real binding site reasonably well. The resulting hypothetical receptor model is named "mini-receptor" or "pseudo-receptor," and can be used to derive 3D quantitative structure–activity relationships (3D-QSAR). The concept was originally developed in the 1980s by several groups [14–17].

C. Importance of the Pharmacophore Concept

A pharmacophore captures the concept of bio-isosterism by not only comparing topological similarities but structural groups at similar locations with the same chemical functionality. It is important to concentrate on the pharmacophoric features, since topological molecule characteristics are often misleading in the superpositioning of two molecules with respect to their binding mode. Figure 21.1 shows the well-known example of dihydrofolate reductase ligands [18]. For the two ligands shown, a topological overlay would result in an incorrect prediction

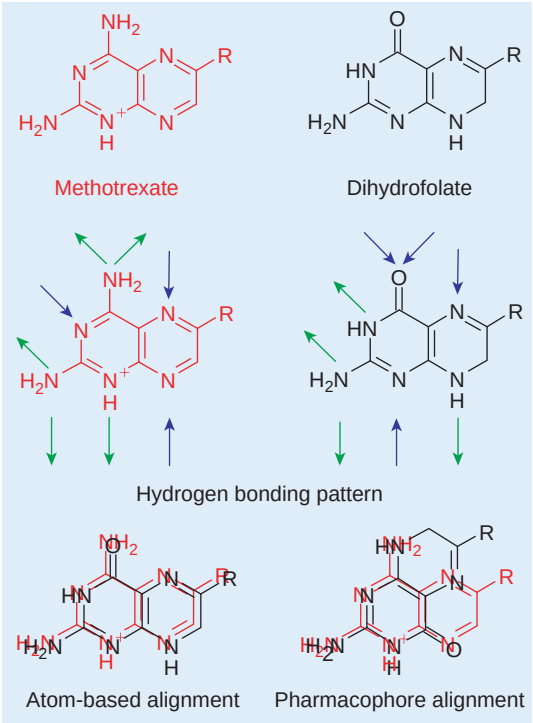


FIGURE 21.1 Comparison of atom-based alignment and experimentally derived position of methotrexate and dihydrofolate in dihydrofolatreductase. Blue arrows indicate hydrogen bond acceptors, and green arrows show hydrogen bond donors.

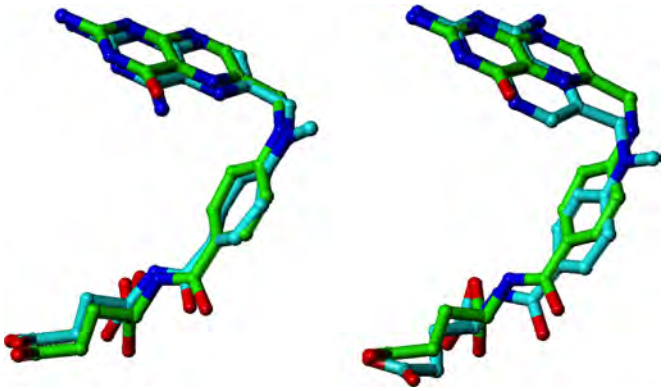


FIGURE 21.2 Comparison of the pharmacophore-based alignment of dihydrofolate reductase ligands (left) and the experimentally derived protein-based alignment (right).

of the binding mode. If the pharmacophoric features (the hydrogen bonding pattern in this example) are taken into account for the superimposition, the correct overlay mode can be deduced. The pharmacophore-based superposition is similar to the binding mode observed in the crystal structures of methotrexate and dihydrofolate with dihydrofolatreductase (Figure 21.2; PDB IDs 1RX2, and 1RB3).

The increasing number of accessible compounds that can currently be used as starting points for biological target screening makes it necessary to have fast and reliable *in silico* screening tools. Structure-based methods are often too slow to screen compound databases with millions of molecules virtually. Beside the speed, other problems in structure-based design and docking programs need to be addressed. For instance, most of the current docking programs do not take into account protein flexibility. Only recently were programs developed (e.g., AutoDock4 [19], GOLD Suite 5.2 [20], Glide [21] or FlexE [22]) that consider protein side-chain flexibility for docking. Other problems, which often occur in ligand docking, are the correct placement of water molecules within the binding site (which represents putative ligand binding partners), the treatment of solvation effects (on the ligand and protein site), and consideration of the internal strain of a docked ligand. Structure-based approaches are able to provide important information about the interaction between a ligand and a macromolecule, but the accurate prediction of the binding affinity is still an unsolved problem. A detailed discussion about the limitations of docking and scoring programs can be found in several reviews [23–28].

Another reason pharmacophore-based approaches are often used in drug design is the missing 3D structure of many interesting macromolecules. Many current drug targets are membrane-bound and, despite recent progress in crystallizing GPCRs [29–31], only a small fraction of membrane proteins have been successfully crystallized. An additional challenge for some membrane-bound receptors is that ligands bind in the extracellular region, causing conformational changes in the protein that lead to a signaling response, such as the release of G-proteins in the case of GPCRs. It is still very difficult to capture the "right" conformation in a crystal, especially when it comes to modeling ligand function (e.g., agonists or antagonists). In the absence of an experimentally determined 3D protein structure, the use of indirect ligand-based approaches—including pharmacophores—is the only way to design novel bioactive molecules rationally [32].

D. Application of Pharmacophores

Pharmacophore modeling in computer-aided drug design is generally applied in three domains. The first is the definition of relevant pharmacophoric features in a drug molecule necessary to achieve a certain biological effect and to establish clear structure–activity relationships. A well-developed pharmacophore model, preferentially including information about the dimension of the receptor binding cavity, may be employed to design novel and more active molecules that fit the model. Often, such pharmacophore models are the starting point for 3D-QSAR analysis (e.g., CoMFA [33]), by which quantitative predictions may be made. The second is scaffold hopping, the practice if detecting molecules with different scaffolds (novel chemotypes) by virtually screening large compound libraries [34]. The third domain is the use of parallel pharmacophore-based screening in order to predict pharmacological profiles for lead structures *in silico*. The use of 3D pharmacophore models can hopefully predict unwanted side effect in very early stages of the drug-discovery process and therefore reduce the risk of late failure of drug candidates [35].

II. METHODOLOGY

A. Pharmacophore Modeling

To end up with a predictive pharmacophore model, it is necessary to start with reliable structural and biological data. First of all, it is important to have correct 3D structures of all compounds under study. Thus, atomic valences, bond orders, protonation state, and stereochemistry have to be checked carefully. Also, the consideration of different possible tautomers is necessary when the bioactive form is not known exactly. Another prerequisite is the existence of a similar binding mode of all ligands under study. Experimental data, from competition experiments or protein–ligand crystal structures, can clearly point out that the ligands interact with the same binding epitope in a similar way instead of on distinct binding sites.

The four steps in the development of a pharmacophore model are: (a) selection of a set of active ligands known to bind to the same target (same binding site); (b) conformational analysis for all ligands; (c) assignment of pharmacophoric features; and (d) molecular superimposition of the ligand conformations to develop a common 3D-pharmacophore. The majority of automated pharmacophore generation programs use qualitative pharmacophore models that do not consider the activity of the ligands. The ultimate goal of all these programs is to search for a unique conformation of all congeners, where most if not all assigned pharmacophoric features of the ligands are presented in a superimposed manner. Most of the programs are based on minimizing the root-mean-square (RMS) superposition error between conformations of the ligands under study while trying to increase the fit of the pharmacophoric features. To compare the different conformations for a data set of given active molecules, a superpositioning procedure is needed. The assignment of the pharmacophoric features and the generation of the ligand alignment is carried out in an automated way by most of the current pharmacophore modeling programs (e.g., Catalyst [36], DISCO [37], Galahad [38], LigandScout [39], Phase [40], MOE [41]). The scope of this chapter is not to describe all available software packages in detail, but to illustrate the different steps of the pharmacophore development process. For a recently published overview of current pharmacophore modeling programs, the reader is referred to the literature [10,42–44].

1. *Conformational Analysis of Ligand Molecules and Bioactive Conformation*

Since molecules are flexible and not static, a conformational analysis has to be carried out first to generate an ensemble of low-energy conformations. This is probably one of the most critical steps in the pharmacophore discovery process, since the goal is not only to consider the global minima of a molecule but also to include the bioactive conformation as part of an ensemble of low-energy conformations.

In order to bind to a protein with high affinity, a ligand must match the binding pocket. The steric match will thereby depend primarily on the ligand conformation. Within a binding pocket, the ligand will not necessarily be present in its lowest energy conformation, as the gain in interaction energy with the receptor can compensate for a conformation with higher energy [45]. Still, it can be expected that for a high-affinity ligand, the bioactive conformation is at least energetically favorable, as otherwise the conformational energy cost would reduce binding affinity. The relation between a high energetic binding conformation and the loss of free energy of binding ΔG is given by Equation 21.1:

$$\Delta G = -2.303RT \log Ki \tag{21.1}$$

Under physiological conditions (T = 310 K), the free energy (in kcal mol^{-1}) and the binding affinity are related by

$$\Delta G = -1.42 \log Ki \tag{21.2}$$

Thus, if a compound binds in a conformation that deviates 1.42 kcal mol^{-1} from the global minimum structure, its affinity will be decreased by one order of magnitude. High-affinity compounds can thus be expected to bind in an energetically favourable conformation. To analyze the conformational space of molecules experimental and theoretical approaches are applied. Experimental techniques like NMR only provide information on one or a few conformations of a molecule. A complete overview about the conformational space of molecules can be gained only by theoretical techniques [46]. Correspondingly a variety of theoretical methods for conformational analysis has been developed. The most general conformational analysis methods are those that are able to identify all minima on the potential energy surface. However, as the number of minima dramatically increases with

the number of rotatable bonds, an exhaustive detection of all minima becomes a difficult and time-consuming task. Commonly used methods for this purpose are listed below (described in depth in [47]):

- Systematic search: Each bond is rotated incrementally and the resulting structures are minimized. Systematic search algorithms have the advantage of sampling the conformational space very well. In cases with a high number of rotatable bonds, this method may be computationally impracticable [48].
- Random search: In a random search, one can move from one region of the energy surface to a completely unconnected region in a single step. A commonly applied method is the Metropolis Monte Carlo scheme that starts with a minimized conformation A of a molecule. A random move on the energy-landscape is carried out (e.g., torsion angles are rotated by a random amount), and the structure is minimized. The potential energy of the output structure B is evaluated. If $E_{pot}(B) < E_{pot}(A)$, the new conformation is accepted. If $E_{pot}(B) > E_{pot}(A)$, the move may still be accepted depending on the transition probability, which in turn depends on the temperature. Monte Carlo methods efficiently sample the conformational space, but there is no guarantee—as with all random search tools—that the entire energetic landscape will be sampled. Another sampling technique applied to the problem of improved conformational searching is known as Poling [49]. Poling is implemented within Catalyst [36] allowing the generation of large multi-conformer virtual screening databases in a reasonable amount of time.
- Simulated annealing or molecular dynamics (MD) simulations: The aim of MD simulations is to reproduce the time-dependent motional behavior of a molecule. MD is based on molecular mechanics. It is assumed that the atoms in the molecule interact with each other according to the rules of an employed force field. MD simulations generate an ensemble of coordinates that does not only contain minimized structures, but rather provides a (limited) sampling of conformational space. In a simulated annealing MD protocol, the system temperature is periodically increased, resulting in a significant rise of kinetic energy, which makes it easier to overcome barriers of potential energy. Subsequently, the system is cooled down, thereby trapping the molecule in an energetically favorable conformation. MD simulation techniques for sampling the conformational space are quite time-consuming and are therefore used only for smaller ligand data sets. Again, there is no guarantee of sampling the entire potential energy surface [50].

There is an ongoing discussion in the literature about which ligand conformations (i.e., within which energy range) have to be considered in a pharmacophore generation process. Several recent studies on protein−ligand X-ray structures have shown that many conformational search tools yield ensembles, including the experimentally observed bioactive conformation [51]. The energy difference between the co-crystallized conformation of a ligand and its global minimum calculated with molecular mechanic programs is dependent on the force field employed. Therefore, a general energy range to be considered cannot be defined [52].

Which conformational analysis performs best? A clear-cut answer cannot be given, as it depends on the individual data set to be studied and the problems to be addressed. If only a limited number of ligands is considered, more computationally intensive methods such as the systematic search can be applied. If a compound library with hundreds of thousands of entries has to be converted into a multi-conformer database, faster simplified approaches have to be used (e.g., in Catalyst [36] or Omega [53]) [54].

2. *Pharmacophore-Ligand Superposition Techniques*

Three-dimensional pharmacophore−ligand superpositioning has to deal with the challenge of conformational flexibility. One possibility for addressing the problem is to perform the identification of common chemical features and the conformational search simultaneously (flexible alignment). Other approaches pre-generate conformations (rigid-body alignment), which makes the overlay algorithm faster, but the inclusion of relevant conformations must be ensured.

A well-known pharmacophore elucidation program using a flexible approach is GASP [55,56], which was initially developed by Jones and co-workers in the mid-1990s. The software is based on a genetic algorithm that simulates evolution by randomly combining and mutating chromosomes of an initial population. Each chromosome represents a potential flexible pharmacophore by encoding all torsion angles and by listing all feature mappings to a manually selected rigid reference compound. In each run, highest scoring chromosomes are selected according to a simple fitness function, and those are then mutated by applying random torsional rotations to cover conformational space during the alignment process. Today, most program use rigid-body alignment techniques, because conformers only have to be calculated once and can be stored in a database, which saves computational time during the alignment.

The wide field of rigid-body superpositioning algorithms (also referred to as "alignment techniques") can roughly be divided into 3D geometry-based and linearized fingerprint-based (descriptor-based approaches). Fingerprint-based approaches create a linearized bit sequence representing the chemical feature properties of the alignment partners. They allow for fast computational similarity assessment. A very advanced example for fingerprint approaches is the Chemically Advanced Template Search (CATS) developed in Gisbert Scheider's group [57]. However, no real 3D overlay is produced using such an implementation. With geometry-based approaches, chemical features represented as 3D points with optional geometry constraints (such as vectors or planes) are assigned to 3D conformations of the ligands. In a subsequent step, these algorithms attempt to minimize distances between those points while considering the assigned constraints like the parallel orientation of the planes or the overlap of the projection point of a vector. Computationally expensive solutions to this problem have been proposed relatively early and range from 3D maximum clique detection algorithms [58] as used in DISCO [59,60] to the sequential build-up of increasingly larger common feature configurations as employed in Catalyst [36], Phase [40], or MOE [41]. This approach becomes problematic if pharmacophore point tolerances shall be subsampled, and it results in geometric fuzziness of the resulting alignment. LigandScout [61] uses a novel and computationally more efficient pattern-matching technique [62] to identify an initial alignment. With this technique, it is possible to perform a geometrically more accurate alignment that also subsamples chemical feature point tolerances when performing high-throughput virtual screening.

3. *Assignment of Pharmacophoric Elements*

The assignment of pharmacophoric features shall be described using as an example the histamine H_3 receptor antagonist shown in Figure 21.3 [63]. Table 21.1 lists the pharmacophoric features assigned in the ligand structure by comparison with other known active antagonists. Thus, the protonated nitrogen atom of the piperidyl moiety can be translated into a center of a sphere with coordinates corresponding to the location of the nitrogen atom

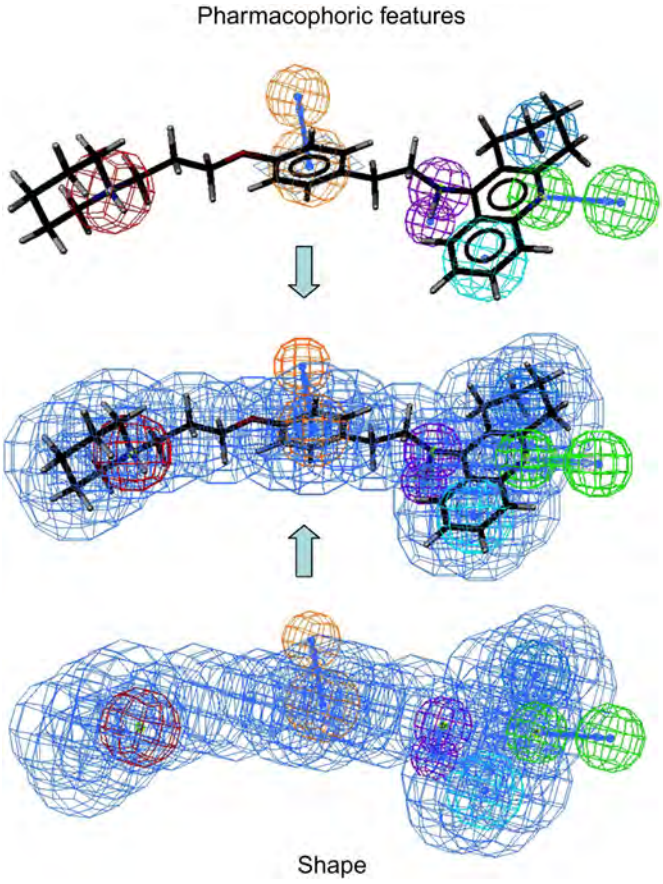


FIGURE 21.3 (top) Pharmacophoric features observed in the ligand (by comparison with other known active ligands). (bottom) The molecule's shape can serve as an additional constraint in pharmacophore searches. The resulting pharmacophore is based on features and shape (middle).

TABLE 21.1 Pharmacophoric features observed in the ligand shown in Figure 21.3

Feature	Color	Representation
positive charge	red	sphere
H-bond donor	magenta	sphere-vector-sphere
H-bond acceptor	green	sphere-vector-sphere
hydrophob. aliphatic	blue	sphere
aromatic ring	orange	plane, center of plane, vector
hydrophobic	light blue	sphere

and a radius defining a volume around this atom. If a molecule is compared to this pharmacophore model and its protonated nitrogen atom lies within the sphere, this pharmacophoric feature will be said to be matched. The bigger the sphere, the easier it will be for a ligand conformation to match the pharmacophoric features. Similarly, an aromatic or a hydrophobic aliphatic moiety can be defined by a center of sphere and radius. Hydrogen bond acceptors and donors are represented by vectors in order to account for the directionality of H-bonds, while aromatic rings can be either defined by spheres or the combination of center, plane, and vector. When defined this way, the orientation of the aromatic plane in respect to the rest of the molecule is considered, too.

Again, the shape of the molecule can be incorporated into a pharmacophore by translating the van der Waals volume into an additional feature. If the ligand is known to fill the binding pocket well, the available volume can be taken into account. The abstract definition of a molecule in form of a pharmacophore as defined in Figure 21.3 facilitates comparison with other molecules. In the given example, most features of the antagonist were considered for the generation of the pharmacophore model resulting in an fingerprint of the molecule. Depending on the number of features included in the model and the tolerances defined, it will be more or less difficult for other molecules to match the pharmacophore model. Matching a pharmacophore additionally depends strongly on the conformation adopted by the molecule that is compared to the pharmacophore model. Even a different conformation of the ligand might not match the pharmacophore model defined above. Thus, in order to increase the likelihood of a specific molecule fitting a pharmacophore model, each molecule of interest is associated with a conformational ensemble. When searching for similarities with the pharmacophore model, all conformations of a molecule are tested on the pharmacophore before the best fit is evaluated. The difficulty in defining a useful pharmacophore model lies in the restriction to essential pharmacophoric features observed in the active ligands.

4. Model Quality and Pharmacophore-Based Virtual Database Screening

If a pharmacophore is used for database screening in order to retrieve new compounds based on the similarity of pharmacophoric features, a model is useful when it is able to identify known actives among a number of inactive molecules. In order to screen commercial compound databases with a pharmacophore model, a so-called multi-conformer database must first be generated. This means a set of conformations must be generated for all compounds deposited in the compound databases. Since such databases can include millions of compounds, fast algorithms are paramount. In addition, the conformation database should not lead to an explosion in storage requirements for the millions of conformers. Finally, the database program should be able to handle the pharmacophore search within a reasonable amount of time. The most widely available commercial programs for building large multi-conformation databases are Catalyst [36], UNITY [64], Omega [53], and MOE [41]. Whereas Catalyst, UNITY, and MOE are also used to carry out pharmacophore generation and pharmacophore searches, Omega can only be used to generate multi-conformer databases. A comparison of the performance of the different programs can be found in the literature [51].

Recently, freeware alternatives to the above-mentioned programs have become available, but they are yet to be evaluated in terms of their performance when used with pharmacophore development and virtual screening [65].

The pre-calculation of conformations bears the important advantage that the screening process is considerably faster and avoids a dramatic reduction of the conformational search space by falling into a local minimum [66,44]. With current computer hardware, the additionally required storage space no longer represents a limitation, and screening databases with pre-generated conformations are clearly preferred. These databases can be generated once and reused for subsequent virtual screening runs, which results in a considerable speed-up of the overall screening procedure.

Typically, a 3D pharmacophore is first created as a hypothesis and analyzed retrospectively to assess its predictive power before being used for a prospective virtual screening that should predict whether new molecules with a certain pharmacophore pattern actually bind to the protein under investigation. To assess retrospective screening performance, several virtual screening metrics have been established that can be used to evaluate model quality [67–69]. These metrics describe the ability of a pharmacophore model to match bioactive molecules and thus include them in the virtual hit list versus the ability to exclude biologically inactive molecules. The most frequently used metrics are described below. For a more extensive overview, the reader is referred to reviews on this topic [70,71].

Sensitivity (Se) describes the ratio of the retrieved true positive compounds (TP) in relation to all biologically active compounds in the database that is the sum of TP and the number of false negatives (FN). Sensitivity values range from 0 to 1, where a value of 0 means that the search did not yield any actives in the database and a value of 1 indicates that all active compounds could be retrieved.

$$Se = \frac{TP}{TP + FN}$$

Specificity (Sp) describes the portion of rejected biologically inactive compounds (true negatives, TN) divided by the sum of TN and the number of false positives (FP). Specificity ranges from 0 to 1, where a value of 0 means that none of the inactive com- pounds could be excluded and a value of 1 indicates that no inactive compounds fit the pharmacophore hypothesis.

$$Sp = \frac{TN}{TN + FP}$$

Yield of Actives (Ya) sets the amount of true positives (TP) in relation to the size of the hit list (n).

$$Ya = \frac{TP}{N}$$

Enrichment Factor (EF) measures the Ya proportionally to the fraction of actives in the database. In the formula below, A is the number of actives in the database and N is the total number of molecules in the database.

$$EF = \frac{Ya}{(A/N)}$$

Convenient and meaningful tools for the assessment of screening performance are Receiver Operating Characteristic (ROC) curves [72]. A ROC curve displays the sensitivity on the y-axis versus (1-Specificity) in the x-axis and thus gives a good impression of the quality of a model. An ideal curve would rise vertically along the y-axis until it reaches the maximum true positive rate (1), and then continue horizontally to the right, which means that the hit list contains all active compounds in the database and that none of the hits is a false positive. The diagonal line between the lower left and the upper right corner of the graph would represent the ROC curve of a random database search. Figure 21.4 shows an example for a ROC curve.

III. ADVANCED APPROACHES

A. Structure-Based Pharmacophores

If the 3D structure of a protein–ligand complex is known from either X-ray crystallography or protein NMR, the most obvious way of deriving a picture of the relevant ligand interactions is to analyze the molecule's complementarity within the corresponding protein binding site. A commonly used structure-based design approach is the previously mentioned molecular docking of ligands into a target binding pocket, assuming that the binding site is more or less rigid whereas the ligand is flexible. Molecular docking is still the most popular method for structure-based drug design. However, pharmacophore-based approaches have shown clear advantages regarding the computational demand and accuracy for virtual screening [61]. Especially with regard to the number of false positives—which are often observed in classical docking-based virtual screening—the idea of combining structural information derived from a protein–ligand complex and the use of a rapid pharmacophore-based screening technique is obvious. There is no competition between ligand-based and structure-based pharmacophore modeling, and both approaches can be used fruitfully in a complementary manner [73]. In fact, a variety of pharmacophore

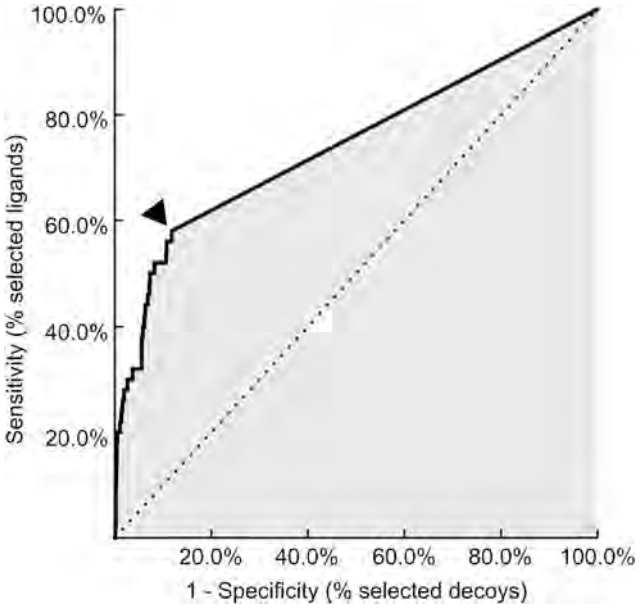


FIGURE 21.4 Example ROC plot for a pharmacophore screening. In this virtual screening run, approximately 60 percent of the biologically active molecules were identified by the model (i.e., part of the virtual hit list), while approximately 85 percent of the inactive compounds could be excluded (i.e., were not part of the virtual hit list). The black arrow illustrates the end of the virtual hit list.

modeling programs allow one to take advantage of additional information provided by a protein or protein–ligand complex structure to help improve the reliability of the generated model. The development of several novel programs for deriving structure-based pharmacophores in the last few years has clearly shown that pharmacophore-based virtual screening is very successful in identifying novel bioactive molecules [74–78]. On the other side, it was also recognized that the consideration of pharmacophores in docking programs can increase reliability and accuracy. Several docking programs are now available which apply the pharmacophore concept to discriminate better between false and real binding modes (e.g., Glide [21], FlexX-Pharm [22], GOLD [20]).

As an illustration, the generation of a structure-based pharmacophore and its application for virtual screening of ABL tyrosine kinase inhibitors is given. STI-571 (Gleevec®) has been approved for the treatment of chronic myelogenous leukemia (CML) and was the first antitumor drug from the family of tyrosine kinase inhibitors [79]. Several crystal structures of STI-571 in complex with different tyrosine kinases (ABL, c-KIT, SYK) have been obtained in the last few years, showing that the compound can bind in varying conformations (open and closed conformation) to different forms of tyrosine kinases. In the case of ABL tyrosine kinase, STI-571 binds to the inactive enzyme form and prevents activation [80]. Several pharmacophores from the available X-ray structures of ABL in complex with STI-571 and analogues (PDB IDs 1IEP, 1FPU and 1OPJ) were generated [81]. In a straightforward approach, the different pharmacophore models were merged using the program LigandScout. The merged pharmacophore contained four lipophilic aromatic areas, two acceptor features, and eight excluded volume spheres. As an example, the structure-based pharmacophore extracted from the X-ray structure 1iep is shown in Figure 21.5. Subsequently a virtual screening was carried out using two different ligand databases. The first one was a collection of 2,765 drug-like ligands from the complexes in the Protein Databank (PDB): the second one was the Maybridge compound library (containing ~59,000 molecules). The pharmacophore model was able to identify all STI-571 entries from the PDB database and did not result in false positives. In addition, seven compounds from the Maybridge database were identified that might represent potential lead structures for the development of novel ABL tyrosine kinase inhibitors.

Several successful applications of the LigandScout program have been reported recently and have supported the feasibility of structure-based pharmacophores to identify novel active molecules [74–77].

B. Pseudo-Receptor Models

Starting in the 1980s, a combination of pharmacophore modeling and structure-based design was introduced and referred to as "pseudo-receptor modeling" or "receptor mapping." [82] Based on a preliminary pharmacophore

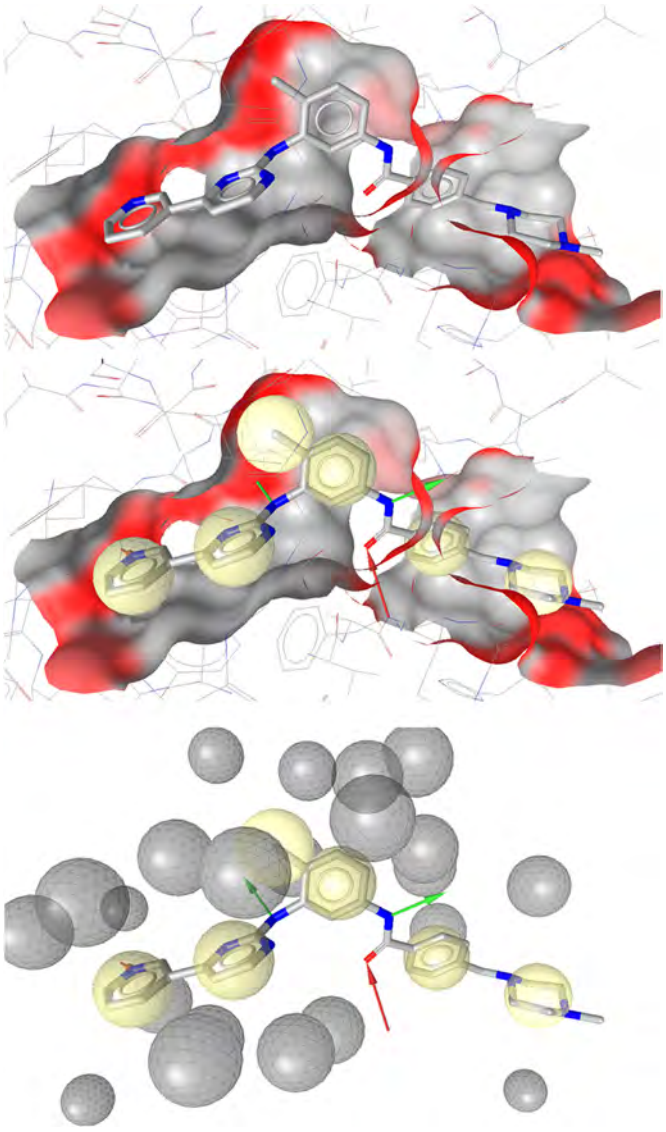


FIGURE 21.5 3D structure of Gleevec® (capped sticks) bound to ABL kinase (top). The structure-based pharmacophore generated with the program LigandScout is shown in the middle, and the extracted pharmacophore together with the excluded volumes (grey spheres) is shown at the bottom. Yellow spheres = hydrophobic features, green arrows = hydrogen bond donors, and red arrows = hydrogen bond acceptors.

model, a hypothetical receptor consisting of individual amino acid residues is constructed, surrounding a set of superimposed ligands. The placement of the individual amino acid residues is guided by experimental data (e.g., from site-directed mutagenesis data). Höltje successfully applied the receptor-mapping technique to several target proteins for which no 3D structure was available [83–85]. Using a data set of twenty 5-HT_{2A} receptor antagonists from different chemical families, a pharmacophore was generated that was able to explain the SAR of the ligand [86]. The receptor mapping (i.e., the placement of the individual amino acids) was based on a homology model of the 5-HT_{2A} receptor generated on the basis of the low-resolution 3D structure of bacteriorhodopsin (a related membrane protein). Using the derived pseudo-receptor, a predictive QSAR model could be obtained that was subsequently applied to design novel potent antagonists [86].

1. *Yak, PrGen, Flo*

Whereas the first pseudo-receptor models were generated more or less intuitively "by hand," which sometimes resulted in irreproducible results, a broader distribution of this concept was achieved by the commercial software packages Yak and PrGen [15]. Both programs allow the generation of a pseudo-receptor in a more or

less automated way. In addition, guided by extensive correlation of experimental and model-derived free energies of binding, a host–guest system is created, mimicking the interaction at a real binding site reasonably well. The fundamental basis of a pseudo-receptor is the placement of the individual amino acid residues. In Yak and PrGen, ligand-specific interaction vectors (the pharmacophoric elements) are calculated and saturated with individual residues from a database of pre-calculated conformations of amino acids. Subsequently, a receptor minimization is carried out by relaxing all residues, keeping the position, orientation, and conformation of the ligands unchanged. To achieve a correlation between the experimentally derived binding affinities (or other biological data) and the calculated interaction energies, a coupling constant is introduced and the system is minimized (correlation-coupled minimization). In a next step, the ligand alignment is allowed to relax within the fixed pseudo-receptor (ligand relaxation). This process (i.e., correlation-coupled minimization followed by unconstrained ligand relaxation) is repeated several times until a highly correlated pseudo-receptor is obtained. To validate the generated pseudo-receptor, its ability to predict the binding affinities of novel ligands must be examined [85]. Therefore, classical QSAR methods such as cross-validation via leave-one-out and/or prediction of external test set compounds are applied. In case of test set or novel ligands, the molecules have to be placed equally to the training set molecules in the pseudo-receptor and have to be minimized applying the same protocol as for the training set ligands.

Another pseudo-receptor modeling approach has been developed by Bohacec et al [87]. Their program, Flo, generates an ensemble of low-energy conformers of each compound of a training set. The conformations are then optimized to minimize the internal energy and maximize the match of chemically similar moieties simultaneously. Then, a pseudo-receptor is composed of functional groups that will mimic the binding cavity. For example, a guanidinium group is selected to form hydrogen bonds with an acidic group of the ligands. The selected residues are positioned around the aligned training set ligands and anchored to the chemically complementary ligand atoms, applying a distance constraint. The remaining volume of the pseudo-receptor is filled with propane molecules to mimic a binding site's hydrophobic surface. In the last step, the pseudo-receptor is equilibrated—comparable to the PrGen approach—by applying several rounds of dynamics. While a pseudo-binding site is quite artificial, the method has the advantage of allowing the binding site to be visualized and used for ligand docking and structure-based design.

2. *Quasar and Raptor*

A further development of Vedani et al was the simplification of the atomistic pseudo-receptor concept (Yak and PrGen) to a quasi-atomistic receptor approach (named Quasar) [88]. Similar to the approach of Walters et al, who developed the program GERM [89], Quasar uses a 3D binding-site surrogate surrounding the ligands instead of a shell of amino acid residues. Each of the virtual particles bears relevant atomistic properties (e.g., H-bond donor, hydrophobic particle). Quasar not only takes into account one conformer per ligand but represents each ligand by an ensemble of low-energy conformations (called "fourth dimension"), thereby reducing the bias associated with the selection of a putative bioactive conformation. Binding of ligand molecules to a macromolecular binding pocket is often facilitated by an induced fit (i.e., the adaptation of a protein to the ligand topology). This effect, which is not considered in most of the pharmacophore and 3D-QSAR approaches, is considered by Quasar and Raptor [90] (the so-called "fifth dimension"). Quantitative models generated with these programs have therefore been named 4D- or 5D-QSARs [91].

3. *Application of Pseudo-Receptor Models*

The pseudo-receptor concept has been applied in recent years to analyze crucial ligand–receptor interaction sites and to establish 3D-QSARs for the prediction of biological activities of ligands [92]. A variety of application studies have shown that the pseudo-receptor concept is a versatile tool in establishing 3D-QSAR models that are often better in their predictive behavior compared to results obtained from classical 3D-QSAR approaches (e.g., CoMFA) [93]. Several application studies have been published that have shown the value—and the limitations—of this approach [82,94].

In a recently published study by Bohacec et al, the pseudo-receptor concept was successfully applied to identify novel small-molecule inducers of fetal hemoglobin [95]. Four available active compounds (Figure 21.6) were selected based on activity and diversity for the construction of an initial pharmacophore.

The initial pharmacophore was constructed using the Flo molecular modeling software [87]. The derived pharmacophore was then successfully tested on a larger ligand data set to see if it could distinguish between active and inactive compounds. Satisfied with the preliminary evaluation of the pharmacophore template, the authors used the model to design novel compounds. The model was sufficiently well defined to allow docking of 630

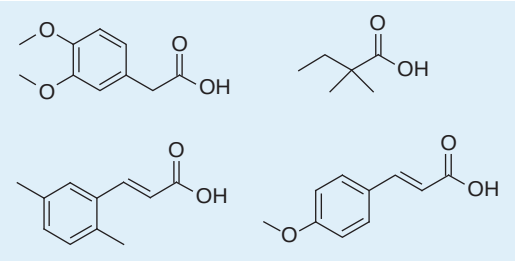


FIGURE 21.6 Molecules used to generate the pharmacophore for inducers of fetal hemoglobin.

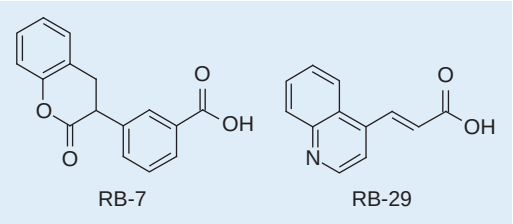


FIGURE 21.7 Molecular structures of the two most active inducers identified by the pseudo-receptor modeling.

compounds and the selection of thirty compounds for testing. Of the twenty-six compounds acquired and tested, four displayed significantly greater activity than previously identified ligands, showing the feasibility of using pseudo-receptor and docking to identify novel bioactive molecules. The structures of the two most potent molecules are shown in Figure 21.7.

When working with pseudo-receptors and in general with quantitative-structure–activity relationships (QSAR) of any dimension, a word of caution is necessary with respect to the biological data that are used. These should preferably constitute binding affinities from a single laboratory, a prerequisite that is also true for all QSAR studies. Since the receptor models simulate interaction events (ΔH) in a highly simplified manner, the experimental data that are combined with them in a correlation analysis must be as close to the molecular level as possible. It is therefore nonsense to correlate the calculated interaction energies with biological *in vivo* data, because the receptor interaction can be blurred or even completely hidden by transport and other pharmacokinetic processes. Sometimes, even the use of *in vitro* data is dangerous if a reaction cascade separates the measured event from the receptor binding interaction. Also, the combination of biological data (e.g., IC_{50} values) from different laboratories or assays is extremely dangerous. The reliability and meaning of any QSAR model (3D-QSAR, pseudo-receptor, 4D-QSAR, 5D-QSAR, 6D-QSAR [96]) should always be assessed by the ultimate test of usefulness, the prediction of new compounds [97]. Very often, QSAR models are internally validated but never tested on whether they are useful in designing novel, more potent compounds [98].

Recently, a novel pseudo-receptor modeling method has been developed named Surflex-QMOD [99,100]. Two datasets, CDK2 inhibitors and muscarinic antagonists, have been chosen to test the performance of the method. Interestingly, the used dataset of congeneric CDK2 inhibitors showed that induced binding pockets can be quite congruent with the enzyme's active site but that model predictivity within a chemical series does not necessarily depend on congruence.

IV. APPLICATION STUDY: NOVEL HISTAMINE H_3-RECEPTOR ANTAGONISTS

A. Pharmacophore-Based Screening

An example from the author's laboratory shall give the reader an informative picture of the pharmacophore generation process and its application to develop novel bioactive compounds [63]. The example deals with antagonists of the human histamine H_3 receptor (hH_3R). hH_3R is a GPCR for which no exact 3D structures is available, as is the case for many other GPCRs. Although we have a crystal structure for the H_1 receptor in the meantime (PDB ID 3RZE), it is still challenging to derive information about the right physiological receptor

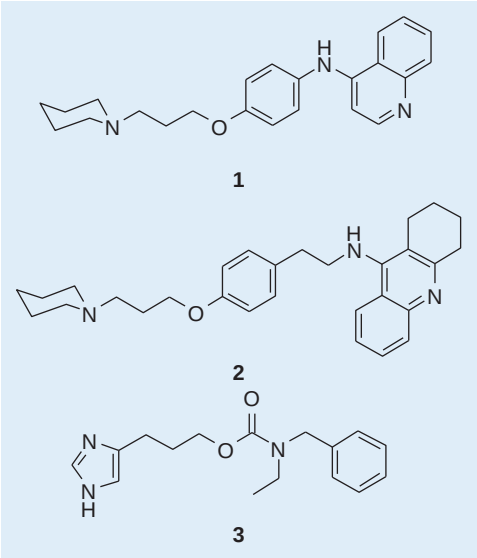


FIGURE 21.8 Molecular structures of hH_3R antagonists used for the pharmacophore development.

conformation by structure-based modeling techniques. The H_3 receptor modulates the release of various neurotransmitters in the central and peripheral nervous system, and therefore is a potential target in the therapy of numerous diseases [101]. Although ligands addressing this receptor are already known, the discovery of alternative lead structures represents a challenging goal in drug design [102]. Experimental structure–activity data for the hH_3R antagonists can be summarized as follows. The pharmacological results suggest that a protonatable nitrogen atom (either in an aromatic imidazole or in a saturated ring system) and an aromatic system separated by a certain distance seem to constitute a potent hH_3R antagonist. Additional polar moieties in the spacer can enhance the antagonistic activity (Figure 21.8).

B. Pharmacophore Determination Process

Due to their high flexibility and huge structural diversity, hH_3R antagonists also provide difficulties in the generation of pharmacophore models by standard means, which normally include the identification of common features required for binding from a ligand set. A dataset of 418 ligands for which hH_3R binding affinities were determined in a $[^3H]N^{\alpha}$-methylhistamine assay is available (pK_i from 5 to 10). A pharmacophore able to discriminate between active and inactive antagonists should be developed on the basis of the known antagonists and be used virtually to screen compound libraries for novel structurally diverse hH_3R antagonists.

For the available ligand dataset, a multi-conformer database was generated using the Catalyst software. An energy cut-off of 20 kcal mol^{-1} from each energetic minimum structure was set in order to avoid high-energy structures. In a first step, three individual pharmacophore models were generated based on the potent antagonists **1**, **2**, and **3**. The bioactive conformation of the ligands was deduced from a conformational analysis of semi-rigid hH_3R antagonists and an extensive docking study carried out on a homology model of the hH_3R (Figure 21.9; for details see [103]). The docking study showed that the homology model is able to explain the interaction of the ligands, which is in accordance with known biochemical data (e.g., site-directed mutagenesis data). However, a receptor-based virtual screening was not very successful in discriminating active from inactive antagonists. Therefore, the idea was to carry out a pharmacophore-based virtual screening.

Defining a pharmacophore model upon a ligand has the advantage that the individual features are already correctly aligned in space. In order to account for the great structural variability of hH_3R antagonists, the pharmacophores were defined as loosely as possible in order to retrieve most of the validated hH_3R ligands as hits.

Once a pharmacophore capable of retrieving known hH_3R antagonists had been defined, it can be used in subsequent screening procedures of commercial compound libraries. As an example, the pharmacophore generated on the basis of compound **1** is shown in Figure 21.10.

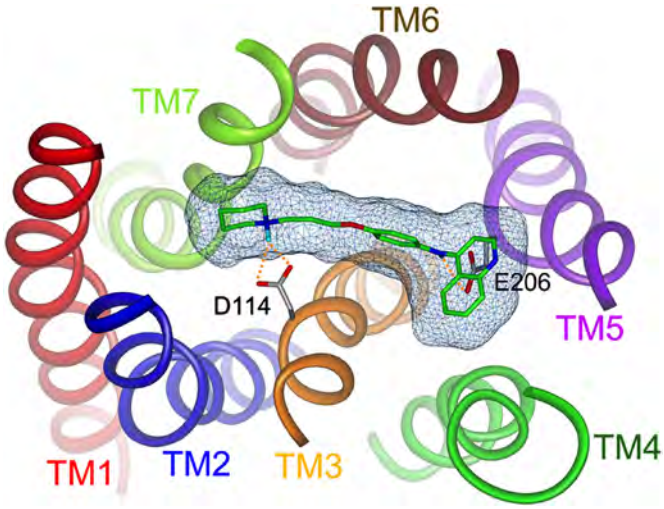


FIGURE 21.9 Interaction of compound **1** with the $hH3_R$ binding site as obtained from the docking study. Only the two important acidic amino acid residues of the binding site are shown for clarity. H-bonds are shown as orange dotted lines.

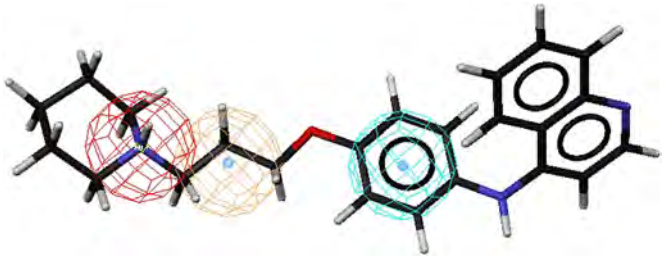


FIGURE 21.10 Pharmacophoric features defined based upon compound **1**. Red sphere = any positively charged element; orange sphere = aromatic or hydrophobic group; cyan sphere = aromatic ring.

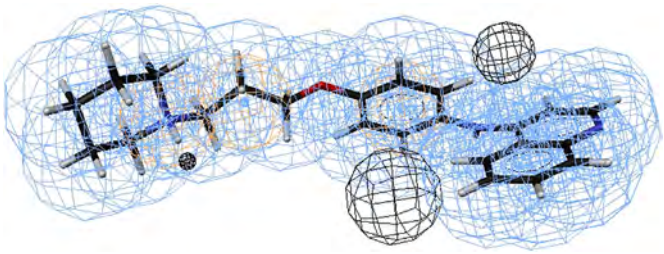


FIGURE 21.11 Pharmacophore model based on compound **1** including the shape feature (van der Waals volume) and the two forbidden volumes (black spheres).

The choice of chemical features was based on functionalities observed in validated hH_3R antagonists and inspection of the binding pocket of the homology model. The linker moiety and the adjacent hydrophobic/π-electron-rich system of the ligands lie in a cleft between trans-membrane region (TM) 3, 6, and 7 of the hH_3R. In this region, several aromatic residues border the binding site that are able to interact with the electron-rich system in the hH_3R antagonists. No pharmacophoric features were defined upon the 4-aminoquinoline moiety, as a high degree of chemical diversity is observed in active ligands within this region. Any restriction of chemical features was thus avoided.

Apparently, the derived pharmacophore model is too loose fitting for screening a compound database. Thus, the van der Waals volume of ligand **1** was included as an additional constraint into the pharmacophore model. Default parameters were used for the definition of the shape query. Finally, forbidden volumes (black spheres) were defined in order to account for the fact that some ligands extending into these areas were inactive, although they resembled other active compounds. Figure 21.11 shows the ligand **1** fitted into the complete pharmacophore model.

Using this model, 316 compounds from the 418 ligand dataset were found as hits in a pharmacophore search using the Catalyst [36] program. Ninety-three percent of the ligands with highest activity were retrieved by the pharmacophore model; less satisfactorily, 54 percent of a set of inactive compounds could also pass the pharmacophoric filter (Figure 21.11, top). Application of the pharmacophore filter for screening the Maybridge Database (MDB) and the World Drug Index (WDI) resulted in 249 and 929 hits, respectively. Thus, 70 percent of the active and moderate active hH_3R ligands (with a $pKi > 7$) were retrieved by the pharmacophore. Meanwhile, from the pool of MDB and WDI ligands (MDB: 59,000 compounds; WDI: 48,000 compounds), 98.9 percent could be excluded.

The filter was still quite loose, however, so that a subsequent definition of further pharmacophoric features could result in a better separation of in/actives. In order to further increase the percentage of active hH_3R ligands found during the virtual screening, further pharmacophore models were defined in a similar way based upon compounds **2** ($Ki = 0.33$ nM) and **3** ($Ki = 69$ nM). For the definition of the pharmacophore derived from compound **2**, the three features described above were again used in combination with a shape query and forbidden volumes. The third individual pharmacophore model was defined based on ligand **3** (see Figure 21.12), capable of retrieving 68 percent of ligands deposited in the hH_3R database. By combining the three pharmacophore models, 369 of 398 (93 percent) hH_3R ligands with a $pKi > 7$ could be obtained, while only 2668 (2.5 percent) compounds were obtained as hits when screening the MDB and WDI database with 107,599 total structures deposited. The small percentage of structures from commercial databases matching the pharmacophores showed that the generated models were stringent enough for a reasonable screening.

C. Pharmacophore-Based Screening of Compound Libraries

For a more stringent screening, a leave-one-out (LOO) filter was defined on the pharmacophoric features of **1.** The Catalyst LOO model consisted of a combination of five individual pharmacophore models, each lacking one pharmacophoric feature found in compound **1** at a time, with the exception of the positive ionizable group and the spacer moiety that were required in all models. The screening of the 2,668 WDI and MDB compounds with the LOO filter reduced the number of hits to 320. In order to ensure that compounds selected by the pharmacophore-based screening could be accommodated into the hH_3R binding site, the 320 hits were docked into the hH_3R binding site using the GOLD [20] program and ranked according to their docking scores. From the top-ranked complexes, seven MDB compounds were selected for experimental testing. The selection of the seven compounds was guided by a cluster analysis in order to select the most structurally diverse compounds among the top-ranked molecules. All compounds showed affinity for the hH_3R with binding affinities ranging from 79 nM to 6.3 μM, thereby showing that the pharmacophore and hH_3R binding site model can be used to identify novel active antagonists. Two compounds, BTB-08079 and RJC-03033, were found to be active in the nanomolar range (Figure 21.13) [63].

In order to determine the structural similarity between the seven retrieved MDB compounds and the 418 hH_3R ligands, we calculated similarity indices on the basis of different fingerprint systems (MACCS keys and graph-3-point pharmacophore fingerprints in MOE [41]). Using the different fingerprint systems, low similarities were observed between the seven MDB compounds and the original hH_3R antagonists. For the most potent hit (BTB-08079; 79 nM), the lowest similarity to the original hH_3R ligand structures was observed. The dimethylaminofuran fragment, which was already known from the potent histamine H_2 receptor antagonist Ranitidine, was not reported before as a structural element of potent H_3R antagonists.

Compared to the receptor-based virtual screening, application of the pharmacophore-based search resulted in significantly improved results. In the docking approach, 66.6 percent of the hH_3R ligands were retrieved, limiting the number of WDI and MDB compounds to approximately 1,720 structures, but application of a pharmacophore-based search allowed retrieval of 93 percent of active compounds, while reducing the number of WDI and MDB structures to 2,668 compounds (2.5 percent). The ideal strategy for the flexible hH_3R ligand data set, however, appeared to be a combined approach comprising a pre-screening of commercial databases with relatively loose pharmacophore models that mainly reflect the available volume in the binding site (e.g., by considering shape queries of sterically demanding ligands and forbidden volumes derived from ligand superposition) and some essential requirements for binding such as the protonated head group. In order to ensure that compounds selected by the pharmacophore-based screening fit into the binding site, docking of this subset of ligands resulted in a selection of candidates for biological testing.

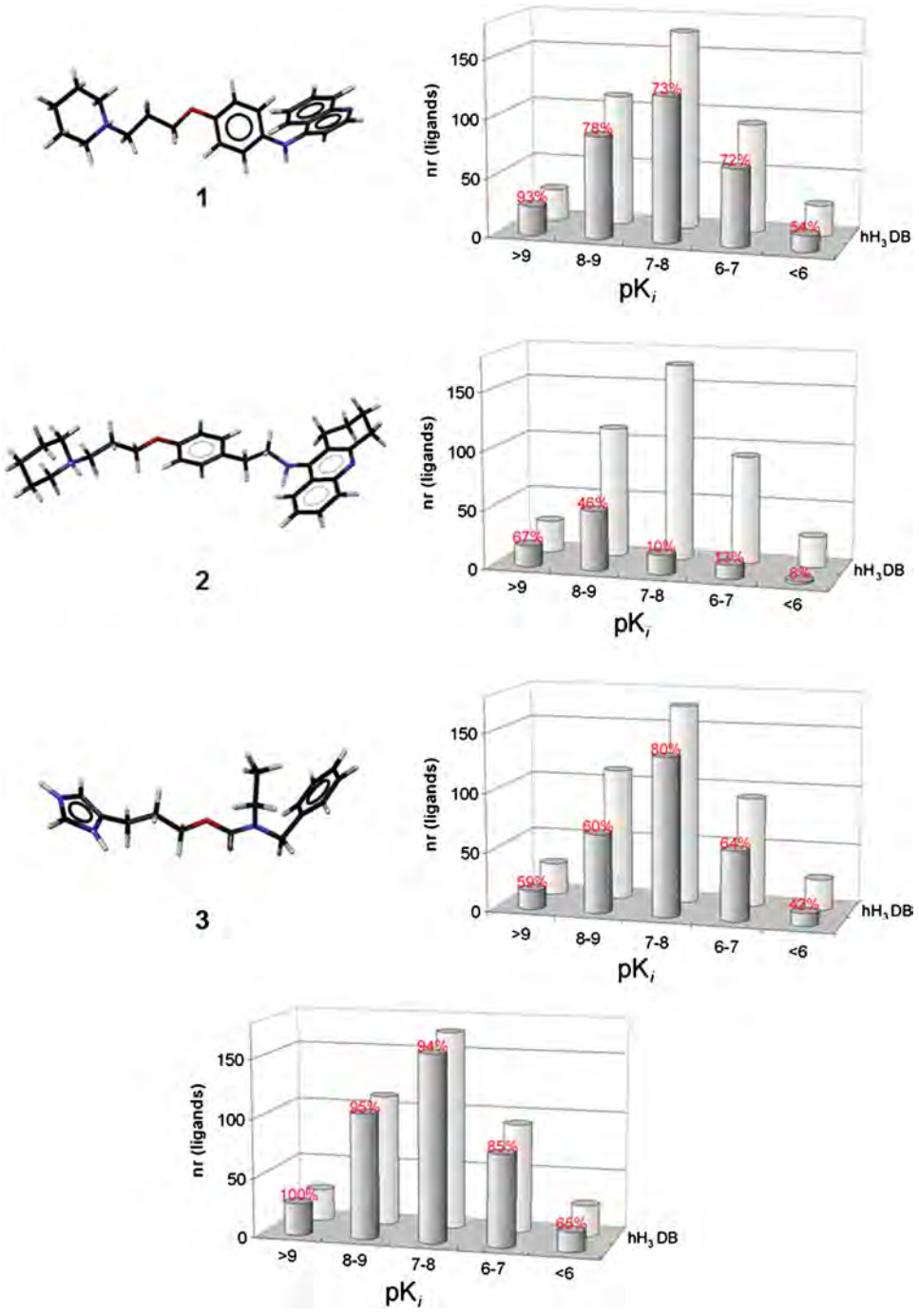


FIGURE 21.12 Enrichment of hH_3R ligands by pharmacophore search based on compound **1**, **2**, and **3**. The percentage of hH_3R ligands retrieved by the individual pharmacophore model within each p*Ki*-cluster is depicted. The percentage of ligands found in each cluster (dark columns) is written in red numbers and compared to the population of p*Ki* clusters of all hH_3R compounds in the 418 ligand data set (light grey columns).

V. RECENT DEVELOPMENTS AND OUTLOOK

3D pharmacophores have evolved as important tools for describing protein−ligand interactions, and the number of examples that successfully predict biological activity using pharmacophore methods is constantly growing. A multitude of therapeutic areas is covered, such as anti-viral drug discovery [104,105], the discovery of novel anti-bacterial agents [106,107], and modeling of GPCR ligand interaction (Figure 21.14) [108,109], among others.

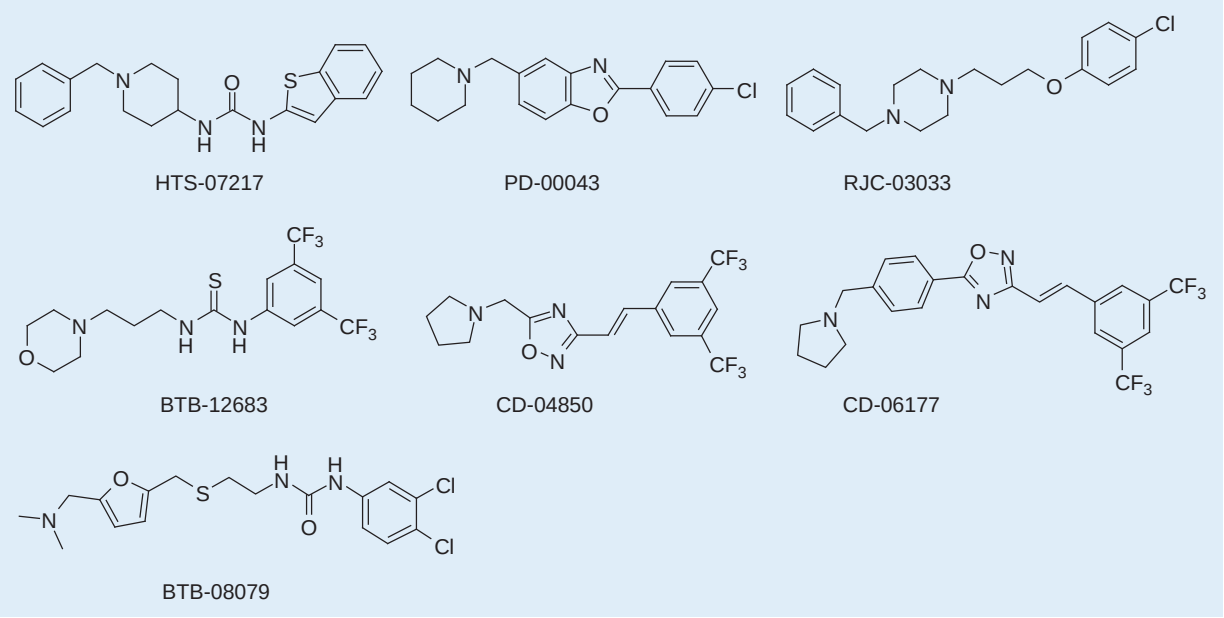


FIGURE 21.13 Selected hits from the pharmacophore-based virtual screening.

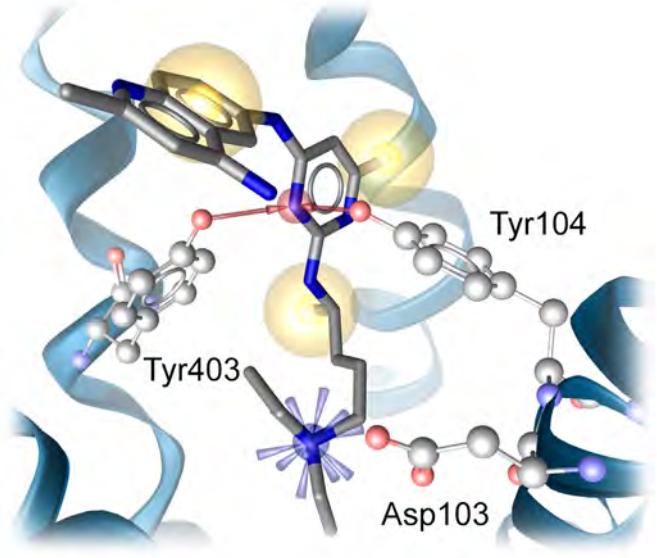


FIGURE 21.14 LigandScout 3D pharmacophore model example for the newly discovered muscarinic acetylcholine receptor antagonist NSC23766. Red arrows = H-bond donors; yellow spheres = lipophilic contact areas; blue rays = positively charged groups.

Technically, all these approaches follow the sequence of careful model design, prediction, and selection of existing compounds from libraries with previously unknown biological activity or *de novo* design to fulfil pharmacophoric requirements. The crucial step remains model design and careful retrospective validation before performing a prospective virtual screening, which is followed by purchasing compounds and biological tests. If model building is done carefully, it can lead to successful predictions and rationalization of a possible ligand-binding mode. Mainly due to the better usability of recent programs, 3D pharmacophores are now also used aside of the typical workflow of virtual screening against a single protein-binding site. The program Ligandscout, for example, now includes a module for fragment-based screening that aims at supporting fragment-based *de novo* design [110].

Another important new application is parallel screening against several pharmacophores to predict multi-target effects [111,112]. Although this can also be done using various other molecular modeling techniques, such as classical QSAR or descriptor approaches [113], multi-target predictions using 3D pharmacophores have the advantage that results remain easy to interpret and provide intuitive starting points for further optimization.

As previously mentioned pharmacophore screening comparisons [43] show, algorithm development for compound/pharmacophore mapping remains a challenge. This leaves room for further algorithmic improvements, such as the implementation of advanced filter methods like bloom filtering as implemented in recent tools like Pharmer [114] or LigandScout [44].

Another recent approach developed by Rognan's group uses 3D fingerprints based on pharmacophoric features to compare protein-binding sites in terms of ligand binding properties (including druggability). In a comparable fashion, pharmacophore fingerprints can be used to describe subpocket similarity [115].

VI. CONCLUSIONS

In spite of the recent success and popularity of pharmacophore-based drug design, one should not forget the limitations of pharmacophore modeling. As with any other model, we should be aware of the abstraction that is applied to generate these models. All pharmacophore approaches are based on molecular mechanical abstractions. Thus, properties associated with the interaction of electrons (e.g., polarization effects) are not considered. Another limitation in many pharmacophore-based approaches is the neglect of the dynamic nature of protein–ligand interaction. Although novel pharmacophore generation programs allow the parallel consideration of multiple/alternative pharmacophores (e.g., Catalyst [38], LigandScout [39]), modeling different binding modes is still a challenge. It is becoming increasingly clear that for some protein binding sites, one has to be prepared to consider different binding modes and therefore different pharmacophores [116–118].

Whereas in the past, pharmacophore models have been mainly generated using ligand-based strategies, novel programs have been developed and applied successfully in the last few years by combining structure-based and pharmacophore-based approaches. This is mainly influenced by the rapidly growing number of protein–ligand 3D structures that are the basis for such combined approaches. Closely related to this, one can observe a general merging of different techniques in molecular modeling studies—pharmacophore modeling, 3D-QSAR, *de novo* design, and docking [119,120]—that might be helpful for future drug design studies.

References

[1] Wermuth CG, Ganellin CR, Lindberg P, Mitschler LA. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1997). Annu Rep Med Chem 1998;33:385–95.

[2] Langley JN. On the reaction of cells and nerve-endings to certain poissons, chiefly as regards the reaction of striated muscle to nictoine and to cuari. J Physiol 1905;33:374–413.

[3] Ehrlich P, Morgenroth J. Über Haemolysine. Dritte Mitteilung [About haemolysins. Third report]. Berl Klin Wochenschr 1900;37:453–7. [in German].

[4] Fischer E. Einfluss der Konfiguration auf die Wirkung der Enzyme [Influence of configuration on the action of enzymes]. Ber Dtsch Chem Ges 1894;27:2985–93. [in German].

[5] Perutz MF. X-ray analysis of haemoglobin. Science 1963;140:863–9.

[6] Marshall GR. Binding-site modelling of unknown receptors. In: Kubinyi H, editor. 3D QSAR in drug design: theory methods and applications. Leiden: ESCOM; 1993. p. 80–116.

[7] Barry CD, Ellis RA, Graesser S, Marshall GR. Display and manipulation in three dimensions. In: Faiman M, Nievergelt J, editors. Pertinent concepts in computer graphics. Chicago: University of Illinois Press; 1969. p. 104–53.

[8] Kier LB, Aldrich HS. A theoretical study of receptor site models for trimethylammonium group interactions. J Theor Biol 1974;46:529–41.

[9] Gund P, Wipke WT, Langridge R. Computer searching of a molecular structure file for pharmacophoric patterns. Comput Chem Res Educ Technol 1974;3:5–21.

[10] Poptodorov K, Liu T, Hoffmann RD. Pharmacophore model generation software tools. In: Langer T, Hoffmann RD, editors. Pharmacophores and pharmacophore searches. Weinheim, Germany: Wiley-VCH; 2006. p. 17–47.

[11] Sanders MP, Barbosa AJ, Zarzycka B, Nicolaes GA, Klomp JP, de Vlieg J, et al. Comparative analysis of pharmacophore screening tools. J Chem Inf Model 2012;52:1607–20.

[12] Evans BE, Rittle KE, Bock MG, DiPardo RM, Freidinger RM, Whitter WL, et al. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. J Med Chem 1988;31:2235–46.

[13] Wermuth CG. Pharmacophores: historical perspective and viewpoint from a medicinal chemist. In: Langer T, Hoffmann RD, editors. Pharmacophores and pharmacophore searches. Weinheim, Germany: Wiley-VCH; 2006. p. 1–13.

[14] Snyder JP, Rao SN, Koehler A, Vedani A. Minireceptors and pseudoreceptors. In: Kubinyi H, editor. 3D QSAR in drug design: theory methods and applications. Leiden: ESCOM; 1993. p. 336–54.

[15] Vedani A, Zbinden P, Snyder JP. Pseudo-receptor modelling: a new concept for the three-dimensional construction of receptor binding sites. J Recept Res 1993;13:163–77.
[16] Höltje HD, Anzali A. Molecular modelling studies on the digitalis binding site of the Na+/K(+)-ATPase. Pharmazie 1992;47:691–8.
[17] Frühbeis H, Klein R, Wallmeier H. Computergestütztes Moleküldesign (CAMD) – ein Überblick [Computer aided molecular design (CAMD): An overview]. Angew Chem Int Ed Engl 1987;26:403–18. [in German].
[18] Böhm HJ, Klebe G, Kubinyi H. Wirkstoffdesign: Der Weg zum Arzneimittel [Drug design: The path to drugs]. Heidelberg, Germany: Spektrum Akademischer Verlag; 1996. [in German].
[19] Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. Proteins 2002;46:34–40.
[20] Gold, Cambridge Crystallographic Data Centre, Cambridge, UK. Available from: <http://www.ccdc.cam.ac.uk/>.
[21] Glide, Schrödinger, Portland, OR. Available from: <http://www.schroedinger.com/>.
[22] FlexE, BiosolveIT, St. Augustin, Germany. Available from: <http://www.biolsolveit.de/>.
[23] Klebe G. Virtual ligand screening. Drug Discov Today 2006;11:580–94.
[24] Tame JR. Scoring functions: the first 100 years. J Comput Aided Mol Des 2005;19:445–51.
[25] Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, et al. A critical assessment of docking programs and scoring functions. J Med Chem 2005;49:5912–31.
[26] Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. Br J Pharmacol 2008;153:S7–26.
[27] Yuriev E, Ramsland PA. Latest developments in molecular docking: 2010–2011 in review. J Mol Recognit 2013;26:215–39.
[28] Sousa SF, Ribeiro AJ, Coimbra JT, Neves RP, Martins SA, Moorthy NS, et al. Protein-ligand docking in the new millennium: a retrospective of 10 years in the field. Curr Med Chem 2013;20:2296–314.
[29] Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature 2013;494:185–94.
[30] Siu FY, He M, de Graaf C, Han GW, Yang D, Zhang Z, et al. Structure of the human glucagon class B G-protein-coupled receptor. Nature 2013;499:444–9.
[31] Hollenstein K, Kean J, Bortolato A, Cheng RKY, Doré AS, Jazayeri A, et al. Structure of class B GPCR corticotropinreleasing factor receptor 1. Nature 2013;499:438–43.
[32] Hackling A, Ghosh R, Perachon S, Mann A, Höltje H-D, Wermuth CG, et al. N-(4-(2-Methoxyphenyl)piperazin-1-yl)alkylcarboxamides as dopamine D_2- and D_3-receptor ligands. J Med Chem 2003;46:3883–99.
[33] Cramer III RD, Patterson DE, Bunce JD. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. J Am Chem Soc 1988;110:5959–67.
[34] Ahlstrom MM, Ridderstrom M, Luthman K, Zamora I. Virtual screening and scaffold hopping based on GRID molecular interaction fields. J Chem Inf Model 2005;45:1313–23.
[35] Steindl TM, Schuster D, Wolber G, Laggner C, Langer T. High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening. J Comput Aided Mol Des 2006;20:703–15.
[36] Catalyst, Accelrys Software, San Diego, CA. Available from: <http://www.accelrys.com/>.
[37] DISCOtech, St. Louis, MO. Available from: <http://www.tripos.com/>.
[38] Galahad, Tripos, St. Louis, MO. Available from: <http://www.tripos.com/>.
[39] LigandScout, Inteligand, Wien, Austria. Available from: <http://www.inteligand.com/>.
[40] Phase, Schrödinger, Portland, OR. Available from: <http://www.schroedinger.com/>.
[41] MOE, Chemical Computing Group, Montreal, Canada. Available from: <http://www.chemcomp.com/>.
[42] Patel Y, Gillet VJ, Bravi G, Leach A. A comparison of the pharmacophore identification programs: catalyst, DISCO, and GASP. J Comput Aided Mol Des 2002;16:653–81.
[43] Spitzer G, Heiss M, Mangold M, Kirchmair J, Wolber G, Liedl K. One concept, three implementations of 3D pharmacophore-based virtual screening: distinct coverage of chemical search space. J Chem Inf Model 2010;50:1241–7.
[44] Seidel T, Ibis G, Bendix F, Wolber G. Strategies for 3D pharmacophore-based virtual screening. Drug Discov Today Technol 2010;7:e221–8.
[45] Boström J, Norrby P-O, Liljefors T. Conformational energy penalties of protein-bound ligands. J Comput Aided Mol Design 1998;12:383–96.
[46] Agrafiotis DK, Gibbs AC, Zhu F, Izrailev S, Martin E. Conformational sampling of bioactive molecules: a comparative study. J Chem Inf Model 2007;47:1067–86.
[47] Höltje H-D, Sippl W, Rognan D, Folkers G. Molecular modelling: basic principles and applications. 3rd ed. Weinheim, Germany: Wiley-VCH; 2007.
[48] Marshall GR, Barry CD, Bosshard HE, Dammkoehler RA, Dunn DA. The conformational parameter in drug design: the active analog approach. In: Olsen EC, Christoffersen RE, editors. Computer-assisted drug design. ACS Monograph 112. Washington DC: American Chemical Society; 1979. p. 205–26.
[49] Smellie A, Teig S, Towbin P. Poling: promoting conformational variation. J Comp Chem 1994;16:171–87.
[50] Kearsley SK, Smith GM. An alternative method for the alignment of molecular structures: maximizing electrostatic and steric overlap. Tetrahedron Comput Methodol 1990;3:615–33.
[51] Kirchmair J, Wolber G, Laggner C, Langer T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. J Chem Inf Model 2006;46:1848–61.
[52] Boström J. Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. J Comput Aided Mol Design 2001;15:1137–52.
[53] Omega, OpenEye Scientific Software, Santa Fee, NM. Available from: <http://www.eyesopen.com/>.

[54] Boström J, Greenwood JR, Gottfries J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. J Mol Graph Mod 2003;21:449–62.
[55] Jones G, Willet P. GASP: genetic algorithm superimposition program. In: Güner OF, editor. Pharmacophore perception, development, and use in drug design. La Jolla. CA: International University Line; 2000. p. 85–106.
[56] Jones G, Willett P, Glen RC. A genetic algorithm for flexible molecular overlay and pharmacophore elucidation. J Comput Aided Mol Des 1995;9:532–49.
[57] Reutlinger M, Koch CP, Reker D, Todoroff N, Schneider P, Rodrigues T, et al. Chemically advanced template search (CATS) for scaffold-hopping and prospective target prediction for "orphan" molecules. Mol Inf 2013;32:133–8.
[58] Brint AT, Willet P. Algorithms for the identification of three-dimensional maximal common substructures. J Chem Inf Comput Sci 1987;27:152–8.
[59] Martin YC. DISCO: what we did right and what we missed. In: Güner OF, editor. Pharmacophore perception, development, and use in drug design. La Jolla. CA: International University Line; 2000. p. 51–66.
[60] Martin YC, Bures MG, Danaher EA, DeLazzer J, Lico I, Pavlik PA. A fast new approach to pharmacophore mapping and its application to dopaminergic and benzodiazepine agonists. J Comput Aided Mol Des 1993;7:83–102.
[61] Wolber G, Langer T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. J Chem Inf Model 2005;45:160–9.
[62] Wolber G, Dornhofer AA, Langer T. Efficient overlay of small organic molecules using 3D pharmacophores. J Comput Aided Mol Des 2006;20:773–88.
[63] Schlegel B, Meier R, Laggner C, Schnell D, Langer T, Seifert R, et al. Generation of a homology model of the human histamine H3 receptor for ligand docking and pharmacophor based screening. J Comput Aided Mol Des 2007;21:437–53.
[64] UNITY, Tripos, St. Louis, MO. Available from: <http://www.tripos.com/>.
[65] Ebejer JP, Morris GM, Deane CM. Freely available conformer generation methods: how good are they? J Chem Inf Model 2012;52:1146–58.
[66] Leach AR, Gillet VJ, Lewis RA, Taylor R. Three-dimensional pharmacophore methods in drug discovery. J Med Chem 2010;53:539–58.
[67] Kirchmair J, Ristic S, Eder K, Markt P, Wolber G, Laggner C, et al. Fast and efficient in silico 3D screening: toward maximum computational efficiency of pharmacophore-based and shape-based approaches. J Chem Inf Model 2007;47:2182–96.
[68] Laggner C, Wolber G, Kirchmair J, Schuster D, Langer T. Pharmacophore-based virtual screening in drug discovery. In: Varnek A, Tropsha A, editors. Chemoinformatics: an approach to virtual screening. Royal Society of Chemistry; 2008. p. 76–119.
[69] Langer T, Wolber G. Pharmacophore definition and 3D searches. Drug Discov Today Technol 2004;1:203–7.
[70] Truchon J-F, Bayly CI. Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. J Chem Inf Model 2007;47:488.
[71] Kirchmair J, Markt P, Distinto S, Wolber G, Langer T. Evaluation of the performance of 3D virtual screening protocols: RMSD comparisons, enrichment assessments, and decoy selection—What can we learn from earlier mistakes? J Comput-Aided Mol Des 2007;22:213–28.
[72] Triballeau N, Acher F, Brabet I, Pin JP, Bertrand HO. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach: application to high-throughput docking on metabotropic glutamate receptor subtype 4. J Med Chem 2005;48:2534–47.
[73] Hindle SA, Rarey M, Buning C, Lengauer T. Flexible docking under pharmacophore type constraints. J Comput Aided Mol Des 2002;16:129–49.
[74] Barreca ML, De Luca L, Iraci N, Rao A, Ferro S, Maga G, et al. Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors. J Chem Inf Model 2007;47:557–62.
[75] Krovat EM, Fruhwirth KH, Langer T. Pharmacophore identification, in silico screening, and virtual library design for inhibitors of the human factor Xa. J Chem Inf Model 2005;45:146–59.
[76] Maccari R, Ottana R, Ciurleo R, Vigorita MG, Rakowitz D, Steindl T, et al. Evaluation of in vitro aldose redutase inhibitory activity of 5-arylidene-2,4-thiazolidinediones. Bioorg Med Chem Lett 2007;17:3886–93.
[77] Laggner C, Schieferer C, Fiechtner B, Poles G, Hoffmann RD, Glossmann H, et al. Discovery of high-affinity ligands of sigma1 receptor, ERG2, and emopamil binding protein by pharmacophore modelling and virtual screening. J Med Chem 2005;48:4754–64.
[78] Spannhoff A, Heinke R, Bauer I, Trojer P, Metzger E, Gust R, et al. Target-based approach to inhibitors of histone arginine methyltransferases. J Med Chem 2007;50:2319–25.
[79] Capdeville R, Buchdunger E, Zimmermann J, Matter A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. Nat Rev Drug Discov 2002;7:493–502.
[80] Mol CD, Fabbro D, Hosfield DJ. Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors. Curr Opin Drug Discov Devel 2004;5:639–48.
[81] Wolber G, Kosara R. Pharmacophores from macromolecular complexes with ligandscout. In: Langer T, Hoffmann RD, editors. Pharmacophores and pharmacophore searches. Weinheim, Germany: Wiley-VCH; 2006. p. 131–50.
[82] Schleifer KJ. Concept and applications of pseudoreceptors. In: Langer T, Hoffmann RD, editors. Pharmacophores and pharmacophore searches, 2006. Weinheim, Germany: Wiley-VCH; 2006. p. 117–30.
[83] Höltje H-D. Theoretical studies on drug-receptor model complexes. Pharm Acta Helv 1979;54:125–34.
[84] Marrer S, Höltje H-D. Molecular modelling study of calcium channel activating and blocking 1,4-dihydropyridine derivatives: a hypothetical molecular receptor model. J Protein Chem 1989;8:420–1.
[85] Sippl W, Stark H, Höltje H-D. Development of a binding site model for histamine H_3-receptor agonists. Pharmazie 1998;53:433–7.
[86] Höltje H-D, Jendretzki UK. Construction of a detailed serotoninergic 5-HT_{2a} receptor model. Arch Pharm (Weinheim) 1995;328:577–84.
[87] McMartin C, Bohacek RS. QXP: powerful, rapid computer algorithms for structure-based drug design. J Comput Aided Mol Des 1997;11:333–44.

[88] Vedani A, Zbinden P. Quasi-atomistic receptor modelling: a bridge between 3D QSAR and receptor fitting. Pharm Acta Helv 1998;73:11–18.
[89] Walters DE, Hinds RM. Genetically evolved receptor models: a computational approach to construction of receptor models. J Med Chem 1994;37:2527–36.
[90] Lill MA, Vedani A, Dobler M. Raptor: combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modelling: application toward the simulation of structurally diverse ligand sets. J Med Chem 2004;47:6174–86.
[91] Lill MA, Vedani A. Combining 4D pharmacophore generation and multidimensional QSAR: modelling ligand binding to the bradykinin B2 receptor. J Chem Inf Model 2006;46:2135–45.
[92] Hahn M, Rogers D. Receptor surface models. 2. Application to quantitative structure-activity studies. J Med Chem 1995;38:2091–102.
[93] Schleifer KJ. Pseudoreceptor model for ryanodine derivatives at calcium release channels. J Comput Aided Mol Des 2000;14:467–75.
[94] Vedani A, Dobler M, Dollinger H, Hasselbach KM, Birke F, Lill MA. Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. J Med Chem 2005;48:1515–27.
[95] Bohacek R, Boosalis MS, McMartin C, Faller DV, Perrine SP. Identification of novel small-molecule inducers of fetal hemoglobin using pharmacophore and "PSEUDO" receptor models. Chem Biol Drug Des 2006;67:318–28.
[96] Vedani A, Dobler M, Lill MA. Combining protein modeling and 6D-QSAR: simulating the binding of structurally diverse ligands to the estrogen receptor. J Med Chem 2005;48:3700–3.
[97] Sippl W, Contreras JM, Parrot I, Rival YM, Wermuth CG. Receptor-based QSAR and design of novel acetylcholinesterase inhibitors. J Comput Aided Mol Des 2001;15:395–410.
[98] Doweyko AM. 3D-QSAR illusions. J Comput Aided Mol Des 2004;18:587–96.
[99] Varela R, Cleves AE, Spitzer R, Jain AN. A structure-guided approach for protein pocket modeling and affinity prediction. J Comput Aided Mol Des 2013;27:917–34.
[100] Jain AN. QMOD: physically meaningful QSAR. J Comput Aided Mol Des 24:865–878.
[101] Ligneau X, Perrin D, Landais L, Camelin JC, Calmels TP, Berrebi-Bertrand I, et al. BF2.649 [1-{3-[3-(4-Chlorophenyl) propoxy]propyl} piperidine, hydrochloride], a nonimidazole inverse agonist/antagonist at the human histamine H_3 receptor: preclinical pharmacology. J Pharmacol Exp Ther 2007;320:365–75.
[102] Stark H, Kathmann M, Schlicker E, Schunack W, Schlegel B, Sippl W. Medicinal chemical and pharmacological aspects of imidazole-containing histamine H_3 receptor antagonists. Mini Rev Med Chem 2004;4:965–77.
[103] Schlegel B, Sippl W, Höltje H-D. Molecular dynamics simulations of bovine rhodopsin: influence of protonation states and different membrane-mimicking environments. J Mol Model 2005;12:49–64.
[104] Murgueitio MS, Bermudez M, Mortier J, Wolber G. In silico virtual screening approaches for anti-viral drug discovery. Drug Discov Today Technol 2012;9:e219–25.
[105] Kirchmair J, Distinto S, Liedl K, Markt P, Rollinger JM, Schuster D, et al. Development of anti-viral agents using molecular modeling and virtual screening techniques. Infect Disord Drug Targets 2011;11:64–93.
[106] Perdih A, Kovaç A, Wolber G, Gobec S, Solmajer T. Discovery of novel benzene 1,3-dicarboxylic acid inhibitors of bacterial MurD and MurE ligases by structure-based virtual screening approach. Bioorg Med Chem Letters 2009;19:2668–73.
[107] Perdih A, Wolber G, Solmajer T. Molecular dynamics simulation and linear interaction energy study of d-Glu-based inhibitors of the MurD ligase. J Comput Aided Mol Des 2013;27:723–38.
[108] Levay M, Krobert KA, Wittig K, Voigt N, Bermudez M, Wolber G, et al. NSC23766, a widely used inhibitor of Rac1 activation, additionally acts as a competitive antagonist at muscarinic acetylcholine receptors. J Pharmacol Exp Ther 2013;347:69–79.
[109] Leschner J, Wennerberg G, Feierler J, Bermudez M, Wolber G, Welte B, et al. Interruption of the ionic lock in the bradykinin B2 receptor results in constitutive internalization and turns several antagonists into strong agonists. J Pharmacol Exp Ther 2013;344:85–95.
[110] Mortier J, Rakers C, Frédérick R, Wolber G. Computational tools for in silico fragment-based drug design. Curr Top Med Chem 2012;12:1935–43.
[111] Schuster D. 3D pharmacophores as tools for activity profiling. Drug Discov Today Technol 2010;7:e205–11.
[112] Meslamani J, Li J, Sutter J, Stevens A, Bertrand H-O, Rognan D. Protein–ligand-based pharmacophores: generation and utility assessment in computational ligand profiling. J Chem Inf Model 2012;52:943–55.
[113] Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. Nat Biotech 2007;25:197–206.
[114] Koes DR, Camacho CJ. Pharmer: efficient and exact pharmacophore search. J Chem Inf Model 2011;51:1307–14.
[115] Wood DJ, de Vlieg J, Wagener M, Ritschel T. Pharmacophore fingerprint-based approach to binding site subpocket similarity and its application to bioisostere replacement. J Chem Inf Model 2012;52:2031–43.
[116] Steuber H, Zentgraf M, La Motta C, Sartini S, Heine A, Klebe G. Evidence for a novel binding site conformer of aldose reductase in ligand-bound state. J Mol Biol 2007;369:186–97.
[117] Steuber H, Zentgraf M, Gerlach C, Sotriffer CA, Heine A, Klebe G. Expect the unexpected or caveat for drug designers: multiple structure determinations using aldose reductase crystals treated under varying soaking and co-crystallisation conditions. J Mol Biol 2006;363:174–87.
[118] Specker E, Bottcher J, Brass S, Heine A, Lilie H, Schoop A, et al. Unexpected novel binding mode of pyrrolidine-based aspartyl protease inhibitors: design, synthesis, and crystal structure in complex with HIV protease. Chem Med Chem 2006;1:106–17.
[119] Sippl W. Receptor-based 3D QSAR analysis of estrogen receptor ligands: merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods. J Comput Aided Mol Des 2000;14:559–72.
[120] Evers A, Gohlke H, Klebe G. Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. J Mol Biol 2003;334:327–45.

CHAPTER

22

Protein Crystallography and Drug Discovery

Jean-Michel Rondeau[1] *and Herman Schreuder*[2]

[1]Novartis Institutes for BioMedical Research, Basel, Switzerland; [2]Sanofi, R&D LGCR/Struct., Design & Informatics FF, Frankfurt am Main, Germany

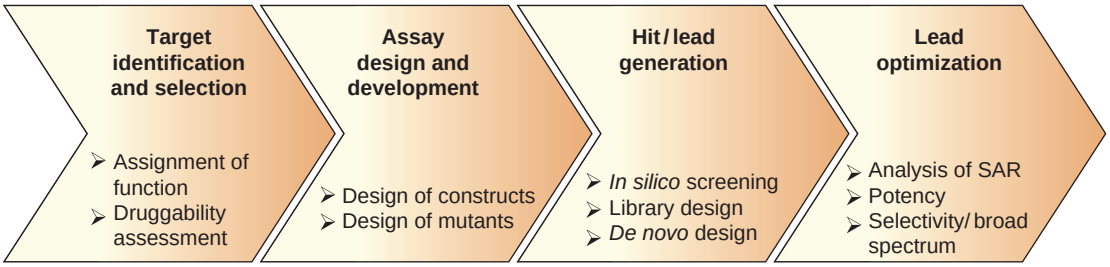
OUTLINE

If you can look into the seeds of time
And say which grain will grow and which will not
Speak then to me ... ***Shakespeare, Macbeth*** *[1]*

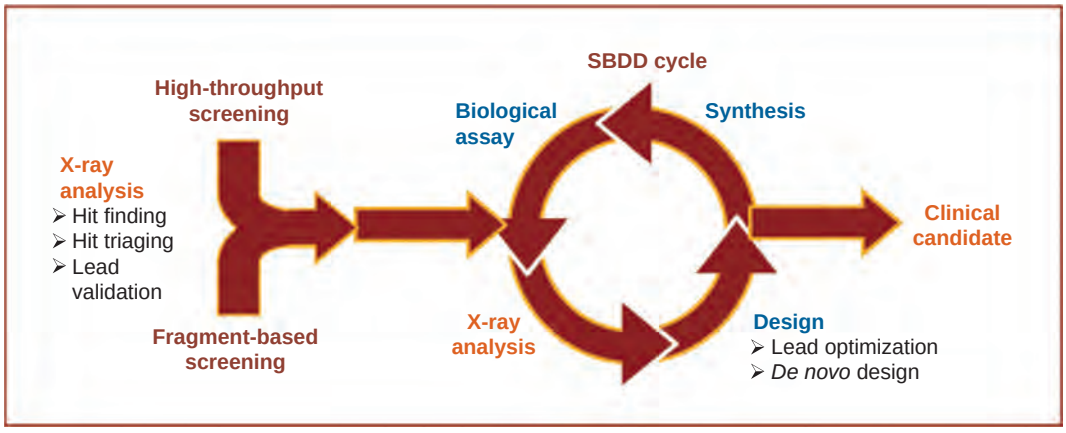
I. INTRODUCTION

Protein crystallography is the only technology to date that allows us to "see" how a ligand (hit, lead) is bound to its target protein. No wonder, therefore, that it has had a profound influence on pharmaceutical research since its inception during the 1970s [2]. Today, the use of structural information pervades all phases of pre-clinical research: target identification and validation, the development of *in vitro* assays, finding the best hit/lead finding strategies, and the entire lead optimization phase [3,4] (Figure 22.1a). The impact of structural biology on the daily work of medicinal chemists by replacing traditional trial and error methods by structure based design has been particularly strong [5,6].

However, exploiting this detailed 3D structural information is not trivial, in part because it must be combined with other constraints such as synthetic accessibility, absorption, distribution, metabolism, and excretion (ADME) properties, toxicology, and intellectual property. Nonetheless, the rewards are immense. Structural information not only clarifies structure–activity relationships, reveals binding modes and bioactive conformations, and



(a)



(b)

FIGURE 22.1 (a) Contributions of X-ray analysis to the drug discovery value chain; (b) Integration of X-ray analysis into hit/lead finding, triaging, validation, and optimization. SBDD: Structure-based drug design.

unveils new binding pockets or allosteric binding sites but also opens new and diverse drug discovery avenues, such as *in silico* screening, design of focused chemical libraries, and *de novo* design of new ligand scaffolds. Having access to such information provides a strong competitive advantage and makes the professional life of medicinal chemists highly stimulating and often very gratifying.

A hundred years after the recording of the first diffraction image, X-ray crystallography is still rapidly evolving with large scale structural genomics efforts focusing on human proteins of medical importance and on potential drug targets from clinically-relevant pathogens [7]. Advances in technologies and methods have led to a breakthrough in the crystallization of membrane proteins and a drastic reduction in the time needed to generate crystal structures. These developments not only produce a wealth of structural data but also enable high-throughput methods in protein crystallography [8] (e.g., hit triaging and *de novo* hit finding by fragment-based screening; Figure 22.1b), which allow the identification of novel, chemically attractive leads and their successful optimization to highly potent drug candidates [9].

In this chapter, we will describe how crystallographic data contribute to the different phases of pharmaceutical research. We will emphasize not only the strengths but also the technical limitations of protein crystallography, so that any medicinal chemist can gauge if and how a project could benefit from this technology. A brief outline of the basic principles and methods of protein crystallography is also provided. To make proper use of structural data, it is essential to be aware of the limitations and potential uncertainties associated with X-ray structures. We also hope that this chapter will contribute to more effective communication between chemists and their fellow crystallographers.

II. HISTORICAL BACKGROUND

A. The Early Days of Crystallography

Crystallography made its first notable contributions to the progress of biology and medicine well before the elucidation by Kendrew and Perutz in 1958–1960 of the first protein structures, myoglobin [10] and hemoglobin [11]. The preparation of "blood crystals"—in fact, hemoglobin crystals—was first reported by Hünefeld in 1840

[12]. During the second half of the nineteenth century, this initial observation sparked considerable interest in the crystallization of hemoglobins and other proteins, mainly from plant seeds [12]. This groundwork set the stage for the first major achievement in this field, which took place 1926–1935: the demonstration of the molecular nature of enzymes and viruses through their isolation in crystalline form by Sumner, Northrop, and Stanley [12]. The second major contribution was made in the early 1950s, when X-ray diffraction photographs of DNA produced by Franklin [13] and Wilkins [14] could be used as a guide by Crick and Watson, ultimately leading to their discovery of the double helical structure of DNA [15]. The next achievement—the determination of the myoglobin and hemoglobin structures by Kendrew and Perutz—revealed for the very first time the intricacies of the architecture of proteins, while also shedding light on the molecular basis of sickle cell anemia [16]. Since then, these milestone studies have been followed by a rich crop of other stunning crystallographic feats. The 3D structures of the human common cold virus [17], the photosynthetic reaction center [18], the F1-ATP synthase [19], the proteasome [20], the nucleosome core particle [21], the 30S and 50S ribosomal particles [22–25], the RNA polymerase II [26], potassium channels [27,28], and the β2 adrenergic G-protein-coupled receptor [29] have all been solved using X-ray diffraction methods, in spite of their daunting size and biochemical complexity.

B. The Current State-of-the-Art

Today, more than 100,000 crystal structures are publicly available from the Protein Data Bank (see Box 22.3), comprising about 40,000 unique sequences, a number that is rapidly increasing due to efforts in academia, structural genomics consortia, and pharmaceutical and biotech companies. This means that for most drug targets, either the structure itself or at least the structure of a homologous protein is available for use in structure-based drug discovery.

On-going developments in miniaturization and robotics allow an extensive screening for crystallization conditions with only a few milligrams of protein, while microfocus beamlines and ever brighter synchrotron sources allow the collection of data from smaller and smaller crystals. Automatic sample mounting and fast hybrid-pixel detectors allow fast data collection on large numbers of crystals, while software pipelines automatically process data [30], calculate electron-density maps, and do the initial ligand fitting [31,32]. Only at this stage does the crystallographer need to examine the structure and manually continue the fitting and refinement process.

X-ray crystallography is now routinely used in drug discovery projects involving soluble targets, and the same may soon be true for membrane proteins. However, a prerequisite for using this method is the availability of suitable crystals. For high-throughput methods such as hit triaging and fragment-based methods, the crystallization process should be robust and routinely produce well-diffracting crystals. This means that in the majority of cases, the main bottleneck does not reside in the X-ray analysis itself but rather in the identification and production of a stable, well-behaved recombinant version of the protein of interest, which is amenable to crystallization.

Today's crystallographers—particularly those working in industry—are often faced with difficult to produce, poorly behaved, poorly characterized targets. Such challenging proteins require a lot of biochemical ingenuity and cannot be conquered without a dedicated, appropriately resourced effort in protein production and characterization. For academic users, large facilities are available to generate and test thousands of clones for expression and solubility [33]. It would be helpful if industrial crystallographers could get access to such facilities as well.

C. Examples of Structure-Based Drug Discovery

1. Captopril

Drugs with anti-sickling properties were the first drugs ever to be studied in complex with their protein target by means of X-ray analysis [16]. These seminal studies spurred the first attempts at designing improved compounds via a structure-based approach [34]. Soon after, the discovery of captopril [35] (Figure 22.2a), an anti-hypertensive agent and the first marketed, orally active inhibitor of the human angiotensin-converting enzyme (ACE), hailed the beginning of a new era for pharmaceutical research. For the first time, a drug had been rationally designed on the basis of structural information, hence providing the first compelling demonstration of the power of the structure-based approach.

Interestingly, the successful design of captopril used simple chemical concepts guided by a hypothetical "paper-and-pencil" model of substrate and inhibitor binding to the enzyme active site that had been inferred from the crystal structure of bovine carboxypeptidase A. The X-ray structure of human ACE became available only in 2003, twenty-five years after the discovery of the captopril class of drugs. While the crystallographic analysis of the ACE complex with captopril [36] confirmed the designed mode of interaction, it revealed little structural similarity overall with carboxypeptidase A.

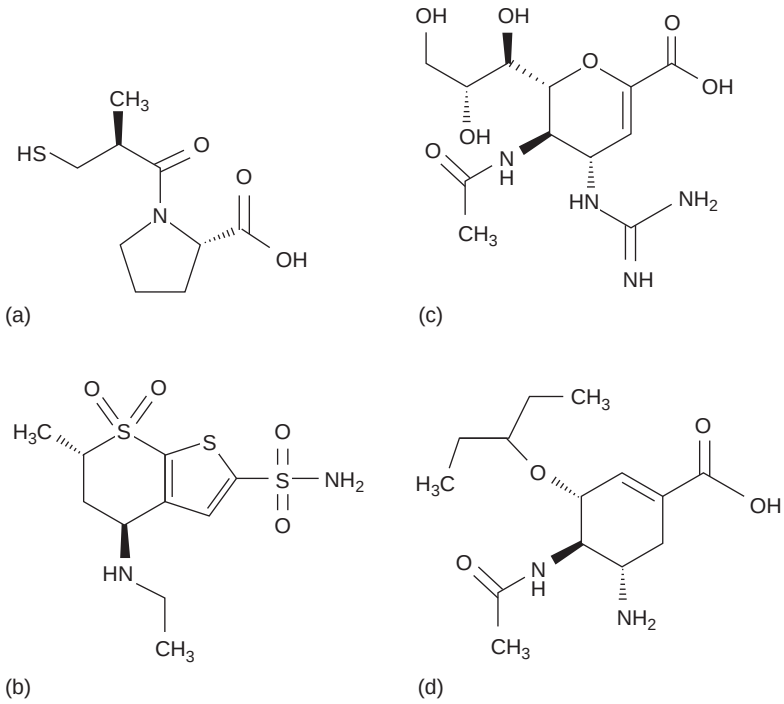


FIGURE 22.2 The first marketed drugs derived from structure-based design: (a) captopril (Capoten™); (b) dorzolamide (Trusopt™); (c) zanamivir (Relenza™); and (d) oseltamivir (Tamiflu™).

2. *Dorzolamide*

Dorzolamide (Figure 22.2b) is the first example of an approved drug that benefited from the complete armament of structure-based design (i.e., multiple X-ray analyses with the human enzyme target [37] combined with sophisticated molecular modeling studies, including in-depth conformational analyses using *ab initio* quantum chemistry calculations) [38]. Dorzolamide is a subnanomolar carbonic anhydrase II inhibitor that was developed in the early 1990s as a topical agent for the treatment of glaucoma.

3. *Relenza and Tamiflu*

The discovery of oseltamivir (Tamiflu™TM; Figure 22.2d), an inhibitor of *influenza* neuraminidase, is another early example of successful structure-based drug design. Interestingly, long standing efforts to identify neuraminidase inhibitors via random screening or the rational design of transition-state analogues had failed to produce any potent compounds [39] until the crystal structure of neuraminidase became available in 1983 [40,41]. A GRID [42] analysis of the sialic acid complex immediately suggested a simple modification of a known sialic acid analogue with low micromolar affinity. Remarkably, only two compounds were synthesized, and both turned out to be extremely potent inhibitors, with Ki values of 50 nM and 0.2 nM, respectively [43]. The most potent compound, zanamivir (Relenza™; Figure 22.2c), became the first marketed neuraminidase inhibitor. Further structure-based design concentrated on the development of an analogue with improved stability and lipophilicity. These efforts very quickly resulted in oseltamivir, a second generation, orally bioavailable drug [44] (Figure 22.2d).

III. BASIC PRINCIPLES AND METHODS OF PROTEIN CRYSTALLOGRAPHY

A. Crystallization

1. *What are Protein Crystals?*

Protein crystals (Figure 22.3), like any crystal of organic or inorganic compounds, are regular 3D arrays of identical molecules or molecular complexes (Figure 22.4). Depending on the symmetry of this arrangement (described by the space group), all molecules in a crystal have a limited number of unique orientations with respect to the crystal lattice. The diffraction of all individual molecules adds up to yield intensities that are sufficiently strong to be measured, the crystal lattice thus acting as an amplifier. An explanation of some common crystallographic terms is given in Box 22.1.

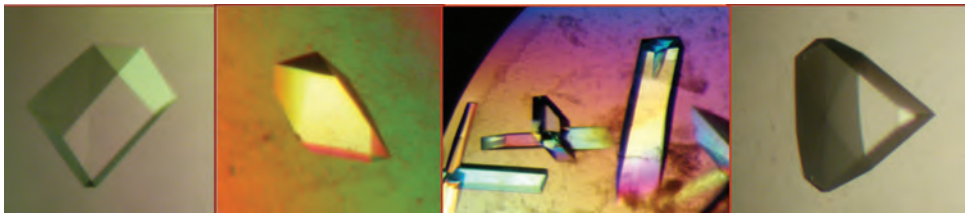


FIGURE 22.3 Examples of protein crystals. From left to right: β-secretase inhibitor complex; human farnesyl pyrophosphatase in complex with zoledronic acid; abl kinase domain in complex with imatinib (courtesy of SW Cowan-Jacob, Novartis); cdk2 inhibitor complex.

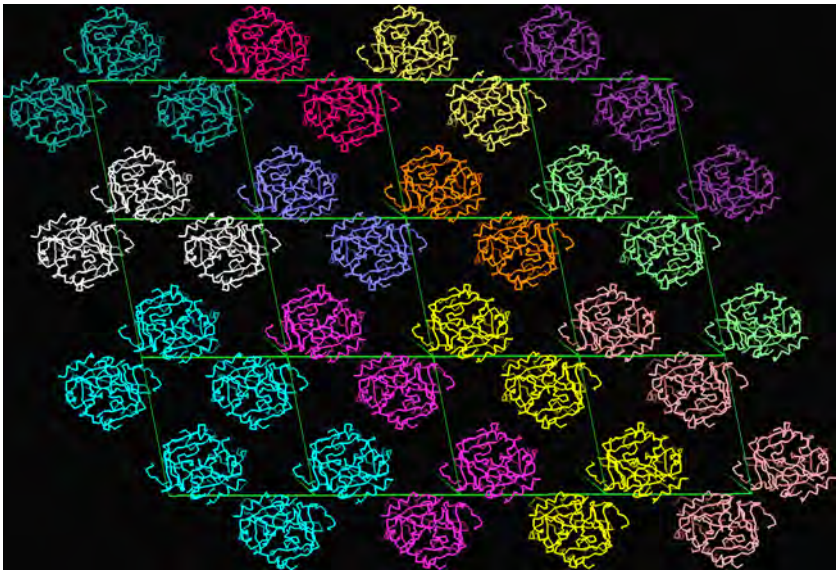


FIGURE 22.4 Crystal packing of a human thrombin complex. Twelve unit cells with one layer of molecules are shown. By looking carefully, one can see that the two molecules in each unit cell are rotated 180° with respect to each other. Protein crystals used for X-ray diffraction extend into three dimensions and consist of many layers of molecules. The next layer of thrombin molecules fits into the holes present in the layer shown.

BOX 22.1

SOME COMMON CRYSTALLOGRAPHIC TERMS

Space group: The group of symmetry operators that describe the symmetry of the crystal. Since biological molecules are optically active, their crystals belong to one of the sixty-five noncentrosymmetric space groups.

Unit cell: The basic building block of a crystal. The whole crystal can be generated by repeated unit translations of the cell in three dimensions. The unit cell is characterized by its axes a, b, c, and the angles (α, β, γ) between them.

Asymmetric unit: The smallest motif from which the whole unit cell can be generated by applying the symmetry operators of the space group. The asymmetric unit may contain one or more copies of the protein or complex under study. In the case of oligomeric particles, the asymmetric unit may contain one or more complete particles or only one or more subunits if some symmetry axes of the particle coincide with some symmetry axes of the crystal.

Reflection: A diffracted beam of X-ray, characterized by its indices h,k,l, and caused by reflection from the lattice planes making intercepts a/h, b/k, and c/l with the unit cell axes. Each reflection contains information on the entire structure. Reflections occurring at high scattering angles have high indices and carry high resolution information (i.e., they correspond to a fine sampling of the structure), while those observed close to the direction of the incident beam have small indices and carry low resolution information (i.e., they correspond to a coarse sampling of the structure).

Resolution: The resolution limit corresponds to the highest scattering angle at which reflections can still be measured (cf. Box 22.2). Individual atoms can be fully resolved when the resolution is better than 1.0Å.

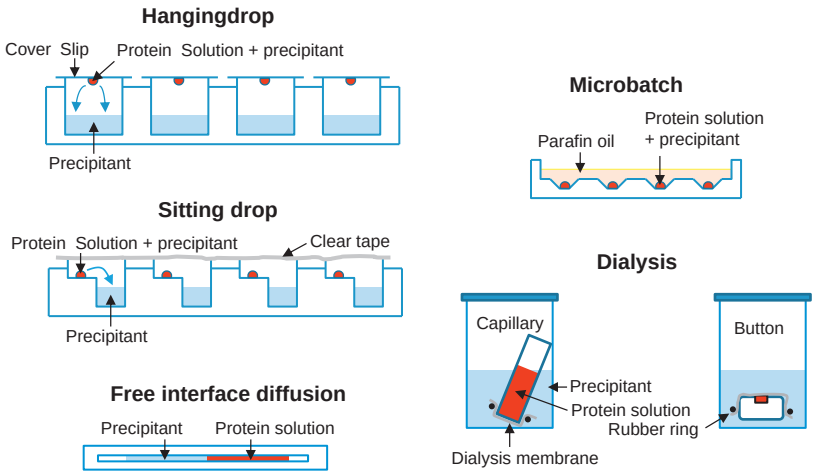


FIGURE 22.5 Methods used to crystallize proteins. For the hanging drop, sitting drop, and microbatch methods, a protein solution is mixed with a precipitant solution (usually in a 1:1 ratio) and set to equilibrate. With the hanging drop and sitting drop methods, water from the less concentrated protein drop will transfer to the more concentrated precipitant solution via the vapor phase, as indicated by blue arrows. This makes the protein drop shrink, thereby increasing the protein and precipitant concentration. With the other two methods, the protein solution is brought into contact with a precipitant solution either directly (free interface diffusion) or via a dialysis membrane (dialysis method). In all cases, adding precipitant to the protein solution creates a supersaturated state, causing to protein to precipitate and—if one is lucky—form crystals.

One notable difference between crystals of small molecules and macromolecular crystals is the very large solvent content of the latter. Protein crystals typically contain 30–80 percent (v/v) solvent (in fact, aqueous crystallization buffer) [45]. Only a fraction of the protein surface is involved in crystal contacts, the rest being fully solvated, pretty much as in solution. As a consequence, protein crystals are very soft and fragile. But on the positive side, low molecular weight ligands, co-factors, and substrates can diffuse from the surrounding mother liquor into the solvent channels in the crystal. If their binding site is not occluded by crystal contacts, the complex can be formed *in situ*. Usually, small conformational changes can take place within the crystal lattice without damaging the crystal, and sometimes very large structural changes can be accommodated as well. Therefore, enzyme crystals are very often active as catalysts.

2. *How do we Get Crystals?*

Crystals are produced by slowly driving a concentrated protein solution into a state of supersaturation [46,47]. Under the right conditions, the protein will not form an amorphous precipitate but will instead settle into a well-ordered crystalline array. Methods for achieving a high degree of supersaturation involve dialyzing away the salt, if the salt has a strong solubilizing (or "salting-in") effect, concentrating a nearly saturated protein solution by evaporation (usually in a hanging or sitting drop set-up; see below) and adding "precipitants" such as poly (ethylene glycol) or high salts such as ammonium sulfate, which has a strong "salting-out" effect on proteins. Other possibilities, which are less often used, are temperature and pH gradients. The methods most often used for crystallization screening are shown in Figure 22.5.

Because of the low protein consumption (100 nl protein solution per drop) and compatibility with crystallization robots and automatic crystal imagers, most crystallization experiments are performed as sitting drop experiments in 96-well microtiter plates [48,49]. Manual crystallizations are usually done in 24-well microtiter plates using the hanging drop set-up with 1 μl protein solution per drop. Other crystallization techniques have been developed, such as crystallization under oil (microbatch) and free interface diffusion using capillaries or microfluidic chips [46,47,49]. Microfluidic methods are also used for determining phase diagrams.

3. *Specific Problems and Solutions*

Crystallization conditions are published. When crystallization conditions are already known, 1 mg of protein may be enough to produce a series of crystals with different inhibitors. One should bear in mind, however, that published crystallization protocols are often difficult to reproduce. It is wise, in a first step, to follow as closely as possible the published expression, purification, and crystallization protocols. Particular attention should be paid to the protein construct, since minor changes to the amino-acid sequence can have a dramatic influence on the solubility, stability, and crystallization behavior.

De novo crystallization. Obtaining X-ray quality crystals is usually the most difficult and time-consuming step of a new structure determination project, notably in the case of a novel, poorly characterized gene product [49]. Modern crystallization robots and miniaturization have considerably simplified the process and shortened the time needed to set up extensive crystallization screening experiments, while dramatically reducing the amount of material needed. By using a protein solution of 10 mg/ml and 100 nl protein solution per drop, 0.1 mg of protein is sufficient to screen 100 conditions, and with a few mg of protein, one can screen several thousand individual

crystallization conditions. Nevertheless, this is often not enough to obtain suitable crystals for new and difficult targets [49] since the protein (construct) may simply not be crystallizable. Several protein constructs or protein variants may have to be generated to increase the chances of finding one that is amenable to crystallization. For this reason, strong and dedicated support in molecular biology, protein expression, and biochemistry are an absolute must for a successful protein crystallography laboratory. Furthermore, in an industrial setting, these activities should be initiated as early as possible, to ensure that crystals are available before chemistry activities are started.

When novel protein targets cannot be produced or crystallized, one should consider the possibility of using a known homolog or anti-target instead. If the binding sites are sufficiently similar, the binding modes of some key compounds or scaffolds can be deciphered, and this information can be fed into the drug-design process.

Conformational heterogeneity. Recombinant proteins designed for assay purposes are often not suitable for crystallization experiments when they contain fusion partners, long tags, floppy ends, disordered or intrinsically unstructured regions, or loosely linked domains [50] that in general prevent crystallization. Intrinsically unstructured regions can usually be identified from the amino-acid sequence as polypeptide segments with low sequence diversity [51]. Domain boundaries can be pinned down by limited proteolysis or with the help of homology modeling, and constructs can be made expressing only a single domain. If tags or fusion partners are needed for enhanced expression and/or ease of purification, then a protease recognition sequence should be engineered to allow their removal before crystallization. Conformational heterogeneity can also be reduced by buffer additives or ligands [52]. Biophysical techniques such as thermal-shift assays and nuclear magnetic resonance (NMR) can be used for the identification of suitable ligands and additives.

Glycosylation. Glycosylated proteins often give poorly diffracting crystals. To circumvent this problem, several strategies can be used: [53] glycosylation can be chopped off enzymatically with PNGase F or other endoglycosidases; glycosylation sites can be mutated away; or a non-glycosylated form can be produced using a prokaryotic expression system. Glycosylation also increases the solubility and stability of proteins, so crystallization of the glycosylated protein should be tried as well.

Proteolytic cleavage. Proteases cleave other proteins, including themselves, and even trace amounts of a contaminating protease may wreak havoc during the time it takes (days to weeks) for crystals to grow. When crystallizing a protease for the first time, it is always a good idea to add the most potent inhibitor available to the crystallization set-up. If a contaminating protease is a problem, one could add a protease inhibitor cocktail or a general broad-spectrum inhibitor like PMSF.

Phosphorylation. Protein kinases and other proteins involved in signaling are often produced from eukaryotic expression hosts as a mixture of inactive (unphosphorylated) and active (with one or more phosphorylations) species, which causes conformational heterogeneity. One possible workaround is mutating the phosphorylation site (s) to glutamate, which mimics the phosphorylation, thus producing a constitutionally active kinase. Other options include mutating the phosphorylation site away or expression in the presence of a ligand or inhibitor [52,54]. Co-expression with a phosphatase is another strategy that may reduce the heterogeneity of the phosphorylation [55].

Membrane proteins The crystallization of membrane proteins is particularly challenging [56] and has long been off-limits for industrial crystallographers. Traditionally, only the soluble domains (catalytic, ligand binding) of multidomain membrane proteins were expressed and crystallized [57]. In recent years, however, the structure of a large number of integral membrane proteins has been published [58], and crystallography of membrane proteins is entering drug discovery laboratories.

Factors that have enabled this breakthrough are improved expression and purification of membrane proteins, optimized detergents, robotics and miniaturization that allow extensive screening of crystallization conditions using very little protein, and microfocus beam lines at synchrotrons that allow the collection of data from tiny crystals. Also new water-detergent phases such as bicelles [59] and lipidic cubic phases [60] have had a large impact on the crystallizing success of membrane proteins. However, the greatest breakthrough came from overcoming the inherent flexibility of many membrane proteins, especially G-protein-coupled receptors (GPCRs).

Most GPCRs such as the β_2-adrenergic receptor (β_2AR) are present as a mixture of conformations ranging from the inactive to the active conformation [61] (Figure 22.6). Adding an agonist merely shifts the equilibrium toward the active conformation, but only binding of a G-protein stabilizes the active state. For crystallization, GPCRs are stabilized by the removal of flexible loops, replacing the flexible intracellular loop three by a stable and soluble protein like T4 lysozyme, the addition of antibodies (Fab fragments or single-chain camelid antibody fragments), and systematic scanning mutagenesis or random evolutionary mutagenesis [58]. In general, these methods stabilize either the active or the inactive conformation, so for studying agonists one would need different antibodies or a different construct than for studying antagonists [62]. At the end of this chapter, the active and inactive conformations of β_2AR will be compared and discussed.

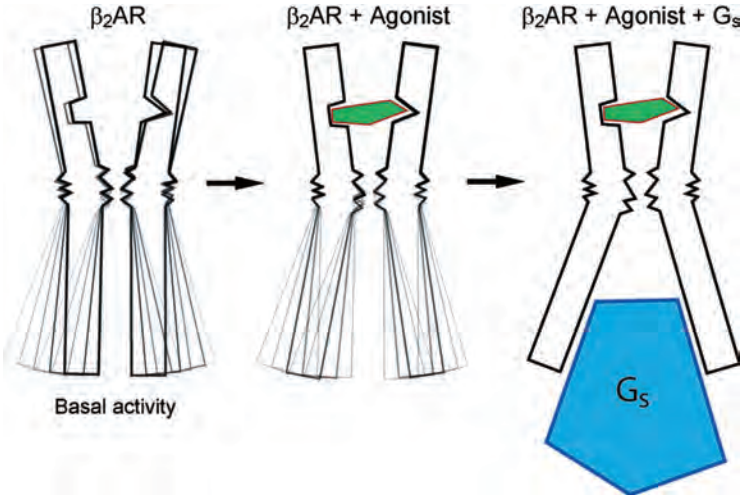


FIGURE 22.6 Cartoon illustrating the dynamic character of the β_2AR by Brian Kobilka [61]. Both the free and the agonist-bound form show a range of conformations. Only binding of the G-protein G_s fully stabilizes the receptor. © *The Nobel Foundation 2012.*

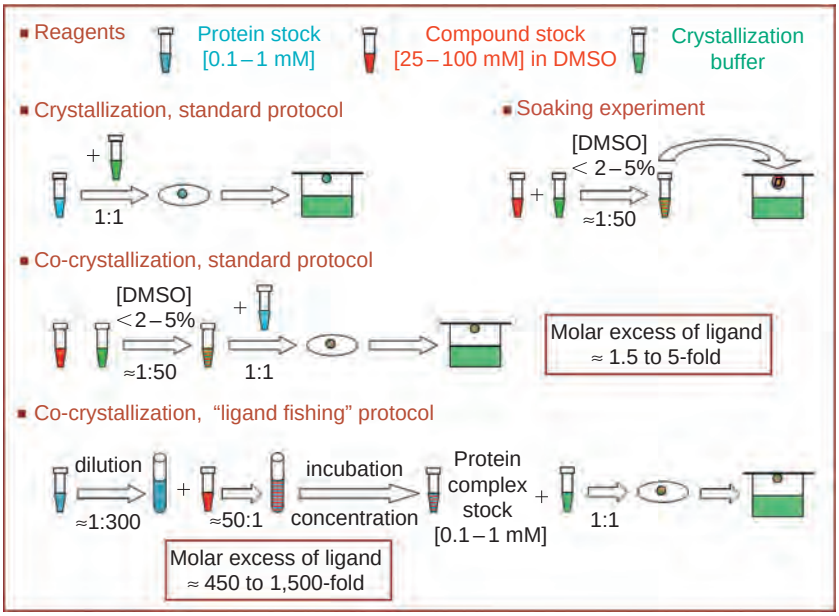


FIGURE 22.7 Protocols for preparing crystals of protein–ligand complexes.

4. *Preparation of Protein–Ligand Complexes*

An important aspect of protein crystallography in the context of drug design concerns the determination of protein–ligand complexes [63]. In many cases, getting crystals of protein–ligand complexes is not trivial, even if the apo protein has been crystallized. The ligand binding site may be blocked by crystal contacts, the ligand may not be soluble in the crystallization buffer, and the ligand may cause conformational changes in the protein that are incompatible with the crystal packing. Several methods exist to obtain crystals of protein–ligand complexes:

Soaking. This is the fastest and easiest way. The ligand is simply added to preformed crystals. If the ligand is a relatively small molecule, the solvent channels in protein crystals are usually large enough to allow the diffusion of the ligand to its binding site. A soaking experiment requires little material (one micromole of compound is usually plenty), but the solubility of the compound under the crystallization conditions is often an issue. The high protein content of the crystallization drop usually requires ligand concentrations in the range of 0.5–5.0 mM. A typical soaking protocol involves the preparation of a concentrated (50–100 mM) stock solution of the ligand, usually in a suitable organic solvent such as dimethylsulfoxide. This solution is then mixed with a crystallization buffer to a final concentration of solvent of up to 5 percent, and a few microliters of this mixture are added to the crystallization drop (Figure 22.7). Compound purity or the use of a diastereomeric mixture may

not be an issue if only one component in the sample binds to the protein. However, the chemical structure of the ligand must be known, since ultra-high resolution would otherwise be required for the unambiguous identification of an unknown binder.

Soaking has some practical advantages, particularly when the crystallization is not robust and good crystals are difficult to prepare. Overnight soaking is usually sufficient, allows a fast feed-back to modeling and chemistry, and opens up the possibility of using X-ray analyses for hit triaging and validation as well as for fragment-based screening. However, the soaking method also has drawbacks. Conformational changes induced by ligand binding may be hindered in the crystal, or access to the binding site may be restricted by protein–protein contacts. As a result, the ligand may not bind at all, or it may adopt an artificial mode of binding. Moreover, the diffraction quality can sometimes suffer from the soaking procedure, or the crystals may even crack or dissolve upon soaking. In such cases, gentle cross-linking of the crystals using the method of Lusty [64] may prove useful, but validation of the soaking approach with a co-crystallization experiment would then be worthwhile.

Co-crystallization. With this method, the ligand is added to the protein in solution, which is subsequently crystallized. When the crystallization is reasonably fast and robust, this is the method of choice and is recommended even in cases where soaking would be possible. The risk of artifacts is minimized, but at the expense of speed, particularly when the crystallization is very slow. A further disadvantage is that, for each and every new complex, crystallization conditions may have to be optimized again, or a full crystallization screening may be required, since crystallization conditions are sometimes very sensitive to changes in the ligand. Seeding is frequently used to accelerate co-crystallization experiments and improve their reproducibility. With modern robotics, this can be done automatically using very little protein [65].

Ligand fishing. Biological assays are usually performed in the presence of a large excess of ligand. This is particularly true for weak ligands, which are assayed at concentrations in the micromolar range while the protein concentration is typically in the nanomolar or sub-nanomolar range. Medicinal chemists should always bear in mind that weak biological activity may sometimes be due to trace amounts of a highly potent compound "contaminating" an otherwise inactive sample. For instance, an IC_{50} of 10 μM could be due to 0.5 percent of an impurity with a potency of 50 nM. This situation is not uncommon in programs where inactive derivatives are sometimes obtained from very potent precursors. Trace impurities of 1.0 percent or less are usually not detected by routine analytical techniques, but may give rise to apparent micromolar activity. These impurities will not be detected by crystallography either, if only a small excess of compound (2 to 5-fold) is added to a concentrated protein aliquot, as is usually the case (Figure 22.7). But if a very large excess of compound (say 500 to 1,500-fold) is added to a diluted sample of protein, there may be enough active impurity to saturate or nearly saturate the protein. The complex can then be concentrated using standard ultrafiltration techniques and crystallization experiments performed. This procedure, which is often referred to as "ligand fishing" (Figure 22.7), is more time-consuming and requires larger amounts of compound (5–10 mgs), but it has the ability to detect very low amounts (down to approximately 0.1 percent) of a potent ligand in a mixture. It may prove useful in cases where a weakly active compound whose structure is at odds with the established structure–activity data could not be observed using the routine crystallization procedure. In addition, this protocol can also be used when ligands are very poorly soluble in the crystallization buffer.

B. Data Collection

Protein crystals contain on average 50 percent solvent and—when exposed to air—they dry out and disintegrate. Moreover, when exposed to high-intensity X-rays at room temperature, they lose their diffraction power very quickly, owing to radiation damage. In order to prolong crystal lifetimes and improve data quality, X-ray measurements are routinely performed at 100 K [66]. Crystals are first mounted on 10–20 μm thin nylon loops and then flash-frozen by immersion into liquid nitrogen. To prevent the formation of ice crystals, it is often necessary to add a cryo-protectant such as glycerol, low molecular weight poly(ethylene glycol), or high salt to the surrounding mother liquor.

For data collection, the crystal is then placed on a goniometer, a device that controls the rotation of the crystal in the X-ray beam, while the temperature is kept at 100 K by blowing dry nitrogen over the crystal (Figure 22.8b). Large, strongly diffracting crystals can be measured in the lab with a rotating anode X-ray generator, but tiny or weakly diffracting crystals must be measured at a synchrotron source, such as the Swiss Light Source (SLS) in Villigen, Switzerland, or the European Synchrotron Radiation Facility (ESRF) in Grenoble, France (Figure 22.8a). For industrial projects, frozen crystals are often sent to the synchrotron by courier service in special cryo-containers. Data collection is remotely controlled by the scientists from their home laboratory or is done by a scientist at the synchrotron for a service fee.

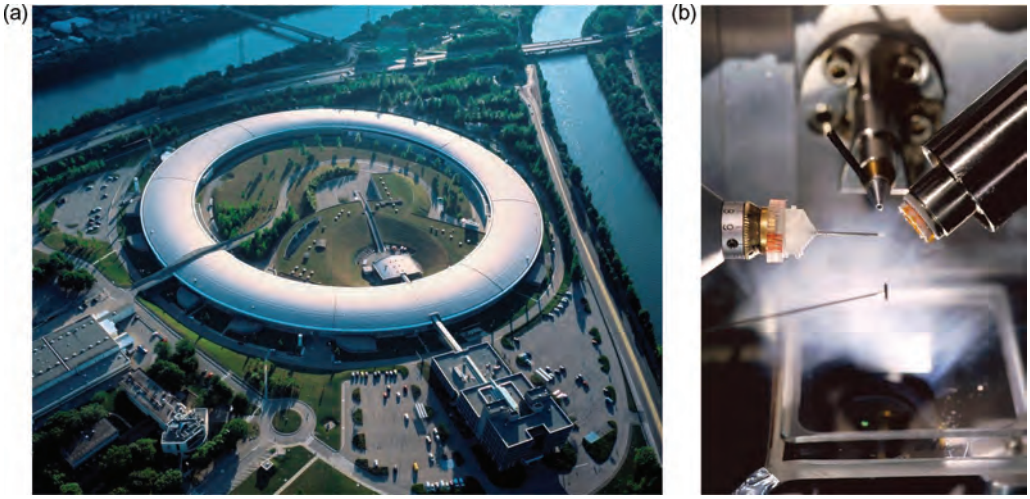


FIGURE 22.8 (a) Aerial view of the ESRF, located between the rivers Isère and Drac in Grenoble, France. Electrons circle around in a large ring inside the circular building, and when they pass bending magnets or assemblies of magnets called undulators they emit powerful X-rays. (b) Crystal being exposed, viewed from the position of the detector. X-rays emanating from the narrow steel tube in the back hit a frozen crystal in the cryo-loop in the center of the picture. The direct beam is stopped by a beam-stop, the small piece of metal just below the center of the picture. On the left is the goniometer, which is used to rotate the crystal, and from the nozzle on the right cold (100 k) nitrogen gas is blown over the crystal. *Pictures courtesy of ESRF/Morel.*

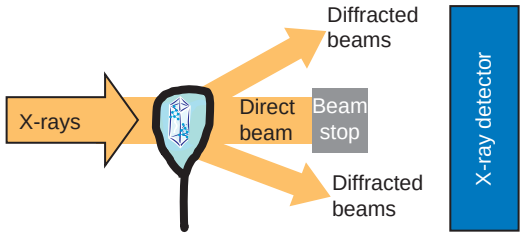


FIGURE 22.9 Schematic picture of an X-ray diffraction experiment. A real-life set-up is shown in Figure 22.8b. Monochromatic X-rays coming from the left hit the crystal—which is usually not longer than 100 μm—and some X-rays are diffracted. Most X-rays pass straight through and are stopped by a small piece of lead, the beam-stop. The diffracted X-rays are detected by a two-dimensional X-ray detector.

During data collection, the crystal is slowly rotated to bring all reflections into diffracting condition (see Box 22.2 about Bragg's law). The diffraction spots are usually recorded by CCD or hybrid-pixel detectors (Figure 22.9). The time to collect complete high quality X-ray data sets from single crystals ranges from a few minutes for decent crystals and high-intensity synchrotron radiation to a few days for weakly diffracting crystals and a conventional X-ray generator. The diffraction images from these detectors (Figure 22.10) are fed directly into a computer, which produces a list of reflection intensities. Ten thousand to several hundred thousand reflections are recorded per crystal, depending on the quality of the crystal and the size of the unit cell.

C. From Diffraction Intensities to a Molecular Structure

Light Microscopy and X-ray Crystallography Share the same Basic principle.

A light microscope allows us to study small objects like insects or cell slices in great detail, but it is physically impossible to resolve any details that are smaller than half the wavelength of the light used. For blue light, this limit is about 200 nm. To resolve atomic details, which are on the order of 1–5Å (0.1–0.5 nm), electromagnetic radiation with a much shorter wavelength than light is required (i.e., X-rays). A light microscope and an X-ray set-up share the same basic principle, although the practical implementation is quite different, owing to the different properties of X-rays and visible light.

In a microscope, light from a light source shines on the sample and is scattered in all directions. A set of lenses is used to reconstruct from this scattered light an enlarged image of the original sample. In an X-ray experiment,

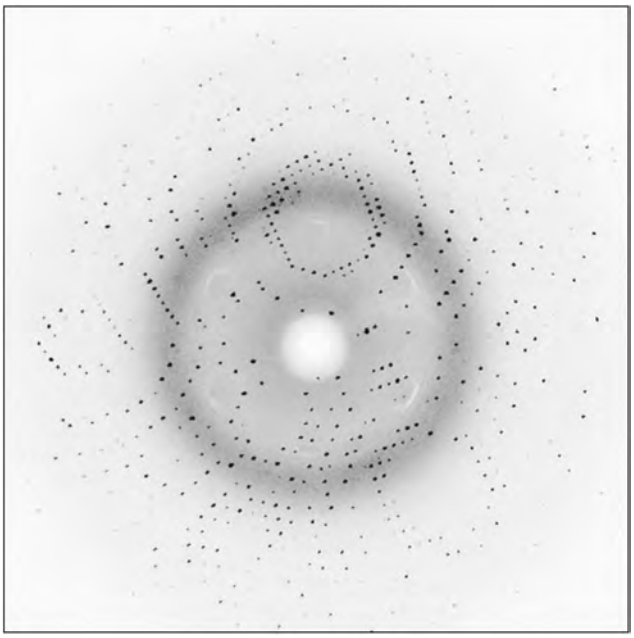
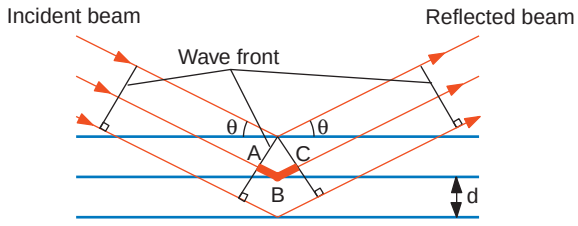


FIGURE 22.10 Example of an X-ray diffraction image.

BOX 22.2

BRAGGS LAW



Observable diffraction is only obtained when waves reflected from adjacent planes reinforce each other, that is, when the path difference ABC in the figure is an integer multiple n of the wavelength λ of the X-ray radiation used. This leads to Bragg's equation:

$$n\lambda = 2d \sin \theta$$

This means that to resolve closely spaced planes (small d), we need to measure high angle (large θ) reflections.

X-rays from an X-ray source hit the crystal and are scattered in all directions, just as with the light microscope. Unfortunately, no lenses can be made which are able to bring the scattered X-rays into focus to reconstruct an enlarged image of the sample. All the crystallographer can do is to record directly the scattered X-rays (the diffraction pattern; see Figure 22.10) and to use computers to reconstruct an enlarged image of the sample.

1. X-rays are Scattered by Electrons

Although X-rays interact only weakly with matter, they are occasionally absorbed by electrons, which start to oscillate. These oscillating electrons serve as X-ray sources that can radiate the X-ray wave in any direction. Waves scattered from different parts of the crystal have to add up constructively in order to produce a measurable intensity. The condition under which the scattered X-rays add up constructively is laid down in Bragg's law, which treats crystals in terms of sets of parallel planes (Box 22.2).

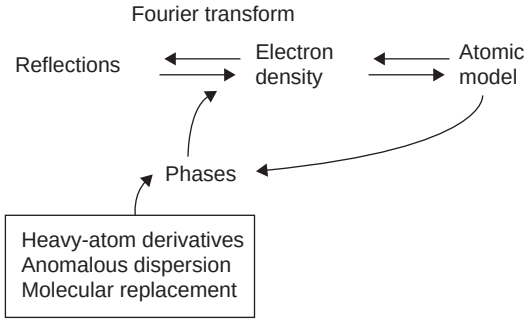


FIGURE 22.11 The phase problem. The experimental data obtained in an X-ray experiment are the intensities of the reflections. By using an inverse Fourier transform, it is possible to calculate electron-density maps from the amplitudes derived from these intensities. However, it is essential for this calculation to know the phase associated with each reflection. Approximate initial phases can be obtained from heavy-atom derivatives, anomalous dispersion, or molecular replacement (see text). More accurate phases can be derived from the refined model, once it has been obtained.

2. *The Diffraction Pattern Corresponds to the Fourier Transform of the Crystal Structure*

Each diffraction spot is caused by reflection of X-rays by a particular set of planes in the crystal. If the crystal contains layers of atoms with the same spacing and orientation as a particular set of planes that would satisfy Bragg's law (if the set of planes is physically present), the corresponding diffraction spot will be strong. On the other hand, if only few atoms in a crystal correspond to a particular set of planes, the corresponding reflection will be weak. The complicated structure present in the crystal is transformed by the diffraction process into a set of diffraction spots that correspond to sets of planes (more precisely, sinusoidal density waves), just as our ear converts a complicated sound signal into a series of (sinusoidal) tones when we listen to music. This conversion of a complicated function into a series of simple sine and cosine functions is called a Fourier transformation.

3. *The Phase Problem*

The original function—in our case the electron-density distribution in the crystal—can be reconstructed by performing the inverse Fourier transformation (i.e., by summing together the corresponding density waves for all reflections; see Figure 22.11). However, in order to make this summation, we need to know not only the amplitude of the density wave but also its relative position with respect to all other density waves (the phase). The amplitude, usually referred to as structure factor or F, can be measured because it is calculated from the intensity of the corresponding diffraction spot, but there is currently no practical way to measure the phases directly. This so-called "phase problem" can be solved by one of the following techniques:

Multiple isomorphous replacement (MIR). Crystals are soaked in solutions with "heavy" atom salts (Hg, Pt, Au, etc.), in the hope that a few heavy atoms will bind to some well-defined sites on the protein molecule. The heavy-atom positions are then found by analyzing the differences between the diffraction pattern of the native and of the soaked crystals. When two or more suitable heavy-atom derivatives are found, phase estimates and an electron-density map can be calculated.

Anomalous scattering (AS). This method makes use of the fact that some inner electrons of the heavier elements have absorption edges in the range of X-ray wavelengths. The method is used to supplement the phase information of a single heavy-atom derivative [67], but also to obtain full phase information from proteins which are labeled with selenomethionine, a selenium-containing amino acid [68]. This method (called "MAD" for multiple wavelength anomalous dispersion) has become the preferred method for the fast structure determination of novel proteins. Other anomalous methods have recently been proposed. For instance, the "halide-soak" approach uses short soaks in solutions containing 0.5−1.0M bromine or iodine, and the anomalous signal of the bound halide ions is then exploited to solve the structure [69]. For well-diffracting crystals, it is also possible to use the sulfur anomalous signal from the cysteines and methionines present in the native protein [70].

The MAD method is performed on a single crystal, but it requires access to tunable radiation (synchrotron source). Moreover, selenomethionine-labeled protein must be produced, purified, and crystallized. This is more easily done for proteins which can be expressed in *E. coli*.

Molecular replacement. When a suitable model of the unknown crystal structure is available, it can be used to solve the phase problem [71]. Examples are the use of the structure of human thrombin to solve the structure of bovine thrombin, the use of a known antibody fragment to solve the structure of an unknown antibody, or the

use of the structure of an enzyme to solve the structure of an inhibitor complex of the same enzyme in a different crystal form. The model is oriented and positioned in the unit cell of the unknown crystal with the use of rotation and translation functions, and the oriented model is subsequently used to calculate phases and an electron-density map.

Molecular replacement is usually straightforward and performed within minutes. However, when only low resolution data and a poor search model are available, model bias can become an issue and experimental phasing may be needed. Nevertheless, if a suitable model is present, which is increasingly likely given the ever increasing number of crystal structures available, molecular replacement is the method of choice to solve the phase problem.

4. Model Building and Refinement

Once a first electron-density map is obtained, it is interpreted by the crystallographer. In the case of a MIR (AS) map, a complete model of the protein has to be fitted to the electron density. The Cα atoms are placed first (chain tracing), and subsequently the complete main-chain and side-chains are built, a process that has become increasingly automated in recent years, particularly when high resolution data are available [72]. In the case of molecular replacement, the search model needs to be updated to reflect the molecule present in the crystal. The model is usually of a similar protein, and the possible changes include the substitution of some amino acids, the introduction of insertions and deletions, and the modification of some loops.

After the (re)building step, the model is refined. Refinement is an iterative procedure that aims at minimizing the differences between the observed diffraction amplitudes (F_o) and the diffraction amplitudes calculated from the model (F_c), while simultaneously optimizing the geometry of the structure. Because of the unfavorable ratio between observations and parameters, a free atom refinement is not possible in protein crystallography, and it is necessary to restrain the bond lengths, valence angles, and dihedral angles toward ideal values. Phases calculated from the refined model at the end of each refinement cycle are then used for the calculation of improved electron-density maps, which are again analyzed by the crystallographer to improve the model further. Cycles of refinement and rebuilding are repeated until convergence is reached. The final set of coordinates is then ready for deposition with the Protein Data Bank (PDB) [73].

5. Most Used Types of Electron-Density Maps

The direct experimental result of a crystallographic analysis is an electron-density map, while the model is derived from a (subjective) interpretation of this map. It is therefore useful to refer to the original data—the electron density—as often as possible. In the following paragraph, we will discuss the different types of electron-density maps most commonly used.

F_o-F_c or difference maps. These maps are obtained after subtracting the calculated structure factors (F_c) from the observed structure factors (F_o), an operation that is—in a first approximation—equivalent to subtracting the calculated electron density from the observed electron density. Features that are present in the "observed" density but not in the calculated density will give peaks, while atoms present in the model (in the F_c), but not in the "observed" electron density will result in holes (Figure 22.12). These maps are frequently used to detect errors in the model and can also be used to obtain an unbiased electron density of a bound inhibitor, for example, by completely removing the inhibitor from the model. In this case, the resulting electron density for the inhibitor is entirely caused by the experimental data and not by any model bias present in the phases. These maps are often referred to as "omit maps."

$2F_o$-F_c maps. These are the standard electron-density maps (Figure 22.12). Because of model bias, maps calculated with F_o and model phases tend to show only electron density associated with the model. As discussed above, F_o-F_c maps show everything that is in F_o but not in the model. By combining a F_o map with a F_o-F_c map, a $2F_o$-F_c electron-density map is obtained, which shows both electron density for the model and electron density for features that are not yet accounted for in the model, such as bound water molecules, carbohydrates, and other molecules associated with the protein. Several weighting schemes exist to minimize model bias. Examples are σ_A [74] and maximum-likelihood [75] weighting.

D. Information Content and Limitations of Protein Crystal Structures

Most chemists are familiar with X-ray analyses of small molecules, which are typically performed at a resolution better than 0.80Å. These subatomic resolution studies deliver highly accurate geometric parameters (bond lengths,

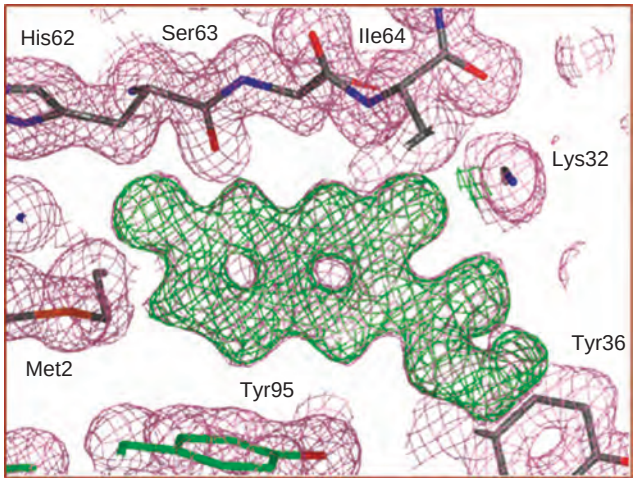


FIGURE 22.12 Close-up view of a protein complex at 1.5Å resolution showing the initial $2F_o$-F_c electron-density map (magenta mesh, 1.0σ contour), as well as the initial F_o-F_c map (green mesh, 3.0σ contour). The ligand has not yet been included in the model. Therefore, it appears as a strong positive difference density in the initial F_o-F_c map.

valence, and dihedral angles), as well as anisotropic displacement parameters ("temperature ellipsoids"). This is made possible by the very favorable observation to parameter ration (typically 50:1) resulting from the ultra-high resolution. Usually, protein crystals do not diffract to atomic or subatomic resolution. The vast majority of protein X-ray studies are performed at much lower resolution (between 3.0Å and 1.50Å), where this level of structural detail is not attainable. In particular, stereochemical parameters such as bond lengths and angles are restrained to standard dictionary values, both for the protein part and any low molecular weight ligand(s), prosthetic group, or post-translational modification. The protonation state and the exact orientation of some amino-acid side-chains (His, Asn, Gln) can only be inferred from potential H-bonded interactions. Substantially fewer solvent molecules and alternate conformations are observed than in the case of ultra-high resolution studies [76,77].

1. *Quality of the Experimental Data*

The quality of a crystal structure cannot be better than the quality of the experimental data upon which it is based. The following criteria are commonly used statistical indicators of the quality of the diffraction data:

Resolution. This corresponds to the shortest spacing of planes (d) whose reflections have been used in map calculation and refinement (see Box 22.2). The smaller this spacing, the sharper and more detailed the electron-density maps will be. The resolution is probably the single most important criterion determining the quality of a crystal structure. At high resolution (better than 2.0Å), the protein and bound water molecules are well defined, and it is unlikely that the structure will contain any serious errors. At low resolution (2.8−3.5Å), it is usually not possible to assign bound waters with certainty, and significant errors can remain unnoticed due to the problem of model bias.

Completeness of the data. One can calculate the total number of reflections to a certain resolution. Ideally, one would like to measure them all. For various reasons, however, it is in practice often not possible to measure all reflections. If only a small fraction of the reflections is missing (~10 percent), and the missing reflections are weak, the electron-density maps will hardly be affected. However, if a significant fraction of the reflections is missing, this may lead to artifacts in the electron-density maps, and the problem of model bias will become more severe.

R_{sym}. This reflects the inconsistency of multiple measurements of the same reflection. The lower the R_{sym}, the better. R_{syms} up to 15 percent are tolerable. Although the PDB still uses R_{sym}, other measures such as R_{meas} are more appropriate [78].

2. *Quality of the Model*

The global quality indicators listed below are commonly reported for refined crystal structures, but many more exist:

R-factor. This is a measure of the disagreement between the observed amplitudes (F_o) and the amplitudes calculated from the model (F_c). Depending on the resolution and quality of the diffraction data, well-refined structures have *R*-factors below 20−25 percent.

Free R-factor. Since refinement programs aim at minimizing the difference between observed and calculated amplitudes (hence the *R*-factor), an unbiased indicator is needed to monitor the progress of refinement. Brünger proposed excluding a subset of reflections from refinement and using these reflections only for the calculation of a "free" *R*-factor [79]. If refinement is progressing correctly, the free *R*-factor will drop as well. But if the model contains serious errors, it will remain stalled above ~35 percent. For correct structures, the free *R*-factor is generally below 30 percent.

Deviations from ideality of bond lengths and bond angles. A correctly fitted model is generally not strained. Significant deviations from ideal values for bond lengths and bond angles usually point to problems with the structure. Root-mean-square (r.m.s.) deviations from ideality should not be much larger than 0.02Å for bond lengths, and 3° for bond angles. The bond lengths and angles are biased toward the target values that are used during refinement. Accurate, unbiased values for these parameters can only be derived when ultra-high resolution (0.85Å or better) is available.

φ,ψ plot. Because of steric hindrance, only certain combinations of the main-chain dihedral angles φ and ψ are "allowed." The protein fold may force some residues to assume unallowed φ,ψ values, and this may have functional significance for some active site residues [80,81]. However, if more than a few percent of all the residues have φ,ψ values completely outside allowed regions, one should suspect errors.

3. Errors in Crystal Structures

Serious errors in crystal structures are rare and are usually associated with the first structure determination of a novel target, particularly when only low resolution data are available (3.0−5.0Å). Small errors and inaccuracies, however, are very common and virtually unavoidable. These errors are often underestimated, and small details of crystal structures are frequently overinterpreted by noncrystallographers. Medicinal chemists making use of crystal structures should be well aware of their limitations [82].

A major source of errors in macromolecular crystallography results from our inability to detect and model "disorder" appropriately [83], owing to the limited resolution and unfavorable parameter-to-observation ratio. Crystallographic refinement often attempts to fit a single model to some blurred electron density originating from several distinct but overlapping conformational states. This may lead to distorted geometry or to several distinct but equally valid interpretations.

A second important source of errors results from the fact that hydrogen atoms cannot be detected and atom types cannot be assigned at the resolution that is typically attainable with most protein crystals (1.5−3.0Å). This leads to ambiguities in the exact orientation of some groups, such as the side-chain amide of Asn and Gln residues or the imidazole ring of histidine side-chains.

Errors affecting the ligand. The exact orientation of one or more ligand groups can sometimes be uncertain. The choice of sensible geometric restraints for the refinement of nonstandard groups—in particular the ligands—is not always trivial and constitutes a potential source of errors [84]. For instance, the nitrogen atom of a tertiary amine bearing one aromatic substituent is usually planar, but it can also be pyramidal. At high resolution (better than 2.0Å), it may be possible to select the appropriate geometric restraints on the basis of the electron density. At lower resolutions, the refined model may mainly reflect the arbitrary choice of geometric restraints.

Errors affecting the solvent model. Water molecules are usually identified on the basis of residual electron-density peaks that meet certain criteria, such as the peak height, the distance, and the angle with respect to H-bond donor or acceptor groups. Since atom types and protonation states cannot be determined, a "water" may as well be a hydroxide, a hydroxonium, an ammonium, a sodium, or a magnesium ion. The assignment of metal ions becomes more reliable when the resolution of the data is good enough to reveal the coordination sphere or when the anomalous signal of the metal can be used.

4. Flexibility and Temperature Factors

Proteins are flexible molecules [85], and they usually retain a substantial degree of flexibility in the crystalline state. The mobility of the atoms in a crystal is expressed in terms of "temperature factors" or "B-factors," which are optimized during refinement. The relationship between mean total displacement and B-factors is given in Figure 22.13. The mean displacement of atoms with B-factors in excess of 60Å^2 is larger than 1.5Å, which is the length of a carbon−carbon bond. These atoms are generally poorly defined in the electron-density maps (Figure 22.14). For functional analysis, one should bear in mind that these flexible surface residues are either put in an arbitrary, low-energy conformation or deleted from the coordinate file. Not taking this into account could lead to serious artifacts, especially with electrostatic calculations.

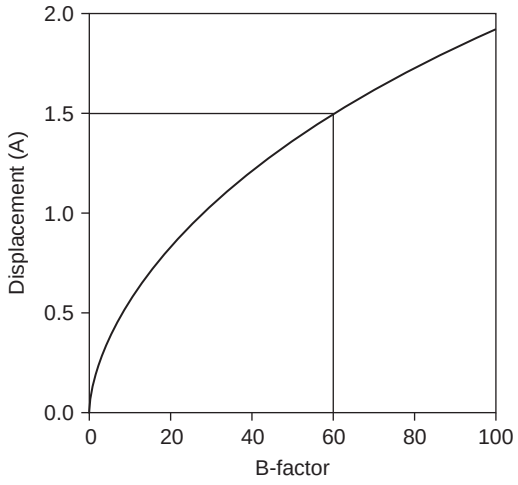


FIGURE 22.13 Relationship between mean total displacement and temperature factor B. At temperature factors of 60Å^2 and higher, the displacement becomes larger than 1.5Å and the electron density becomes very poor (see Figure 22.14). The formula used in the figure is derived from the relationship $B = 8\pi^2 < u^2 >$ where $< u^2 >$ represents the displacement perpendicular to the diffracting planes. The total mean square displacement $< u^2_{tot} > = 3 < u^2 >$, hence $< u_{tot} > = \sqrt{(3B/8\pi^2)}$.

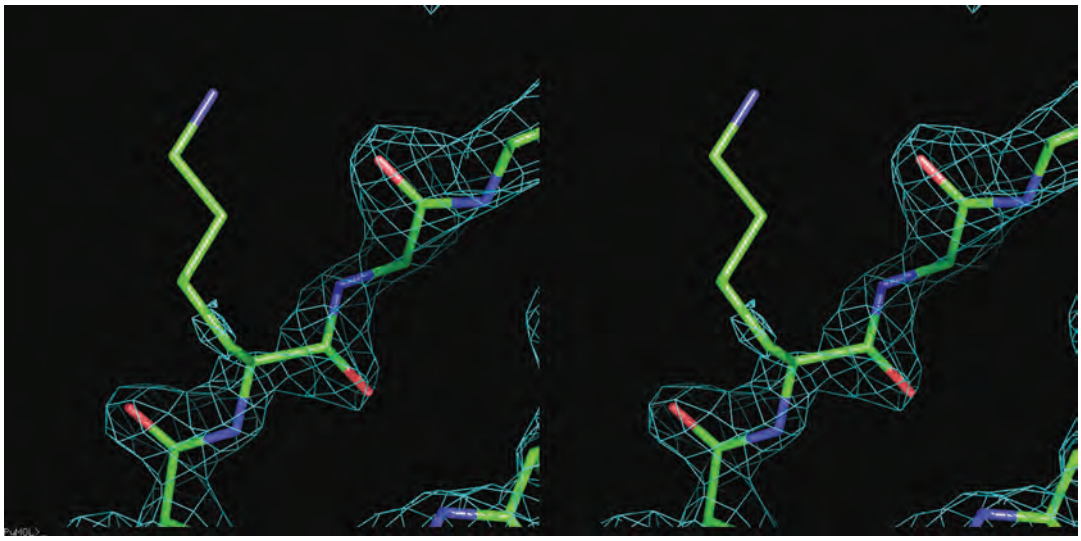


FIGURE 22.14 Long and flexible side-chains (such as Arg, Lys, Glu, and Gln) that are exposed to the solvent often move around freely. As a result, these side-chains have very high temperature factors, are very poorly or not at all defined in the electron-density maps, and are often fitted in an arbitrary, low-energy conformation. Lys87, located at the surface of human thrombin, is shown as an example. If one uses protein crystal structures for drug design, one should bear in mind that many exposed surface residues do not have a well-defined conformation.

5. *Misinterpretations of Electron-Density Maps*

Protein crystal structures are the result of a human interpretation of electron-density maps that are biased by the very model one is building. It is therefore no surprise that misinterpretations occur. Reasons for these errors include insufficient resolution or data quality, the presence of multiple overlapping binding modes, binding of several buffer components or fragments to the same site, and the lack of experience of the crystallographer.

False positives. False positives occur when a ligand is fitted to electron density belonging to bound solvent atoms, buffer components, or PEG molecules. Deleting the inhibitor and running a few rounds of refinement usually reveals this type of error clearly. Alternatively, one could run the Twilight script [86].

False negatives. There are two types of false negatives. With the first type, no bound ligand is found, although the ligand should bind according to biochemical and other binding assays [87]. In this case, crystal packing or the crystallization conditions (pH, high salt, high PEG) may prevent binding, or compound solubility may have

BOX 22.3

USING PDB FILES—TIPS AND TRICKS

Crystal structures are stored in formatted text files called "PDB" files. These files can be freely downloaded from the RCSB Protein Data Bank at http://www.pdb.org/. Information on how to search and navigate the PDB is available on the PDB home page.

Tip1: Always download a complete biological assembly. PDB files usually contain only the portion of the structure forming the asymmetric unit of the crystal. The asymmetric unit may include only a fraction of the functional biological molecule, for instance a single subunit of a homodimer. In such a case, information derived from viewing a single subunit may be very misleading, since binding sites or active sites are sometimes located at the interface between two or more subunits. Although a complete biological assembly can be generated with help of crystallographic software, it is also possible to download the corresponding file directly from the PDB (Download Files → Biological Assembly).

Tip2: Check all molecules of the asymmetric unit. Sometimes the asymmetric unit of the crystal contains several copies of the molecule or complex of interest. In such cases, individual copies of the biological assembly can be downloaded individually as separate files or all together in the original PDB file. It is very important to inspect them all, since significant differences can exist between these molecules due to different crystal contacts, disorder, partial occupancy of a ligand or co-factor, or as a consequence of different conformational states.

Tip3: Do not look only at the 3D model. Check the actual experimental information as well: the electron-density map. Electron-density maps contain more information than can possibly and accurately be included in an atomic model, even after careful refinement by an experienced crystallographer. For instance, some alternate conformations may not have been modeled (also note that many graphic programs ignore alternate conformations and do not display them). Some portions of a ligand molecule may be disordered, but coordinates for the complete molecule have been included. These and other important details can be revealed by displaying the electron-density map together with the atomic model. For structures which have been deposited with the PDB together with the corresponding diffraction data, electron-density map files can be downloaded from the Uppsala electron-density server [89] at http://eds.bmc.uu.se/eds/.

Tip4: Browse through your PDB file to find out more about its content. While the 3D structures encoded in the PDB files are best visualized using a graphic program (some interactive viewers are directly accessible from the PDB web pages), bear in mind that PDB files are simply text files that can also be displayed using a text editor. Browsing through PDB files can reveal some important information, notably on the method used to derive the structure (NMR, X-ray, or modeling), some data statistics, the amino-acid sequence with comments about engineered residues, the numbering of the protein residues and associated co-factors, ligands and solvent molecules, and more.

been too low to form enough complex. Hence, the failure to observe binding in an X-ray experiment does not necessarily disqualify a compound from being a genuine ligand. Before abandoning dubious but interesting hits, one should first verify them using other methods, such as protein NMR [88], mass spectroscopy, surface plasmon resonance, microcalorimetry, or thermophoresis.

In the second type of false negatives, compounds do bind but are not recognized as such. For example, at resolutions normally used in protein crystallography, the electron densities of ammonia, water, and sodium ions are virtually indistinguishable. Also bound buffer components and side products of the synthesis of the compound might not be recognized as such due to disorder or just because the exact chemical nature of the molecule is not known. In most cases, water molecules get fitted to these unknown densities.

IV. APPLICATIONS

A. Target Identification and Selection

1. Assignment of Function

Current low molecular-weight medicines exploit a fraction of all potential drug targets [90,91]. However, the large scale sequencing of whole genomes, including the human genome, has uncovered thousands of previously

unknown genes or "open reading frames." Functional annotation of these novel gene products is mainly based on sequence homologies to previously known proteins. For distant relatives, these homologies are often limited to a few short—but usually characteristic—sequence motifs. Tentative assignments are reinforced by sophisticated and powerful approaches, such as threading techniques, which verify the compatibility of a given amino-acid sequence with a 3D fold [92].

However, a substantial fraction of the novel genes code for proteins with no apparent relationship to any of the currently known ones. Structural genomic centers are solving the crystal structures of many of these novel proteins. In many cases the protein family and sometimes also the function can be deduced from the 3D structure, for example, from the presence of certain structural motifs with known catalytic functions like the Ser-His-Asp catalytic triad of serine proteases [93].

2. *Druggability (ligandability) Assessment*

While biology plays a key role in the selection of new targets, chemistry must have a strong say too, for pursuing a non-druggable target is a waste of time and resources. To get potent, selective, and orally active drugs, binding pockets with suitable properties must be present on the protein target. Potency and selectivity are usually achieved by optimizing the fit of the ligand to its receptor site, while oral bioavailability requires certain criteria to be met, such as Lipinski's "rule of five" [94]. Hence, a druggable target may be defined as a protein with a binding site of suitable size (that can accommodate compounds of MW $<$ 500Da), appropriate lipophilicity, and sufficient H-bonding potential [91], which can be deduced from crystal structures. Also, the presence of allosteric binding sites [95] and the existence of distinct structural conformations [85] can greatly increase the odds of finding a drug.

Most receptors and enzymes possess beautiful binding sites and are druggable [90,91]. In contrast, many protein−protein interaction sites are large and flat, and are therefore hopeless drug targets. However, because of their relevance to many diseases, protein−protein interactions are still attracting considerable interest, and a few may ultimately turn out to be druggable [96].

As a final remark, the methods mentioned above only estimate whether a ligand with a particular size and particular physiochemical properties is likely to bind to the target. They do not say whether such a compound will be a good drug that can be used to treat patients. For that reason, it might be more appropriate to speak of ligandability assessment instead of druggability assessment [97].

B. Hit/Lead Generation

1. *Structure-Based De Novo Drug Design*

The *de novo* design of novel scaffolds usually starts with careful scrutiny of multiple X-ray structures of the target in complex with a variety of ligands or tool compounds that in themselves are not attractive for chemical optimization due to issues with—for example—ease of synthesis and derivatization, intellectual property, and drug-likeness.

However, even the most promising templates designed with the most sophisticated computational tools are likely to have very weak potencies during the early stages of the *de novo* design process, so that standard biochemical assays may not be appropriate to evaluate these prototypic compounds. Protein crystallography may be of great help here, since in favorable cases it can detect high micromolar or even low millimolar binders. Once a first co-crystal structure with the designed template is obtained, subsequent optimization is usually straightforward.

2. *In Silico Screening*

Structure-based virtual screening, also called "high-throughput docking," involves the automatic docking and scoring of thousands of compounds to binding sites on protein targets [98−100]. Although the method has some shortcomings, like imperfect handling of receptor plasticity and reliability of scoring functions, its high-throughput and relatively low cost combined with its versatility outweigh these deficiencies. Most importantly, several recent success stories demonstrate that these methods do indeed deliver useful hits [101,102]. Obviously the effectiveness of high-throughput docking critically depends on the amount and quality of the structural information that is available for the drug target [102]. The outcome may further be improved by using target-based scoring and an expert system [103]. Particularly important is the understanding of the relevant conformational states and possible induced-fit mechanisms of the receptor binding site. Multiple co-crystal structures of the target of interest with different chemotypes, as well as any X-ray structures of related targets, contribute to this

understanding. Furthermore, it is essential that the most critical interaction sites or binding-site "hot spots" are identified [99]. Example of such key interaction sites include the hinge region of protein kinases and the flap and catalytic aspartates of aspartic proteinases. Sometimes, one or more conserved water molecules have been found to play an important role in ligand recognition and binding. Since the incorporation of such waters can strongly influence both the docking and the scoring steps, it is wise to search the available crystallographic data for the presence of conserved waters at critical locations within the receptor binding site [99,104]. Last but not least, it is also important to be aware of the limitations and uncertainties of crystal structures that can affect the virtual screening experiment. These uncertainties include the protonation state of protein residues and the exact orientation of some donor/acceptor groups, such as imidazole side-chains and the side-chain amide groups of asparagine and glutamine residues. Moreover, some important protein loops lining a binding site may not have well-defined electron density due to partial disorder (multiple conformations are present in the crystal), or the observed conformation may be influenced by the crystallization conditions or protein−protein contacts.

3. *Fragment-Based Screening*

While the two previous methods were computational (virtual), fragment-based screening (FBS) is an experimental method. Here, small fragments the size of a decorated benzene ring are screened for binding to the target protein. FBS emerged out of a need to overcome the current shortcomings of existing experimental or computational hit-finding approaches. The rationale behind the FBS strategy is well known: because the likelihood of a compound fitting a binding site decreases exponentially as the size increases [105], high-throughput screening approaches often fail to deliver hits, or they provide hits that are difficult to optimize owing to their low ligand efficiency [106,107] and "drug-like" rather than "lead-like" properties [108].

In contrast, the aim of FBS is to find hits that are easy to optimize by using a carefully selected fragment library [109]. Because of their small size and the fact that the entropic penalty associated with the loss of rigid-body translational and rotational freedom upon complex formation is independent of molecular weight [110,111], small fragments bind weakly, even when their ligand efficiency is high. Consequently, highly sensitive robust experimental techniques are needed to detect these weak binders. Historically, NMR has played a pioneering role in the development of FBS [9], but other technologies are applied as well, such as mass spectroscopy, surface plasmon resonance, and protein crystallography [100]. We hope that the reader will not take it amiss if we concentrate below on protein crystallographic applications to FBS. More general information on FBS can be found in [112]) or chapter 8 of this book.

To optimize FBS hits, it is essential to know their binding mode for a couple of reasons. First, the biophysical techniques used in FBS detect binding and not biochemical or biological activity and binding may be anywhere. Second, even if we know the binding site from competition experiments, modeling or docking of small fragments is usually ambiguous.

When one has access to a large compound store, it is often possible to dig out analogues with improved potency by substructure or similarity searches, and in this way to generate structure−activity data easily. For this reason, such an approach is often called "SAR by inventory." Nevertheless, in the absence of more detailed structural information, the optimization of weak FBS hits into potent leads can be a lengthy and cumbersome process.

A variety of NMR techniques exist to infer structural details on protein−ligand interactions [88]. However, protein crystallography remains the preferred approach for elucidating binding modes with certainty and guiding the hit-to-lead phase. Unfortunately, experience shows that only a fraction of FBS hits discovered by NMR or other biophysical techniques can be observed by protein crystallography.

Fragment-based screening by X-ray crystallography. When a suitable crystallization platform is available, one may consider using protein crystallography as the main FBS screening technique. Since crystallographic information is usually essential for the subsequent hit optimization, the use of X-ray analysis from the start can save time and certainly avoids the frustration of finding hits that cannot be reproduced later by crystallography.

Before an FBS by X-ray campaign can be launched, an initial investment in the preparation of suitable crystals may be needed [113], since crystals originally used for the first structure determination of a new drug target may not be suited. They may be difficult to grow or not diffract well enough, or the binding site of interest may be occupied by a strong ligand. For FBS by X-ray, it is essential that the crystals diffract to high resolution (better than 2.5Å, preferably 2.0Å or better) and are amenable to soaking, which implies that the targeted binding site is free and accessible. In cases where the crystallization is particularly robust, co-crystallization with the fragment cocktails can be attempted, but this strategy is usually less effective than the soaking approach. High crystal symmetry is not a must but makes data collection faster. When suitable crystals are not available, it may be necessary to engineer and produce new protein variants.

Fragments should be highly soluble under crystallization conditions. With typical protein concentrations in a crystallization experiments in the 0.1−1.0 mM range, fragments should be soluble up to concentrations of 1−10 mM. This is particularly critical when crystals are grown under high salt conditions. Apolar and aromatic scaffolds should feature one or more solubilizing group, such as a carboxylic or ammonium group. The risk that electrostatic interactions dominate binding is largely alleviated under high salt conditions, which strengthen hydrophobic interactions at the expense of the electrostatic ones.

With current technology, a library of 500 to 1,000 fragments split up into cocktails of five to ten compounds can be screened by X-ray crystallography within reasonable timelines. The cocktails should be designed in such a way that each component of a mixture has a distinct shape to allow unambiguous identification of any bound fragment on the basis of the shape of the electron density.

Over the past decade, FBS by X-ray has made notable contributions to the overall success of the fragment-based screening approach. It has provided novel, chemically attractive leads for some notoriously difficult targets, such as β-secretase [114,115], and these hits could be successfully optimized to highly potent drug candidates, hence fulfilling the initial promise of this approach.

4. *Triaging and Validation of HTS Hits*

Protein crystallography plays an important role in hit validation and selection for further optimization. Whereas before, only a few selected HTS hits could be analyzed crystallographically, high-throughput crystallography allows a more systematic approach. It is now possible to analyze many hits, including some of those that in the past would have been discarded. This analysis can provide highly valuable information regarding novel binding sites or subsites, alternative binding modes, privileged interaction patterns, and protein conformational substates. This information can then be fed into the structure-based design process, even when some of these hits are not pursued any further. Moreover, among the weak hits many compounds are often "fragment-like," with molecular weights in the 150−250 Da range. It may be of particular interest to investigate this region of chemical space where HTS meets FBS.

Finally, having a co-crystal structure of a hit bound to its target protein provides definitive experimental proof that the compound was not a false positive of some sort [116] and reassures the chemists that fast, structure-based optimization of the compound will be feasible.

C. Lead Optimization

Crystallographic information greatly enhances the speed and efficiency of lead optimization. However, crystal structures only show part of the picture and do not provide information on factors such as physicochemical properties, toxicity issues, metabolic weak points, thermodynamic parameters such as entropy and enthalpy, or protonation states of active site residues. For successful lead optimization, data obtained from many different sources need to be brought together.

1. *Optimizing Potency*

Large weakly binding compounds have very poor ligand efficiencies and are generally difficult to optimize. In most cases, fragments with a high ligand efficiency are much better starting points. The identification of key interaction sites (or "hot spots") within the binding pocket is a first and essential step when an enhancement in potency is sought [117]. To this end, an experimental fragment-based approach can be used [118] or computational methods can be utilized [42,119,120]. Interactions with the binding site hot spots should be maximized through the introduction of new substituents or the replacement of functionalities making sub-optimum contacts.

Protein ligands rarely bind in their lowest energy conformation [121]. When present, unfavorable strain energy should be detected and minimized. An analysis of the conformation of related compounds in the Cambridge data bank can guide this process [122], and *ab initio* calculations are often useful [122,123]. Compounds requiring minimal conformational reorganization on enzyme binding should be favored. Small-molecule ligands frequently adopt an extended conformation in the bound state [121]. Hence, hydrophobic ligands exhibiting a folded conformation in solution may incur a high reorganization energy cost on binding. Introduction of conformational restraints through (macro)cyclization [124] or the introduction of rigid linkers [125] is another strategy that has been successfully used in many cases to minimize entropic penalties.

2. *Optimizing Selectivity*

Exquisite selectivity can often be achieved by exploiting binding subsites or pockets adjacent to the main binding site that are not involved in the normal biological function of the drug target and are thus poorly conserved in other family members [126,127]. Likewise, taking advantage of the flexibility of the protein by targeting an unusual conformational state is an excellent means to achieving high selectivity. Protein kinase inhibitors provide numerous examples of this kind [128,129].

One disadvantage of the above approaches is that resistance mutations are more likely to emerge when nonfunctional states or cavities are used by the drug [127]. Designing more flexible compounds has also been proposed as a possible strategy for achieving broad-spectrum activity [130], since rigid molecules are less likely to adapt to structural changes. The entropic cost of the built-in flexibility needs to be compensated by a larger enthalpic contribution to binding through an optimization of all available polar interactions. Hence, enthalpic optimization of the binding affinity has been proposed as a better alternative to potency enhancement through hydrophobic binding and rigid fit [131]. In doing so, however, it is important to ensure that the strongest interactions involve residues with a low probability to mutate [132].

3. *Optimizing ADME Properties*

A general recipe for turning a potent lead into a real drug candidate does not exist, but some guidelines are available, such as the well-known "rule of five" [94]. Structure-guided design can aid in achieving the right balance between lipophilicity and polar surface area by guiding the introduction or replacement of heteroatoms, polar groups, and other solubilizing groups. Essential hydrogen-bonded interactions can be identified along with dispensable acceptor/donor groups. Minimizing molecular weight is frequently an effective strategy to achieving good oral bioavailability [122,133] and X-ray structures can identify groups and substituents that do not contribute much to binding and can therefore be replaced or removed. Suitable replacements for functionalities that are detrimental to ADME properties can be sought using focused chemical libraries or a FBS approach. Ultimately, it may be necessary to switch to a different chemotype. The availability of multiple X-ray structures of diverse molecular scaffolds may allow the combination of two different scaffolds into a new one [134], or the grafting of one particular motif from one inhibitor onto another [135,136].

V. TWO SELECTED EXAMPLES

A. Imatinib (Gleevec™)

The development of protein kinase inhibitors targeting the ATP binding site was initially received with great skepticism, on the grounds that it would not be possible to achieve a sufficient level of selectivity to turn them into useful therapeutic agents. In view of the large size of the human kinome [137] (518 genes) and the high conservation of the ATP binding site, this criticism was well founded. However, the discovery of imatinib [138,139] (Glivec®, Gleevec™), an inhibitor of the tyrosine kinase activity of the Bcr-Abl oncogene and an effective, frontline therapy for chronic myelogenous leukemia, provided compelling evidence for the viability of this approach.

The X-ray structure of the abl kinase domain in complex with des-methylpiperazinyl imatinib became available in 2000 [140], soon followed by the imatinib complex [141,142]. The N-methylpiperazine moiety of imatinib had been introduced during the lead-optimization phase to improve solubility, at a point in time where the exact binding mode of the drug was not known. Unexpectedly, the X-ray analyses revealed that the drug was binding to an inactive conformation of the kinase, with the benzamide and piperazinyl groups accessing a channel at the back of the ATP site (Figure 22.15). A conformational switch of the DFG motif of the kinase was responsible for the formation of this channel, which is therefore referred to as the "DFG-out" pocket. In this mode of binding, the N-methylpiperazine moiety was only partially exposed to solvent and strongly interacted with the kinase [127]. More importantly, several structural features of the inactive state of the abl kinase were important for imatinib binding, and detailed structural comparisons indicated that these features were poorly conserved in other protein kinases, thus explaining the high selectivity of this compound [127]. In addition, these and follow-up structures provided a platform for the analysis of resistance mutants [127,128]. The concept of DFG-in and DFG-out conformations has become a central theme in the search for kinase inhibitors.

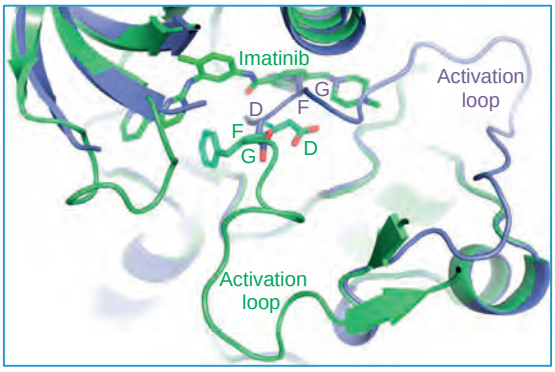


FIGURE 22.15 Close-up of a superposition of unligated abl kinase [127] (blue; pdb code 2hz4, chain A; DFG-in) and abl kinase complexed with imatinib [127] (green, pdb code 2hyy, chain A; DFG-out). In the DFG-in conformation, the position of the Phe of the DFG motif completely overlaps with the imatinib molecule bound to the DFG-out conformation. Also, the activations loops, which contain the DFG motif, assume completely different conformations.

B. The β_2-Adrenergic Receptor

The human genome contains at least 800 GPCRs that respond to a broad range of molecules and other entities, including photons, protons, odorants, neurotransmitters, hormones, and glycoproteins [61]. Since GPCRs are involved in most physiological processes, they represent the largest class of drug targets [58]. However, due to extreme difficulties in crystallizing them, very little structural information was available on GPCRs until recently. Structure-based drug design on GPCRs was mainly done using homology models, constructed on the basis of the rhodopsin crystal structure [143].

The β_2AR is not a drug target, but the closely related β_1AR is the target of beta blockers, a class of drugs widely used to treat heart patients. For that reason, a rich diversity of commercial ligands (full, partial, and inverse agonists, and neutral antagonists) was available. Using such compounds, Brian Kobilka and colleagues solved the mechanism of GPCRs by determining crystal structures of active and inactive states of β_2AR. For this work, Brian Kobilka—together with Robert Lefkowitz—was awarded the 2012 Nobel Prize in chemistry.

Crucial for this success was the stabilization of the inherently flexible GPCR (see Figure 22.6). The inactive state was crystallized using the potent inverse agonist carazolol and an antibody fragment binding the flexible loop between helices M5 and M6 [29], or, alternatively, by replacing this loop with the stable protein T4 lysozyme [144]. To crystallize the active state [145], the ultra-high affinity agonist BI-167107 from Boehringer-Ingelheim was used with either the G-protein mimicking nanobody [146] (single-chain camelid antibody fragment) Nb80 or a combination of inserted T4 lysozyme, Gs protein, and the nanobody Nb35.

Comparison of the inactive and active states (Figure 22.16) reveals that small differences in the ligand binding pocket due to antagonist or agonist binding are amplified via a repacking of Ile121, Pro211, Phe282, and Asn318 in the core of the β_2AR molecule. A rotation of helix TM6 results in a 14Å outward movement of the tip of this helix and causes conformational changes in the associated G-protein, ultimately resulting in the exchange of GDP by GTP and activation of the G-protein.

VI. OUTLOOK

A sequence-based search in the January 2014 release of the PDB with the sequence of the β_1-adrenergic receptor yielded 101 GPCR structures corresponding to thirty-one unique sequences. This number is rapidly increasing, paving the way for true structure-based GPCR drug discovery.

Finally, the construction of free electron lasers all over the world may again radically change the way protein crystallography is done. By using extremely short (10−200 fs) and extremely bright ($>10^{12}$ photons) X-ray pulses, it is possible to record useful X-ray diffraction before the atoms in the crystal have had time to move (i.e., before any radiation damage could occur). In a test experiment, a full 8.5Å data set was recently collected from nanocrystals of Photosystem I [147], a large membrane protein complex with a molecular mass of 1 MDa, 36 proteins, and 381 co-factors. Diffraction was observed from crystals smaller than ten unit cells on a side. Being able to use nanocrystals would remove one of the biggest hurdles in protein crystallography, namely the preparation of large, well-diffracting crystals.

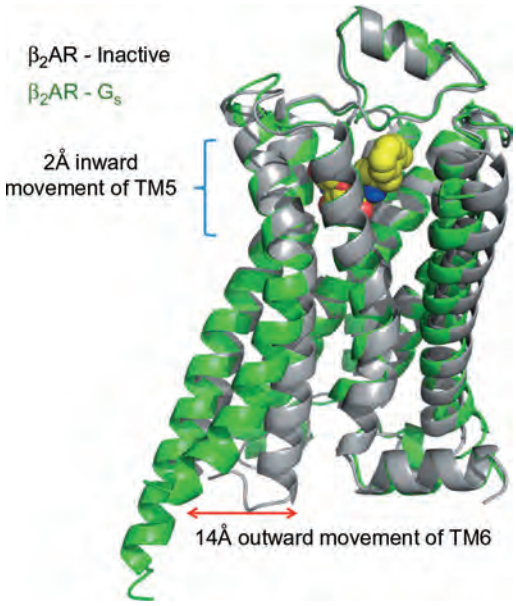


FIGURE 22.16 A comparison of the carazolol-bound, inactive-state structure of the β_2AR [144] (grey) and the active-state structure of the β_2AR (green) from the β_2AR–Gs complex [145]. *Figure by Brian Kobilka [61].* © *The Nobel Foundation 2012.*

References

[1] Shakespeare W. Macbeth. Paris: Editions Aubier Montaign; 1977.
[2] Congreve M, Murray CW, Blundell TL. Keynote review: structural biology and drug discovery. Drug Discov Today 2005;10(13):895–907.
[3] Scapin G. Structural biology and drug discovery. Curr Pharm Des 2006;12(17):2087–97.
[4] Tintelnot-Blomley M, Lewis RA. A critical appraisal of structure-based drug design. IDrugs 2006;9(2):114–18.
[5] Lombardino JG, Lowe JA. The role of the medicinal chemist in drug discovery—then and now. Nat Rev Drug Discov 2004;3(10):853–62.
[6] Williams SP, Kuyper LF, Pearce KH. Recent applications of protein crystallography and structure-guided drug design. Curr Opin Chem Biol 2005;9(4):371–80.
[7] Weigelt J. Structural genomics: impact on biomedicine and drug discovery. Exp Cell Res 2010;316(8):1332–8.
[8] Blundell TL, Jhoti H, Abell C. High-throughput crystallography for lead discovery in drug design. Nat Rev Drug Discov 2002;1(1):45–54.
[9] Hajduk PJ, Greer J. A decade of fragment-based drug design: strategic advances and lessons learned. Nat Rev Drug Discov 2007;6 (3):211–19.
[10] Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by X-Ray analysis. Nature 1958;181(4610):662–6.
[11] Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North ACT. Structure of haemoglobin: a three-dimensional fourier synthesis at 5.5-Å resolution, obtained by X-ray analysis. Nature 1960;185(4711):416–22.
[12] McPherson A. A brief history of protein crystal growth. J Cryst Growth 1991;110(1–2):1–10.
[13] Franklin RE, Gosling RG. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. Nature 1953;172(4369):156–7.
[14] Wilkins MHF, Seeds WE, Stokes AR, Wilson HR. Helical structure of crystalline deoxypentose nucleic acid. Nature 1953;172(4382): 759–62.
[15] Watson JD, Crick FHC. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature 1953;171(4356):737–8.
[16] Perutz M. Protein structure: new approaches to disease and therapy. New York: WH Freeman and Company; 1992.
[17] Rossmann MG, Arnold E, Erickson JW, Frankenberger EA, Griffith JP, Hecht H-J, et al. Structure of a human common cold virus and functional relationship to other picornaviruses. Nature 1985;317(6033):145–53.
[18] Deisenhofer J, Epp O, Miki K, Huber R, Michel H. Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3Å resolution. Nature 1985;318(6047):618–24.
[19] Abrahams JP, Leslie AGW, Lutter R, Walker JE. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature 1994;370(6491):621–8.
[20] Lowe J, Stock D, Jap B, Zwickl P, Baumeister W, Huber R. Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. Science 1995;268(5210):533–9.
[21] Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8A resolution. Nature 1997;389(6648):251–60.
[22] Brodersen DE, Clemons Jr WM, Carter AP, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. Cell 2000;103(7):1143–54.
[23] Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science 2000;289(5481):905–20.
[24] Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. Science 2000;289 (5481):920–30.

[25] Pioletti M, Schlünzen F, Harms J, Zarivach R, Glühmann M, Avila H, et al. Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine, and IF3. EMBO J 2001;20(8):1829–39.
[26] Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: RNA polymerase II at 2.8 Ångstrom resolution. Science 2001;292(5523):1863–76.
[27] Doyle DA, Cabral JM, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, et al. The structure of the potassium channel: molecular Basis of K+ conduction and selectivity. Science 1998;280(5360):69–77.
[28] Jiang Y, Lee A, Chen J, Cadene M, Chait BT, MacKinnon R. Crystal structure and mechanism of a calcium-gated potassium channel. Nature 2002;417(6888):515–22.
[29] Rasmussen SGF, Choi H-J, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, et al. Crystal structure of the human [bgr]2 adrenergic G-protein-coupled receptor. Nature 2007;450(7168):383–7.
[30] Kroemer M, Dreyer MK, Wendt KU. APRV: a program for automated data processing, refinement, and visualization. Acta Crystallogr D 2004;60(9):1679–82.
[31] Mooij WTM, Hartshorn MJ, Tickle IJ, Sharff AJ, Verdonk ML, Jhoti H. Automated protein–ligand crystallography for structure-based drug design. Chem Med Chem 2006;1(8):827–38.
[32] Evrard GX, Langer GG, Perrakis A, Lamzin VS. Assessment of automatic ligand building in ARP/wARP. Acta Crystallogr D 2007;63(1):108–17.
[33] Yumerefendi H, Tarendeau F, Mas PJ, Hart DJ. ESPRIT: an automated, library-based method for mapping and soluble expression of protein domains from challenging targets. J Struct Biol 2010;172(1):66–74.
[34] Beddell CR, Goodford PJ, Norrington FE, Wilkinson S, Wootton R. Compounds designed to fit a site of known structure in human haemoglobin. Br J Pharmacol 1976;57(2):201–9.
[35] Ondetti M, Rubin B, Cushman D. Design of specific inhibitors of angiotensin-converting enzyme: new class of orally active antihypertensive agents. Science 1977;196(4288):441–4.
[36] Natesh R, Schwager SLU, Evans HR, Sturrock ED, Acharya KR. Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. Biochemistry 2004;43(27):8718–24.
[37] Eriksson AE, Jones TA, Liljas A. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. Proteins: structure. Funct Bioinform 1988;4(4):274–82.
[38] Baldwin JJ, Ponticello GS, Anderson PS, Christy ME, Murcko MA, Randall WC, et al. Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. J Med Chem 1989;32(12):2510–13.
[39] Wade RC. 'Flu' and structure-based drug design. Structure 1997;5(9):1139–45.
[40] Varghese JN, Laver WG, Colman PM. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. Nature 1983;303(5912):35–40.
[41] Colman PM, Varghese JN, Laver WG. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. Nature 1983;303(5912):41–4.
[42] Goodford PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 1985;28(7):849–57.
[43] von Itzstein M, Wu W-Y, Kok GB, Pegg MS, Dyason JC, Jin B, et al. Rational design of potent sialidase-based inhibitors of influenza virus replication. Nature 1993;363(6428):418–23.
[44] Kim CU, Lew W, Williams MA, Liu H, Zhang L, Swaminathan S, et al. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J Am Chem Soc 1997;119(4):681–90.
[45] Matthews BW. Solvent content of protein crystals. J Mol Biol 1968;33(2):491–7.
[46] Ducruix A, Giegé R. Crystallization of nucleic acids and proteins: a practical approach. 2nd ed. Oxford: Oxford University Press; 1993.
[47] McPherson A. Crystallization of biological macromolecules. Cold Spring Harbor: Laboratory Press; 1999.
[48] Kundrot CE. Which strategy for a protein crystallization project? Cell Mol Life Sci 2004;61(5):525–36.
[49] Pusey ML, Liu Z-J, Tempel W, Praissman J, Lin D, Wang B-C, et al. Life in the fast lane for protein crystallization and X-ray crystallography. Prog Biophys Mol Biol 2005;88(3):359–86.
[50] Derewenda ZS. The use of recombinant methods and molecular engineering in protein crystallization. Methods 2004;34(3):354–63.
[51] Wright PE, Dyson HJ. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. J Mol Biol 1999;293(2):321–31.
[52] McPherson A, Cudney B. Searching for silver bullets: an alternative strategy for crystallizing macromolecules. J Struct Biol 2006;156(3):387–406.
[53] Chang VT, Crispin M, Aricescu AR, Harvey DJ, Nettleship JE, Fennelly JA, et al. Glycoprotein structural genomics: solving the glycosylation problem. Structure 2007;15(3):267–73.
[54] Strauss A, Fendrich G, Horisberger MA, Liebetanz J, Meyhack B, Schlaeppi J-M, et al. Improved expression of kinases in Baculovirus-infected insect cells upon addition of specific kinase inhibitors to the culture helpful for structural studies. Protein Expr Purif 2007;56(2):167–76.
[55] Brading RL, Abbott WM, Green I, Davies A, McCall EJ. Co-expression of protein phosphatases in insect cells affects phosphorylation status and expression levels of proteins. Protein Expr Purif 2012;83(2):217–25.
[56] Wiener MC. A pedestrian guide to membrane protein crystallization. Methods 2004;34(3):364–72.
[57] Minor Jr DL. The neurobiologist's guide to structural biology: a primer on why macromolecular structure matters and how to evaluate structural data. Neuron 2007;54(4):511–33.
[58] Yong K, Scott D. Engineering stabilised G protein-coupled receptors for biochemical and structural studies. Aust Biochem 2013;44(2):7–12.
[59] Faham S, Bowie JU. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. J Mol Biol 2002;316(1):1–6.
[60] Caffrey M. Membrane protein crystallization. J Struct Biol 2003;142(1):108–32.

[61] Kobilka B. The structural basis of G-protein-coupled receptor signaling (Nobel Lecture). Angew Chem Int Ed 2013;52(25):6380–8.
[62] Congreve M, Rich RL, Myszka DG, Figaroa F, Siegal G, Marshall FH. Fragment screening of stabilized G-protein-coupled receptors using biophysical methods N.p. In: Lawrence CK, editor. Methods in enzymology. Academic Press; 2011. p. 115–36
[63] Hassell AM, An G, Bledsoe RK, Bynum JM, Carter HL, Deng S-JJ, et al. Crystallization of protein–ligand complexes. Acta Crystallogr D 2007;63(1):72–9.
[64] Lusty CJ. A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography. J Appl Crystallogr 1999;32 (1):106–12.
[65] D'Arcy A, Villard F, Marsh M. An automated microseed matrix-screening method for protein crystallization. Acta Crystallogr D 2007;63 (4):550–4.
[66] Garman EF, Owen RL. Cryocooling and radiation damage in macromolecular crystallography. Acta Crystallogr D 2006;62(1):32–47.
[67] Dauter Z, Dauter M, Dodson E. Jolly SAD. Acta Crystallogr D 2002;58(3):494–506.
[68] Ealick SE. Advances in multiple wavelength anomalous diffraction crystallography. Curr Opin Chem Biol 2000;4(5):495–9.
[69] Dauter Z, Dauter M. Entering a new phase: using solvent halide ions in protein structure determination. Structure 2001;9(2):R21–6.
[70] Debreczeni JÉ, Bunkóczi G, Ma Q, Blaser H, Sheldrick GM. In-house measurement of the sulfur anomalous signal and its use for phasing. Acta Crystallogr D 2003;59(4):688–96.
[71] Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. J Appl Crystallogr 1997;30(6):1022–5.
[72] Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D 2006;62(9):1002–11.
[73] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. Nucleic Acids Res 2000;28(1):235–42.
[74] Read RJ. Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr A 1986;42(3):140–9.
[75] McCoy AJ. Liking likelihood. Acta Crystallogr D 2004;60(12–1):2169–83.
[76] Petrova T, Podjarny A. Protein crystallography at subatomic resolution. Rep Prog Phys 2004;67(9):1565.
[77] Howard EI, Sanishvili R, Cachau RE, Mitschler A, Chevrier B, Barth P, et al. Ultrahigh resolution drug design. I. Details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. Proteins: structure. Funct Bioinform 2004;55(4):792–804.
[78] Diederichs K, Karplus PA. Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat Struct Biol 1997;4 (4):269–75.
[79] Brunger AT. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 1992;355(6359):472–5.
[80] Jia Z, Vandonselaar M, Quail JW, Delbaere LTJ. Active-centre torsion-angle strain revealed in 1.6 A-resolution structure of histidine-containing phosphocarrier protein. Nature 1993;361(6407):94–7.
[81] Chevrier B, Schalk C, D'Orchymont H, Rondeau J-M, Moras D, Tarnus C. Crystal structure of Aeromonas proteolytica aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. Structure 1994;2(4):283–91.
[82] Davis AM, Teague SJ, Kleywegt GJ. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. Angew Chem Int Ed 2003;42(24):2718–36.
[83] DePristo MA, de Bakker PIW, Blundell TL. Heterogeneity and inaccuracy in protein structures solved by X-Ray crystallography. Structure 2004;12(5):831–8.
[84] Kleywegt GJ. Crystallographic refinement of ligand complexes. Acta Crystallogr D 2007;63(1):94–100.
[85] Teague SJ. Implications of protein flexibility for drug discovery. Nat Rev Drug Discov 2003;2(7):527–41. Available from: http://dx.doi.org/10.1038/nrd1129.
[86] Weichenberger CX, Pozharski E, Rupp B. Visualizing ligand molecules in twilight electron density. Acta Crystallogr F 2013;69 (2):195–200.
[87] Chung C-W. The use of biophysical methods increases success in obtaining liganded crystal structures. Acta Crystallogr D 2007;63(1): 62–71.
[88] Jahnke W, Widmer H. Protein NMR in biomedical research. Cell Mol Life Sci 2004;61(5):580–99.
[89] Kleywegt GJ, Harris MR, Zou J-Y, Taylor TC, Wählby A, Jones TA. The uppsala electron-density server. Acta Crystallogr D 2004;60 (12–1):2240–9.
[90] Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006;5(12):993–6.
[91] Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov 2002;1(9):727–30.
[92] Blundell TL, Sibanda BL, Montalvão RW, Brewerton S, Chelliah V, Worth CL, et al. Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. Philos Trans R Soc B Biol Sci 2006;361(1467):413–23.
[93] Laskowski RA, Watson JD, Thornton JM. Protein function prediction using local 3D templates. J Mol Biol 2005;351(3):614–26.
[94] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;23(1–3):3–25.
[95] Hardy JA, Wells JA. Searching for new allosteric sites in enzymes. Curr Opin Struct Biol 2004;14(6):706–15.
[96] Chène P. Drugs targeting protein–protein interactions. Chem Med Chem 2006;1(4):400–11.
[97] Edfeldt FNB, Folmer RHA, Breeze AL. Fragment screening to predict druggability (ligandability) and lead discovery success. Drug Discov Today 2011;16(7–8):284–7.
[98] Shoichet BK. Virtual screening of chemical libraries. Nature 2004;432(7019):862–5.
[99] Klebe G. Virtual ligand screening: strategies, perspectives, and limitations. Drug Discov Today 2006;11(13–14):580–94.
[100] Rees DC, Congreve M, Murray CW, Carr R. Fragment-based lead discovery. Nat Rev Drug Discov 2004;3(8):660–72.
[101] Cavasotto CN, Orry AJ. Ligand docking and structure-based virtual screening in drug discovery. Curr Top Med Chem 2007;7 (10):1006–14.
[102] Alvarez JC. High-throughput docking as a source of novel drug leads. Curr Opin Chem Biol 2004;8(4):365–70.
[103] Jansen JM, Martin EJ. Target-biased scoring approaches and expert systems in structure-based virtual screening. Curr Opin Chem Biol 2004;8(4):359–64.
[104] Zentgraf M, Steuber H, Koch C, La Motta C, Sartini S, Sotriffer CA, et al. How reliable are current docking approaches for structure-based drug design? Lessons from aldose reductase. Angew Chem Int Ed 2007;46(19):3575–8.

[105] Hann MM, Leach AR, Harper G. Molecular complexity and its impact on the probability of finding leads for drug discovery. J Chem Inf Comput Sci 2001;41(3):856–64.
[106] Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today 2004;9(10):430–1.
[107] Abad-Zapatero C. Ligand efficiency indices for effective drug discovery. Expert Opin Drug Discov 2007;2(4):469–88.
[108] Teague SJ, Davis AM, Leeson PD, Oprea T. The design of leadlike combinatorial libraries. Angew Chem Int Ed 1999;38(24):3743–8.
[109] Congreve M, Carr R, Murray C, Jhoti HA. "Rule of Three" for fragment-based lead discovery?. Drug Discov Today 2003;8(19):876–7.
[110] Finkelstein AV, Janin J. The price of lost freedom: entropy of bimolecular complex formation. Protein Eng 1989;3(1):1–3.
[111] Murray C, Verdonk M. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. J Comput Aided Mol Des 2002;16(10):741–53.
[112] Jahnke W, Erlanson DA, editors. Fragment-based approaches in drug discovery. Weinheim: Wiley-VCH; 2006.
[113] Nienaber V, Wang J, Davidson D, Henkin J. Re-engineering of human urokinase provides a system for structure-based drug design at high resolution and reveals a novel structural subsite. J Biol Chem 2000;275(10):7239–48.
[114] Murray CW, Callaghan O, Chessari G, Cleasby A, Congreve M, Frederickson M, et al. Application of fragment screening by X-ray crystallography to β-secretase†. J Med Chem 2007;50(6):1116–23.
[115] Congreve M, Aharony D, Albert J, Callaghan O, Campbell J, Carr RAE, et al. Application of fragment screening by X-ray crystallography to the discovery of aminopyridines as inhibitors of β-secretase§. J Med Chem 2007;50(6):1124–32.
[116] Böcker A, Bonneau PR, Edwards PJ. HTS Promiscuity analyses for accelerating decision making. J Biomol Screening 2011;16(7):765–74.
[117] Brenk R, Klebe G. "Hot spot" analysis of protein-binding sites as a prerequisite for structure-based virtual screening and lead optimization. In: Langer T, Hoffmann RD, editors. Pharmacophores and pharmacophore searches, vol. 32. Weinheim: Wiley-VCH; 2006. p. 171–92.
[118] Ciulli A, Williams G, Smith AG, Blundell TL, Abell C. Probing hot spots at protein − ligand binding sites: A fragment-based approach using biophysical methods. J Med Chem 2006;49(16):4992–5000.
[119] Gohlke H, Hendlich M, Klebe G. Predicting binding modes, binding affinities and "hot spots" for protein–ligand complexes using a knowledge-based scoring function. Perspect Drug Discov Des 2000;20(1):115–44.
[120] Verdonk ML, Cole JC, Watson P, Gillet V, Willett P. Superstar: improved knowledge-based interaction fields for protein binding sites. J Mol Biol 2001;307(3):841–59.
[121] Perola E, Charifson PS. Conformational analysis of drug-like molecules bound to proteins: An extensive study of ligand reorganization upon binding. J Med Chem 2004;47(10):2499–510.
[122] Kim EE, Baker CT, Dwyer MD, Murcko MA, Rao BG, Tung RD, et al. Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. J Am Chem Soc 1995;117(3):1181–2.
[123] Greer J, Erickson JW, Baldwin JJ, Varney MD. Application of the three-dimensional structures of protein target molecules in structure-based drug design. J Med Chem 1994;37(8):1035–54.
[124] MacPherson LJ, Bayburt EK, Capparelli MP, Bohacek RS, Clarke FH, Ghai RD, et al. Design and synthesis of an orally active macrocyclic neutral endopeptidase 24.11 inhibitor. J Med Chem 1993;36(24):3821–8.
[125] Kim S-H, Pudzianowski AT, Leavitt KJ, Barbosa J, McDonnell PA, Metzler WJ, et al. Structure-based design of potent and selective inhibitors of collagenase-3 (MMP-13). Bioorg Med Chem Lett 2005;15(4):1101–6.
[126] Rahuel J, Rasetti V, Maibaum J, Rüeger H, Göschke R, Cohen NC, et al. Structure-based drug design: the discovery of novel nonpeptide orally active inhibitors of human renin. Chem Biol 2000;7(7):493–504.
[127] Cowan-Jacob SW, Fendrich G, Floersheimer A, Furet P, Liebetanz J, Rummel G, et al. Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia. Acta Crystallogr D 2007;63(1):80–93.
[128] Noble MEM, Endicott JA, Johnson LN. Protein kinase inhibitors: insights into drug design from structure. Science 2004;303 (5665):1800–5.
[129] Cowan-Jacob SW. Structural biology of protein tyrosine kinases. Cell Mol Life Sci 2006;63(22):2608–25.
[130] Das K, Clark AD, Lewi PJ, Heeres J, de Jonge MR, Koymans LMH, et al. Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (Etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. J Med Chem 2004;47(10):2550–60.
[131] Velazquez-Campoy A, Todd MJ, Freire E. HIV-1 Protease inhibitors: Enthalpic versus entropic optimization of the binding affinity. Biochemistry 2000;39(9):2201–7.
[132] Ohtaka H, Velázquez-Campoy A, Xie D, Freire E. Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. Protein Sci 2002;11(8):1908–16.
[133] Brown FJ, Andisik DW, Bernstein PR, Bryant CB, Ceccarelli C, Damewood Jr JR, et al. Design of orally active, non-peptidic inhibitors of human leukocyte elastase. J Med Chem 1994;37(9):1259–61.
[134] Pierce AC, Rao G, Bemis GW. BREED: Generating novel inhibitors through hybridization of known ligands. Application to CDK2, P38, and HIV protease. J Med Chem 2004;47(11):2768–75.
[135] Anderson M, Beattie JF, Breault GA, Breed J, Byth KF, Culshaw JD, et al. Imidazo[1,2-a]pyridines: a potent and selective class of cyclin-dependent kinase inhibitors identified through structure-based hybridisation. Bioorg Med Chem Lett 2003;13(18):3021–6.
[136] Terasaka T, Kinoshita T, Kuno M, Nakanishi I. A highly potent non-nucleoside adenosine deaminase inhibitor: efficient drug discovery by intentional lead hybridization. J Am Chem Soc 2003;126(1):34–5.
[137] Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002;298 (5600):1912–34.
[138] Zimmermann J, Buchdunger E, Mett H, Meyer T, Lydon NB. Potent and selective inhibitors of the Abl-kinase: phenylamino-pyrimidine (PAP) derivatives. Bioorg Med Chem Lett 1997;7(2):187–92.
[139] Capdeville R, Buchdunger E, Zimmermann J, Matter A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. Nat Rev Drug Discov 2002;1(7):493–502.

[140] Schindler T, Bornmann W, Pellicena P, Miller WT, Clarkson B, Kuriyan J. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. Science 2000;289(5486):1938–42.
[141] Nagar B, Bornmann WG, Pellicena P, Schindler T, Veach DR, Miller WT, et al. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and Imatinib (STI-571). Cancer Res 2002;62(15):4236–43.
[142] Manley PW, Cowan-Jacob SW, Buchdunger E, Fabbro D, Fendrich G, Furet P, et al. Imatinib: a selective tyrosine kinase inhibitor. Eur J Cancer 2002;38:S19–27.
[143] Bourdon H, Trumpp-Kallmeyer S, Schreuder H, Hoflack J, Hibert M, Wermuth C-G. Modelling of the binding site of the human m1 muscarinic receptor: experimental validation and refinement. J Comput Aided Mol Des 1997;11(4):317–32.
[144] Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, et al. High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. Science 2007;318(5854):1258–65.
[145] Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, et al. Crystal structure of the [bgr]2 adrenergic receptor-Gs protein complex. Nature 2011;477(7366):549–55.
[146] Muyldermans S. Nanobodies: natural single-domain antibodies. Annu Rev Biochem 2013;82(1):775–97.
[147] Chapman HN, Fromme P, Barty A, White TA, Kirian RA, Aquila A, et al. Femtosecond X-ray protein nanocrystallography. Nature 2011; 470(7332):73–7.

CHAPTER

23

Physiological Aspects Determining the Pharmacokinetic Properties of Drugs

Koen Boussery[1], *Frans M. Belpaire*[2] *and Johan Van de Voorde*[3]

[1]Laboratory of Medical Biochemistry and Clinical Analysis, Faculty of Pharmaceutical Sciences, Gent University, Gent, Belgium; [2]Heymans Institute for Pharmacology, Gent, Belgium; [3]Vascular Research Unit, Ghent University, Gent, Belgium

OUTLINE

To explain all nature is too difficult a task for any one man or even for any one age. 'Tis much better to do a little with certainty, and leave the rest for others that come after you, than to explain all things. ***Isaac Newton***

I. INTRODUCTION

In order to produce its intended effect, a drug must be present at an appropriate concentration in the fluid surrounding the effect site, that is, the biophase. Only rarely can drugs be applied directly to the biophase. In most cases, drugs need to be transferred from the site of administration to the biophase. Usually, this translocation involves two steps: absorption and distribution. During absorption, the drug passes from its site of administration (e.g., the gastrointestinal (GI) tract when the drug is taken orally) into the systemic circulation. Subsequently,

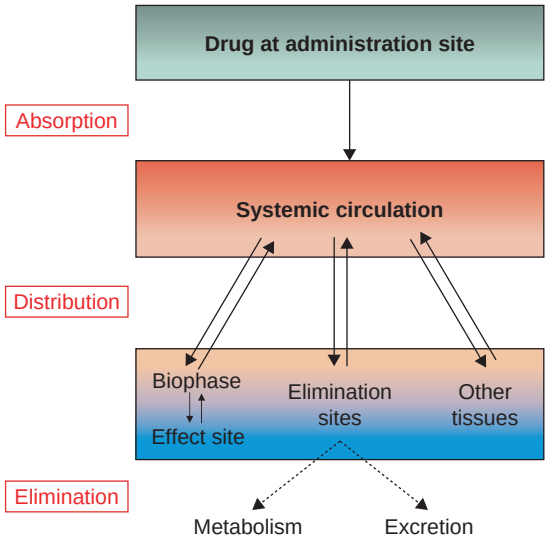


FIGURE 23.1 Schematic representation of drug absorption, distribution, and elimination.

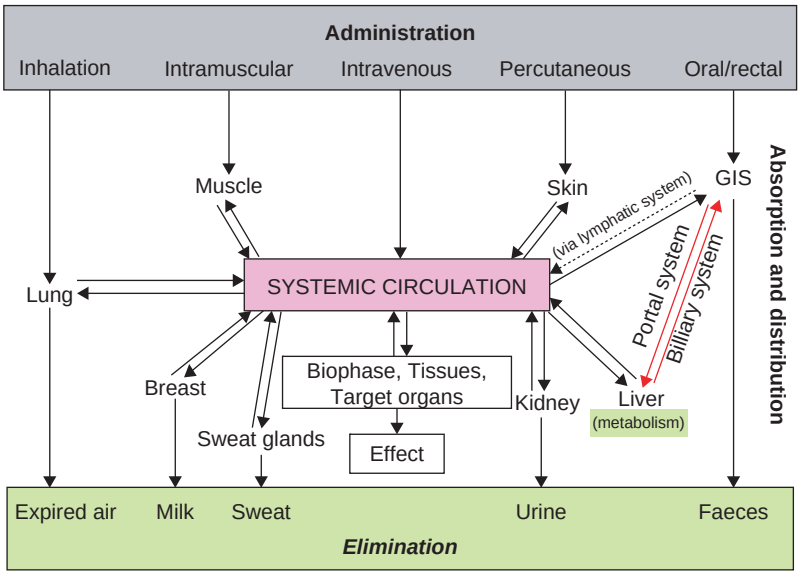


FIGURE 23.2 Schematic representation of the main routes of drug absorption, distribution, and elimination (GIS = GI system). The red arrows represent the enterohepatic cycle (see Section V.A.2.).

the drug is distributed via the circulating blood plasma (the fluid portion of the blood) to the different parts of the organism, including the organ(s) in which the biophase for the drug is localized. Each drug molecule that reaches the target site can add to the intended pharmacological effect of the drug. However, a portion of the drug molecules in the body is always distributed to organs and tissues that account for an irreversible loss of drug molecules from the body (drug elimination) by either biotransformation (the conversion of one chemical entity to another) or excretion. This causes a decrease in the concentration of the drug in the body and, consequently, also in the biophase. Figure 23.1 shows a schematic representation of the processes involved in the journey of a drug molecule through the human body. Figure 23.2 shows a more detailed scheme of the main routes of drug absorption, distribution, and elimination. Pharmacokinetics is the study of the drug concentrations in the different parts of the organism as a function of time. These concentrations depend on the dose administered and upon the rate and extent of absorption, distribution, and elimination.

In the first part of this overview (Sections II, III, IV, and V), some physiological aspects of drug absorption, distribution, and elimination will be discussed. The second part (Sections VI and VII) will briefly focus on some pharmacokinetic parameters and terminology, and on variability in pharmacokinetics. Due to its limited size, this chapter only includes some basic and general information. For more elaborate information, the reader is referred to some excellent textbooks in the further reading section.

II. PASSAGE OF DRUGS THROUGH BIOLOGICAL BARRIERS

On its journey through the body, a drug needs to cross different biological barriers. These barriers can be a single layer of cells (e.g., the intestinal epithelium), several layers of cells (e.g., in the skin), or the cell membrane itself (e.g., to reach an intracellular receptor). A drug can cross a cell layer either by traveling through the cells (transcellular drug transport) or through gaps between the cells (paracellular drug transport). The mechanisms by which a drug can cross the cell membrane will be discussed, together with transcellular drug transport.

A. Transcellular Drug Transport

In order to travel through a cell or to reach a target inside a cell, a drug molecule must be able to traverse the cell membrane. The cell membrane (also called plasma membrane) is a lipid bilayer interspersed with carbohydrates and proteins. Although cell membranes vary in their permeability characteristics depending on the tissue, the main mechanisms of drugs passing through the cell membrane are passive diffusion, carrier-mediated processes, and vesicular transport.

1. Passive Diffusion

Passive diffusion is the process by which molecules spontaneously diffuse from a region of higher concentration (e.g., outside of the cell) to a region of lower concentration (e.g., inside the cell), and it is the main mechanism for passage of drugs through membranes. Lipid-soluble drugs penetrate the lipid cell membrane with ease, and can pass the cell membrane by passive diffusion. Polar molecules and ionized compounds, on the other hand, partition poorly into lipids and are either unable to diffuse through the cell membrane or do so at a much lower rate. Also, large molecules, such as proteins and protein-bound drugs, cannot diffuse through the cell membrane.

Transmembrane diffusion is driven by the concentration gradient of the drug over the cell membrane. The rate of diffusion depends—apart from the lipid/water partition coefficient of the drug (P) and the concentration gradient ($C_{out}-C_{in}$)—on membrane properties such as the membrane area (A) and thickness (h), and the diffusion coefficient (D) of the drug in the membrane, according to Fick's Law (Equation 23.1).

$$\text{Rate of diffusion} = \frac{DAP(C_{out}-C_{in})}{h} \tag{23.1}$$

Many drugs are acidic or basic compounds that are ionized to a certain degree in aqueous medium. Their degree of ionization depends on their dissociation constant (pK_a) and the pH of the solution, according to the Henderson–Hasselbach Equation (Equations 23.2 and 23.3).

For acidic drugs:

$$\log\frac{\text{Ionized concentration}}{\text{Unionized concentration}} = \text{pH}-pK_a \tag{23.2}$$

For basic drugs:

$$\log\frac{\text{Unionized concentration}}{\text{Ionized concentration}} = \text{pH}-pK_a \tag{23.3}$$

Very weak acids with pK_a values higher than 7.5, are essentially unionized at physiological pH values. For these drugs, diffusion over the cell membrane is rapid and independent of pH changes within the body, provided the unionized form of the drug is lipid soluble. For acidic drugs with a pK_a value between 3.0 and 7.5, the fraction of unionized drug varies with the changes in pH encountered in the organism. For these drugs, the pH of the extracellular environment is critical in determining the diffusion across the cell membrane. For acidic drugs with a pK_a lower than 2.5, the fraction of unionized drug is low at any physiological pH, resulting in very slow diffusion across membranes. A similar analysis can be made for bases.

At the diffusion equilibrium, the concentrations of unionized molecules on both sides of a biological barrier are equal. If the pH on both sides of the barrier is equal, then the concentration of ionized molecules—and consequently the total concentration of the molecules—will be the same on both sides of the barrier. However, if there is a difference in pH, as for example between blood plasma (pH 7.4) and stomach contents (pH 1–3), the

concentration of the ionized molecules at equilibrium—and therefore the total concentration—will be much higher on one side of the barrier than on the other. This phenomenon is called "ion trapping."

2. *Carrier-Mediated Processes*

Many cell membranes possess specialized transport mechanisms that regulate entry and exit of physiologically important molecules and drugs. Such transport systems involve a carrier molecule, that is, a transmembrane protein that binds one or more molecules and releases them on the other side of the membrane. Such systems may operate passively (without any energy source) and along a concentration gradient; this is called "facilitated diffusion." However, facilitated diffusion seems to play only a minor role in drug transport. An example is the transport of vitamin B_{12} across the GI membrane. Alternatively, the system may spend energy obtained from the energy rich molecule adenosine triphosphate (ATP) required to pump molecules against a concentration gradient. This mechanism is called "active transport."

At high drug concentrations, the carrier sites become saturated, and the rate of transport does not further increase with concentration. Furthermore, competitive inhibition of transport can occur if another substrate for this carrier is present. In recent years, several transporters have been described to be present in various organs and tissues throughout the body and to determine absorption, distribution, and elimination of compounds that are substrates for these transporters. Although some transporters mediate the uptake of compounds in the cell (influx transporters), others may mediate secretion back out of the cell (efflux transporters). Transporters in the intestinal membrane affect the absorption of drugs (see Section III.C.), while transporters in the liver and kidney influence elimination by mediating transport into and out of cells responsible for biotransformation (hepatocytes; see Section V.B.) or excretion (e.g., renal tubule cells in the kidneys; see Section V.A.1.). Furthermore, efflux transporters may limit the penetration of compounds into certain areas of the body, such as the cerebrospinal fluid and blood cells. Chapter 26 of this textbook presents a more elaborate description of drug transport mechanisms.

3. *Vesicular Transport*

During vesicular transport, the cell membrane forms a small cavity that gradually surrounds particles or macromolecules, thereby internalizing them to the cell in the form of a vesicle or vacuole. Vesicular transport is the proposed process for the absorption of orally administered Sabin polio vaccine and of various large proteins. It is called endocytosis when moving a macromolecule into a cell, exocytosis when moving a macromolecule out of a cell, and transcytosis when moving a macromolecule across a cell.

B. Paracellular Drug Transport

Drugs can also cross a cell layer through the small aqueous contact points (cell junctions) between cells. This paracellular drug transport can be initiated by a concentration gradient over the cell layer (passive diffusion) or by a hydrostatic pressure gradient across the cell layer (filtration). The size and characteristics of cell junctions vary widely between different barriers to drug transport. For example, the endothelium of glomerular capillaries in the kidney (see Section V.A.1.a.) forms a leaky barrier that is very rich in intercellular pores. Therefore, this membrane is very permeable and permits filtration of water and solutes. On the other hand, endothelial cells of brain capillaries are sealed together by tight junctions, practically eliminating the possibility of paracellular drug transport.

III. DRUG ABSORPTION

Absorption can be defined as the passage of a drug from its site of administration into the systemic circulation. If a drug is administered directly into the systemic circulation by intravenous (i.v.) administration, absorption is not needed. Drugs can be administered by enteral and parenteral routes. Enteral administration occurs through the GI tract, by contact of the drug with the mucosa in the mouth (buccal or sublingual), by swallowing (oral), or by rectal administration. In parenteral administration the GI tract is bypassed; examples are the i.v. (direct injection into the systemic circulation) and intramuscular (injection into a muscle) routes. Drugs can also be absorbed through the skin or through the mucosa of various organs (e.g., bronchi, nose, vagina). In some cases, a drug is applied for a local effect, and no absorption is intended (e.g., antacids that neutralize stomach acid). In this chapter, we will describe drug administration by the oral route, which is the most common and popular route of

TABLE 23.1 Common Routes of Drug Administration

1. Parenteral routes
   - Intravenous bolus (i.v.)
     - Direct injection of complete dose into the systemic circulation.
     - Complete and instantaneous bioavailability; no need for absorption.
     - Often used for immediate effect.
     - Main disadvantages: technique requires extensive training; some complications may have serious consequences; formulation must be sterile.
   - Intravenous infusion (i.v.)
     - Similar to i.v. bolus, but dose is injected slowly into the systemic circulation at a constant rate (controlled by an infusion pump).
     - Plasma drug levels are more precisely controlled.
     - Larger fluid volumes can be injected.
   - Intramuscular injection (IM)
     - Injection of a dose into a muscle, from where it is absorbed due to the perfusion of the muscle by blood.
     - Easier than i.v. injection.
     - Rapid absorption from aqueous solution; slower from nonaqueous (oil) solutions.
     - Main disadvantages: irritating drugs may be very painful; different rates of absorption depending on muscle group injected and blood flow.
   - Subcutaneous injection (s.c.)
     - Injection of a dose into the s.c. tissue layer immediately beneath the skin.
     - Main disadvantage: drug absorption is relatively slow and depends on local blood flow; s.c. tissues are often adipose and poorly perfused.
     - Used for insulin replacement therapy in diabetic patients.
2. Enteral routes
   - Buccal or sublingual (SL) drug delivery
     - A drug formulation is brought in close contact with the mucous membranes inside the mouth (lining the cheeks (buccal) or under the tongue (SL)).
     - No first-pass effects.
     - Mainly passive diffusion; only small lipophilic drugs are absorbed.
     - Main disadvantage: some drug may be swallowed.
   - Oral drug delivery
     - A drug formulation is swallowed; absorption from the GI tract.
     - Absorption may vary in rate and extent.
     - Safest and easiest route of drug administration.
     - Main disadvantage: some drugs may have erratic absorption, be unstable in the GI tract, or be metabolized by liver prior to systemic absorption.
   - Rectal drug delivery
     - Absorption from suppository may vary; more reliable absorption from enema (solution).
     - Useful when patient cannot swallow medication (e.g., elderly and very young patients; vomiting patients).
     - Used for both local and systemic effects.
     - First-pass metabolism in the liver is partly avoided.
     - Main disadvantages: absorption may be erratic; suppository may migrate to different position; some patient discomfort.
3. Other routes
   - Transdermal/percutaneous drug delivery
     - Generally, drug absorption through the skin is slow; absorption can be increased by occlusive dressing.
     - Permeability of skin varies with condition, anatomic site, age, and gender.
     - Easy to use (e.g., patches).
     - First-pass metabolism in the liver is avoided.
     - Main disadvantage: possible irritation of the skin by patch or drug.
   - Intranasal drug delivery
     - Primarily used for local effects, but also developed as a route for systemic effects: absorption through the nasal mucosa.
     - First-pass metabolism in the liver is avoided.
     - Especially attractive for the delivery of peptides (e.g., desmopressin).
   - Pulmonary drug delivery/inhalation
     - Primarily used for local effects: particle size of drug determines deposition site in respiratory tract.
     - Inhaled drugs can be absorbed from their deposition site in various parts of the respiratory tract (large surface area + blood supply).
     - Main disadvantages: may stimulate cough reflex; some drug may be swallowed.

drug dosing. Some characteristics of other common routes of drug administration are listed in Table 23.1. For more details on these various routes, the reader is referred to the further reading list at the end of this chapter.

The enteral system consists of the GI tract from the mouth to the anus. With respect to drug absorption after oral dosing, the stomach, the small intestine, and the large intestine (or colon) are the major components (Figure 23.3). The small intestine includes the duodenum, jejunum, and ileum. These segments differ from one

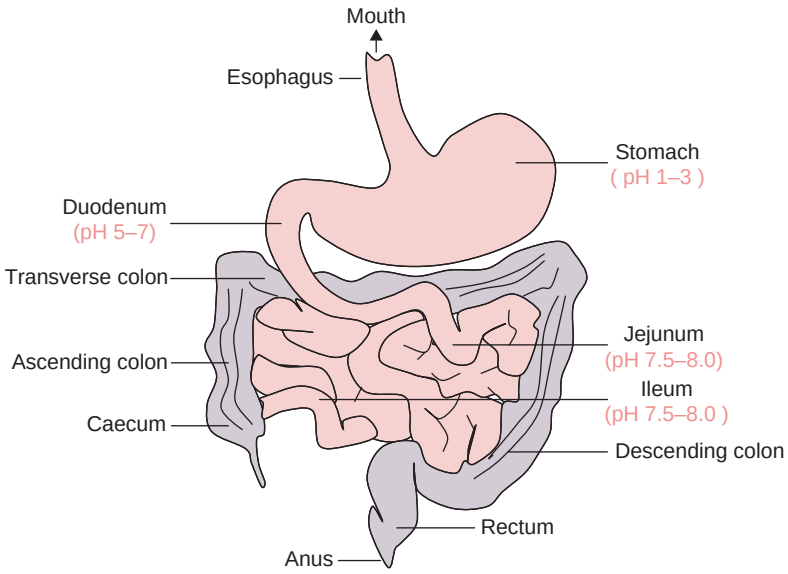


FIGURE 23.3 Schematic representation of the GI tract.

another anatomically and morphologically, as well as with respect to secretions and pH. As orally administered drugs move through the GI tract, they encounter environments that vary in pH (Figure 23.3), enzyme composition, and fluidity of contents, as well as in the surface area available for absorption. The most important area for drug absorption is the duodenum, the upper portion of the small intestine. The small intestine is also the region that is responsible for almost all digestion and absorption of nutrients, and its structural adaptation to this task makes it very suitable for the absorption of drugs. Its length alone provides a large surface area for absorption, and that area is further increased by circular folds, villi (finger-like projections of the intestinal wall), and microvilli (small projections of villi). If the small intestine is imagined as a hollow cylinder, its net increase in total surface area due to these folds, villi, and microvilli is 600-fold (versus that of a smooth cylinder with the same length). The total surface area of the human small intestine is approximately 200 m^2, or the surface area of a doubles tennis court.

The main determinants of the rate and extent of absorption after oral administration are:

- the dosage form of the drug;
- GI motility and gastric emptying;
- GI permeability to the drug;
- perfusion of the GI tract; and
- the first-pass effect.

A. Dosage Form of the Drug

For a drug to be absorbed from the GI tract, it has to be dissolved in the aqueous medium of the stomach and the intestine. However, many drugs are administered as tablets or capsules that first have to disintegrate to release the drug. Therefore, liquid dosage forms are, in general, more rapidly absorbed than solid forms. Disintegration can be controlled, and various drug products have been modified to alter the timing of the release of the active drug from the drug product. The term "controlled release" is used for various types of oral extended release rate dosage forms (e.g., sustained release, prolonged release) and delayed release rate dosage forms (e.g., enteric coated). One example of the systems for controlled release is the osmotic pump, in which drug delivery is driven by an osmotically controlled device that pumps a constant amount of water through the system, dissolving and releasing a constant amount of drug per time unit.

B. GI Motility and Gastric Emptying

Once a drug is given orally, GI motility tends to move the drug through the GI tract from mouth to anus. The drug rapidly reaches the stomach, which subsequently empties its contents into the small intestine. The

residence time of the drug in the stomach varies from a few minutes to several hours and is dependent on a range of factors, such as the volume, viscosity, and composition of the stomach content. The surface area of the stomach is limited in comparison with the small intestine, and under normal conditions gastric emptying is rapid. Therefore, the stomach's role in drug absorption is—in general—rather modest. Factors that influence gastric emptying can, however, influence the absorption rate of most drugs, but not necessarily the total amount of drug eventually absorbed. For example, consumption of a meal (especially a meal high in fat) reduces gastric emptying. Therefore, a drug taken with food will stay longer in the stomach, which could reduce the absorption rate of that drug. The motility of the intestine can also influence the absorption. It mixes the contents of the duodenum, bringing the drug in close contact with the intestinal wall (the biological barrier the drug needs to cross to be absorbed). When intestinal motility increases, disintegration of the formulation and dissolution of the drug are often accelerated. On the other hand, high motility of the intestinal tract (as with diarrhea) can result in a very short residence time in the small intestine and less opportunity for adequate absorption as a consequence.

C. GI Permeability to the Drug

Once a drug is dissolved in the aqueous medium in the GI tract, it can pass into the capillaries of the GI wall. The drug needs to have a reasonable amount of lipid solubility to enable absorption across the lipid membrane, but a reasonable amount of water solubility is also necessary to dissolve in the GI system. A drug that is too lipid soluble will exhibit impaired dissolution in the GI system, which will lead to reduced absorption. On the other hand, a drug that is too water soluble will have adequate dissolution but reduced absorption, due to its inability to cross the lipid membrane. General rules for the intestinal absorption of a drug include:

- Small amphiphatic drugs move efficiently through the transcellular route by partitioning into and out of lipid bilayers.
- Small hydrophilic drugs are restricted to the paracellular route or to aqueous routes that normally absorb nutrients, vitamins, or cofactors.
- Peptides and proteins are poorly absorbed intact and require the application of enhancing agents or special uptake mechanisms.
- In general, the permeability for a drug decreases along the intestine, but this is obviously very dependent on the drug and the route of transport.

For acids and bases, only the non-ionized molecules can be absorbed. At all physiological pH values, weak acids and bases exist mostly in the un-ionized form and can be absorbed as well from the stomach as from the intestine. In theory, weakly acidic drugs are better substrates for passive diffusion at the pH of the stomach than at that of the intestine. However, the limited residence time of the drug in the stomach and the relatively small surface area of the stomach more than balance the influence of pH in determining the optimal site of absorption. Strong bases such as the quaternary ammonium compounds are, to a large extent, ionized at all physiological pHs and are hardly absorbed at all.

For many years, the rate and extent of absorption in the small intestine were thought to be determined solely by the lipid/water solubility and membrane permeability characteristics of the drug. While this relatively simplistic model worked for many drugs, there are a number of exceptions to this rule, suggesting other forces are at work within the GI system to control the absorption of drugs. It is now known that a complex system of transporter proteins and metabolic enzymes is present within the GI system. Expression of influx transporters in the intestinal epithelial cells can increase absorption of drugs that are substrates for these transporters, whereas efflux transporters can reduce oral absorption of these drugs. In particular, the impact of the P-glycoprotein multidrug transporter (P-gp) on drug absorption has been studied extensively. Since P-gp is located on the epithelium of intestinal cells, it can act as a counter-transport pump that transports its substrates back into the intestinal lumen as they begin to be absorbed across the intestinal wall.

D. Perfusion of the GI Tract and the First-Pass Effect

The villi in the duodenal region are highly perfused with a network of capillaries and lymphatic vessels. The capillaries in the villi are fenestrated (i.e., they have large pores) and have a large surface area, so drugs absorbed from the small intestine can easily enter these capillaries. The drug is then transported to the portal vein and to

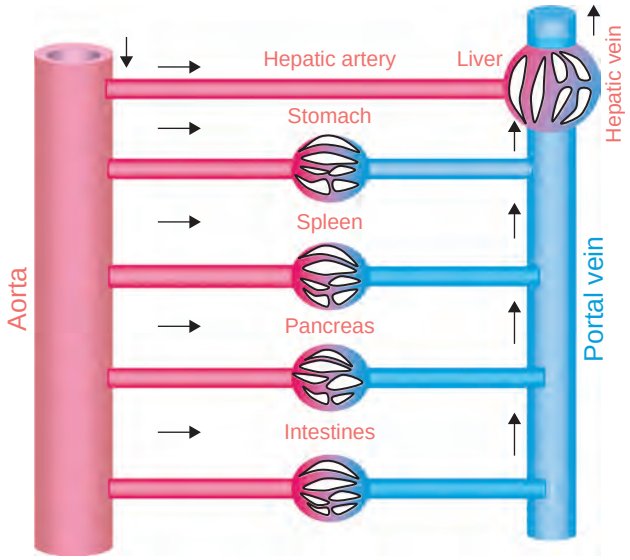


FIGURE 23.4 Schematic representation of the splanchnic circulation.

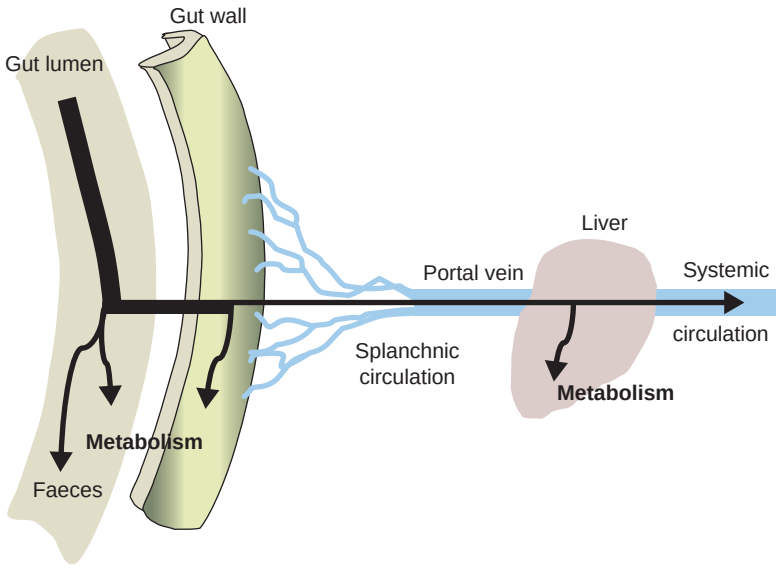


FIGURE 23.5 The first-pass effect: an orally administered drug must pass through different potential sites of elimination before it reaches the systemic circulation.

the liver prior to reaching the systemic circulation. Figure 23.4 shows a schematic representation of the splanchnic circulation, which includes the blood flow through the stomach, small and large intestines, pancreas, spleen, and liver. This splanchnic circulation receives about 28 percent of the cardiac output via the abdominal aorta, a fraction that is significantly elevated for two to four hours after a meal. Any change in blood flow to the GI tract will affect the rate of drug absorption from the intestinal tract.

As is clear from Figure 23.4, the liver receives most (approximately 75 percent) of its blood supply through the portal vein, which carries the venous blood draining from all of the organs in the splanchnic circulation except the liver itself. As a consequence, drugs that are given orally first pass through the liver before being distributed to the rest of the body (i.e., before entering the systemic circulation). This is an important issue for some drugs that are highly metabolized by the liver. When administered orally, a substantial fraction of these drugs will be metabolized before reaching systemic circulation. Such a loss when a drug passes through sites of elimination during absorption is known as a first-pass effect. Besides the hepatic first-pass effect, biotransformation during absorption can also occur in the lumen of the intestine and by enzymes that are present in the gut wall

(Figure 23.5). For example, CYP3A—a major subfamily of Phase I drug-metabolizing enzymes in humans (see Section V.B.)—has been shown to be present at high levels in the intestinal wall, which limits oral drug delivery of its substrates.

A first-pass phenomenon may also occur after intraperitoneal and, partially, after rectal administration. It does not occur for parental routes of administration or after buccal or sublingual administration. For some orally administered drugs with high lipid solubility, absorption via the lymphatic system is also possible. Drugs are absorbed through the lymphatic vessels under the microvilli. Absorption of drugs through the lymphatic system bypasses the hepatic first-pass effect, because drug absorption through the portal vein is avoided (the lymph delivers these substances to the systemic circulation via the thoracic duct).

IV. DRUG DISTRIBUTION

After absorption into the systemic circulation, drugs are distributed to the various organs and tissues in the body. The blood plasma carries the drug molecules to the effect site for drug action, as well as to other tissues where side effects or adverse reactions may occur. The rate and extent of distribution depend on blood flow to different organs, tissue size, the binding of drugs to plasma proteins and tissue components, and the permeability of tissue membranes. The latter factor is related to the physicochemical properties of the drug, as described above (see Section II). For lipid-soluble drugs, tissue membranes represent no barrier, and distribution depends essentially on the perfusion rate of the tissue. For these drugs, rapid equilibration occurs between blood and tissues such as lungs, kidney, liver, heart, and brain (i.e., organs with a high blood flow). Less rapid equilibration is found for skeletal muscle, bone, and adipose tissue, which receive a considerably smaller volume of blood per unit mass. This is called "perfusion limited distribution," since blood flow is the rate-limiting step in the distribution of the drug. In contrast, if drug distribution is limited by the slow passage of drug across the membrane in the tissue, this is called "permeability limited distribution." Tissue uptake of a drug continues until equilibrium is reached between the diffusible form of the drug in the tissue and the blood (i.e., until the free concentrations in blood plasma and tissue water are equal). Drugs can be present in tissues in higher concentrations than in blood plasma as a consequence of pH-gradients but mainly because of a high affinity for that particular type of tissue. This is called drug accumulation. On the other hand, drugs can be present in high concentrations in blood plasma due to a high plasma protein binding.

A. Plasma Protein Binding

Many drugs are bound to some extent to plasma proteins. It may be important to know to what extent a certain drug is bound to plasma proteins, since a protein-bound drug is a large complex that cannot easily cross the biological barrier and therefore has a restricted distribution. Furthermore, the protein-bound drug is usually pharmacologically inactive. The plasma protein binding is expressed as "fraction bound," that is, the ratio of bound concentration over total (bound plus free) concentration, or as "percentage bound" if this value is multiplied by 100. The free fraction equals one minus the bound fraction. Many acidic drugs bind to albumin, the major component of plasma proteins responsible for reversible drug binding (normal plasma concentrations for albumin range from 35 to 40 g/L). α_1-Acid glycoprotein is an acute phase reactant, a group of plasma proteins that changes in concentration following tissue injury or inflammation. It primarily binds to basic drugs such as propranolol and imipramine. The plasma concentration of α_1-acid glycoprotein is low (0.4−1 g/L), but its concentration in plasma rises in inflammation. Binding to other macromolecular components in the blood (including lipoproteins, immunoglobulins, and erythrocytes) generally occurs to a much smaller extent.

For most drugs, the binding of drugs to plasma proteins is a reversible process with extremely rapid rates of association and dissociation that can be described by the law of mass action. The degree of binding is determined by affinity (expressed as the association constant), capacity (the number of binding sites per molecule protein), protein concentration, and drug concentration. At therapeutic drug concentrations, usually only a small fraction of the available binding sites is occupied. For a given protein concentration, the free fraction of the drug is then rather constant and independent of drug concentration. In some instances, the drug concentrations are so high that most binding sites are occupied, and the free fraction becomes concentration-dependent. Concentration-dependent changes in drug binding are most likely to occur with drugs that have a high affinity for the proteins and that are given in large doses, for example, acetylsalicylic acid, phenylbutazone, some penicillins, and cephalosporins.

The plasma protein binding of drugs is altered in some physiological and pathological conditions, often as a result of changes in plasma protein concentration or as a result of competition for common binding sites by another (endogenous or exogenous) compound. In various disease states (such as renal failure, liver disease, inflammation), in pregnancy, and in the neonatal period, hypoalbuminemia is observed; α_1-acid glycoprotein concentrations rise in inflammatory diseases, stress, and malignancy, and fall in liver disease. Free fatty acids bind strongly to albumin. When their concentration in plasma increases due to fasting, exercise, or infection, albumin-bound drugs can be displaced from their binding sites. In renal failure, waste products that accumulate in the blood may compete for plasma protein binding. Besides endogenous compounds, other administered drugs may compete for plasma protein binding. Such an interaction is to be expected when the "displacer" is present in the same concentration range as the binding sites at the proteins. This situation may result in a decrease of the binding sites available for the "displaced" drug. The changes in actual free plasma concentration will always be smaller than the changes in free fraction because of redistribution of the displaced drug to the tissues and its more rapid elimination.

B. Drug Accumulation

Drugs may accumulate in body tissues because of a high affinity for that particular type of tissue. For example, drugs with a high fat/water partition coefficient are very fat soluble and tend to accumulate in body fat. Accumulation in body fat is important for only a few drugs, mainly because the fat/water partition coefficient is relatively low for most drugs. Morphine, for example, though lipid soluble enough to cross the blood–brain barrier, has a fat/water partition coefficient of only 0.4, and sequestration of the drug by body fat is of little importance. With thiopentone, on the other hand (fat/water partition coefficient approximately 10), accumulation in body fat is considerable. Accumulation of drugs in body fat is also limited by a low blood supply to body fat—less than 2 percent of the cardiac output. As blood flow to body fat is limited, drugs are delivered to body fat slowly, and the equilibrium distribution between fat and body water is also approached slowly. As a result, accumulation in body fat is only important when lipid-soluble drugs are given chronically (e.g., benzodiazepines). Only for highly lipid-soluble drugs (e.g., general anesthetics), partition into body fat is also important at first dose. Drugs may also accumulate in tissues by binding reversibly or irreversibly to tissue constituents. For example, tetracyclines, a class of antibiotics, bind with calcium to form an insoluble chelate and therefore accumulate irreversibly in growing bones and teeth.

The body tissues in which a drug accumulates are potential reservoirs for the drug. If a stored drug is in equilibrium with that in plasma and is released as the plasma concentration declines, drug concentration in plasma and at the biophase is sustained, and pharmacological effects of the drug are prolonged. However, if the reservoir for the drug has a large capacity and fills rapidly, a larger initial dose is required to reach a therapeutically effective concentration at the biophase after the first administration.

C. The Blood–Brain Barrier

Delivery of drugs to the brain by the systemic circulation is difficult due to the presence of the so-called blood–brain barrier. This barrier acts as a self-defense mechanism by preventing the passage of many potentially harmful substances from blood into brain tissue. It is formed by the brain capillary endothelial cells that are sealed together by tight junctions and closely surrounded by processes of large numbers of astrocytes (a type of supporting cell in the brain), thereby eliminating the possibility of paracellular transport. Furthermore, different efflux transporters are present that remove drugs from the brain and transfer them to the systemic circulation. This is why brain penetration of most drugs is markedly restricted. The pathways for drug uptake in the brain are mainly limited to active transport and simple diffusion. As a consequence, only molecules that are either a substrate for an influx transporter or are highly lipid soluble and of relatively low molecular weight can cross the blood–brain barrier. A challenging area of research in drug delivery attempts to make use of influx transporters to develop blood–brain barrier vectors that can improve drug uptake in the brain.

V. DRUG ELIMINATION

Drug elimination refers to the irreversible removal of drug from the body and covers both excretion (i.e., disappearance of unchanged drug from the body) and biotransformation (the process by which the drug is biochemically converted to a metabolite). Excretion of drugs through sweat and tears is quantitatively unimportant. The

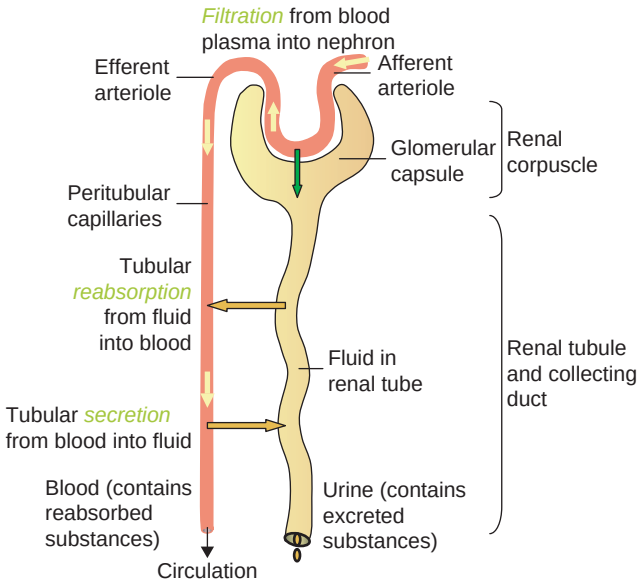


FIGURE 23.6 Schematic representation of the renal excretion of drugs.

concentration of some drugs in saliva parallels that in plasma. Therefore, saliva is sometimes a useful biological fluid to determine drug concentrations. However, this is not a real route of excretion since drugs in saliva are swallowed. The excretion of drugs and toxic compounds in breast milk is of importance in relation to their potential (toxic) effect in the nursing infant. The fact that molecules can also be excreted through the loss of hair, nails, and skin is of toxicological and forensic significance, as sensitive methods can detect traces of toxic metals, for example, in hair (arsenic and mercury). General anesthetics, on the other hand, often leave the body in expired air. However, the major routes of drug elimination are renal excretion and hepatic biotransformation. This chapter will discuss these two processes, as well as biliary excretion.

A. Excretion

1. Renal Excretion

The kidneys function as a filter, aiming to clear metabolic products and toxins from the blood and to excrete them through the urine. Efficient clearance is promoted by a high blood flow to the kidneys (20 percent of the total body blood flow for only 0.5 percent of the total body weight).

The basic functional unit of the kidney is the nephron (Figure 23.6). Blood arriving in the kidney is first filtered in the glomerulus of the nephron. The primary urine formed by filtration flows from the glomerular capsule through the renal tubule and the collecting ducts. The urine drains from the collecting ducts in the renal pelvis and through the ureters in the bladder. The composition of the urine is unaltered after leaving the nephron. Blood that is not filtered in the glomerulus flows through peritubular capillaries situated along the renal tubule. This allows exchange of molecules between the blood and the fluid in the renal tubule (Figure 23.6).

The renal urinary excretion is the ultimate result of three processes: glomerular filtration, tubular reabsorption, and active tubular secretion (Figure 23.6).

A. GLOMERULAR FILTRATION

Blood flow to the kidneys is about 1.2–1.5 L/min. About 10 percent of this volume is filtered through the pores in the glomeruli, which amounts to a filtrate (primary urine) of about 125 mL/min or 180 L per 24 h. The pores of the glomerulo-capillary membrane are sufficiently large to permit passage of small molecules and most drug molecules, but do not allow passage of blood cells and of large molecules (>60 kDa), such as plasma proteins. Therefore, drug molecules bound to plasma proteins are not eliminated by glomerular filtration.

B. TUBULAR REABSORPTION

More than 99 percent of the original 180 L of protein-free filtrate is reabsorbed via the tubular cells; only about 1.5 L (per 24 h) is excreted as final urine. Solutes and drugs dissolved in the filtrate can also be reabsorbed. Glucose, for example, is filtered from the blood but is completely reabsorbed in the renal tubule by carriers in the tubular cells. For different drugs, tubular reabsorption varies from almost absent to almost complete. For most drugs, reabsorption is a passive process (passive diffusion). If the tubular wall is freely permeable for the molecule, more than 99 percent of the filtered molecule will be reabsorbed passively. Drugs with high lipid solubility—and hence high tubular permeability—are therefore excreted slowly. The drugs diffuse from tubular fluid to plasma in accordance to their concentration gradient, lipid/water partition coefficient, degree of ionization, and molecular weight. The pH of the urine varies between 4.5 and 7.0, and changes in pH can influence passive reabsorption and thus the excretion of the drug (see the Henderson−Hasselbalch Equation in Section II.A.). Acidifying the urine favors the reabsorption of weak acids, such as salicylates, and retards their excretion, whereas the reverse is true for weak bases. Alkalinization of the urine increases the excretion of weak acids. For example, it is possible to accelerate the excretion of phenobarbital (a weak acid) in an intoxicated patient by administration of sodium bicarbonate. On the other hand, the urinary excretion of weak bases is low in alkaline urine.

Increased urine flow by forced intake of fluids or co-administration of a diuretic drug can increase the excretion of some drugs by decreasing the time for drug reabsorption.

C. ACTIVE TUBULAR SECRETION

Considering most blood (90 percent) leaves the glomerulus unfiltered, most of the drug delivered to the kidneys reaches the peritubular capillaries. Here, drugs can be transferred to the tubular lumen by relatively non-selective carrier systems. These carriers can transport molecules against a concentration gradient from the blood capillaries across the tubular membranes to the tubular fluid. There are at least two active renal secretion systems: one that normally secretes naturally occurring organic acids (such as uric acid) and one that secretes naturally occurring organic bases (such as choline or histamine). Acid drugs such as penicillins, indomethacin, and glucuronides are transported by the first system, whereas the second system transports bases such as morphine, procaine, and quaternary ammonium compounds. These transport systems can be saturated, and competition for the active transport systems can occur, leading to desirable or undesirable drug interactions. This characteristic has been used to decrease the urinary excretion of penicillin (and thereby prolonging its effect) by co-administering probenicid, another weak organic acid that competes for the acid transport system in the tubulus. Also, the P-glycoprotein multidrug transporter is present in the brush border of the renal tubules and can play a role in the active tubular secretion of exogenous substances. It is involved in tubular secretion of, for example, digoxin, and can be inhibited by quinidine or verapamil. Co-administration of quinidine therefore decreases the renal clearance (Cl_R) of digoxin, leading to an increase in digoxin serum concentrations.

Plasma protein binding does not limit the rate of active tubular secretion, as the affinity of the drugs is much higher for the carrier than for the plasma proteins. Tubular secretion is potentially the most effective mechanism for the elimination of drugs by the kidney. Penicillin, for example—though about 80 percent protein bound and therefore slowly cleared by glomerular filtration—is almost completely removed from blood by passage through the kidney due to efficient tubular secretion.

2. *Biliary Excretion*

The liver can also be considered an excretory organ. The liver is responsible for the formation of the bile fluid that drains in the gut and is (at least in part) removed along with the feces. The brown color of the feces is due to pigments of the bile. Bile secreted by the hepatocytes of the liver enters bile canaliculi (narrow intercellular canals) that empty into bile ducts that drain the bile fluid into the gallbladder. The gallbladder stores and concentrates the bile. When the smooth muscle cells in the gallbladder contract, the bile is delivered in the small intestine.

Each day hepatocytes secrete about 1 L of bile, consisting mostly of water, ions, bile salts (important for the absorption of lipids), cholesterol, and bile pigments. Bile formation by hepatocytes requires the active secretion of inorganic and organic solutes into the canalicular lumen, followed by the movement of water. Other solutes can be carried with this movement of water through the tight junctions between hepatocytes.

Some drugs are actively secreted into the bile and pass as such into the intestine. In humans, the molecular weight threshold for appreciable biliary excretion is in the order of 400−500 Da. In order to be excreted into the bile, drugs usually require a strong polar group. Many drugs excreted into bile are metabolites, often glucuronide conjugates.

A drug (and/or its metabolites) entering the intestine through bile may be excreted in the feces. However, it can also be reabsorbed from the intestine and thus undergo "enterohepatic cycling" (cf. Figure 23.2). Drug conjugates (e.g., glucuronides) can be hydrolyzed in the gut by bacteria, resulting in the liberation and reabsorption of the parent drug. In particular, this has been found for chloramphenicol and for steroids. These compounds may undergo extensive biliary cycling, with final excretion by the kidney.

B. Biotransformation

As was described above, most drugs need to pass through biological membranes to reach their site of action. Therefore, most drugs have lipophilic characteristics and are only partially ionized at the pH values encountered in the organism. These characteristics also favor reabsorption from the renal tubules after glomerular filtration (see Section V.A.1.c.). As a consequence, renal excretion often plays only a modest role in the total elimination of therapeutic agents from the body. For these compounds, biotransformation into metabolites that are more hydrophilic in nature is essential, because it allows their excretion by the kidneys. Biotransformation reactions take place mainly in the liver (hepatic biotransformation) but can also occur in intestinal mucosa, lungs, kidneys, skin, placenta, and plasma. Within a given cell, most of the biotransformation enzymes are found in the endoplasmic reticulum (a network of folded membranes inside the cell) and the cytosol (the fluid inside the cell). When liver tissue (or any other tissue) is homogenized and differential centrifugation is applied, the endoplasmic reticulum of the cells breaks up. Fragments of the endoplasmic reticulum then form microvesicles called microsomes. Therefore, the biotransformation enzymes in the endoplasmic reticulum are often referred to as microsomal enzymes.

For hepatic biotransformation to occur, a drug must enter the hepatocytes that contain the biotransformation enzymes. Polar molecules do this more slowly than nonpolar molecules, except where specific transport mechanisms exist. As a consequence, hepatic metabolism is in general more important for lipid-soluble drugs than for polar drugs. Renal excretion and biotransformation can therefore be regarded as two additional and synergistic elimination pathways, guaranteeing an efficient elimination of a whole range of hydrophilic as well as lipophilic substances from the body. Biotransformation usually inactivates a drug, but in some cases metabolites with biological activity or toxic properties are formed. For some drugs, the activity may reside wholly in one or more metabolites. Drugs that only become active after biotransformation are termed "prodrugs." Prodrugs are sometimes developed to improve absorption, often because of better lipid solubility than the active metabolite. After absorption, these prodrugs are rapidly converted to the active metabolite in the gut wall or in the liver. An example is pivampicillin, an ester of ampicillin, which is rapidly and completely hydrolyzed to ampicillin during absorption.

Two phases can be distinguished in the pathways of biotransformation. Phase I involves addition of functionally reactive groups by oxidation, reduction, or hydrolysis. These products are sometimes more chemically reactive and therefore more toxic than the parent drug. The Cytochrome P450 (CYP) enzyme family is involved in most (but not all) Phase I reactions. It comprises a large group of enzymes localized in the endoplasmic reticulum of numerous tissues. The CYP enzymes are grouped in families denoted by an Arabic numeral (e.g., the CYP3 family), within which the amino acid sequence homology is higher than 40 percent. Each P450 family is further divided into subfamilies denoted by a capital letter (e.g., the CYP3A subfamily), with greater than 55 percent amino acid sequence homology. Finally, another Arabic numeral represents the individual enzyme (e.g., CYP3A4). The main CYP enzymes involved in drug metabolism are: CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP2D6, CYP2E1, and CYP3A4. CYP3A4 accounts for 30 percent of total P450 enzyme in the liver and is clinically the most important isoenzyme present in the liver. Nearly 50 percent of all clinically used medications are metabolized by CYP3A4.

Phase II biotransformation consists of conjugation of reactive groups present either in the parent molecule or after Phase I transformation. Phase II conjugates are usually pharmacologically inactive. They are more hydrophilic than the parent compounds and are easily excreted by the kidneys or the bile. Phase I and Phase II reactions often, though not invariably, occur sequentially. Phenytoin, for example, is first hydroxylated by a Phase I reaction and subsequently conjugated with glucuronic acid. Enzymes involved in Phase I reactions are located primarily in the endoplasmic reticulum, while the Phase II conjugation enzyme systems are mainly cytosolic. For more in-depth information on biotransformation and on biological factors influencing biotransformation, the reader is referred to Chapter 24 of this textbook. For further discussion on toxic metabolites, see Chapter 25.

VI. SOME PHARMACOKINETIC PARAMETERS AND TERMINOLOGY

A. Plasma Concentration–Time Curve

As was described earlier, a drug often reaches its site of action after it is absorbed into the systemic circulation and distributed to both the target tissue and other tissues. Ideally, the drug concentrations in the target tissue should remain above the minimum effective concentration (the lowest concentration that results in the desired pharmacological response) as long as the pharmacological effect is desired, whereas drug concentrations in all tissues should remain below the minimum toxic concentration (the lowest concentration that results in a toxic effect) at all times. As it is often very difficult—if not impossible—to measure drug concentrations in the target tissue, this technique is rarely used to ascertain whether the drug reaches the target tissue at a proper concentration. Instead, the pharmacokinetics of a drug is assessed by measuring drug concentrations at an alternative and more accessible site—the plasma.

A plasma concentration–time curve can be obtained by measuring the drug concentration in plasma samples taken at various time intervals after a drug is administered. These concentrations are then plotted against the corresponding time at which the plasma sample was taken. The profile of such a concentration–time curve is, in fact, determined by the complex interplay between the processes described earlier in this chapter: absorption, distribution, and elimination of the drug and, more precisely, by the rate at which these processes occur. Usually, (but not always) absorption, distribution, and elimination are assumed to be first-order processes, meaning their rate at all times is proportional to the amount of drug involved (e.g., the rate of first-order absorption is at all times proportional to the amount of drug present at the absorption site). As a consequence, concentration–time curves often show an exponential profile.

When a drug is administered as an i.v. bolus, the entire dose of the drug is injected straight into the blood. Therefore, the absorption process is considered to be completed immediately, and the concentration–time profile of the drug in plasma will be determined by the rate of distribution and elimination. When the distribution of the drug is very fast, the plasma concentration–time curve is determined only by the elimination rate and shows a mono-exponential (first-order) decline (a theoretical example is shown in Figure 23.7a; Figure 23.7b shows the same data in a semi-logarithmic graph). For many drugs, however, distribution occurs more slowly and contributes to the profile of the plasma concentration–time curve. After i.v.-bolus administration of such a drug, the plasma concentration–time curve declines bi-exponentially as the sum of two first-order processes: distribution and elimination. A theoretical example of such a plasma concentration–time profile is presented in Figure 23.8a (on a linear scale) and Figure 23.8b (on a semi-log scale). Note the bi-exponential decline in the plasma concentration in Figure 23.8b (both distribution and elimination contribute to the profile), in contrast to the mono-exponential decline in Figure 23.7b (very fast distribution; therefore only elimination contributes to the profile). The plasma concentration–time curve in Figure 23.8b may be divided into two parts: a distribution phase and an elimination phase. The distribution phase is the initial more-rapid decline that is mainly due to the distribution of drug from plasma to the tissues. Once equilibration between the drug concentrations in plasma and in tissues has occurred, both plasma and tissue concentrations decline in parallel due to elimination. This decline is often referred to as the elimination phase.

When a drug is not injected directly into the plasma, drug absorption from the site of administration also adds to the profile of the plasma concentration–time curve (besides distribution and elimination). Figure 23.9a shows an idealized example of a plasma concentration–time curve after a single oral administration of a drug. Initially, drug concentration at the absorption site is high, and the rate at which the drug is absorbed into the systemic circulation exceeds its rate of elimination from the body. Therefore, the drug concentration in plasma rises and the drug can be distributed to the tissues. As the drug is absorbed into the systemic circulation, its rate of absorption decreases (due to a decrease of the drug concentration at the absorption site), whereas its rate of elimination increases (due to the increase in plasma concentration). As a consequence, the difference between these rates diminishes. However, as long as the rate of absorption exceeds that of elimination, the plasma concentration continues to rise. This rising portion of the curve is often called the absorption phase. At peak concentration, both rates are equal. Afterwards, the rate of drug elimination exceeds the rate of its absorption, and the concentration of drug in both the plasma and the tissues starts to decline. This declining part of the curve is often called the elimination phase.

The time to reach the peak plasma concentration (t_{max}) is a rough marker for the rate of drug absorption, whereas the peak plasma concentration (C_{max} or maximum drug concentration in the plasma) itself is related to the dose, the rate of absorption, and the rate of elimination. The area under the curve (AUC) is related to the total amount of drug that reaches systemic circulation. Figure 23.9b shows the same data on a semi-log scale.

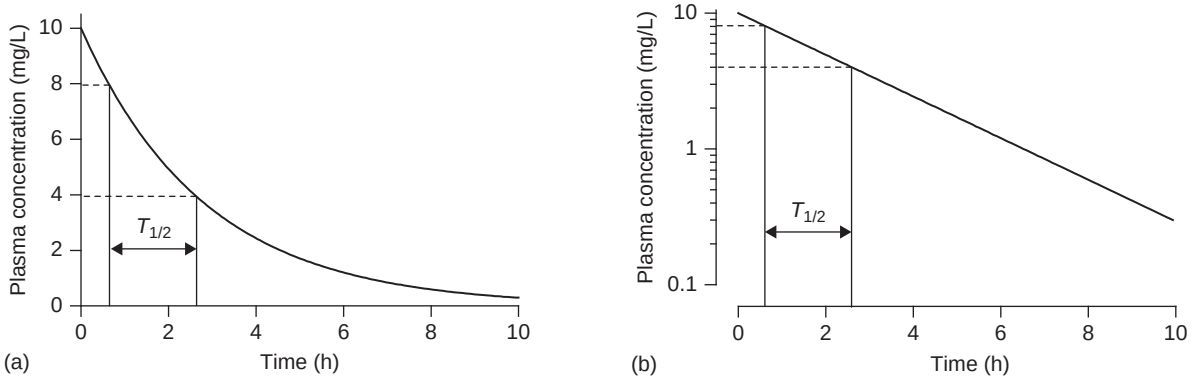


FIGURE 23.7 Plasma concentration–time curve after i.v. administration of an imaginary drug with a very high distribution rate: (a) linear scale; (b) semi-log scale. $T_{1/2}$ is the elimination half-life as derived from the plasma concentration–time curve (see Section VI.D).

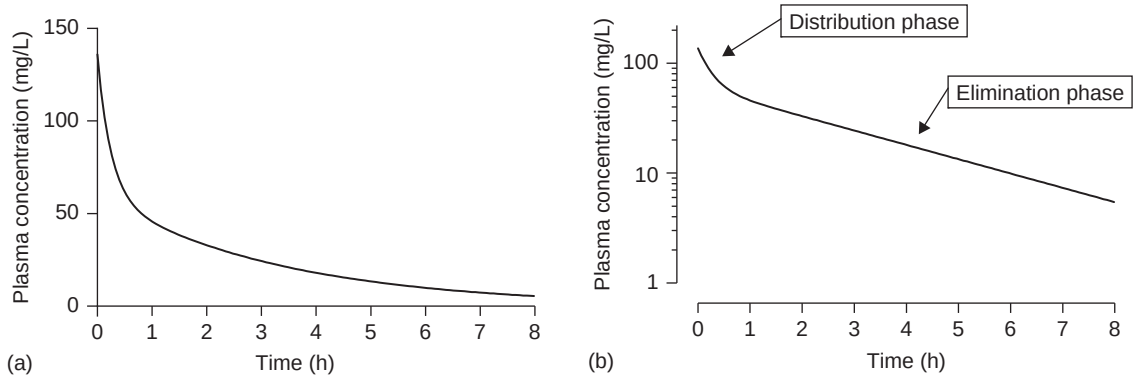


FIGURE 23.8 Plasma concentration–time curve after i.v. administration of an imaginary drug for which also distribution adds to the profile: (a) linear scale; (b) semi-log scale.

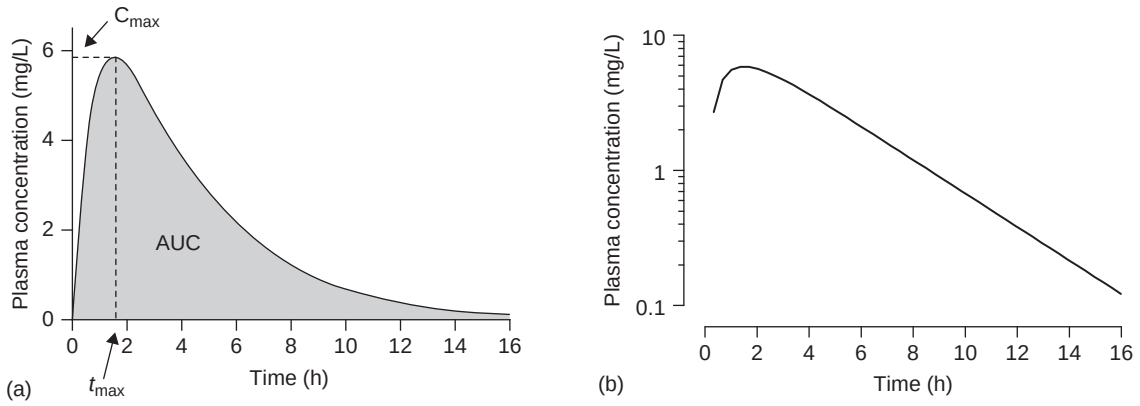


FIGURE 23.9 Plasma concentration–time curve after oral administration of an imaginary drug: (a) linear scale;(b) semi-log scale.

B. Volume of Distribution

The volume of distribution is not a "real" volume. It is a proportionality constant, relating the total amount of drug present in the organism to its plasma concentration at the same moment. It is the fluid volume in which the total amount of drug in the body should be dissolved to give rise to the same concentration as measured in the plasma. This calculated volume does not necessarily correspond to an identifiable physiological volume and can be much larger than the volume of total body water. It is therefore called "apparent" volume of distribution.

Total body water consists of the body fluids within cells (intracellular fluid) and the fluids outside the body cells (extracellular fluid). The extracellular fluid that fills the narrow spaces between cells of tissues is known as interstitial or intercellular fluid, whereas the extracellular fluid within the blood vessels is termed plasma. In a normal 70-kg man, total body water volume is about 42 L (or 60 percent of the body weight), consisting of about 3 L plasma, 11 L interstitial fluid, and 28 L intracellular fluid. If a drug is not bound in plasma or tissues and distributes over total body water, the apparent volume of distribution will be 42 L per 70 kg. This is the case, for example, for antipyrine. If a drug is likewise not bound in plasma and tissues but does not penetrate cells, the distribution will be limited to the extracellular fluid (plasma plus interstitial fluid), equaling 14 L. The apparent volumes of distribution of such drugs approximate their true volume of distribution. However, most substances bind to plasma and tissue proteins. For a drug that preferentially binds to plasma proteins, the plasma concentration will be higher than the concentration in the interstitial and intracellular fluid. In this case, the apparent volume of distribution will be smaller than 42 L. If, however, a drug binds preferentially to tissue proteins, the total drug concentration in plasma will be lower than in tissues and the apparent volume of distribution will be larger than 42 L. A typical example is digoxin, which is highly bound in muscle and has an apparent volume of distribution of about 600 L.

Equation (23.4) describes the relationship between apparent volume of distribution, drug binding and anatomical volumes:

$$V_d = V_p + V_T \frac{f_p}{f_T} \tag{23.4}$$

where V_d is the apparent volume of distribution, V_p is the plasma volume, V_T is the extravascular volume (the sum of the interstitial fluid volume and the intracellular fluid volume), and f_p and f_T are the free fractions of drug in plasma and extravascular space, respectively. The apparent volume of distribution increases with increases in anatomical volumes or tissue binding and decreases with increases in plasma or blood binding.

Many acidic drugs (e.g., salicylates, sulfonamides, penicillins, and anticoagulants) are highly bound to plasma proteins or are not lipophilic enough to distribute into cells. These drugs, therefore, have small volumes of distribution (<20 L). Basic drugs, on the other hand, often exhibit a large apparent volume of distribution because they tend to be highly distributed to tissues, and their plasma concentrations remain relatively low. Table 23.2 shows the apparent volume of distribution for some drugs.

C. Clearance

The term clearance describes the process of drug elimination from the body or from a single organ regardless of the mechanism involved. It can be defined as the volume of biological fluid, such as blood or plasma, that would have to be completely freed of the drug per unit of time to account for the elimination. The units for clearance are, therefore, milliliters per minute (mL/min) or liters per hour (L/h). As was described earlier, total elimination of a drug from the body may be a result of processes that occur in the kidney, liver, and other organs. Clearance by means of these various organs of elimination is additive. The systemic or total body clearance (Cl_T) is the sum of these respective organ clearances and considers the entire body as a drug-eliminating system from which many elimination processes may occur. As liver and kidneys are the major organs for drug elimination, we will elaborate on renal and hepatic clearance.

TABLE 23.2 Apparent Volumes of Distribution for Some Drugs in L/kg

0.11	Warfarin	2.1	Cimetidine
0.14	Ibuprofen	3.9	Propranolol
0.17	Salicylic acid	8.0	Digoxin
0.25	Gentamicin	30.0	Imipramine
0.51	Digitoxin	235	Chloroquine
0.70	Atenolol		

1. Renal Clearance

The efficiency of the renal excretion of a drug is expressed as "renal clearance" (Cl_R). The Cl_R of a drug is the volume of plasma that is cleared of that drug by the kidneys per time unit. Substances such as inulin and creatinine are eliminated by glomerular filtration, are not subject to either tubular secretion or reabsorption, and are not bound to plasma proteins. Their Cl_R in adults with normal renal function will be around 125 mL/min, corresponding to the volume of plasma that is subjected per minute to glomerular filtration. The clearance for inulin or creatinine can therefore be used as an index of the glomerular filtration rate. For substances that are filtered but also actively secreted, Cl_R is higher than 125 mL/min and can be as high as 650 mL/min, which is the total plasma flow through the kidneys. Such values are found for *para*-aminohippuric acid and penicillins, for example. For drugs that are filtered but also reabsorbed or for drugs bound to plasma proteins, clearance values can be lower than 120 mL/min. Relating the Cl_R of a drug to the glomerular filtration rate can therefore provide information on the mechanisms of renal excretion.

The Cl_R of a drug can be calculated by dividing the amount of drug excreted in the urine over a given time interval by the concentration of the drug in blood or plasma at the time corresponding to the midpoint of the urine collection interval.

2. Hepatic Clearance

The hepatic clearance (Cl_H) of a drug can be defined as the volume of blood that is cleared of the drug by the liver per unit of time. The pharmacokinetic concept of Cl_H takes into consideration the anatomical and physiological facts that drugs reach the liver through the portal vein and the hepatic artery and leave it through the hepatic vein. Unbound drugs in plasma diffuse through the liver cell membrane to reach the metabolic enzymes. Therefore, at least three major parameters ought to be considered when quantifying drug elimination by the liver: blood flow through the organ (Q), which reflects transport to the liver; free fraction of the drug in blood (f_u) which affects access of the drug to the enzymes; and intrinsic ability of the hepatic enzymes to metabolize the drug, expressed as intrinsic clearance (Cl'_{int}). Cl'_{int} is the ability of the liver to remove a drug in the absence of flow limitations and blood binding. Taking these three parameters into account, the Cl_H can be expressed by the following equation:

$$Cl_H = Q\left[\frac{f_u \cdot Cl'_{int}}{Q + f_u \cdot Cl'_{int}}\right] \tag{23.5}$$

It is obvious that the Cl_H cannot exceed the total volume of blood reaching the liver per unit of time, that is, the liver blood flow Q. The ratio of the Cl_H of a drug to the hepatic blood flow is called the extraction ratio of the drug (E). The value of the extraction ratio can vary between 0 and 1. It is zero when $f_u Cl'_{int}$ is zero, that is, when the drug is not metabolized in the liver. It is 1 when the Cl_H equals the hepatic blood flow (about 1.5 L/min in humans).

When $f_u Cl'_{int}$ is very small in comparison to hepatic blood flow ($f_u Cl'_{int} < Q$), Equation 23.5 reduces to the following equation:

$$Cl_H = f_u \cdot Cl'_{int} \tag{23.6}$$

In that case, clearance is not blood-flow-dependent but depends on enzymatic activity and on plasma protein binding. Binding to plasma proteins will limit the elimination. This is called "restrictive elimination." Drugs with a restrictive elimination have a low extraction ratio (<0.3). Examples are antipyrine, phenytoin, and warfarin.

When $f_u Cl'_{int}$ is very large in comparison to hepatic blood flow ($f_u Cl'_{int} > Q$), Equation 23.5 reduces to the following equation:

$$Cl_H = Q \tag{23.7}$$

In this case, clearance is dependent on hepatic blood flow and independent of Cl'_{int} and f_u. This is called "blood-flow-dependent" or "nonrestrictive" elimination (e.g., nitroglycerin, propranolol, and lidocaine). Drugs with a nonrestrictive elimination have a high extraction ratio (>0.7), and bound as well as free molecules are eliminated, since the affinity of the drug for the hepatic enzymes exceeds its affinity for the plasma proteins. As clearance of nonrestrictively eliminated drugs depends on hepatic blood flow, a decrease in hepatic blood flow (e.g., in hepatic disease or cardiac failure) will reduce clearance of these drugs.

D. Elimination Half-Life ($T_{1/2}$)

The elimination half-life ($T_{1/2}$) is the time it takes for the elimination processes to reduce the plasma concentration or the amount of drug in the body by 50 percent. Elimination half-life is a composite pharmacokinetic parameter determined by both clearance and volume of distribution (V_d), as described by the following equation:

$$T_{1/2} = 0.7\frac{V_d}{Cl} \tag{23.8}$$

Elimination half-life is increased by an increase in volume of distribution or a decrease in clearance, and vice versa. This is because a decrease in the efficiency of elimination (and therefore in clearance) would, of course, cause an increase in the time needed to reduce the plasma concentration by 50 percent. On the other hand, the larger the volume of distribution, the more the drug is concentrated in the tissues rather than in the plasma. It is, however, the drug in plasma that is exposed to the elimination mechanisms. Therefore, an increase in volume of distribution also increases elimination half-life. For the simplest cases, elimination half-life may be used to make decisions about drug dosage, and can be derived from the plasma concentration–time profile as the time it takes for a random plasma concentration in the elimination phase to be halved (see Figure 23.7). It does not matter at what concentration half-life is measured, as long as it is measured in the mono-exponential elimination phase of the curve. Therefore, the time for the plasma concentration to drop from 10 to 5 mg/L is the same as from 8 to 4 mg/L or from 2 to 1 mg/L. It becomes more complicated when the plasma concentration follows a multiexponential pattern of decline and two or more half-lives can be calculated. This situation is left out of the discussion in this chapter, but the interested reader can refer to the textbooks on pharmacokinetics mentioned at the end of the chapter.

E. Bioavailability

Bioavailability is (1) the fraction of an administered dose of a drug that reaches the systemic circulation as intact drug (expressed as F) and (2) the rate at which this occurs. As an i.v. dose is injected directly into the systemic circulation, the bioavailability of an i.v. dose is by definition 100 percent ($F = 1$). For all other routes of administration, bioavailability is determined by the extent of drug absorption (being the result of both drug uptake from the administration site and possible first-pass effects; see Section III.D.), and varies between 0 and 100 percent ($0 < F < 1$). For example, orally administered morphine has a bioavailability of about 25 percent due to significant first-pass metabolism in the liver. Therefore, the dose of morphine given orally is usually 3–5 times larger than an i.v. dose of morphine.

The usual method for measuring bioavailability (also called absolute bioavailability) of an oral formulation is to give a group of volunteers an i.v. administration of the drug and the oral formulation on separate occasions and to determine the respective area under the plasma concentration–time curves. Since the AUC is a measure of the total amount of unaltered drug that reaches the systemic circulation (see Section VI.A.), the bioavailability of the oral formulation can subsequently be determined by comparing these respective AUCs, as described by the following equation:

$$\text{Absolute bioavailability} = F = \frac{AUC_{oral}/dose_{oral}}{AUC_{i.v.}/dose_{i.v.}} \tag{23.9}$$

For example, if the AUC_{oral} is 25 percent of the $AUC_{i.v.}$, the bioavailability of the oral formulation is 25 percent ($F = 0.25$).

Sometimes the bioavailability of a new formulation is not assessed against an i.v. formulation but against another (reference) formulation. This is referred to as measuring relative bioavailability, and it provides a measure of the relative performance of two formulations (new formulation A and reference formulation B) in getting the drug absorbed into the systemic circulation (see Equation 23.10).

$$\text{Relative bioavailability} = \frac{AUC_A/dose_A}{AUC_B/dose_B} \tag{23.10}$$

Obviously, the relative bioavailability of a formulation is not equal to F (the fraction of the dose that reaches the systemic circulation), as the absolute bioavailability of the reference formulation might be quite low due to poor absorption and/or first-pass metabolism.

VII. VARIABILITY IN PHARMACOKINETICS

When a plasma concentration curve is constructed for different patients that have been given an identical dose of an identical drug, inter-individual differences will be noted. In some cases, plasma concentrations in one patient may remain below the minimal effective concentration, whereas the plasma concentration in another patient reaches the minimum toxic concentration. Besides some very obvious causes, such as body weight and body composition, some other factors involved in the inter-individual variability in pharmacokinetics are concisely described below.

A. Genetic Factors

Studies on identical and nonidentical twins have shown that much inter-individual pharmacokinetic variability is determined genetically. Pharmacokinetic variability may be caused by genetic polymorphism (the situation where several functionally distinct genes are common in a population) in genes involved in drug absorption, distribution, and elimination. In recent years, several polymorphisms in genes encoding for transporter proteins have been described. These polymorphisms could alter the absorption, distribution, and elimination of compounds that are substrates for these transporters. However, much work remains to be done to understand the clinical implications of these polymorphisms.

Genetic polymorphism of genes involved in drug metabolism is regarded as one of the major sources of variability in pharmacokinetics. On the other hand, renal excretion of drugs does not appear to be prone to genetic polymorphism. The Cl_R for any drug tends to be similar in age- and weight-matched healthy subjects. As a consequence, drugs that are predominantly excreted unchanged tend to show less inter-individual variability than extensively metabolized ones.

B. Age

The main reason for age affecting drug action is that renal excretion is less efficient in neonates and elderly people, so that renally cleared drugs commonly produce stronger and more prolonged effects at the extremes of life. Glomerular filtration rate in the newborn (normalized to body surface area) is only about 20 percent of the rate in adults, and tubular function is also reduced. Accordingly, elimination half-lives of renally eliminated drugs are longer in newborns than in adults. In babies born at term, renal function increases to values similar to those in young adults in less than a week and continues to increase to a maximum of approximately twice the adult value at six months of age. From about twenty years of age, glomerular filtration rate starts to decline slowly, falling by about 25 percent at fifty years and by 50 percent at seventy-five years.

In the developing child, biotransformation of drugs is also altered. Several drugs that are eliminated primarily through hepatic metabolism have exhibited a higher clearance in children than in adults. A number of factors contribute to these changes during development, such as relative liver size (relative to total body size) and the maturation profile of different drug metabolizing enzymes from birth onwards. In neonates, distinct patterns of development in drug-metabolizing enzymes have been observed. Some drug-metabolizing enzymes show an onset of activity within hours after birth or in the first week of life; others appear to approach full competence only after several months.

C. Drug Interactions

The pharmacokinetics of a drug can be influenced by the concurrent administration of another drug that affects its absorption, distribution, metabolism, and/or excretion.

The GI absorption of drugs can be influenced by agents with a large surface area upon which the drug can be absorbed, bound, or chelated, or it can be influenced by agents that alter GI motility and thereby alter the rate or extent of absorption. Drugs that lower local blood flow can slow down the absorption. The addition of adrenaline to local anesthetic injections results in local vasoconstriction and slower absorption of the anesthetic, thus prolonging its local effect at the injection site.

Drug distribution can be altered by competition of drugs for plasma protein binding or by displacement of a drug from tissue binding sites. A competition for plasma protein binding can increase the free fraction and—temporarily—the free concentration of the drug in plasma, and can therefore result in an increased distribution

of the drug toward the biophase, the elimination sites, and other tissues. Displacement of a drug from tissue binding sites can induce a temporary increase in the plasma concentration, since it allows redistribution of the displaced drug from the tissue toward the plasma. This increase often increases drug elimination so that a new steady state is reached. For some drugs, the temporary rise in plasma concentration before a new steady state is reached may cause toxicity.

Metabolism is the major factor leading to clinically significant drug−drug interactions. Hepatic microsomal drug-metabolizing enzymes can be induced (leading to higher production of enzyme) after chronic administration of, for example, phenobarbital, phenytoin, rifampicin, or St. John's wort. This produces more rapid metabolism of drugs. A very potent enzyme inducer, such as rifampicin, can markedly alter enzyme activity within 48 h after its administration, while for most inducers the maximal effect is obtained only after 7−10 days. An equal or even longer time after stopping the enzyme inducer is required to dissipate the induction. Hepatic microsomal metabolism may also be inhibited by exogenous and endogenous compounds, resulting in a slower rate of metabolism. This enzyme inhibition generally occurs more quickly than enzyme induction and may begin as soon as a sufficient hepatic concentration of the inhibitor is achieved. The most common mechanism is competitive inhibition; any two drugs that are metabolized by the same enzyme may compete with each other for binding to the enzyme, thereby slowing down each other's metabolism. Moreover, some drugs act as a competitive inhibitor for a particular enzyme, although they are not metabolized by that particular enzyme. This is the case in humans for quinidine, which selectively inhibits CYP2D6, although it is not metabolized by that enzyme. Also, noncompetitive inhibition can occur. In this case, there is no direct competition between the substrate and the inhibitor for the enzyme, but the inhibitor deactivates the enzyme by binding to other parts of the enzyme. For example, macrolide antibiotics such as erythromycin are metabolized by CYP3A4 to a reactive metabolic intermediate that forms a stable, inactive complex with the enzyme.

The renal excretion of certain drugs that are weak acids or weak bases may be influenced by other drugs that affect urinary pH. This is due to changes in ionization of the drug and thus to alteration of its lipid solubility and the ability to be absorbed back into the blood from the kidney tubule (see Section V.A.1.b.). Also, active secretion into the renal tubules can be inhibited by concurrent drug therapy, thus increasing serum drug levels and pharmacologic response (see Section V.A.1.c.). A drug can also affect the rate of renal excretion by altering the protein binding and, hence, filtration.

D. Disease State

Several diseases can cause variations in pharmacokinetics. Renal or hepatic insufficiency predisposes to toxicity by causing intense or prolonged drug effects as a result of increased plasma levels following a standard dose regimen. Drug absorption is slowed in conditions causing gastric stasis (e.g., migraine) and may be incomplete in patients with diarrhea or with malabsorption due to diseases of the pancreas or gut or due to edema of ileal mucosa. Nephrotic syndrome is characterized by a heavy loss of proteins in urine (proteinuria) and thus a reduced concentration of albumin in plasma and edema. Edema of intestinal mucosa alters drug absorption, while changes in binding to plasma albumin changes drug disposition. An impaired functioning of the blood−brain barrier occurs in meningitis. Hypothermia (lowered body temperature, often in elderly persons) markedly reduces the clearance of many drugs.

E. Pregnancy

Pregnancy is associated with numerous alterations in physiology that can influence pharmacokinetics. Plasma concentration of albumin is diminished in the mother, resulting in alteration in drug−protein binding. The increase in glomerular filtration, aimed to help excrete the increased amount of waste products, also results in an increased renal elimination of drugs. The blood of the mother and the fetus are separated by the placental barrier (cf. blood−brain barrier). This barrier could allow some drugs to be administered to the mother without influencing the fetus. The placental barrier is, however, rapidly crossed by lipophilic molecules that can have effects on the fetus. For example, some drugs are known to cause abnormal development of the fetus (teratogenic effect). When transferred to the fetus, drugs are usually slowly eliminated by the fetus. Drug-metabolizing enzymes in the fetal liver are less active than in adults, and elimination through fetal kidney is not efficient since fetal urine drains in the amniotic fluid, which is swallowed by the fetus.

Further Reading

The following textbooks and reviews were used to compile this chapter. These references can offer a good start for the interested reader looking for more elaborate and/or more in-depth information.

Textbooks on Pharmacokinetics

Birkett DJ. Pharmacokinetics made easy. Rev. ed.. Australia: McGraw-Hill; 2002.
Rowland M, Tozer TN. Clinical pharmacokinetics: Concepts and applications. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 1995.
Shargel L, Wu-Pong S, Yu ABC. Applied biopharmaceutics and pharmacokinetics. 5th ed. The New York: McGraw-Hill Companies; 2005.
Washington N, Washington C, Wilson CG. Physiological pharmaceutics: Barriers to drug absorption. 2nd ed. London: Taylor and Francis; 2001.

General Textbooks on Physiology and Pharmacology

Boron WF, Boulpaep EL, editors. Medical physiology. Updated ed.. Philadelphia, PA: Elsevier Saunders; 2005.
Ganong WF. Review of medical physiology. 22nd ed. New York: McGraw-Hill; 2005.
Hardman JG, Limbird LE, Molinoff PB, Ruddon RW. Goodman & Gilman's The pharmacological basis of therapeutics. 9th ed. New York: McGraw-Hill; 1996.
Rang HP, Dale MM, Ritter JM, Flower R. Rang & Dale's Pharmacology. 6th ed. New York: Churchill Livingstone; 2007.
Tortora GJ, Derrickson B. Principles of anatomy and physiology. 11th ed. Chichester: John Wiley & Sons; 2006.

Reviews on Selected Topics

Beringer PM, Slaughter RL. Transporters and their impact on drug disposition. Ann Pharmacother 2005;39:1097–108.
Daugherty AL, Mrsny RJ. Transcellular uptake mechanisms of the intestinal epithelial barrier: Part one. PSTT 1999;2(4):144–51.
Mizuno N, Niwa T, Yotsumoto Y, Sugiyama Y. Impact of drug transporter studies on drug discovery and development. Pharmacol Rev 2003;55(3):425–61.
Zhang Y, Benet LZ. The gut as a barrier to drug absorption. Clin Pharmacokinet 2001;40(3):159–68.

CHAPTER

24

Biotransformation Reactions and their Enzymes

Bernard Testa[1] *and Bernd Clement*[2]

[1]Department of Pharmacy, Lausanne University Hospital (CHUV), Lausanne, Switzerland; [2]Pharmaceutical Institute, Christian Albrechts-University, Gutenbergstraße, Kiel, Germany

OUTLINE

Nothing that is worth knowing can be taught. **Oscar Wilde**
(a quotation engraved on his monument in Merrion Square, Dublin)

I. INTRODUCTION

Drug metabolism—and more generally xenobiotic metabolism—has become a major pharmacological and pharmaceutical science with particular relevance to biology, therapeutics, and toxicology. As such, it is also of great importance in medicinal chemistry since it influences the deactivation, activation, detoxification, and toxification of most drugs. This chapter is written for medicinal chemists and will offer structured knowledge rather than encyclopedic information. Readers who wish to go further in the study of xenobiotic metabolism may consult various books [1–18] and reviews [19–32].

The metabolism of drugs and other xenobiotics is often a biphasic process in which the compound may first undergo a functionalization reaction (phase I reaction) of oxidation, reduction, or hydrolysis. This introduces or unveils a functional group, such as a hydroxyl or amino group, suitable for coupling with an endogenous molecule or moiety in a second metabolic step known as a conjugation reaction (phase II reaction). In a number of cases, phase I metabolites may be excreted prior to conjugation, while many xenobiotics can be directly

conjugated. Furthermore, reactions of functionalization may follow some reactions of conjugation (e.g., some conjugates are hydrolyzed and/or oxidized prior to their excretion) [19–32].

The major function of xenobiotic metabolism is the elimination of physiologically useless compounds, some of which may be harmful; witness the many toxins produced by plants. This evolutionary function justifies the designation of detoxification originally given to reactions of xenobiotic metabolism. However, numerous xenobiotics and even a marked number of drugs are known to yield toxic metabolites, a situation known as toxification [33]. In pharmacological terms, a drug may or may not have active metabolites. The former case is rather frequent, especially with phase I metabolites. Examples of drugs that do not yield any active metabolites include the soft drugs. The reverse case is that of prodrugs, which are inactive therapeutic agents whose clinical activity is due to an active metabolite [34].

The present chapter aims at laying the foundations of drug metabolism by offering an analytical view of the field. In other words, the focus will be on metabolic reactions, the target groups they affect, and the enzymes by which they are catalyzed. Further information can be found in accompanying chapters. Also, the analytical information presented here needs to be complemented by a synthetic view as provided by metabolic schemes. These show at a glance the competitive and sequential reactions undergone by a given drug, and they bring logic and order to what may appear as a random presentation. As an example of a metabolic scheme, Figure 24.1 presents the biotransformation of propranolol (**1**) in humans [35]. There are relatively few studies as comprehensive and clinically relevant as this one, which remains as current today as it was when published in 1985. Indeed, over 90 percent of a dose was accounted for and consisted mainly of products of oxidation and conjugation. The missing 10 percent may represent other minor and presumably quite numerous metabolites, such as those resulting from ring hydroxylation at other positions or from the progressive breakdown of glutathione conjugates.

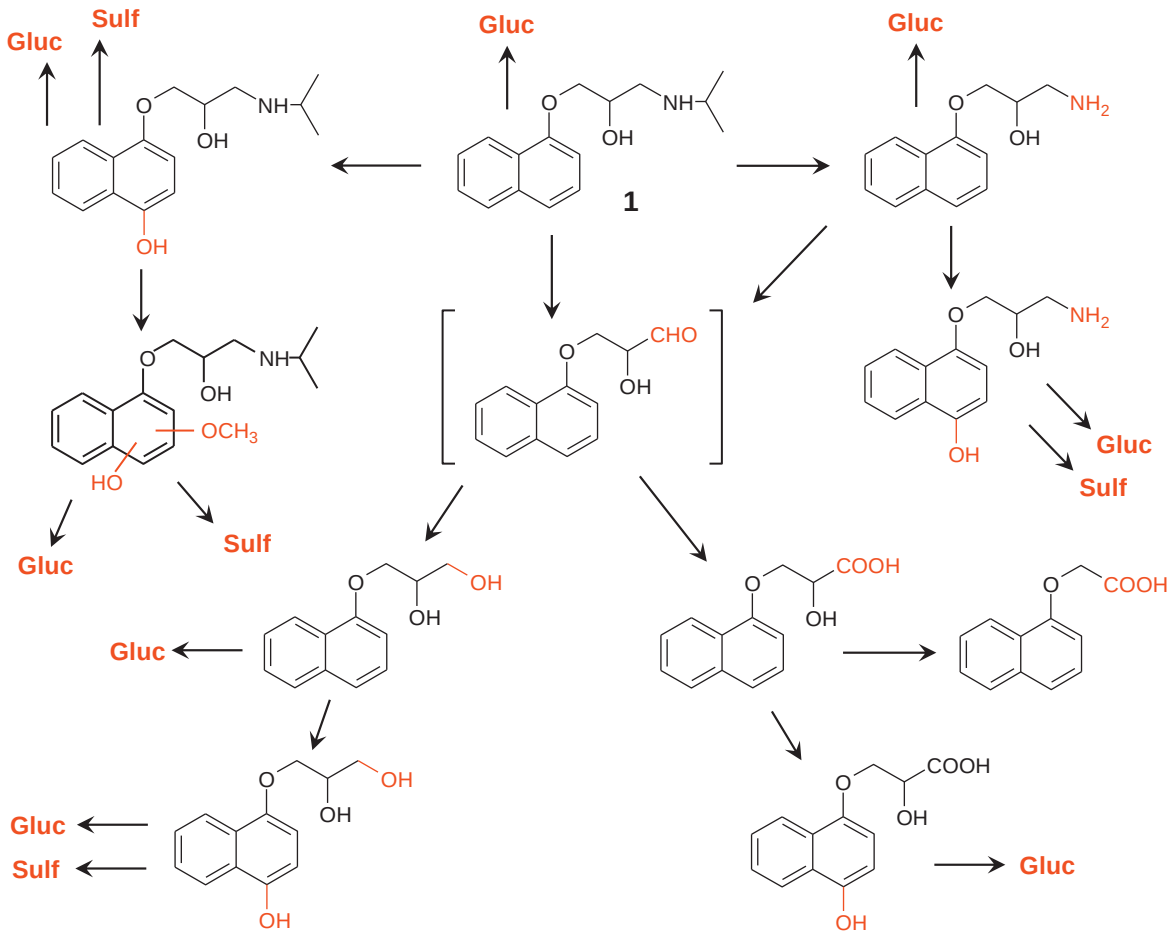


FIGURE 24.1 The metabolism of propranolol (**1**) in humans, accounting for more than 90 percent of the dose; Gluc = glucuronide(s); Sulf = sulfate(s) [35].

II. FUNCTIONALIZATION REACTIONS

Reactions of functionalization comprise oxidations (electron removal, dehydrogenation, and oxygenation), reductions (electron addition, hydrogenation, and removal of oxygen), and hydrations/dehydrations (hydrolysis and addition or removal of water). The reactions of oxidation and reduction are catalyzed by a very large variety of oxidoreductases, while various hydrolases catalyze hydrations. A large majority of the enzymes involved in xenobiotic functionalization are briefly reviewed below before we examine in detail the metabolic reactions and pathways of functionalization.

A. Enzymes Catalyzing Functionalization Reactions

1. *Monooxygenases*

Monooxygenation reactions are of major significance in drug metabolism and are mediated mainly by two enzymes that differ markedly in their structure and properties, namely the cytochromes P450 (CYPs) and the flavin-containing monooxygenases (FMOs) (see Table 24.1) [1,3,9–12,19,21,23,26,36–40]. Among these, the most important as far as xenobiotic metabolism is concerned are the CYPs, a very large group of enzymes belonging to heme-coupled monooxygenases. Cytochrome P450 is the major drug-metabolizing enzyme system, playing a key role in detoxification and toxification. It is of additional significance in medicinal chemistry because several CYP enzymes are drug targets (e.g., thromboxane synthase (CYP5A1) and aromatase (CYP19A1)). The CYPs are encoded by the *CYP* gene superfamily and are classified in families and subfamilies. A total of fifty-seven human *CYP* genes are known to date. The three CYP families mostly involved in xenobiotic metabolism are CYP1 to CYP3 (Table 24.2). The relative importance of the major human CYPs is given in Table 24.3. The endobiotic-metabolizing CYPs are in families 4, 5, 7, 8, 11, 17, 19–21, 24, 26, 27, 39, 46 and 51.

An understanding of the regiospecificity and broad reactivity of cytochrome P450 requires a presentation of its catalytic cycle (Figure 24.2). The enzyme in its ferric (oxidized) form exists in equilibrium between two spin states: a hexacoordinated low-spin form whose reduction requires a high-energy level, and a pentacoordinated high-spin form. Binding of the substrate to enzyme induces a shift to the reducible high-spin form (reaction **a**).

TABLE 24.1 A Survey of Oxidoreductases Playing a Role in Drug Metabolism (Listed in the Order they Appear in the Text) [12,19,36]

Enzymes	EC numbers	*Gene root* (or *gene*) and major human enzymes
Cytochrome P450	Mainly EC 1.14.13 and 1.14.14.1	*CYP* (see Table 24.2)
Flavin-containing monooxygenases	EC 1.14.13.8	*FMO* (FMO1 to FMO5)
Monoamine oxidases	EC 1.4.3.4	*MAO* (MAO-A and MAO-B)
Copper-containing amine oxidases	EC 1.4.3.6	*AOC* (DAO and SSAO)
Aldehyde oxidase	EC 1.2.3.1	*AOX1* (AO)
Xanthine oxidoreductase	EC 1.17.1.4 and 1.17.3.2	*XOR* (XDH and XO)
Various peroxidases	EC 1.11.1.7 and 1.11.1.8	e.g., *EPO* (EPO), *MPO* (MPO) and *TPO* (TPO)
Protaglandin G/H synthase	EC 1.14.99.1	*PTGS* (COX-1 and COX-2)
Alcohol dehydrogenases	EC 1.1.1.1	*ADH* (ADH1A, ADH1B and ADH1C, ADH4, ADH5, ADH6, and ADH7)
Aldehyde dehydrogenases	EC 1.2.1.3 and 1.2.1.5	*ALDH* (e.g., ALDH1A1, 1A2 and 1A3, 1B1, 2, 3A1, 3A2, 3B1, 3B2, 8A1 and 9A1)
Aldo-keto reductases	In EC 1.1.1 and 1.3.1	*AKR* (e.g., ALR1, ALR2, DD1, DD2, DD3, DD4, AKR7A2, 7A3 and 7A4)
Carbonyl reductases	EC 1.1.1.184	*CBR* (CR1, CR3)
Quinone reductases	EC 1.6.5.2 and 1.10.99.2	*NQO* (NQO1 and NGO2)
Mitochondrial amidoxime reducing component (mARC)	EC 1.-.-.-	*mARC1* and *mARC2* (*MOSC1* and *MOSC2*)

TABLE 24.2 The Three Human *CYP* Gene Families Mostly Involved in Xenobiotic Metabolism Along with their Subfamilies of Gene Products

Families	Subfamilies	Human gene products
CYP 1 FAMILY (ARYL HYDROCARBON HYDROXYLASES; XENOBIOTIC METABOLISM; INDUCIBLE)		
	CYP 1A	CYP1A1, CYP1A2
	CYP 1B	CYP1B1
CYP 2 FAMILY (XENOBIOTIC METABOLISM; CONSTITUTIVE AND XENOBIOTIC-INDUCIBLE)		
	CYP 2A	CYP2A6, CYP2A7, CYP2A13
	CYP 2B	CYP2B6
	CYP 2C	CYP2C8, CYP2C9, CYP2C18, CYP2C19
	CYP 2D	CYP2D6
	CYP 2E	CYP2E1
	CYP 2F	CYP2F1
	CYP 2J	CYP2J2
	CYP 2R	CYP2R1
	CYP 2S	CYP2S1
	CYP 2U	CYP2U1
	CYP 2W	CYP2W1
CYP 3 FAMILY (XENOBIOTIC AND STEROID METABOLISM; STEROID-INDUCIBLE)		
	CYP 3A	CYP3A4, CYP3A5, CYP3A7 (fetal CYP enzyme), CYP3A43

TABLE 24.3 Levels and Variability of Human CYP Enzymes Involved in Drug Metabolism [4,27,28,38,41]

CYP	Level in liver (% of total)	Variability range	Percent of drugs being substrates
1A1			3
1A2	ca. 13	ca. 40-fold	10
1B1	<1		1
2A6	ca. 4	ca. 30- to 100-fold	3
2B6	<1	ca. 50-fold	4
2C	ca. 18	25- to 100-fold	25
2D6	up to 2.5	>1000-fold	15
2E1	up to 7	ca. 20-fold	3
3A4	up to 28	ca. 20-fold	36

The first electron then enters the enzyme-substrate complex (reaction **b**), reducing the enzyme to its ferrous form, which has a high affinity for diatomic gases such as CO (a strong inhibitor of cytochrome P450) and dioxygen (reaction **c**). The cycle continues with a second electron entering via either F_{P1} or F_{P2} and reducing the ternary complex (reaction **d**). Electron transfer within the ternary complex generates bound peroxide anion (O_2^{2-}). The bound peroxide anion is split by the addition of two protons, liberating H_2O (reaction **e**). The exact electronic nature of this decisive reactive entity is still a matter of debate. However, in agreement with data from different sources, the formation of a high-valent iron(IV)-oxo-porphyrin radical cation is highly feasible. This is the reactive form of oxygen that will attack the substrate. The binary enzyme-product complex dissociates, thereby regenerating the initial state of cytochrome P450 (reaction **f**) [42,43].

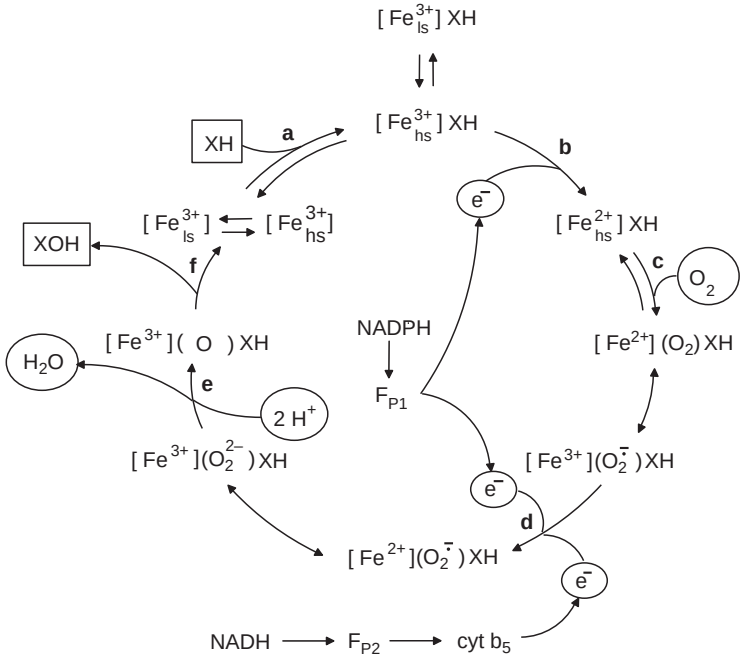


FIGURE 24.2 Catalytic cycle of cytochrome P450 associated with monooxygenase reactions. [Fe^{3+}] = ferricytochrome P450; hs = high spin; ls = low spin; [Fe^{2+}] = ferrocytochrome P450; F_{P1} = flavoprotein 1 = NADPH-cytochrome P450 reductase; F_{P2} = NADH-cytochrome b_5 reductase; cyt b_5 = cytochrome b_5; XH = substrate. *Modified from [23].*

A major challenge is still the prediction of the sites of metabolism and the products formed. Current computational approaches to address xenobiotic metabolism have been summarized recently [44].

2. *Other Oxidoreductases*

Other oxidoreductases that can play a role in drug oxidation are (Table 24.1):

- Monoamine oxidases, which are essentially mitochondrial enzymes;
- The broad group of copper-containing amine oxidases, which contain diamine oxidase (DAO) and semicarbazide-sensitive amine oxidases (SSAO);
- The cytosolic molybdenum hydroxylases, namely aldehyde oxidase and xanthine oxidoreductase, which exist in a dehydrogenase form (XDH) and an oxidase form (XO);
- Various peroxidases, such as eosinophil peroxidase (EPO), myeloperoxidase (MPO) and thyroid peroxidase (TPO) (note that several cytochrome P450 enzymes have been shown to have peroxidase activity);
- Prostaglandin G/H synthase, which is able to use a number of xenobiotics as cofactors in a cooxidation reaction.

Dehydrogenases/reductases involved in reactions of oxidation (dehydrogenation) and/or reduction (hydrogenation) are:

- Alcohol dehydrogenases (ADH), which are zinc enzymes found in the cytosol of the mammalian liver and in various extrahepatic tissues;
- Aldehyde dehydrogenases (ALDH), a large superfamily of enzymes produced by nineteen human genes in eleven families and thirteen subfamilies;
- The aldo-keto reductases (AKR), a complex superfamily of enzymes, which includes aldehyde reductases (ALR) and dihydrodiol dehydrogenase (DD);
- Carbonyl reductases (CR) and quinone reductases (NQO);
- Mitochondrial amidoxime reducing component (mARC).

3. *Hydrolases*

Hydrolases constitute a very complex ensemble of enzymes, many of which are known or suspected to be involved in xenobiotic metabolism (Table 24.4) [1,7,20,24]. Relevant enzymes among the serine hydrolases

TABLE 24.4 A Survey of Hydrolases Playing a Role in Drug Metabolism (Listed in the Order they Appear in the Text) [7,20,24,36]

Classes of hydrolases	Examples of enzymes (with some *gene roots* and human enzymes)
EC 3.1.1: Carboxylic Ester Hydrolases	EC 3.1.1.1: Carboxylesterase (*CES*) CES1A1, CES2, CES3
	EC 3.1.1.2: Arylesterase (*PON*, see 3.1.8.1)
	EC 3.1.1.8: Cholinesterase (*BCHE*)
EC 3.1.2: Thiolester Hydrolases	EC 3.1.2.20: Acyl-CoA hydrolase
EC 3.1.3: Phosphoric Monoester Hydrolases	EC 3.1.3.1: Alkaline phosphatase (*ALP*)
	EC 3.1.3.2: Acid phosphatase (*ACP*)
EC 3.1.6: Sulfuric Ester Hydrolases	EC 3.1.6.1: Arylsulfatase
EC 3.1.8: Phosphoric Triester Hydrolases	EC 3.1.8.1: Paraoxonase (*PON*) PON1, PON2, PON3
	EC 3.1.8.2: Diisopropyl-fluorophosphatase
EC 3.2: Glycosylases	EC 3.2.1.31: β-Glucuronidase (*GUSB*)
EC 3.3.2: Ether Hydrolases	EC 3.3.2.9: Microsomal epoxide hydrolase (*EPHX1*) mEH
	EC 3.3.2.10: Soluble epoxide hydrolase (*EPHX2*) sEH
EC 3.4.11: Aminopeptidases	EC 3.4.11.1: Leucyl aminopeptidase (*LAP*)
EC 3.4.13 and 3.4.14: Peptidases acting on di- and tripeptides	EC 3.4.14.5: Dipeptidyl-peptidase IV (*DPP4*)
EC 3.4.16 to 3.4.18: Carboxypeptidases	EC 3.4.16.2: Lysosomal Pro-Xaa carboxypeptidase
	EC 3.4.17.1: Carboxypeptidase A (*CPA*)
EC 3.4.21 to 3.4.25: Endopeptidases	EC 3.4.21.1: Chymotrypsin (*CTRB*)
	EC 3.4.24.15: Thimet oligopeptidase (*THOP*)
EC 3.5.1: Hydrolases acting on linear amides	EC 3.5.1.4: Amidase
	EC 3.5.1.39: Alkylamidase
EC 3.5.2: Hydrolases acting on cyclic amides	EC 3.5.2.1: Barbiturase
	EC 3.5.2.2: Dihydropyrimidinase (*DPYS*)
	EC 3.5.2.6: β-Lactamase

include carboxylesterases, arylesterases, cholinesterase, and a number of serine endopeptidases [EC 3.4.21]. The roles of arylsulfatases, phosphatases, β-glucuronidases, epoxide hydrolases, and some endopeptidases are also significant.

B. Reactions of Carbon Oxidation and Reduction

As is usual, we distinguish here between reactions targeting sp^3-, sp^2-, and sp-carbon atoms [19,23].

1. sp^3-Carbon Atoms

The most important reactions of oxidation of sp^3-carbon atoms are schematized in Figure 24.3. In the simplest cases, a nonactivated carbon atom in an alkyl group undergoes CYP-catalyzed hydroxylation (Box **A** in Figure 24.3). The penultimate position is a preferred site of attack, but hydroxylation can also occur at the terminal position or at another position in case of steric hindrance or with some highly regiospecific cytochromes P450. Dehydrogenation by dehydrogenases can then yield a carbonyl derivative that is either an aldehyde or a ketone. Note that these reactions may involve not only metabolites but also xenobiotic alcohols, and are reversible since dehydrogenases catalyze the reactions in both directions [23,45,46]. And while a xenobiotic ketone is very seldom oxidized further, aldehydes are good substrates for aldehyde dehydrogenases and lead irreversibly to carboxylic acid metabolites. A classic example is that of ethanol, which in the body exists in redox equilibrium with acetaldehyde, this metabolite being rapidly and irreversibly oxidized to acetic acid [47].

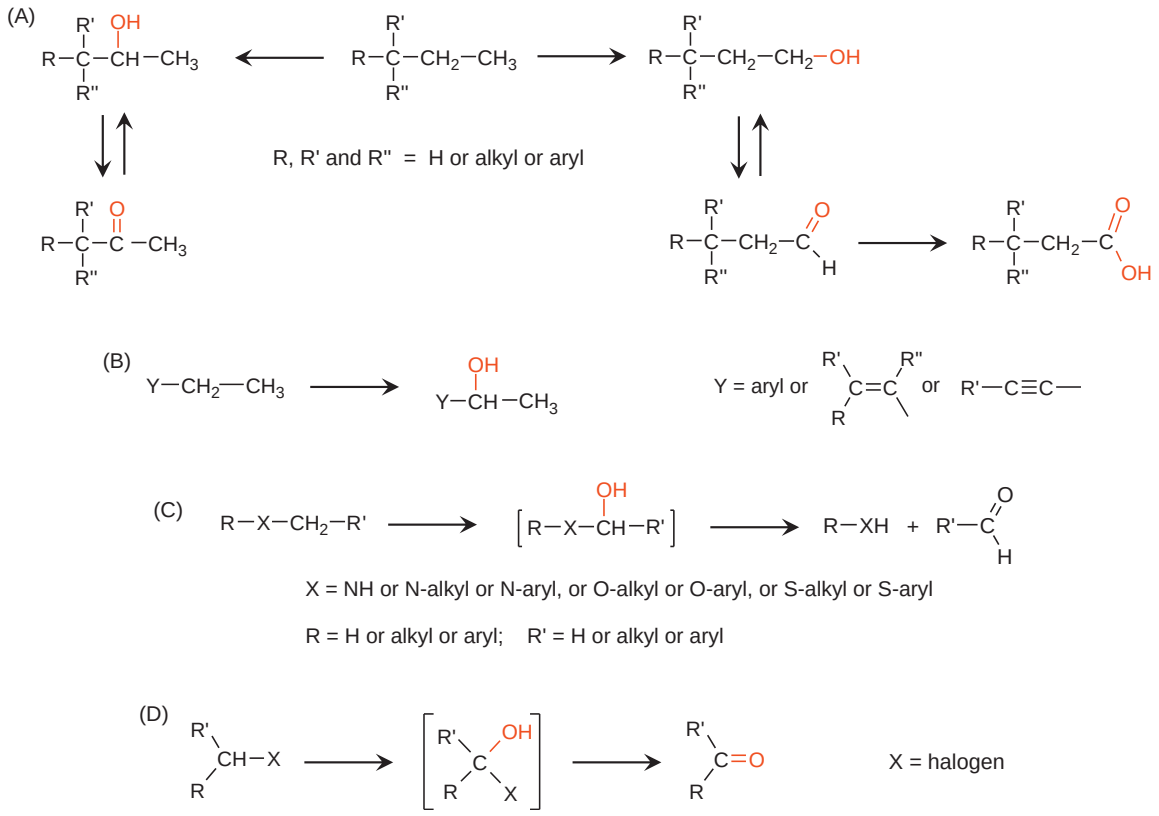


FIGURE 24.3 Major reactions of oxidation involving an sp^3-carbon in substrate molecules. **A**: CYP-catalyzed hydroxylation of an alkyl group, followed by reversible dehydrogenation to a carbonyl, and followed for aldehydes by irreversible oxidation to a carboxylic acid. **B**: The priviledged CYP-catalyzed hydroxylation of benzylic, allylic, and propargylic positions. **C**: CYP-catalyzed hydroxylation alpha to a heteroatom leads to spontaneous C-heteroatom cleavage (i.e., N-dealkylation or deamination, O-dealkylation, and S-dealkylation). **D**: CYP-catalyzed reactions of oxidative dehalogenation.

There is a known regioselectivity in CYP-catalyzed hydroxylations for carbon atoms adjacent (alpha) to an unsaturated system or a heteroatom such as N, O or S. In the former cases (Box **B** in Figure 24.3), hydroxylation can easily be followed by dehydrogenation (not shown). In the latter cases (Box **C** in Figure 24.3), however, the hydroxylated metabolite is usually unstable and undergoes a rapid postenzymatic elimination. Depending on the substrate, this pathway produces a secondary or primary amine, an alcohol or a thiol, while the alkyl group is released as an aldehyde or a ketone. Such reactions of deamination and N-dealkylation constitute a very common pathway as far as drug metabolism is concerned, since it underlies some well-known metabolic reactions of N-C cleavage discussed later.

Aliphatic carbon atoms bearing one or more halogen atoms (mainly chlorine or bromine) can be similarly metabolized by hydroxylation and the loss of HX to dehalogenated products (Box **D** in Figure 24.3).

For alkyl hydroxylation, one proposed mechanism is the abstraction of a hydrogen atom by $[Fe^{3+}](O)$ to form an alkyl radical and $[Fe^{3+}](OH)$, followed by "oxygen rebound," where the C-centered radical combines with a hydroxyl radical from $[Fe^{3+}](OH)$ to yield the alcohol and Fe^{3+} of heme [42].

2. *sp^2- and sp-Carbon Atoms*

Reactions at sp^2-carbons are characterized by their own pathways, catalytic mechanisms, and products (Figure 24.4). Thus, the oxidation of aromatic rings generates a variety of (usually stable) metabolites. Their common precursor is often a reactive epoxide (Box **A** in Figure 24.4), which can either be hydrolyzed by epoxide hydrolase to a dihydrodiol or rearrange under proton catalysis to a phenol. The production of a phenol is a common metabolic reaction for drugs containing one or more aromatic rings (e.g., the metabolism of propranolol; Figure 24.1). The *para*-position is the preferred position of hydroxylation for unsubstituted phenyl rings, but the regioselectivity of the reaction becomes more complex with substituted phenyl or with other aromatic rings [42].

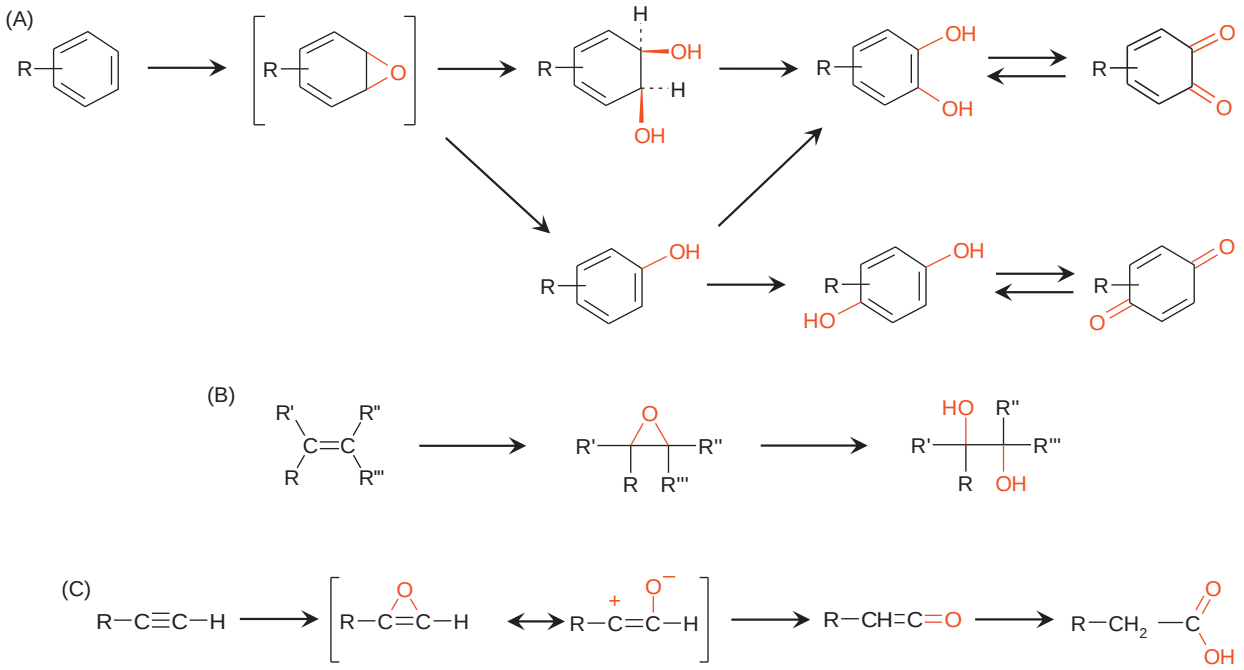


FIGURE 24.4 Major functionalization reactions involving an sp^2- or sp-carbon in substrate molecules. These reactions are oxidations (oxygenations and dehydrogenations), reductions (hydrogenations) and hydrations, plus some postenzymatic rearrangements. Their target sites are aromatic rings (**A**), carbon–carbon double bonds (**B**), and carbon–carbon triple bonds (**C**).

Dihydrodiols are seldom observed, as are catechol metabolites produced by dehydrogenation, catalyzed by dihydrodiol dehydrogenase. The further oxidation of phenols and phenolic metabolites to a catechol or hydroquinone is also possible, the rate of reaction and the nature of products depending on the ring and on the nature and position of its substituents. In a few cases, catechols and hydroquinones have been found to undergo further oxidation to quinones by two single-electron steps. The intermediate in this reaction is a semiquinone. Both quinones and semiquinones are reactive—in particular toward biomolecules—and have been implicated in many toxification reactions. For example, the high toxicity of benzene in bone marrow is believed to be due to the oxidation of catechol and hydroquinone catalyzed by myeloperoxidase [48].

The oxidation of diphenols to quinones is reversible; a variety of cellular reductants are able to mediate the reduction of quinones either by a two-electron mechanism or by two single-electron steps. The two-electron reduction can be catalyzed by carbonyl reductase and quinone reductase, while cytochrome P450 and some flavoproteins act by single-electron transfers. The nonenzymatic reduction of quinones can occur, for example, in the presence of $O_2^{\bullet -}$ or some thiols such as glutathione.

Olefinic bonds in xenobiotic molecules can also be targets of cytochrome P450-catalyzed epoxidation (Box **B** in Figure 24.4). In contrast to arene oxides, the resulting epoxides are fairly stable and can be isolated and characterized. But like arene oxides, they are substrates of epoxide hydrolase to yield dihydrodiols. This is exemplified by carbamazepine, whose 10,11-epoxide is a major and pharmacologically active metabolite in humans, and is further metabolized to the inactive dihydrodiol [49].

The few drugs that contain an acetylenic moiety are also targets for cytochrome P450-catalyzed oxidation (Box **C** in Figure 24.4). Oxygenation of the C-C triple bond yields an intermediate, which—depending on the substrate—can react in a number of ways, for example, by binding covalently to the enzyme or forming a highly reactive ketene whose hydration produces a substituted acetic acid [50].

C. Oxidation and Reduction of N- and S-Containing Moieties

The main metabolic reactions of oxidation and reduction of N- or S-containing functional groups in organic molecules are summarized in Figure 24.5. The reactions of oxidation are catalyzed mainly by cytochromes P450 and/or flavin-containing monooxygenases, whereas the enzymes catalyzing reductions are not always characterized and can be CYPs, NADPH-CYP reductase, and a variety of other reductases [12,19,23].

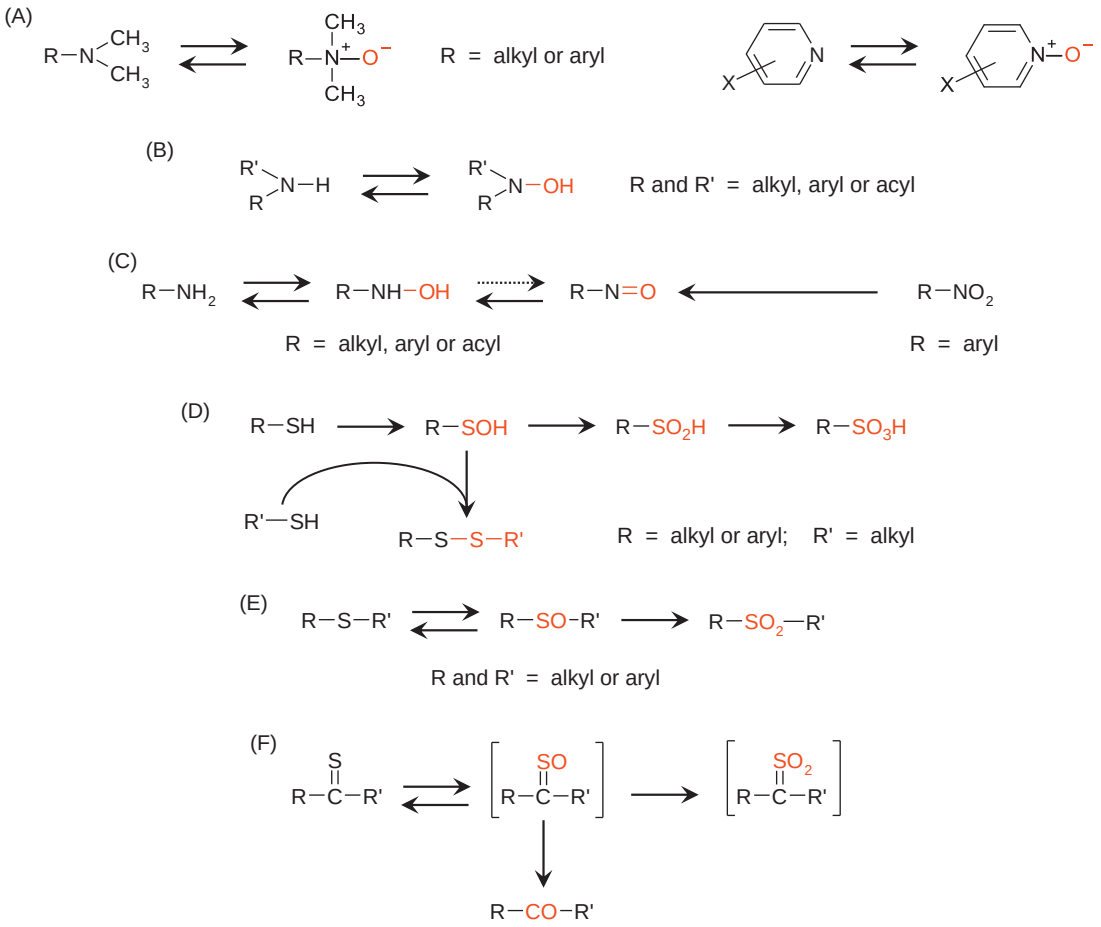


FIGURE 24.5 Major functionalization reactions involving N- or S-containing moieties in xenobiotics. The reactions shown here are mainly oxidations (oxygenations and dehydrogenations) and reductions (deoxygenations and hydrogenations).

Nitrogen oxygenation is a straightforward metabolic reaction of tertiary amines (Box **A** in Figure 24.5), both aliphatic or aromatic. Numerous drugs undergo this reaction, and the resulting N-oxide metabolite is more polar and hydrophilic than the parent compound. Identical considerations apply to pyridines and analogous aromatic azaheterocycles. These reactions can be reversible. Secondary and primary amines (Boxes **B** and **C** in Figure 24.5) also undergo N-oxygenation. The first isolable metabolites are hydroxylamines. Again, reversibility is documented. These compounds can be aliphatic or aromatic amines, and the same metabolic pathway occurs in secondary and primary amides, while tertiary amides appear resistant to N-oxygenation. The oxidation of secondary amines and amides usually stops at the hydroxylamine-hydroxylamide level. As opposed to secondary amines, primary amines (Box **C** in Figure 24.5) can sometimes be further oxidized to nitroso metabolites, but oxidation of the latter metabolites to nitro compounds does not seem to occur *in vivo*. In contrast, aromatic nitro compounds are readily reduced to primary amines. This is the case for numerous chemotherapeutic drugs such as metronidazole [23].

The "mitochondrial amidoxime reducing component" (mARC) is a newly discovered fourth molybdenum enzyme in mammals. All hitherto analyzed mammals express two mARC proteins, referred to as mARC1 and mARC2. Together with their electron-transport proteins cytochrome b5 and NADH-cytochrome b5 reductase, they form a three-component enzyme system and catalyze the reduction of N-hydroxylated structures (e.g., amidoximes used as prodrugs of amidines; Figure 24.6) [51].

A limited number of drugs contain a sulfur atom. Thus, thiol compounds (Box **D** in Figure 24.5) can be oxidized to sulfenic acids, to sulfinic acids, and finally to sulfonic acids. Depending on the substrate, the pathway is mediated by CYP and/or FMO. Another route for oxidation of thiols is to disulfides, usually by dehydration between a thiol and a sulfenic acid. The metabolism of sulfides (thioethers) is rather straightforward (Box **E** in Figure 24.5). Besides S-dealkylation reactions discussed earlier, these compounds can also be oxygenated by monooxygenases to sulfoxides and then to sulfones. Sulfoxides can be reduced to sulfides, whereas the formation of sulfones is irreversible.

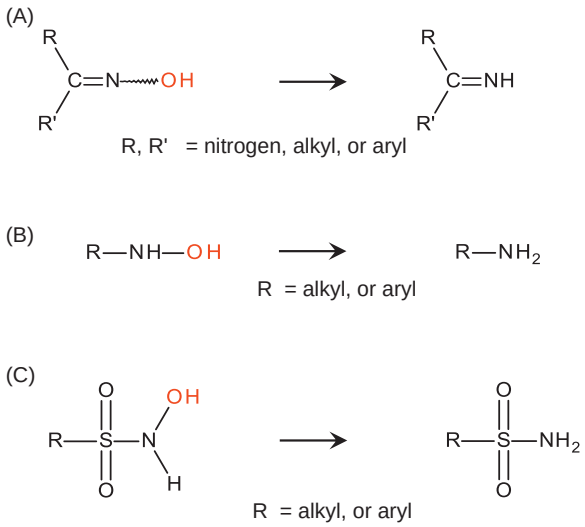


FIGURE 24.6 Examples of reactions catalyzed by the mitochondrial amidoxime reducing component (mARC): **A**: Reduction of N-hydroxylated amidines, guanidines and oximes. **B**: Reduction of hydroxylamines (sulfamethoxazole, N-hydroxylated base analogs). **C**: Reduction of sulfohydroxamic acids.

Thiocarbonyl compounds (Box **F** in Figure 24.5) are also substrates of monooxygenases, forming S-monoxides (sulfines) and then S-dioxides (sulfenes). As a rule, these metabolites cannot be identified as such due to their reactivity. Thus, S-monoxides rearrange to the corresponding carbonyl by expelling a sulfur atom, a reaction known as oxidative desulfuration and occurring in thioamides and thioureas (e.g., thiopental). As for the S-dioxides, they are strong electrophiles that react very rapidly with nucleophiles and particularly with nucleophilic sites in biological macromolecules. This covalent binding results in the formation of adducts of toxicological significance. Such a mechanism is believed to account for the carcinogenicity of a number of thioamides.

Some other elements besides carbon, nitrogen, and sulfur can undergo metabolic redox reactions. The direct oxidation of oxygen atoms in phenols and alcohols is well documented for some substrates. Thus, the oxidation of secondary alcohols by some peroxidases can yield a hydroperoxide and ultimately a ketone. Some phenols are known to be oxidized by cytochrome P450 to a semiquinone and ultimately to a quinone. A classic example is that of the analgesic drug acetaminophen, a minor fraction of which is oxidized by CYP2E1 to a highly reactive and toxic quinone imine [19]. Additional elements of limited significance in medicinal chemistry able to enter redox reactions are silicon, phosphorus, arsenic, and selenium, among others.

D. Reactions of Hydration and Hydrolysis

The two terms of hydrolysis and hydration both imply bond breakage with addition of a molecule of H_2O. In this text, we prefer to apply the term "hydrolysis" to the cleavage of esters (carboxylesters, lactones, and inorganic esters), amides (e.g., carboxamides, sulfamates, phosphoamides, and lactams), and glycosides. In contrast, the term "hydration" will be restricted to epoxides, although the enzymes catalyzing this reaction are also classified as hydrolases (see Table 24.4). More extensive treatments of hydrolases and their metabolic reactions can be found elsewhere [1,7,22].

1. Esters and Amides

The hydrolysis of esters and amides—including peptides—have been published. Some of these are discussed in Chapter 30. Here, two drugs are presented to illustrate these two chemical classes. (-)-Cocaine (**2**) has two ester groups, whose hydrolysis (Box **A** in Figure 24.7) is a route of detoxification, which accounts for as much as 90 percent of the dose in humans [7,52,53]. Three human enzymes are now known to be involved in the hydrolysis of cocaine. One is the liver carboxylesterase hCE-1, which catalyzes the hydrolysis of the methyl ester group. As for the benzoyl ester goup, it is hydrolyzed by the liver carboxylesterase hCE-2 and serum cholinesterase.

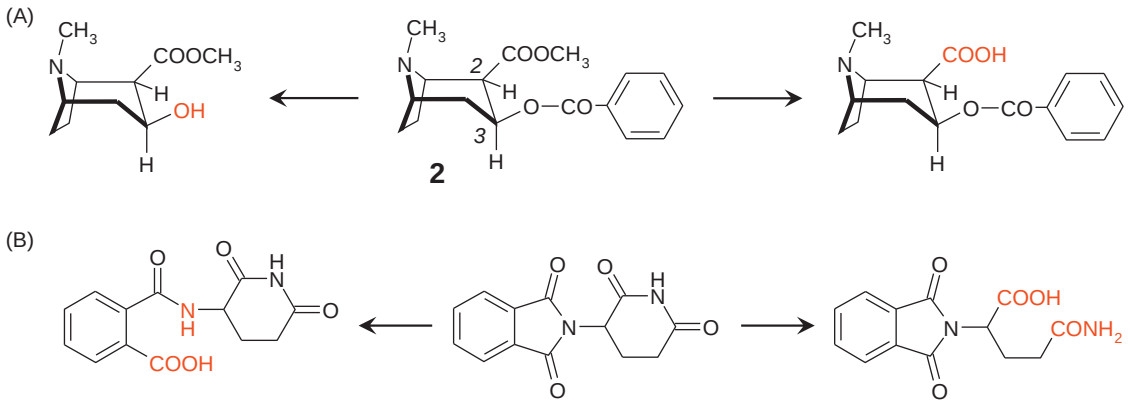


FIGURE 24.7 Examples of hydrolytic reactions: **A**: The diester cocaine (**2**). **B**: The diimide thalidomide (**3**).

Among amides, we turn our attention to a cyclic analog, namely the infamous thalidomide (**3**), which bears two imide rings. Parallel to rapid inversion of configuration and very low rates of hydroxylation, thalidomide is rapidly hydrolyzed to ring-opened products (Box **B** in Figure 24.7) [54,55]. All four imide bonds of the molecule are susceptible to hydrolytic cleavage at pH $>$ 6, and the reactions are nonenzymatic and base-catalyzed. The two main urinary metabolites in humans are shown here, each accounting for about 30–50 percent of a dose.

2. *Epoxides*

The overall reaction catalyzed by epoxide hydrolases (Table 24.4) is the addition of a water molecule to an epoxide. Together with glutathione conjugation, hydration is a major pathway in the inactivation and detoxification of arene oxides. As a rule, these are good substrates of microsomal epoxide hydrolase, yielding *trans*-dihydrodiols (see Box **A** in Figure 24.4). In phenyl and naphthyl rings, the proton-catalyzed isomerization of epoxides to phenols is an extremely fast reaction that markedly reduces the likelihood of the epoxide being hydrated by epoxide hydrolase. This chemical instability decreases for chemicals with three or more fused rings, but such compounds are no longer of medicinal interest. Yet despite the high reactivity of benzene epoxides, the characterization of a dihydrodiol metabolite has been achieved for a limited number of phenyl-containing drugs, particularly for neurodepressant drugs such as hypnotics (e.g., glutethimide) and antiepileptics (e.g., ethotoin and phenytoin).

Alkene oxides are generally quite stable chemically, indicating a much-reduced chemical reactivity compared to arene oxides. Under physiologically relevant conditions, they have little capacity to undergo rearrangement reactions and are resistant to uncatalyzed hydration. In contrast, they are often good substrates of epoxide hydrolases, yielding diols (see Box **B** in Figure 24.4). A well-known example is that of the anticonvulsant drug carbamazepine. Thus, the 10,11-epoxide and the 10,11-dihydrodiol are urinary metabolites in humans and rats given the drug.

III. CONJUGATION REACTIONS

A. Introduction

Conjugation reactions (also known as phase II reactions) are of critical significance in the metabolism of endogenous compounds, as witnessed by the impressive battery of enzymes that have evolved to catalyze them. Conjugation is also of great importance in the biotransformation of xenobiotics, involving parent compounds or metabolites thereof [2,15,20,25]. Conjugation reactions of xenobiotics and their metabolites are characterized by a number of criteria:

A. Substrates are coupled covalently to an endogenous molecule or moiety.
B. That molecule or moiety is polar (hydrophilic).
C. Conjugation reactions are catalyzed by enzymes known as transferases.
D. They involve a cofactor that binds to the enzyme in close proximity to the substrate and carries the endogenous molecule or moiety to be transferred.

It is important from a biochemical and practical viewpoint to note that criterion A is essential in defining conjugation reactions of xenobiotics. In contrast, criteria B to D are neither sufficient nor necessary. They are not sufficient, since in hydrogenation reactions (i.e., typical reactions of functionalization) the hydride is also transferred from a cofactor (NADPH or NADH). They are not necessary, since all the above criteria suffer from some important exceptions discussed below (e.g., some glutathione conjugations are nonenzymatic). A survey of transferases involved in drug metabolism is offered in Table 24.5.

TABLE 24.5 A Survey of Transferases (EC 2) [2,20,25,36]

Cofactor	Examples of enzymes (with some *gene roots* and human enzymes)
METHYLTRANSFERASES (EC 2.1.1)	
S-Adenosyl-L-methionine (SAM)	EC 2.1.1.6: Catechol O-methyltransferase (*COMT*)
	EC 2.1.1.1: Nicotinamide N-methyltransferase (*NNMT*)
	EC 2.1.1.8: Histamine N-methyltransferase (*HNMT*)
	EC 2.1.1.28: Noradrenaline N-methyltransferase (*PNMT*)
	EC 2.1.1.49: Arylamine N-methyltransferase, indolethylamine N-methyltransferase (*INMT*)
	EC 2.1.1.9: Thiol S-methyltransferase (TMT)
	EC 2.1.1.67: Thiopurine S-methyltransferase (*TPMT*)
SULFOTRANSFERASES (EC 2.8.2) (SULT)	
3′-Phosphoadenosine 5′-phosphosulfate (PAPS)	EC 2.8.2.1: Aryl sulfotransferase (SULT1A1, 1A2 and 1A3)
	EC 2.8.2.4: Estrogen sulfotransferase (SULT1E1)
	EC 2.8.2.14: Bile salt sulfotransferase (SULT2A1)
	EC 2.8.2.2: Alcohol sulfotransferase (SULT2B1)
	EC 2.8.2.15: Steroid sulfotransferase
	EC 2.8.2.18: Cortisol sulfotransferase
	EC 2.8.2.3: Amine sulfotransferase (SULT3)
UDP-GLUCURONOSYLTRANSFERASES (2.4.1.17) (UGT)	
Uridine-5′-diphospho-α-D-glucuronic acid (UDPGA)	Subfamily UGT1: UGT1A1, 1A3, 1A4, 1A5 to 1A10
	Subfamily UGT2A: UGT2A1 to 2A3
	Subfamily UGT2B: UGT2B4, 2B7, 2B10, 2B15, 2B17, 2B28
	Subfamily UGT3A: UGT3A1, 3A2
	Subfamily UGT8A: UGT8A1
ACETYLTRANSFERASES	
Acetylcoenzyme A (AcCoA)	EC 2.3.1.5: N-Acetyltransferase (*NAT*) NAT1 and NAT2
	EC 2.3.1.56: Aromatic-hydroxylamine O-acetyltransferase
	EC 2.3.1.118: N-Hydroxyarylamine O-acetyltransferase
ACYL-COA SYNTHETASES	
Coenzyme A (CoA)	EC 6.2.1.1: Short-chain fatty acyl-CoA synthetase (*ACSS*)
	EC 6.2.1.2: Medium-chain acyl-CoA synthetase
	EC 6.2.1.3: Long-chain acyl-CoA synthetase (*ACSL*)
	EC 6.2.1.7: Cholate-CoA ligase
	EC 6.2.1.25: Benzoyl-CoA synthetase

(*Continued*)

TABLE 24.5 (Continued)

Cofactor	Examples of enzymes (with some *gene roots* and human enzymes)
ACYLTRANSFERASES	
Xenobiotic acyl-Coenzyme A	EC 2.3.1.13: Glycine N-acyltransferase (*GLYAT*)
	EC 2.3.1.71: Glycine N-benzoyltransferase
	EC 2.3.1.14: Glutamine N-phenylacetyltransferase
	EC 2.3.1.68: Glutamine N-acyltransferase
	EC 2.3.1.65: Cholyl-CoA glycine-taurine *N*-acyltransferase (*BAAT*)
GLUTATHIONE S-TRANSFERASES (EC 2.5.1.18) (GST)	
(Glutathione)	Microsomal GST superfamily (homotrimers): *MGST*: GST1 to GST3
	Cytoplasmic GST superfamily (homodimers, and a few heterodimers): *GST A*: Alpha class, GST A1-1, A1-2, A2-2, A3-3, A4-4, A5-5
	GST K: Kappa class, GST K1-1
	GST M: Mu class, GST M1-1, M2-2, M3-3, M4-4, M5-5
	GST O: Omega class, GST O1-1, O2
	GST P: Pi class: GST P1-1
	GST T: Theta class, GST T1-1, T2
	GST Z: Zeta class, GST Z1

B. Methylation

Reactions of methylation imply the transfer of a methyl group from the cofactor S-adenosyl-L-methionine (SAM). A number of methyltransferases are able to methylate xenobiotic molecules and metabolites [56]. They target the *ortho*-diphenol moiety (COMT), N-atoms in aromatic azaheterocycles such as pyridines (NNMT), primary arylamines (INMT), endocyclic secondary amines (HNMT), and a variety of thiols including heterocyclic ones (TMP, TPMT). O-methylation (Box **A** in Figure 24.8) is a common reaction of compounds containing a catechol moiety (**4**), with a usual regioselectivity for the meta position. The substrates can be xenobiotics and particularly drugs, L-dopa being a classic example. Frequently, O-methylation occurs as a late event in the metabolism of aryl groups, after they have been oxidized to catechols. A toxicologically relevant reaction of N-methylation (Box **B** in Figure 24.8) is that of theophylline (**5**) to yield caffeine [57]. This reaction is not seen in adult humans but is effective in neonates (5–10 percent of a dose of theophylline), where it causes unwanted side effects.

C. Sulfonation

Sulfonation reactions (also less correctly known as sulfation reactions) consist in a sulfate being transferred from the cofactor 3′-phosphoadenosine 5′-phosphosulfate (PAPS) to the substrate under catalysis by a sulfotransferase. Sulfotransferases, which catalyze a variety of physiological reactions, are soluble enzymes [24,58–62]. The most significant for drug metabolism are listed in Table 24.5. The sulfate moiety in PAPS is linked to a phosphate group by an anhydride bridge whose cleavage is exothermic and supplies enthalpy to the reaction. The nucleophilic -OH or -NH- site in the substrate will react with the leaving SO_3^- moiety, forming an ester sulfate or a sulfamate. Some of these conjugates are unstable under biological conditions and will form electrophilic intermediates of considerable toxicological significance [2].

The sulfoconjugation of alcohols leads to metabolites of different stabilities. Endogenous hydroxysteroids (i.e., cyclic secondary alcohols) form relatively stable sulfates, while some secondary alcohol metabolites of allylbenzenes (e.g., safrole and estragole) form highly genotoxic carbenium ions. In contrast to alcohols, phenols form stable sulfate esters. The reaction is usually of high affinity (i.e., rapid), but the limited availability of PAPS restricts the amounts of conjugate being produced. Typical drugs undergoing limited sulfonation (Box **C** in Figure 24.8) include paracetamol and diflunisal (**6**). Aromatic hydroxylamines and hydroxylamides are good

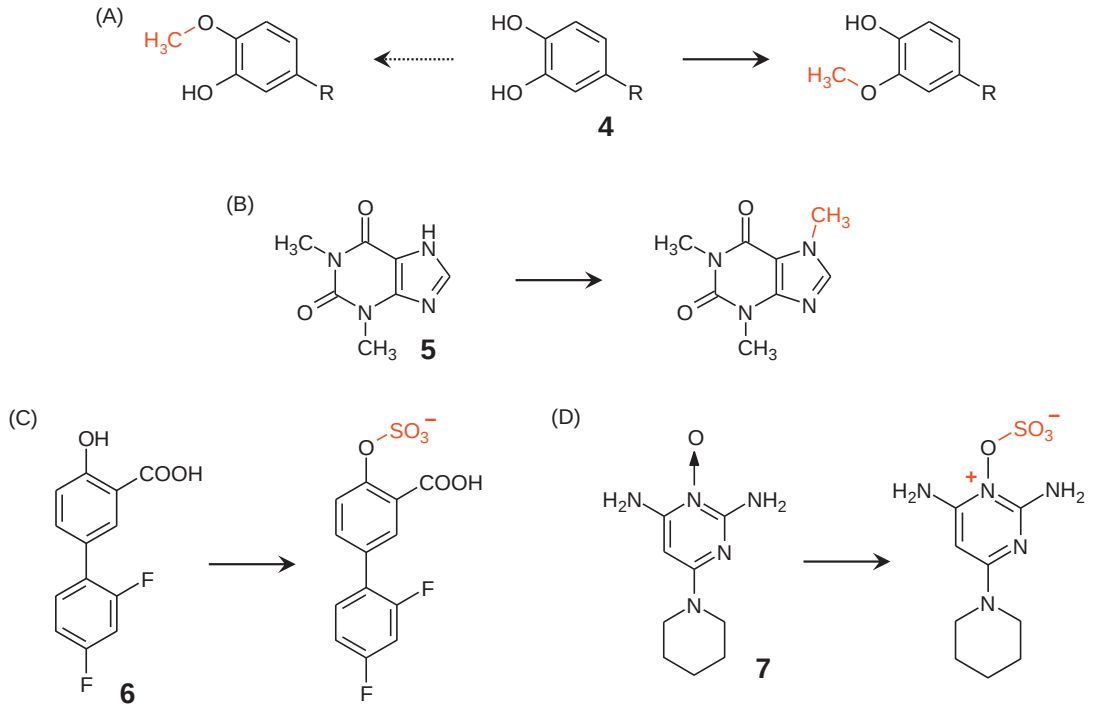


FIGURE 24.8 A few examples of reactions of methylation (**A** and **B**) and sulfonation (**C** and **D**). The substrates are catechols (**4**), theophylline (**5**), diflunisal (**6**), and minoxidil (**7**).

substrates for some sulfotransferases and yield reactive sulfate esters [48]. In contrast, significantly more stable products are obtained upon N-sulfoconjugation of amines. An intriguing and rare reaction of conjugation (Box **D** in Figure 24.8) occurs for minoxidil (**7**). This drug is an N-oxide, and the actual active form responsible for the different therapeutic effects is the N,O-sulfate ester.

D. Glucuronidation

Glucuronidation is a major and very frequent reaction of conjugation. It involves the transfer to the substrate of a molecule of glucuronic acid from the cofactor uridine-5′-diphospho-α-D-glucuronic acid (UDPGA). As listed in Table 24.5, the enzyme catalyzing this reaction consists of a number of proteins coded by genes of the *UGT* superfamily [25,63–66]. The human UDPGA transferases (UGT) known to metabolize xenobiotics are the products of two gene families, *UGT1* and *UGT2*. Glucuronic acid exists in UDPGA in the 1α-configuration, but the products of conjugation are β-glucuronides (**8**; Figure 24.9). This is due to the mechanism of the reaction being a nucleophilic substitution with inversion of configuration (SN_2). Indeed, all functional groups able to undergo glucuronidation are nucleophiles, a common characteristic they share despite their great chemical variety.

O-glucuronidation is a frequent metabolic reaction of xenobiotic phenols and alcohols, yielding polar metabolites excreted in urine and/or bile. An important example is that of morphine (**9**), which is conjugated on its phenolic and secondary alcohol groups to form the 3-O-glucuronide (a weak opiate antagonist) and the 6-O-glucuronide (a strong opiate agonist), respectively [67]. O-glucuronidation is often in competition with O-sulfonation, with the latter reaction predominating at low doses and the former at high doses. In biochemical terms, glucuronidation is a reaction of low affinity and high capacity, while sulfonation displays high affinity and low capacity. In general, glucuronidation is the preferred reaction in humans.

An important pathway of O-glucuronidation is the formation of acyl-glucuronides. Substrates are anti-inflammatory arylacetic acids and aliphatic acids such as valproic acid (**10**). These metabolites are quite reactive, rearranging to positional isomers and binding covalently to plasma and seemingly also to tissue proteins [68]. Thus, acyl-glucuronide formation cannot be viewed solely as a reaction of inactivation and detoxification.

Second in importance to O-glucuronides are the N-glucuronides formed from amides and amines. The reaction has special significance for antibacterial sulfanilamides such as sulfadimethoxine (**11**), since it produces highly

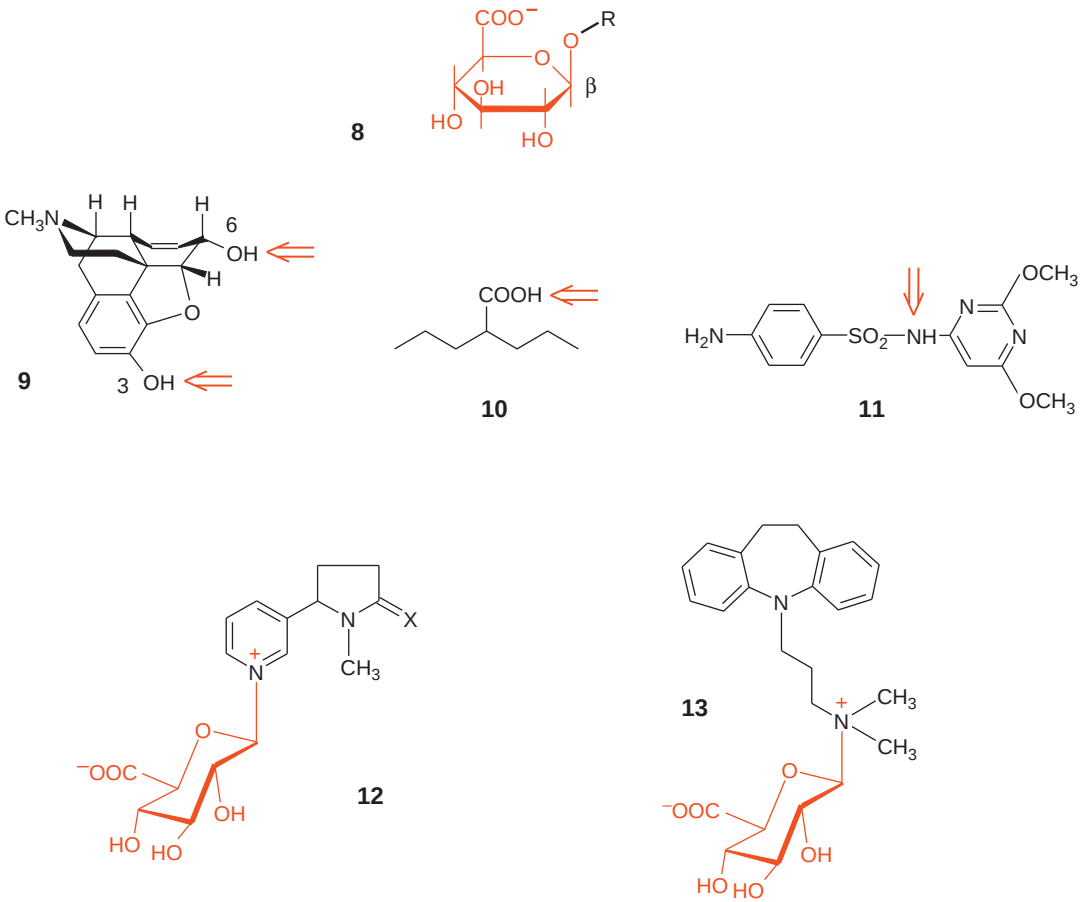


FIGURE 24.9 The generic structure of O-β-D-glucuronides is shown as **8**. Morphine (**9**) exemplifies the glucuronidation of phenols and alcohols, whereas valproic acid (**10**) forms an acylglucuronide. Sulfadimethoxine (**11**) exemplifies the N-glucuronidation of amides. The zwitterionic N-glucuronides are those of nicotine (**12**; X = 2H), cotinine (**12**; X = O), and imipramine (**13**). The double arrows point to the target sites.

water-soluble metabolites that show no risk of crystallizing in the kidneys. For amines, a number of observations show that pyridine-type nitrogens can be N-glucuronidated, as is illustrated by the N-glucuronide of nicotine (**12**, X = 2H) and cotinine (**12**, X = O). Another reaction of significance is the N-glucuronidation of lipophilic basic tertiary amines. More and more drugs of this type (e.g., antihistamines and neuroleptics) are found to undergo this reaction to a marked extent in humans, as illustrated by the N-glucuronide of imipramine (**13**) [25].

E. Acetylation

The major enzyme system catalyzing acetylation reactions is arylamine N-acetyltransferase. Two enzymes have been characterized, NAT1 and NAT2, the latter showing considerably reduced levels in slow acetylators (i.e., subjects expressing a mutated NAT2 protein) [69–71]. Two other activities, aromatic-hydroxylamine O-acetyltransferase and N-hydroxyarylamine O-acetyltransferase, are also involved in the acetylation of aromatic amines and hydroxylamines (Table 24.5). The coenzyme in acetylation reactions is acetyl-Coenzyme A. The substrates of acetylation are mainly amines of medium basicity. Very few basic amines (primary or secondary) of medicinal interest have been reported to form N-acetylated metabolites, and when they did, the yields were low. In contrast, a large variety of primary aromatic amines are N-acetylated. Thus, several drugs such as sulfonamides and *para*-aminosalicylic acid (**14**; Figure 24.10) are acetylated to large extents, not to mention various carcinogenic amines such as benzidine. The same is true of hydrazines and hydrazides, such as isoniazid (**15**).

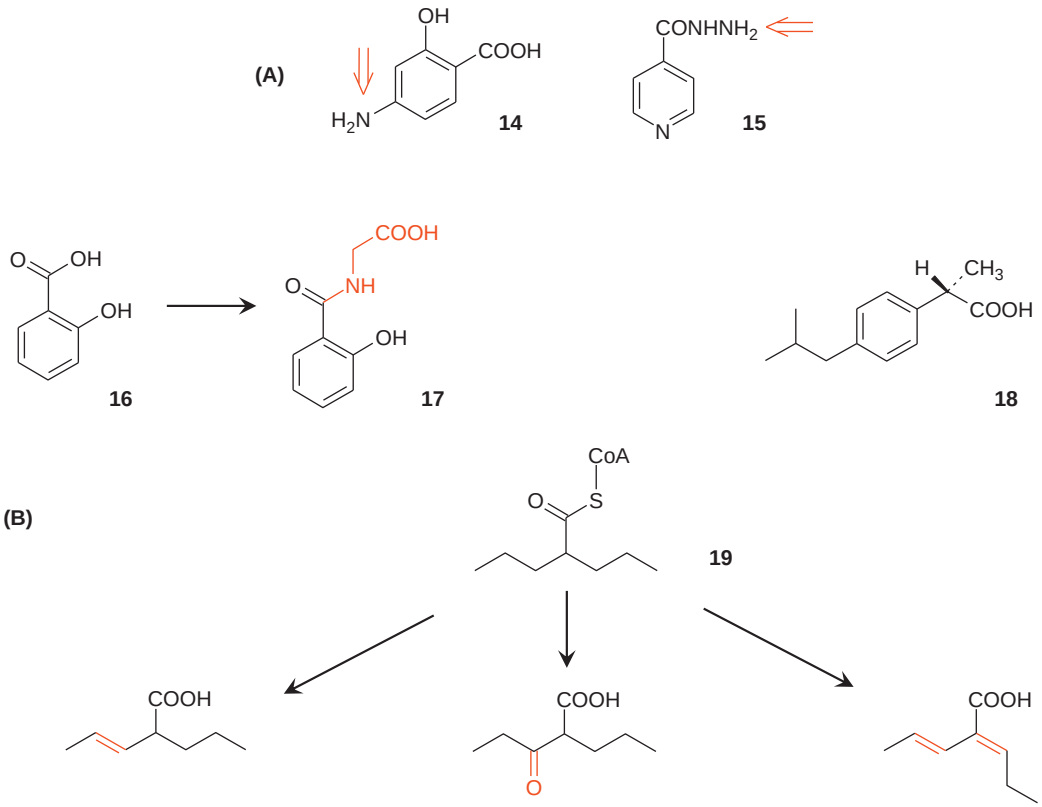


FIGURE 24.10 Metabolic reactions having in common the involvement of Coenzyme A: **A**: Two examples of N-acetylated substrates, namely *para*-aminosalicylic acid (**14**) and isoniazid (**15**). The double arrows point to the target sites. **B**: Some reactions consecutive to the formation of xenobiotic acyl-CoA conjugates. Well-known substrates are salicylic acid (**16**), which forms salicyluric acid (**17**), (*R*)-ibuprofen (**18**), and valproic acid, whose intermediate acyl-CoA thioester is shown here (**19**).

F. Conjugation with Coenzyme A and Subsequent Reactions

The reactions described in this subsection all have in common the fact that they involve xenobiotic carboxylic acids (R-COOH) forming an acyl-CoA thioester (R-CO-S-CoA) as the metabolic intermediate and as a cofactor. The reaction requires ATP and is catalyzed by various acyl-CoA synthetases also known as acyl-CoA ligases (Table 24.5) of overlapping substrate specificity. The acyl-CoA conjugates thus formed are seldom excreted, but they can be isolated and characterized relatively easily in *in vitro* studies. In the present context, the interest of acyl-CoA conjugates is their further transformation by a considerable variety of pathways [25,72–74], as summarized in Table 24.6.

Amino-acid conjugation is a major route for a number of small aromatic acids and involves the formation of an amide bond between the xenobiotic acyl-CoA and the amino acid. Glycine is the amino acid most frequently used for conjugation, as illustrated by the formation of salicyluric acid (**17**) from salicylic acid (**16**; Figure 24.10), although some glutamine and taurine conjugates have also been characterized in humans. The enzymes catalyzing these transfer reactions are various N-acyltransferases listed in Table 24.5.

Incorporation of xenobiotic acids into lipids forms highly lipophilic metabolites that may burden the body as long retained residues. In the majority of cases, triacylglycerol analogs or cholesterol esters are formed. One telling example is that of ibuprofen (**18** in Figure 24.10), a much-used anti-inflammatory drug whose (*R*)-enantiomer forms hybrid triglycerides detectable in rat liver and adipose tissues.

In addition, the lesser active (*R*)-ibuprofen enantiomer (distomer; see Chapter 18) and a few other arylpropionic acids (i.e., profens) undergo an intriguing metabolic reaction such that the (*R*)-enantiomer is converted to the (*S*)-enantiomer, while the reverse reaction is negligible. This unidirectional chiral inversion is thus a reaction of bioactivation [75].

In some cases, acyl-CoA conjugates formed from xenobiotic acids can also enter the physiological pathways of fatty acids catabolism or anabolism. A few examples are known of xenobiotic alkanoic and arylalkanoic acids

TABLE 24.6 Metabolic Consequences of the Conjugation of Xenobiotic Acids with Coenzyme A [25,72–74]

Initial reaction	**R-COOH → R-CO-S-CoA → •**
Subsequent metabolic options	• Hydrolysis (futile cycle) • Formation of amino acid conjugates (glycine, glutamic acid, taurine, . . .) • Formation of hybrid triglycerides • Formation of cholesteryl and bile acid esters • Formation of acyl-carnitines • Unidirectional chiral inversion of arylpropionic acids (profens) • Dehydrogenation and β-oxidation • 2-Carbon chain elongation

undergoing two-carbon chain elongation or two-, four-, or even six-carbon chain shortening. In addition, intermediate metabolites of β-oxidation may be seen, as illustrated in Figure 24.10 with valproic acid, whose acyl-CoA intermediate (**19**) is a substrate for some first steps of β-oxidation [76].

G. Conjugation Reactions of Glutathione

1. *Introduction*

Glutathione (**20** in Figuer 24.11; GSH) is a thiol-containing tripeptide of major significance in the detoxification and toxification of drugs and other xenobiotics. In the body, it exists in a redox equilibrium between the reduced form (GSH) and an oxidized form (GS-SG). The metabolism of glutathione (i.e., its synthesis, redox equilibrium, and degradation) is quite complex and involves a number of enzymes [25,77–79]. Glutathione reacts in a variety of ways, one of which is its redox capacity. Indeed, GSH can reduce peroxides (a reaction catalyzed by glutathione peroxidase) and organic nitrates. In its GSSG form, glutathione can oxidize the superoxide anion-radical. Of major significance in detoxification reactions is the capacity of GSH (and other endogenous thiols including albumin) to scavenge free radicals, especially reactive oxygen species (ROSs; e.g., $HO^{\bullet}$, $HOO^{\bullet}$, $ROO^{\bullet}$). As such, glutathione and other thiols have a critical role to play in cellular protection.

In this chapter, we focus on the conjugation reactions of glutathione as catalyzed by glutathione transferases. The glutathione transferases are multifunctional proteins coded by two multigene superfamilies (Table 24.5) [80–84]. Seven classes are now known in humans. The conjugating reactivity of glutathione is due to its thiol group (pK_a 9.0), which makes it a highly effective nucleophile. This nucleophilic character is greatly enhanced by deprotonation to a thiolate. In fact, an essential component of the catalytic mechanism of glutathione transferases is the marked increase in acidity (pK_a decreased by 2–3 units) experienced by the thiol group upon binding of glutathione to the active site of the enzyme [81]. As a result, GSTs transfer glutathione to a very large variety of electrophilic groups. Depending on the nature of the substrate, the reactions can be categorized as nucleophilic substitutions or nucleophilic additions. With compounds of sufficient reactivity, these reactions can also occur nonenzymatically [81,85].

Once formed, glutathione conjugates may be excreted as such (they are best characterized *in vitro* or in the bile of laboratory animals), but they usually undergo further biotransformation prior to urinary or fecal excretion. Cleavage of the glutamyl and glycyl residues by γ-glutamyltransferase and dipeptidases leave a cysteine conjugate, which is further N-acetylated by cysteine-S-conjugate N-acetyltransferase (EC 2.3.1.80) to yield an N-acetylcysteine conjugate. The latter type of conjugates are known as mercapturic acids (**21**; Figure 24.11). This, however, does not imply that the degradation of unexcreted glutathione conjugates must stop at this stage, since cysteine conjugates can be substrates of cysteine-S-conjugate β-lyase (EC 4.4.1.13) to yield thiols (R-SH). These, in turn, can rearrange, be oxidized, or be S-methylated and then S-oxygenated to yield thiomethyl conjugates (R-S-Me), sulfoxides (R-SO-Me), and sulfones ($R\text{-}SO_2\text{-}Me$).

2. *Substrates of Glutathione Transferases*

Most known cases of glutathione conjugation are nucleophilic attacks at electron-deficient carbon atoms, but attack of a nitrogen atom (e.g., in an aromatic nitroso group), a sulfur atom (in thiols), or an oxygen

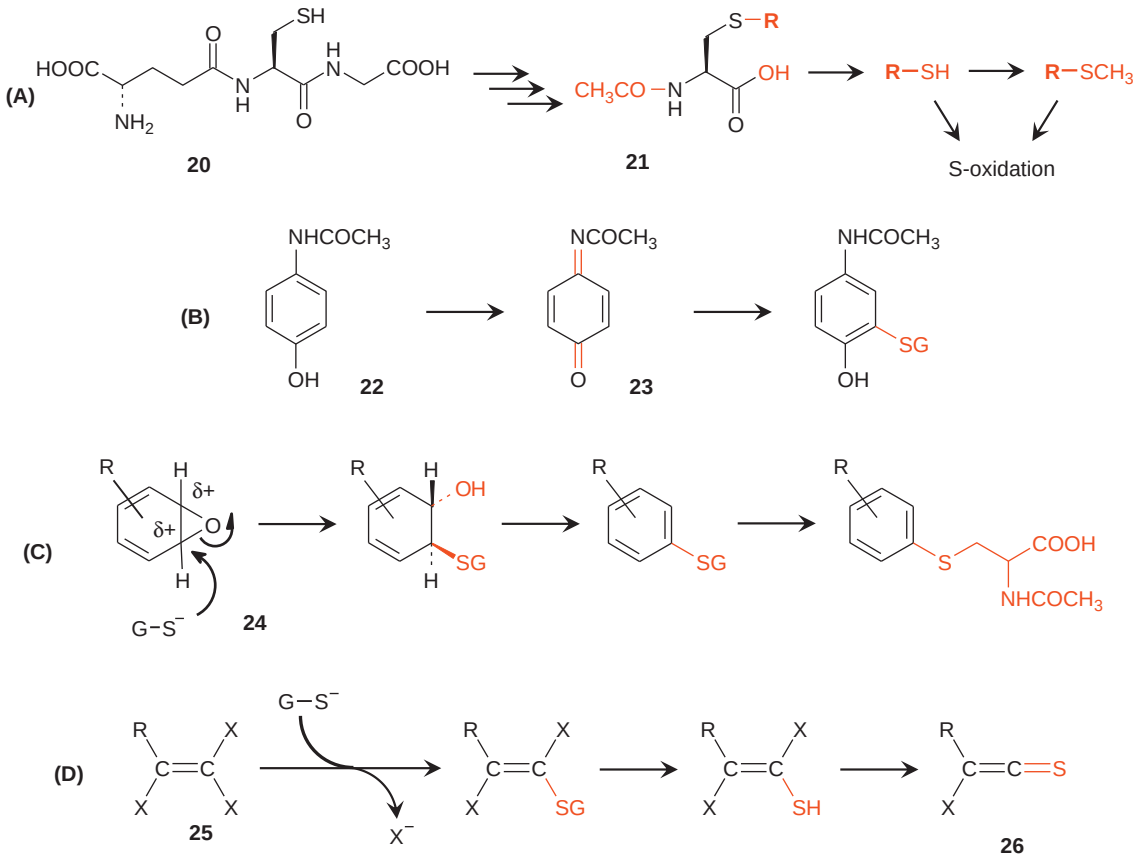


FIGURE 24.11 **A**: The structure of glutathione (**20**), mercapturic acids (**21**), and further degradation products [25]. **B**: The metabolism of acetaminophen (**22**) to its glutathione conjugate via the intermediate quinone imine **23**. **C**: The conjugation of arene oxides (**24**) to mercapturic acids. **D**: The glutathione-mediated toxification of haloalkenes (**25**) to thioketenes (**26**).

(in hydroperoxides) is also documented. Nucleophilic additions can involve metabolites arising from oxidation reactions, but they can also occur as primary metabolic reactions. Frequent cases of GSH addition are to α,β-unsaturated carbonyls; a typical xenobiotic substrate is the toxin acrolein ($CH_2 = CH\text{-}CHO$). Attack occurs at the activated CH_2 group. Quinones (*ortho-* and *para-*) and quinone imines are structurally very similar to α,β-unsaturated carbonyls. The reaction has physiological significance, since endogenous metabolites such as quinone metabolites of estrogens are conjugated to glutathione. A medicinal example is provided by the toxic quinoneimine metabolite (**23**; Figure 24.11) of acetaminophen (**22**). Its glutathione conjugate is not excreted as such in humans dosed with the drug, but as the mercapturic acid. The reaction is one of major detoxification, the quinone imine being extremely hepatotoxic and resulting in liver necrosis, liver failure, and even death when produced at levels and rates that oversaturate the GSH conjugation pathway. Nevertheless, the GSH conjugation of quinones and quinone imines is not always a reaction of detoxification, as some of these conjugates are known to undergo further transformations leading to reactive products [86].

An important role of GSH is in the conjugation of arene oxides, particularly those that rearrange slowly to the phenol and are poor substrates of epoxide hydrolase. The first reaction is again a nucleophilic addition to the epoxide (**24**; Figure 24.11). The resulting non-aromatic conjugate then dehydrates to an aromatic GSH conjugate, followed by a cascade leading to the mercapturic acid as also shown in Box **C** (Figure 24.11). This is a common reaction of metabolically produced arene oxides, as documented for naphthalene, numerous drugs, and xenobiotics containing a phenyl moiety. Note that the same reaction can also occur readily for epoxides of olefins.

Glutathione conjugations occurring by a mechanism of nucleophilic substitution (including addition-elimination) are documented for a number of industrial xenobiotics as well as drugs. This is the case for compounds having an activated alkyl moiety, such as the $-CH_2Cl$ group of nitrogen mustards, which yields conjugates with structure $-CH_2SG$. Haloalkenes (**25**; Figure 24.11) are a special group of substrates of GS-transferases, as they may react with GSH either by substitution to form an alkene conjugate as shown or by addition to form an

alkane conjugate (not shown). Formation of mercapturic acids occurs as for other glutathione conjugates, but in both routes S-C cleavage of the S-cysteinyl or N-acetyl-S-cysteinyl conjugates by renal β-lyase yields thiols of significant toxicity. Indeed, these thiols rearrange by hydrohalide expulsion to form highly reactive thioketenes (**26**) and/or thioacyl halides [87].

IV. BIOLOGICAL FACTORS INFLUENCING DRUG METABOLISM

A variety of genetic, physiological, and pathological factors influence xenobiotic metabolism and hence the wanted and unwanted activities associated with a drug. In this brief section, we restrict ourselves to a conceptual overview and invite readers to personal study [2,6,8,21,27,28,41].

The major difference is between inter-individual and intra-individual factors that influence drug metabolism (Table 24.7). The former remain constant throughout the life span of an organism and are the expression of its genome. In other words, they are written in the genome and the epigenome [27]. In contrast, intra-individual factors vary depending on time (age, time of day), pathological states, or external factors (nutrition, pollutants, drug treatment) [28].

Inter-individual factors include species differences, genetic differences between organisms in a given species (including genetic polymorphism and ethnic differences), and sex-related differences. Species differences have been well known since the beginnings of the science of drug metabolism. In contrast, genetic differences (pharmacogenetics) have gained significance only during the last two or three decades. Sex-related differences are well documented in laboratory rodents but until recently have received comparatively less attention in humans.

Intra-individual factors are those that change as a function of time for a given organism. Among physiological factors, the first is obviously age, with major differences seen in humans between infants, children, adults, and elderly persons. Biological rhythms include the 24-hour cycle (circadian rhythm), the 28-day cycle in women, and the yearly cycle; their study is known as chronopharmacology. Pregnancy is another physiological factor well studied in laboratory animals. For obvious reasons, little is known in humans. Pathological factors are the many diseases (e.g., fever, infections, cardiac or renal conditions) and abnormal conditions (e.g., stress) that perturb physiological equilibria and indirectly influence the body's response to drugs. And finally, we find the external influences. Differences due to diets (nutritional factors) remain modest. In contrast, enzyme induction and inhibition play a major role in increasing or decreasing (often strongly) the biotransformation of numerous drugs. This is a huge and critical issue in drug–drug interactions, whose study and significance keep growing with new drugs entering the market. Much attention has been paid to developing new drug candidates that are not metabolized by P450 enzymes and thus avoiding P450 related drug–drug interactions. As a consequence, however, metabolism by other drug-metabolizing enzymes like molybdenum-containing catalysts has become more important [17,51].

Note that several of the factors listed above influence not only biotransformation but can also affect absorption, distribution, and excretion by interacting with transporters and therapeutic effects by influencing drug targets.

TABLE 24.7 Biological Factors Affecting Xenobiotic Metabolism [2,27,28]

Inter-individual Factors	**Animal species**	
	Genetic and ethnic factors (genetic polymorphism)	
	Sex	
Intra-individual Factors	Physiological changes	Age
		Biological rhythms
		Pregnancy
	Pathological changes:	Disease
		Stress
	External influences	Nutrition
		Enzyme induction by drugs and other xenobiotics
		Enzyme inhibition by drugs and other xenobiotics

V. WHAT IS THE RELATIVE SIGNIFICANCE OF THESE MANY TYPES OF METABOLIC REACTIONS?

What is the medicinal chemist or the biochemist to do with the depth and breadth of our knowledge of xenobiotic metabolism, a knowledge we have tried to summarize, structure, and conceptualize in the previous pages? A first cause of perplexity for drug researchers may well be the relative importance of the numerous metabolic reactions, be it in quantitative or toxicological terms. Indeed, an informed view of the relative importance of drug-metabolizing reactions and enzymes has long been lacking, but a recent meta-analysis of the primary literature may begin to shed some light on these issues [33]. Specifically, 903 papers published in three peer-reviewed primary journals during the years 2004–2009 were selected according to objective criteria. These papers were analyzed and the results classified using a set of strict rules. The experimental metabolic data revealed 1,171 different substrates yielding a total of 6,767 different metabolites (a mean of 5.78 per substrate).

The primary objective of this meta-analysis was to offer an overview of the relative quantitative importance of biotransformation reactions in the metabolism of medicinal compounds and other xenobiotics. Table 24.8 breaks down the major reaction classes into individual reaction types as defined beforehand. As shown, reactions of Csp3-, Csp2-, and Csp-oxidation together accounted for 34.5 percent of all metabolites, while redox reactions to and from the carbonyl group formed 8.3 percent of metabolites. Redox reactions at nitrogen-containing groups led to ca. 5.5 percent of metabolites, whereas redox reactions at sulfur-containing groups formed 2.8 percent of metabolites. Redox reactions to form or reduce quinones and analogs accounted for 4.0 percent of all metabolites. Only 1.7 percent of metabolites were formed by unclassified redox reactions, mainly reduction of olefinic bonds.

Reactions of hydrolysis and hydration were mainly ester hydrolyses (ca. 3.8 percent), while unclassified reactions such as oxime or imine hydrolysis, hydration of iminium groups or other electrophiles, and hydrolytic dehalogenation or hydration of antitumor platinum compounds accounted for almost 2.8 percent of metabolites. Conjugations reactions are dominated by glucuronidations (ca. 14.1 percent), followed by enzymatic (or nonenzymatic) conjugations with glutathione (or in a few cases with cysteine or N-acetyl-Cys; 8.0 percent). Sulfonations of hydroxy or amino groups represented 4.8 percent of metabolites.

Another objective was to obtain an estimate of the relative toxicological significance of the various reaction types listed in Table 24.8. As seen, the last column of the table reports the detection and detailed breakdown of 473 toxic or reactive metabolites, accounting for almost 7.0 percent of the total number of metabolites. To be counted in this class, metabolites had to be either reported or known to be toxic or reactive. In many cases, chemical reactivity (mostly strong electrophilicity) was demonstrated by the formation of adducts with macromolecules or nucleophiles such as thiols.

Csp2- and Csp-oxidations accounted for 15 percent of reactive metabolites, whereas N- and S-oxidations together yielded a total of 18.3 percent. However, the most remarkable result to emerge from this table is the fact that quinones and analogs accounted for almost 41 percent of all reactive metabolites. This figure should send a strong warning to all medicinal chemists that the possible *para*- or *ortho*-hydroxylation of phenols or arylamines is just a single two-electron step removed from an strong adduct-forming electrophile. Such a reaction of toxification is often catalyzed by cytochromes P450, but the significance of peroxidases such as prostaglandin G/H synthase and myeloperoxidase—which is responsible for the final reaction of toxification of benzene to benzoquinone—is often neglected. This is all the more worrying since the tissue distribution of peroxidases differs markedly from that of CYPs, not to mention that these enzymes are far from receiving the attention they deserve.

Conjugation reactions play a comparatively modest role in toxification reactions, as they account for 10.6 percent of the toxic/reactive metabolites seen in this analysis. A limited number of reactive metabolites are known to be produced by sulfonation of an alcohol group, the resulting sulfate ester being an electrophilic alkylating agent. Some acylglucuronides are known to be reactive, as are a few glutathione conjugates.

The results confirm the primary role of CYP-catalyzed oxidations and UGT-catalyzed glucuronidations in xenobiotic metabolism, but they also document the marked significance of several other reactions and enzymes. The formation of quinones and analogs emerges as a major cause of potential toxicological problems.

The implications drawn from these results include a need for a number of drug discovery scientists to have a better grasp of the variety of drug metabolism reactions and enzymes and their consequences. This includes an understanding of when and how reactive metabolites are formed and when and how to best identify them.

TABLE 24.8 Distribution of Metabolites According to Reaction Types [33][a]

	Percent metabolites[b]	Toxic/reactive metabolites[c]
REDOX REACTIONS		
Oxidation of Csp^3	**20.6%**	0.44%
Oxidation of Csp^2 and Csp	**13.9%**	**1.05%**
-CHOH ↔ >C = O or >C = O → -COOH	**8.3%**	0.18%
Oxidation of R_3N	1.4%	0.03%
Reduction of N-oxides	0.1%	0
Oxidation of >NH or >NOH	1.4%	**0.58%**
Reduction of $-NO_2$ or −N = O or >NOH	2.6%	0.09%
Oxidation to quinones or analogs	3.4%	**2.85%**
Reduction of quinones or analogs	0.6%	0
Oxidation of S atoms	2.6%	**0.67%**
Reduction of S atoms	0.2%	0
Other redox reactions	1.7%	0.06%
HYDROLYTIC REACTIONS		
Hydrolysis of esters, lactones, or inorganic esters	3.8%	0.10%
Hydrolysis of amides, lactams, or peptides	1.7%	0.01%
Hydration of epoxides	1.6%	0.03%
Other reactions of hydration or reactions of dehydration	2.8%	0.16%
CONJUGATION REACTIONS		
O-Glucuronidations or glycosylations	**12.7%**	0.10%
N-, S-, or C-Glucuronidations	1.4%	0
Sulfonations	4.8%	0.27%
Conjugations with glutathione and sequels (incl. reductions)	**8.0%**	0.19%
N- and O-Acetylations	1.5%	0.10%
CoAs-Ligation + aminoacyl conjugations	1.1%	0
CoAS-Ligation + other sequels (formation of hybrid lipids, β-oxidations, 2C-elongations, etc.)	0.7%	0.04%
O-, N- and S-Methylations	2.2%	0.03%
Other conjugations (phosphorylations, carbonyl additions, . . .)	0.9%	0.01%
	100.0%	**6.99%**

[a] *A total of 1171 different substrates were included, which yielded 6,767 metabolites. Among these, 473 metabolites (i.e., 6.99% of all metabolites) were reactive and/or toxic.*
[b] *Values > 5% are bold.*
[c] *Values > 0.50% are bold.*

VI. CONCLUDING REMARKS

Our objective in writing this chapter was to present structured data, namely, a reasoned classification of metabolic reactions and their enzymes. In this way, the vast diversity of metabolic reactions and of xenobiotic-metabolizing enzymes ceases to be a vague notion and can begin to be grasped as a structured and integrated whole made of many interacting parts.

The implicit objective of the chapter is to warn medicinal chemists against the danger of over-simplification. For too many drug discoverers, biotransformation begins and ends with CYP-catalyzed oxidations. As a result

of this narrow view, little is done computationally and experimentally before the development phases to obtain a comprehensive view of the biotransformation of lead and preclinical candidates. One can only wonder about the proportion of metabolism-related side effects that could have been avoided during clinical phases had medicinal chemists been more attentive to potential toxification reactions and drug–drug interactions caused by non-CYP reactions.

References

[1] Testa B, Krämer SD. The biochemistry of drug metabolism. Vol. 1. Principles, redox reactions, hydrolyses. Zurich: Verlag Helvetica Chimica Acta; 2008.
[2] Testa B, Krämer SD. The biochemistry of drug metabolism. Vol. 2. Conjugations, consequences of metabolism, influencing factors. Zurich: Verlag Helvetica Chimica Acta; 2010.
[3] Utrecht JP, Trager WF. Drug metabolism: chemical and enzymatic aspects. New York: Informa; 2007.
[4] Phillips IR, Shephard EA, editors. Cytochrome P450 protocols. Humana Press; 2006.
[5] Lash LH, editor. Drug metabolism and transport. Humana Press; 2005.
[6] Boullata J, Armenti VT, editors. Handbook of drug-nutrient interactions. Totowa, NJ: Humana Press; 2004.
[7] Testa B, Mayer JM. Hydrolysis in drug and prodrug metabolism: chemistry, biochemistry, and enzymology. Wiley-VHCA: Zurich; 2003.
[8] Rodrigues AD, editor. Drug–drug interactions. New York: Dekker; 2002.
[9] Ioannides C, editor. Enzyme systems that metabolise drugs and other xenobiotics. Chichester: Wiley; 2002.
[10] Woolf TF, editor. Handbook of drug metabolism. New York: Dekker; 1999.
[11] Ortiz de Montellano PR, editor. Cytochrome P450: structure, mechanism, and biochemistry. 2nd ed. New York: Plenum Press; 1996.
[12] Testa B. The metabolism of drugs and other xenobiotics: biochemistry of redox reactions. London: Academic Press; 1995.
[13] Anders MW, Dekant W, editors. Conjugation-dependent carcinogenicity and toxicity of foreign compounds. San Diego: Academic Press; 1994.
[14] Welling PG, Balant LP, editors. Pharmacokinetics of drugs. Heidelberg: Springer; 1994.
[15] Kauffman FC, editor. Conjugation-deconjugation reactions in drug metabolism and toxicity. Berlin: Springer; 1994.
[16] Silverman RB. The organic chemistry of drug design and drug action. 2nd ed. San Diego: Academic Press; 2004.
[17] Kalgutkar AS, Dalvie D, Obach RS, Smith DA. Reactive drug metabolites. Weinheim: Wiley-VCH; 2012.
[18] Anzenbacher P, Zanger UM, editors. Metabolism of drugs and other xenobiotics. Weinheim: Wiley-VCH; 2012.
[19] Trager WF. Principles of drug metabolism 1: redox reactions. In: 2nd ed. Testa B, van de Waterbeemd H, editors. ADME-tox approaches. Comprehensive medicinal chemistry, vol. 5. Oxford: Elsevier; 2007. p. 87–132.
[20] Testa B. Principles of drug metabolism 2: hydrolysis and conjugation reactions. In: 2nd ed. Testa B, van de Waterbeemd H, editors. ADME-tox approaches. Comprehensive medicinal chemistry, vol. 5. Oxford: Elsevier; 2007. p. 133–66.
[21] Testa B. Principles of drug metabolism. In: Abraham DJ, Rotella DP, editors. Burger's medicinal chemistry. vol. 1. Drug discovery and development. 7th ed. Hoboken, NJ: John Wiley & Sons; 2010. p. 127–66.
[22] Testa B, Krämer SD. The biochemistry of drug metabolism: an introduction. Part 1: principles and overview. Chem Biodiver 2006;3:1053–101.
[23] Testa B, Krämer SD. The biochemistry of drug metabolism: an introduction. Part 2: redox reactions and their enzymes. Chem Biodiver 2007;4:257–405.
[24] Testa B, Krämer SD. The biochemistry of drug metabolism: an introduction. Part 3: reactions of hydrolysis and their enzymes. Chem Biodiver 2007;4:2031–122.
[25] Testa B, Krämer SD. The biochemistry of drug metabolism: an introduction. Part 4: reactions of conjugation and their enzymes. Chem Biodiver 2008;5:2171–336.
[26] Testa B, Krämer SD. The biochemistry of drug metabolism: an introduction. Part 5: metabolism and bioactivity. Chem Biodiver 2009;6:591–684.
[27] Krämer SD, Testa B. The biochemistry of drug metabolism: an introduction. Part 6: inter-individual factors affecting drug metabolism. Chem Biodiver 2008;5:2465–578.
[28] Krämer SD, Testa B. The biochemistry of drug metabolism: an introduction. Part 7: intra-individual factors affecting drug metabolism. Chem Biodiver 2009;6:1477–660.
[29] Kadlubar S, Kadlubar FF. Enzymatic basis of phase I and phase II drug metabolism. In: Pang KS, Rodrigues AD, Peter RM, editors. Enzyme- and transporter-based drug–drug interactions. New York: Springer; 2010. p. 3–25.
[30] Johnson CH, Patterson AD, Idle JR, Gonzalez FJ. Xenobiotic metabolomics: major impact on the metabolome. Annu Rev Pharmacol Toxicol 2012;52:37–56.
[31] Guengerich FP, Rendic S. Update information on drug metabolism systems 2009. Part I. Curr Drug Metab 2010;11:1–3.
[32] Rendic S, Guengerich FP. Update information on drug metabolism systems 2009. Part II: summary of information on the effects of diseases and environmental factors on human cytochrome P450 (CYP) enzymes and transporters. Curr Drug Metab 2010;11:4–84.
[33] Testa B, Pedretti A, Vistoli G. Reactions and enzymes in the metabolism of drugs and other xenobiotics: a meta-analysis of the primary literature. Drug Discovery Today 2012;17:549–60.
[34] Ettmayer P, Amidon GL, Clement B, Testa B. Lessons learned from marketed and investigational prodrugs. J Med Chem 2004;47:2393–404.
[35] Walle T, Walle UK, Olanoff LS. Quantitative account of propranolol metabolism in urine of normal man. Drug Metab Dispos 1985;13:204–9.
[36] Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). Enzyme Nomenclature, Available from: <http://www.chem.qmul.ac.uk/iubmb/enzyme/>; 2014 [accessed 14.01].

[37] Nelson DR, Zeldin DC, Hoffman SMG, Maltais LJ, Wain HM, Nebert DW. Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes, and alternative splice variants. Pharmacogenetics 2004;14:1–18.
[38] Directory of P450–containing systems. Available from: <http://www.icgeb.org/~p450srv/>; 2009 [accessed 14.02].
[39] Brenda: the comprehensive enzyme information system. Available from: <http://www.brenda-enzymes.info/>; 2014 [accessed 14.01].
[40] Palfey BA, McDonald CA. Control of catalysis in flavin-dependent monooxygenases. Arch Biochem Biophys 2010;493:26–36.
[41] Totah RA, Rettie AE. Principles of drug metabolism 3: enzymes and tissues. In: 2nd ed. Testa B, van de Waterbeemd H, editors. ADME-tox approaches. Comprehensive medicinal chemistry, vol. 5. Oxford: Elsevier; 2007. p. 167–91.
[42] Brown CM, Reisfeld B, Mayeno AN. Cytochromes P450: a structure-based summary of biotransformations using representative substrates. Drug Metab Rev 2008;40:1–100.
[43] Sbaragli L, Woggon WD. On the mechanism of cytochrome P450-catalyzed oxidations: reaction of a new enzyme model with a radical clock. Synthesis 2005;1538–42.
[44] Kirchmair J, Williamson MJ, Tyzack JD, Tan L, Bond PJ, Bender A, et al. Computational prediction of metabolism: sites, products, SAR, P450 enzyme dynamics, and mechanisms. J Chem Inf Model 2012;52:617–48.
[45] Skarydová S, Wsól V. Human microsomal carbonyl reducing enzymes in the metabolism of xenobiotics: well-known and promising members of the SDR superfamily. Drug Metab Rev 2012;44:1173–91.
[46] Malátková P, Wsól V. Carbonyl reduction pathways in drug metabolism. Drug Metab Rev 2014;46:96–123.
[47] Guengerich FP, Sohl CD, Chowdhury G. Multi-step oxidations catalyzed by cytochrome P 450 enzymes: processive vs. distributive kinetics and the issue of carbonyl oxidation in chemical mechanisms. Arch Biochem Biophys 2011;507:126–34.
[48] Rendic S, Guengerich FP. Contributions of human enzymes in carcinogen metabolism. Chem Res Toxicol 2012;25:1316–83.
[49] Rambeck B, May T, Juergens U. Serum concentrations of carbamazepine and its epoxide and diol metabolites in epileptic patients: the influence of dose and comedication. Ther Drug Monit 1987;9:298–303.
[50] Subramanian R, Tam J, Aidasani D, Reid D, Skiles GL. Novel cytochrome P450 bioactivation of a terminal phenyl acetylene group: formation of a one-carbon loss benzaldehyde and other oxidative products in the presence of N-acetylcysteine or glutathione. Chem Res Toxicol 2011;24:677–86.
[51] Havemeyer A, Lang J, Clement B. The fourth mammalian molybdenum enzyme mARC: current state of research. Drug Metab Rev 2011;43:524–39.
[52] Jindal SP, Lutz T. Mass spectrometric studies of cocaine disposition in animals and humans using stable isotope-labeled analogues. J Pharm Sci 1989;78:1009–14.
[53] Brzezinski MR, Spink BJ, Dean RA, Berkman CE, Cashman JR, Bosron WF. Human liver carboxylesterase hCE-1: binding specificity for cocaine, heroin, and their metabolites and analogs. Drug Metab Dispos 1997;25:1089–96.
[54] Reist M, Carrupt PA, Francotte E, Testa B. Chiral inversion and hydrolysis of thalidomide: mechanisms and catalysis by bases and serum albumin, and chiral stability of teratogenic metabolites. Chem Res Toxicol 1998;11:1521–8.
[55] Lu J, Helsby N, Palmer BD, Tingle M, Baguley BC, Kestell P, et al. Metabolism of thalidomide in liver microsomes of mice, rabbits, and humans. J Pharmacol Exp Ther 2004;310:571–7.
[56] Weinshilboum RM, Otterness DM, Szumlanski CL. Methylation pharmacogenetics: catechol O-methyltransferase, thiopurine methyltransferase, and histamine N-methyltransferase. Annu Rev Pharmacol Toxicol 1999;39:19–52.
[57] Ginsberg G, Hattis D, Russ A, Sonawane B. Physiologically-based pharmacokinetic (PBPK) modeling of caffeine and theophylline in neonates and adults: implications for assessing children's risks from environmental agents. J Toxicol Environ Health A 2004;67:297–329.
[58] Nagata K, Yamazoe Y. Pharmacogenetics of sulfotransferase. Annu Rev Pharmacol Tox 2000;40:159–76.
[59] Banoglu E. Current status of the cytosolic sulfotransferases in the metabolic activation of promutagens and procarcinogens. Curr Drug Metab 2000;1:1–30.
[60] Kauffman FC. Sulfonation in pharmacology and toxicology. Drug Metab Rev 2004;36:823–43.
[61] Leyh TS, Cook I, Wang T. Structure, dynamics, and selectivity in the sulfotransferase family. Drug Metab Rev 2013;45:423–30.
[62] James OM, Ambadapadi S. Interactions of cytosolic sulfotransferases with xenobiotics. Drug Metab Rev 2013;45:401–14.
[63] Tukey RH, Strassburg CP. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. Annu Rev Pharmacol Toxicol 2000;40:581–616.
[64] Fisher MB, Paine MF, Strelevitz TJ, Wrighton SA. The role of hepatic and extrahepatic UDP-glucuronosyltransferases in human drug metabolism. Drug Metab Rev 2001;33:273–97.
[65] Bock KW. Vertebrate UDP-glucuronosyltransferases: functional and evolutionary aspects. Biochem Pharmacol 2003;66:691–6.
[66] Maruo Y, Iwai M, Mori A, Sato H, Takeuchi Y. Polymorphism of UDP-glucuronosyltransferase and drug metabolism. Curr Drug Metab 2005;6:91–9.
[67] Milne RW, Nation RL, Somogyi AA. The disposition of morphine and its 3- and 6-glucuronide metabolites in humans and animals, and the importance of the metabolites to the pharmacological effects of morphine. Drug Metab Rev 1996;28:345–472.
[68] Spahn-Langguth H, Benet LZ. Acyl glucuronides revisited: is the glucuronidation process a toxification as well as a detoxification mechanism? Drug Metab Rev 1992;24:5–48.
[69] Grant DM, Blum M, Meyer UA. Polymorphisms of N-acetyltransferase genes. Xenobiotica 1992;22:1073–81.
[70] Agundez JAG, Olivera M, Martinez C, Ladero JM, Benitez J. Identification and prevalence study of 17 allelic variants of the human NAT2 gene in a white population. Pharmacogenetics 1996;6:423–8.
[71] Kawamura A, Graham J, Mushtaq A, Tsiftsoglou SA, Vath GM, Hanna PE, et al. Eukaryotic arylamine N-acetyltransferase: investigation of substrate specificity by high-throughput screening. Biochem Pharmacol 2005;69:347–59.
[72] Knights KM, Drogemuller CJ. Xenobiotic-CoA ligases: kinetics and molecular characterization. Curr Drug Metab 2000;1:49–66.
[73] Kasuya F, Igarashi K, Fukui M. Participation of a medium chain acyl-CoA synthetase in glycine conjugation of the benzoic acid derivatives with the electron-donating groups. Biochem Pharmacol 1996;51:805–9.

[74] Brugger R, Reichel C, Garcia Alia B, Brune K, Yamamoto T, Tegeder I, et al. Expression of rat liver long-chain acyl-CoA synthetase and characterization of its role in the metabolism of R-ibuprofen and other fatty acid-like xenobiotics. Biochem Pharmacol 2001;61:651–6.
[75] Mayer JM, Testa B. Pharmacodynamics, pharmacokinetics, and toxicity of ibuprofen enantiomers. Drugs of the Future 1997;22:1347–66.
[76] Gopaul VS, Tang W, Farrell K, Abbott FS. Amino acid conjugates: metabolites of 2-propylpentanoic acid (valproic acid) in epileptic patients. Drug Metab Dispos 2003;31:114–21.
[77] Sies H. Glutathione and its role in cellular functions. Free Rad Biol Med 1999;27:916–21.
[78] Dickinson DA, Forman HJ. Cellular glutathione and thiols metabolism. Biochem Pharmacol 2002;64:1019–26.
[79] Pompella A, Visvikis A, Paolicchi A, De Tala V, Casini AF. The changing faces of glutathione, a cellular protagonist. Biochem Pharmacol 2003;66:1499–503.
[80] Hayes JD, Pullford DJ. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit Rev Biochem Molec Biol 1995;30:445–600.
[81] Armstrong RN. Structure, catalytic mechanism, and evolution of the glutathione transferases. Chem Res Toxicol 1997;10:2–18.
[82] Sheenan D, Meade G, Foley VM, Dowd C. Structure, function, and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. Biochem J 2001;360:1–16.
[83] Jowsey IR, Thomson AM, Flanagan JU, Murdock PR, Moore GBT, Meyer DJ, et al. Mammalian class sigma glutathione S-transferases: catalytic properties and tissue-specific expression of human and rat GSH-dependent prostaglandin D2 synthases. Biochem J 2001;359:507–16.
[84] Schnekenburger M, Morceau F, Duvoix A, Delhalle S, Trentesaux C, Dicato M, et al. Increased glutathione S-transferase P1-1 expression by mRNA stabilization in hemin-induced differentiation of K562 cells. Biochem Pharmacol 2004;68:1269–77.
[85] Ketterer B. The role of nonenzymatic reactions of glutathione in xenobiotic metabolism. Drug Metab Rev 1982;13:161–87.
[86] Monks TJ, Lau SS. Biological reactivity of polyphenolic-glutathione conjugates. Chem Res Toxicol 1997;10:1296–313.
[87] Anders MW. Glutathione-dependent bioactivation of haloalkanes and haloalkenes. Drug Metabol Rev 2004;36:583–94.

CHAPTER

25

Biotransformations Leading to Toxic Metabolites: Chemical Aspects

Anne-Christine Macherey[1] *and Patrick M. Dansette*[2]

[1]Prévention du Risque Chimique, Gif sur Yvette, France;
[2]Laboratoire de Chimie et Biochimie Pharmacologique et Toxicologique, Université Paris Descartes, Sorbonne Paris Cité, Paris

OUTLINE

La matière demeure et la forme se perd.
Matter remains and Form is lost. ***Pierre De Ronsard***

I. HISTORICAL BACKGROUND

As drugs are usually foreign chemicals, the history of concern for the biotransformations of drugs leading to toxic metabolites formation is intrinsically linked to the history of xenobiotic metabolism studies. The International Society for the Study of Xenobiotics (ISSX) website (http://www.issx.org/) presents an overview of the field history where some key figures may be pointed out.

One is probably Richard Tecwyn Williams, who introduced the Phase I and II classification of xenobiotics metabolism reactions. Although his emblematic book [1] was called "Detoxication mechanisms," he estimated that in some cases metabolism may increase toxicity. He also considered that this "bioactivation" may occur during the Phase II reactions (usually considered as detoxification reactions) and not only that of Phase I (functionalization reactions).

Quite at the same time, Bernard Brodie studied the antimalarial atabrine (quinacrine) metabolism in order to avoid the toxic side effects of the drug. He also developed some new analytical methods necessary for metabolic studies. Then he set up a group of researchers—including Julius Axelrod, James Gillette, and many others—in this field, and they published many studies of great significance related to drug metabolism, probably the most famous ones concerning acetaminophen (paracetamol) [2,3]. During this period, these scientists developed the concept of covalent binding, which provides an explanation for the toxic side effects of numerous drugs. Following the work of James and Elizabeth Miller in the 1940s on the covalent binding to DNA of electrophilic metabolites of polycyclic aromatic hydrocarbons [4], Brodie and co-workers suggested that *in vivo* bioactivation may lead to the formation of electrophilic entities, capable of linking with biological macromolecules, thus inducing disturbances in cellular functions [2].

The discovery of mixed function oxidases during the 1950s and the characterization of cytochrome P-450 by Omura and Sato [5] were a "revolution" in the field of xenobiotic metabolism. Remmer discovered that cytochrome P450 may be induced by phenobarbital, and Conney characterized the induction with 3-methylcholanthrene [6]. These works [7] were of great importance for mechanistic studies of drug metabolism. A new step occured in 1999 with the first crystallization of a mammalian cytochrome P-450 by Johnson, which provided new perspectives in safer drug design [8].

The induction of cytochrome P450 synthesis suggests that xenobiotics may exert genomic effects promoting genomics and proteomics to a new challenge for predictive toxicology in drug design.

II. INTRODUCTION

Toxicity is the outcome of the more or less harmful action of chemicals on a living organism. Toxicology, the study of toxicity, is situated at the borders of chemistry, biology, and, in some cases, physics. Molecular toxicology tries to elucidate the mechanisms by which chemicals exert their toxic effects. Because many foreign chemicals enter the body in inert but unexcretable forms, biotransformations are an important aspect of the fate of xenobiotics [9,10]. In the case of drugs, metabolic conversions may be required for therapeutic effect ("prodrugs"; see Chapter 28 for a detailed discussion of prodrugs). In other cases, metabolism results in a loss of the biological activity. Biotransformations sometimes produce toxic metabolites, a process called toxification or bioactivation. It should be emphasized that the general principles of pharmacology embrace the occurrence of toxic events: although biotransformation processes are often referred to as detoxification, in a number of cases the metabolic products are more toxic than the parent compounds. For drugs, whether biotransformations lead to the formation of toxic metabolites or to variations in therapeutic effects depends on intrinsic (e.g., the genetic polymorphism of some metabolism pathways) and extrinsic (e.g., dose, route, or duration of administration) factors. The biochemical conversions are usually of an enzymatic nature and yield reactive intermediates that may be implicated in the toxicity of proximal or distal metabolites. The primary events that constitute the beginning of the toxic effect may result after metabolism from an inhibition of a specific (and in most cases enzymatic) cellular function, an alkylating attack, or an oxidative stress.

With regard to the toxicity arising from metabolites ("indirect toxicity"), three cases may be distinguished (Figure 25.1):

- Biotransformation begins with the transient formation of a reactive intermediate whose lifetime is long enough to allow an attack on cellular components. This occurs when a reactive intermediate (generally radicals or electrophiles, such as a carbonium ion) is formed and reacts rapidly with cellular macromolecules (e.g., unsaturated lipids, proteins, nucleic acids), thus leading to their degradation and finally to cellular necrosis.
- The first step of the metabolic process yields a primary metabolite, which can in some cases accumulate in the cell and react with cellular components before being transformed.
- The final metabolites, when in excess, may accumulate and react with cellular macromolecules.

Usually, metabolic conversions are divided into two major types of reactions [11] (see Chapter 24 for a detailed discussion of metabolic biotransformations). Phase I reactions, or functionalization reactions, involve the introduction of a polar functionality such as a hydroxyl group into the xenobiotic structure. During Phase II reactions, this group is subsequently coupled (or conjugated) with an endogenous cofactor that contains a functional group that is usually ionized at physiological pH. This ionic group facilitates active excretion into the urinary and/or hepatobiliary system. The elimination by transport mechanisms is sometimes also called "Phase III."

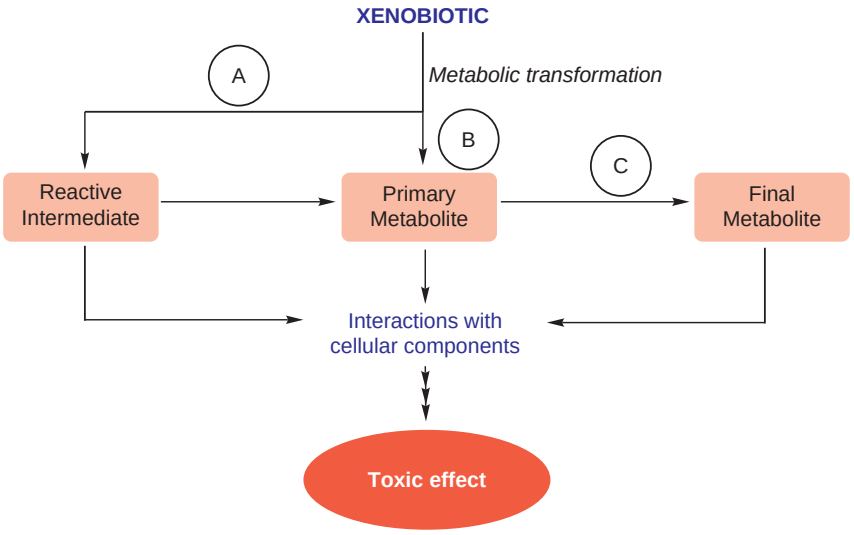


FIGURE 25.1 Indirect toxicity.

Because bioactivation is mainly an activation of xenobiotics to electrophilic forms, which are entities capable of reacting irreversibly with tissue nucleophiles, biotransformations leading to toxic metabolites are in most cases Phase I reactions, but Phase II reactions may also give rise to toxic phenomena (e.g., when conjugation produces a toxic metabolite, or when it is responsible for a specific target organ toxicity by acting as a delivery form to particular sites in the body where it is hydrolyzed and exerts a localized effect). Also, the final toxic metabolite may be formed by combinations of several Phase I and Phase II reactions. Because of the increasing understanding of drug metabolizing enzymes, some authors [6] claim that Williams "Phase I and II" classification is now inaccurate and even misleading. Pointing out the fact that Williams only introduced the classification at the end of his book and did not use it in his monograph, they consider it would be now wiser to avoid using any special category.

III. REACTIONS INVOLVED IN BIOACTIVATION PROCESSES

During biotransformations affecting xenobiotics, five major kinds of chemical reactions may occur: oxidations (by far the most important), reductions, hydrolysis, substitutions, and eliminations. As Phase I and II reactions are part of this classification, each class of reactions can give rise to toxic metabolites.

A. Oxidation

Several enzymatic systems are involved during the oxidative transformations of xenobiotics. Whether substances act upon one enzyme rather than another depends not only on their specific functions but also on the electromolecular environment. The most important is the microsomal drug-metabolizing system known as cytochrome P-450 monooxygenase, which is localized mainly in the liver and is involved in most biological oxidations of xenobiotics [12–14]. These include *C*-, *N*-, and *S*-oxidations, *N*-, *O*-, and *S*-dealkylation, deaminations, and certain dehalogenations. Under anaerobic conditions, it can also catalyze reductive reactions. The cytochrome P-450 monooxygenase system is a multienzymatic complex constituted by the cytochrome P-450 hemoprotein, the flavoprotein enzyme NADPH cytochrome P450 reductase, and the unsaturated phospholipid phosphatidylcholine. The isoforms involved in xenobiotic metabolism are membrane bound enzymes situated in the endoplasmic reticulum. After cell lysis for *in vitro* studies, they are found in the microsomal fraction. There are numerous isoforms (more than 20,000 known in all species as of July 2013). Thus, a nomenclature based on their sequence similarity has been designed, and they are classified into families and subfamilies. For instance, CYP3A4 is the major human cytochrome P-450, where CYP is for cytochrome P-450, 3 for the number of the family (more than 40 percent sequence identity with at least one member), A for the letter of the subfamily (more than 55 percent sequence identity), and 4 the number in the subfamily. The human genome shows fifty-seven complete CYP sequences plus a number of pseudogenes. The CYPs involved in xenobiotic metabolism [14] (about fifteen) belong to families 1–4. The catalytic mechanism of cytochrome P-450 involves a formal $(FeO)^{3+}$ complex formed by the elimination of H_2O from the iron site after two electrons have been added (Figure 25.2).

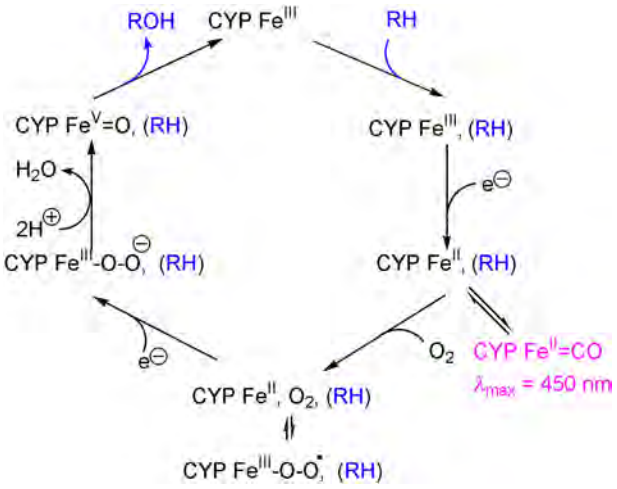


FIGURE 25.2 Catalytic cycle of cytochrome P-450 monooxygenase.

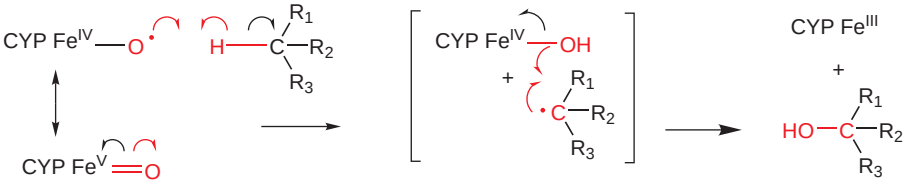


FIGURE 25.3 Cytochrome P-450 oxidation process.

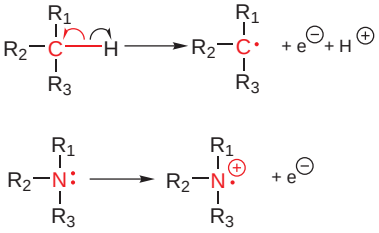


FIGURE 25.4 One-electron oxidation.

Another oxidative enzyme is the FAD-containing monooxygenase, which is capable of oxidizing nucleophilic nitrogen, sulfur, and organophosphorus compounds. The flavoprotein binds NADPH, oxygen, and then the substrate. The oxidized metabolite is released, followed by NADP. Alcohol dehydrogenase and aldehyde dehydrogenase catalyze the oxidation of a variety of alcohols and aldehydes into aldehydes and acids, respectively, in the liver. Xanthine oxidase oxidizes several purine derivatives such as theophylline. Monoamine oxidase (MAO) and diamine oxidase convert amines into alkyl or aryl aldehydes by oxidation of the amine to an imine followed by subsequent hydrolysis. Peroxidases are oxidative enzymes, which couple the reduction of hydrogen peroxide and lipid hydroperoxides to the oxidation of other substrates. This co-oxidation is responsible for the production of reactive electrophiles from aromatic amines (e.g., the highly carcinogenic benzidine), phenols, hydroquinones, and polycyclic aromatic hydrocarbons.

The oxidation reactions can be described in terms of a rather common chemistry that involves the abstraction of either a hydrogen atom or a nonbonded (or π) electron by the iron–oxo porphyrin complex (Figure 25.3). The high-valent complex electronic configuration is unknown but is usually written as $Fe^{V} = 0$.

The one-electron oxidation yields transient radicals (Figure 25.4) that are transformed into more stable forms.

These radicals can incorporate an oxygen atom by abstraction of a hydroxyl group from the cytochrome P-450 iron–oxo species. This yields an oxidized derivative that may sometimes be more toxic than the parent compound or susceptible to further metabolic conversions. Free radicals may also bind to the site of their formation, thus leading to inhibition or inactivation of the enzyme. When the radical is not efficiently controlled by the

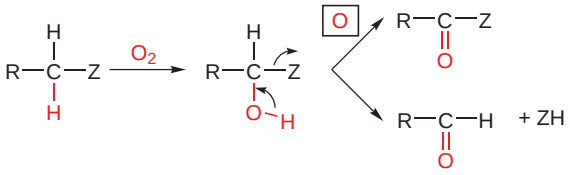


FIGURE 25.5 C–H bond oxidation in the α-position to a heteroatom.

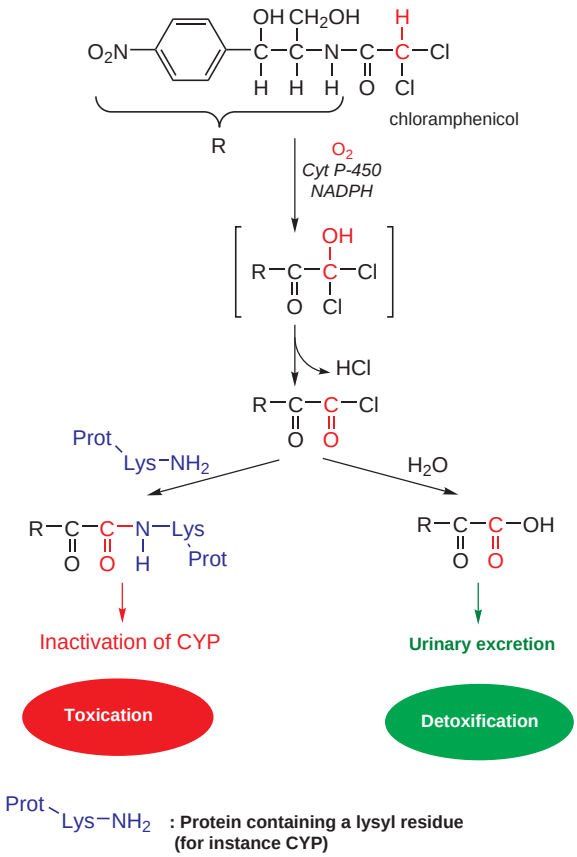


FIGURE 25.6 Metabolic activation of chloramphenicol.

iron, it may leave the active site. The subsequently released radical is able to produce damage to unsaturated fatty acids, thus leading to lipid peroxidation and destruction of the cellular structure. Another mode of the radical stabilization is a second one-electron oxidation, which consists of the loss of another electron. The fate of free radicals is now extensively studied because of their great capacities for forming covalent bonds with cellular macromolecules [15–17].

1. C–H Bond Oxidations

These oxidations, which are usually catalyzed by cytochrome P-450 monooxygenases, produce hydroxylated derivatives [18]. When the C–H bond is located in the α position to a heteroatom (e.g., O, S, N, halogen), the α-hydroxylated derivative obtained is usually unstable and may be further oxidized or cleaved (Figure 25.5).

The antibiotic chloramphenicol is metabolized by cytochrome P-450 monooxygenase to chloramphenicol oxamyl chloride formed by the oxidation of the dichloromethyl moiety of chloramphenicol followed by elimination of hydrochloric acid (Figure 25.6) [19]. The reactive metabolite reacts with the ε-amino group of a lysine residue in cytochrome P-450 [20] and progressively inhibits the enzymatic reaction with time. This type of inhibition is a time-dependent inhibition or a mechanism-based inhibition or inactivation, and the substrate involved has historically been called a "suicide substrate," because the enzymatic reaction yields a reactive metabolite that destroys the enzyme [21].

In the case of chloroform, the unstable trichloromethanol loses hydrochloric acid and forms phosgene, which is very reactive (Figure 25.7) [22].

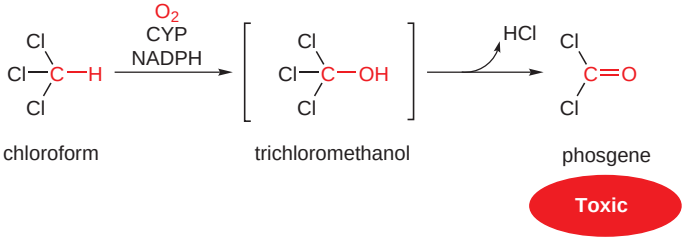


FIGURE 25.7 Oxidation of chloroform.

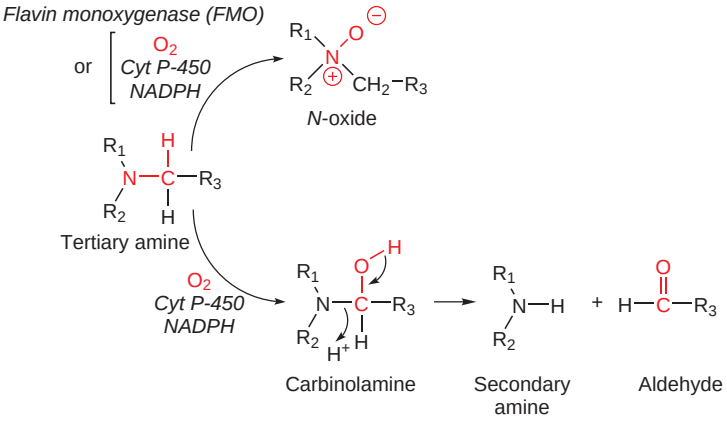


FIGURE 25.8 Oxidation of a tertiary amine.

Tertiary amines containing at least one hydrogen on the α carbon may either be *N*-oxidized (leading to an *N*-oxide in the case of tertiary amines) or *C*-oxidized (leading to a carbinolamine). The latter—usually being unstable—splits into a secondary amine and an aldehyde moiety (Figure 25.8). Several electron transfer mechanisms have been proposed [12–14].

During the oxidation of nitrosamines, the hydroxylated derivative formed cleaves spontaneously into highly reactive metabolites capable of alkylating nucleophilic sites in the cellular components.

2. *Unsaturated Bond Oxidations*

Double bonds are oxidized by cytochrome P-450 monooxygenases into epoxides, which are generally very reactive. Epoxides are considered responsible for the toxicity of unsaturated compounds.

The hepatocarcinogenicity of aflatoxin B_1 (AFB_1) is known to be due to the epoxide (AFB_1-oxides) formed, which binds directly with the N-7 atom of a guanine molecule in DNA (Figure 25.9) [23].

Aromatic chemicals are metabolized into unstable arene-oxides, which—as epoxides—are comparable to potentially equivalent electrophilic carbocations. These metabolites react easily with thiol groups derived from proteins, leading, for example, to hepatotoxicity. Bromobenzene seems to target a large group of functionally diverse hepatic proteins, as demonstrated recently in a proteomic analysis [24]. The chemical is oxidized (Figure 25.10) into a 3,4-epoxide, which does not exhibit mutagenic or carcinogenic activity but reacts non-enzymatically with liver proteins and produces hepatic necrosis [25]. The isomeric 2,3-epoxide rearranges very quickly to 2-bromophenol and is less toxic. A secondary P450-catalyzed oxidation to hydroquinone and benzoquinone can also occur. In this alternative pathway, conjugation with glutathione can lead to the formation of products that may elicit their toxicity in sites other than the liver, especially in the kidneys [25].

3. N-*oxidations*

Tertiary amines are transformed into *N*-oxides (generally less toxic), but primary and secondary amines are oxidized into hydroxylated derivatives (hydroxylamines). This oxidation is responsible for the hepatotoxicity and mutagenicity of acetamino-2-fluorene (Figure 25.11) [26].

Nitrenium ions may occur during bioactivation of aromatic amines and amides, which are usually *N*-oxidized into *N*-hydroxylated derivatives. By sulfation or esterification followed by elimination of the newly formed leaving group, the latter may be transformed into highly reactive nitrenium ions. In the case of aromatic nitrenium ions, they are in equilibrium with their tautomeric aromatic carbocations, which react with cellular nucleophilic macromolecules (e.g., nucleic acids).

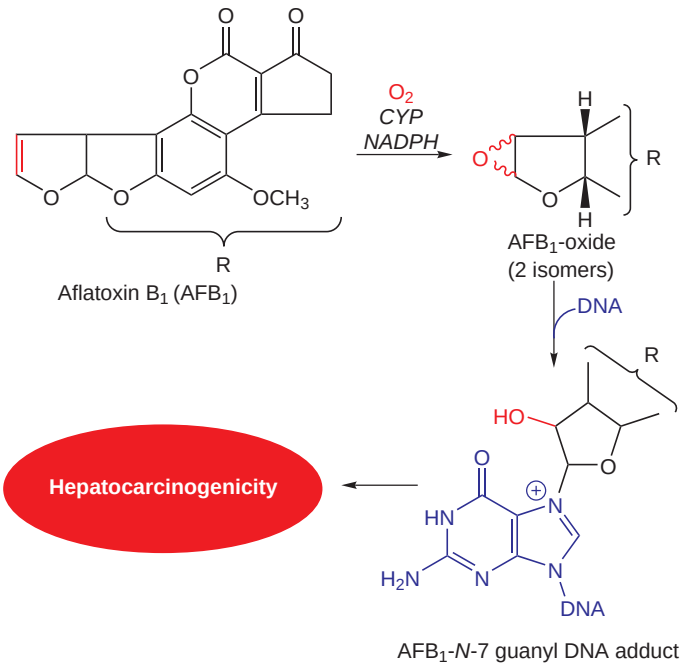


FIGURE 25.9 Oxidation of aflatoxin B_1.

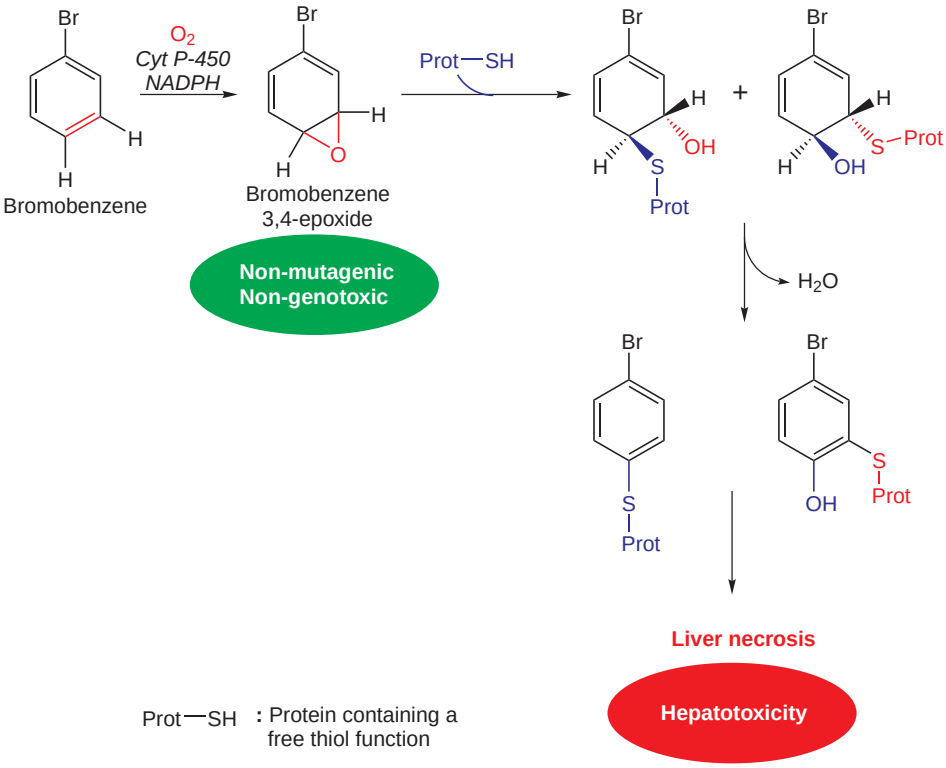


FIGURE 25.10 Metabolism of bromobenzene.

4. *Heteroatom Oxidations*

Heteroatoms such as nitrogen or sulfur are oxidized at their nonbonded peripheric electrons, as described for thiophene (Figure 25.12) [27]. Thiophene is oxidized to thiophene sulfoxide, which is unstable and dimerizes spontanously to thiophene *S*-oxide dimers through a Diels–Alder reaction [28,29]. They also react with nucleophiles like the thiol group of glutathione or proteins, giving glutathione or protein adducts. In addition, thiophenes are oxidized to unstable thiophene epoxides, which rearrange spontanously to thiolenones as found recently for 2- and

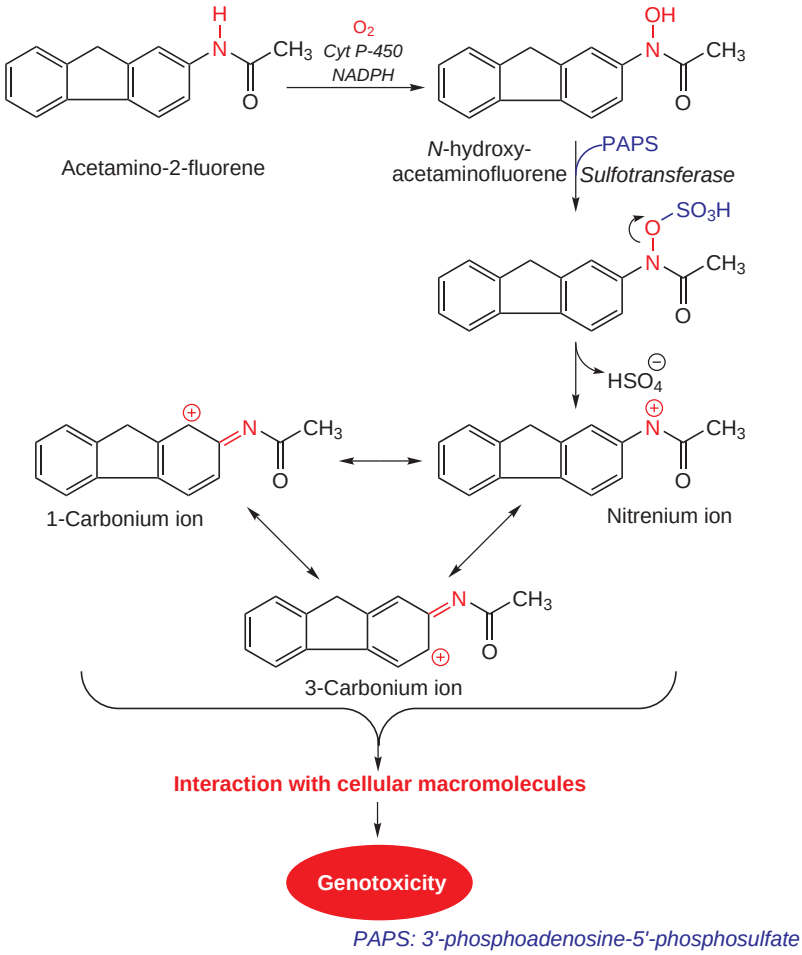


FIGURE 25.11 *N*-oxidation of acetamino-2-fluorene.

3-phenylthiophenes (Figure 25.12a). In fact, there is a competition between *S*-oxidation (sulfoxide pathway) and double-bond oxidation (epoxide pathway). In the presence of glutathione, adducts formed from both reactive intermediates have been found, in addition to thiophene *S*-oxide dimers and the thiolenones tautomers of hydroxythiophenes [30,31].

Sulfur oxidation of thioester or thiocarbamates lead to very reactive acyl-sulfoxides, which can be attacked by water or nucleophiles to give reactive sulfenic acids (Figure 25.13a) [32,33]. Moreover, cyclic thio-acyl compounds can be oxidized to acyl-sulfoxide, which are bis-alkylating compounds (Figure 25.13b) [33–35].

Halogenated aromatic compounds may also be oxidized by cytochrome P-450 monooxygenases, yielding hypervalent halogenated compounds.

B. Oxidative Stress

Oxidative stress has been defined as a disturbance in the pro-oxidant–antioxidant balance in favor of the pro-oxidant state resulting from alterations in the redox state of the cell. The stepwise reduction of oxygen into superoxide anion, hydrogen peroxide, hydroxyl radical, and finally water, which accounts for about 5 percent of the normal oxygen reduction (versus 95 percent by means of the mitochondrial electron-transport chain), may be increased by the redox cycling of some xenobiotics, such as quinones or nitro-aromatic derivatives. These compounds are susceptible to one-electron reduction, which yields radical structures that may be back-oxidized to the parent compound. During this reoxidation, oxygen is reduced into superoxide anion. The oxygen reduction products are highly reactive entities that attack all cellular components, especially when their normal degradation systems (superoxide dismutase, glutathione peroxidase, catalase) are overloaded. The polyunsaturated lipids are especially sensitive to these attacks because they are susceptible to a membrane-degrading peroxidation.

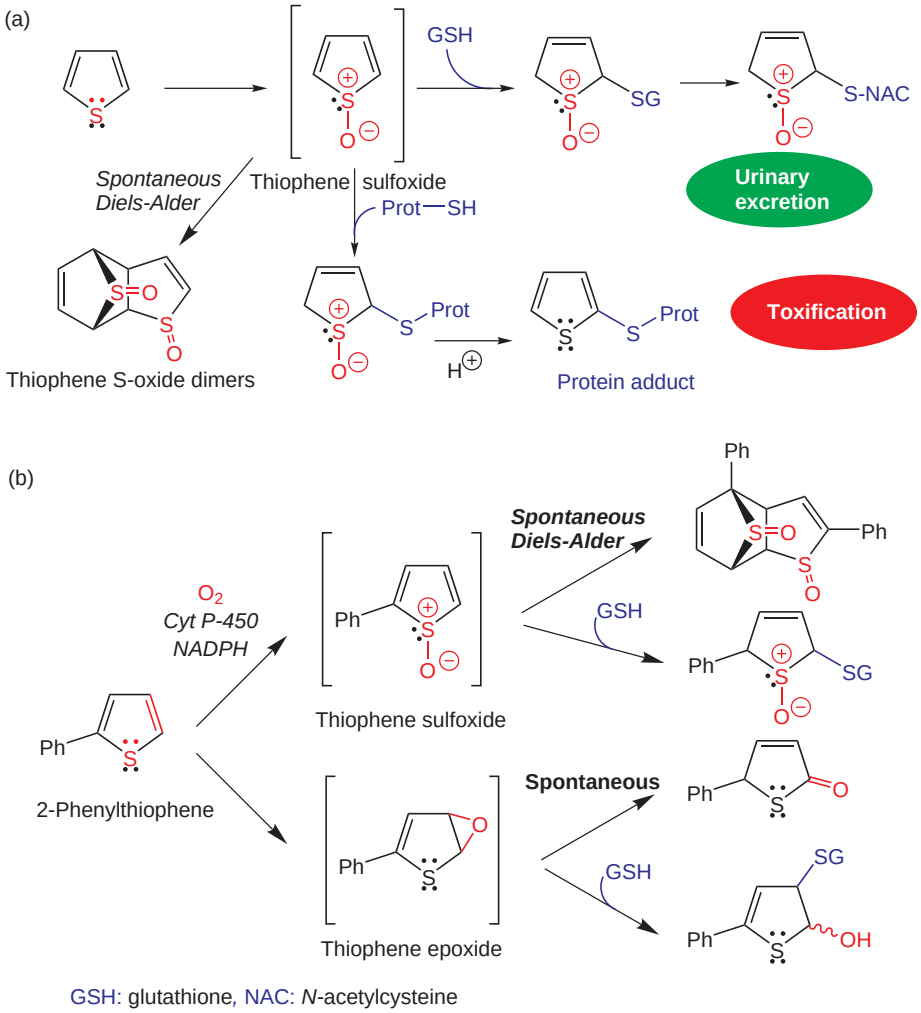


FIGURE 25.12 (a) oxidation of thiophene; (b) oxidation of 2-phenylthiophene.

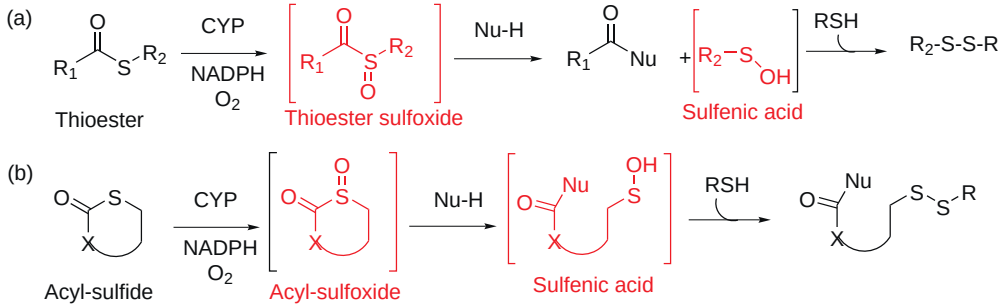


FIGURE 25.13 (a) oxidation of a thioester into a sulfenic acid; (b) oxidation of a cyclic acyl-sulfide into acyl-sulfoxide opened by nucleophilic attack into a reactive sulfenic acid.

C. Reduction

Reductive biotransformations of several compounds such as polyhalogenated, keto, nitro, and azo derivatives are catalyzed by a variety of enzymes that differ according to the substrates and the species. The liver cytochrome P-450-dependent drug-metabolizing system is capable of reducing *N*-oxide, nitro, and azo bonds, whereas the cytosolic nitrobenzene reductase activity is mainly due to cytochrome P-450 reductase, which transforms nitrobenzene into its hydroxylamino derivative. NADPH cytochrome c reductase is also able to catalyze

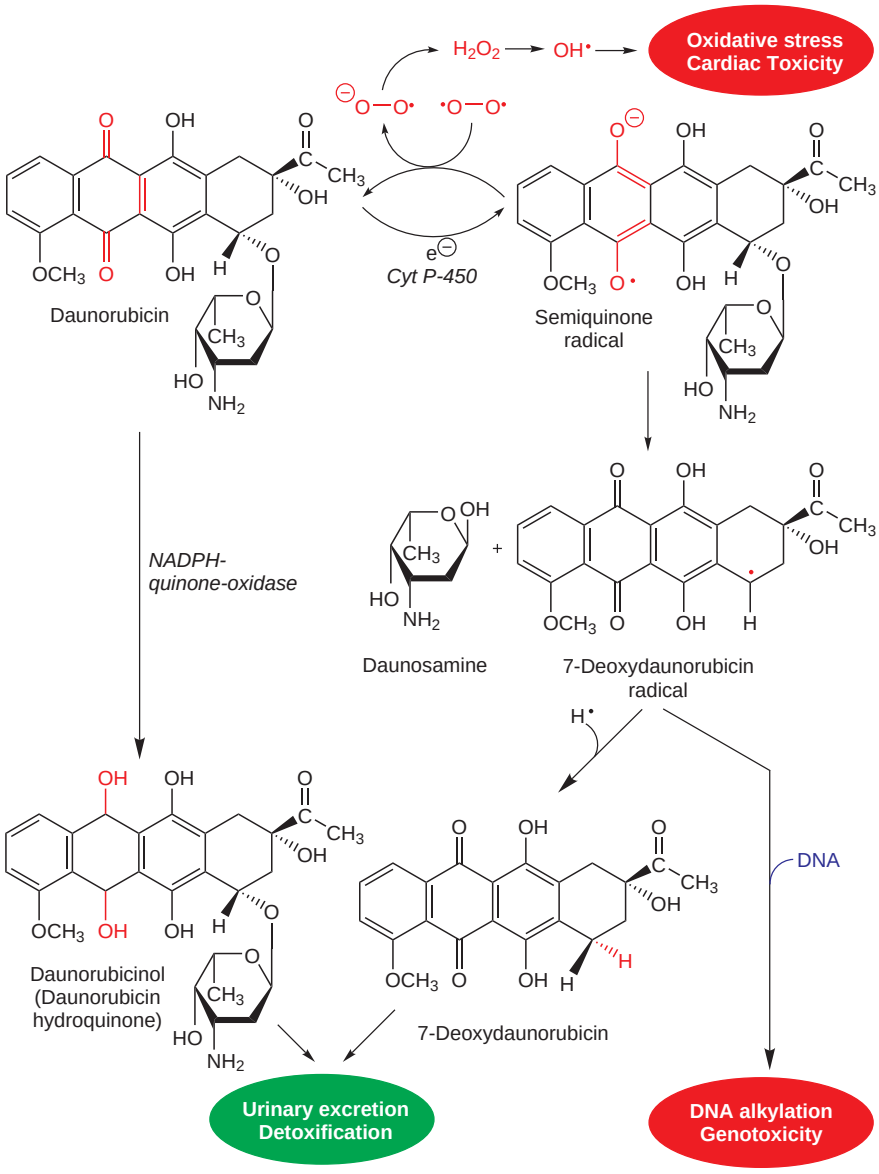


FIGURE 25.14 Biotransformations of daunorubicin.

the reduction of nitro compounds. These metabolic conversions may also be brought about by gastrointestinal anaerobic bacteria. Reductive processes that occur during the metabolism of xenobiotics involve either one-electron reduction or a two-electron transfer.

Ionic reduction using a hydride occurs *in vivo* during the reduction catalyzed by NADH or NADPH enzymes, whereas one-electron reduction releases a radical structure, which may contribute to the toxic effect. Figure 25.14 illustrates the biotransformations affecting the anthracycline antitumor drug daunomycin [36]. Recent studies suggest that nitric oxide synthases may contribute to cardiotoxicity, probably because of their structural similarities with P-450 reductase [37].

1. *Reduction of Polyhalogenated Compounds*

Some polyhalogenated compounds, such as CCl_4, $BrCCl_3$, and halothane (CF_3-CHBrCl), when in the presence of the reduced form of cytochrome P-450, may undergo one-electron reduction [38] (Figure 25.15), leading to a radical that may be transformed by different pathways.

The radical formed may add directly on the unsaturated lipid bonds, initiate lipid peroxidation, or undergo another one-electron reduction. The last reaction yields a carbene that can form a complex with the iron of the reduced form of cytochrome P-450. Reduction of polyhalogenated compounds gives rise to several reactive intermediates, such as radicals, carbenes, and peroxides, whose participation to toxic effect varies greatly [18,39].

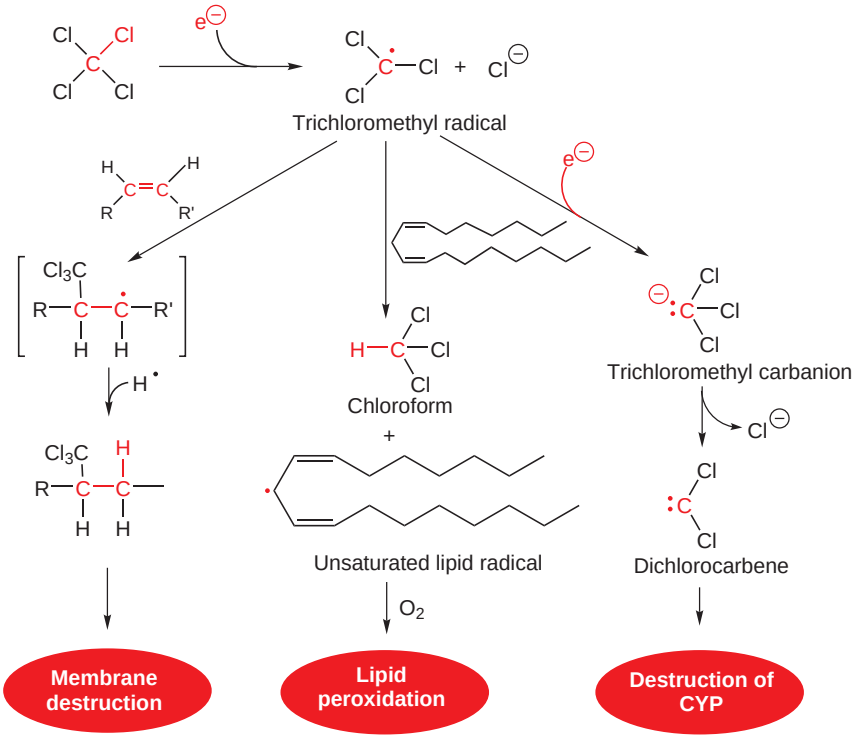


FIGURE 25.15 Reduction of polyhalogenated compounds.

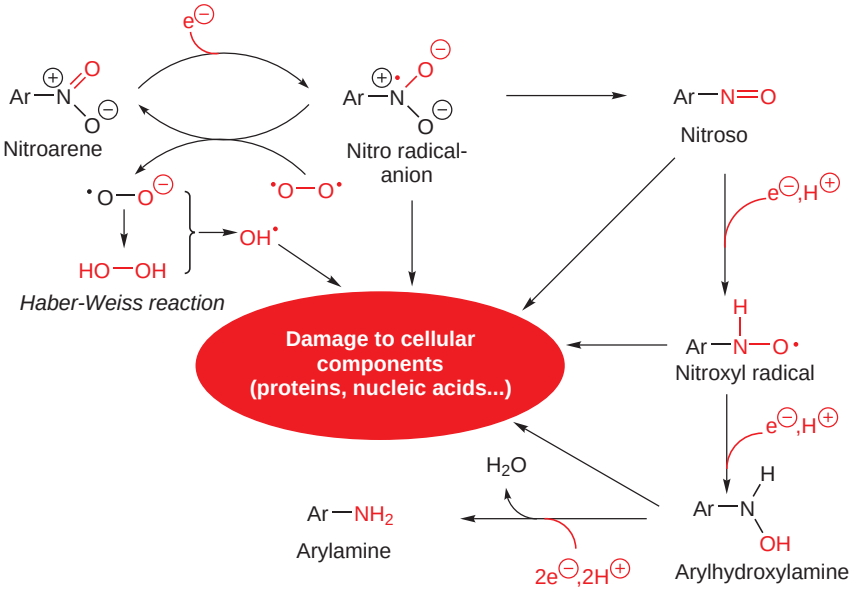


FIGURE 25.16 Reductive biotransformation of nitro arene compounds.

2. *Reduction of Nitro Compounds*

The different steps of the biotransformations that produce a primary amine from an aromatic nitro compound involve a nitro radical-anion, a nitroso derivative, a nitroxyl radical, a hydroxylamine, and then the primary amine (Figure 25.16).

Each of these different intermediates may contribute to toxicity. Hydroxylamines are often responsible for methemoglobinemia [40], whereas mutagenic and carcinogenic activity may be due to the combination of nitro radical-anion, nitroso derivatives, or esterified hydroxylamine (such as sulfate derivatives) with cellular macromolecules.

Carcinogenicity may also be the result of the oxidative stress subsequent to the formation of oxygen reduction products (superoxide anion, hydrogen peroxide, hydroxyl radical) during redox cycling of the nitro radical-anion, which restores the parent nitro compound.

3. *Reduction of Azo Compounds*

Azo compounds are susceptible to reduction, first to hydrazo intermediates, which are reductively cleaved into the appropriate amines. It has been proposed that the first step, as with nitro compounds, is the formation of an azo-anion radical [41].

D. Hydrolysis and Conjugation

Hydrolysis of ester and amide hydrolysis produces more polar compounds than the substrate and are often involved during detoxification processes. Both specific enzymatic and chemical hydrolysis may occur. Acid-catalyzed reactions may occur in the stomach and the kidneys, whereas base-catalyzed reactions may be assisted by the alkaline pH of the intestine.

Phase II or conjugation reactions are also substitution reactions that proceed by means of an endogenous and generally activated electrophile. In mammals, five major conjugation reactions of xenobiotics exist and are mediated by transferase enzymes. Acid compounds, through their acylCoA ester, may also be conjugated with amino acids such as glycine, glutamine, and taurine. The specificity for the endogenous agent is high, but the specificity for the xenobiotic is broader.

To a great extent, conjugation produces excretable and nontoxic metabolites and thus is referred to as detoxification, but exceptions exist in each class of conjugation reaction. A more in-depth discussion of Phase II metabolism can be found in Chapter 24.

1. *Glucuronic acid Conjugation*

This substitution involves the transfer of a glucuronic acid from uridine diphosphate glucuronic acid (UDPGA) to a functional group in the xenobiotic substrate. The group may be a hydroxyl, carboxylic acid, or an amino or sulfur functional group. Most glucuronides are not implicated in toxicity. However, acyl-glucuronide can rearrange in acidic medium and lead to reactive intermediates that have been implicated in toxicity of profens and diclofenac [42,43]. Aromatic amines may be converted in the liver into hydroxylamine *O*-glucuronides, which are excreted in the urine and broken down in the bladder (if its pH is acidic) to liberate the proximate hydroxylamine carcinogen.

2. *Sulfate Conjugation*

Sulfate conjugation gives a polar and ionized conjugate by means of the esterification of a hydroxyl group with sulfate ion (transferred from 3′-phosphoadenosine-5′-phosphosulfate or PAPS). The reaction is catalyzed by a hydrosoluble sulfotransferase. Sulfation sometimes gives rise to reactive intermediates that may undergo further reactions to yield electrophilic metabolites. In the case of 2-acetaminofluorene, the *O*-sulfate moiety is a facile leaving group, and this cleavage produces nitrenium ions, which act as alkylating agents for DNA (Figure 25.12).

3. *Acetylation*

Acetylation is a very common metabolic reaction that occurs with amino, hydroxyl, or sulfhydryl groups. The acetyl group is transferred from acetyl-coenzyme A, and the reaction is catalyzed by acetyltransferases. An important aspect of this kind of substitution is the genetic polymorphism of one-acetyltransferase in humans, who are divided into fast and slow acetylators. In a few cases, the conjugates are further metabolized to toxic compounds, as is seen with isoniazid. Some evidence exists that acetylation of the antitubercular isoniazid leads to enhanced hepatotoxicity of the drug [44,45]. Acetylation followed by hydrolysis and cytochrome P450-dependent oxidation yields free acetyl radicals [46] or acylium cations, which may acetylate the nucleophilic macromolecule functions (Figure 25.17).

4. *Glutathione conjugation*

Substitution reactions of xenobiotics with glutathione are the most important conjugation reactions and contribute efficiently to detoxification. Nevertheless, in some cases, such as vicinal dihalogenated compounds, glutathione conjugation produces monosubstituted derivatives, which may cyclize into a highly electrophilic episulfonium ion (Figure 25.18) [47].

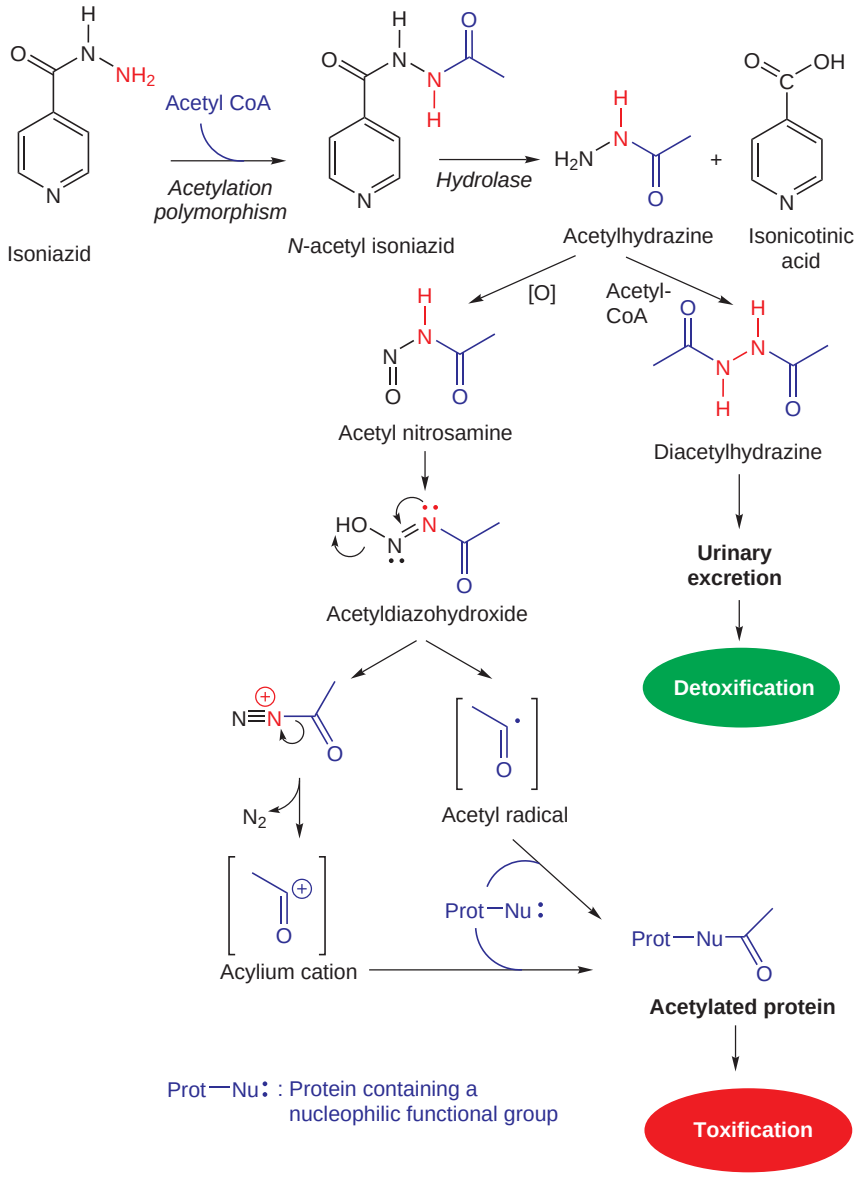


FIGURE 25.17 Bioactivation of isoniazid.

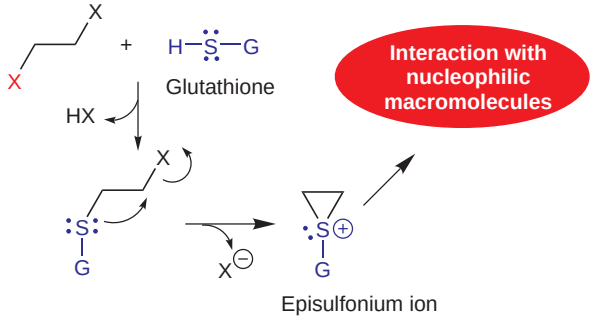


FIGURE 25.18 Bioactivation to episulfonium ion.

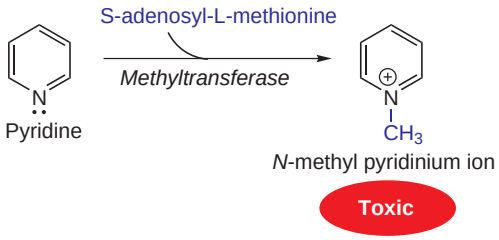


FIGURE 25.19 Bioactivation of pyridine.

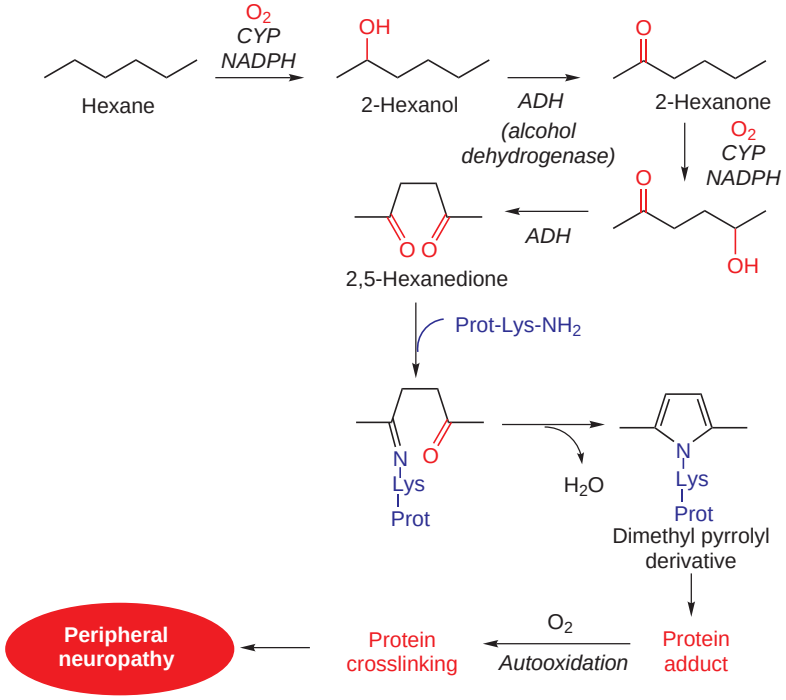


FIGURE 25.20 Bioactivation of hexane.

5. *Methylation*

Methylation is rarely of quantitative importance in the metabolism of xenobiotics. The methyl group is transferred from the nucleotide *S*-adenosyl-L-methionine (SAM) by means of a methyltransferase. The functional groups that undergo methylation include primary, secondary, and tertiary amines, pyridines, phenols, catechols, and thiophenols. The azaheterocycle pyridine is metabolized to the *N*-methylpyridinium ion, which is more polar than pyridine itself (Figure 25.19) [48]. Some pyridinium derivatives, like paraquat or MPTP, can be easily reduced to radical anions, causing oxidative stress and toxicity.

E. Eliminations

Eliminations of hydrogen and a halogen sometimes occur during the metabolism of halogenated xenobiotics and lead to an alkene. The double bond may be oxidized into an epoxide by means of oxidative enzyme systems as discussed above. Dehydrogenation, dehydrochlorination, and dechlorination are (with oxidation) the different metabolic pathways of the γ-isomer of the insecticide hexachlorocyclohexane (lindane) [49].

F. Further Biotransformations Leading to the Ultimate Toxicant

Other reactions in addition to the major ones described above must be mentioned. These reactions may be responsible for the transformation of a toxic metabolite into the ultimate toxicant [4]. Rearrangements and cyclizations are examples of reactions involved in these processes. In the case of the solvent hexane (Figure 25.20), the toxic metabolite 2,5-hexanedione is formed by four successive oxidations of the molecule. The condensation of the γ-diketone with the lysyl amino group of a neurofilament protein is followed by a

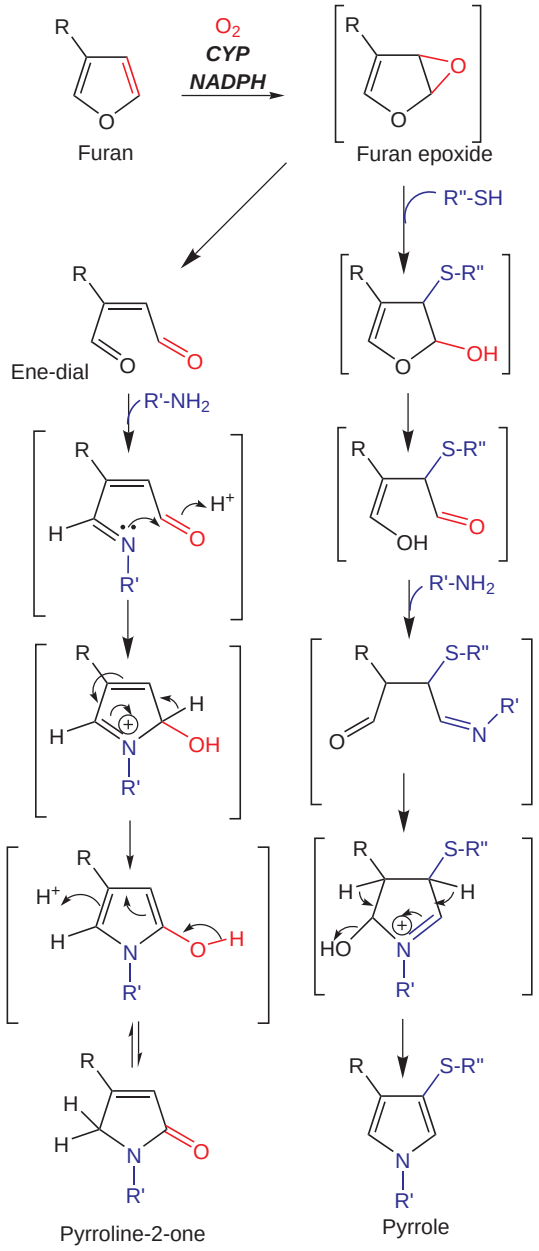


FIGURE 25.21 : Bioactivation of furans.

Paal–Knorr cyclization reaction. This is the initial process that explains the hexane-induced neurotoxicity [50]. A further auto-oxidation of the *N*-pyrrolyl derivatives leads to the cross-linking of the axonal intermediate filament proteins and the subsequent occurrence of peripheral neurotoxicity [51].

Analogous pyrrolyl derivatives are also found as furan metabolites. Furans are oxidized by cytochrome P-450 to reactive furan-epoxides, which rearrange to ene-dial or ene-keto-aldehyde metabolites (Figure 25.21) [31,52,53]. After reaction with thiols and amines like lysine, they form stable pyrrolic derivatives. This first depletes the cell of glutathione before creating cross-links in proteins and eliciting toxicity.

IV. EXAMPLES OF METABOLIC CONVERSIONS LEADING TO TOXIC METABOLITES

The formation of toxic metabolites and/or intermediates during the metabolism of drugs may result from a considerable variety of pathways mediated by several enzyme systems. The following six examples do not

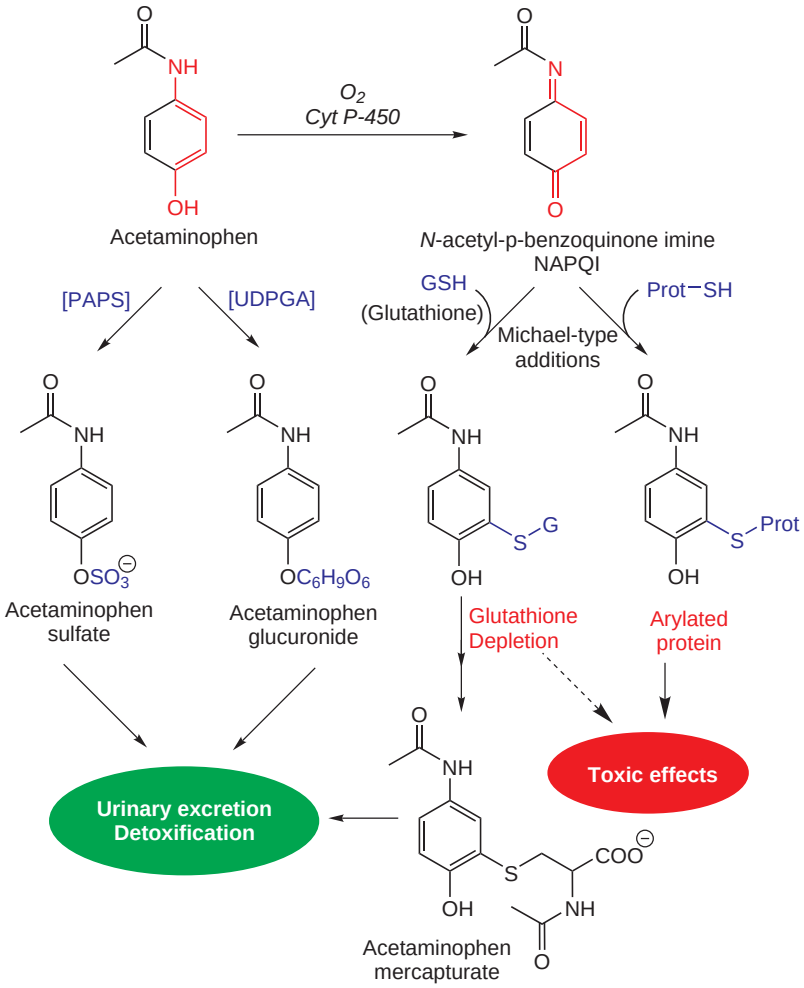


FIGURE 25.22 Biotransformation pathway of acetaminophen.

represent an exhaustive list of bioactivation processes, but are samples of original, significant, or well-known drugs whose biotransformations lead to toxic compounds by the main types of reactions discussed above. Two of them (acetaminophen and tienilic acid) are cytochrome P-450-mediated oxidations. Halothane acts through both oxidative and reductive biotransformations. Valproic acid is toxic through its elimination product. The toxicity of troglitazone seems to involve two distinct metabolic pathways, leading to both alkylating and oxidative stresses. Tetrahydrothienopyridine need two successive P-450-mediated reactions.

A. Acetaminophen

The analgesic acetaminophen (4-hydroxyacetanilide, paracetamol) exhibits lethal hepatotoxicity when administered in very high doses (approximately 250 mg kg^{-1} in rat and about 13 g for a 75 kg human) [54]. The metabolite responsible is known to be the *N*-acetyl-*p*-benzoquinone imine (NAPQI; Figure 25.22) [55].

The formation of NAPQI may proceed via CYP2E1 [56] and via peroxidases such as prostaglandin hydroperoxidase. The most commonly described mechanism proposes that metabolic activation occurs through *N*-oxidation of acetaminophen to *N*-hydroxyacetaminophen followed by dehydration to NAPQI (Figure 25.23) [57].

It seems, however, that *N*-hydroxyacetaminophen is not a major intermediate in the oxidation of acetaminophen. The formation of *N*-acetyl-*p*-benzo-quinone imine probably proceeds by two successive one-electron oxidations (Figure 25.24) [58].

During the first step, a one-electron oxidation yields a phenoxy radical (Ar-$O^{\bullet -}$) [59]. The presence of this radical has been supported by fast-flow ESR spectroscopy in the presence of horseradish peroxidase. In the second one-electron oxidation step, the phenoxy radical is oxidized to NAPQI. As described in Figure 25.22, the highly electrophilic NAPQI may easily react with glutathione or protein thiol groups according to a Michael-type addition.

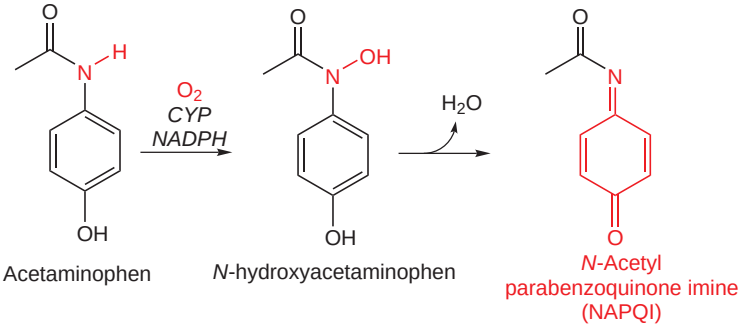


FIGURE 25.23 Oxidation of acetaminophen according to the *N*-hydroxyacetamin-ophen pathway.

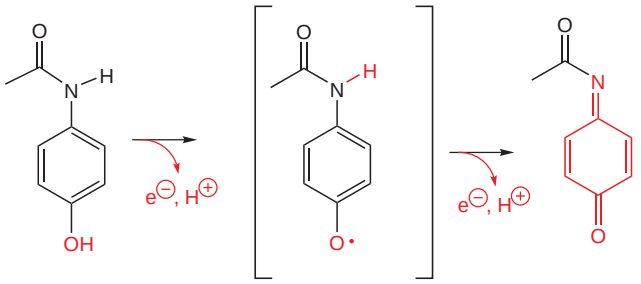


FIGURE 25.24 Oxidation of acetaminophen by means of the phenoxy radical.

The attack of liver protein thiol groups and the subsequent adduct formation is frequently mentioned in the mechanism of acetaminophen hepatotoxicity. In mice, a number of proteins were identified, such as glyceraldehyde-3-phosphate dehydrogenase [60], calreticulin, and the thiol: protein disulfide reductases Q1 and Q5 [61], and this number is increasing with the advances of proteomics [62].

Another hypothesis for the mechanism of toxicity is supported by the oxidative potency of NAPQI, but still suffers from lack of evidence [63]. NAPQI is a good oxidant for thiols of cellular components and pyridine nucleotides. Moreover, it may undergo redox cycling with formation of superoxide anion by means of an oxygen one-electron reduction (Figure 25.25).

The stepwise reduction of oxygen produces hydrogen peroxide and then a hydroxyl radical, a strong oxidant implicated in cellular oxidative stress. This oxidative stress causes glutathione depletion, a disruption of the cellular calcium regulation and modifications of cellular proteins, thus leading to cell death. Some biochemical parameters related to necrotic and apoptotic processes are affected in acetaminophen-exposed PC12 cells transfected with CYP2E1 [64,65].

It therefore appears that both covalent (e.g., alkylation) and noncovalent (e.g., oxidative stress) interactions play major roles in the pathogenesis of acute lethal cell injury caused by NAPQI [66]. At present, it is not possible to identify which of these two interactions is the critical event in initiating acetaminophen hepatotoxicity, even if some authors suggest that the characteristic features of oxidative stress are more likely the consequences of damage mediated by protein adduction [67].

B. Tienilic Acid

Tienilic acid is a uricosuric diuretic drug that may cause immunoallergic hepatitis in 1 in 10,000 patients, a side effect that resulted in its withdrawal from the market. The immunoallergic hepatitis was associated with the appearance of circulating antireticulum antibodies called anti-LKM_2 antibodies, which are directed toward a liver and kidney endoplasmic reticulum protein [68–70]. From these observations, the mechanism of the immunotoxicity associated with the prolonged use of tienilic acid was elucidated by the Mansuy group [69,71,72].

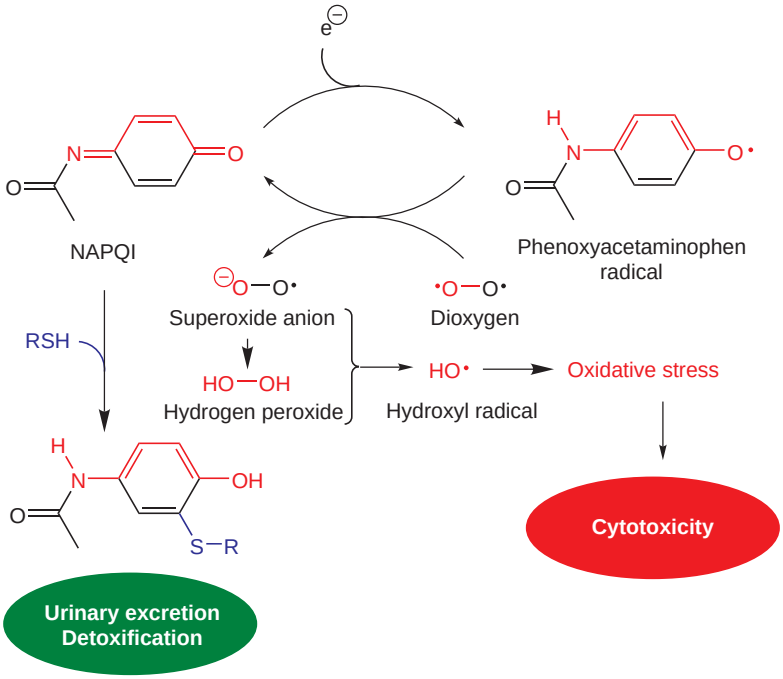


FIGURE 25.25 Redox cycling of *N*-acetyl-*p*-benzoquinone imine.

Tienilic acid is oxidized in the liver by cytochrome P-450 monooxygenase to 5-hydroxytienilic acid, which is the major urinary metabolite (about 50 percent in human). In humans, the bioactivation of tienilic acid depends on CYP2C9. This isoform is one of the major forms of cytochrome P-450 in the human liver. This oxidation occurs through an electrophilic intermediate capable of very specifically alkylating cytochrome P-450, leading to its inactivation [71,73,74]. This mechanism-based inactivation is also observed with many xenobiotics, such as alkenes with terminal unsaturation, alkynes, strained cycloalkylamines, 4-alkyldihydropyridines, benzodioxoles, and some tertiary amines [21,75]. The irreversible binding of the compound with cytochrome P-450 leads to an immune response and to generation of antibodies against both the modified protein and its native form. In fact, the autoantibodies anti-LKM2 present in hepatitis patients recognize CYP2C9 both as native protein and as modified protein. In addition, patient sera contain antibodies to tienilic acid-modified proteins. It has also been demonstrated in a rat model that tienilic acid-modified CYP2C9 is exported to the plasma membrane of hepatocytes [76]. Rabbit hepatocytes, when first treated with tienilic acid then incubated with anti-LKM2, were lysed by human NK cells [77]. Thus, it is hypothesized that vesicular transfer of tienilic acid-modified proteins on the plasma membrane of hepatocytes triggers their cytolysis.

In the case of tienilic acid, the electrophilic reactive species is unknown but is believed to be either thiophene sulfoxide, as has been demonstrated for its 3-isomer [78], or a thiophene epoxide (Figure 25.26) [30,79]. In both cases, the electrophilic character of the intermediate is enhanced by the presence of an activating 2-keto group. In any event, this electrophilic species reacts with CYP2C9 where it is produced and inactivates it efficiently (one inactivation event every thirteen turnovers) [71]. The covalent binding of tienilic acid to CYP2C9 has been directly observed by mass spectrometry [80]. This reaction occurs in all patients with active CYP2C9 using this drug, but very few produce anti-LKM2 and have hepatitis, which suggests some specificity in their immune response.

C. Halothane

Halothane is a widely used anesthetic drug that occasionally results in severe hepatitis. About 60 to 80 percent of the dose is eliminated in unmetabolized form during the twenty-four hours following administration to patients. This compound is metabolized in the presence of cytochrome P450 monooxygenase CYP2E1 according to the two main pathways depicted in Figure 25.27 [18].

The major biotransformation pathway involves an oxidative step with introduction of an oxygen atom and subsequent formation of halohydrin. The unstable halohydrin loses hydrobromic acid to yield trifluoroacetyl chloride, which in turn is hydrolyzed to trifluoroacetic acid. This final metabolite is found in the urine [81].

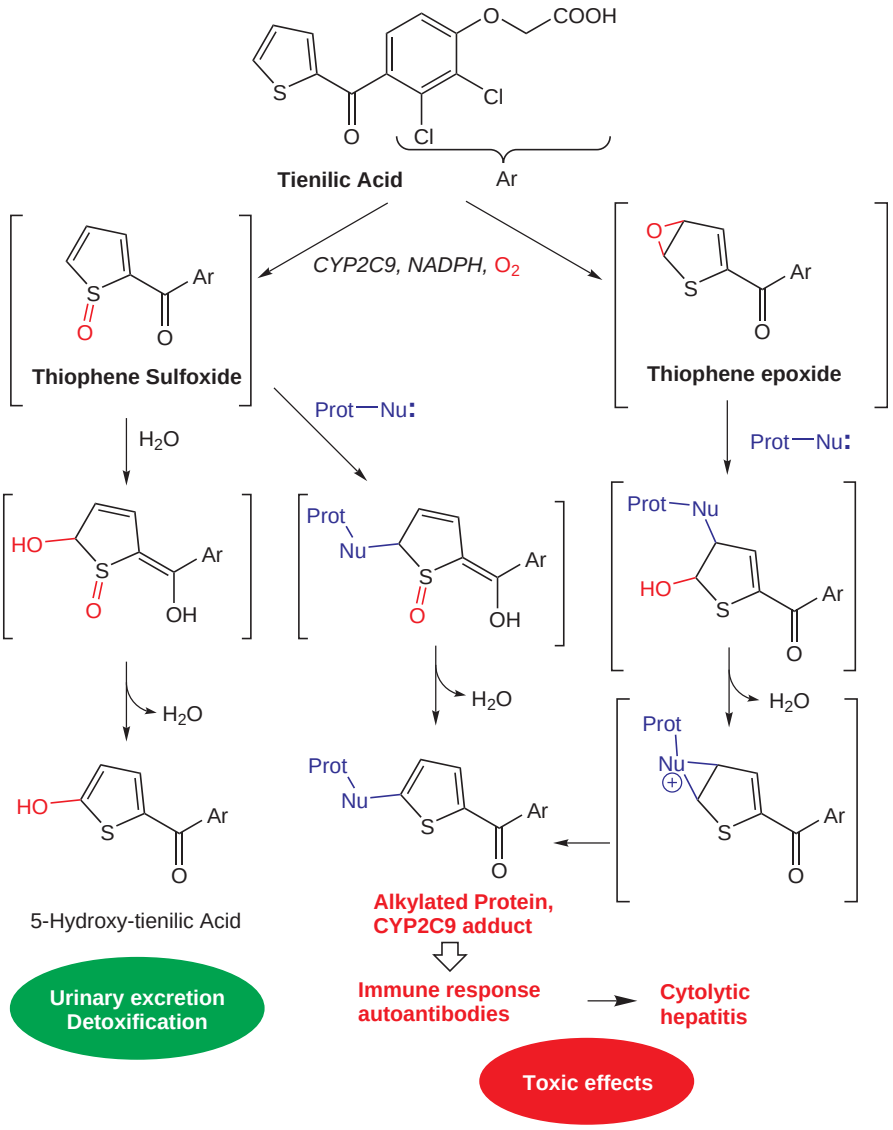


FIGURE 25.26 Tienilic acid biotransformation to reactive intermediates and stable metabolites.

In conditions of low levels of oxygen, a reductive pathway (10 percent) is enhanced and yields a free radical intermediate characterized as 1-chloro-2,2,2-trifluoroethyl radical. Another one-electron reduction produces the 1-chloro-2,2,2-trifluoroethyl carbanion, which may undergo two possible kinds of eliminations.

One is the abstraction of a fluoride ion according to a E1cB elimination, which yields 1-chloro-2,2-difluoroethylene. This metabolite is eliminated by exhalation. Early studies suggested that a second elimination process might be an α-elimination of a chloride ion, which produces trifluoromethylcarbene [82], but this was later reconsidered [83]. It was hypothesized that a carbene complex with the Fe^{II} in the active site might lead to inactivation of the cytochrome P450, but this inactivation is now thought to be due to the formation of an iron-σ-alkyl complex derived from the 1-chloro-2,2,2-trifluoroethyl radical.

The initially formed 1-chloro-2,2,2-trifluoroethyl radical may also cause a radical attack of polyunsaturated lipids, which produces 1-chloro-2,2,2-trifluoroethane (Figure 25.27). This mechanism is similar to the pathway described with the trichloromethyl radical formed during the one-electron reduction of carbon tetrachloride (Figure 25.14). The trichloromethyl radical may initiate a peroxidation of unsaturated lipids from the membrane with subsequent liberation of chloroform.

Several studies have demonstrated that halothane hepatotoxicity is mainly due to an immune reaction to modified proteins of the liver. In fact, these proteins are trifluoroacetylated on their ε-NH_2 lysyl residue by the

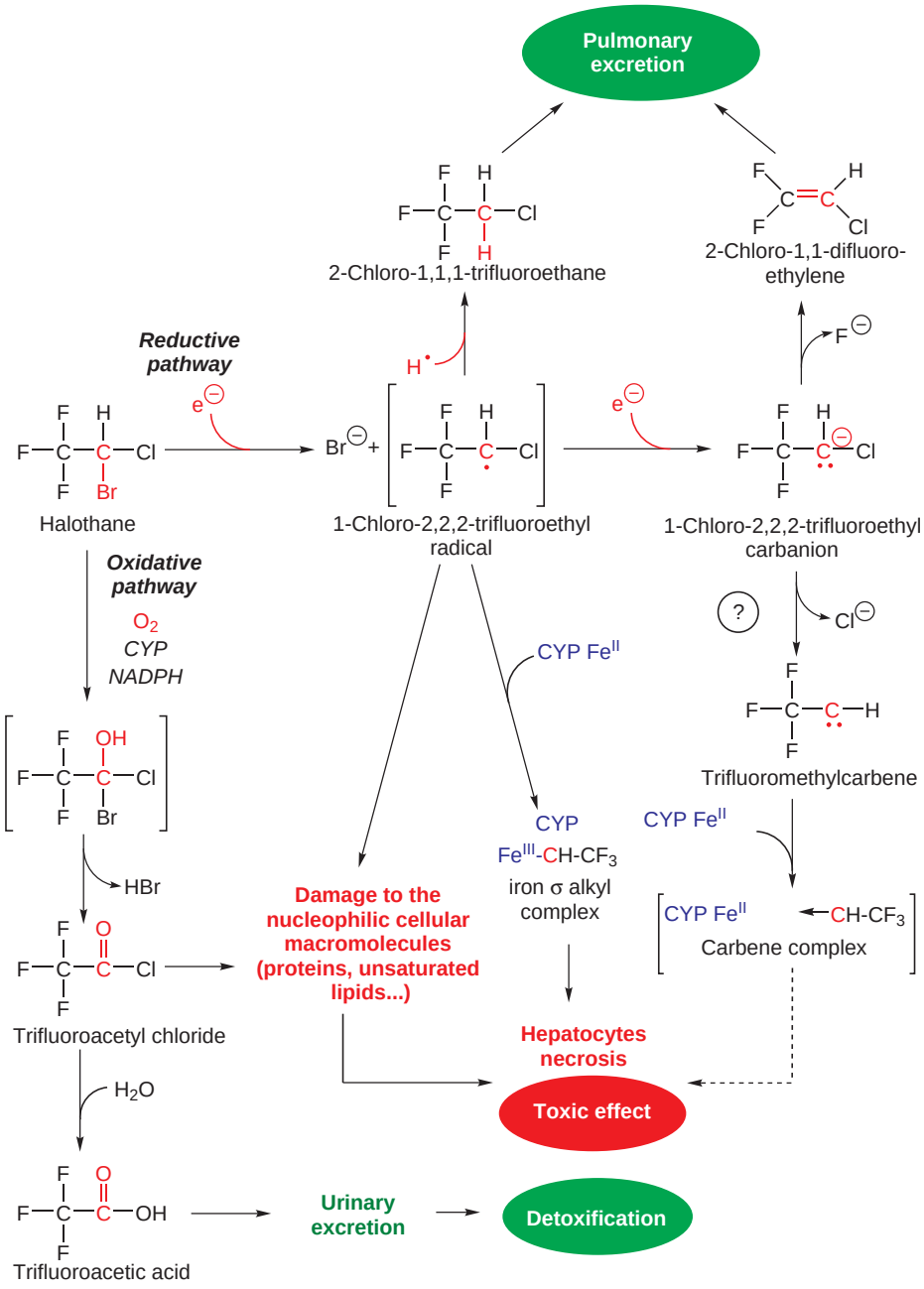


FIGURE 25.27 The major metabolic pathways of halothane.

trifluoroacetyl chloride formed during the oxidative metabolism of halothane [84,85]. The product of the reaction can act as a foreign epitope, and the drug–protein conjugate, called neoantigen, elicits an immune response toward the liver (Figure 25.28) [86].

A related fluorocarbon used in air conditioning systems, HCFC 1,2,3, is metabolized to the same acyl halide and was recently implicated in an epidemic of liver disease in nine workers in a Belgian factory [87]. All patients had serum antibodies to trifluoroacetylated proteins.

D. Valproic Acid

Valproic acid is an anticonvulsant agent used for the therapy of epilepsy, which occasionally results in hepatotoxicity in young children. The toxicity is characterized by mitochondrial damage, impairment of fatty acid β-oxidation and lipid accumulation.

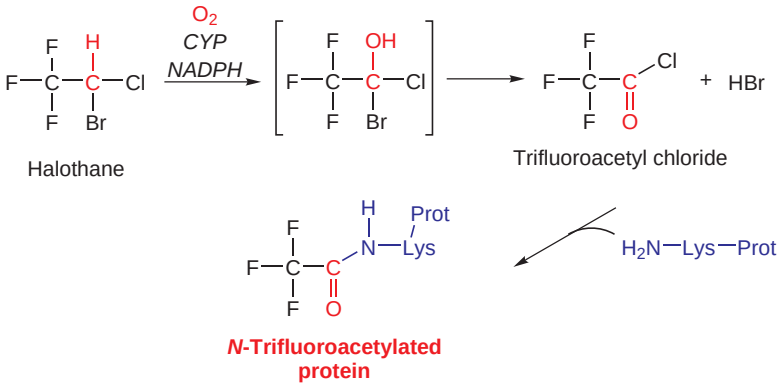


FIGURE 25.28 Biotransformation of halothane to trifluoroacetyl chloride and the subsequent binding to protein.

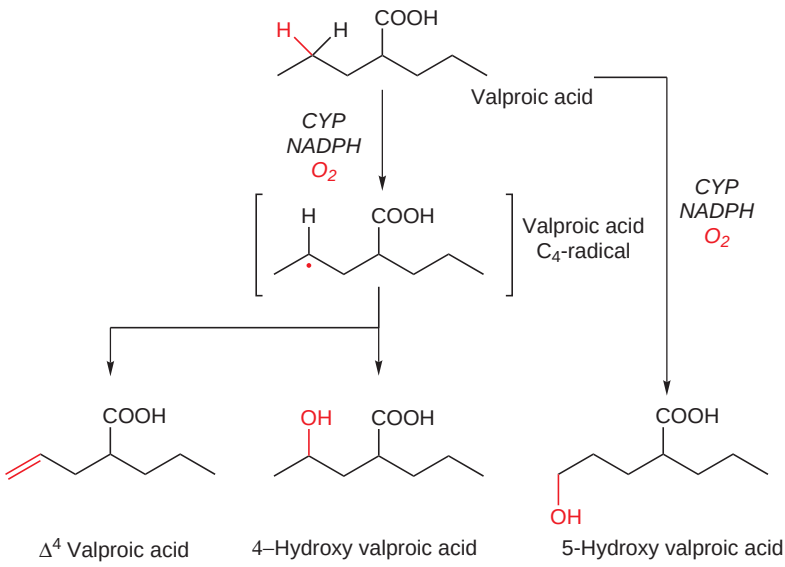


FIGURE 25.29 Bioactivation of valproic acid to Δ^4VPA.

It has been proposed that hepatotoxicity is a consequence of the further biotransformation of the valproic acid metabolite 2-propyl-4-pentenoic acid (also called Δ^4VPA) [88].

As depicted in Figure 25.29, Δ^4VPA is not formed by dehydration of 4- or 5-hydroxy valproic acids, which are—with the glucuronide conjugate—the major metabolites of valproic acid [89]. The mechanism is proposed to involve an initial hydrogen abstraction to generate a transient free radical intermediate. It has been demonstrated that the carbon-centered radical was localized at the C4 position. The radical undergoes both recombination (which yields 4-hydroxy valproic acid) and elimination (which produces the unsaturated derivative Δ^4VPA). The formation of these metabolites is catalyzed in rats by CYP4B1 [90] and in humans by CYP2C9 [14]. Δ^4VPA is a hepatotoxic and strong teratogenic compound in animal models. In addition to that metabolic pathway, valproic acid undergoes biotransformation leading to *(E)*-Δ^2VPA, which is devoid of embryotoxic effect in rodents [91].

Further biotransformations of Δ^4VPA involve both the liver microsomal cytochrome P-450 enzymes and the fatty acid β-oxidation pathway (Figure 25.30). The mixed-function oxidase system metabolizes the unsaturated metabolite to a γ-butyrolactone [92] derivative through a chemically reactive entity that is a mechanism-based inhibitor of cytochrome P-450. The alkylation of the prosthetic heme by means of the radical occurs prior to formation of the epoxide [93]. Thus, the epoxide is not involved in cytochrome P450 inhibition.

The β-oxidation cycle activates Δ^4VPA to its coenzyme A derivative and, through sequential steps of β-oxidation, yields the coenzyme A ester of 3-oxo 2-propyl-4-pentenoic acid [94]. This final metabolite is believed to be a reactive electrophilic species that alkylates 3-ketoacyl-CoA thiolase (the terminal enzyme of β-oxidation) by means of a Michael-type addition through nucleophilic attack at the olefinic terminus [95]. Oxidative stress may also be implicated, at least in part, in valproic acid hepatotoxicity, as suggested by experimental data on the drug's effect on reactive oxygen species [96].

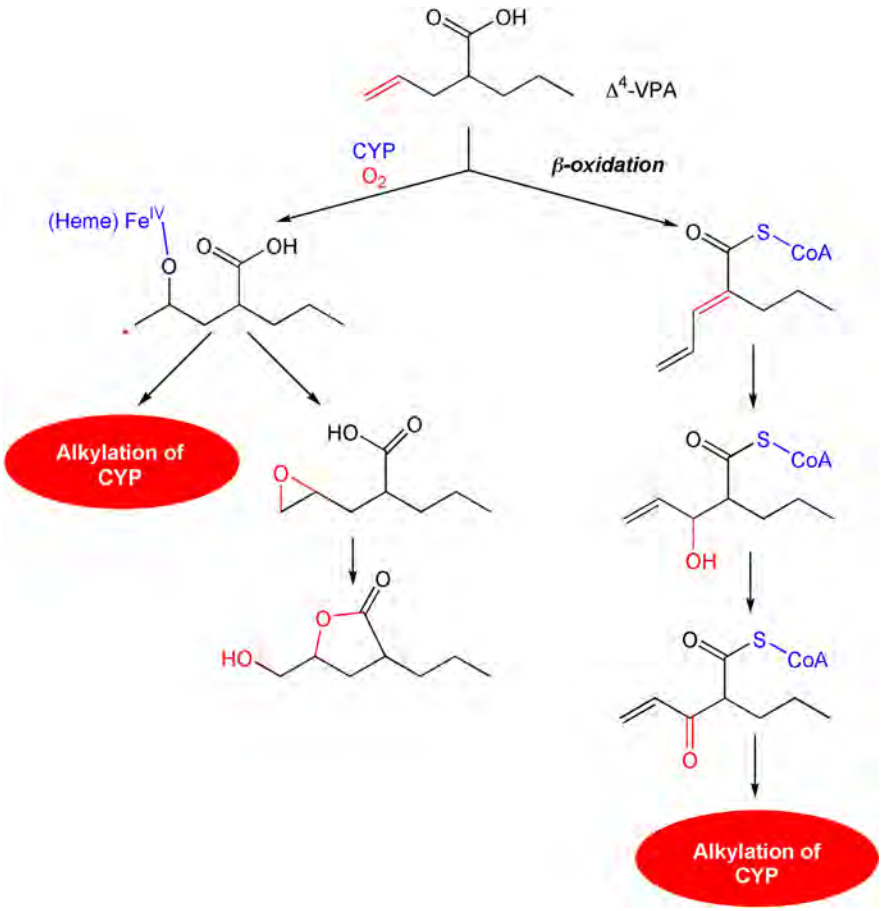


FIGURE 25.30 Bioactivation of Δ^4VPA.

E. Troglitazone

Troglitazone ((±)-5-[4-(6-hydroxy-2,5,7,8-tetramethyl-chroman-2-ylmethoxy) benzyl]-2,4-thiazolidinedione) is an oral insulin sensitizer belonging to the thiazolidinedione class of compounds used for the treatment of type II diabetes. Its withdrawal from the US market was the consequence of the recent occurrence of hepatic failure, sometimes leading to death.

It was first demonstrated that troglitazone is metabolized mainly to sulfate and glucuronide conjugates [97]. Troglitazone is also an inducer of CYP3A [98]. Its mechanism of toxicity is still unclear but seems to proceed according to two distinct pathways. This is supported by the demonstration that incubation of troglitazone with P-450 isoforms in the presence of glutathione give rise to at least five GSH conjugates [35,99]. Identification of these adducts provided evidence for the two pathways described in Figures 25.31 and 25.32.

As described in Figure 25.31, oxidative cleavage of the thiazolidinedione ring is probably oxidized into a highly electrophilic acyl-sulfoxide cleaved into reactive α-ketoisocyanate and sulfenic acid intermediates. This P-450 3A-mediated oxidation would afford a reactive sulfoxide intermediate that undergoes a spontaneous ring opening. The sulfonic acid protein adduct also shows that the intermediate acyl-sufoxide can be attached by nucleophiles on the carbonyl function.

The second pathway (Figure 25.32) consists of a CYP3A-mediated [100] one-electron oxidation of the phenolic hydroxyl group leading to an unstable hemiacetal, which opens spontaneously to form the quinone metabolite. This undergoes thiazolidinedione-ring oxidation according to the pathway shown in Figure 25.31. Alternatively, a P-450-mediated hydrogen abstraction may occur on the phenoxy radical, leading to an *o*-quinone methide derivative.

It is now well established that troglitazone undergoes several metabolic transformation mediated by CYP3A4, leading to numerous electrophilic species [101]. Thus, toxicity probably acts both by covalent binding to hepatic proteins and by oxidative stress through a redox cycling process. The implication of the thiazolidinedione moiety

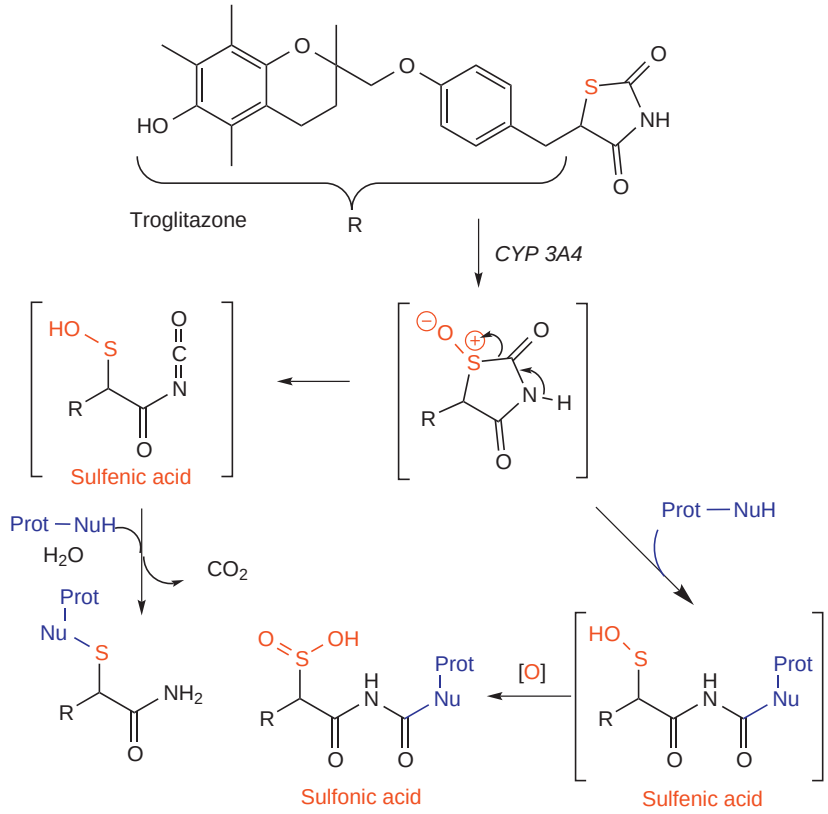


FIGURE 25.31 Oxidation of the thiazolidinedione ring of troglitazone.

is less likely, since the more recent drugs of this series seems devoid of toxicity. Recent studies using mitochondrial manganese superoxide dismutase partially deficient mice also suggested that genetic deficiencies may be, at least partially, responsible for the liver failure in troglitazone-treated patients [102,103].

F. Tetrahydro-Thienopyridines

Tetrahydro-thienopyridines antiaggregant prodrugs belong to a new class of "covalent drugs." [104] These antithromboplastic compounds must be activated to an active thiol metabolite, which inactivates the ADP platelet receptor P2Y12. This is a four-step reaction [105,106]. The first step of ticlopidine and clopidogrel activation is a CYP-dependent oxidation of the thiophene ring to a thiolenone. The first step of prasugrel activation is hydrolysis of the acetoxy-thiophene by a carboxylesterase to a similar thiolenone. The common second step is CYP-catalyzed oxidation of the thiolenone to a very reactive thiolenone sulfoxide, which is hydrolyzed by water into an even more reactive sulfenic acid. In the presence of thiol, this sulfenic acid intermediate yield a mixed disulfide (third step) that can be reduced enzymatically or nonenzymatically by excess thiol or other reducing agents to the active thiol metabolite (fourth step). Sulfenic acid has also be trapped *in vitro* by specific sulfenic trapping agents like dimedone and cyclopentane1,3-dione [105,106]. This thiol binds covalently with a cysteine of the P2Y12 receptor and inactivates it (Figure 25.33). Thus, the receptor can no longer bind ADP and aggregation is suppressed.

However, the thiolenone sulfoxide intermediate is a cyclic electrophile that can react with water, leading first to a carboxylic acid-sulfenic acid then to adducts with thiol (e.g., P2Y12, giving a disulfide) or with sulfenic trapping agents like dimedone or cyclopentane-1,3-dione (Figure 25.33) [34,35].

This thiolenone sulfoxide can also react with other nucleophiles, like amine (primary and secondary), leading to amide-sulfenic acids, or with thiol, leading to thioester-sulfenic acid then to thioester-sulfenic acid adducts. If the first nucleophile is a protein cysteine or lysine, reaction of the sulfenic acid with a thiol of the same or of another protein may lead to protein crosslinks and possibly toxicity (Figure 25.34).

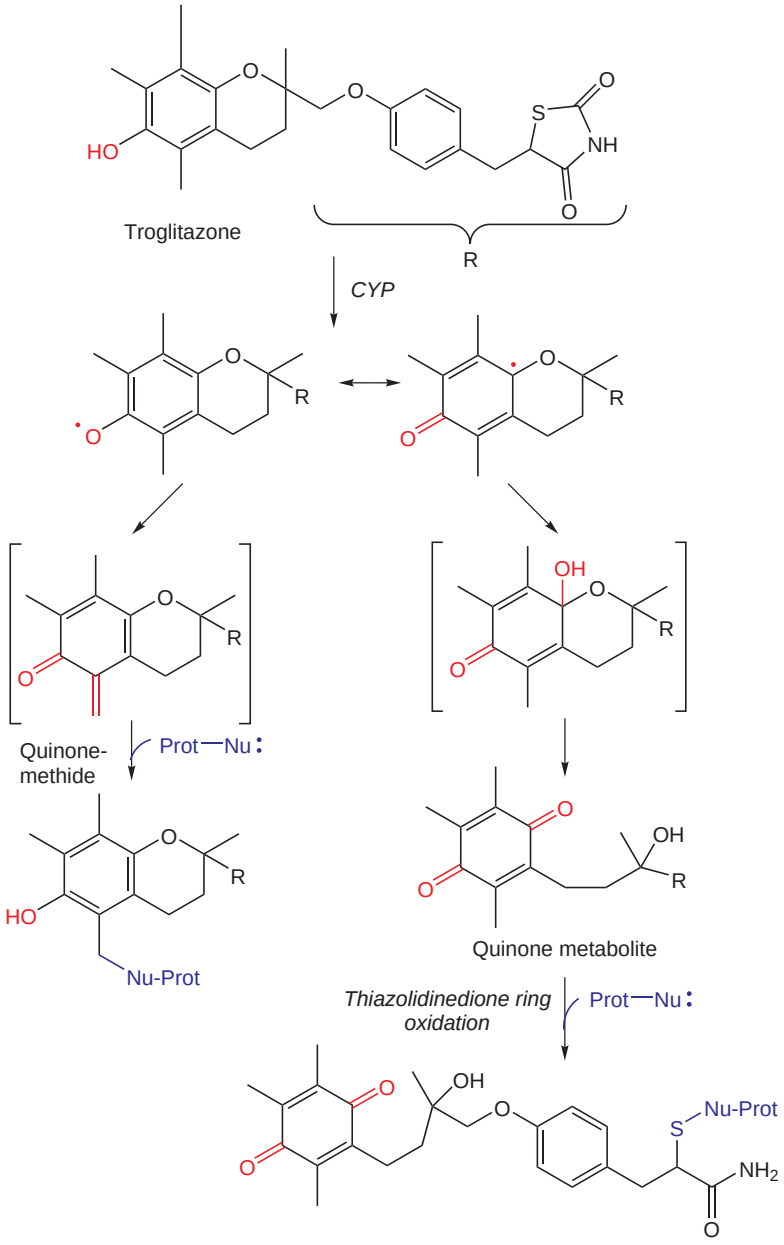


FIGURE 25.32 Oxidation of the chromane ring of troglitazone.

V. CONCLUSION

The foregoing review has emphasized that almost all metabolic reactions are capable of producing reactive metabolites. This bioactivation yields toxic compounds that may act directly or indirectly [66] (Figure 25.35). The emergence of toxicity may be the outcome of the interactions of metabolites or reactive intermediates with biological targets such as cellular macromolecules. Some compounds exhibit their toxicity by inducing the generation of reactive oxygen species, thus producing alterations in the redox state of the cell. Often, covalent bonds are formed during a phenomenon that may be referred to as "alkylating stress" (Figure 25.35). Bioactivation of drugs followed by drug protein adduction is then considered a key sequence in the occurrence of toxic side effects [107] (Figure 25.35). Because the precise damages of adducts on cellular functions are not fully understood, the formation of electrophilic metabolites is to be avoided in drug design. Proteome profiling (proteomics) may help to identify and compare proteins implicated in alkylating stress due to drugs, but this field remains to be developed and methods must be validated. The specific inhibition of an enzyme by its own substrate is a peculiar

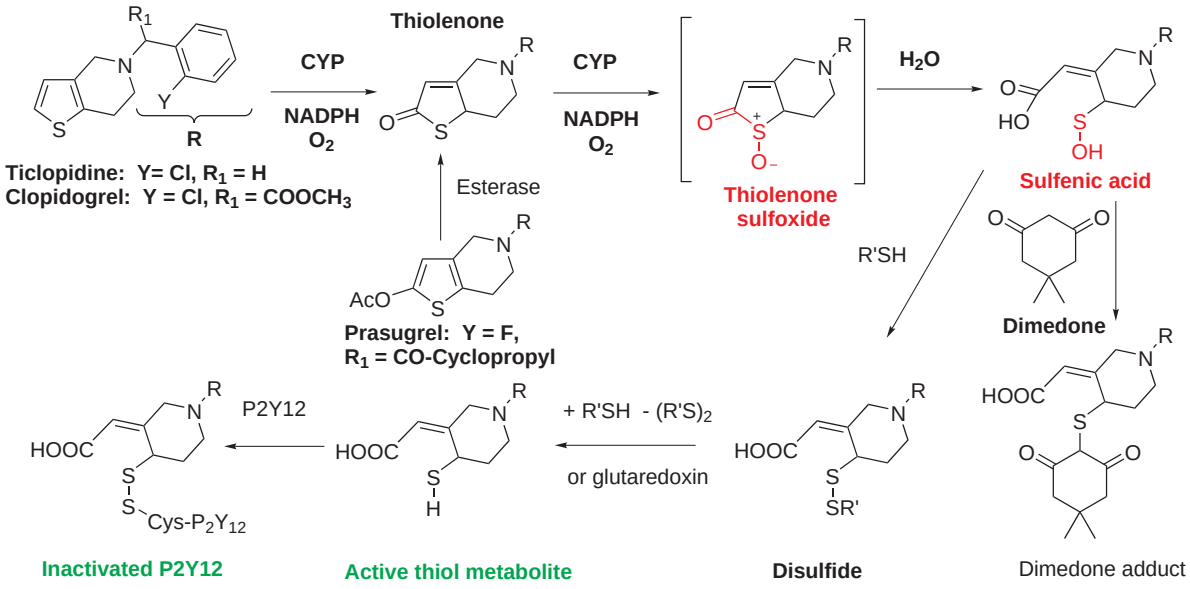


FIGURE 25.33 Activation of thienopyridine drugs ticlopidine, clopidogrel and prasugrel into the pharmacologically active thiol metabolite and trapping of the sulfenic acid intermediate with dimedone.

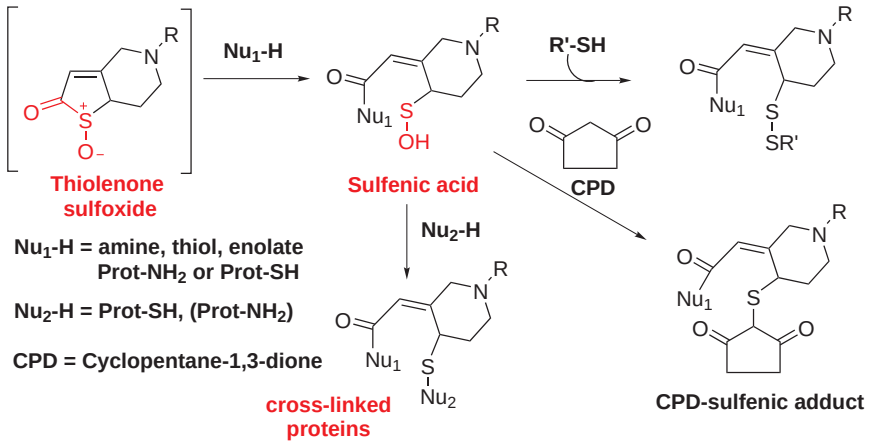


FIGURE 25.34 Reactive thiolenone sufoxide intermediates are bis-alkylating and eventually crosslinking agents.

feature of alkylating stress. Determination and monitoring of drug protein adducts have important implications in drug development, for example, in identifying CYP3A4 inactivation, since this cytochrome P-450 isoform is responsible for the metabolism of about 50 percent of therapeutic drugs [108]. Thus, medicinal chemists have set a threshold of acceptable covalent binding when developing a new drug. For example, the value for covalent binding levels to liver proteins was less than 50 pmol-equiv/mg protein under standard conditions at Merck [109]. This can be subject to discussion on a case-by-case basis. This target represents about 1/20th of the level of binding for model hepatotoxins. Often the molecule can be modified to decrease this type of unwanted reaction without losing too much pharmacological activity [110–112].

Such a variety of mechanisms makes it difficult to point to molecular functions susceptible to producing toxic effects through bioactivation. However, some major toxophoric groups may be highlighted (Table 25.1). They may be implicated in acute or chronic toxicity. These patterns must be of particular concern in drug design. A number of recent papers on these matters have been published on how to avoid those toxic events in drug design [112–114].

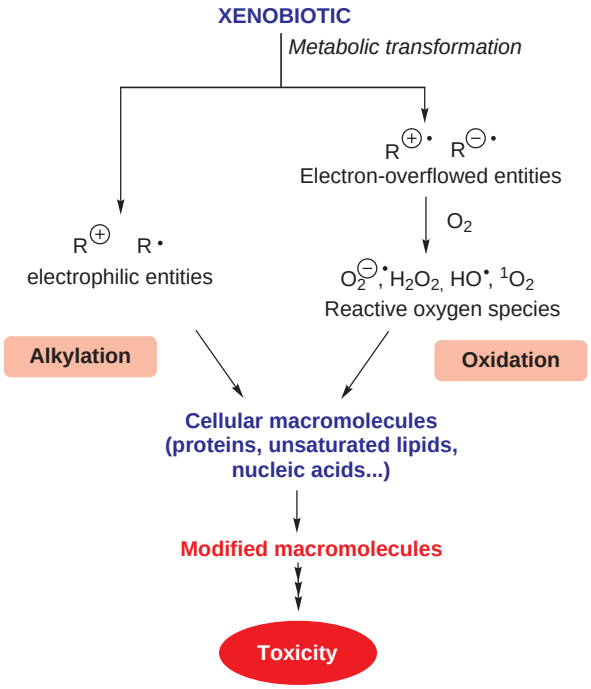


FIGURE 25.35 Alkylating and oxidative stresses.

TABLE 25.1 Some Major Toxophoric Groups and Their Bioactivation Mechanisms

Toxophoric group	Bioactivation mechanism
Azocompounds	Nitrenium ions, tautomeric carbonium ions
Acetamides	
Aromatic/heterocyclic amines	
Nitro compounds	
Nitroaromatic compounds	radical formation / oxidative stress
Bromoarenes	arene oxide formation
Ethinyl	ketene formation / heme destruction
Furanes	furane epoxide and ene-dial formation
Pyrroles	pyrrole oxide
Nitrogen mustard	aziridium ions
Nitroso compounds	diazonium ions / heme adduct / radical formation
Hydrazines	
Nitrosamines	carbenium ions / DNA alkylation
Polyhalogenated compounds	radical and carbene formation / episulfonium with GSH
Quinone, (quinone imine or methide)	semiquinone radical anion/ oxidative stress / thiol trapping
Thioamides	thiourea formation
Thioester, Thiolactone	Thioester, thiolactone sulfoxide, sulfenic acid
Thiophene	thiophene sulfoxide or thiophene epoxide formation
Vinyl	epoxidation / haem destruction

Generally, the formation of toxic metabolites is not the only pathway of biotransformation, and overall metabolism is constituted toward detoxication and bioactivation processes. The toxic metabolites are themselves often further detoxified. The duality between a beneficial detoxification phenomenon (metabolism, drug resistance) and the occurrence of a toxic effect represents the cost for adaptability of metabolic enzymes to the diversity of xenobiotics. For those interested, a recent review applies the above chemistry to predict drug safety [115].

Acknowledgments

We thank Professor Bernard Testa for helpful corrections and comments.

References

[1] Williams R. Detoxication mechanisms. New York: J Wiley & Sons; 1959.
[2] Jollow DJ, Mitchell JR, Potter WZ, Davis DC, Gillette JR, Brodie BB. Acetaminophen-induced hepatic necrosis. II. Role of covalent binding in vivo. J Pharmacol Exp Ther 1973;187(1):195–202.
[3] Potter WZ, Davis DC, Mitchell JR, Jollow DJ, Gillette JR, Brodie BB. Acetaminophen-induced hepatic necrosis. 3. Cytochrome P-450-mediated covalent binding in vitro. J Pharmacol Exp Ther 1973;187(1):203–10.
[4] Miller EC, Miller JA. Mechanisms of chemical carcinogenesis. Cancer 1981;47(5 Suppl):1055–64.
[5] Omura T, Sato R. The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. J Biol Chem 1964;239:2370–8.
[6] Estabrook RW. A passion for P450s (remembrances of the early history of research on cytochrome P450). Drug Metab Dispos 2003;31(12):1461–73.
[7] Snyder R. Microsomal enzyme induction. Toxicol Sci 2000;55(2):233–4.
[8] Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE. Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. J Inorg Biochem 2000;81(3):183–90.
[9] Sipes G, Gandolfi AJ. Biotransformation of toxicants. In: Amdur MO, Doull J, Klaassen CD, editors. Casarett and doull's toxicology: the basic science of poisons. New York: Pergamon Press; 1991. p. 88–126.
[10] Alvares AP, Pratt WB. Pathways of drug metabolism. In: Pratt WB, Taylor P, editors. Principles of drug action: the basis of pharmacology. New York: Churchill Livingstone; 1990. p. 227–300.
[11] Josephy PD, Guengerich FP, Miners JO. "Phase I and phase II" drug metabolism: terminology that we should phase out? Drug Metabolism Reviews 2005;37(4):575–80.
[12] Guengerich FP. Cytochrome P450 oxidations in the generation of reactive electrophiles: epoxidation and related reactions. Arch Biochem Biophys 2003;409(1):59–71.
[13] Guengerich FP. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. Chem Res Toxicol 2001;14(6):611–50.
[14] Isin EM, Guengerich FP. Complex reactions catalyzed by cytochrome P450 enzymes. Biochim Biophys Acta 2007;1770(3):314–29.
[15] Mason RP, Chignell CF. Free radicals in pharmacology and toxicology—Selected topics. Pharmacol Rev 1981;33(4):189–211.
[16] Aust SD, Chignell CF, Bray TM, Kalyanaraman B, Mason RP. Free radicals in toxicology. Toxicol Appl Pharmacol 1993;120(2):168–78.
[17] Singal PK, Petkau A, Gerrard JM, Hrushovetz S, Foerster J. Free radicals in health and disease. Mol Cell Biochem 1988;84(2):121–2.
[18] Anders MW, Pohl LR. Halogenated alkanes. In: Anders MW, editor. Bioactivation of foreign compounds. Orlando: Academic Press; 1985. p. 284–315.
[19] Pohl LR, Nelson SD, Krishna G. Investigation of the mechanism of the metabolic activation of chloramphenicol by rat liver microsomes: identification of a new metabolite. Biochem Pharmacol 1978;27(4):491–6.
[20] Halpert JR, Miller NE, Gorsky LD. On the mechanism of the inactivation of the major phenobarbital-inducible isozyme of rat liver cytochrome P-450 by chloramphenicol. J Biol Chem 1985;260(14):8397–403.
[21] Fontana E, Dansette PM, Poli SM. Cytochrome p450 enzymes mechanism based inhibitors: common sub-structures and reactivity. Curr Drug Metab 2005;6(5):413–54.
[22] Cresteil T, Beaune P, Leroux JP, Lange M, Mansuy D. Biotransformation of chloroform by rat and human liver microsomes; In vitro effect on some enzyme activities and mechanism of irreversible binding to macromolecules. Chem Biol Interact 1979;24(2):153–65.
[23] Benasutti M, Ejadi S, Whitlow MD, Loechler EL. Mapping the binding site of aflatoxin B1 in DNA: systematic analysis of the reactivity of aflatoxin B1 with guanines in different DNA sequences. Biochemistry 1988;27(1):472–81.
[24] Koen YM, Gogichaeva NV, Alterman MA, Hanzlik RP. A proteomic analysis of bromobenzene reactive metabolite targets in rat liver cytosol in vivo. Chem Res Toxicol 2007;20(3):511–19.
[25] Rietjens IM, den Besten C, Hanzlik RP, van Bladeren PJ. Cytochrome P450-catalyzed oxidation of halobenzene derivatives. Chem Res Toxicol 1997;10(6):629–35.
[26] Verna L, Whysner J, Williams GM. 2-Acetylaminofluorene mechanistic data and risk assessment: DNA reactivity, enhanced cell proliferation, and tumor initiation. Pharmacol Ther 1996;71(1–2):83–105.
[27] Dansette PM, Thang DC, el Amri H, Mansuy D. Evidence for thiophene-S-oxide as a primary reactive metabolite of thiophene in vivo: formation of a dihydrothiophene sulfoxide mercapturic acid. Biochem Biophys Res Commun 1992;186(3):1624–30.
[28] Treiber A, Dansette PM, El Amri H, Girault J-P, Ginderow D, Mornon J-P, et al. Chemical and biological oxidation of thiophene: Preparation and complete characterization of thiophene S-oxide dimers and evidence for thiophene S-oxide as an intermediate in thiophene metabolism in vivo and in vitro. J Am Chem Soc 1997;119(7):1565–71.

[29] Treiber A, Dansette PM, Mansuy D. Mechanism of the aromatic hydroxylation of thiophene by acid-catalyzed peracid oxidation. J Org Chem 2002;67(21):7261–6.
[30] Dansette PM, Bertho G, Mansuy D. First evidence that cytochrome P450 may catalyze both S-oxidation and epoxidation of thiophene derivatives. Biochem Biophys Res Commun 2005;338(1):450–5.
[31] Dalvie DK, Kalgutkar AS, Khojasteh-Bakht SC, Obach RS, O'Donnell JP. Biotransformation reactions of five-membered aromatic heterocyclic rings. Chem Res Toxicol 2002;15(3):269–99.
[32] Mansuy D, Dansette PM. Sulfenic acids as reactive intermediates in xenobiotic metabolism. Arch Biochem Biophys 2011;507(1):174–85.
[33] Kalgutkar AS, Ryder TF, Walker GS, Orr STM, Cabral S, Goosen TC, et al. Reactive metabolite trapping studies on imidazo- and 2-methylimidazo[2,1-b]thiazole-based inverse agonists of the ghrelin receptor. Drug Metab Dispos 2013;41(7):1375–88.
[34] Dansette PM, Levent D, Hessani A, Bertho G, Mansuy D. Thiolactone sulfoxides as new reactive metabolites acting as bis-electrophiles: implication in clopidogrel and prasugrel bioactivation. Chem Res Toxicol 2013;26(5):794–802.
[35] Kassahun K, Pearson PG, Tang W, McIntosh I, Leung K, Elmore C, et al. Studies on the metabolism of troglitazone to reactive intermediates in vitro and in vivo: evidence for novel biotransformation pathways involving quinone methide formation and thiazolidinedione ring scission. Chem Res Toxicol 2001;14(1):62–70.
[36] Gaudiano G, Koch TH. Redox chemistry of anthracycline antitumor drugs and use of captodative radicals as tools for its elucidation and control. Chem Res Toxicol 1991;4(1):2–16.
[37] Garner AP, Paine MJ, Rodriguez-Crespo I, Chinje EC, Ortiz DeMontellano P, Strafford IJ, et al. Nitric oxide synthases catalyze the activation of redox cycling and bioreductive anticancer agents. Cancer Res 1999;59(8):1929–34.
[38] Butler TC. Reduction of carbon tetrachloride in vivo and reduction of carbon tetrachloride and chloroform in vitro by tissues and tissue constituents. J Pharmacol Exp Ther 1961;134:311–19.
[39] Mico BA, Branchflower RV, Pohl LR. Formation of electrophilic chlorine from carbon tetrachloride—Involvement of cytochrome P-450. Biochem Pharmacol 1983;32(15):2357–9.
[40] Gross P. Biologic activity of hydroxylamine: a review. Crit Rev Toxicol 1985;14(1):87–99.
[41] Mason RP. Redox cycling of radical anion metabolites of toxic chemicals and drugs and the Marcus theory of electron transfer. Environ Health Perspect 1990;87:237–43.
[42] Stachulski AV, Meng X. Glucuronides from metabolites to medicines: a survey of the in vivo generation, chemical synthesis, and properties of glucuronides. Nat Prod Rep 2013;30(6):806–48.
[43] Dickinson RG. Iso-glucuronides. Current Drug Metabolism 2011;12(3):222–8.
[44] Grant DM, Hughes NC, Janezic SA, Goodfellow GH, Chen HJ, Gaedigk A, et al. Human acetyltransferase polymorphisms. Mutat Res 1997;376(1–2):61–70.
[45] Timbrell JA, Mitchell JR, Snodgrass WR, Nelson SD. Isoniazid hepatoxicity: the relationship between covalent binding and metabolism in vivo. J Pharmacol Exp Ther 1980;213(2):364–9.
[46] Sinha BK. Activation of hydrazine derivatives to free radicals in the perfused rat liver: A spin-trapping study. Biochim Biophys Acta 1987;924(2):261–9.
[47] Weber GL, Steenwyk RC, Nelson SD, Pearson PG. Identification of N-acetylcysteine conjugates of 1,2-dibromo-3-chloropropane: evidence for cytochrome P450 and glutathione mediated bioactivation pathways. Chem Res Toxicol 1995;8(4):560–73.
[48] D'Souza J, Caldwell J, Smith RL. Species variations in the N-methylation and quaternization of [14C]pyridine. Xenobiotica 1980;10(2):151–7.
[49] Videla LA, Barros SB, Junqueira VB. Lindane-induced liver oxidative stress. Free Radic Biol Med 1990;9(2):169–79.
[50] DeCaprio AP, Strominger NL, Weber P. Neurotoxicity and protein binding of 2,5-hexanedione in the hen. Toxicol Appl Pharmacol 1983;68(2):297–307.
[51] Genter St Clair MB, Amarnath V, Moody MA, Anthony DC, Anderson CW, Graham DG. Pyrrole oxidation and protein cross-linking as necessary steps in the development of gamma-diketone neuropathy. Chem Res Toxicol 1988;1(3):179–85.
[52] Chen LJ, Hecht SS, Peterson LA. Characterization of amino acid and glutathione adducts of cis-2-butene-1,4-dial, a reactive metabolite of furan. Chem Res Toxicol 1997;10(8):866–74.
[53] Baer BR, Rettie AE, Henne KR. Bioactivation of 4-ipomeanol by CYP4B1: adduct characterization and evidence for an enedial intermediate. Chem Res Toxicol 2005;18(5):855–64.
[54] Thomas SH. Paracetamol (acetaminophen) poisoning. Pharmacol Ther 1993;60(1):91–120.
[55] Dahlin DC, Miwa GT, Lu AY, Nelson SD. N-acetyl-p-benzoquinone imine: a cytochrome P-450-mediated oxidation product of acetaminophen. Proc Natl Acad Sci USA 1984;81(5):1327–31.
[56] Chen W, Koenigs LL, Thompson SJ, Peter RM, Rettie AE, Trager WF, et al. Oxidation of acetaminophen to its toxic quinone imine and nontoxic catechol metabolites by baculovirus-expressed and purified human cytochromes P450 2E1 and 2A6. Chem Res Toxicol 1998;11(4):295–301.
[57] Mitchell JR, Jollow DJ, Gillette JR, Brodie BB. Drug metabolism as a cause of drug toxicity. Drug Metab Dispos 1973;1(1):418–23.
[58] Ramakrishna Rao DN, Fischer V, Mason RP. Glutathione and ascorbate reduction of the acetaminophen radical formed by peroxidase: detection of the glutathione disulfide radical anion and the ascorbyl radical. J Biol Chem 1990;265(2):844–7.
[59] Fischer V, West PR, Harman LS, Mason RP. Free-radical metabolites of acetaminophen and a dimethylated derivative. Environ Health Perspect 1985;64:127–37.
[60] Dietze EC, Schafer A, Omichinski JG, Nelson SD. Inactivation of glyceraldehyde-3-phosphate dehydrogenase by a reactive metabolite of acetaminophen and mass spectral characterization of an arylated active site peptide. Chem Res Toxicol 1997;10(10):1097–103.
[61] Zhou L, McKenzie BA, Eccleston Jr ED, et al. The covalent binding of [14C]acetaminophen to mouse hepatic microsomal proteins: the specific binding to calreticulin and the two forms of the thiol:protein disulfide oxidoreductases. Chem Res Toxicol 1996;9(7):1176–82.
[62] Welch KD, Reilly TP, Bourdi M, Hays T, Pise-Masison CA, Radonovich MF, et al. Genomic identification of potential risk factors during acetaminophen-induced liver disease in susceptible and resistant strains of mice. Chem Res Toxicol 2006;19(2):223–33.
[63] Rosen GM, Singletary Jr WV, Rauckman EJ, Killenberg PG. Acetaminophen hepatotoxicity: an alternative mechanism. Biochem Pharmacol 1983;32(13):2053–9.

[64] Holownia A, Mapoles J, Menez JF, Braszko JJ. Acetaminophen metabolism and cytotoxicity in PC12 cells transfected with cytochrome P4502E1. J Mol Med 1997;75(7):522–7.
[65] Dai Y, Cederbaum AI. Cytotoxicity of acetaminophen in human cytochrome P4502E1-transfected HepG2 cells. J Pharmacol Exp Ther 1995;273(3):1497–505.
[66] Nelson SD, Pearson PG. Covalent and noncovalent interactions in acute lethal cell injury caused by chemicals. Annu Rev Pharmacol Toxicol 1990;30:169–95.
[67] Josephy PD. The molecular toxicology of acetaminophen. Drug Metabolism Reviews 2005;37(4):581–94.
[68] Homberg JC, Andre C, Abuaf N. A new anti-liver-kidney microsome antibody (anti-LKM2) in tienilic acid-induced hepatitis. Clin Exp Immunol 1984;55(3):561–70.
[69] Beaune P, Dansette PM, Mansuy D, Kiffel L, Finck M, Amar C, et al. Human anti-endoplasmic reticulum autoantibodies appearing in a drug-induced hepatitis are directed against a human liver cytochrome P-450 that hydroxylates the drug. Proc Natl Acad Sci USA 1987;84(2):551–5.
[70] Bonierbale E, Valadon P, Pons C, Desfosses B, Dansette PM, Mansuy D. Opposite behaviors of reactive metabolites of tienilic acid and its isomer toward liver proteins: use of specific anti-tienilic acid-protein adduct antibodies and the possible relationship with different hepatotoxic effects of the two compounds. Chem Res Toxicol 1999;12(3):286–96.
[71] Lopez-Garcia MP, Dansette PM, Mansuy D. Thiophene derivatives as new mechanism-based inhibitors of cytochromes P-450: inactivation of yeast-expressed human liver cytochrome P-450 2C9 by tienilic acid. Biochemistry 1994;33(1):166–75.
[72] Lecoeur S, Bonierbale E, Challine D, Gautier JC, Valadon P, Dansette PM, et al. Specificity of in vitro covalent binding of tienilic acid metabolites to human liver microsomes in relationship to the type of hepatotoxicity: comparison with two directly hepatotoxic drugs. Chem Res Toxicol 1994;7(3):434–42.
[73] Mansuy D. Molecular structure and hepatotoxicity: compared data about two closely related thiophene compounds. J Hepatol 1997;26 (2 Suppl):22–5.
[74] Lopez-Garcia MP, Dansette PM, Valadon P, Amar C, Beaune PH, Guengerich FP, et al. Human-liver cytochromes P-450 expressed in yeast as tools for reactive-metabolite formation studies: oxidative activation of tienilic acid by cytochromes P-450 2C9 and 2C10. Eur J Biochem 1993;213(1):223–32.
[75] Masubuchi Y, Horie T. Toxicological significance of mechanism-based inactivation of cytochrome p450 enzymes by drugs. Crit Rev Toxicol 2007;37(5):389–412.
[76] Robin MA, Maratrat M, Le Roy M, Le Breton FP, Bonierbale E, Dansette P, et al. Antigenic targets in tienilic acid hepatitis: both cytochrome P450 2C11 and 2C11-tienilic acid adducts are transported to the plasma membrane of rat hepatocytes and recognized by human sera. J Clin Invest 1996;98(6):1471–80.
[77] Neuberger J, Williams R. Immune mechanisms in tienilic acid associated hepatotoxicity. Gut 1989;30(4):515–19.
[78] Valadon P, Dansette PM, Girault JP, Amar C, Mansuy D. Thiophene sulfoxides as reactive metabolites: formation upon microsomal oxidation of a 3-aroylthiophene and fate in the presence of nucleophiles in vitro and in vivo. Chem Res Toxicol 1996;9(8):1403–13.
[79] Rademacher PM, Woods CM, Huang Q, Szklarz GD, Nelson SD. Differential oxidation of two thiophene-containing regioisomers to reactive metabolites by cytochrome P450 2C9. Chem Res Toxicol 2012;25(4):895–903.
[80] Koenigs LL, Peter RM, Hunter AP, Haining RL, Rettie AE, Friedberg T, et al. Electrospray ionization mass spectrometric analysis of intact cytochrome P450: Identification of tienilic acid adducts to P450 2C9. Biochemistry 1999;38(8):2312–19.
[81] Harris JW, Pohl LR, Martin JL, Anders MW. Tissue acylation by the chlorofluorocarbon substitute 2,2-dichloro-1,1,1-trifluoroethane. Proc Natl Acad Sci USA 1991;88(4):1407–10.
[82] Mansuy D, Nastainczyk W, Ullrich V. The mechanism of halothane binding to microsomal cytochrome P450. Naunyn Schmiedebergs Arch Pharmacol 1974;285(4):315–24.
[83] Ahr HJ, King LJ, Nastainczyk W, Ullrich V. The mechanism of reductive dehalogenation of halothane by liver cytochrome P450. Biochem Pharmacol 1982;31(3):383–90.
[84] Pohl LR. An immunochemical approach of identifying and characterizing protein targets of toxic reactive metabolites. Chem Res Toxicol 1993;6(6):786–93.
[85] Kenna JG, Neuberger J, Williams R. Evidence for expression in human liver of halothane-induced neoantigens recognized by antibodies in sera from patients with halothane hepatitis. Hepatology 1988;8(6):1635–41.
[86] Pohl LR, Kenna JG, Satoh H, Christ D, Martin JL. Neoantigens associated with halothane hepatitis. Drug Metab Rev 1989;20 (2–4):203–17.
[87] Hoet P, Graf ML, Bourdi M, Pohl LR, Duray PH, Chen W, et al. Epidemic of liver disease caused by hydrochlorofluorocarbons used as ozone-sparing substitutes of chlorofluorocarbons. Lancet 1997;350(9077):556–9.
[88] Baillie TA. Metabolic activation of valproic acid and drug-mediated hepatotoxicity: role of the terminal olefin, 2-n-propyl-4-pentenoic acid. Chem Res Toxicol 1988;1(4):195–9.
[89] Rettie AE, Rettenmeier AW, Howald WN, Baillie TA. Cytochrome P-450-catalyzed formation of delta 4-VPA, a toxic metabolite of valproic acid. Science 1987;235(4791):890–3.
[90] Rettie AE, Sheffels PR, Korzekwa KR, Gonzalez FJ, Philpot RM, Baillie TA. CYP4 isozyme specificity and the relationship between omega-hydroxylation and terminal desaturation of valproic acid. Biochemistry 1995;34(24):7889–95.
[91] Kassahun K, Baillie TA. Cytochrome P-450-mediated dehydrogenation of 2-n-propyl-2(E)-pentenoic acid, a pharmacologically-active metabolite of valproic acid, in rat liver microsomal preparations. Drug Metab Dispos 1993;21(2):242–8.
[92] Prickett KS, Baillie TA. Metabolism of unsaturated derivatives of valproic acid in rat liver microsomes and destruction of cytochrome P-450. Drug Metab Dispos 1986;14(2):221–9.
[93] Ortiz de Montellano PR, Yost GS, Mico BA, Dinizo SE, Correia MA, Kumbara H. Destruction of cytochrome P-450 by 2-isopropyl-4-pentenamide and methyl 2-isopropyl-4-pentenoate: Mass spectrometric characterization of prosthetic heme adducts and nonparticipation of epoxide metabolites. Arch Biochem Biophys 1979;197(2):524–33.
[94] Rettenmeier AW, Gordon WP, Prickett KS, Levy RH, Baillie TA. Biotransformation and pharmacokinetics in the rhesus monkey of 2-n-propyl-4-pentenoic acid, a toxic metabolite of valproic acid. Drug Metab Dispos 1986;14(4):454–64.

[95] Rettenmeier AW, Prickett KS, Gordon WP, Bjorge SM, Chang SL, Levy RH, et al. Studies on the biotransformation in the perfused rat liver of 2-n-propyl-4-pentenoic acid, a metabolite of the antiepileptic drug valproic acid. Evidence for the formation of chemically reactive intermediates. Drug Metab Dispos 1985;13(1):81–96.
[96] Chang TK, Abbott FS. Oxidative stress as a mechanism of valproic acid-associated hepatotoxicity. Drug Metab Rev 2006;38(4):627–39.
[97] Kawai K, Kawasaki-Tokui Y, Odaka T, Tsuruta F, Kazui M, Iwabuchi H, et al. Disposition and metabolism of the new oral antidiabetic drug troglitazone in rats, mice, and dogs. Arzneimittelforschung 1997;47(4):356–68.
[98] Ramachandran V, Kostrubsky VE, Komoroski BJ, Zhang S, Dorko K, Esplen JE, et al. Troglitazone increases cytochrome P-450 3A protein and activity in primary cultures of human hepatocytes. Drug Metab Dispos 1999;27(10):1194–9.
[99] Prabhu S, Fackett A, Lloyd S, McClellan HA, Terrell CM, Silber PM, et al. Identification of glutathione conjugates of troglitazone in human hepatocytes. Chem Biol Interact 2002;142(1–2):83–97.
[100] Yamazaki H, Shibata A, Suzuki M, Nakajima M, Shimada N, Guengerich FP, et al. Oxidation of troglitazone to a quinone-type metabolite catalyzed by cytochrome P-450 2C8 and P-450 3A4 in human liver microsomes. Drug Metab Dispos 1999;27(11):1260–6.
[101] Smith MT. Mechanisms of troglitazone hepatotoxicity. Chem Res Toxicol 2003;16(6):679–87.
[102] Jaeschke H. Troglitazone hepatotoxicity: are we getting closer to understanding idiosyncratic liver injury? Toxicol Sci 2007;97(1):1–3.
[103] Ong MM, Latchoumycandane C, Boelsterli UA. Troglitazone-induced hepatic necrosis in an animal model of silent genetic mitochondrial abnormalities. Toxicol Sci 2007;97(1):205–13.
[104] Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. Nat Rev Drug Discov 2011;10(4):307–17.
[105] Dansette PM, Libraire J, Bertho G, Mansuy D. Metabolic oxidative cleavage of thioesters: evidence for the formation of sulfenic acid intermediates in the bioactivation of the antithrombotic prodrugs ticlopidine and clopidogrel. Chem Res Toxicol 2009;22(2):369–73.
[106] Dansette PM, Thebault S, Bertho G, Mansuy D. Formation and fate of a sulfenic acid intermediate in the metabolic activation of the antithrombotic prodrug prasugrel. Chem Res Toxicol 2010;23(7):1268–74.
[107] Zhou S, Chan E, Duan W, Huang M, Chen YZ. Drug bioactivation, covalent binding to target proteins, and toxicity relevance. Drug Metab Rev 2005;37(1):41–213.
[108] Yang XX, Hu ZP, Chan SY, Zhou SF. Monitoring drug–protein interaction. Clin Chim Acta 2006;365(1–2):9–29.
[109] Evans DC, Watt AP, Nicoll-Griffith DA, Baillie TA. Drug–protein adducts: an industry perspective on minimizing the potential for drug bioactivation in drug discovery and development. Chem Res Toxicol 2004;17(1):3–16.
[110] Tang C, Subramanian R, Kuo Y, Krymgold S, Lu P, Kuduk SD, et al. Bioactivation of 2,3-diaminopyridine-containing bradykinin B1 receptor antagonists: irreversible binding to liver microsomal proteins and formation of glutathione conjugates. Chem Res Toxicol 2005;18(6):934–45.
[111] Doss GA, Baillie TA. Addressing metabolic activation as an integral component of drug design. Drug Metab Rev 2006;38(4):641–9.
[112] Baillie TA, Cayen MN, Fouda H, Gerson RJ, Green JD, Grossman SJ, et al. Drug metabolites in safety testing. Toxicol Appl Pharmacol 2002;182(3):188–96.
[113] Obach RS, Walsky RL, Venkatakrishnan K. Mechanism-based inactivation of human cytochrome p450 enzymes and the prediction of drug–drug interactions. Drug Metab Dispos 2007;35(2):246–55.
[114] Kalgutkar AS, Obach RS, Maurer TS. Mechanism-based inactivation of cytochrome P450 enzymes: chemical mechanisms, structure–activity relationships, and relationship to clinical drug–drug interactions and idiosyncratic adverse drug reactions. Curr Drug Metab 2007;8(5):407–47.
[115] Guengerich FP, MacDonald JS. Applying mechanisms of chemical toxicity to predict drug safety. Chem Res Toxicol 2007;20(3):344–69.

CHAPTER

26

Drug Transport Mechanisms and their Impact on the Disposition and Effects of Drugs

Jean-Michel Scherrmann

Department of Pharmaceutical Sciences, University Paris Descartes, Paris, France

OUTLINE

Science may set limits to knowledge but should not set limits to imagination. ***Bertrand Russell***

I. INTRODUCTION

The exchange of solutes between body compartments depends to a considerable extent on the properties of the body that allow easy communication between tissues and compartments via pores and fenestra on the walls of the blood vessel or gap junctions between the cells of many epithelia. These features allow free solutes to move in both directions through biological membranes by the paracellular pathway. But the organs of the body and pharmacological targets at the biophase are not readily accessible to exogenous molecules because of the integrity of the lipid bilayer membranes that protect the interiors of cells. Some physiological barriers, like the blood–brain barrier (BBB), the blood–placenta barrier, and the blood–testis barrier, are so impermeable that solutes can only cross the lipid bilayer by a transcellular pathway. It has been established for some time that only small, nonpolar, uncharged molecules like oxygen, carbon dioxide, water, and ethanol can easily diffuse through membranes if there is an appropriate gradient, whereas charged small ions like sodium and potassium and molecules like glucose (180 Da) cross membranes considerably less readily than water. As the delivery of many polar molecules—such as anions and cations, vitamins, sugars, nucleosides, amino acids, peptides, bile

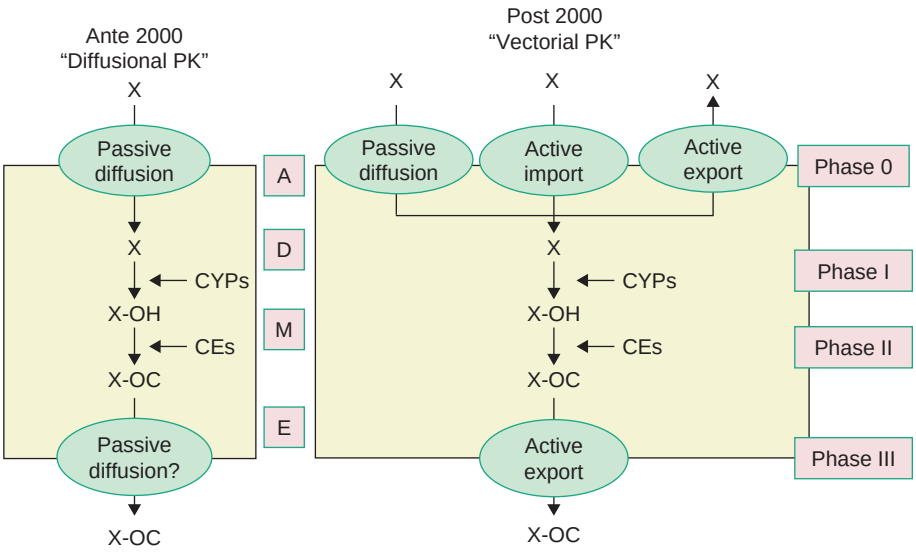


FIGURE 26.1 Schematic representations of the A, D, M, E processes of drugs and xenobiotics in virtual biological systems (e.g., body, organ, cell) illustrating the fate of a drug where membrane permeation is either mediated by passive diffusion (diffusional pharmacokinetics) or by a combination of passive diffusion and active transport processes (vectorial pharmacokinetics). The coordinated activities between transporters and metabolizing enzymes has led to the identification of these sequential events as Phases 0, I, II, and III (X = drug or xenobiotic; CYPs = cytochromes P-450; CEs = conjugation enzymes).

acids, and porphyrins—to cells is essential for life, essential transporter proteins anchored in the lipid bilayer have evolved to permit their exchange between cells and their environment.

Pharmacokinetics is now challenged by the growing importance of transporters, a relatively new and potentially major factor in drug absorption, distribution, metabolism, and excretion (the ADME process). Several years ago, passive diffusion was the main advanced process by which xenobiotics were believed to move through body membranes. The recent intrusion of drug transporters means that there is no single mechanism by which drugs penetrate the membranes. The presence of transporters at membranes that facilitate the movement of solutes into cells (influx or import transporters) and of transporters that remove substances from the cytosol of cells (efflux or export transporters) modulates the traditional theory of "diffusional pharmacokinetics" toward "vectorial pharmacokinetics" in which ADME processes are more deterministically governed (Figure 26.1).

We can now prepare a fairly complete list of drug transporters, the tissues in which they occur and function, how they are regulated or mutated, and the clinical relevance of their presence in normal and diseased tissues.

II. BIOLOGY AND FUNCTION OF TRANSPORTERS

A. Modes of Active Transport

Several types of transporters have been identified; they differ in their energy source and the direction of transport. The primary active transport systems are coupled to an energy source, like the hydrolysis of adenosine triphosphate (ATP) by ion pumps (ATPases), and the ABC transporters move their substrate in one specific direction. Movement is independent of the solute-concentration gradient. They are primary transporters because no additional biochemical step is needed for solute transport. The second group of transporters are the co-transporters. These use a voltage and/or ion gradient to transport both ions and solutes together. They are uniporters when only one species is transported, symporters when both species are transported in the same direction, and antiporters when transporting solutes and ions in opposite directions. The H^+ ion is the most common form of energy in prokaryotes, while Na^+ is more frequently encountered in eukaryotic cells. Other sources of energy involve HCO_3^-, glutathione (GSH), and the dicarboxylate α-ketoglutarate (α-KG). Some of these only transport solutes in a direction defined by the solute concentration gradient by the facilitative diffusion pathway. The main role of the Na^+, K^+-ATPase system is to activate the cascades of multiple co-transport processes. These co-transporters are also called secondary or even tertiary transporters, because the machinery of ion transport must be activated by one or two pumps before solute transport occurs (Figure 26.2).

B. Genes and Classification

Following the determination of the prokaryotic and eukaryotic genomes, it was predicted that 15 percent of the 23,000 genes in the human genome code for transport proteins—nearly 3,500 transporters. They are clustered

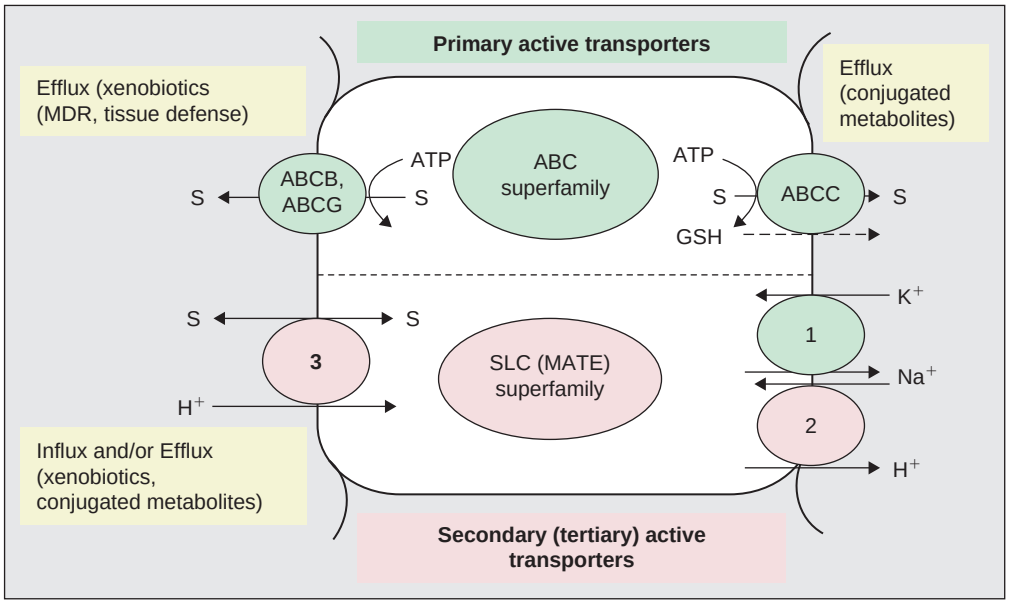


FIGURE 26.2 ABC transporters (green) that transport the substrate (S) in one defined direction are called "primary transporters," because no additional biochemical step other than the ATP hydrolysis or GSH co-transport is needed for (S) transport. SLC transporters (**3**; pink) need to be activated by one or two ion transporters before S transport occurs. In this model, **1** is the Na^+, K^+-ATPase and **2** is the Na^+, H^+ antiporter, providing the H^+ driving force for S transport by **3**.

in several superfamilies, and only the members of three superfamilies are presently known to affect drug transport. They are the ABCs, the SLCs, and the multidrug and toxin extrusion (MATE) transporters. ABC proteins are widespread in all organisms from bacteria to mammals, with about 600 referenced transporters. Only forty-eight genes have been identified in humans, and no more than around nine ABCs have been shown to affect drug pharmacokinetics and pharmacodynamics [1]. The SLC family, which may have about 2,000 members, is presently known to have forty-six families, including 475 transporter genes with documented transport functions [2]. The MATE emerged very recently, and only two proteins (MATE1 and MATE2) are presently known to efflux drugs in mammals, whereas 861 related sequences are found in the three kingdoms of living organisms (Eukarya, Archaea, and Eubacteria) [3]. The transporters have not yet been completely assigned to superfamilies, families, and subfamilies. The Human Genome Organisation (HUGO) Nomenclature Committee Database has provided a list of genes and defined the families of these transporters (see http://www.gene.ulc.ac.uk/nomenclature/). MATE1 and MATE2 were recently ranked by the HUGO as members of the SLC family (SLC47) and not as an independent superfamily. This review uses the HUGO as the primary reference for identifying genes and proteins. Human proteins (genes) are shown in capitals (e.g., ABCB1 (*MDR1*)), while rat and mouse proteins (genes) are indicated by an initial capital followed by small letters (proteins) and small letters (genes) (e.g., ABCC1 (*mdr1*)).

C. Basic Structure

Transporters are integral membrane proteins that typically have twelve transmembrane domains (TMDs), although some have six, eight, ten, eleven, thirteen, or even seventeen TMDs. The TMDs are folded in α-helical structures within the membrane and linked at both sides by amino acid sequences floating in the internal or external cell environment. The amino acids in the external loop domains are frequently *N*-glycosylated, while those of the intracellular loops of SLC, ABC, and MATE proteins bear phosphorylation sites. The ABCs also have one or two ATP-binding domains. The 3D structure of TMDs is a crown shape, and they look like a channel allowing communication between the two fluid spaces separated by the lipid bilayer (Figure 26.3). Many SLC and MATE transporters have 300−800 amino acid residues and a molecular mass of 40−90 kDa, while the ABC transporters are larger, with 1,200−1,500 residues and a mass of 140−190 kDa. The amino acid sequence can be used to classify transporters by family and subfamily according to the degree of amino acid homology. For example, a transporter protein is assigned to a specific SLC family if it has an amino acid sequence that is at least

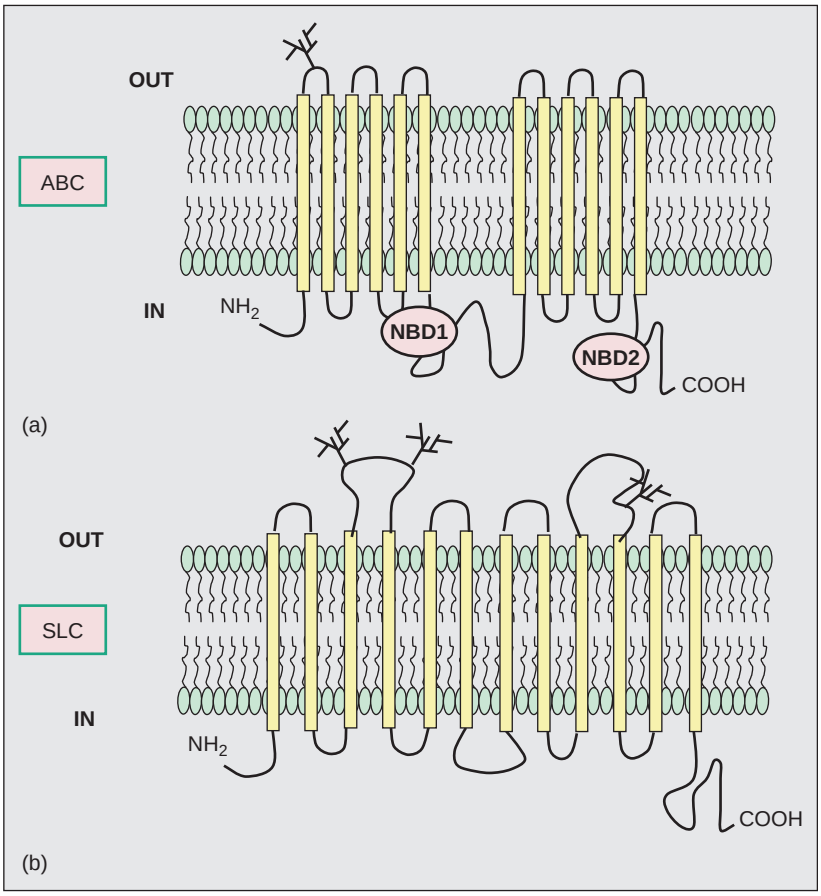


FIGURE 26.3 Topological models inserted in the lipid bilayer membrane of the best-characterized ABC and SLC transporters that transport drugs and toxicants. (A) Model of MDR1 (P-gp) resembles MRP4, 5, 8 and (B) model of the human *SLCO*, the organic anion polypeptide transporter (OATP). Features common to all members of the OCT, OAT, OATP transporter family include twelve transmembrane spanning domains (TMDs; represented by yellow rectangular bars) with intracellular N and C termini. Features common to all members of the ABC include the nucleotide binding domain (NBD) *N*-glycosylation sites (indicated by branches) are present on extracellular protein loops. Cytoplasmic (IN) and extracellular (OUT) orientations are indicated.

20–25 percent identical to those of other members of that family. A new nomenclature system was recently proposed that is based on the classification of drug-metabolizing enzymes. The transport proteins of a superfamily are arranged in clusters of families (≥40 percent identity) and subfamilies (≥60 percent identity). Amino acid sequences are also extremely helpful for assessing the effect of a single mutation within the sequence that can change the conformation of the transport protein and alter its transport functions.

D. Distributions and Properties of Transporters in Tissues

1. Cellular and Subcellular Distribution

There are about 200 types of cells in human tissues, and all their plasma membranes and the membranes of their organelles contain transporters. The drug transporters at the organelles may well become most important in the future. This was recently documented during a dramatic Phase II trial in which the nucleoside antiviral fialuridine (FIAU) caused the death of subjects as a result of severe toxicity, including hepatotoxicity, pancreatis, neuropathy, or myopathy. These toxic events were clearly linked to mitochondrial damage due to the transport of FIAU into mitochondria by an SLC transporter in the human mitochondrial membrane. Very little attention has been paid to the processes regulating transport across the endoplasmic reticulum (ER) membrane, although they are required for the activities of intraluminal UDP-glucuronosyltransferases (UGTs). It is now evidenced that the presence of multiple ER glucuronide transporters of different specificities in the ER membrane allow the import of the aglycone substrate and UDP-glucuronic acid and the exit of the conjugated end products, which are impermeable, bulky, polar, and charged molecules, to lipid bilayers (Figure 26.4). The size and shape of the aglycone are critical determinants of transporter specificity, rather than the glucuronic acid moiety and hydrophobicity. The fact that glucuronide transport in the ER membrane is independent of ATP and GSH suggests that the translocation is mainly mediated by several SLC transporters. The presence of transporters in mitochondria, the ER, and other constituents of the cell cytosol opens the possibility that the intracellular kinetic trafficking of xenobiotics and their metabolites may be mediated by active transport processes.

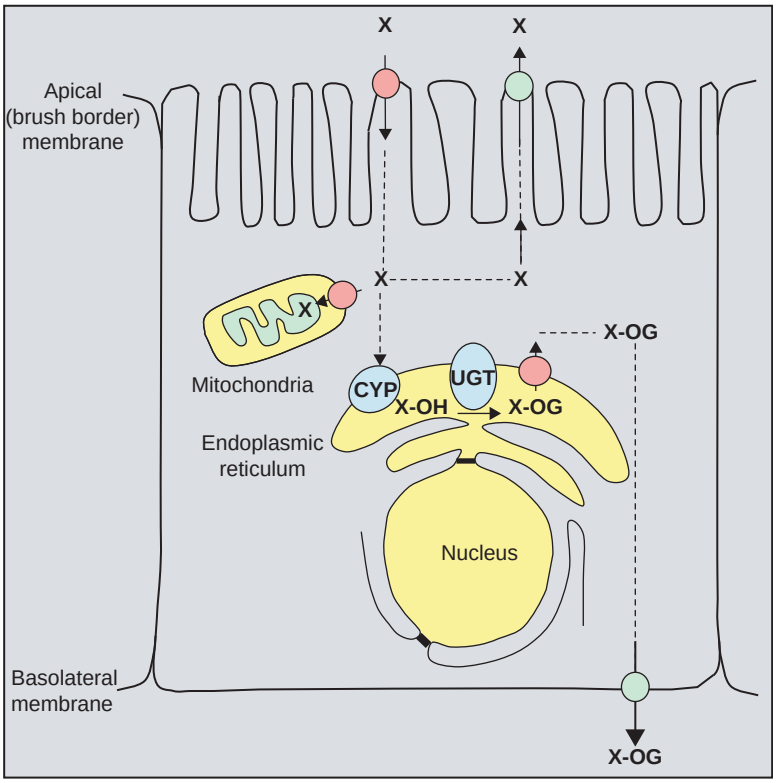


FIGURE 26.4 A schematic proposal for cellular and subcellular distribution of ABC (green) and SLC (pink) transporters on the apical and basolateral membranes of an epithelial cell. SLC import and ABC export the unchanged xenobiotic (X) at the apical membrane, and ABC export the glucuronated-X (X-OG) at the basolateral membrane. Subcellular events include mitochondrial uptake of X by a SLC transporter and X-OG efflux from the lumen of the ER where cytochrome P-450 (CYP) and UDP-glucuronosyltransferase (UGT) produce X-OH and X-OG metabolites.

2. *Polarized Expression of Transporters in Epithelia and Vectorial Transport*

The location of transporters at the cell plasma membrane is a critical issue because most of the cells involved in the A, D, and E pharmacokinetic processes are polarized. Hence, their apical (luminal) and basolateral (abluminal) membranes do not have the same populations of transporters (Figure 26.4). The same transporter is rarely found at both the apical and basolateral membranes, but most of the ABC and SLC transporters are located at either the apical or the basolateral epithelial membranes. Their location helps to define the direction of substrate transport and the resulting pharmacokinetic event.

For example, some SLC on the sinusoidal (basolateral) membrane of hepatocytes take up organic anions, while the ABC on the apical membranes of bile canicular cells excrete them. The combined activities of these two transporters thus results in the vectorial transport of drugs from the blood to the bile. Similarly, the basolateral transporters of the kidney tubular cells act in a coordinated vectorial manner with apical transporters to secrete organic cations (OCs) from the blood to the urine.

3. *Coordination between Transporters and Metabolic Enzymes*

Drug metabolism was considered to be one of the main processes for removing xenobiotics prior to the emergence of transporters. The cytochrome P-450 (CYP) isoenzymes catalyze the first step of biotransformation. This function was called Phase I metabolism, while the subsequent conjugation step was called Phase II metabolism. We now know that these two phases occur in specialized cells like the hepatocytes and enterocytes and that they are preceded and followed by two other phases controlled by transporters. Efflux or influx transporters reduce or increase the uptake of substrates, and these actions help to regulate the amounts of a xenobiotic reaching the enzyme binding sites or the rate at which the metabolites produced are eliminated. The first step has been called "Phase 0" and the second "Phase III," indicating a close relationship between transporters and enzymes (Figures 26.1 and 26.4). They provide the cell with a suite of processes that may operate in parallel and in series. This integrated biological function of combined transport and metabolic processes is strongly supported by the presence of common regulation pathways that act via similar nuclear receptors, such as PXR, RXR, and others, to induce or repress the genes encoding enzymes and transporters.

4. Polyspecific Transport and Inhibition

The substrate specificities of transporters are often very broad, as indicated by the many overlaps of substrates and inhibitors, much like the specificity of the drug metabolism enzymes. Thus, probenecid was initially known to produce many drug interactions by blocking the secretion of many drugs by the kidney, including the penicillins and the antiviral Tamiflu. Probenecid is today known to be a polyspecific inhibitor of several ABCs and SLCs. Thus, all ionized chemicals, peptides, and nucleosides that cannot diffuse freely across membranes are very likely to interact with one or more transporters.

5. Transport Kinetics and Variability

As each transporter's capacity is limited in the same way that the metabolism enzymes, transport kinetics can be saturated by substrate concentrations greater than its K_m. The K_m of transporters can vary from nM to mM values, and the risk of saturating a transport will depend on the amount of substrate in the transporter environment. Transport can also be inhibited in a competitive or noncompetitive manner, in the same way as the drug-metabolizing enzymes, so that transporters can promote drug–drug interactions that were initially thought to be due to the drug-metabolizing enzymes alone. *In vitro* transporter assays are increasingly being used to assess the potential risks of drug–drug interactions mediated by transporters. The *in vitro* inhibition constant (K_i) can be measured and used to predict changes in the clearance or systemic exposure by measuring the area under the curve. Transport kinetics may also depend on the amount of transporter, which will depend on the actions of drugs, nutrients, and disease states on the nuclear receptor pathways mentioned above. The most recent area of variation concerns the presence of genetic polymorphisms. However, studies on the pharmacogenetics of most drug transporters have only recently begun.

III. TRANSPORTERS IN DRUG DISPOSITION

This section covers only those transporters that influence the A, D, or E of drugs and xenobiotics. About forty transporters belonging to the ABC and SLC superfamilies are presently known to influence the pharmacokinetics, pharmacodynamics, and toxicity of drugs and xenobiotics. They are often classified according to the chemical nature of their substrates. Hence, they translocate organic anions or cations, peptides, or nucleosides. Most of them were first named according to their specific chemical substrate, such as the organic cation transporters (OCT) or the organic anion transporters (OAT), before they were named using the HUGO nomenclature rules.

A. ABC Transporters

Most of the ABC transporters in eukaryotes move compounds from the cytoplasm to the outside of cells. This means that they are frequently called efflux pumps. Phylogenetic analysis has grouped the eukaryotic ABC genes into seven subfamilies (A–G). Only three of these subfamilies—B, C, and G—contain transporters that influence drug disposition [4].

1. ABCB Subfamily

ABCB1, also called P-glycoprotein (P-gp), is a 170-kDa protein that was the first human ABC transporter cloned. It is responsible for the MDR phenomenon that occurs with such anticancer agents as the anthracyclines, vinca alkaloids, and taxanes. P-gp is the product of two MDR genes in humans, MDR1 and MDR2 (also called MDR3). Only the MDR1 protein is involved in the MDR phenotype. Two genes, Mdr1a and Mdr1b, result in a similar MDR phenotype in rodents. P-gp is present mainly on the apical membrane of many secretory cells, including those of the intestine, liver, kidney, and adrenal gland. In the placenta, P-gp is found on the apical surface of syncytiotrophoblasts, where it can protect the fetus from toxic xenobiotics. P-gp is also abundant on hematopoietic stem cells, where it may protect the cells from toxins, and on the luminal surface of endothelial cells forming physiological barriers like the blood–testis barrier and the BBB. P-gp transports not only antineoplastic agents but also a wide variety of structurally dissimilar substrates. They are mostly hydrophobic compounds that are either neutral or positively charged and are presented to the transporter binding sites directly from the lipid bilayer. The range of substrates that P-gp recognizes overlaps those of the main drug-metabolizing enzymes CYP3A4/5. These enzymes are known to metabolize about 50 percent of the drugs currently on the market. Immunosuppressive agents (cyclosporin A and its analog PSC833), cardiac glycosides (digoxin), protease inhibitors (saquinavir, indinavir), antibiotics (rifampicin), calcium channel blockers (verapamil), and quinoline (quinidine) have all been found to interact with P-gp as both substrates and inhibitors.

2. ABCC subfamily

Although P-gp is considered to be the major transporter responsible for drug export at the plasma membrane of many cells, MDR phenotypes that are not P-gp-mediated have been linked to several MRPs or ABCCs. At least five of them, MRP1, MRP2, MRP3, MRP4, MRP5, and recently MRP8, are likely to be involved in mediating drug resistance and affecting drug pharmacokinetics. Although several molecules are substrates of P-gp when the unconjugated cationic (vincristine, doxorubicin) and neutral (etoposide) compounds are transported, these MRPs preferentially transport anions (like many Phase II metabolites of drugs) conjugated to GSH, glucuronate, or sulfate. The MRP4 and MRP5 proteins mainly confer resistance to cyclic nucleosides and purine analogs. They transport substrates by a different mechanism from P-gp. There may even be multiple mechanisms that include co-transport with GSH. All these isoforms are concentrated on specific areas of polarized cells, like the epithelial cells of the gut and kidney, and probably also on the endothelial cells of brain microvessels. MRP2, MRP4, and MRP8 are, like P-gp, found in the apical (luminal) membrane, while MRP1, MRP3, MRP4, and MRP5 are found in the basolateral (abluminal) membrane.

3. ABCG Subfamily

There are presently four known human members of the G subfamily: ABCG1, ABCG2, ABCG5, and ABCG8. Three of them (ABCG1, ABCG5, and ABCG8) are all implicated in lipid transport. ABCG2 is important for drug resistance and drug disposition. It was cloned independently by three different groups and called BCRP, mitoxantrone-resistance protein (MXR), and placenta-specific ABC protein (ABCP) before it was designated ABCG2. This second member of the G subfamily confers resistance to anticancer agents like mitoxantrone, topotecan, irinotecan, and flavopiridol, but not to paclitaxel, cisplatin, or vinca alkaloids. BCRP also actively transports structurally diverse organic molecules, both conjugated and unconjugated, such as SN38, the metabolite of irinotecan and its glucuronide conjugate SN38-G, the estrone-3-sulfate, the 17β-E2G, DHEAS, and the organic anions like methotrexate. Other BCRP substrates include tyrosine kinase inhibitors like imatinib mesylate (Gleevec®), which may also be a potent inhibitor of BCRP, the nucleotide reverse transcriptase inhibitors like zidovudine (AZT), its active metabolite AZT5′-monophosphate, lamivudine (3TC), and the proton pump inhibitor pantoprazole. BCRP can transport chemical toxins such as pheophorbide a (a breakdown product of dietary chlorophyll that is phototoxic) and the small heterocyclic amine carcinogen PhIP that causes mammary and prostate cancers.

BCRP inhibitors include the fungal toxin derivative fumitremorgin C (FTC) and several dietary flavonoids. BCRP lies primarily in the plasma membrane and at the apical membrane of polarized epithelia, as does P-gp. High concentrations of BCRP are found in the placental syncytiotrophoblasts, the apical membrane of the epithelium of the small intestine, the membranes of liver canaliculi, and at the luminal surface of the endothelial cells of the brain microvessels that form the BBB. Thus, ABCG2 is found mainly in organs that are important for absorption (proximal part of the small intestine), distribution (placenta and the BBB), and elimination (liver, kidney, and small intestine). BCRP has been found in stem cells where it protects them from cytotoxic substrates. BCRP was also recently shown to secrete drugs or toxins into milk. BCRP lies in the apical membrane of the mammary glands' alveolar epithelial cells, at the main site of milk production. The milk-to-plasma ratios of several drugs, such as acyclovir, cimetidine, and nitrofurantoin, were found to be high even before they were known to be BCRP substrates. The secretion of xenobiotics into milk by BCRP is puzzling, because this function exposes the suckling infant to a range of drugs and toxins.

B. SLC Transporters

The system for naming members of the SLC superfamily differs somewhat from the ABC nomenclature. The genes are usually named using the root symbol *SLC*, followed by a number corresponding to the family (e.g., *SLC22*, solute carrier family 22), the letter A, and finally the number of the individual transporter (e.g., *SLC22A2*), but there may be differences between families. The *SLC21* family encoding the organic anion-transporting (OATP) proteins has been reclassified as a superfamily with families and subfamilies much like the classification of drug-metabolizing enzymes. The gene symbol then becomes *SLCO* (i.e., the "21" and the "A" have been replaced by the letter "O" for organic transporter) and the "OATP" symbol has been kept for protein nomenclature (e.g., *SLCO1A2* for the gene and OATP1A2 for the protein).

1. OATP (SLC21/SLCO) Transporters

A total of eleven human OATPs have been identified to date. The OATPs were originally identified as uptake transporters, although some may function primarily in efflux [5]. The driving force for OATP-mediated transport is

still not clear, but it is independent of ATP or sodium gradients. There is experimental evidence that bidirectional transmembrane transport can be mediated by anion (HCO_3^-) or GSH exchange. It has now been shown that many OATPs are polyspecific OAT with partially overlapping substrate specificities for a wide range of solutes, including bile salts, the organic dye bromosulfopthalein (BSP), steroid conjugates (DHEAS, 17βE2G estrone-3-sulfate (E-3-S)), thyroid hormones, neuroactive peptides ((D-penicillamine 2,5), enkephalin (DPDPE), Leu-enkephalin, and deltorphin II), and numerous drugs and toxins, such as the cardiotonic digoxin, the angiotensin-converting enzyme inhibitors enalapril and temocaprilat, and the 3-hydroxy-3methylglutaryl coenzyme A (HMG-CoA) reductase inhibitor pravastatin. OATP substrates are mainly high molecular weight (>450 Da) amphiphatic molecules, mostly bound to albumin, that have a steroid nucleus or linear and cyclic peptides. Most OATPs, mainly those of the OATP1 family, are found in many tissues and are thought to be part of the body's detoxification system, helping to remove xenobiotics from the systemic circulation (e.g., drug uptake into hepatocytes). The rat Oatp1a1, Oatp1a4, and Oatp1b2, and human OATP1B3, OATP1B1, and OATP2B1 are all found in the sinusoid membrane of hepatocytes, where they are responsible for the uptake of xenobiotics for hepatic clearance. The hepatic OATPs may play a strategic role in drug–drug interactions and hepatotoxicity. For example, rifampicin is a potent inhibitor of both OATP transporters and CYP3A4. Thus, giving rifampicin with OATP substrates may reduce hepatic first-pass clearance, increase the bioavailability of an intrahepatically active drug like pravastatin, and decrease its efficacy. On the other hand, induction of OATP gene expression could increase the hepatic uptake and the total body clearance of the substrate.

2. OCT (SLC22) Transporters

The OCTs include three potential-sensitive proteins (OCT1, OCT2, and OCT3) and three H^+-driven transporters of carnitine and/or cations (OCTN1, OCTN2, and CT2, also known as *SLC22A4*, *SLC22A5*, and either *FLIPT2* or *SLC22A16*, respectively [6]. Both OCT1 and OCT2 are found primarily in the major excretory organs (kidney and liver) and to a smaller extent in the intestine and the brain, while OCT3 is much more widely distributed.

All three OCTs recognize a variety of OCs, including endogenous bioactive amines like acetylcholine, choline, epinephrine, norepinephrine, dopamine, and serotonin, and drugs like cimetidine, quinine, quinidine, prazosin, desipramine, verapamil, and morphine. The nitrogen moiety of the weak bases bears a net positive charge at physiological pH, allowing them to interact electrostatically with the binding sites of the OCTs. The "type 1" and "type 2" classifications of OCs were developed to study their uptake by the liver. Type 1 OCs are small (60–350 Da) monovalent hydrophilic compounds such as tetraethylammonium (TEA) and the parkinsonian neurotoxin 1-methyl-4-phenylpyridinium (MPP^+). In contrast, type 2 OCs are usually bulkier (>500 Da; e.g., anthracyclines) and polyvalent (e.g., d-tubocurarine). This classification helps to differentiate the mechanisms by which they are transported across polarized cells. Type 2 OCs are believed to diffuse across the basolateral membrane and to be exported across the apical membrane by MDR1. In contrast, the basolateral entry of type 1 OCs involves one or more transporters, including OCT1, OCT2, and OCT3, and their efflux at the apex may be mediated by OCTN1 and OCTN2 or MATE1 and MATE2. The OCTs generally mediate the bidirectional transport of substrate molecules. This depends mainly on the membrane potential and not directly on the transmembrane gradients of Na^+ or H^+. Unlike the OCTs, which have a common energy-supply mechanism, the OCTNs differ markedly in their mode of action. OCTN1 supports electroneutral OC/H^+ exchange, OCTN2 supports both Na^+-dependent co-transport (e.g., carnitine) and electrogenic-facilitated diffusion (e.g., TEA and type 1 OCs), and OCTN3 mediates the electrogenic transport of carnitine. OCTN3 and CT2 are present only in the testes of mice and humans, where transported carnitine improves sperm quality and fertility. OCTN1 is most abundant in the kidney (at the apical membrane of tubule cells), small intestine, bone marrow, and fetal liver, but not in the adult liver. OCTN2 is mainly found in the heart, placenta, skeletal muscle, kidney, and pancreas. Both OCTN1 and OCTN2 have a low affinity for MPP^+, cimetidine, and TEA, and OCTN2 plays a major role in carnitine homeostasis.

3. OAT (SLC22) Transporters

As their name implies, small organic anions (300–500 Da) possess a net negative charge at physiological pH, and their transepithelial transport into the negatively charged environment of the cell requires energy. The OATs (*SLC22* family) are found mainly in cells playing a critical role in the excretion and detoxification of xenobiotics. There are six members of the OAT family (OAT1, OAT2, OAT3, OAT4, OAT5, and URAT1), present mainly in the liver, kidney, placenta, brain capillaries, and choroid plexus [7].

The OAT proteins play a critical role in the excretion and detoxification of a wide variety of drugs, toxins, hormones, and neurotransmitter metabolites. A number of common nonsteroid anti-inflammatory drugs (NSAID)—including acetyl salicylate and salicylate, acetaminophen, diclofenac, ibuprofen, ketoprofen, indomethacin, and naproxen—are substrates of one or more OAT isoforms, so that there can be significant interactions between

NSAIDs and other drugs. The β-lactam antibiotics (penicillins, cephalosporins, and penems) and the antiviral nucleosides (adefovir, cidofovir, acyclovir, and AZT) are also substrates of one or more OAT isoforms and are actively excreted in the urine. Toxins like chlorinated phenoxyacetic acid herbicides, mercuric conjugates, cadmium, and ochratoxin A are also transported either into the renal tubule cells or hepatocytes via the OAT network, and this predisposes these tissues to nephrotoxicity or hepatotoxicity.

4. PEPT1 (SLC15A1) and PEPT2 (SLC15A2) Transporters

PEPT1 and PEPT2 translocate dipeptides and tripeptides and their pharmacological importance is due to their ability to transport a wide variety of peptide-mimetic drugs, such as β-lactam antibiotics of the cephalosporin and penicillin classes and drugs like captopril, enalapril, and fosinopril. Other drugs include the dopamine D2 receptor antagonist sulpiride and the peptidase inhibitor bestatin [8]. PHT1 and PHT2 transport histidine and certain di- and tripeptides, but their location on the cell or lysosomal membranes remains as questionable as their implication in pharmacotherapy. PEPT1 is the low-affinity (mM range) high-capacity transporter that is mainly found in the apical membranes of enterocytes in the small intestine, in renal proximal tubule cells of the S1 segment, and in bile duct epithelial cells. In contrast, PEPT2 is a high-affinity (μM range) low-capacity transporter that is more widely distributed in the apical membranes of kidney tubule cells of the S2 and S3 segments, brain astrocytes, and epithelial cells of the choroid plexus. They are involved in the uptake of their substrates, leaving a basal transporter (s) to account for the exit. This basal transporter could be PHT1 and/or PHT2, or the amino acid transporters of the *SLC1* and *SLC7* families. Both PEPT1 and PEPT2 can mediate the renal reabsorption of the filtered compounds in kidney tubules, whereas PEPT2 may be responsible for the removal of brain-derived peptide substrates from the cerebrospinal fluid via the choroid plexus. The pharmaceutical relevance of these peptide transporters is closely linked to the design of drug-delivery strategies mediated by the intestinal PEPT1. One successful approach has been to produce peptide derivatives of parent compounds as substrates for PEPT1. The pharmacophoric pattern for the transporter includes the rules that the peptide bond is not a prerequisite for a substrate and that 5′-amino acid esterification, mostly using L-valine or L-alanine, markedly improves recognition by PEPT1. This prodrug strategy was used to improve the bioavailability of oral enalapril from 3–12 percent to 60–70 percent for the ester enalaprilat, which resembles the structure of the tripeptide Phe-Ala-Pro. The oral bioavailability of the nucleoside antiviral acyclovir (22 percent) was similarly improved by adding a valine residue to give valacyclovir (70 percent). Current studies on the regulation of PEPT1 and PEPT2 synthesis in inflammatory intestinal diseases may provide helpful information on the variations in bioavailability of oral PEPT1 drug substrates.

5. CNT (SLC28) and ENT (SLC29) Transporters

The members of the human *SLC28* and *SLC29* families mainly catalyze the transport of purine and pyrimidine nucleosides [9]. Hydrophilic nucleosides like the purine adenosine are important signaling molecules that control both neurotransmission and cardiovascular activity. They are also precursors of nucleotides, the constitutive elements of DNA and RNA, and are the basic elements of a variety of antineoplastic and antiviral drugs. The *SLC28* proteins in the apical membranes of polarized cells work in tandem with the *SLC29* proteins found in the basolateral membrane. The *SLC28* family consists of three dependent concentrative nucleoside transporters that differ in their substrate specificities. CNT1 transports naturally occurring pyrimidine nucleosides plus the purine adenosine. Several antiviral analogs, like AZT, lamivudine (3TC), and ddC, are substrates of CNT1. The cytotoxic cytidine analogs cytarabine (AraC) and gemcitabine (dFdc) are also transported by CNT1. CNT1 is primarily found at the apical membrane of epithelial cells, including those of the small intestine, kidney, and liver. Human CNT2 is widely distributed in the kidney, liver, heart, brain, intestine, skeletal muscles, pancreas, and placenta. CNT2 transports purine nucleosides and uridine. Pharmaceutical substrates include the antiviral didanosine (ddI) and ribavirin.

Like CNT2, human CNT3 has a wide tissue distribution with high concentrations in the pancreas, bone marrow, and mammary glands. CNT3 is broadly selective and transports both purine and pyrimidine nucleosides in a 2:1 Na^+ nucleoside coupling ratio—in contrast to the 1:1 ratio employed by CNT1 and CNT2. CNT3 transports several anticancer nucleoside analogs including cladrabine, dFdc, fludarabine, and zebularine.

The human *SLC29* family has four members. ENT1 and ENT2 can both transport adenosine but differ in their abilities to transport other nucleosides and nucleobases. ENT1 is almost ubiquitous in human and rodent tissues and transports purine and pyrimidine nucleosides with K_m values of from 50 μM (adenoside) to 680 μM (cytidine). The antiviral drugs ddC and ddI are also poorly transported. ENT2 is present in a wide range of tissues including the brain, heart, pancreas, prostate, and kidney, and is particularly abundant in skeletal muscle. ENT2 differs from ENT1 in that it can also transport nucleobases like hypoxanthine and AZT. ENTs also mediate the uptake and efflux of several nucleoside drugs because of their bidirectional transport property.

6. MATE Transporters

The MATE transporters are involved in transport of many organic cations, via a H^+ or Na^+-coupled antiport mechanism. In humans, two genes encode MATE1 and MATE2. MATE1 is ubiquitous throughout the body but is most abundant in the luminal membrane of the urinary tubules and bile canaliculi in the liver. By contrast, MATE2 is specific to the kidney brush border membranes. Both MATEs are responsible for the final step in the excretion by the kidney and liver of metabolic waste and xenobiotic OCs with very diverse chemical structures.

MATE1 was shown to transport TEA and 1-methyl-4-phenyl pyridinium (MPP^+). MATE2 also transports multiple OCs including TEA, MPP^+, cimetidine, and metformin [3].

IV. ROLES OF TRANSPORTERS IN DRUG PHARMACOKINETICS, PHARMACODYNAMICS AND TOXICOLOGY

Transporters are now recognized to be as important as the metabolizing enzymes in the modulation of the main steps controlling the fate and action of xenobiotics in the body. They affect all the main pharmacokinetic events like the oral bioavailability, distribution, and clearance of substrates. They are presently known to modulate the active drug concentration in all biophases and influence the effects of drugs.

A. Intestinal Absorption

Both influx and efflux transporters modulating drug absorption are present in the epithelium of the various segments of the intestine (Figure 26.5) [10]. PEPT1, OATP1A2, OATP2B1, OATP3A1, and OATP4A1 are all found on the apical membrane and mostly import substrates from the lumen into the circulation. PEPT1 is the best-characterized drug transporter in the small intestine of mammals and is widely used to improve the absorption of poorly absorbed oral drugs using a prodrug strategy. ABC transporters, including MDR1, MRP2, and BCRP, are also present on the apical membrane, where they either limit the intestinal uptake of their substrates or contribute to the active secretion of drugs from the blood to the intestinal lumen. For example, the antineoplastic agent paclitaxel, a P-gp substrate, is poorly absorbed when taken orally by humans (only 5 percent is bioavailable), but when it is administered with the P-gp competitor cyclosporin A, its bioavailability is increased to 50 percent. The roles of basolateral transporters are much less well known. Oct1 is present on the basolateral sides of cells, and studies using Oct1 knockout mice indicate that Oct1 is important for the secretion of OCs into the lumen of the small intestine.

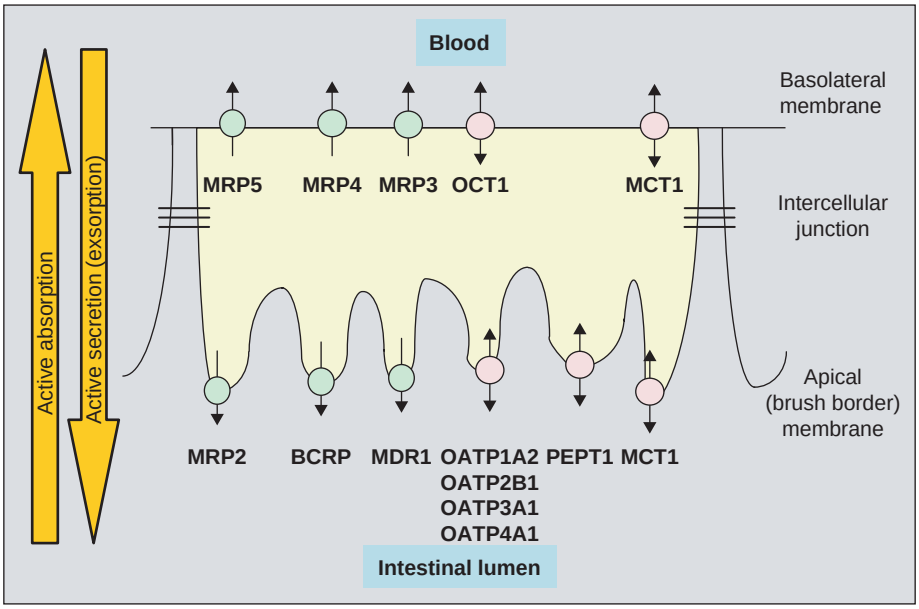


FIGURE 26.5 Distribution of the main drug ABC (green) and SLC (pink) transporters on the apical and basolateral membranes of the human intestinal enterocytes. All apical transporters (except MRP1) lie at the top of the villi. They define active absorption and/or secretion of drugs and xenobiotics through the intestinal epithelium.

The intestinal transporters are not uniformly distributed along the crypt–villus axis. Many of those implicated in the absorption of drugs, like PEPT1, MDR1, BCRP, MRP2, and MRP3, are villus-specific. This restriction of transporters to the villus is also correlated with the presence of CYP3A in intestinal cells, suggesting coordinated Phase 0 and I activities of MDR1 and CYP3A4 in the intestinal first-pass effect. A major concern is the way their densities vary along the gastrointestinal tract. For example, MRP3 is the most abundant ABC protein throughout the human intestine, except at the terminal ileum where MDR1 is most abundant. Similarly, the concentration of MDR1 increases from the duodenum to the colon, whereas BCRP is found throughout the small intestine and colon. MRP2 is most prevalent in the duodenum and becomes undetectable toward the terminal ileum and colon. These diverse distributions of the intestinal transporters may have dramatic pharmaceutical consequences. The pharmaceutical form of an oral drug can vary from a simple solution to a solid controlled-release complex, and this can influence the gastrointestinal site (stomach, duodenum, jejunum, ileum, or colon) at which the active compound is released. Such differences may also influence the efficacy of the carrier-mediated transports, as these may vary from one region of the intestine to another. The great risk of saturating active transport is that it can also affect the kinetics of drug absorption. This can occur when a large amount of drug is rapidly dissolved in the intestinal lumen, ready to be absorbed by a relatively small area of intestine. Active transport can be saturated by a relatively high concentration of substrate, shifting absorption toward diffusion. Here, too, the properties of the oral preparation, like its rate of dissolution, may influence the contribution of active transport to drug absorption.

B. Liver and Hepatic Clearance

Hepatic clearance is a combination of metabolic (Phase I and II) and biliary clearance. As previously indicated, hepatocytes can take up drugs by diffusion or active transport (Phase 0; Figure 26.6) [11]. The basolateral membrane OAT include OAT2, OAT4, OATP1A2, -1B1, -1B3, and -2B1, the OCT1, and the Na-taurocholate co-transporting polypeptide NTCP (SLC10A1). They are responsible for the uptake by the liver of a wide variety of drugs because of their broad, overlapping substrate specificities. Phase III, which follows Phases 0, I, and II, results in the elimination of the intact drug and/or metabolite(s) via efflux transporters on the apical and basolateral membranes. The hepato-biliary transporters include several ABC proteins (MDR1, MDR3, MRP2, BSEP, BCRP, and MATE1) that are the main mediators of the excretion of numerous endogenous conjugated and unconjugated bile salts and drugs via the bile. Phase III also includes the efflux of compounds from hepatocytes back into the systemic circulation via basolateral membrane efflux transporters. Some of the OATPs, OATs, and OCT1 are bidirectional and may facilitate efflux, but the main exporters are the ABC proteins, which transport a wide range of glucuronides and sulfated and GSH conjugates. The main ones are MRP1 and MRP3, whose synthesis is readily induced, and the cyclic nucleoside transporters MRP4 and MRP5.

This huge network of hepato-biliary transporters can give rise to variations in drug disposition between individuals by modulating the uptake or the exit of drugs and their metabolites from hepatocytes. A change in hepatic uptake may have clinical consequences. It may modulate the pharmacological activity of drugs that act via the intrahepatocellular transduction pathways, cause hepatotoxicity, or give rise to drug–drug interactions. The concentration of the cholesterol-lowering HMG-CoA inhibitors in hepatocytes must be adequate for their

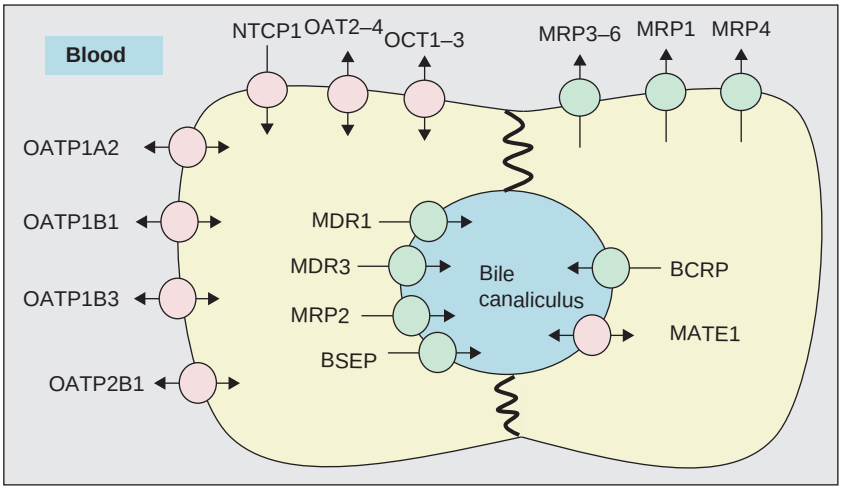


FIGURE 26.6 Distribution of the main drug ABC (green) and SLC (pink) transporters on the basolateral (sinusoid) and apical (bile canaliculus) membranes of the human hepatocytes. SLC transporters at the basolateral membrane mainly define active hepatic uptake, whereas ABC transporters at the basolateral and bile canaliculus membranes efflux drugs and their metabolites in blood or bile, respectively.

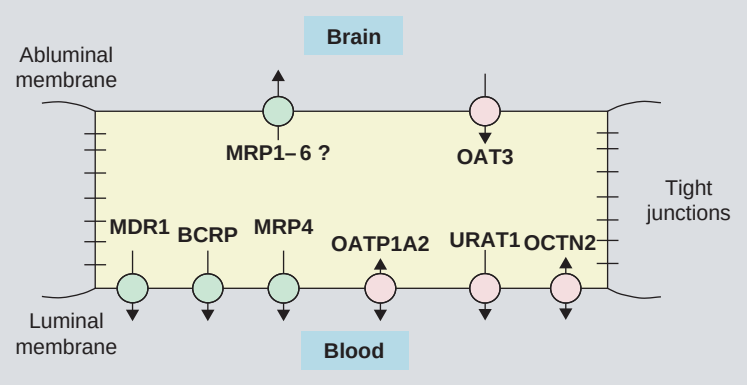


FIGURE 26.7 Distribution of the main ABC (green) and SLC (pink) transporters on the abluminal (facing brain extracellular fluid) and luminal membranes of the brain microvessel endothelial cells constituting the BBB.

pharmacological activity, and most of the statins—like pravastatin, simvastatin, lovastatin, cerivastatin, and pitavastatin—enter hepatocytes via OATP1B1 and to a lesser degree via OAT1B3. Recently identified genetic polymorphisms like the SLCO1B3 haplotype *17 are associated with reduced uptake of statins by the liver and lower concentrations in hepatocytes. They thus have less effect on cholesterol synthesis. Large-scale clinical studies are needed to confirm the impact of OATP1B1 polymorphisms on the considerable variation between individuals to therapy with hypolipidemic agents. Transporters can also mediate hepatotoxicity. For example, the sulfate conjugate of the antidiabetic troglitazone can cause troglitazone hepatotoxicity by inhibiting OATP1B1 and OATP1B3. These hepatic impacts of the basolateral transporters have their counterpart at the apical pole. The multiple ABC transporters may also be responsible for variable drug disposition. For example, giving patients on digoxin the P-gp inhibitor verapamil decreases the biliary clearance of digoxin by 43 percent and increases its plasma concentration by 44 percent.

C. Blood Barriers and Tissue Distribution

The tissue distribution of a drug can be affected by transporters because they lie on the luminal or abluminal membranes of the endothelial cells of the tissue blood vessels or on the membranes of the specific cells of the underlying organ. The transporters on the membranes of the blood vessels may be several key physiological components of blood barriers throughout the human body if tight junctions seal adjacent cells and prevent the paracellular exchange of solutes. In contrast, solutes can freely communicate between extracellular spaces when blood vessels are fenestrated, as in the liver sinusoids, and transporters on the plasma membranes of the tissue cells (e.g., the hepatocyte membranes) become the first barrier regulating the import and export of solutes. Several organs, including the brain, nose, eyes, testes, prostate, and placenta, are protected by endothelial barriers that contain extensive networks of transporters. Figure 26.7 illustrates the luminal and abluminal distributions of several transporters at the BBB [12]. The two ABC proteins MDR1 and BCRP are most abundant on the luminal side of the endothelial cells and are most important for protecting the brain from numerous xenobiotics. Few SLCs have been characterized in the rat BBB, except for the important network of SLC transporters that allows the blood–brain exchange of amino acids and sugars. Rat Oatp1a4 is found on both the luminal and abluminal membranes of the brain capillaries. The human isoform OATP1A2 is also present, but its membrane location has not been determined. Both OATPs can mediate uptake or efflux transport because of their bidirectional transport characteristic. The members of the SLC22, OAT3, OCTN2, and URAT1 have been found in the BBB. OAT3 is abluminal and effluxes benzylpenicillin, cimetidine, PAH, and several acidic metabolites of neurotransmitters from the brain to the blood. The luminal position of URAT1 enables this vectorial transport of the OAT3 substrates. OCTN2, which is believed to be luminal, can simultaneously transport carnitine into the brain and efflux OCs from the brain to the blood. Here, too, drug transporters on the membranes of physiological barriers or on specific membranes of the tissue cells can affect drug distribution and consequently the fraction of the drug available for binding to intracellular receptors or other biological targets.

D. Kidney and Renal Clearance

The presence of a drug in the urine is the net result of filtration, secretion, and reabsorption. Filtration occurs by passive glomerular filtration of unbound plasma solutes, whereas secretion and reabsorption are generally

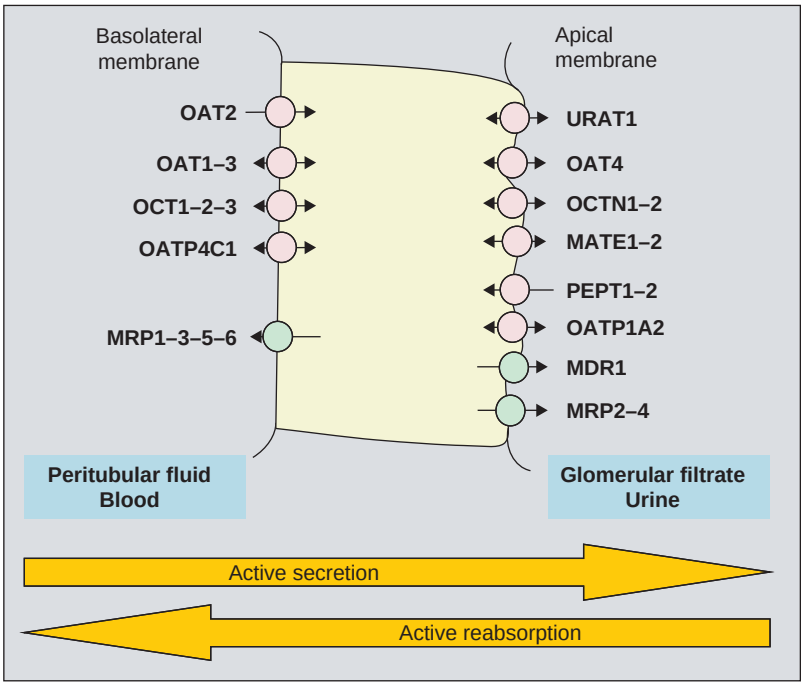


FIGURE 26.8 Distribution of the main drug ABC (green) and SLC (pink) transporters on the basolateral (peritubular fluid) and apical (glomerular filtrate) membranes of the kidney proximal tubule cells. Active secretion and reabsorption help to define the overall renal clearance of drugs and xenobiotics.

carrier mediated. They can occur in the proximal tubule, which has three segments (S1, S2, and S3), the loop of Henle, the distal tubul,e and the collecting tubule. These specific anatomical and functional regions of the kidney must be carefully considered, just like the regions of the intestine, because region-specific distributions of transporters define their action in renal clearance. Most renal transporters lie on the apical and basolateral membranes of the proximal tubule cells, with fewer on the epithelial membrane of the other components of the nephron. The resulting vectorial transport from the peritubular capillaries to the tubule lumen or vice versa can produce either secretion or reabsorption. Figure 26.8 shows the locations of the major drug transporters in the human proximal renal tubule cells [13]. Multiple SLC transporters at their basolateral membrane (close to peritubular capillaries) mediate drug uptake into the tubule cells. Although bidirectional by nature, the direction of the transmembrane-driving gradients favors tubular uptake rather than the efflux of organic anions and cations. Organic anions enter these cells via OAT1, OAT2, and OAT3, and probably via OATP1A2 and OATP4C1, which was recently identified and transports digoxin and methotrexate. OCs are similarly transported by OCT1, OCT2, and OCT3. The efflux transporters MRP1, MRP3, MRP5, and MRP6 mediate their efflux back into the systemic circulation. At the apical membrane, OAT4, URAT1, OCTN1, and OCTN2 can mediate drug transport with bidirectional properties, either secretion or reabsorption. For example, OCTN2 secretes OCs and reabsorbs zwitterions. OATP1A2, PEPT1, and PEPT2 mediate the reabsorption of their substrates from the tubule lumen.

The ABC transporters MDR1, BCRP MRP2, and MRP4 are also present on the apical membrane and efflux compounds by secretion. As indicated above, transporters are not evenly distributed along the nephron. MDR1, MRP2, MRP4, and MRP6 are found mainly within the three segments of the proximal tubule, MRP3 lies in the distal convoluted tubule, and MRP1 is found in the epithelial cells of the loop of Henle and the distal and collecting duct tubule cells, but not in proximal tubule cells. The regional distributions of the SLC transporters are also specific. OAT1 is found only on the basolateral membrane of the S2 segment cells of the proximal tubule, whereas OAT3 is present on the cells of the S1, S2, and S3 segments. This transporter network can be responsible for drug–drug interactions, nephrotoxicity, and drug efficacy mediated by the reabsorptive and secretory capacities of the kidney. If the renal clearance of a drug is equal to or more than the overall body clearance, renal transporters can be important in clinical efficacy or toxicity. For example, the cephalosporin antibiotics are primarily eliminated via the kidney. Creatinine clearance is normally 100–140 mL/min, but the renal clearance of cephalosporins is 16.8–469 mL/min, suggesting that some of them, like cefotaxine and cefadroxil, are excreted into the urine by tubular secretion, whereas others, like ceftriaxone and cefazodin, whose renal clearance is less than that of creatinine, are reabsorbed. OAT1, OAT2, and OAT3 are located on the basolateral side of the proximal tubule and mediate the uptake of most of the cephalosporins into the proximal tubule from the peritubular capillary.

The apical OAT4 mediates both the uptake (reabsorption from the tubular lumen) and the efflux (secretion) of these anionic antibiotics. Like the basolateral transporters of hepatocytes, which can modulate the pharmacological activity of drugs acting via intrahepatocyte targets or induce hepatotoxicity, the basolateral OATs can make some cephalosporins cause nephrotoxicity, which may lead to acute proximal tubular necrosis. This toxicity is mainly due to the accumulation of cephalosporin in the renal cortex because of the lack of efficient vectorial transtubular transport. This transport-mediated nephrotoxicity also results in the adverse effect of cisplatin and related drugs via their basolateral uptake in the proximal tubule by OCT2. The toxic effects depend on the platinium complex used, as does the structure-dependent nephrotoxicity of cephalosporins. Nephrotoxicity also limits the use of the nucleoside phosphonates adefovir and cidofovir in the treatment of human immunodeficiency virus. The toxicity of these drugs appears to be a function of both OAT1-mediated proximal tubular accumulation and decreased efflux at the luminal membrane by MRP2. A small dose of the OAT1 inhibitor probenecid may reduce the nephrotoxicity of cidofovir. The use of transporter inhibitors to reduce nephrotoxicity suggests that drug–drug interactions affecting anionic and cationic drugs can be mediated via competition at the basolateral and luminal tubular transporters. Multiple drug–drug interactions have been reported with probenecid and cimetidine, and there have been fatal cases with methotrexate and NSAID following the inhibition of the basolateral OAT1 and OAT3. Finally, renal transporters can be critical for the action of diuretics. Tubular secretion is the main route by which diuretics act in the kidney and are excreted. Diuretic drugs like the thiazides, the loop diuretics bumetanide and furosemide, and the carbonic anhydrase inhibitors are all competitive inhibitors of the renal OATs, although their affinities and specificities vary.

V. CONCLUSION

The recent expansion of information on drug transporters in pharmacokinetics has added a new layer of complexity to our understanding of the mechanisms underlying the absorption, distribution, and elimination of drugs. New transporters undoubtedly remain to be identified at the plasma membranes of both cells and organelles. Nevertheless, it is clear that drug transporters are significant determinants of variations in drug responsiveness between individuals, drug–drug interactions, drug-induced organ toxicities, and diseases. Detailed knowledge of genetic polymorphisms in transporters and how they affect transporter function will help to optimize drug therapies and identify unknown residual factors that influence subject-to-subject variations [14].

Transporters are now an integral part of the drug-discovery and development processes. They are attractive markers in the creation of drugs that are readily absorbed and accurately targeted. The incorporation of transport properties into structure–activity models should help medicinal chemists design more efficient, safer new medicines.

Acknowledgments

The author is most grateful to Ms. Elisabeth Niel for providing illustrations and technical support, and to Dr. Owen Parkes for reviewing the English text.

References

[1] Dassa E, Bouige P. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. Res Microbiol 2001;152:211–29.
[2] Hediger MA, Romero MF, Peng JB, Rolfs A, Takanaga H, Bruford EA. The ABCs of solute carriers: physiological, pathological, and therapeutic implications of human membrane transport proteins: Introduction. Pflug Arch 2004;447:465–8.
[3] Omote H, Hiasa M, Matsumoto T, Otsuka M, Moriyama Y. The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. Trends Pharmacol Sci 2006;27:587–93.
[4] Leslie EM, Deeley RG, Cole SP. Multidrug resistance proteins: role of P-glycoprotein, MRP1, MRP2, and BCRP (ABCG2) in tissue defense. Toxicol Appl Pharmacol 2005;204:216–37.
[5] Hagenbuch B, Meier PJ. Organic anion transporting polypeptides of the OATP/ SLC21 family: phylogenetic classification as OATP/ SLCO superfamily, new nomenclature and molecular/functional properties. Pflug Arch 2004;447:653–65.
[6] Jonker JW, Schinkel AH. Pharmacological and physiological functions of the polyspecific organic cation transporters: OCT1, 2, and 3 (SLC22A1–3). J Pharmacol Exp Ther 2004;308:2–9.
[7] Sweet DH. Organic anion transporter (Slc22a) family members as mediators of toxicity. Toxicol Appl Pharmacol 2005;204:198–215.
[8] Daniel H, Rubio-Aliaga L. An update on renal peptide transporters. Am J Physiol Renal Physiol 2003;284:F885–92.

[9] Pastor-Anglada M, Cano-Soldado P, Molina-Arcas M, Lostao MP, Larrayoz I, Martinez-Picado J, et al. Cell entry and export of nucleoside analogues. Virus Res 2005;107:151–64.
[10] Takano M, Yumoto R, Murakami T. Expression and function of efflux drug transporters in the intestine. Pharmacol Ther 2006;109: 137–61.
[11] Chandra P, Brouwer KL. The complexities of hepatic drug transport: current knowledge and emerging concepts. Pharm Res 2004;21: 719–35.
[12] Ohtsuki H. New aspects of the blood–brain barrier transporters: its physiological roles in the central nervous system. Biol Pharm Bull 2004;27:1489–96.
[13] Shitara Y, Sato H, Sugiyama Y. Evaluation of drug–drug interaction in the hepatobiliary and renal transport of drugs. Annu Rev Pharmacol Toxicol 2005;45:689–723.
[14] Ho RH, Kim RB. Transporters and drug therapy: implications for drug disposition and disease. Clin Pharmacol Ther 2005;78:260–77.

CHAPTER

27

Strategies for Enhancing Oral Bioavailability and Brain Penetration

Gerhard Gross

Scientific Director at Janssen in Beerse, Belgium

OUTLINE

I. INTRODUCTION

Oral administration is still the preferred way to administer a drug. About 60 percent of all marketed drugs are oral drugs [1]. This is the case for a variety of reasons, among them cost and compliance of patients. For some drugs, more than 90 percent is lost due to lack of significant absorption or presystemic metabolism [2,3]. Poor bioavailability is typically accompanied by high exposure variability, which is a leading cause of failure in clinical trials.

Before being bioavailable, a drug has to cross a number of intestinal membranes and "escape" metabolizing enzymes in the gut and in the liver.

As is evident from Figure 27.1, a number of barriers and hurdles must be passed before a drug enters systemic circulation and reaches its target. An excellent review by Fasinu et al [4] summarizes the attempts to improve bioavailability in recent years. In general terms, a strategy could address all four stages and optimize the relevant parameters. It is important to understand that many of these parameters are working in conflicting directions. For example, higher lipophilicity typically facilitates passive diffusion through the lipid bilayer membrane of the intestinal endothelial cells but also increases presystemic degradation by metabolism.

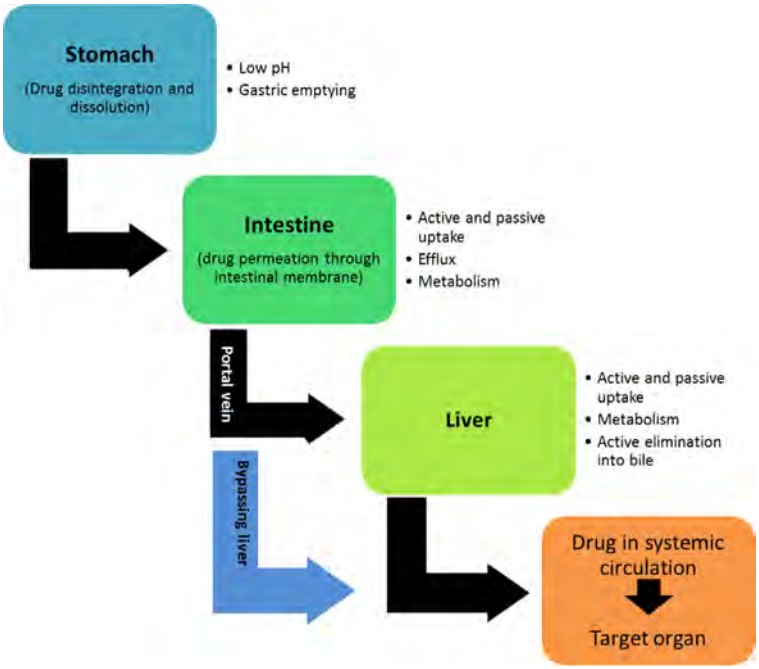


FIGURE 27.1 Barriers for oral bioavailability.

II. ENHANCING ORAL BIOAVAILABILITY

A. Strategic Options and Chemical Space

As outlined during the introduction, there are a number of strategic options on how to improve bioavailability. In principle, it starts with some generic rules like the rule-of-5 (Ro5) [5] and *in silico* calculations to produce drug candidates within a certain chemical space. This is typically followed by *in vitro* data on solubility, permeability, and metabolic stability, followed by *in vivo* studies. Many times the permeability investigation includes information about efflux transporters, passive diffusion, and sometimes active transport mechanisms in the absorption process. The Biopharmaceutics Classification System (BCS) and the newer Biopharmaceutical Drug Disposition Classification System (BDDCS) (Figure 27.2) [6] allows to get an easy orientation, what processes are dominating overall disposition.

FIGURE 27.2 BDDCS system.

TABLE 27.1 Influencing Factors for Oral Bioavialiability

Influenced parameters	Increase	decrease
Fa (fraction absorbed)		Increased molecular weight esp. >500
		Polarity (cLogD , −2)
		PSA (>125 A2)
		Total H-bond donors and acceptors (>9)
		Rotatable bonds (>12)
Fg (fraction escaping gut metabolism)	Most impacted by lipophilicity	
	$cLogD_{pH7.4}$ >3	
Fh (fraction escaping liver metabolism)	Most impacted by lipophilicity	
	$cLogD_{pH7.4}$ >3	

Critical factors influencing gastrointestinal (GI) tract permeability are polar surface area (PSA), ionization state, lipophilicity, molecular weight, solubility, energy required for dissolving into water, and the number of rotatable bonds [7].

In terms of pharmacokinetics, absolute oral bioavailability is defined as:

$$Absolute\ bioavailability = \frac{AUC_{oral}}{AUC_{iv}} \times \frac{Dose_{iv}}{Dose_{oral}}$$

Bioavailability is a function of the fractions of intestinal absorption (Fa), fractions escaping intestinal metabolism (Fg), and hepatic metabolism (Fh):

$$F = Fa \times Fg \times Fh$$

This again demonstrates that the overall process is a combination of several influencing factors that need to be balanced to achieve oral bioavailability.

A recent study [7] based on data for 309 drugs, analyzed the impact of key physicochemical parameters for optimum bioavailability and the impact on fractions absorbed and fractions metabolized (Table 27.1).

Table 27.1 illustrates that intestinal absorption and presystemic metabolism work in the opposite direction. Higher lipophilicity enhances permeability but also increases metabolic degradation.

A study investigating the physicochem parameters for oral bioavailability in rats [8] using a dataset of 1,100 drug candidates came to a similar conclusion. Reduced molecular flexibility (as measured by the number of rotatable bonds) and low PSA or total hydrogen bond count (sum of donors and acceptors) were found to be important predictors of good oral bioavailability, independent of molecular weight.

This establishes the need to identify the issue impacting bioavailability to enable optimization to eliminate or to circumvent the issue.

The chemical space typically associated with oral absorption was outlined by Lipinski's rule-of-5 [5] (i.e., mw <500, a log P <5, H-bond donors <5, and <10 hydrogen bond acceptors).

Figure 27.3 outlines the general ideas behind approaches to enable or enhance oral bioavailability. In principle, they depend on the current phase of a project (e.g., throughput needed) and the strategic decisions being made. The themes are:

- Move the molecules into a chemical space that is generally accepted as being suitable for oral absorption.
- Identify the issue and address by modifying the structure.
- Exploit the possibility of formulation modifications.

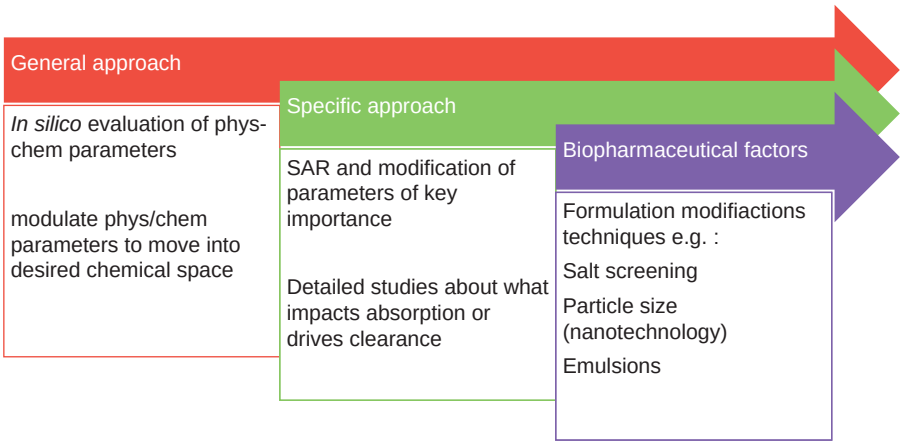


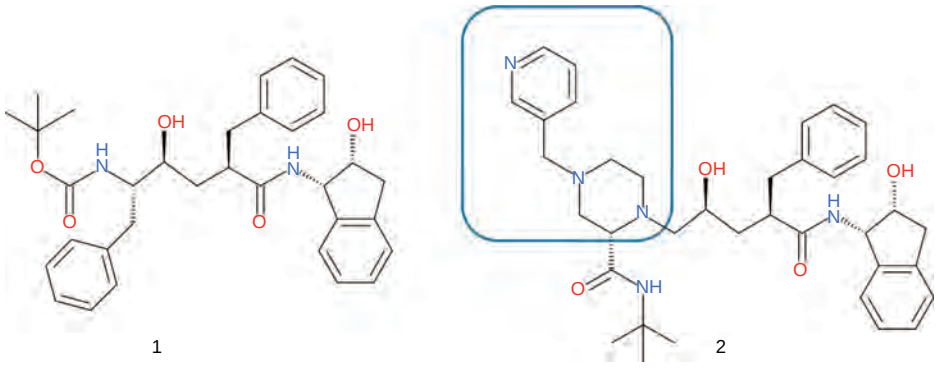
FIGURE 27.3 Themes to improve oral bioavailability.

B. Solubility

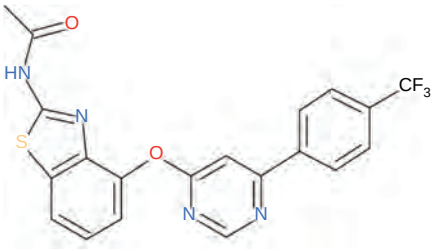
The solubility needed to support a specific human dose is dependent on permeability and the projected dose itself [9,10]. As the human dose is difficult to predict, discovery teams target a kinetic solubility around 60 ug/ml (pH 6.5) [11].

Yalkowsky introduced general solubility equations [12] that include not only logP but also the melting point of a compound. The melting point is indicative of the crystal packing. The equation shows that for solubility in water, the crystallinity of the solute is important, as is the ability to interact with water via ionizable groups. Typically solubility can be increased by adding ionizable groups (e.g., pyridine or carboxylic acids) [13,14].

A key step during the development of Indinavir from compound **1** (L-6854340) was the introduction of the ionizable groups: a piperazine moiety and the pyridine group. These changes enhanced solubility and potency of **2** [15].

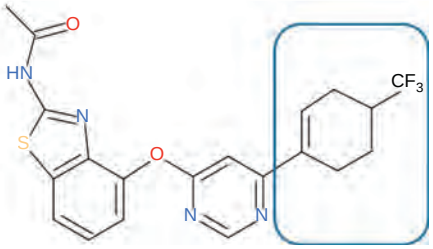


Because crystal packing is influenced by symmetry and planarity, one could expect disruption of symmetry or planarity to disturb crystal packing, thereby lowering the melting point and increasing solubility. The following example shows the impact of removing aromaticity [16,17].



3

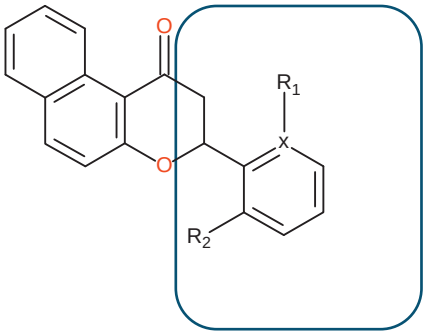
Solubility (0.01 m HCl): <0.1 μg/ml
ClopP: 4.6
Melting point: 219–222 C



4

Solubility (0.01 m HCl): 13 μg/ml
ClogP: 3.7
Melting point: 130–131 C

The next example illustrates how the thermodynamic aqueous solubility has been modified by increasing the dihedral angle of the bicyclic structure [18].



5

Compound	R1	R2	X	Calc dihedral angle	Slubility (ug/ml)	Melting point
5 a	H	H	C	17.8	84, 6	165–167
5 b	H	Me	C	37.9	262	135–137
5 c	Me	Me	C	70	1270	92
5 d	F	H	C	9.1	153	157
5 e	F	F	C	40.5	248	150
5 f	OMe	H	C	18.5	45.8	192–193
5 g		H	N	0	299	187–188

These data strongly suggest that compound 5c, which shows the highest disruption of planarity, exhibits a lower crystal packing energy, indicated by the lower melting point. This, in fact, caused higher solubility.

Compound 5 g shows a higher solubility due to its reduced lipophilicty and higher ionizability.

C. Permeability

A general rule of thumb is that a logP of 2–3 provides a good balance between solubility and permeability [19,20]. Intestinal permeability can be influenced by several distinct processes. For passive diffusion, three approaches are common to increase absorption:

- modification of Mw and/or logP if not in optimum range (Mw: 300–400 Da and logP 2–3);
- reduction of hydrogen bond functionality; and
- reduction of ionization.

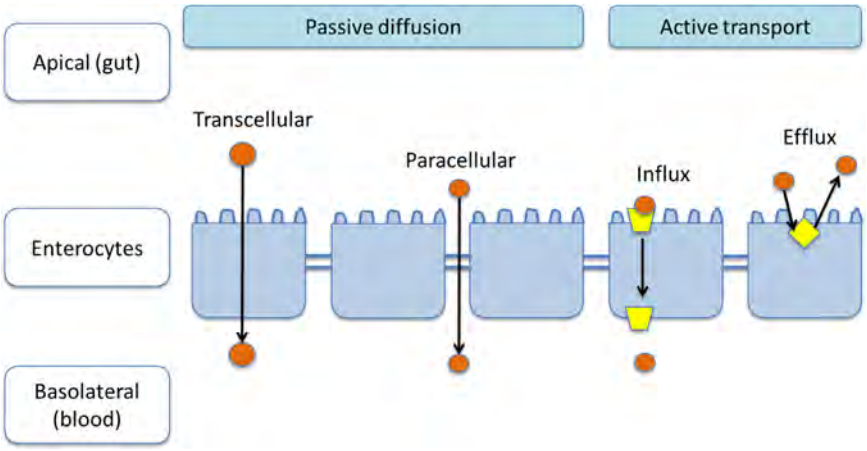
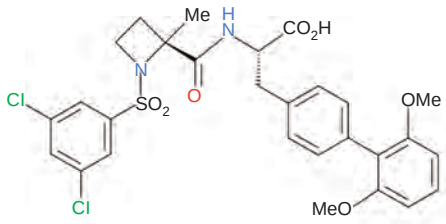


FIGURE 27.4 Principal mechanisms of intestinal permeability.

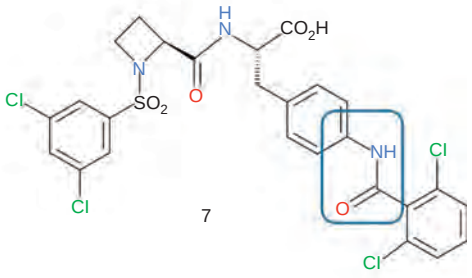
The two main routes how a drug can cross the intestine are via passive diffusion and/or active transport (Figure 27.4) [21]. To optimize a compound structure, it is therefore desirable to know how the compound is absorbed. The type of pathway is determined by the physicochemical properties of the compound. The paracellular pathway is typically a minor pathway due to the small surface area. The transcellular pathway by passive diffusion is the most common pathway. Efflux by P-gp or BCRP might play a key role for some compounds (e.g., Docetaxel [22] and Andrographolide [23]).

For the transcellular pathway, a strong correlation exists between lipophilicity and absorption.

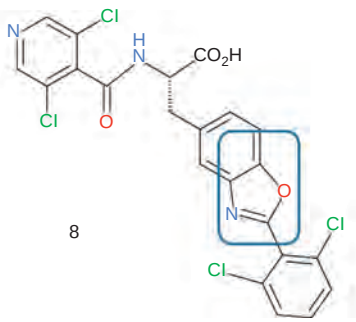
For VLA4 antagonists [24] both lead compounds **6** and **7** showed good pharmacological activity (0.07 and 0.27 nM for the Mn^{2+}activated states and 0.27 and 0.31 nM for the Ca^{2+}/Mg^{2+} unactivated states, respectively). Unfortunately, the PK properties of both compounds were poor. Compound **6** was somewhat absorbed but rapidly cleared, while compound **7** was not absorbed. Bioisosteric replacement of the anilide moiety with benzoxazole led to compound **8**, which showed significant bioavailability.



6

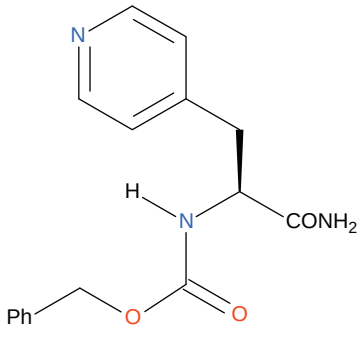


7



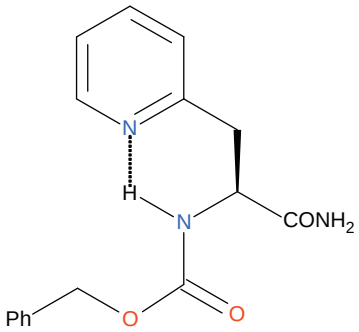
8

Masking the polar groups by complexation is also a potential way to improve permeability, as shown with compounds **9** and **10** [25].



9

Papp (A to B) = 43 nm/sec



10

Papp (A to B) = 177 nm/sec

The introduction of basic and/or hydrophilic atoms generally improves solubility and reduces metabolism, but these improvements come at the expense of membrane permeability. Conversely, masking hydrogen bond donor (HBD) functions by N- or O-alkylation increases lipophilicity and often improves permeability. This can affect binding affinity to the target or decrease solubility.

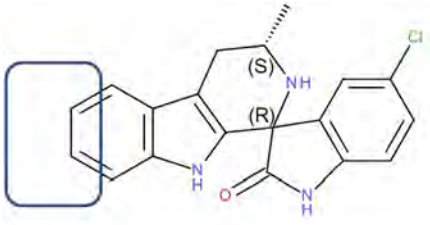
An attractive alternative to any permanent change of HBD functionality is the introduction of a complementary hydrogen bond acceptor (HBA) functionality, thus enabling complexation.

Through the formation of intramolecular hydrogen bonds, the HBD and HBA groups are shielded, thereby reducing the energetic penalty of desolvation required in moving from an aqueous environment to a lipophilic membrane interior [25].

D. Metabolic Stability

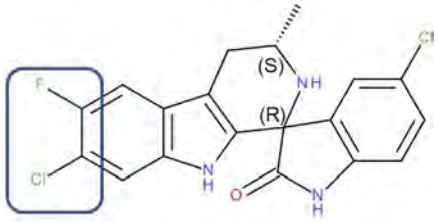
Metabolic stability is a key element of any screening approach. Typically it is investigated in liver microsomes. Hepatocytes provide a clearer picture, as they express fully functional phase II enzymes likes UGTs and cytosolic enzymes like aldehyde oxidase. It is important to emphasize that there are pronounced species differences in drug metabolism. Therefore, multiple species are required to estimate metabolism and exposure in humans.

A typical approach would be to elucidate the soft spots in a molecule that are metabolized and to block them. This approach was followed in the spiroindolone series [26] where halogen substituents were introduced to hinder oxidation.



11

NF54 EC50 = 9 nM
mouse microsomes t1/2= 1.8 min.

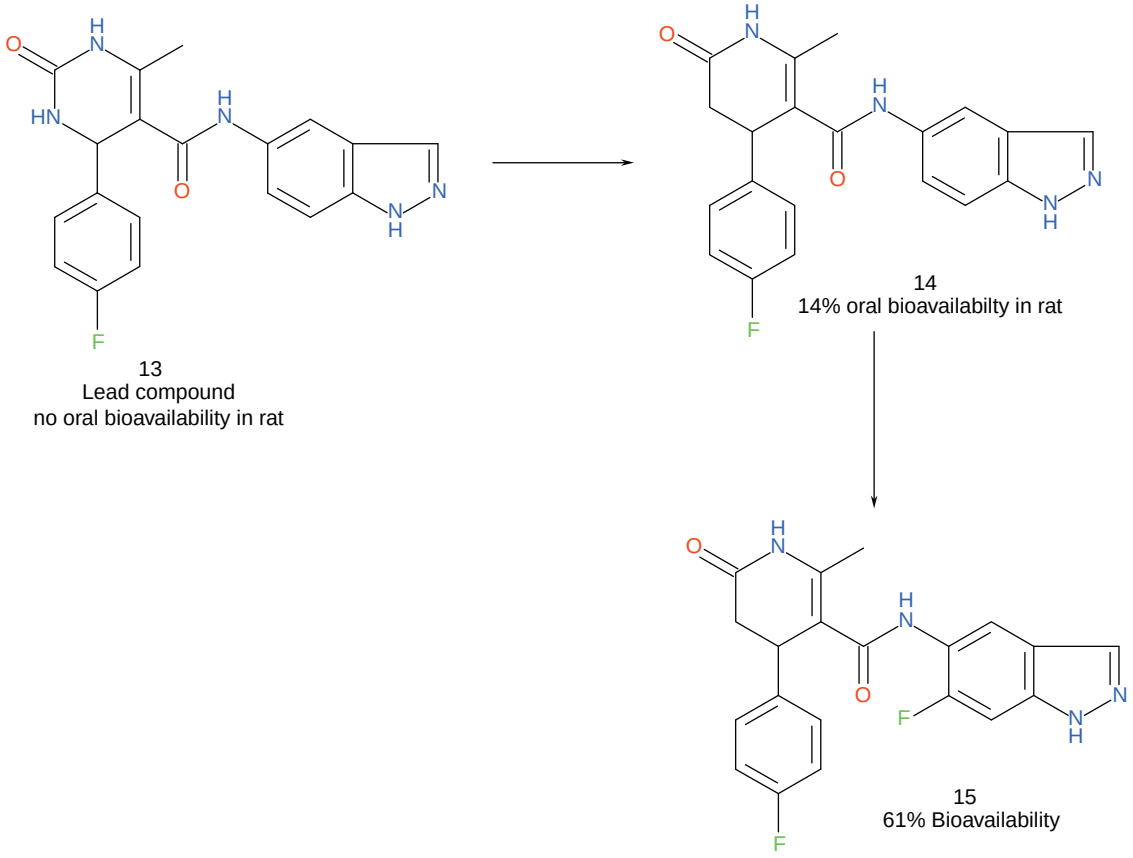


12
(NITD609)

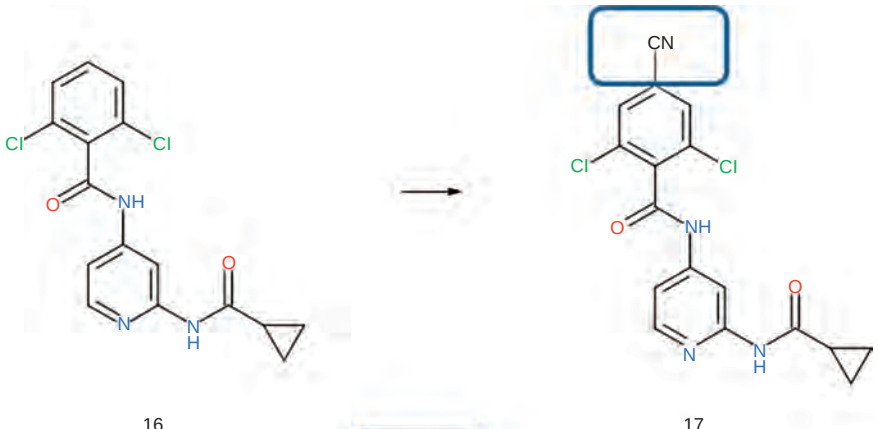
NF54 EC50 = 1 nM
mouse microsomes t1/2= 49 min.

Another example is represented by a series of Dihydropyridone Indazole amides 13–15, which are Rho-Kinase inhibitors [27]. Decreasing the hydrogen bond results in increased bioavailability. As the lead compound shows poor bioavailability, metabolism was also investigated also. Following incubation in rat hepatocytes, LC/MS data suggested significant oxidation at the indazole moiety. Further optimization blocking the C4 position resulted in significantly increased metabolic stability of 15 with highly increased bioavailability.

Metabolic stability and *in vivo* PK of a number of TYK2 inhibitors have been investigated. Differences in drug metabolism across species is illustrated in the following example: [28]



Metabolic stability and *in vivo* PK of a number of analogs have been investigated.

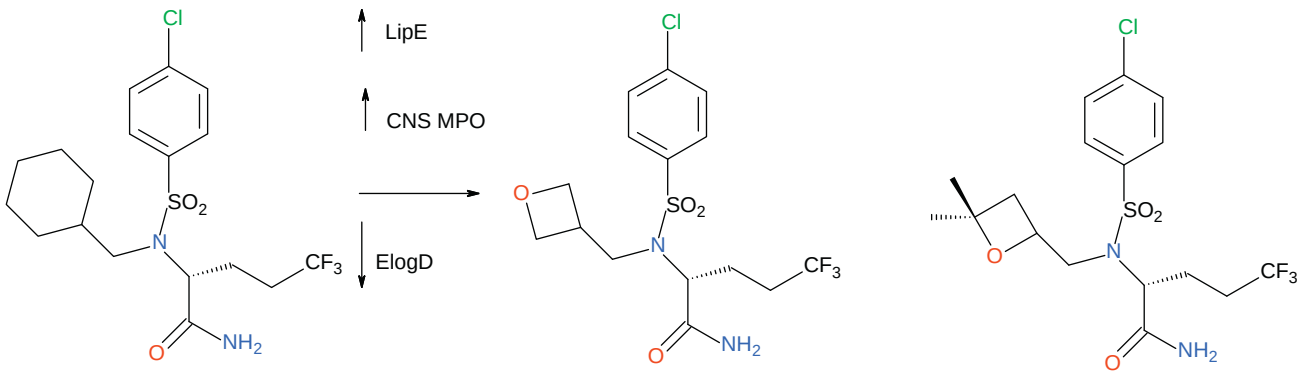


In vitro				In vivo PK									
Stability Liver u-somes (ml/min/kg)				Rat				Mouse				High dose Mouse	
compd	human	rat	mouse	CIp (ml/min/kg))	T1/2 (h)	Vd (L/kg)	F (%)	CIp (ml/min/kg))	T1/2 (h)	Vd (L/kg)	F (%)	AUC (uM/h)	Cmax (uM)
16	7	16	64	21	1.2	1.1	100	65	0.3	0.6	73		
17	8	12	38	31	4.2	0.8	86	13	1	0.9	100	818	919
18	8	11	34	1	40	2.3	70					917	108

The cyanide group was introduced (compounds **16** and **17**) to block oxidative metabolism. However, no significant change was observed in human liver microsomal stability, while the *in vivo* clearance in rats was reduced from 21 to 3.1 ml/min/kg. Introducing a fluorine atom in the cyclopropane ring and keeping the cyanide function (compounds **17** and **18**] resulted in low clearance in rats and an excellent PK profile in mice.

A metabolism based approach was reported for a series of y-secretase inhibitors [29]. Compound **19** showed extensive metabolism in the cycloalkyl motif, as indicated by poor human liver microsomal stability. The ring system was systematically modified to decrease lipophilicity and metabolism. The effort started with tetrahydropyran, and

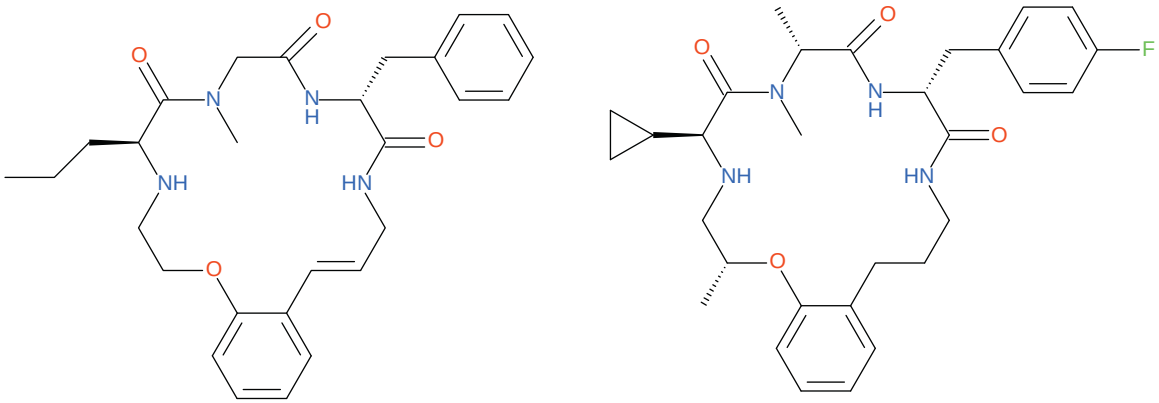
tetrahydrofuran derivatives showed greater microsomal stability. Further variations investigated cyclobutane derivatives, which showed excellent stability while still maintaining pharmacological activity.



19 20 21

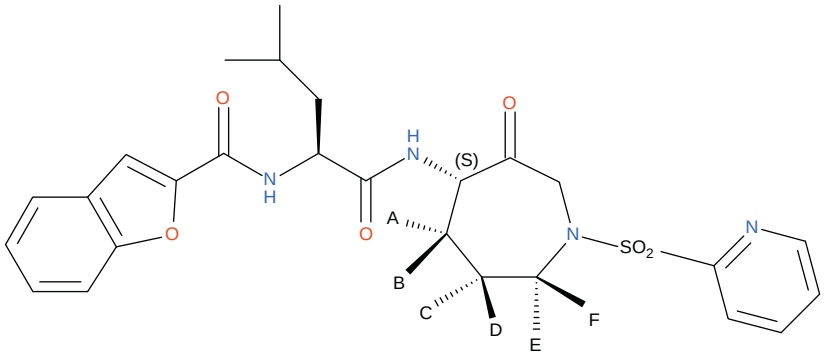
E. Structural Rigidity

An interesting example for changing structural rigidity to enhance bioavailability is the development of the Ghrelin receptor antagonist Ulimorelin and its lead optimization [30]. Starting form early hits like **22**, the lead optimization finally led to **23** (Ulimorelin), which has now progressed into phase II studies for the treatment of postoperative Illeus. As such, macrocyles are typically not really Ro5 compliant, the authors focused on reducing the molecular flexibility as outlined by Veber et al [8]. Extensive SAR studies, including a X-ray structure and NMR studies, resulted in **23** (Ulimorelin), which showed about 24 percent oral bioavailability in rats and monkeys.



22 23 Ulimorelin

A series of cathepsin K inhibitors of azepanone derivatives has been optimized [31] through modification of the azepinone moiety.



24

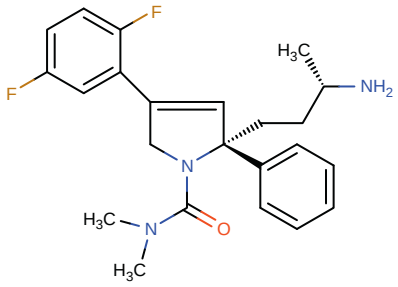
Compound 24	A	B	C	D	E	F
a	H	H	H	H	H	H
b	H	H	H	Me	H	H
c	H	H	H	H	Me	H

Bioavailabilties of 42, 26, and 90 percent of compounds 24a, 24b, and 24c were observed in rats. The authors suggested this was due to a decrease in rotational freedom [8]. Although the number of rotable bonds is equal among the three compounds at eight, the authors comment that 7-methyl-ring substitution should decrease the rotational freedom of the 4-sulfonamide substituent.

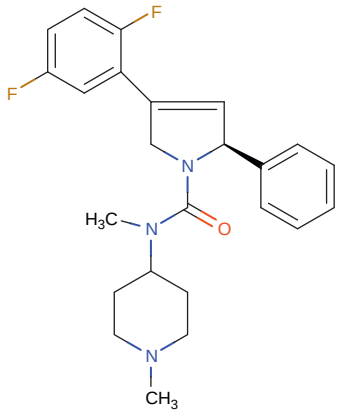
F. pKa

As already outlined in section B about permeability, it is vital to keep the pKa within the optimal range for absorption. Typically, keeping the compound net charge in the pH range of 5.5–7.0 improves intestinal permeability [32].

In some cases, keeping the pKa at a certain level is necessary to reach the target as well [33]. For a number of kinesin spindle protein inhibitors, compound **25** possessed a good *in vitro* profile revealing good target activity but no critical hERG affinity.



25



26

KSP IC_{50} = 5 nM

hERG IC_{50} = 15.9 uM

in vivo EC_{90} = 1425 nM

Poor *in vivo* activity in mouse xenograft model

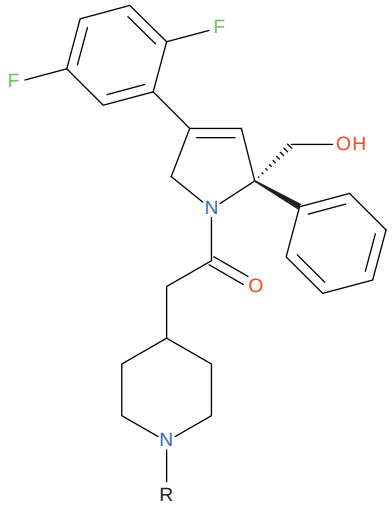
KSP IC_{50} = 4 nM

hERG IC_{50} = 1.2 uM

in vivo EC_{90} = 400 nM

Unacceptable affinity for hERG channel

Compound **26** showed an unacceptable affinity for the hERG channel, however, so derivatives of **27** were investigated.

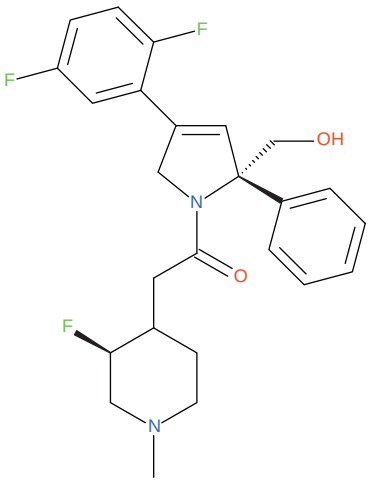


27

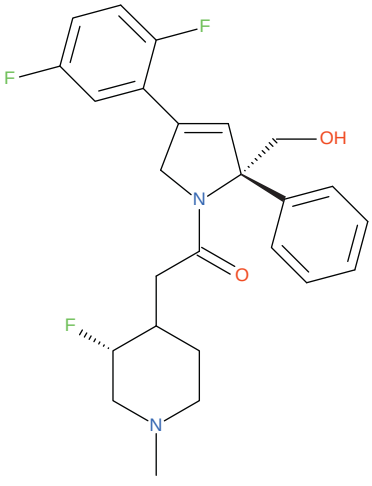
R	Compound 27	KSP IC50 (nM)	MDR ratio	pKa	hERG IC50 (uM)
H	a	7.4	345.1	9.8	19.2
CH2-Cyclopropyl	b	5.1	29.9	9.1	
Me	c	6.2	21.2	8.8	14.6
Fluoroethyl	d	5.0	4.5	7.6	13.8
Cyclopropyl	e	5.9	1.2	7.5	15.2
Difluoroethyl	f	34.1	1.2	4.9	

For KSP inhibition, a pKa range of 6.5–8 was optimal. Any inhibitor with pKa $<$ 6.5 (e.g., compound **27f**) had lower KSP inhibition, while any inhibitor with a pKa $>$ 8 (e.g., compounds **27a**, **27b**, and **27c**) had too high of an MDR ratio. The two inhibitors **27d** and **27e** were found to be optimal.

Compound **27d** showed metabolic dealkylation to **27a** in rat liver microsomes accompanied by fluoroacetaldehdye as a reactive byproduct, which was then further oxidized to fluoroacetic acid. As fluoroacetic acid is a toxic product, another solution had to be found.



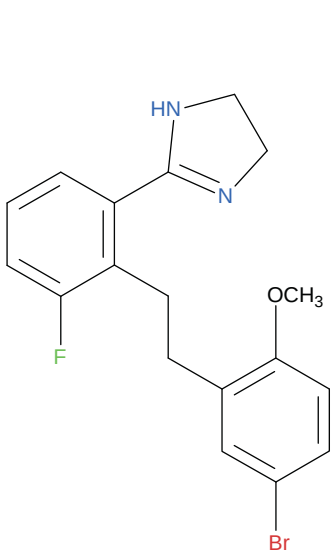
28



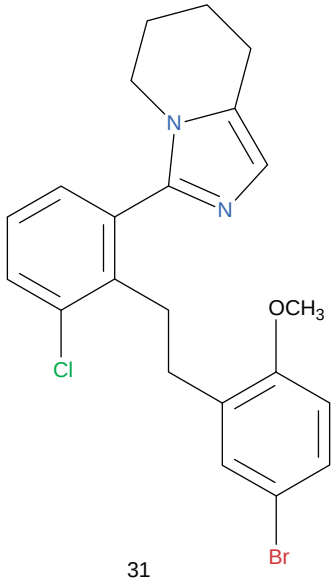
29

By deploying a fluorine atom in the piperidine ring, an elegant and effective solution was found. The effect of the fluorine atom on basicity depends on the stereochemistry. In the trans isomer **29**, the effect was stronger (pKa from 8.8 to 6.6) as in the cis isomer **28** (pKa from 8.8 to 7.6). Compound **28** was advanced into clinical trials.

A series of melanocortin-4 receptor antagonists was investigated [34]. The initial compound **30** showed good activity but demonstrated low exposure in rats with rapid plasma clearance (AUC = 204 nM h, Cl = 13.3 L/h/kg). Reducing the basicity of compound **30** yielded compound **31**, which resulted in significantly improved exposure (Cmax of 6,434 nM compared to 31 nM for compound **30**). MC4-R binding activity of **31** versus **30** was somewhat lower but still in the submicromolar range.



30

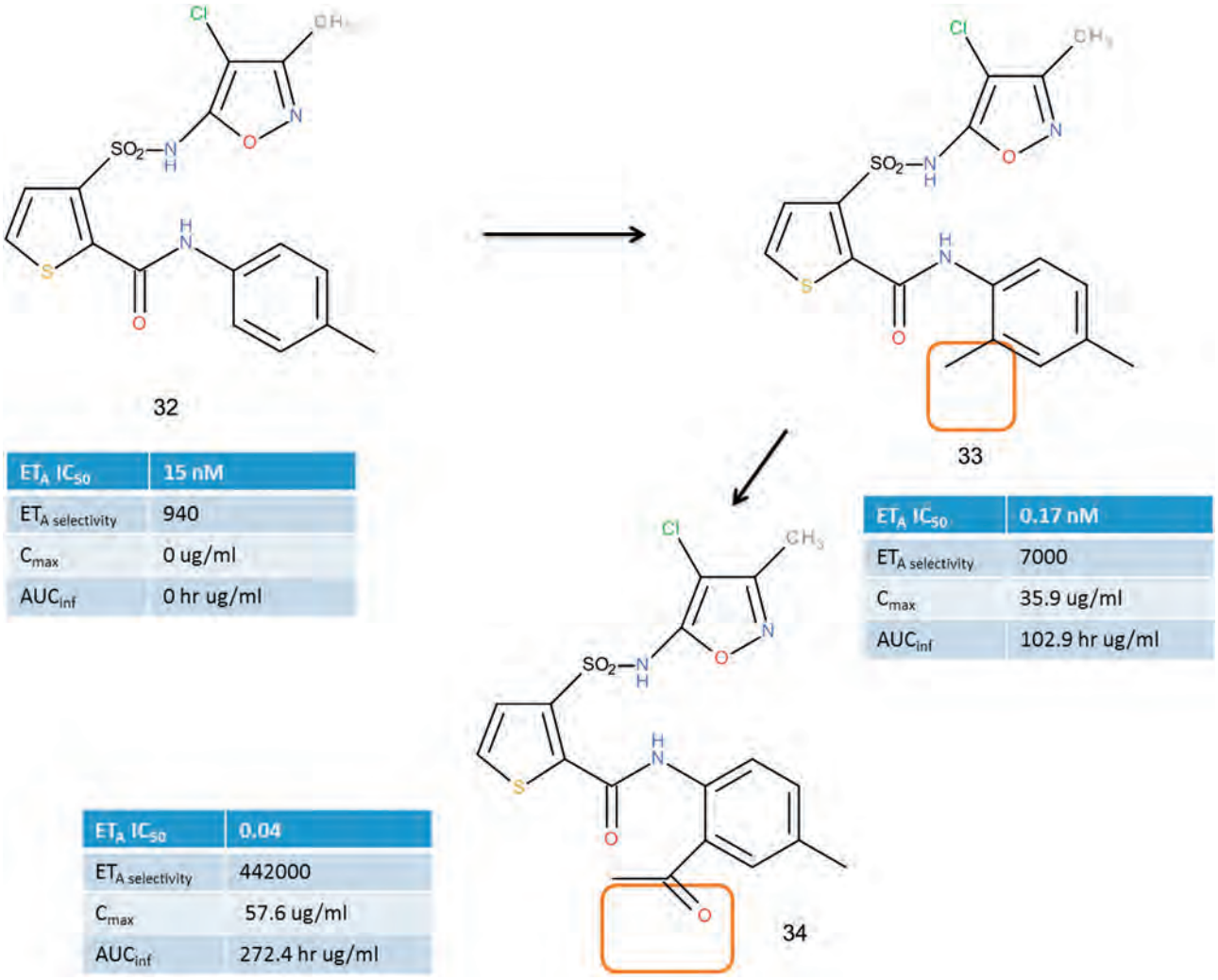


31

G. Hydrogen Bond Interactions

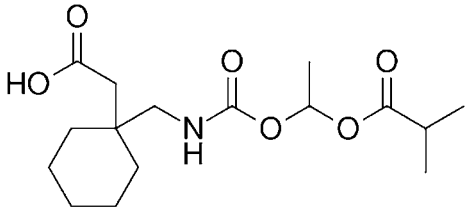
As outlined previously, hydrogen bonding plays a vital role in absorption.

For a series of ET_A-selective endothelin antagonists, the PK properties (e.g. compound 32) were poor as the compound has not been bioavailable. Introduction of a methyl group (compound 33) improved pharmacological activity as well as PK properties. Introduction of the acyl group improved both properties further (compound 34) [35]. The authors suggest that the formation of an internal hydrogen bond between the amide NH and the acyl carbonyl would increase permeability and thus bioavailability.



H. Transporter Based Strategies

Some drugs like valacyclovir are substrates of uptake transporters and show much higher absorption as they are actively transported [36,37]. Another example would be compound **35** (gabapentin enacarbil), a prodrug that is actively taken up by the two high-capacity nutrient transporters MCT1 and SMVT in the intestine [38]. After absorption, enzymatic cleavage releases active drug.



35

I. Prodrugs

Although prodrugs will not be discussed in detail, they should be mentioned here as a strategic option [39,40]. The principle would be typically in order to increase absorption to use a lipophilic prodrug to enable its passage through the intestinal wall and following that its cleaved to active drug.

III. ENHANCING BRAIN PENETRATION

A. Introduction and Chemical Space

The blood–brain barrier is another barrier that has special characteristics [41]. The barrier is formed by capillary endothelial cells which are connected by tight junctions with a very high electrical resistance. It limits the entry of polar compounds when they are not entering by active transport or endocytosis. For low molecular weight compounds, some rules of thumb have been developed.

MW <500, PSA <140, and <10 rotatable bonds have been associated with good oral absorption, while MW <450 and PSA <70 have been indicated as requirements for good CNS penetration [42,43,44] (cf. Table 27.2).

TABLE 27.2 Suggested Physicochemal Property Ranges for Increasing the Potential for Brain Penetration (from [44])

Property	Top 25 CNS drugs mean values	Suggested limits	% of top 25 CNS drugs in suggested range	Preferred range	% of top 25 CNS drugs in Preferred range
PSA (Å^2)	47	<90	96	<70	76
HBD	0.8	<3	100	0–1	92
cLogP	2.8	2–5	68	2–4	52
clogD (pH = 7.4)	2.1	2–5	61	2–4	61
MW	293	<500	100	<450	100

The blood–brain barrier contains a number of transporters and has very tight junctions (Figure 27.5):

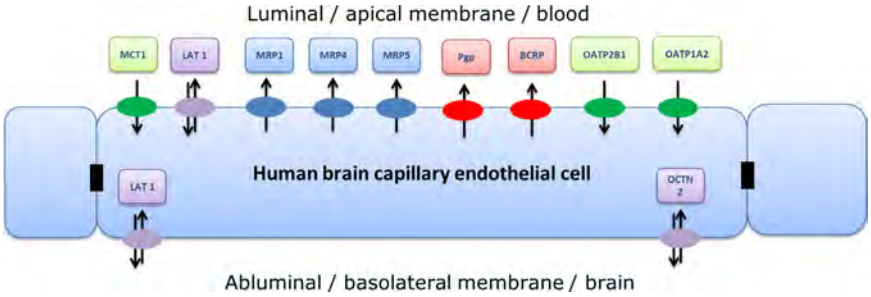


FIGURE 27.5 Transporters at the blood–brain barrier.

Efflux pumps like pGP and BCRP are known to limit brain penetration substantially for many low molecular weight compounds. Based on this understanding, a typical screening cascade in drug discovery also includes assessment of the efflux properties of drug candidates (Figure 27.6).

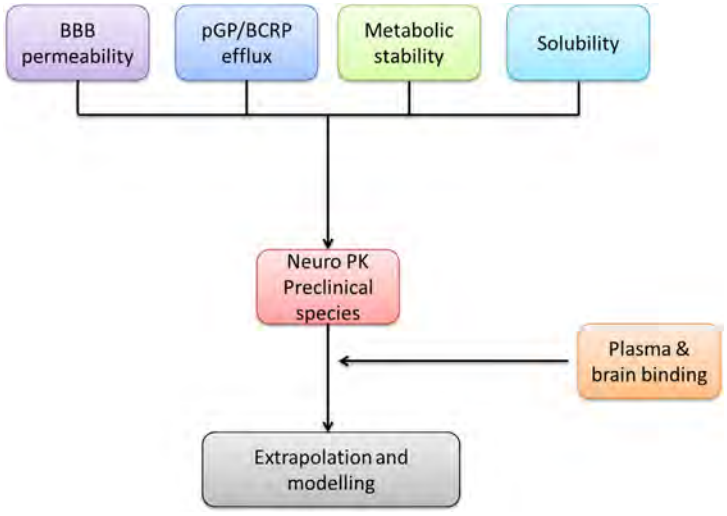
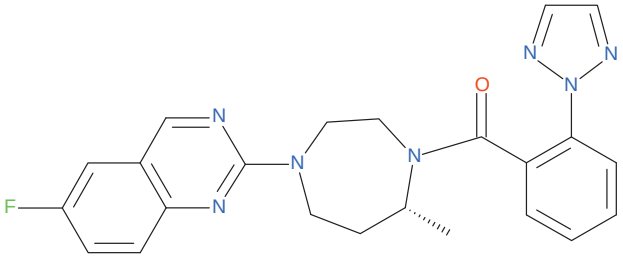


FIGURE 27.6 A typical screening strategy for CNS compounds.

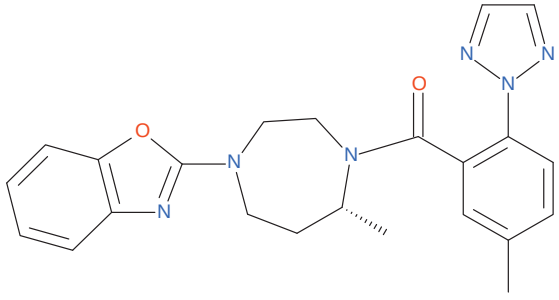
The importance of efflux can be exemplified by the first generation of antihistamines. (+)Chlorpheniramine had good brain penetration, which showed sedation as one of the major side effects. By using pGP mediated efflux in the 2nd generation antihistamine drug ebastine, brain penetration was much more limited with sedative side effects being strongly reduced [45].

B. LogP

A number of potent dual orexin receptor antagonists were investigated [46] for the treatment of insomnia. Hit-to-lead efforts following HTS resulted in the novel diazepane series. Compound **36** showed reactive metabolism that had to be addressed. The optimization efforts resulted in compound **37**, which showed excellent brain permeability due to the added lipophilicity (LogP for **36** was 3.0; for compound **37** it was 3.6).

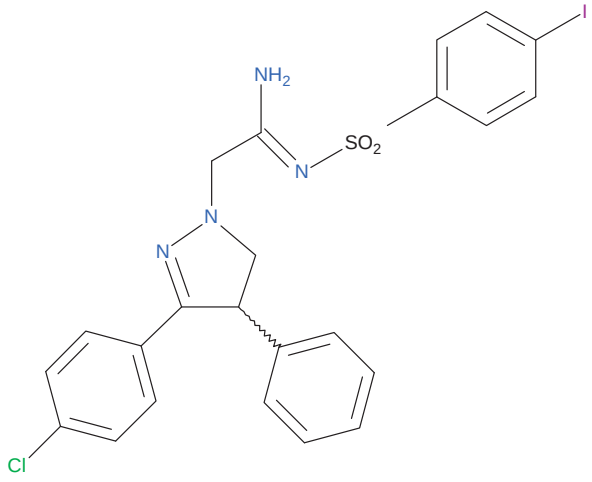


36



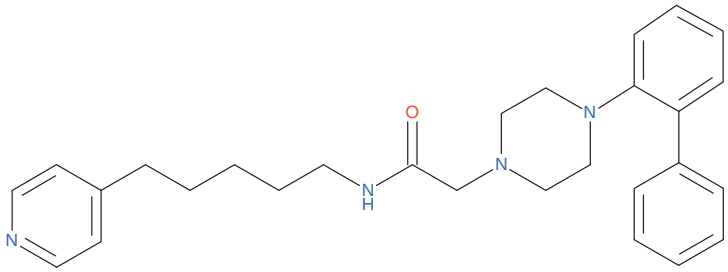
37

A radioligand **38** was specifically designed to stay within a certain logP range (2.0−3.5) to enable rapid brain penetration [47].

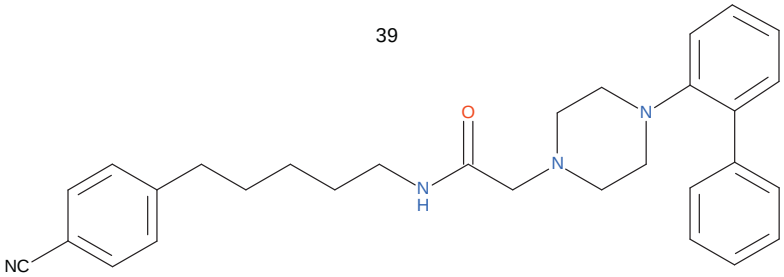


38

A series of *N*-(1,2,3,4-Tetrahydronaphthalen-1-yl)-4-Aryl-1-piperazinehexanamides was structurally modified [48]. Special importance was given to preserving 5-HT7 receptor activity and to keeping the lipophility (logP <5) within a defined range to enable good passive brain penetration.



39



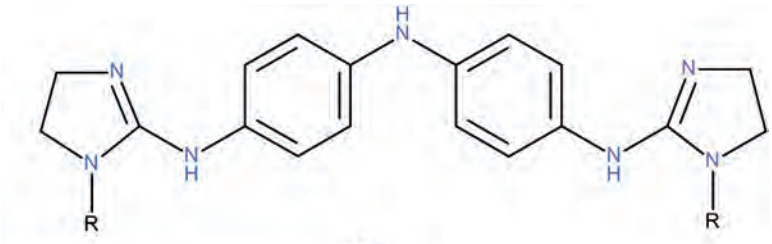
40

Compounds **39** and **40** were among the most promising compounds.

C. pKa

The pKa distribution of 582 drugs has been analyzed [32] regarding CNS and non-CNS drugs. For single acids, a remarkable difference has been found. Only one CNS compound had a pKa below 6.1. For basic compounds a cut-off for CNS compounds of 10.5 was found. No CNS compound had a pKa above this value.

Several compounds [49] have been synthesized that are active against Human African trypanosomiasis (HAT or sleeping sickness). Compound **41a** showed excellent antitrypanosomal and antiplasmodial activity *in vitro*. *In vivo* studies revealed this lead compound as being curative against acute T.b. brucei and T. b. rhodiense infections in murine models following IP administration. As the compound was no longer active in the late stage disease that included CNS infection, the structure was further optimized. The observed lack of efficacy was associated with the high pKa (pKa = 9.9) of the imidazolinylamino group. At physiological pH, the compound would be positively charged and therefore its brain penetration would be very limited.

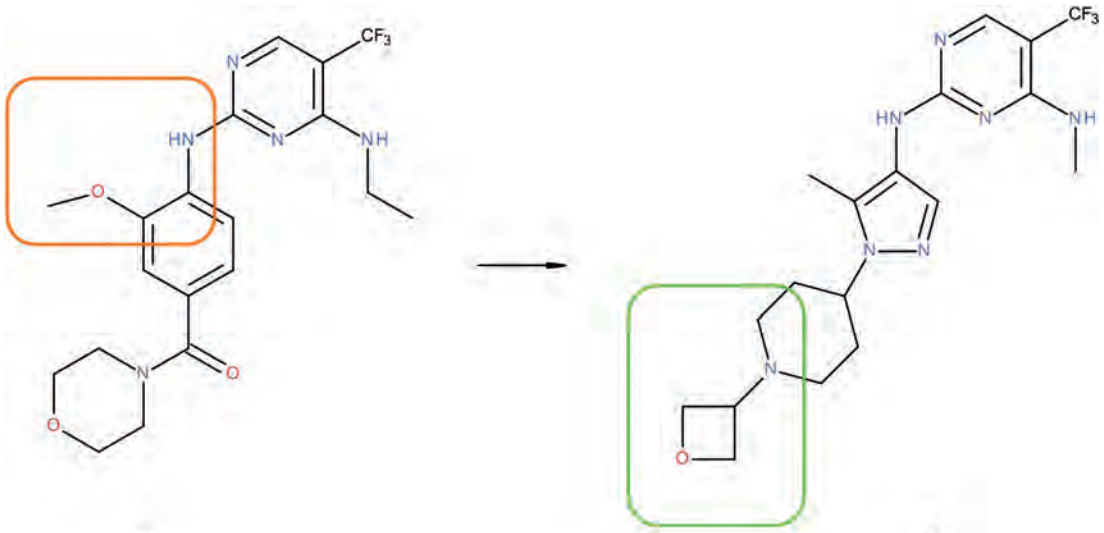


41

Compound 41	R=
a	H
b	OMe
c	OEt
d	OBn
e	OTHP
f	OH

Only compound **41f** had higher brain permeability than compound **41a**. Unfortunately, **41f** had reduced activity compared to **41a** ($IC_{50} = 0.37$ uM). The observed *in vivo* activity was associated with the conversion of **41f** to **41a**. The strategy to lower the pKa of the amidine moiety by attaching OR groups has been successful in additional cases [50,51].

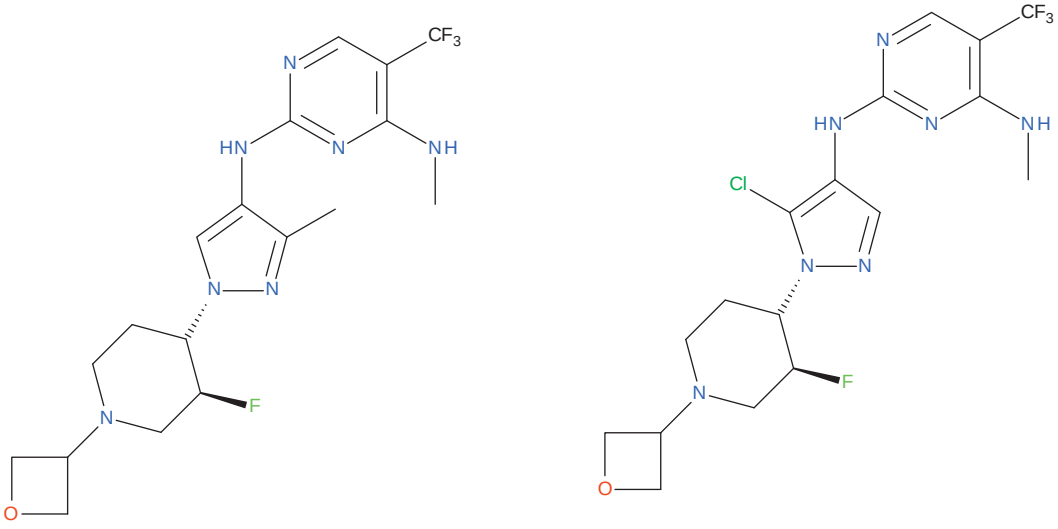
Starting with compound **42** (GNE-7915), successful attempts have been made to further optimize a series of LRRK2 inhibitors [52]. One of the goals was the elimination of the potential for quinoneimine-reactive metabolite formation. One of the promising series was the aminopyrazole derivatives. To control the basicity of the piperidine nitrogen, the electron-withdrawing oxetyl group was added.



42
GNE-7915

43

Despite this, no significant brain penetration was observed, as compound **43** was a substrate of P-gp (ER = 5.1). Compound **44** showed a much lower efflux ratio in MDR-1 overexpressing cells. However, brain penetration was not improved. The authors propose BCRP being the cause for the impaired brain penetration. All compounds were pharmacologically active following small changes to the molecule. Compound **45** showed good brain penetration.



44

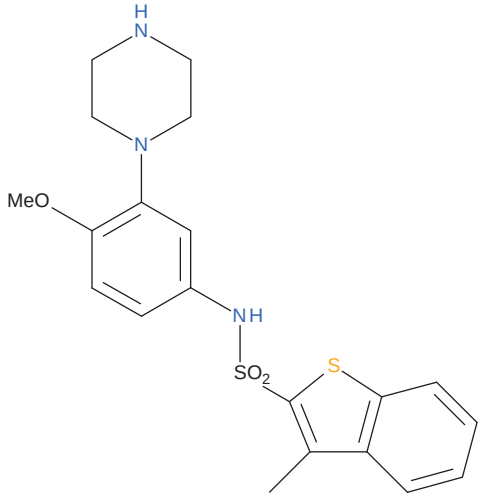
45

Compd	LRRK2 Ki (nM)	pLRRK2 IC50 (nM)	Hep Cl_{hep} (ml/min/kg) Human/rat	MDR1 P-gp ER	Rat Bu/Pu
44	2	14	1/−	2.2	0.08
45	2	19	1/21	0.81	0.51

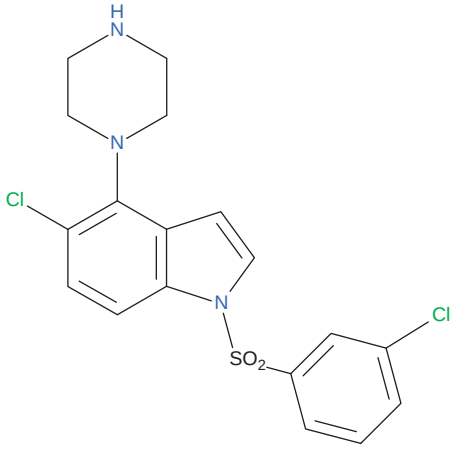
D. Hydrogen Bond Interactions

A study [53] demonstrated that $\Delta\log P_{oct\text{-}tol}$ (difference between $\log P_{octanol}$ and $\log P_{toluene}$) describes a compound's propensity to form intramolecular hydrogen bonds (IMHB) and may be considered a privileged molecular descriptor for use in drug discovery and for prediction of IMHB in drug candidates. Such approaches could also facilitate brain penetration.

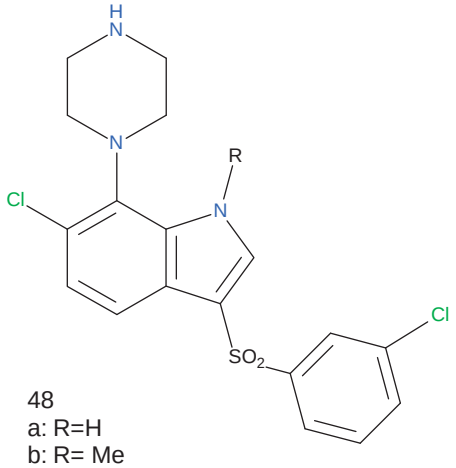
A series of HT6 receptor antagonists was investigated [54]. Compound **46** showed excellent pharmacological data, as it was a selective and highly potent compound. However, the brain/plasma ratio (B/P) was quite low because it was a substrate for pGP related efflux. Reducing the number of H-bond donors and increasing structural rigidity resulted in a compound with much better B/P ratios. All compounds still displayed high affinity for the human 5-HT6 receptor (p*K*i >8.0).



46



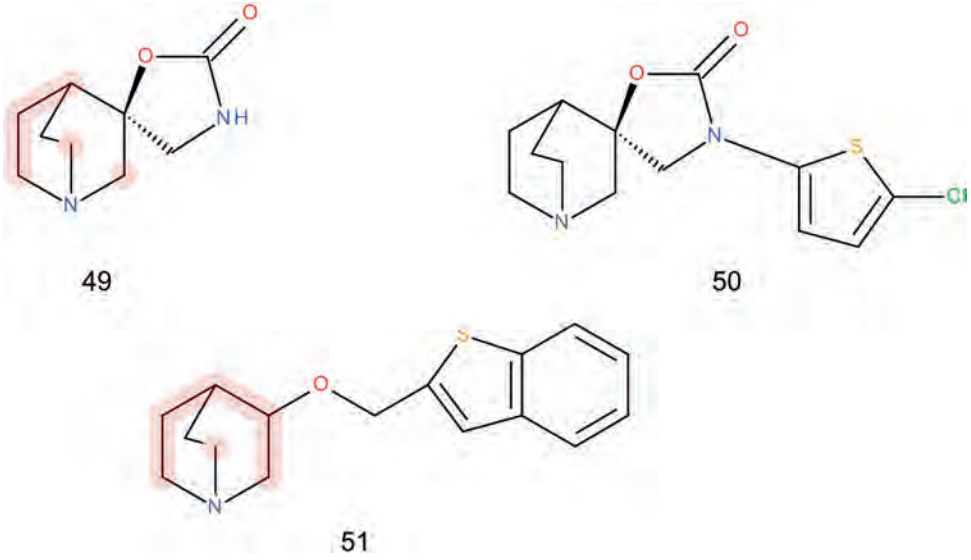
47



48
a: R=H
b: R= Me

Compound	MW	PSA (Å^2)	clogP	clogD	HBD	B/P
46	452	71	4.1	3.6	2	0.05
47	410	54	4.4	3.6	1	3.0
48a	376	65	2.9	1.1	2	0.7
48b	390	54	3.0	1.4	1	2.6

A study investigating the brain penetration of R7 neuronal nicotinic acetylcholine receptors (nAChR) was performed [55,56]. Compound **49** showed only modest activity, while compound **50** showed very good activity but very low PK and brain penetration. Compound **51** showed high pharmacological activity, good bioavailability and excellent brain penetration, mainly due to the low number of HBDs and its increased lipophilicity.

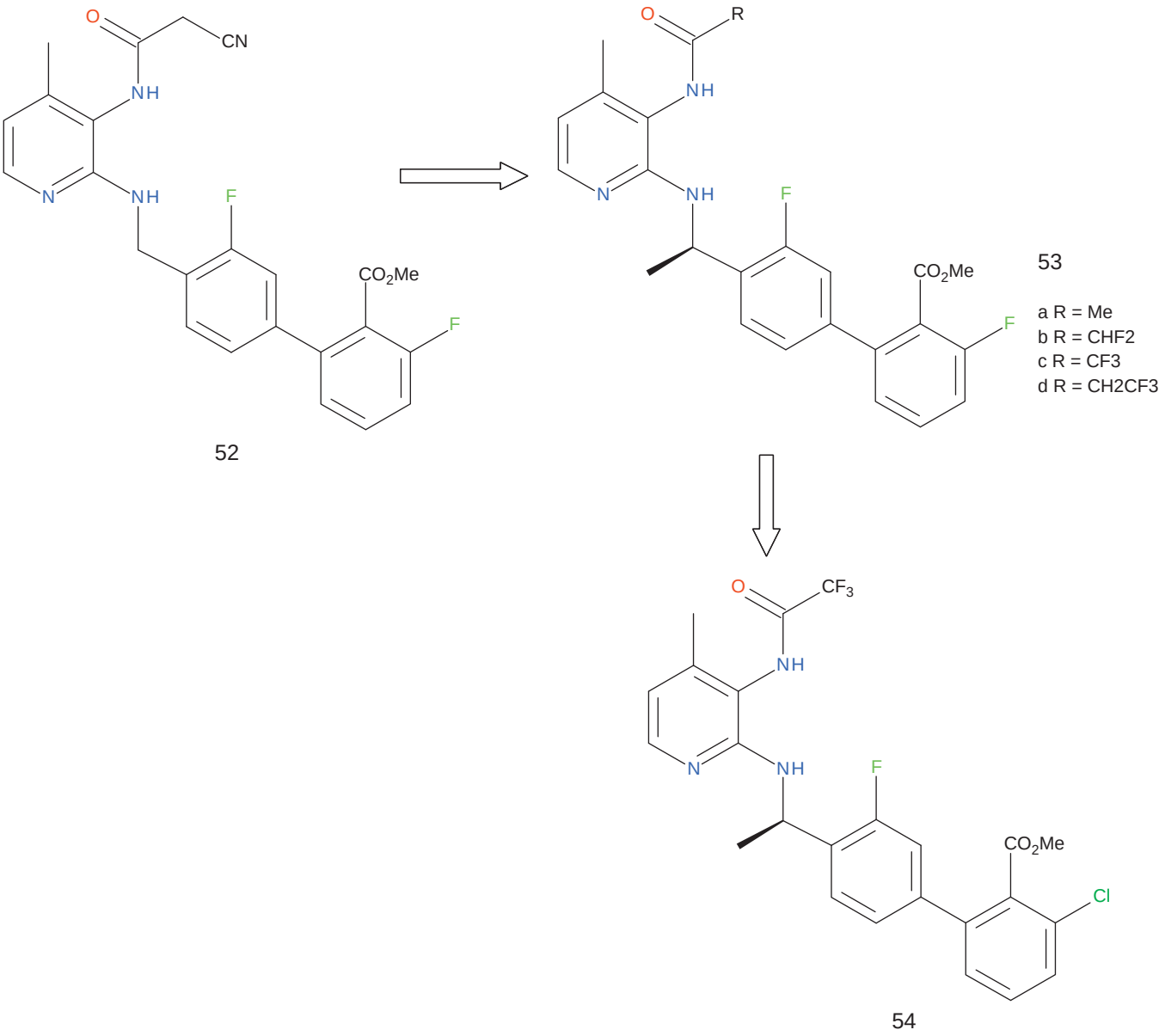


Compound	MW	PSA (Å^2)	cLogP	clogD	HBD	B/P	α7 receptor activity (Ki)
49	182	42	−0.5	−1.5	1	0.76	340 nM
50	273	12	3.5	3.6	0	ND	13 nM
51	299	33	2.1	0.57	0	59	9 nM

E. P-gp Liability

Most drugs acting in the CNS system show low efflux properties and moderate to high permeability. Regarding the efflux properties for the blood−brain barrier, P-gp is the most important efflux pump. Efflux properties for CNS drugs are typically seen as a liability, and they are screened out by a number of *in vitro* models and *in vivo* models.

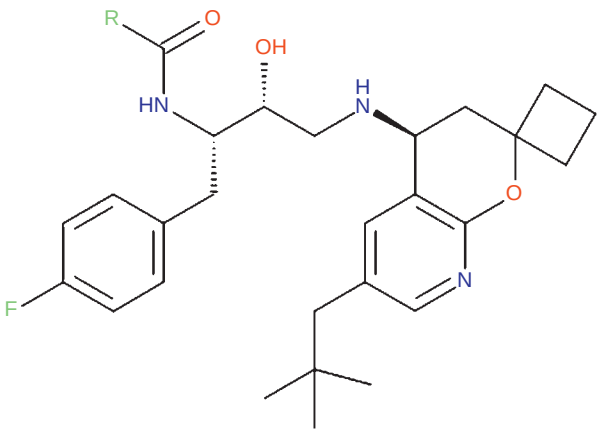
For a series of BACE1 inhibitors [57], the efflux ratio was significantly reduced by modification of the amide group in the molecule while maintaining the hB1 activity in the same range.



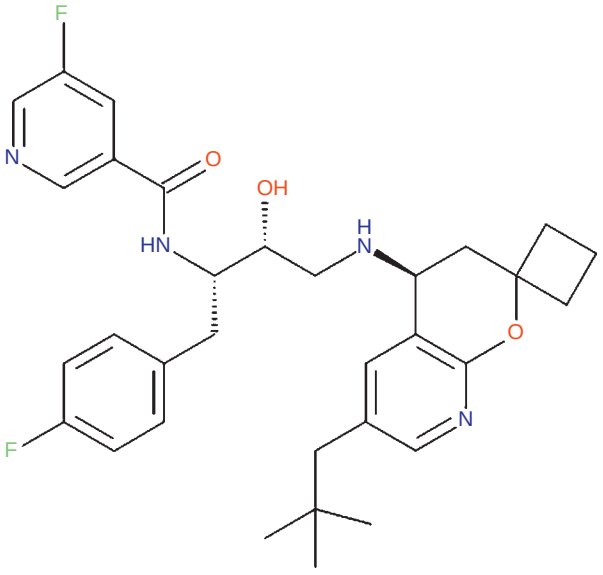
Cmpd	H B1 Ki (nM)	Papp (X 10-6 cm/s)	P-gp (ER)	MW	ClogP	PSA (Å^2)	HBD
52	0.71	NR	22.6	450	3.6	104.1	2
53a	0.93	21	8.6	416	1.0	84.5	2
53b	0.4	31	3.2	452	1.6	84.5	2
53c	0.57	28	2.3	470	2.6	84.5	2
53d	0.13	25	8.6	484	2.3	84.5	2
54	0.44	34	1.9	486	3.1	84.5	2

The modulation of efflux properties was achieved by placing hydrogen acceptor bonds proximal to the amide NH group.

Perspective



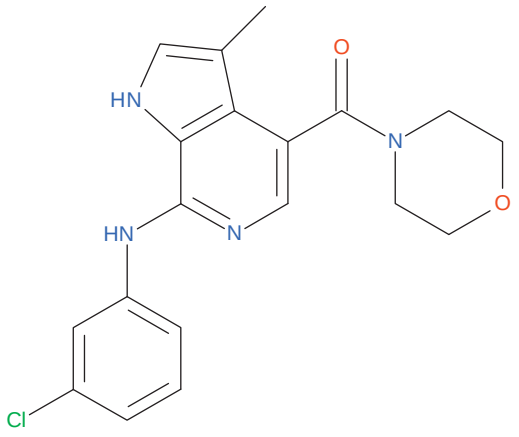
55



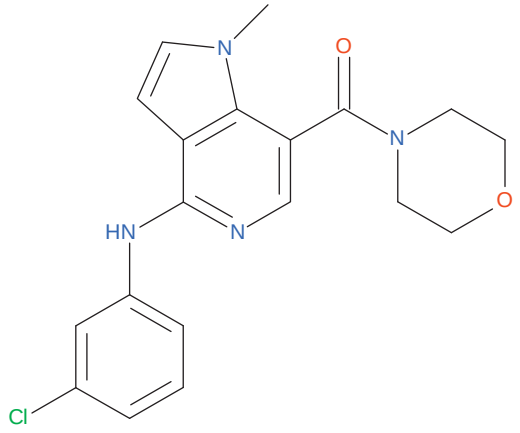
56

For R = Me the efflux ratio for hP-gp was 17 while for compound 56 it was only 1.6. BACE IC50 was 13 nM versus 30 nM for compound 56.

A rather striking example from the same review (57) is the following:



57



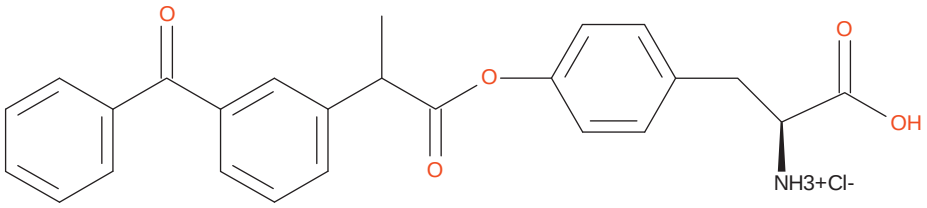
58

cmpd	CB2 EC50 (nM)	P-gp (ER)	Rat B/P	MW	cLogP	cLogD (pH = 7.4)	pKa	PSA (Å^2)	HBD
57	11	74	<0.05	371	2.3	1.4	1.4	70.3	2
58	8	2.9	1.04	371	2.2	2.2	2.2	59.4	1

A dramatic shift of P-gp efflux was achieved during the optimization of azaindole CB2 agonists. Note that a small change of the molecule (reduction of PSA and reduction of HBD from 2 to 1) caused a significant difference in efflux.

F. Miscellaneous

A few attempts to investigate carrier mediated prodrugs should be mentioned here (e.g., using the LAT1 transporter) [58]. The model prodrug **59** used was an Ibuprofen moiety coupled to L-tyrosine. An active brain uptake through the LAT-1 transporter could be demonstrated. Such approaches rely on carrier-mediated uptake into brain and subsequent cleavage to the active moiety in the brain.



59

References

[1] Masaoka Y, Tanaka Y, Kataoka M, Sakuma S, Yamashita S. Site of drug absorption after oral administration: assessment of membrane permeability and luminal concentration of drugs in each segment of gastrointestinal tract. Eur J Pharm Sci 2006;29(3–4):240–50.
[2] Hamman JH, Enslin GM, Kotze AF. Oral delivery of peptide drugs. Biodrugs 2005;19:165–77.
[3] Vaidyanathan S, Jarugula V, Dieterich HA, Howard D, Dole WP. Clinical pharmacokinetics and pharmacodynamics of aliskiren. Clin Pharmacokin 2008;47(8):515–31.
[4] Fasinu P, Pillay V, Ndesendo VMK, du Toit LC, Choonara YE. Diverse approaches for the enhancement of oral drug bioavailability. Biopharm Drug Dispos 2011;32:185–209.
[5] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv 2001;46(1–3):3–26.
[6] Broccatelli F, Cruciani G, Benet LZ, Oprea TI. BDDCS class prediction for new molecular entities. Mol Pharm 2012;9(3):570–80.
[7] Varma MVS, Obach RS, Rotter C, Miller HR, Chang G, Steyn SJ, et al. Physicochemical space for optimum oral bioavailability: contribution of human intestinal absorption and first-pass elimination. J Med Chem 2010;53:1098–108.
[8] Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem 2002;45:2615–23.
[9] Bergstroem CAS, Wassvik CM, Johansson K, Hubatsch I. Poorly soluble marketed drugs display solvation limited solubility. J Med Chem 2007;50:5858–62.
[10] Williams HD, Trevaskis NL, Charman SA, Shanker RM, Charman WN, Pouton CW, et al. Strategies to address low drug solubility in discovery and development. Pharmacol Rev 2013;65(1):315–499.
[11] Kerns EH, Di L, Carter GT. In vitro solubility assays in drug discovery. Curr Drug Metab 2008;9(9):879–85.
[12] Ran Y, Yalkowsky SH. Prediction of drug solubility by the general solubility equation (GSE). J Chem Inf Comput Sci 2001;41(2):354–7.
[13] Scully SS, Tang AJ, Lundh M, Mosher CM, Perkins KM, Wagner BK. Small-molecule inhibitors of cytokine-mediated STAT1 signal transduction in β-Cells with improved aqueous solubility. J Med Chem 2013;56(10):4125–9.
[14] Takahashi K, Kasai M, Ohta M, Shoji Y, Kunishiro K, Kanda M, et al. Novel indoline-based acyl-CoA:cholesterol acyltransferase inhibitor with antiperoxidative activity: improvement of physicochemical properties and biological activities by introduction of carboxylic acid. J Med Chem 2008;51(15):4823–33.
[15] Vacca JP, Dorsey BD, Schleif WA, Levin RB, Mcdaniel SL, Darke PL, et al. L-735,524: an orally bioavailable human immunodeficiency virus type 1 protease inhibitor. J R Proc Natl Acad Sci USA 1994;91:4096.
[16] Ishikawa M, Hashimoto Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. J Med Chem 2011;54(6):1539–54.

[17] Doherty EM, Fotsch C, Bannon AW, Bo Y, Chen N, Dominguez C, et al. Novel vanilloid receptor-1 antagonists: 2. Structure − activity relationships of 4-oxopyrimidines leading to the selection of a clinical candidate. J Med Chem 2007;50:3415−527.
[18] Wang HL, Katon J, Balan C, Bannon AW, Bernard C, Doherty EM, et al. Novel vanilloid receptor-1 antagonists. 3. The identification of a second-generation clinical candidate with improved physicochemical and pharmacokinetic properties. J Med Chem 2007;50:3528−39.
[19] Di L, Kerns EH. Profiling drug-like properties in discovery research. Curr Opin Chem Biol 2003;7(3):402−8.
[20] Thomas VH, Bhattachar S, Hitchingham L, Zocharski P, Naath M, Surendran N, et al. The road map to oral bioavailability: an industrial perspective. Expert Opin Drug Metab Toxicol 2006;2(4):591−608.
[21] Yang Z. The role of membrane transporters on the oral drug absorption. J Mol Pharm Org Process Res 2013;1(1):1−4.
[22] van Waterschoot RA, Lagas JS, Wagenaar E, van der Kruijssen CM, van Herwaarden AE, Song JY, et al. Absence of both cytochrome P450 3A and P-glycoprotein dramatically increases docetaxel oral bioavailability and risk of intestinal toxicity. Cancer Res 2009;69 (23):8996−9002.
[23] Ye L, Wang T, Tang L, Liu W, Yang Z, Zhou J, et al. Poor oral bioavailability of a promising anticancer agent andrographolide is due to extensive metabolism and efflux by P-glycoprotein. J Pharm Sci 2011;100(11):5007−17.
[24] Lin LS, Lanza Jr TJ, Castonguay LA, Kamenecka T, McCauley E, Van Riper G, et al. Bioisosteric replacement of anilide with benzoxazole: potent and orally bioavailable antagonists of VLA-4. Bioorg Med Chem Lett 2004;14(9):2331−4.
[25] Rafi SB, Hearn BR, Vedantham P, Jacobson MP, Renslo AR. Predicting and improving the membrane permeability of peptidic small molecules. J Med Chem 2012;55:3163−9.
[26] Chatterjee AK. Cell-based medicinal chemistry optimization of high-throughput screening (HTS) hits for orally active antimalarials. Part 1: challenges in potency and absorption, distribution, metabolism, excretion/pharmacokinetics (ADME/PK). J Med Chem 2013;56:7741−9.
[27] Goodman KB, Cui H, Dowdell SE, Gaitanopoulos DE, Ivy RL, Sehon CA, et al. Development of dihydropyridone indazole amides as selective rho-kinase inhibitors. J Med Chem 2007;50:6−9.
[28] Liang J, van Abbema A, Balazs M, Barrett K, Berezhkovsky L, Blair W, et al. Lead Optimization of a 4-aminopyridine benzamide scaffold to identify potent, selective, and orally bioavailable TYK2 inhibitors. J Med Chem 2013;56:4521−36.
[29] Stepan AF, Karki K, McDonald WS, Dorff PH, Dutra JK, DiRico KJ, et al. Metabolism-directed design of oxetane-containing arylsulfonamide derivatives as γ-secretase inhibitors. J Med Chem 2011;54:7772−83.
[30] Hoveyda HR, Marsault E, Gagnon R, Mathieu AP, Vézina M, Landry A, et al. Optimization of the potency and pharmacokinetic properties of a macrocyclic ghrelin receptor agonist (Part I): development of Ulimorelin (TZP-101) from hit to clinic. J Med Chem 2011;54:8305−20.
[31] Yamashita DS, Marquis RW, Xie R, Nidamarthy SD, Oh H-J, Jeong JU, et al. Structure activity relationships of 5-, 6-, and 7-methyl-substituted azepan-3-one cathepsin k inhibitors. J Med Chem 2006;49:1597−612.
[32] Manallack DT. The pK_a distribution of drugs: application to drug discovery. Perspect Medicin Chem 2007;1:25−38.
[33] Cox CD, Coleman PJ, Breslin MJ, Whitman DB, Garbaccio RM, Fraley ME, et al. Kinesin spindle protein (KSP) inhibitors. 9. Discovery of (2S)-4-(2,5-Difluorophenyl)-N-[(3R,4S)-3-fluoro-1-methylpiperidin-4-yl]-2-(hydroxymethyl)-N-methyl-2-phenyl-2,5-dihydro-1H-pyrrole-1-carboxamide (MK-0731) for the treatment of taxane-refractory cancer. J Med Chem 2008;51:4239−52.
[34] Marsiljea TH, Rosesa JB, Calderwooda EF, Strouda SG, Forsytha NE, Blackburna C, et al. Synthesis and biological evaluation of imidazole-based small molecule antagonists of the melanocortin 4 receptor (MC4-R). Bioorganic Med Chem Letters 2004;14:3721−5.
[35] Wu C, Decker ER, Blok N, Li J, Bourgoyne AR, Bui H, et al. Acyl substitution at the ortho position of anilides enhances oral bioavailability of thiophene sulfonamides: TBC3214, an ETA selective endothelin antagonist 1. J Med Chem 2001;44:1211−16.
[36] Burnette TC, Harrington JA, Reardon JE, Merrill BM, de Miranda P. Purification and characterization of a rat liver enzyme that hydrolyzes valaciclovir, the L-valyl ester prodrug of acyclovir. J Biol Chem 1995;270(26):15827−31.
[37] Beauchamp LM, Orr GF, de Miranda P, Burnette T, Krenitsky TA. Amino acid ester prodrugs of acyclovir. Antiviral Chem Chemother 1992;3:157−64.
[38] Cundy KC, Branch R, Chernov-Rogan T, Dias T, Estrada T, Hold K, et al. XP13512 [(+/-)-1-([(alpha-isobutanoyloxyethoxy)carbonyl] aminomethyl)-1-cyclohexane acetic acid], a novel gabapentin prodrug. 1. Design, synthesis, enzymatic conversion to gabapentin, and transport by intestinal solute transporters. J Pharm Exp Ther 2004;311(1):315−23.
[39] Huttunen KM, Raunio H, Rautio J. Prodrugs: from serendipity to rational design. Pharmacol Rev 2011;63:750−71.
[40] Beaumont K, Webster R, Gardner I, Dack K. Design of ester prodrugs to enhance oral absorption of poorly permeable compounds: challenges to the discovery scientist. Curr Drug Met 2003;4:461−85.
[41] Di L, Rong H, Feng B. Demystifying brain penetration in central nervous system drug discovery: miniperspective. J Med Chem 2013;56 (1):2−12.
[42] Pajouhesh H, Lenz GR. Medicinal chemical properties of successful central nervous system drugs. NeuroRx 2005;2:541−53.
[43] van de Waterbeemd H, Camenisch G, Folkers G, Chretien JR, Raevsky OA. Estimation of blood−brain barrier crossing of drugs using molecular size and shape, and H-bonding descriptors. J Drug Target 1998;6:151−65.
[44] Hitchcock SA, Pennington LD. Structure − brain exposure relationships. J Med Chem 2006;49(26):7559−83.
[45] Bartra J, Valero AL, del Cuvillo A, Dávila I, Jáuregui I, Montoro J, et al. Interactions of the H1 antihistamines. Investig Allergol Clin Immunol 2006;16(Suppl 1):29−36.
[46] Cox CD, Breslin MJ, Whitman DB, Schreier JD, McGaughey GB, Bogusky MJ, et al. Discovery of the dual orexin receptor antagonist [(7R)-4-(5-Chloro-1,3-benzoxazol-2-yl)-7-methyl-1,4-diazepan-1-yl][5-methyl-2-(2H-1,2,3-triazol-2-yl)phenyl]methanone (MK-4305) for the treatment of insomnia. J Med Chem 2010;53:5320−32.
[47] Waterhouse RN. Determination of lipophilicity and its use as a predictor of blood−brain barrier penetration of molecular imaging agents. Mol Imaging Biol 2003;5:376−89.
[48] Donohue SR, Pike VW, Finnema SJ, Truong P, Andersson J, Gulyás B, et al. Discovery and labeling of high-affinity 3,4-diarylpyrazolines as candidate radioligands for in vivo imaging of cannabinoid subtype-1 (CB1) receptors. J Med Chem 2008;51:5608−16.

[49] Nieto L, Mascaraque A, Miller F, Glacial F, Martínez CR, Kaiser M, et al. Synthesis and antiprotozoal activity of N-alkoxy analogues of the trypanocidal lead compound 4,4-bis(imidazolinylamino)diphenylamine with improved human blood − brain barrier permeability. J Med Chem 2011;54:485−94.
[50] Clement B. Reduction of N-hydroxylated compounds: amidoximes (N-hydroxyamidines) as pro-drugs of amidines. Drug Met Rev 2002;34:565−79.
[51] Riggs JR, Kolesnikov A, Hendrix J, Young WB, Shrader WD, Vijaykumar D, et al. Factor VIIa inhibitors: a prodrug strategy to improve oral bioavailability. Bioorg Med Lett 2006;16:2224−8.
[52] Estrada AA, Chan BK, Baker-Glenn C, Beresford A, Burdick DJ, Chambers M, et al. Discovery of highly potent, selective, and brain-penetrant aminopyrazole leucine-rich repeat kinase 2 (LRRK2) small molecule inhibitors. J Med Chem 2014;57:921−36.
[53] Shalaeva M, Caron G, Abramov YA, O'Connell TN, Plummer MS, Yalamanchi G, et al. Integrating intramolecular hydrogen bonding (IMHB) considerations in drug discovery using ΔlogP as a tool. J Med Chem 2013;56:4870−9.
[54] Ahmed M, Briggs MA, Bromidge SM, Buck T, Campbell L, Deeks NJ, et al. Bicyclic heteroarylpiperazines as selective brain penetrant 5-HT6 receptor antagonists. Bioorg Med Chem Lett 2005;15:4867−71.
[55] Tatsumi R, Fujio M, Satoh H, Katayama J, Takanashi S-i, Hashimoto K, et al. Discovery of the α7 nicotinic acetylcholine receptor agonists. (R)-3-(5-chlorothiophen-2-yl) spiro-1-azabicyclo[2.2.2]octane-3,5′-[1′,3′]oxazolidin-2′-one as a novel, potent, selective, and orally bioavailable ligand. J Med Chem 2005;48:2678−86.
[56] Tatsumi R, Seio K, Fujio M, Katayama J, Horikawa T, Hashimoto K, et al. (+)-3-[2-(benzo[b]-thiophen-2-yl)-2-oxoethyl]-1-azabicyclo[2.2.2] octane as potent agonists for the R7 nicotinic acetylcholine receptor. Bioorg Med Chem Lett 2004;14:3781−4.
[57] Hitchcock SA. Structural modifications that alter the P-glycoprotein efflux properties of compounds. J Med Chem 2012;55:4877−95.
[58] Gynther M, Laine K, Ropponen J, Leppänen J, Mannila A, Nevalainen T, et al. Large neutral amino acid transporter enables brain drug delivery via prodrugs. J Med Chem 2008;51(4):932−6.

CHAPTER

28

Designing Prodrugs and Bioprecursors

Yong Mi Choi-Sledeski[1] *and Camille G. Wermuth*[2]

[1]Sanofi, Boston, MA, USA; [2]Founder of Prestwick Chemical, Strasbourg, France

OUTLINE

I. INTRODUCTION

Therapeutic approaches based on molecular pharmacology predominantly use *in vitro* models (e.g., membrane or enzyme preparations, cell or microorganism cultures, isolated organs) to identify a lead molecule for optimization. The bioavailability of molecules derived from high-throughput chemical synthesis or chemical libraries can often be low when exclusively screened through *in vitro* assays. Although their potency against the biological target may be high, because of the suboptimal physicochemical property of the molecule, they may be poorly absorbed or incorrectly distributed. As a result of their vulnerability, they may also be subject to early metabolic destructions, such as first-pass effects or other kinds of degradation leading to a short biological half-life. For such molecules, *in vivo* administration is limited to the parenteral route, and their clinical usefulness is thus restricted. Sometimes an adequate pharmaceutical formulation (micro-encapsulation, sustained-release or enterosoluble preparations) can compensate for these drawbacks, but often formulation efforts alone may not be enough to overcome the problem and a chemical modification of the active molecule is necessary to correct its pharmacokinetic insufficiencies. This chemical formulation process, whose objective is to convert an interesting active molecule into a clinically acceptable drug, involves prodrug design [1–7].

II. THE DIFFERENT KINDS OF PRODRUGS

A. Definitions and Classifications

Initially, the term prodrug was introduced by Albert to describe "any compound that undergoes biotransformation prior to exhibiting its pharmacological effects." [8] Such a broad definition includes accidental historic prodrugs (salicin and salicylic acid), active metabolites (imipramine and des-methylimipramine), and compounds intentionally prepared to improve the pharmacokinetic profile of an active molecule. From this point of view, the term "drug latentiation" proposed by Harper [9] is more appropriate for prodrug design, as it indicates that there is an intention. Drug latentiation is defined as "the chemical modification of a biologically active compound to form a new compound that, upon *in vivo* enzymatic attack, will liberate the parent compound." Even this definition is too broad, and a careful survey of the specialized literature led us to divide prodrugs into two main classes: carrier prodrugs and bioprecursors [3,10,11].

Carrier prodrugs result from a temporary linkage of the active molecule with a transport moiety that is frequently of lipophilic nature. A simple hydrolytic reaction cleaves this transport moiety at the correct moment (e.g., pivampicillin, bacampicillin). Such prodrugs are, *per se*, less active than the parent compounds or even inactive. The transport moiety (carrier group) is chosen for its nontoxicity and its ability to ensure the release of the active principle with efficient kinetics.

Bioprecursors do not contain a temporary linkage between the active principle and a carrier group but result from a molecular modification of the active principle itself. This modification generates a new compound, which is a substrate for metabolizing enzymes, leading to a metabolite that is the expected active principle. This approach exemplifies the active metabolite concept in a prospective way (e.g., sulindac, fenbufen, acyclovir, losartan).

B. The Carrier Prodrug Principle

The carrier prodrug principle (Figure 28.1) consists of "the attachment of a carrier group to the active drug to alter its physicochemical properties and then the subsequent enzyme attack to release the active drug moiety." [9] "Prodrugs can thus be viewed as drugs containing specialized nontoxic protective groups used in a transient manner to alter or eliminate undesirable properties in the parent molecule." [2]

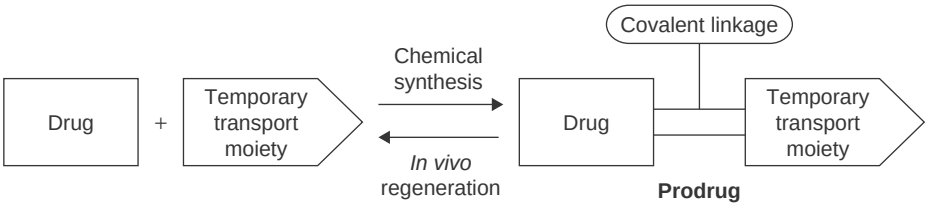


FIGURE 28.1 The carrier prodrug principle [12].

A well-designed carrier prodrug satisfies the following criteria: [1,13]

1. The linkage between the drug substance and the transport moiety is usually a covalent bond.
2. As a rule, the prodrug is inactive or less active than the parent compound.
3. The linkage between the parent compound and the transport moiety must be broken *in vivo*.
4. The prodrug as well as the *in vivo* released transport moiety must be nontoxic.
5. The generation of the active form must take place with rapid kinetics to ensure effective drug levels at the site of action and to minimize either direct prodrug metabolism or gradual drug inactivation.

An example of prodrug design taking these criteria into account is found in orally active ampicillin derivatives [14–16]. Ampicillin is one of the main β-lactam antibiotics. It is widely used as a broad spectrum antibiotic, but it suffers from insufficient absorption when administered orally: approximately 40 percent of the drug is absorbed. In other words, to achieve the same clinical efficiency and the same blood levels, one must give two to three times more ampicillin by mouth than by intramuscular injection. The clinical tolerance of orally given ampicillin

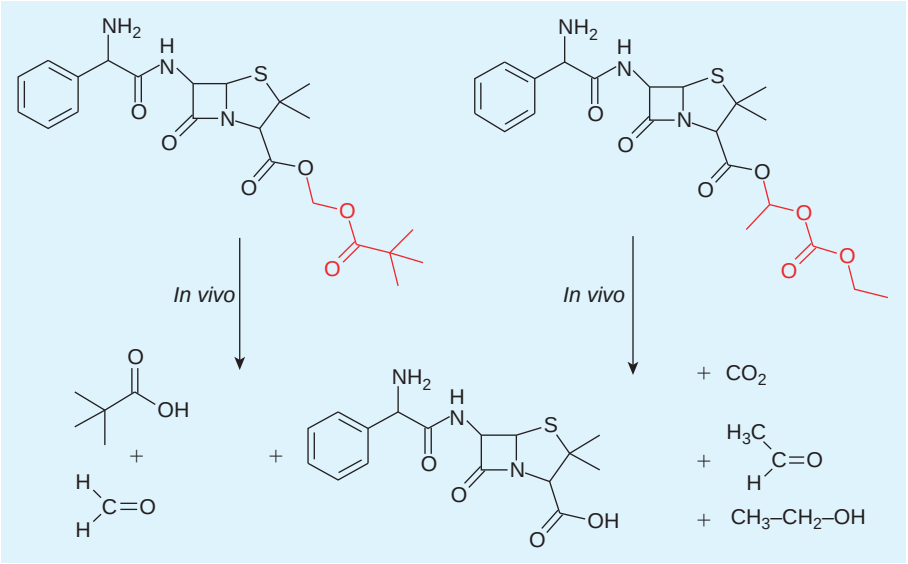


FIGURE 28.2 Prodrugs derived from ampicillin [14–16].

may be affected, with the non-absorbed part of the drug destroying the intestinal flora. Despite the good absorption of ampicillin compared to other drugs in general, improved absorption would be advantageous for better efficacy in this case.

Figure 28.2 represents two prodrugs of ampicillin: pivampicillin and bacampicillin. They both result from esterification of the polar carboxylic group with a lipophilic, enzymatically labile ester. The main properties of these prodrugs can be summarized as follows:

1. Absorption of these compounds is nearly quantitative (98–99 percent).
2. Generation of free ampicillin in the blood stream is rapid (less than 15 minutes).
3. The released carrier molecules are formaldehyde and pivalic acid (trimethyl acetic acid) for pivampicillin and acetaldehyde, ethanol, and carbon dioxide in the case of bacampicillin. These three latter compounds are natural metabolites in the human body, which may explain the better tolerance of bacampicillin compared to pivampicillin.
4. The serum levels attained following oral administration of bacampicillin are similar to those obtained after intramuscular injection of an equimolecular amount of free ampicillin.
5. The clinical trials confirm the efficiency and the safety of the prodrugs. Due to their good absorption, the drugs are given at lower dosage than ampicillin; 0.8–1.0 g daily is sufficient in common infections, as compared to 2.0 g daily for ampicillin.
6. It has been shown—and this seems to be a rule for prodrugs—that pivampicillin and bacampicillin are inactive *per se*, the antibiotic potency appearing only *in vivo* after the release of free ampicillin.

C. The Bioprecursor Prodrug Principle

As stated above, bioprecursor prodrugs result from molecular modification of the active principle, generating a new compound that is a substrate for metabolizing enzymes, with the metabolite being the expected active principle. The bioprecursor prodrug approach exemplifies the active metabolite concept. A survey of a great number of examples of active metabolites shows that they belong exclusively to Phase I metabolism products and result from one of the reactions mentioned in Table 28.1. As such, reactions follow some general rules, and they can often be predicted. Taking into account the common metabolic pathways, one can design a given molecule so that it will be converted *in vivo* into the desired pharmacologically active compound by one or more of the Phase I reactions. In other words, the active metabolite concept can be used in a forward-looking way ("metabolic synthesis"). By analogy to the retrosynthetic reasoning used in organic chemistry, we can utilize retrometabolic reasoning in prodrug design. Such reasoning can lead to a particular group of prodrugs for which the name bioprecursors or bioprecursor prodrugs was proposed [10,11].

TABLE 28.1 Phase I Reactions [17].

OXIDATIVE REACTIONS
Oxidation of alcohol, carbonyl and acid functions, hydroxylation of aliphatic carbon atoms, hydroxylation of alicyclic carbon atoms, oxidation of aromatic carbon atoms, oxidation of carbon–carbon double bonds, oxidation of nitrogen-containing functional groups, oxidation of silicon, phosphorus, arsenic and sulfur, oxidative *N*-dealkylation, oxidative *O*- and *S*-dealkylation, oxidative deamination, other oxidative reactions
REDUCTIVE REACTIONS
Reduction of carbonyl groups, reduction of alcoholic groups and carbon–carbon double bonds, reduction of nitrogen-containing functional groups, other reductive reactions
REACTIONS WITHOUT CHANGE IN THE STATE OF OXIDATION
Hydrolysis of esters and ethers, hydrolytic cleavage of C—N single bonds, hydrolytic cleavage of nonaromatic heterocycles, hydration and dehydration at multiple bonds, new atomic linkages resulting from dehydration reactions, hydrolytic dehalogenation: removal of hydrogen halide molecules, various reactions.

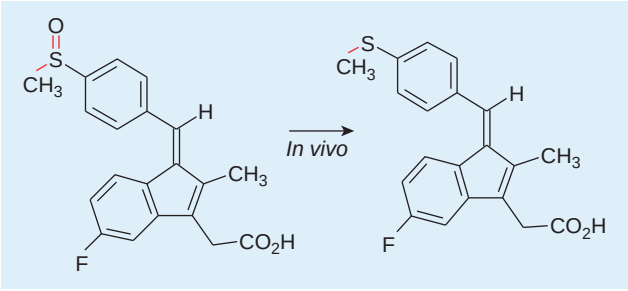


FIGURE 28.3 Reductive bioactivation of sulindac [22].

A typical example of an effective bioprecursor prodrug is the anti-inflammatory drug sulindac. Sulindac, cis-5-fluoro-2-methyl-1-[p-(methylsulfinyl) benzylidene] indene-3 acetic acid [18], is a nonsteroidal anti-inflammatory drug (NSAID) having a broad spectrum of activity in animal models and in humans. The two quantitatively significant biotransformations undergone by sulindac in laboratory species [19] and in humans [20,21] involve only changes in the oxidation state of the sulfinyl substituent, namely irreversible oxidation of the parent (sulindac) to sulfone and reversible reduction to sulfide (Figure 28.3), the latter being the active species [22]. In two *in vitro* models of inflammation, prostaglandin synthase inhibition and inhibition of platelet aggregation, the sulfide has activities comparable to those of indomethacin, whereas sulindac itself is devoid of activity. Nevertheless, sulindac is the preferred compound for clinical applications. Oral administration of this inactive bioprecursor circumvents initial exposure of the active drug to gastric and intestinal mucosa, thus reducing the gastrointestinal irritation often observed with NSAIDs and providing a therapeutic advantage over the sulfide.

D. Novel Categories of Prodrugs

A number of innovative ideas in prodrug design have been proposed by researchers in academia and industry. Among them are the design of soft drugs, codrugs, pro-prodrugs, cascade prodrugs, and transporter-targeted prodrugs. Areas of intense investigation relate to site-specific prodrug delivery systems and antibody–drug conjugates in particular. The interest in these approaches lies in the potential to increase efficacy and improve the safety window by delivering the drug to the site of action. There are macromolecular prodrugs in clinical studies that show promise, and a few have reached the market, but patient outcomes will determine whether these novel approaches can lead to transformative new medicines. Key examples will be highlighted and the approaches discussed in Section IV.

E. Prodrug Design to Improve Parent Drug Shortcomings

The domain of application of the prodrug approach is illustrated in Figure 28.4. In practice, prodrugs usually achieve one of the five following goals: increased lipophilicity, increased duration of pharmacological effects,

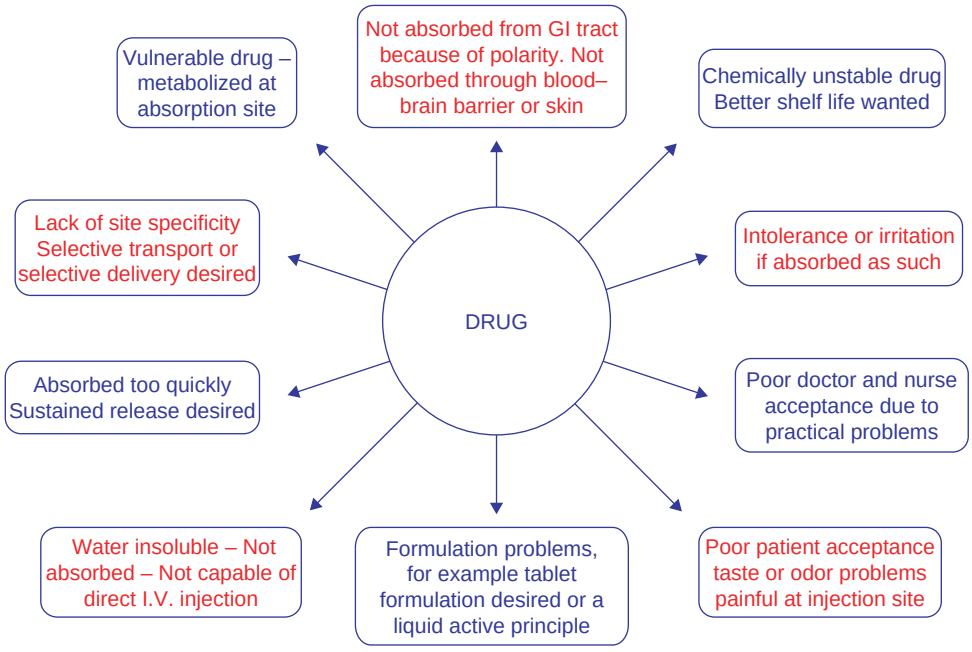


FIGURE 28.4 Shortcomings that may be overcome through chemical modification [1,12,23].

increased site-specificity, decreased toxicity and adverse reactions, or improvement in drug formulation (stability, water solubility, suppression of an undesirable organoleptic, or physicochemical property).

III. PRACTICAL APPLICATIONS OF CARRIER PRODRUGS

Bioactive compounds and drugs usually bear a limited number of polar functional groups suitable for carrier prodrug synthesis. Among these, the most frequent are alcoholic and phenolic hydroxyls, amino groups, and the carboxylic function. The aim of the following section is to illustrate how such groups can be used to prepare carrier prodrugs with improved pharmacokinetic (PK) and/or pharmacodynamic properties. First, a brief discussion on the biomembrane passage is necessary to understand the mechanism by which drugs are absorbed or transported in the intestinal wall.

A. Improvement of the Bioavailability and the Biomembrane Passage

The biomembrane passage of a drug depends primarily on its physicochemical properties, especially its partition coefficient (Chapter 15). Thus, the transient attachment of a lipophilic carrier group to an active principle can provide better bioavailability, mostly by facilitating cell membrane crossing by passive diffusion. Peroral absorption, as well as rectal absorption, ocular drug delivery, and dermal drug delivery, is dependent on passive diffusion. In addition to improving passive diffusion by modifying the physicochemical properties (lipohilicity) of the drug molecule, a successful strategy to overcome biological barriers is to design prodrugs targeting active transport mechanisms [24–26]. Preferential site-directed delivery of the drug molecule may be envisioned by taking advantage of the differential expression levels of the transporter in cells/tissues in combination with appropriate routes of administration (e.g., topical, intravenous) to increase drug exposure. Besides differential tissue expression, other important considerations in choosing a particular transporter are: (1) whether it is a high-capacity transporter, which is dictated by both expression levels and intrinsic transporter capacity (turnover); (2) whether it has broad substrate specificity to allow flexibility in recognition of the prodrug; (3) whether there are species differences to establish *in vivo* models predictive of human pharmacokinetics (PK); (4) whether there are genetic polymorphisms; and (5) whether there are known drug–drug interactions involving the transporter that can have direct and adverse effects on therapeutic safety and efficacy. Achieving targeted drug delivery in general remains a major challenge in drug discovery.

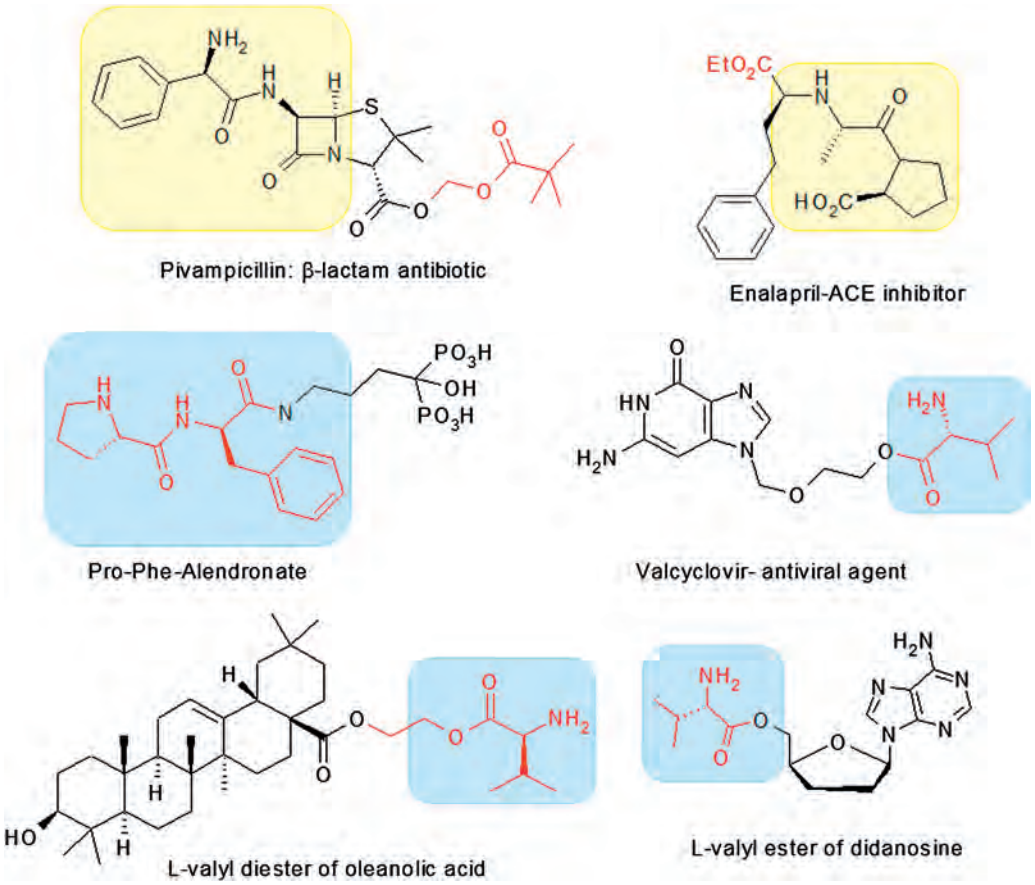


FIGURE 28.5 Peptidyl substrates of PepT1: recognition element inherent to pharmacological activity or conjugated as a promoiety (drug molecule in black; promoiety in red) [25].

Among the many membrane transporters (hASBT, MCT1, OATPs, GluT1, SVCT, LAT1) that have been identified as suitable targets, PepT1 is the most well characterized and amenable to prodrug design [27], particularly since the elucidation of its X-ray crystal structure in 2011 [28]. The nutrient transporter PepT1 is a low-affinity, high-capacity transporter with broad substrate specificity and low occurrence of functional polymorphism. It is located predominantly on the luminal cell membrane of the small intestine to affect oral bioavailability and on the corneal epithelium to improve ocular bioavailability. Its physiological function is the absorption of di- and tripeptide nutrients. Targeting this transporter proved to be a successful strategy in the areas of β-lactam antibiotics, antivirals, renin inhibitors, and ACE inhibitors that were retrospectively found to be substrates for PepT1. Figure 28.5 illustrates classical and recent examples of prodrugs where the PepT1 peptidic recognition element is part of the drug pharmacophore (highlighted in yellow) and examples where the recognition element is incorporated as a promoiety (highlighted in blue). The diversity of structures that are recognized by PepT1 is noteworthy, and the dipeptidyl prodrug of alendronate is a particularly interesting example.

Fosamax (alendronate) has very low oral bioavailability (F = 0.76 percent in postmenopausal women) [29], but it is an effective osteoporosis drug because of its major uptake in the bone and long residence time in the skeleton. In addition to the low oral bioavailability, the food effect thought to be associated with the decrease in solubility of the divalent metal-complexed bisphosphonic acid functionality led to a search for prodrugs masking this group (discussed in the prodrugs of carboxylic acid section below). In contrast to traditional phosphonate esters, the PepT1 recognition element of the Pro-Phe alendronate prodrug (Figure 28.5) was incorporated at the amino end to exploit the PepT1 transporter. Oral administration of the Pro-Phe prodrug to rats resulted in a 3-fold increase in drug levels in the tibia compared to dosing with the parent drug [30]. Relatively high levels of the prodrug also found in the kidney and liver corroborate the observed improvement in absorption of the prodrug. Given that the prodrug has similar lipophilicity and higher molecular

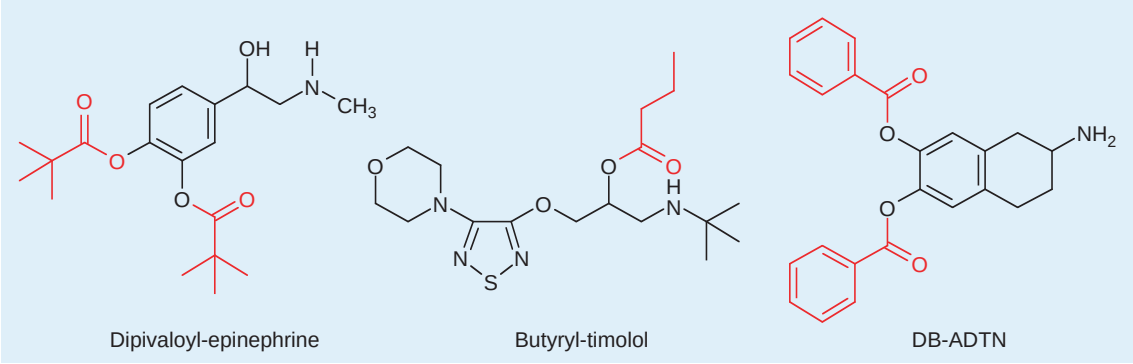


FIGURE 28.6 Lipophilic prodrugs of hydroxyl compounds with facilitated membrane penetration [33–36].

weight than the parent drug, the researchers reasoned the increased absorption of the prodrug could not be due to increased passive diffusion. Indirect evidence supporting the hypothesis that the improved absorption was due to active transport mechanism was obtained by competition studies with Pro-Phe in perfusion and pylorus-ligated rats [30].

As with many other cases, the alendronate prodrug example underscores the importance of both absorption and efficient release of the parent drug. A series of stereoisomeric valine–valine-based dipeptide ester prodrugs of acyclovir were synthesized to determine the effects of chirality on absorption and release. The apparent permeability of the peptide promoiety was found to be modestly dependent on the stereochemistry of the amino acids, with the L-isomer imparting better absorption properties. The bioconversion of the dipeptide prodrug by hydrolytic enzymes, however, is stereospecific and has strong preferential affinity for L-isomers [31,32]. This observation is useful in designing a prodrug targeting PepT1 that balances *in vivo* stability and efficient release of the drug.

Finally, lipophilic carriers and transporter-targeted prodrugs not only improve absorption of poorly absorbed drug molecules but can also serve to reduce first-pass metabolism [33] by masking or blocking the site of metabolism and/or shielding the drug from efflux pumps.

B. Classification of Prodrugs by Functional Groups

The sections below describe the types of promoieties utilized for the various functional groups.

1. Derivatization of Drugs Containing Alcoholic or Phenolic Hydroxy Groups

Starting with hydroxylic derivatives, high lipophilicity can simply be obtained by esterification with lipophilic carboxylic acids. Dipivaloyl-epinephrine, for example (Figure 28.6), crosses the cornea and is used in the treatment of glaucoma [34]. Note that prodrugs that liberate pivalate have been shown to result in toxicity consistent with tissue carnitine depletion impacting carnitine homeostasis. Thus, chronic systemic treatment or high dosage may require carnitine supplementation [37]. The development of pivalate-containing prodrugs should be weighed by the intended use, route of administration, efficacious dose, and superiority versus other potential carboxylates or alternative derivatives. This example highlights the safety considerations of the released by-products when designing prodrugs.

The β-blocker timolol contains a secondary amino group with a pKa of 9.2. Since this group is highly protonated at pH 7.4, the compound shows low lipophilicity at physiological pH (log P = −0.04), which is unfavorable for corneal penetration. The corresponding butyryl ester has an increased lipophilicity (log P = 2.08) and causes a 4–6-fold increase in the corneal absorption of timolol following topical administration to rabbits [33]. In a similar manner, dibenzoyl-2-amino-6,7-dihydroxy-tetrahydronaphthalene (DB-ADTN) reaches the central nervous system (CNS), whereas the parent dopamine agonist ADTN does not [35,36]. For dipivaloyl-epinephrine and dibenzoyl-ADTN, the selective acylation of the phenolic hydroxyl groups was achieved in a strong acidic medium, the amino function being protected by protonation [35,38]. Acylated thymidine analogs, such as 3′-O-hexyl-5′-amino- 2′-deoxythymidine, are prodrugs for topical application against herpes simplex type 1 (HSV-1) viruses [39]. Diacetyl and dipropionyl guanine derivatives given orally to mice provided concentrations of the parent drug that were more than 15-fold higher than those observed after dosing the non-acylated

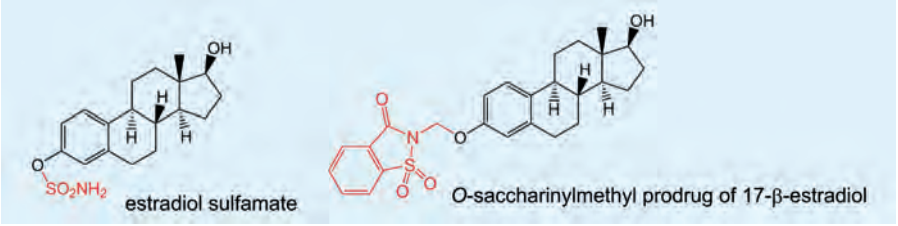


FIGURE 28.7 Prodrugs of steroids.

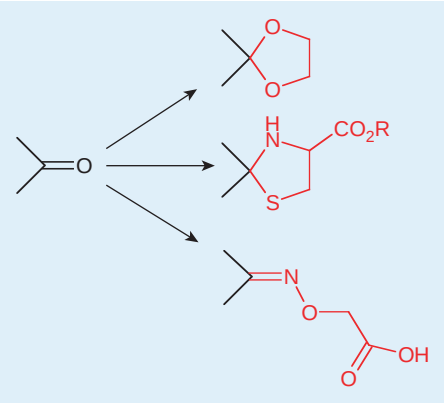


FIGURE 28.8 Prodrug possibilities starting from aldehyde or ketones.

parent drug [40]. In augmenting the lipophilicity and simultaneously destroying the crystal lattice energy, the 2′,3′-diacetate of the antiviral agent 6-methoxypurine arabinoside allowed a 5-fold increase in bioavailability and a 3-fold increase in water solubility in comparison to the non-acetylated drug [41]. As a consequence, an intravenous formulation could be developed.

Many drugs containing a phenolic group [42] cannot be administered orally due to their rapid metabolism in the liver and intestinal tract. Ester prodrugs may undergo rapid hydrolysis by the gut enzymes, thus releasing the active drug prematurely. Alternative prodrug types may be necessary to overcome this problem, and some noteworthy examples are observed in the area of steroids (Figure 28.7). Orally administered estrogens undergo accumulation and extensive metabolism in the liver, resulting in high hepatotoxicity and low bioavailability. Incorporation of a sulfamate group masks 17-β-estradiol [43] so that it is trafficked into the erythrocytes and passes the liver without extraction or hepatic metabolism. The prodrug is subsequently released from the erythrocytes and converted to the parent drug by systemic hydrolysis. Erythrocyte transport of the prodrug is responsible for the improved level of systemic estradiol without adverse effect on erythrocyte functions [43]. Alternatively, an *O*-(saccharinylmethyl) prodrug of 17-β-estradiol designed to undergo chemical hydrolysis by Sn2 mechanism with less sensitivity to enzymatic hydrolysis demonstrated improved oral potency and decreased first-pass metabolism [44,45].

Esterification of hydroxylic functions with polyacids (e.g., succinic acid, phosphoric acid) represents an excellent way to prepare water-soluble prodrugs if poor solubility is limiting absorption (see Chapter 30).

2. *Derivatization of Aldehydes and Ketones*

The ethylene ketal derivative of prostaglandin E2 (dinoprostone) possesses much improved solid-state stability. Functionalized spirothiazolidines of hydrocortisone and hydrocortisone 21-acetate (Figure 28.8), prepared with cysteine esters or related β-aminothiols, have shown improved topical anti-inflammatory activity. It is speculated that a Schiff base intermediate formed upon ring-opening results in the accumulation in the skin by binding (through its SH function) to thiol groups in the skin [46].

Simple and substituted oximes are biostable unless intramolecular assistance is provided. This is the case for the oximes derived from oxyamino acetic acid that are possible water-soluble prodrugs of ketones and aldehydes.

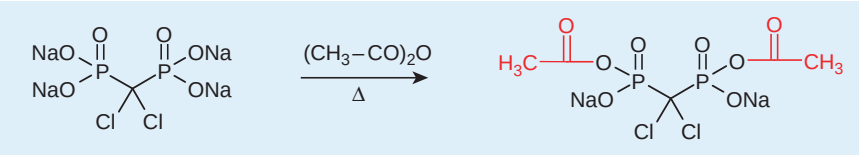


FIGURE 28.9 Use of anhydrides as lipophilic prodrugs of phosphonates.

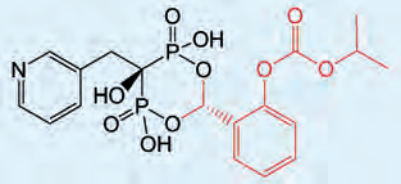


FIGURE 28.10 Novel cyclic acetal prodrugs of bisphosphonate [69].

3. *Derivatization of Drugs Containing a Carboxylic Acid Functional Group*

Lipophilic prodrugs can also be derived from a carboxylic functional group, the most commonly used derivatives being carboxylic esters. Simple esters of aliphatic alcohols are attractive, as they are cheap to prepare, chemically stable, and yield harmless hydrolysis products [47]. Typical representatives of such prodrugs are tyrosine methyl ester [48], levodopa ethyl ester [49], nipecotic acid ethyl ester [50], enalaprilat ethyl ester [51,52], trandolapril [53], γ-aminobutyric acid (GABA) cetyl ester [54,55], and methotrexate cetyl ester [56].

Lipoidal prodrugs, in which the carboxyl function esterifies the free alcoholic hydroxyl of 1,2- or 1,3-diglyceride, are well absorbed and show high lymphotropism [57,58]. Due to their greasy properties and difficult purification, these compounds have limited industrial application.

The widespread use of acyloxymethyl esters in antibiotic chemistry, as illustrated above for bacampicillin, was initiated by Jansen and Russel [59] at Wyeth Laboratories and successfully applied to pivampicillin [14], talampicillin [16], and cephalosporins [60]. In each of these cases, the oral absorption of the antibiotic was improved by some 2–3-fold over that of the parent compound. The acyloxymethyl derivatization was also extended to amino acids such as α-methyldopa [61], isoguvacine [62], or tranexamic acid [63], to anti-inflammatory drugs such as niflumic acid [64] or indomethacin [65], and to quinolone antibacterials such as norfloxacin [66].

Mixed anhydrides represent original attempts to prepare prodrugs of carboxylic or phosphonic acids [67]. Clodronic acid dianhydrides (Figure 28.9), for example, were shown to be novel bioreversible prodrugs of clodronate. They are more lipophilic than the parent clodronate, stable against chemical hydrolysis, and hydrolyze enzymatically to clodronate in human serum [68].

A novel prodrug class introduced by P&G researchers to increase lipophilicity (absorption) and reduce food effects hypothesized to be caused by Ca^{++2} complexation of bisphosphonic acid groups comprises the cyclic acetal prodrugs. Intraduodenal administration of the prodrug (Figure 28.10) to fed rats resulted in greater risedronate exposure compared to dosing risedronate based on urinary recovery [69]. Systemic rate of conversion of the cyclic acetal to risedronate can be modulated by substitution on the phenyl ring. Preliminary experimental studies suggested that oral administration of bisphosphonate cyclic acetals can improve bisphosphonate exposure even under fed conditions, as long as these prodrugs are stable enough in the stomach to ensure premature release of parent drug does not occur.

In contrast to modifications on the bisphosphonate side chain described earlier, Medivir scientists incorporated distil PepT1 recognition elements to the phosphonic acid moiety to mask the highly ionized functionality and promote carrier-mediated transport [70]. Absolute bioavailability (BA) was determined as the ratio of the area under the curves (AUCs) of the prodrug administered orally against the parent drug administered i.v., and the most promising examples are shown in Figure 28.11. These data suggest significant enhanced oral bioavailability of these prodrugs relative to alendronate. The efficiency of bioconversion, however, is not described.

Methoxy-imino bioisosteres of carboxylic acid anhydrides, which can be prepared by *O*-acylation of the corresponding hydroxamates, represent another way to prepare prodrugs of carboxylic acids (Figure 28.12). One of these derivatives, compound FOX 988, a diabetes drug, was designed as a prodrug of 4-(2,2-dimethyl-1-

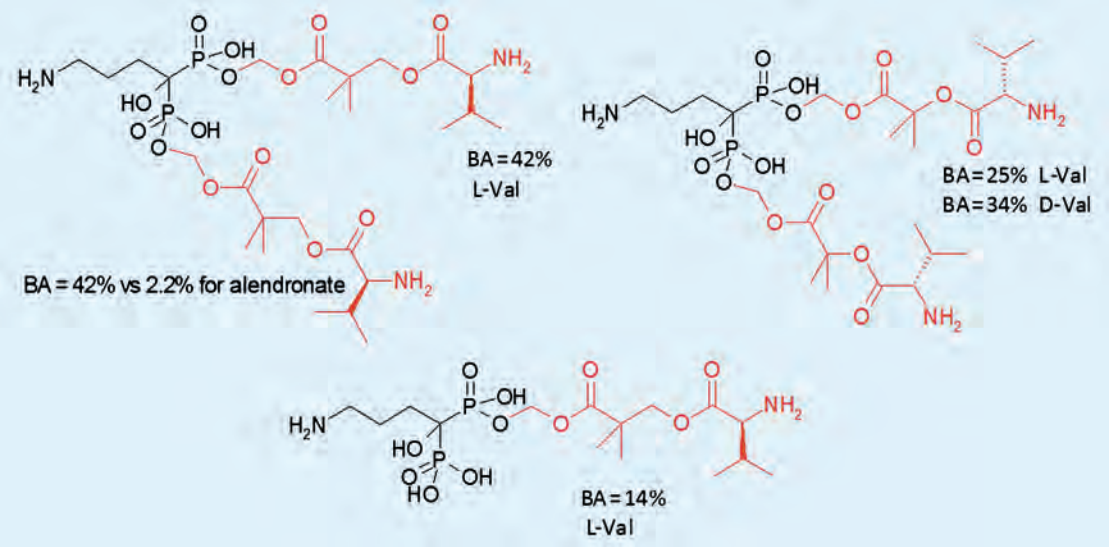


FIGURE 28.11 Acyloxyalkylester prodrugs of bisphosphonate suggestive of PepT1 mediated transport [70].

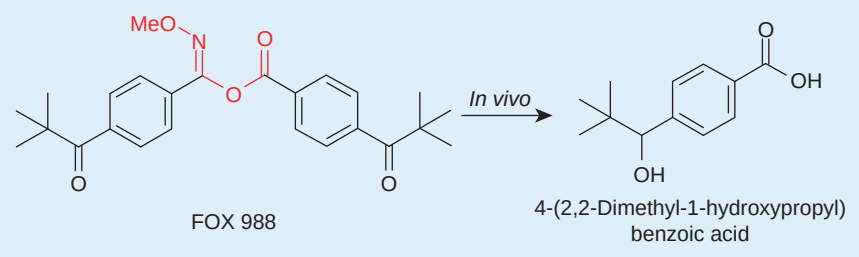


FIGURE 28.12 Methoxy-imino bioisosteres of carboxylic acid anhydrides as prodrugs for carboxylic acids.

hydroxypropyl) benzoic acid. In the liver this prodrug is metabolized at a rate sufficient to possess hypoglycemic potency (an ED_{50} of 65 μmol/kg, 28 mg/kg/day, for glucose lowering) [71] but avoids significant release of the active metabolite to the systemic circulation to avoid testicular toxicity at doses up to 1,500 μmol/kg/day. The mechanism of action of 4-(2,2-dimethyl-1-hydroxypropyl) benzoic acid is the sequestration of coenzyme A (CoA) in the mitochondria, thus inhibiting medium-chain acyltransferase and, as a consequence, causing hepatic neoglucogenesis [72].

Primary amides of carboxylic acids are easily converted in humans to the corresponding acid (e.g., depamide, progabide) by amidases and can thus be used in prodrug design. Amides of ketoprofen-derived arylacetic acids possess a therapeutic index one order of magnitude greater than that of indomethacin [73].

4. *Derivatization of Amines*

Due to the slow *in vivo* cleavage rate of the *N*-substituted amides, acylation of amines is generally not recommended. Better possibilities are offered by activated amides, peptides, imines, and soft quaternary ammonium salts. However, the use of simple *N*-acyl derivatives must not systematically be discarded. The *N*-benzoyl- or *N*-pivaloyl derivatives of the inhibitory neurotransmitter GABA are examples of compounds able to penetrate the blood–brain barrier (BBB) and to abolish pentetrazole- and bicuculline-induced convulsions. It was also demonstrated in rats that, following subcutaneous injection, rat-brain homogenates liberate free GABA from these amides [74]. Some biochemical and pharmacological evidence suggests that *N*-pivaloyltaurine crosses the BBB [75].

The compound XP13512 illustrates a successful example of a prodrug that utilizes carrier-mediated transport to increase bioavailability. The oral bioavailability of gabapentin in animals and humans is not dose dependent due to saturation of its absorption pathway through a low-capacity amino acid transporter localized in the upper small intestine. Consequently, plasma exposure to gabapentin in patients is unpredictable, and a sustained-release

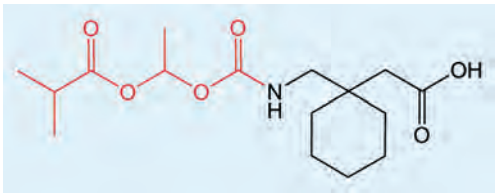


FIGURE 28.13 XP13512 (Horizant), prodrug of gabapentin.

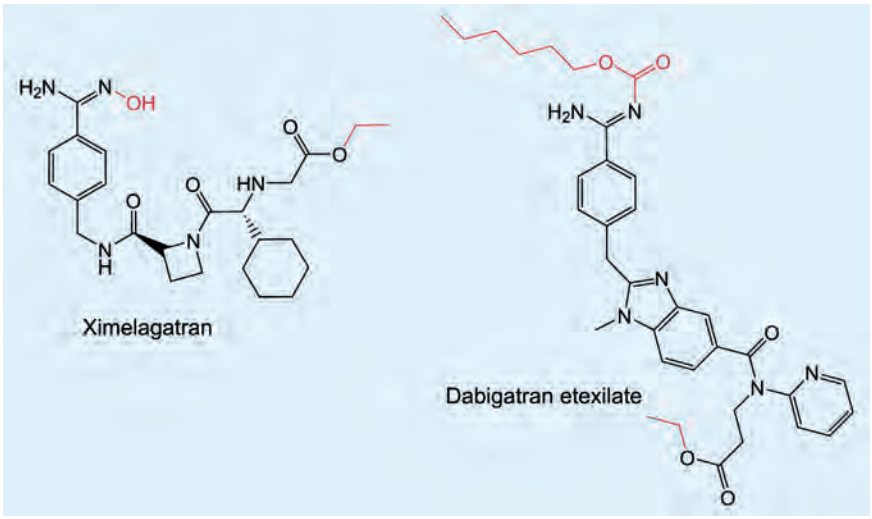


FIGURE 28.14 Recently approved anticoagulants.

formulation of the drug is not available. A carbamate prodrug (Figure 28.13) of gabapentin (XP13512) was designed by Xenoport to be actively absorbed from the intestine by a high-capacity nutrient transporter, MCT-1, which is expressed in all segments of the colon. The prodrug is chemically stable and is rapidly converted to gabapentin by nonspecific esterases following oral absorption to liberate gabapentin, isobutyrate, acetaldehyde, and CO_2. The oral bioavailability of gabapentin and XP13512 at 10 mpk is 54 percent and 94 percent, respectively, in male cynomolgus monkeys [76]. It was shown to be effective in rat models for epilepsy and neuropathic pain. The US FDA approved Horizant (gabapentin enacarbil) extended-release tablets in 2012 for restless leg syndrome.

5. *Derivatization of Amidines*

The amidine moiety is associated with poor pharmacokinetic properties due to its highly charged nature (pKa ~ 12). Consequently, a prodrug approach may be a solution to increase oral bioavailability. A prodrug strategy to mask the polar basic group could improve absorption of the molecule from the gastrointestinal tract into the circulation, where the active drug molecule is then released by chemical or enzymatic cleavage. This approach has been successfully demonstrated in the area of thrombosis by the direct thrombin inhibitors (Figure 28.14) Ximelagatran (Exanta) and, more recently, Dabigatran etexilate (Pradaxa). In the case of Ximelagatran, the bioconversion of the double prodrug to the active component takes place in the liver and many other tissues via ester hydrolysis and reduction. The bioavailability of Melagatran, the active drug molecule, following oral administration of Ximelagatran was 5–10 percent in rats and ~20 percent in humans, with low intersubject variability [77]. Dabigatran etexilate, a recently approved antithrombotic agent, is primarily metabolized by esterase-mediated hydrolysis to Dabigatran. *In vitro* studies using human liver microsomes and specific esterase inhibitors suggest that hydrolysis is mediated by microsomal carboxylesterases. The bioavailability of Dabigatran after p.o. administration of Dabigatran etexilate is low (7.2 percent), but pharmacologically efficacious concentrations of Dabigatran can be achieved [78].

6. *Prodrugs for Compounds with Acidic NH Functions.*

Prodrugs obtained by N-alkoxycarbonyloxymethylation of 5-fluorouracil show improved delivery properties. Both 1- and 3-alkoxycarbonyloxymethyl derivatives are hydrolyzed quantitatively to 5-fluorouracil, but the

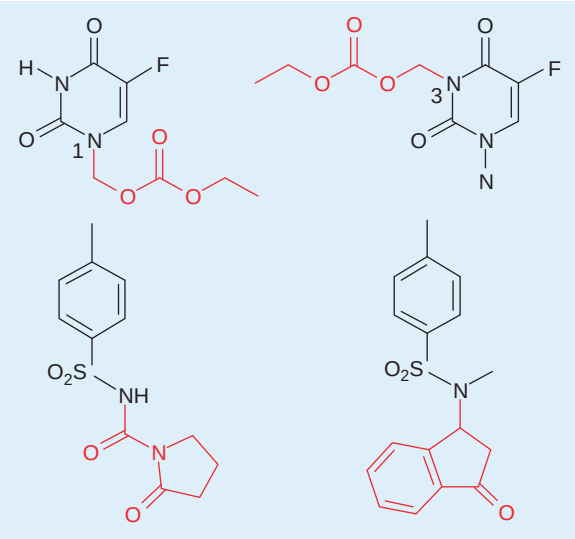


FIGURE 28.15 Prodrugs of acidic NH functions.

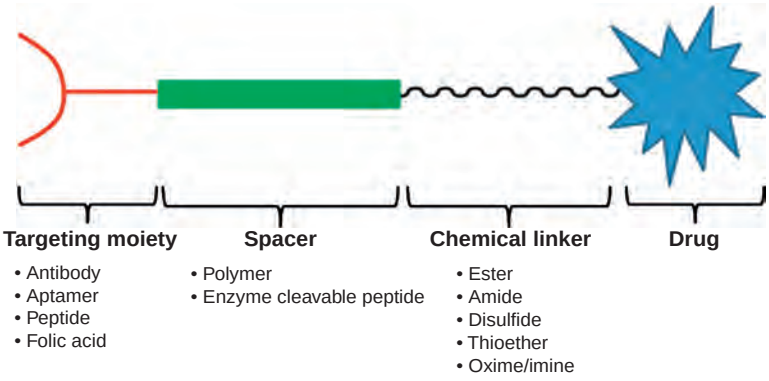


FIGURE 28.16 General design aspects of a targeted drug conjugate.

3-substituted derivatives show a greater promise as prodrugs since they combine adequate stability in aqueous solution with a high susceptibility to hydrolysis in plasma [79]. Sulfonamides, as well as carboxamides, carbamates, and other NH-acidic compounds (Figure 28.15), can be acylated with various groups [80] or converted into phthalidyl derivatives [81].

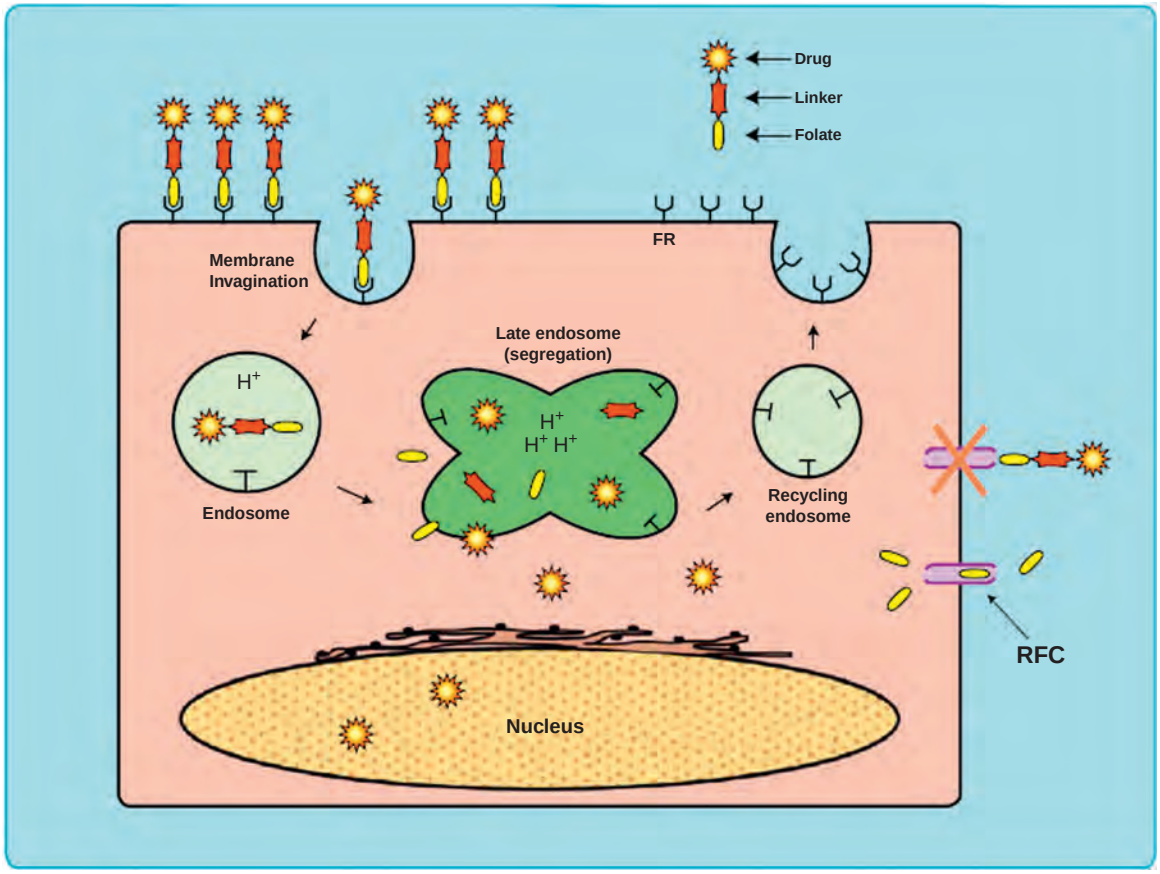
IV. UNIQUE APPROACHES TO CARRIER PRODRUG DESIGN

A. Site-specific Delivery

Tissue-selective drug delivery has enormous potential to improve the safety and efficacy profile of drug molecules. Prodrug strategies have been investigated to target cell/tissue delivery. In principle, two targeting possibilities can be considered: (1) site-directed drug delivery [81], whereby prodrugs/conjugates are transported to target cells/tissues by tissue-specific transporters or cell surface receptors; and (2) site-specific drug release, in which the prodrug is distributed everywhere but is activated by enzymes preferentially localized in target cells/tissues. In either case, there are several components to consider when designing a prodrug, such as the targeting moiety, spacer and/or cleavable linker, and the drug molecule, as illustrated in Figure 28.16.

1. Site-Directed Drug Delivery

Most of the successes in achieving site-directed drug delivery through prodrugs have been through localized delivery of lipophilic prodrugs (eyes, skin) with increased permeability characteristics. Systemic site-directed



SCHEME 28.1 **Representation of targeted delivery and folic acid-drug conjugate uptake via endocytosis.** The folic acid–drug conjugate binds specifically to the FR with high affinity. The plasma membrane invaginates to form an intracellular vesicle that migrates toward the perinuclear region and fuses with other endosomal elements. The mature endosome becomes acidic, and the receptor changes conformation and releases the conjugate, which cleaves to yield the drug molecule [85]. *Reprinted with permission from Elsevier. Copyright 2004.*

delivery, that is, to a specific internal site or organ through a selective transport, is very difficult to achieve. These difficulties (e.g., technical challenges/complexities, cost, scalability) are many fold, notwithstanding the fact that no targeting vector has yet been found to be specific to only one desired site. Local enrichment or privileged entry to a given organ or into the CNS by systemic targeting, however, is possible, and this prospect has the potential to revolutionize medicine. A number of innovative prodrug conjugates targeting specific cells/tissues are in the pipeline, and some have now reached the market. The discussion below is meant to highlight key findings.

A. PRODRUGS AIMED AT TUMOR DELIVERY

Most examples of site-directed delivery research are described in the oncology arena, since the drug molecules are usually indiscriminate cytotoxic agents, and their toxicity to healthy tissues is a major issue. Selective delivery to only cancerous cells would yield a transformative therapeutic. A promising strategy identified by researchers at Endocyte targets folate receptors (FR), which are limited to tissues that are responsible for retention of folates such as kidney, placenta, and liver, and more importantly to specific pathological tissues such as tumors (ovarian, endometrial, renal, lung, and breast carcinomas) and chronic inflammatory sites [82–85]. This differential tissue expression provides the potential for increased therapeutic index and reduced toxicity.

The FR is a glycosylphosphatidylinositol-linked protein that captures ligands from the extracellular environment and transports them inside the cell via a non-destructive, recycling endosomal pathway (Scheme 28.1). The design of the folate–drug conjugates by Endocyte takes into consideration the highly lipophilic nature of cytotoxic agents, which is associated with off-target toxicities due to entrance into all cells and the difficulty of formulation for parental administration. Hydrophilic spacers are incorporated to confer additional aqueous solubility to the water-soluble folic acid. In addition, cleavable linkers (hence prodrugs) between the spacer and the drug

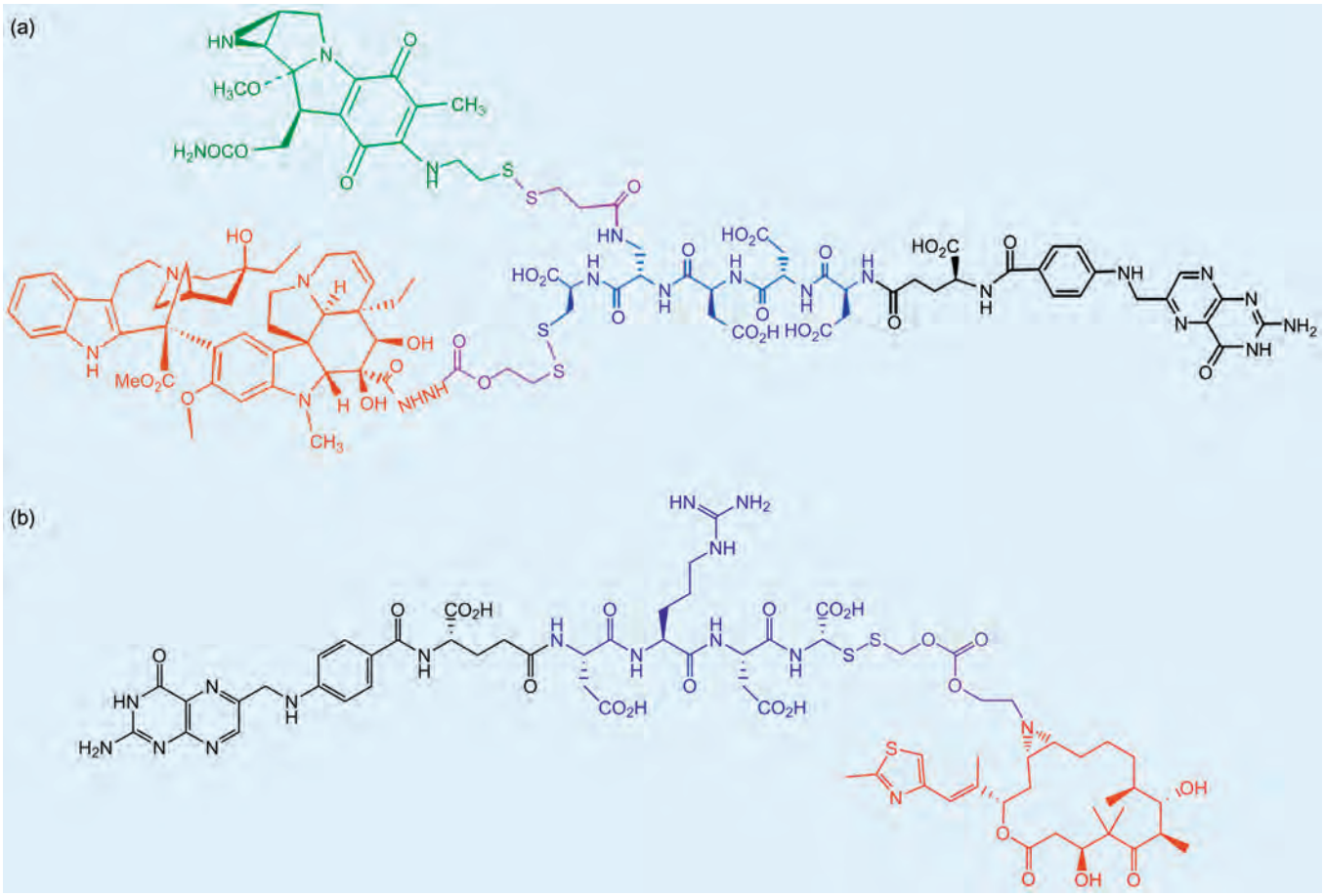


FIGURE 28.17 (a) Chemical structure of dual drug conjugate, EC0255. Folic acid is shown in black, the hydrophilic peptide spacer is shown in blue, the biologically cleavable linker is shown in purple, desacetyl vinblastine is shown in red and mitomycin is shown in green [81–85,87]. *Reprinted with permission from reference. Copyright 2007 American Chemical Society. (b) Epothilone–folate conjugate (epothilone in red).*

were found that effectively release the drug from the targeting component upon entry into the endosome [86]. Consequently a series of folate-drug conjugates have been put forward.

EC145 is the first folic acid–drug conjugate to be evaluated in clinical trials and is being evaluated in Phase III as a treatment for platinum resistant ovarian and lung cancer by Endocyte. The folate–conjugate construct allows tethering of multiple drug molecules via cleavable disulfides, and this concept is demonstrated by EC0225, a conjugate of folic acid and the small molecule anticancer drugs desacetylvinblastine and mitomycin (Figure 28.17a). Following internalization of the prodrug, the high glutathione concentration within the cells reduces the linker and the active drug is released in tumor cells. EC0225 produced potent, dose-dependent activity *in vitro*, and a curative outcome was observed against FR-positive syngeneic and xenograft tumors following an i.v. dosing regimen. The macromolecular prodrug, EC0225, was tested in Phase I study to evaluate safety and tolerability of escalating doses in patients with refractory or metastatic tumors. Along with EC0225, several other antitumor-drug conjugates (e.g., epothilone–folate conjugate; [84] Figure 28.17b) have gone into human trials. Given the large total number of tumors that express folate receptors, FR-targeted prodrug strategies could prove hugely beneficial for patients diagnosed with FR-positive cancer.

Another tumor-targeting strategy utilizes the arginine-glycine-aspartic acid (RGD) peptide motif, which displays a strong binding affinity and selectivity to integrin $\alpha_v\beta_3$. Integrins are a family of heterodimeric transmembrane glycoproteins involved in cell-to-extracelluar matrix and cell-to-cell interactions. The ability of integrins to be internalized by cells on activation with anchoring ligands suggests they can be used to facilitate the delivery of therapeutics. Integrin $\alpha_v\beta_3$ plays a critical role as an indicator in tumor angiogenesis, tumor progression, invasion, and metastasis. They are overexpressed on proliferating endothelial cells as well as on various tumor cell types including breast, glioma, melanoma, prostate, and ovarian carcinoma, while being absent in pre-existing endothelial cells and normal tissues [88–91]. Given these attributes, integrin $\alpha_v\beta_3$ may be an ideal target for chemotherapeutics.

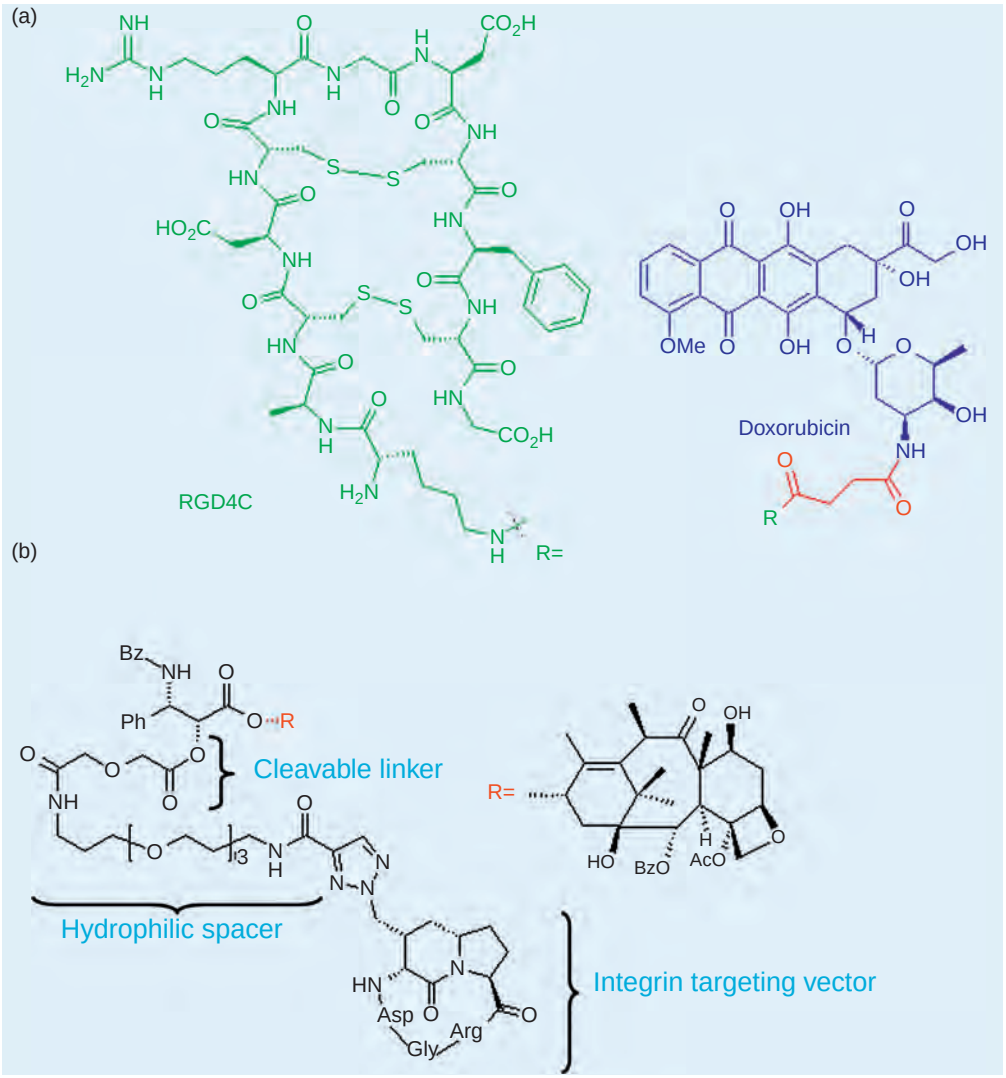


FIGURE 28.18 Integrin-targeting prodrugs utilizing the RGD peptide motif [90,93].

The most-studied integrin-targeting prodrug involves the anticancer drug doxorubicin (DOX) conjugated with the RGD peptide sequence [90]. X-ray crystal structures of the integrin $\alpha_v\beta_3$ complexed with RGD ligands [92] have allowed optimization of the ligand to bind preferentially to the $\alpha_v\beta_3$ integrin family and helped identify regions suitable to link the drug molecule without affecting affinity. Induction of multidrug resistance by doxorubicin in combination with its nonspecific toxicities has restricted DOX-based chemotherapy, and thus a targeting approach may expand its use. Among the different RGD peptide analogs, doxorubicin–RGD4C conjugate (Figure 28.18a) improved inhibition of tumor growth and spreading of pulmonary metastases with less toxicity to liver and heart than free doxorubicin in a mouse MDA-MB-435 breast cancer model in which integrin $\alpha_v\beta_3$ is expressed by the endothelium in angiogenic blood vessels and by the tumor cells [90].

Besides doxorubicin, paclitaxel was covalently attached via a cleavable ester linkage to an integrin-recognizing cyclic RGD vector (Figure 28.18b). The ability of the conjugate to inhibit growth of tumor cells was assessed using a growth inhibition assay in which the effect of free paclitaxel was compared. The conjugate demonstrated low nanomolar IC_{50} values that are comparable or even superior to paclitaxel [93]. In an ovarian carcinoma model xenografted in immunodeficient mice, the targeted prodrug administered i.v. every four days for four times attained better antitumor activity than paclitaxel, as shown by marked induction of aberrant mitoses consistent with the mechanism of action of the drug [93]. Whether the drug is prematurely released outside the cell and then internalized through cell membrane diffusion or is actively targeted and released inside the cell by integrin-mediated endocytosis could not be concluded from their studies.

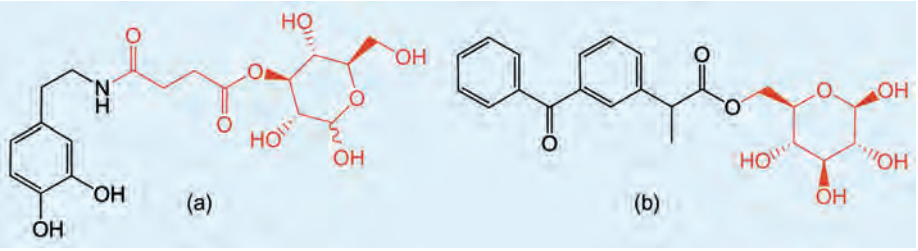


FIGURE 28.19 GLUT-1 targeting sugar conjugates.

Besides targeting integrin $\alpha_v\beta_3$ for drug delivery to tumor cells, the epidermal growth factor receptor (EGFR) is another potential site to target. EGFR is overexpressed in a number of epithelial malignancies and consequently was used for EGFR-targeted therapy to improve antitumor efficacy. Using an EGFR-binding peptide (NH_2-CMYIEALDKYAC-CO_2H, EBP) and cleavable ester linker to carry doxorubicin (DOX), Ai et al [94,95] demonstrated the significantly higher survival rate of tumor-bearing mice treated with DOX-EBP compared to free DOX. They further demonstrated the lower distribution of DOX in normal tissues, in particular with no significant accumulation in the heart which is especially sensitive to DOX toxicity.

As these types of conjugates enter clinical trials, it remains to be seen whether prodrugs aimed at targeting tumor cells will lead to a viable transformative therapeutic, but clearly the area of oncology is where this approach has the potential to make a tremendous impact.

B. PRODRUGS AIMED AT BRAIN DELIVERY

Crossing the BBB remains a key obstacle in the development of drugs for brain diseases. The capillaries of the brain have evolved to constrain the movement of molecules and cells between blood and brain, providing a natural defense against circulating toxins. The relative impermeability of the BBB is due to tight junctions between capillary endothelial cells that are more tightly arranged as compared to normal capillaries. Moreover, brain endothelial cells also possess few alternate transport pathways, unlike those operating in the GI tract, and express high levels of active efflux transport proteins (e.g., P-glycoprotein (P-gp)), Multidrug Resistance Protein-1 (MRP-1), and Breast Cancer Resistance Protein (BCRP)) to remove foreign molecules [96]. The BBB also has additional enzymatic processes that serve to protect the brain. Thus, creative solutions are required to overcome these obstacles and improve brain penetration.

Two interesting prodrug strategies utilize transporter-mediated and receptor-mediated delivery. GLUT-1 is a transporter that mediates the uptake of glucose from the bloodstream to the CNS and is located in the membrane of brain capillary endothelial cells. L-Dopa, a dopamine precursor that (unlike dopamine) crosses the BBB, is the standard treatment for Parkinson's disease. Unfortunately, the premature conversion of L-dopa to dopamine in the peripheral tissues is the cause of unwanted side effects. Thus a dopamine prodrug using a succinyl linker to D-glucose was prepared (Figure 28.19a) [97]. Using human retinal pigment epithelium cells (HRPE) as a model for cells lining the BBB (because these cells also express GLUT1), Pavan et al [97] showed that uptake of the prodrug was greatly reduced in the presence of glucose, whereas dopamine had no effect. Transporter-mediated uptake may explain why glu-dopamine is able to be absorbed into the CNS from the bloodstream while dopamine is not. Gynther et al [98] demonstrated that the ketoprofen-D-glucose conjugate (Figure 28.19b) possesses good chemical stability in aqueous solution at pH 7.4, its uptake by the brain is mediated by GLUT-1, and the ester prodrug undergoes bioconversion to ketoprofen and glucose in the brain tissue. The conjugate, however, undergoes extensive enzymatic hydrolysis in the liver therefore making it inadequate as an effective oral prodrug for brain delivery. In both examples, linker optimization for the precise release of the drug is necessary for this strategy to fulfill the requirements for brain delivery.

A recent example utilizing receptor-mediated endocytosis to improve brain exposure of poorly absorbed drugs was discovered by Angiochem. Their EPiC (Engineered Peptide Compound) platform is based on a proprietary 19-amino-acid peptide (Angiopep-2) shown in animal models to cross the BBB by transcytosis via low-density lipoprotein receptor-related protein (LRP-1) expressed on the surface of the BBB. ANG1005 (Figure 28.20), their first NCE, is the first agent to reach clinical trials for the treatment of primary and secondary brain cancers. In Phase I human clinical trials, ANG1005 reached therapeutic concentrations in brain tumors and produced significant antitumor responses in patients with recurrent gliomas or brain metastasis who had failed with prior standard therapy

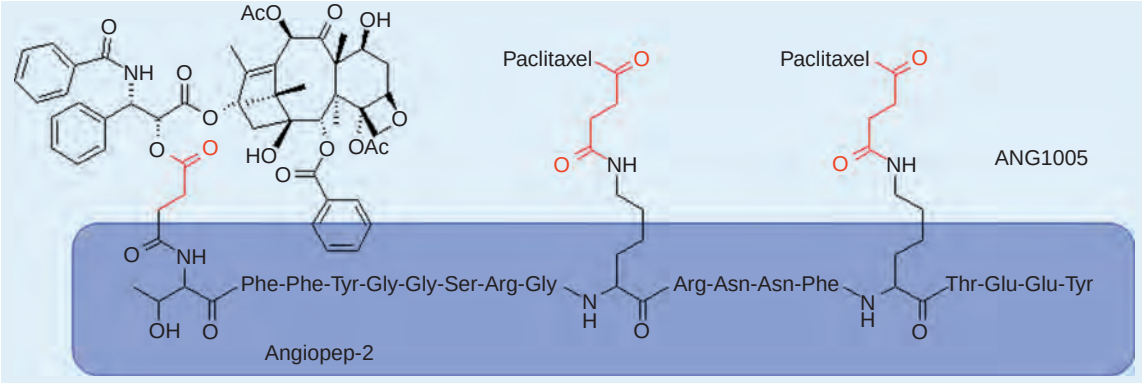


FIGURE 28.20 ANG1005 combines Angiopep-2 with three molecules of paclitaxel using cleavable succinyl linkers (in red).

[99]. Phase II studies of ANG1005 were discontinued due to the inability to achieve a safety window at the dose required to observe intracranial response. In the interim, Angiochem demonstrated the broad applicability of their platform to increase brain penetration by conjugating two other drugs, doxorubicin and etoposide, with Angiopep-2 [100]. A key feature of Angiopep-2 is the high ligand efficiency of Angiopep-2; that is, three drug molecules can be loaded onto one peptide, and it is amenable to combination drug therapy. It remains to be seen whether this targeting peptide prodrug will make it through clinical studies to become a marketed drug.

C. OCULAR DIRECTED PRODRUGS

Optimal delivery of ophthalmic drugs depends on whether the disease affects the anterior or posterior segment of the eye. A noninvasive topical delivery is the favored route for diseases of the anterior of the eye, but poor corneal penetration and loss of drug due to eliminating mechanisms, such as blinking, tear turnover, absorption into intraocular tissues leading to systemic uptake via the nasolacrimal duct, and metabolism by enzymes in the anterior, result in poor bioavailability with < 5 percent reaching the target site [101]. Despite this, treatment for diseases of the anterior segment of the eye is less problematic than the posterior segment. Since drug availability in the anterior segment is governed by ocular barriers and the physicochemical restrictions of the drug molecule, modulating the properties of the drug—perhaps through the design of prodrugs—developing optimum formulations, and utilizing alternative routes of administration can be sufficient to achieve efficacy.

Access to the posterior segments of the eye requires site-specific drug delivery systems to target the vitreous cavity, retinal pigment epithelium (RPE), and choroid [101]. Topical delivery does not provide enough drug to this region, and even systemically administered drugs have limited access to the retina and vitreous due to the blood−retinal barrier (BRB). The BRB is the major barrier to drug delivery to the posterior chamber of the eye. The BRB is composed of the RPE (outer BRB) and endothelial cells of the retinal blood vessels (inner BRB). Transporters expressed in the corneal epithelium and RPE play vital roles in mediating the transport of drugs, xenobiotics, and nutrients into and out of the anterior and posterior regions of the eye (Figure 28.21) [102].

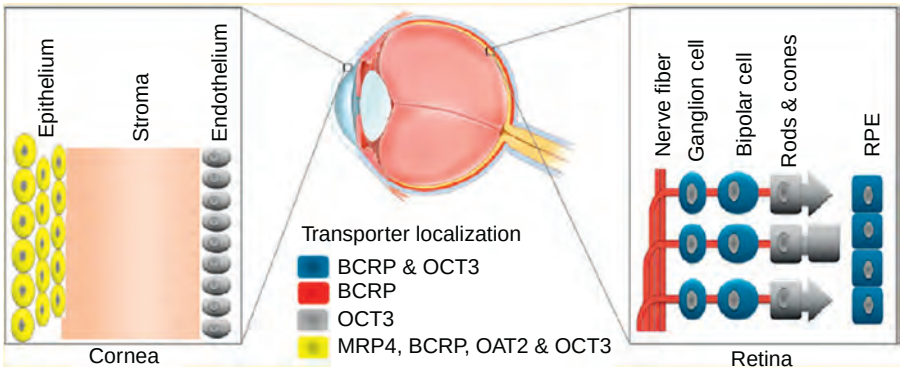


FIGURE 28.21 Barrier epithelial of the cornea and retina and transporter localization [102]. *Reprinted with permission from reference. Copyright 2013 American Chemical Society.*

Recently, Giacomini et al [102] applied array gene expression-profiling techniques to establish the identity and localization of transporters, drug-metabolizing enzymes, transcription factors, and nuclear receptors within the

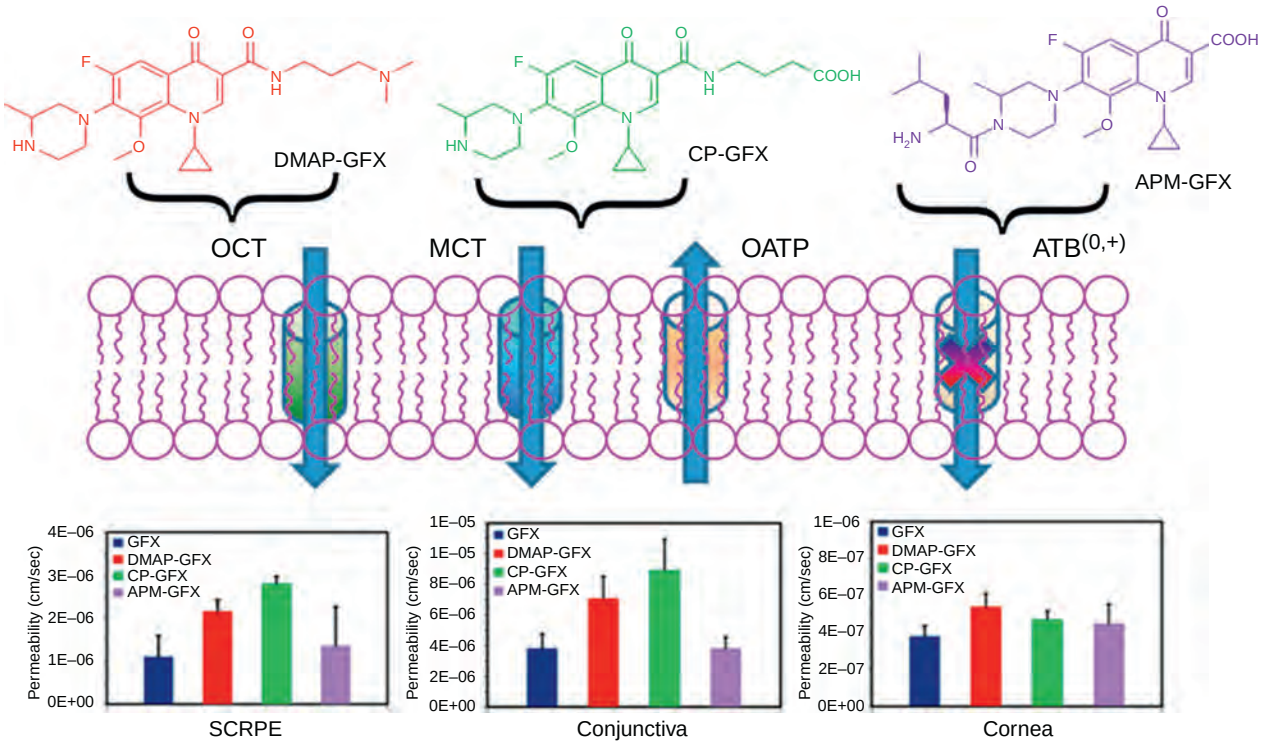


FIGURE 28.22 Permeability profiles of gatifloxacin prodrugs DMAP-GFX, CP-GFX, and APM-GFX designed to target OCT, MCT, and $ATB^{0,+}$ transporters, respectively [104]. *Reprinted with permission from reference. Copyright 2012 American Chemical Society.*

various tissues of the human eye, with particular focus on the cornea and retina. In addition to the peptide and amino acid transporters, such as PepT1, $ATB^{0,+}$, LAT1, and ASC previously discovered on the corneal epithelium [101,103], the organic anion transporter OAT2 and the organic cation transporter OCT3 were found to be localized in the corneal epithelium. Given its role in sensory detection and processing functions, the retina expresses high levels of genes for neurotransmitter transporters and the SLC25A family of mitochondrial carriers. Besides localization in the corneal epithelium and endothelium, OCT3 was also found to be expressed in the RPE, rod, and cones [102]. Along with the transporters MCT, PepT1, and the FR, which are located in the RPE, these influx transporters on the retina and BRB may be exploited to attain higher ocular bioavailability by using prodrugs for enhanced drug delivery in conjunction with optimized modes of administration (intravitreal, intravenous, or transcleral/subconjunctival). Additionally, the derivatization of the parent drug to target influx transporters may also help to evade efflux mechanisms. Careful consideration, however, is needed to ensure that systemically delivered prodrugs reach the posterior chamber with minimum uptake by other tissues also expressing these transporters.

Kompella et al [104] designed prodrugs of gatifloxacin, a fourth-generation antibiotic used to prevent postoperative endophthalmitis after ocular surgery. Gatifloxacin (GFX) achieves rapid and effective vitreous levels after oral dosing. It was removed from the market, however, due to its adverse side effects of hypo- and hyperglycemia. Topical application did not achieve therapeutic levels needed, as endophthalmitis is associated with posterior tissues of the eye. Thus the researchers devised prodrugs based on transporters (OCT, MCT, $ATB^{0,+}$) located in the cornea, conjunctiva, and RPE to improve GFX delivery to the back of the eye. *In vitro* permeability across the cornea, conjunctiva, and sclera-choroid-retinal pigment epithelium (SCRPE) was measured using an Ussing chamber set up, and the mechanism was determined by comparing transport in the presence of transporter specific inhibitors. DMAP-GFX exhibited 1.4-, 1.8- and 1.9-fold improvement in permeability across the cornea, conjunctiva, and SCRPE, respectively, compared to GFX, whereas APM-GFX showed no improvement after topical administration (Figure 28.22) [104]. Moreover, competitive inhibition studies indicated OCT-mediated transport by DMAP-GFX. In an *in vivo* setting, 1.95- and 3.6-fold higher levels of DMAP-GFX were only observed in the choroid-retinal pigment epithelium and vitreous humor, respectively. In all remaining tissues, drug and prodrug dosing resulted in similar exposure. Note that no significant bioconversion was observed in *in vitro* transport studies. but in *in vivo*, significant conversion was observed. This work provides the first example of ocular prodrugs directed at MCT and OCT transporters.

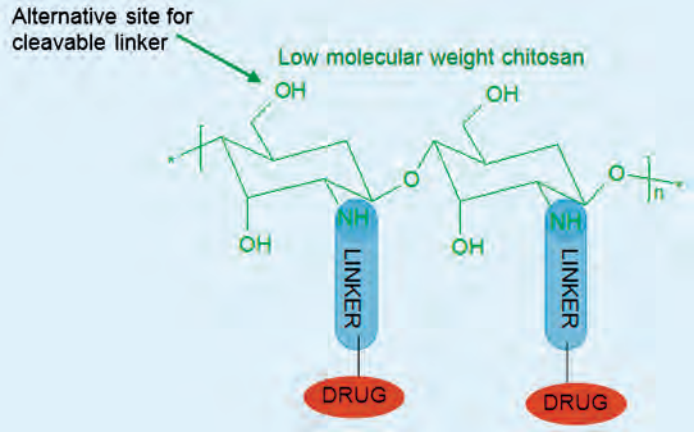


FIGURE 28.23 Chitosan prodrug concept.

D. KIDNEY DIRECTED PRODRUGS

Kidney-targeted delivery is important when trying to reduce extrarenal toxicity of the drug and to improve its therapeutic efficiency for diseases of the kidney. The mesangial cells of the glomerulus, the proximal tubular cells, and the interstitial fibroblasts are principal targets for renal drug delivery since they play a key role in many disease processes in the kidney. Drug delivery to the proximal tubular cell has been the most extensively studied, and of the two main prodrug approaches, the use of low-molecular-weight protein (LMWP) appears most promising. The LMWP approach is based on drug attachment to a protein (<30kD) that is freely filtered through the glomerulus and accumulates specifically in the kidney, in particular in the proximal tubular cells, through a luminal reabsorption mechanism and is stable in circulation but is digested in the lysosomes of the proximal tubular cell to release the drug [105]. Among the different types of LMWPs, lysozyme, cytochrome-c, aprotinin, and chitosan have been shown to be potential renal-specific drug carriers (Figure 28.23). Chitosan has been extensively used in the pharmaceutical field because it is biocompatible, biodegradable, and nontoxic. Low-molecular-weight chitosan is especially useful because it can quickly and reversibly open the tight junctions between intestinal epithelial cells [106], a useful feature for a drug given by oral route of administration. In a recent example, Zhang et al reported that randomly 50-percent N-acetylated low-molecular-weight chitosan (LMWC) selectively accumulated in the kidneys, especially in the renal tubular cells [107,108]. Subsequently, prednisolone was conjugated to the 19 kD molecular weight LMWC via a succinic acid spacer, and the distribution of the prodrug was found to be 13-fold higher in the kidney than the free prednisolone. In a follow up study, it was found that the site-specific uptake of LMWC was mediated by the megalin receptor, whose ligand shares a similar glucosamine unit with LMWC [109]. Using LMWPs as drug carriers, higher renal selectivity can be achieved, and a broad range of drug molecules can be linked as the physicochemical properties of the large protein overrides the properties of the attached drug. Efficient release of the drug from the chitosan conjugate was not described, but it is critical if this approach is to be a feasible kidney-targeting delivery platform.

E. LIVER DIRECTED PRODRUGS

Coupling of drugs to modified bile acids was proposed for liver specific targeting [110]. The rationale is based on the recognition of bile acid-linked drugs by the endogenous bile acid transport system. Chlorambucil (an alkylating cytostatic agent), HR-780 (an inhibitor of hydroxymethylglutaryl-CoA reductase), and an oxaproline peptide (an inhibitor of prolyl-4-hydroxylase) were chosen for conjugation to bile acids (Figure 28.24).

A recent example using a dendrimer as drug carrier exploits a galactose receptor-mediated liver-targeting prodrug with a pH-triggered drug release (Figure 28.25) [111]. Polyamidoamine dendrimer (PAMAM)-drug conjugates with a galactosylated PEG could reach the hepatocytes, due to the affinity of galactosyl residues to the asialo-glycoprotein receptor, where it can be endocytosed and the drug released under the acidic (pH 5.6−6.5) conditions of the lysosomes. The enhanced accumulation and prolonged retention time in hepatoma tissue of a Gal-PEG-b-PAMAM-Doxorubin conjugate using magnetic resonance imaging technique suggests its potential use as an antitumor agent. Because of their monodispersity, multiple sites for drug attachment and vectorization, well-defined size, and shape, dendrimers have received a great deal of attention as targeted-drug-delivery systems [112]. The low solubility of dendrimers can be addressed with the use of PEG to improve water solubility, but the cytotoxicity concerns of dendrimers in general still needs to be addressed.

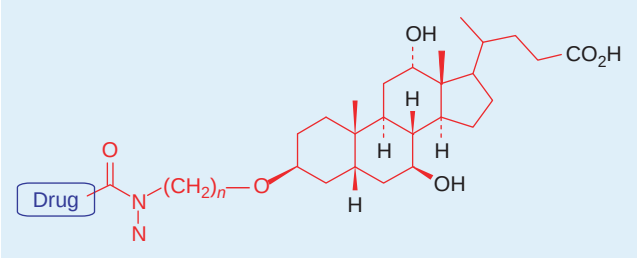


FIGURE 28.24 Bile acids for liver-specific targeting [110].

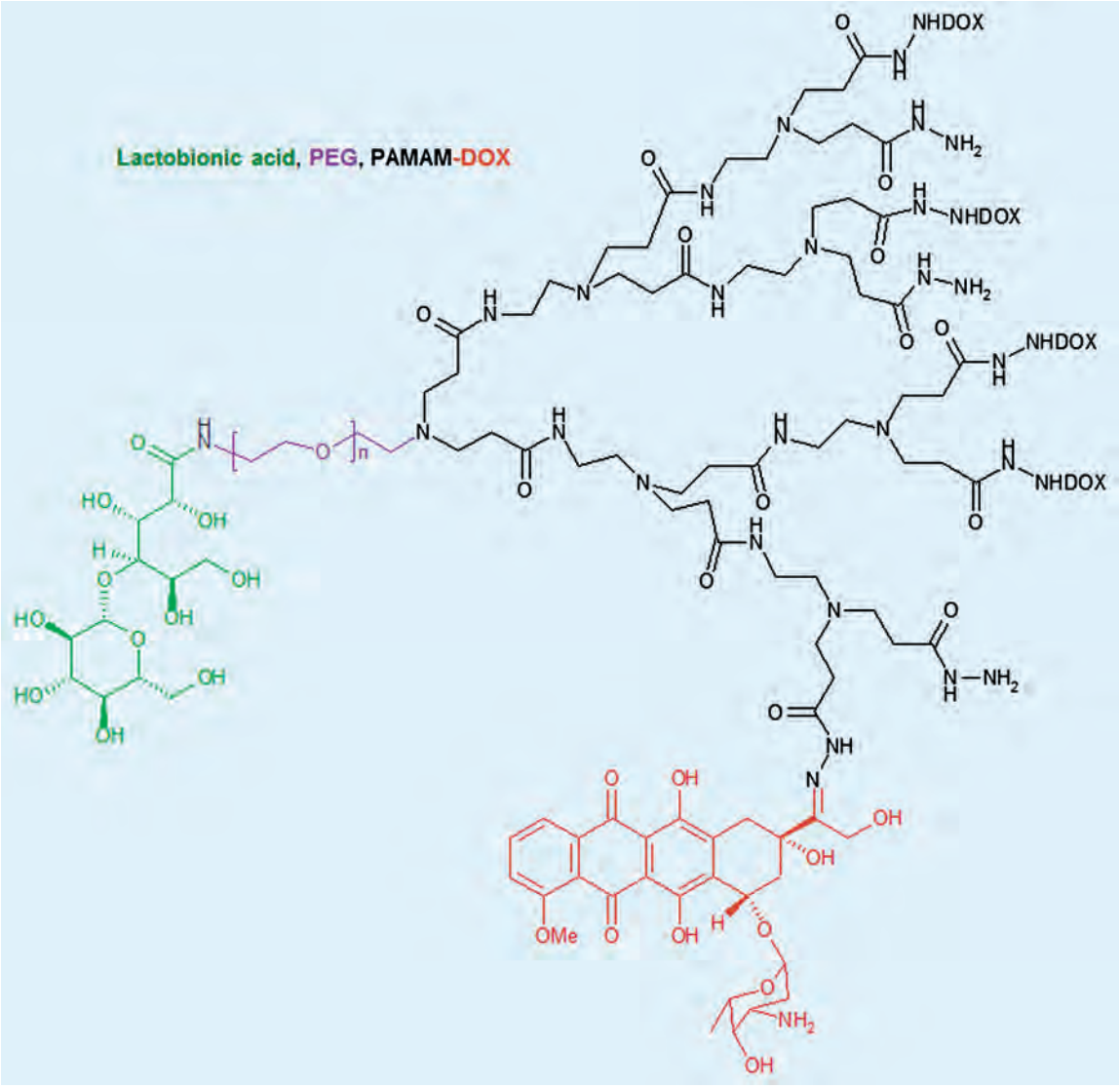


FIGURE 28.25 Gal-PEG-b-PAMAM-DOX$_n$ liver targeting prodrug with pH sensitive hydrazone linker [111].

A simpler liver-targeting prodrug uses the thiazolidine-4-carboxamide group (Figure 28.26) to enhance hepatic exposure of amantadine, an antiviral agent against hepatitis C virus [113]. A comparison of the thiazolidine prodrug and bile acid prodrug of amantadine reveals a reverse pH profile. In general, the thiazolidine prodrug was shown to be susceptible to degradation in acidic media, while only modest degradation of the bile acid prodrug occurred at the higher pH 8. The rate of release of amantadine was 3.5 times higher in rat liver homogenate

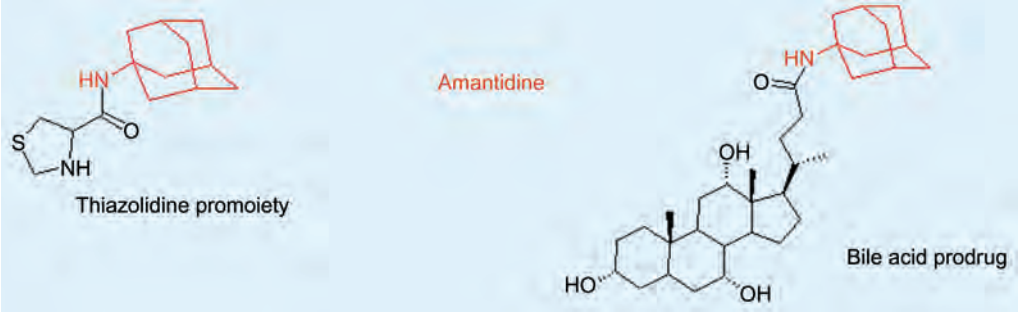


FIGURE 28.26 Liver-targeting prodrugs of Amantidine.

compared to human plasma for both prodrugs, with the efficiency of release in rat liver homogenate being higher for the thiazolidine prodrug. *In vivo* pharmacokinetic study of amantidine versus the thiazolidine prodrug showed sustained levels (T = 1−10 hr) of the parent drug (amantidine) in the liver with i.g. dosing of the prodrug in contrast to dosing the parent drug itself. AUC measurements of amantidine comparing prodrug with amantidine dosing were not provided, but the time-course exposure profile of the parent drug after prodrug administration did not appear particularly compelling. A pharmacokinetic study comparing the drug exposure levels of the two carrier systems could be instructive. These studies, however, demonstrated the possibility of localized delivery of a drug to the liver with this promoiety.

2. *Site-Specific Drug Release*

Targeting drugs to specific organs, tissues, or cells is an attractive strategy for improving drug efficacy and reducing side effects. A complementary approach is to use prodrugs, which distribute widely but cleave intracellularly to the active drug by a tissue-specific enzyme. The whole strategy of site-specific release of a given drug lies in the identification of an enzyme present in high concentrations in the target tissue and nearly absent elsewhere. In addition, the enzyme is required to efficiently cleave a broad range of structurally diverse substrates in order for this approach to be generally applicable. An appropriate prodrug can then be designed using the selective cleavage possibility offered by the enzyme. The examples below demonstrate how this approach has been used to target the liver, kidney, and tumor cells.

A related concept that should be mentioned is antibody-directed enzyme-prodrug therapy (ADEPT) [114]. This approach utilizes an antibody or antibody fragment directed at a tumor-associated antigen to carry an enzyme that has no human homologue to cancer sites. After clearance of the enzyme in the blood, a nontoxic prodrug that is a substrate for the enzyme is given, and a potent cytotoxic agent is released within the tumor site by the catalytic action of the enzyme.

A. RELEASE IN THE LIVER

Attempts to target the liver using liver-specific receptors and transporters have not been as successful as utilizing liver-specific enzymes such as the cytochrome P450 (CYP) enzymes. The liver is the major site for CYP mediated reactions, although CYP enzymes are also expressed to a lesser extent in adrenal gland, small intestine, brain, and kidney [87]. Attention must also be given for potential interpatient variability in exposure of the drug due to differences in expression levels of the enzyme among patients and also for drug−drug interactions that may occur through inhibition or induction of CYPs, especially CYP3A4. These aspects that could affect prodrug conversion must be taken into account when developing prodrugs activated by CYPs. A novel class of phosphate and phosphonate prodrugs, HepDirect prodrugs, are cyclic 1,3-propanyl esters containing a ring system susceptible to oxidative cleavage by CYP3A4. Hydroxylation of the C4 methine results in irreversible ring opening and formation of a transient intermediate that undergoes β-elimination to release the phosphate or phosphonate and a vinyl ketone by-product that is rapidly detoxified by intracellular glutathione. Since the released drug has poor cell permeability, it remains localized in the heptacytes. Pradefovir is the most advanced prodrug utilizing this approach to deliver adefovir to the liver for chronic hepatitis B (Figure 28.27). In a Phase IIb clinical trial, 71 percent of patients treated with pradefovir had undetectable levels of viral DNA, approximately double what was seen with a maximal dose of the marketed hepatitis B drug Hepsera [115]. The HepDirect prodrug demonstrated a higher liver-tokidney exposure ratio which translated to better efficacy in heptatitis B patients [116]. Clinical Phase III study of adefovir had revealed kidney toxicity, which limited its use.

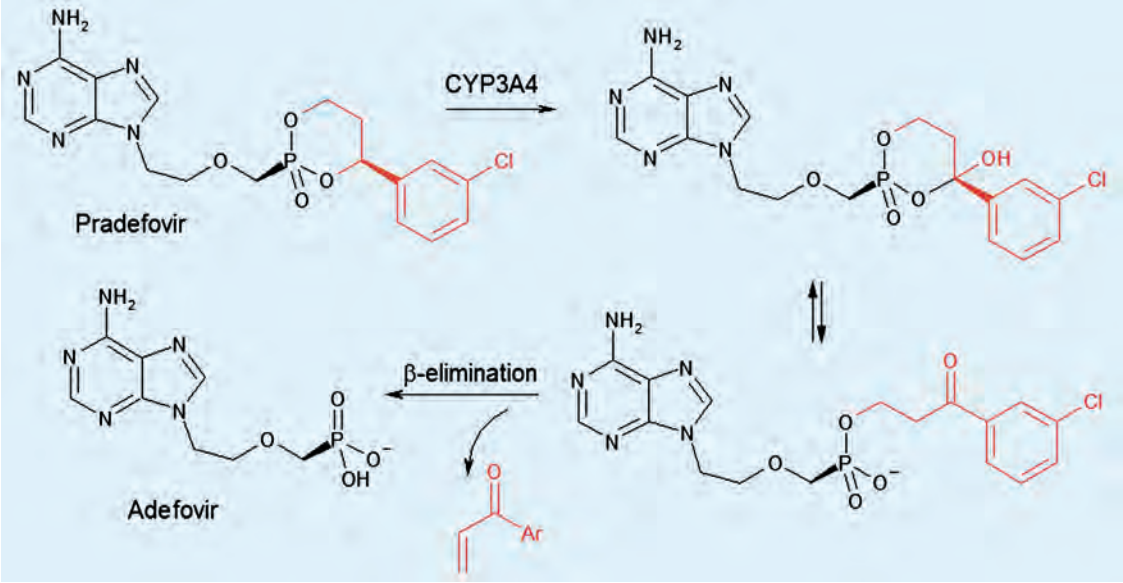


FIGURE 28.27 HepDirect prodrug Pradefovir [87].

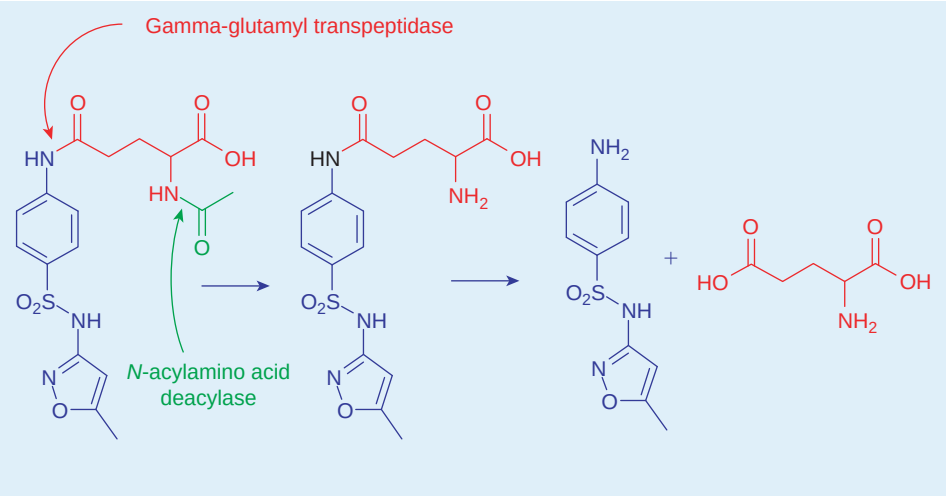


FIGURE 28.28 Kidney-selective release of sulfamethoxazole [117].

B. RELEASE IN THE KIDNEY

It is possible to obtain a kidney-selective accumulation of sulfamethoxazole by administering the drug in the form of N-acetyl-γ-glutamate [117]. The regeneration of the free sulfamide requires the initial deacylation of the glutamic moiety by an N-acylamino acid deacylase, which is also present in the kidneys in high concentrations (Figure 28.28). The γ-glutamyl strategy for confining drug action to the kidney and the urinary tract requires that the prodrug under consideration function as a substrate for γ-glutamyl transpeptidase and, eventually, for N-acylamino acid deacylase.

C. RELEASE IN TUMOR SITES

Enzyme-prodrug approaches to cancer therapy have the potential to achieve tumor-selective drug delivery resulting in less toxic side effects, more effective antitumor activity, and perhaps the delay of the development of drug resistance. Three examples that highlight this concept are the use of uridine phosphorylase, legumain, and β-D-glucuronidase cleavable prodrugs.

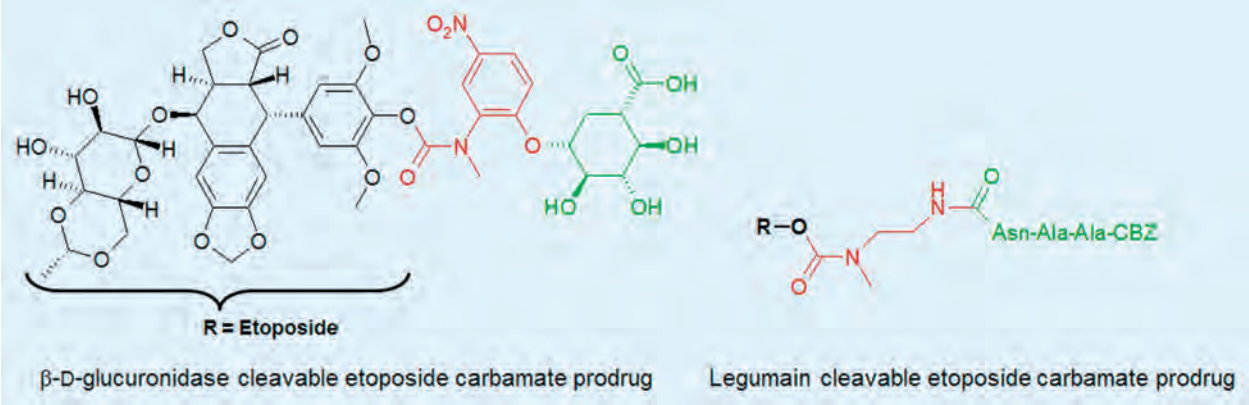


FIGURE 28.29 Enzyme-prodrug approaches to cancer therapy.

In various tumor tissues, the activity of the enzyme uridine phosphorylase is markedly higher than in the surrounding normal tissues. This observation prompted the synthesis of 5-fluorouracil prodrugs. Among them, 5′-deoxy-5-fluorouracil shows high antitumor activity and less host toxicity compared to fluorouracil. This favorable therapeutic index is attributed to a preferential bioactivation by uridine phosphorylase in tumor cells [118,119].

Legumain, a lysosomal endoprotease, has been reported to be overexpressed in the majority of human solid tumors, promote cell migration, enhance tissue invasion and metastases, and cleave asparaginyl bonds. Thus, the legumain-cleavable carbamate ester prodrug of the antineoplastic agent Etoposide (Figure 28.29) was synthesized [120]. The tripeptide legumain substrate was conjugated through an amine spacer to provide the prodrug. The prodrug, however, failed to inhibit the proliferation of 293 HEK cells, which were stably transfected to overexpress legumain constitutively. It was not established whether this failure was due to the inability of the ethylenediamine-etoposide (legumain-cleaved product) to exit the lysosome and reach the cytoplasm or that cyclization did not occur to release etoposide. Another possibility is that the prodrug was not able to enter the cell and therefore did not inhibit cell proliferation. It is also possible that the prodrug did internalize in the cell but failed to enter the lysosome in order to be cleaved. The prodrug, however, was shown to be cleaved by overexpressed and endogeneous legumain. Confocal microscopy study is needed to understand the subcellular trafficking and localization of the prodrug.

A glucuronide based prodrug of etoposide relies on β-D-glucuronidase released extracellularly in high local concentration by necrotic tumors to selectively liberate the active drug. Besides this site, β-D-glucuronidase is localized inside the lysosome, thus allowing a degree of selectivity. The prodrug designed by Schmidt [121] (Figure 28.29) consists of etoposide linked to the β-D-glucuronic group by a self-immolative spacer. *In vitro* kinetic study using Escherichia coli β-D-glucuronidase demonstrated fast enzymatic cleavage of the prodrug ($T_{1/2} < 25$ min) and rapid cleavage of the spacer. Since the concentration of the enzyme in the necrotic area may be variable, it is difficult to predict whether the concentration required for efficient cleavage of the etoposide prodrug will be reached in an *in vivo* setting. Further work is needed to establish the true utility of this approach.

β-Galactosidase responsive prodrugs have been designed for solid tumors in the context of an ADEPT approach. A macromolecular prodrug designed to take advantage of specific tumor-associated receptors and activation by lysosomal enzymes inside malignant cells represents a "simpler" approach to ADEPT [122]. A novel drug-delivery system composed of a galactoside trigger, a FR-targeting ligand, and a potent antimitotic compound MMAE conjugated to a self-immolative linker (Figure 28.30b) was investigated by Papot [122]. Selectivity in this case is achieved by targeting the increased FR expression on tumor cells rather than via site-specific enzyme enrichment, since β-galactosidase is present in the lysosomes of both healthy and malignant cells. The ability of the MMAE prodrug to selectively target tumor cells was investigated using KB and HeLa cells, which overexpress FR and FR-negative A549 cells. The differences in cytotoxicities observed between FR expressing and nonexpressing cells were consistent with the targeting and efficient MMAE release of the prodrug. The antiproliferative activity of the prodrug in KB cells was similar to that of MMAE ($IC_{50} = 0.250$ nM) and corresponded to the FR level in the different cell type, making this prodrug the most potent folate-drug conjugate to date. Confocal microscopy and FACS analysis confirmed that the galactoside prodrug produced a similar effect on cell division as MMAE and that it was a selective receptor-mediated endocytosis process followed by β-galactosidase catalyzed release of MMAE.

Since cancerous tissues are heterogeneous, the selective destruction of a particular population of cells expressing a specific membrane receptor may not be sufficient to eradicate a wide diversity of tumor cells [122].

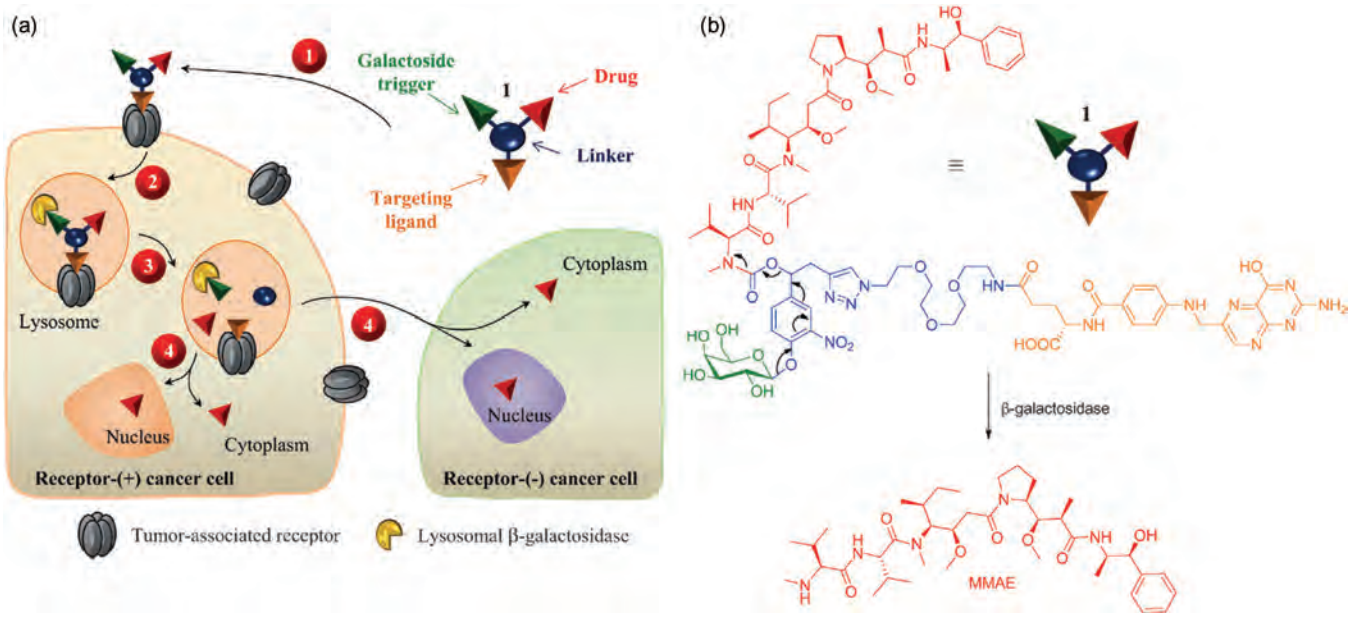


FIGURE 28.30 (a) Principle of tumor targeting. Step 1: selective recognition of receptor-positive cancer cells; step 2: receptor-mediated endocytosis; step 3: β-galactosidase-catalyzed drug release; step 4: diffusion of the drug into the nucleus or cytoplasm of both receptor-positive and receptor-negative cancer cells, leading to death of each type of cell. (b) Structure of the macromolecular prodrug with β-galactosidase-catalyzed MMAE release mechanism [122]. *Reprinted with permission from reference. Copyright 2012 John Wiley and Sons.*

Consequently, Papot studied the effect of MMAE produced from the prodrug by lysosomal β-galactosidase located inside FR-expressing cells on nearby FR-nonexpressing A549 cells. The researchers demonstrated that the prodrug could efficiently release MMAE so as to diffuse out of FR-positive cells and destroy surrounding FR-negative cancer cells. Moreover, treatment of the prodrug (i.v.) in nude mice bearing luciferase-transfected KB xenografts produced significant antitumor effect and improved survival rates without any overt toxicity compared to the MMAE treated mice. This new generation of galactoside prodrugs may be an excellent alternative to the antibody-directed enzyme-prodrug approach and offer selective chemotherapy of solid tumors.

Several interesting examples of prodrugs using a site-specific drug release strategy have been described in the literature. However, a critical component of this approach is the identification of novel enzyme-cleavable linkers. Since a review on this topic was recently published, it will not be discussed here [123].

3. *Antibody-Drug Conjugates as Macromolecular Prodrugs*

Antibody-drug conjugates can be considered both a site-directed and site-specifically released macromolecular prodrug. An antibody–drug conjugate (ADC) is essentially a three component system comprising a potent drug substance linked via a biodegradable linker to an antibody mAb (Figure 28.31a). The antibody binds to specific antigens or receptors at a cell surface where it can be internalized within the cell, the linker cleaved, and the drug released. This approach is highly complex, because it involves multiparameter optimizations beyond the usual ones for a small drug molecule. Besides mAb specificity and internalization capability, linker technology to achieve release at the desired site, efficient drug loading, optimum stoichiometry and homogeneity of the macromolecule are vitally important for attaining good pharmacokinetics, efficacy, and tolerability [125]. The linker can be further divided into four parts [124]. (1) The connection of the drug to the antibody is usually via a lysine or cysteine amino acid residue on the antibody. Antibodies have been engineered with two site-specific (light or heavy chain) cysteine residues to allow two drugs per antibody when a payload connected to a maleimide is added to the thio-engineered antibody. The drug conjugate retains the interchain disulfide bonds, while only the engineered cysteine residue forms a bond to the drug molecule. This process provides a uniform and well-characterized ADC. This is only one example of how this has been done to control drug loading. In most cases, the sites of attachment and the number of attachments vary to give an average drug-to-antibody ratio (DAR). (2) The polar component imparts hydrophilicity to the inherently hydrophobic ADC, thus preventing aggregation and/or precipitation. (3) The trigger is the part of the linker that initiates cleavage and release of the drug. The three trigger mechanisms that have been exploited in the ADCs that went into the clinic are hydrazone hydrolysis in the acidic lysosomal

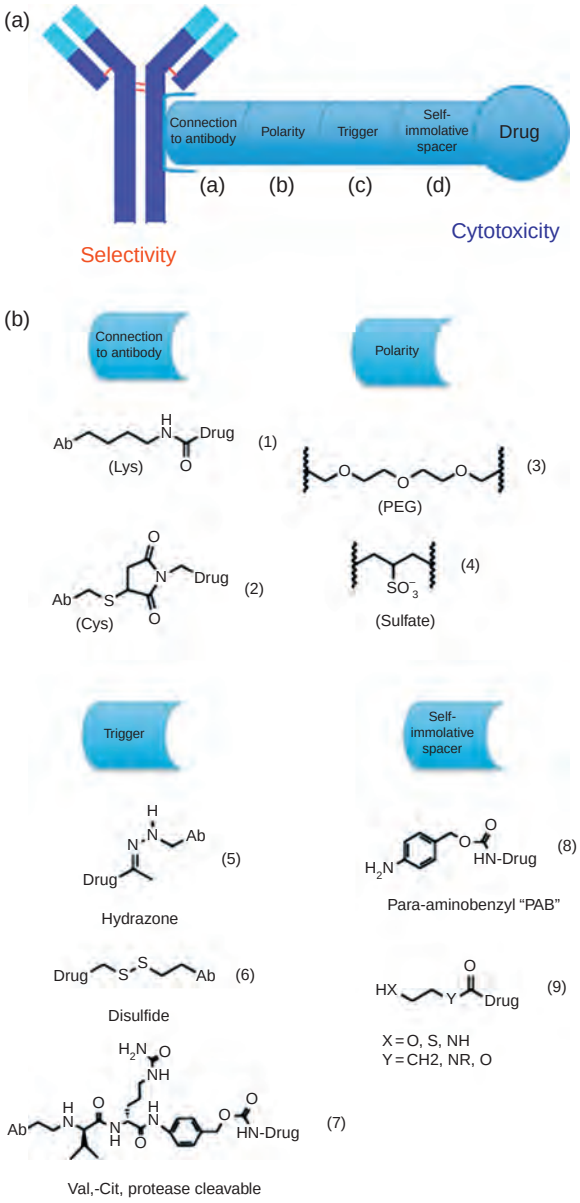


FIGURE 28.31 (**a**) Architecture of an ADC (**b**) The components of the linker portion of the ADC: (a) the connection of the linker to the antibody; (b) the polarity of the linker; (c) the trigger that initiates the cleavage of the drug; and (d) the self-immolative portion that liberates the drug [124]. *Reprinted with permission from reference. Copyright 2013 John Wiley and Sons.*

environment, disulfide reduction by lysosomal glutathione, and a dipeptide cleavable by a lysosomal protease. Finally, (4) the self-immolative spacer is that part of the linker which chemically undergoes a rapid and spontaneous intramolecular reaction liberating the free drug following activation by the trigger. ADCs have been predominantly used in the field of oncology, where the focused delivery of cytotoxic agents to tumor cells while minimizing exposure to normal cells to improve the therapeutic index is advantageous.

Clinical trials on ADCs have been conducted. Few have yielded successful outcomes, however. Exceptions include Mylotarg, SGN-35 (brentuximad vedotin, Adcetris), and most recently TDM-1 (trastuzumab emtansine, Kadcyla) from Roche, which was approved in 2013. Mylotarg was approved under an accelerated approval process by the FDA in 2000 for use in patients over the age of 60 with relapsed acute myelogenous leukemia (AML). However, FDA approval of Mylotarg, a monoclonal antibody to CD33 linked to a cytotoxic agent from the class of calicheamicins, was withdrawn in 2010, when a clinical trial showed that use of the drug was associated with a fatal liver condition. Since then, Pfizer has initiated additional clinical trials that showed the benefit of

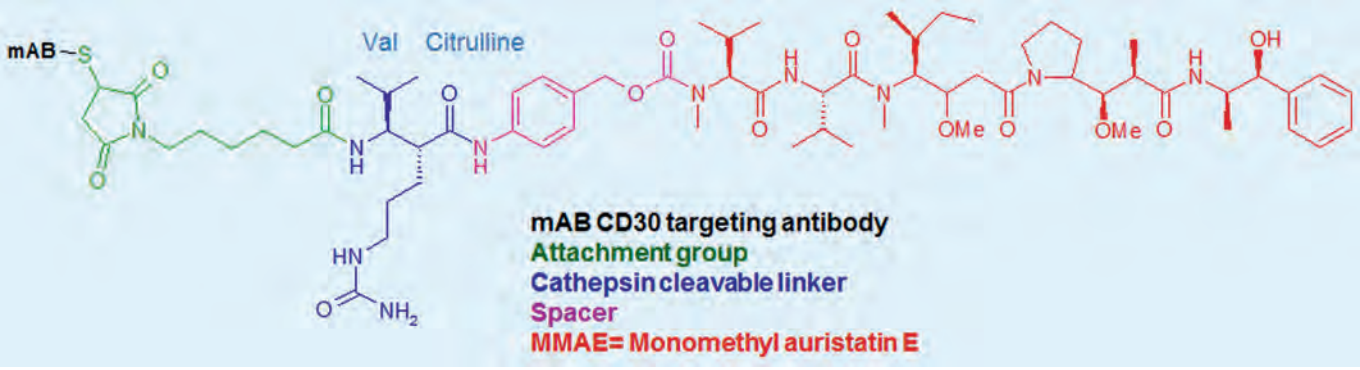


FIGURE 28.32 Structure of brentuximab vedotin (Adcetris).

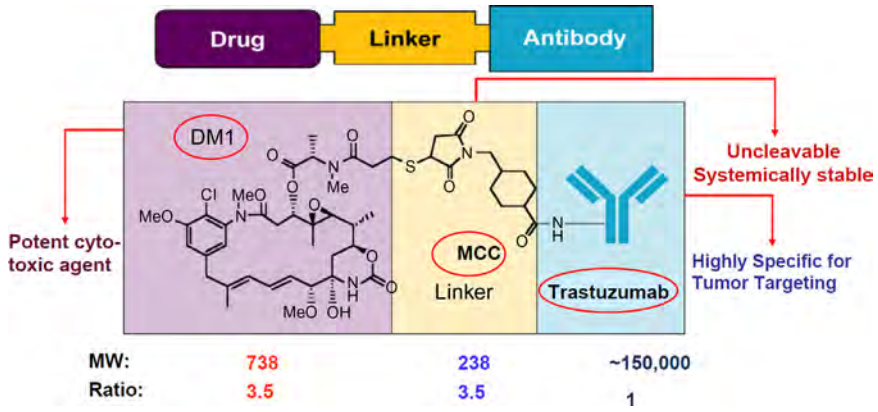


FIGURE 28.33 T-DM1 antibody-directed conjugate [126]. Copyright 2010 Genentech/Roche. DM1 = emtansine = Derivative of Maytansine, a microtubule destabilizing agent. MCC = [maleimidomethyl]cyclohexane-1-carboxylate, a nonreducible thioether linkage. Trastuzumab = Herceptin.

Mylotarg for patients with early diagnosis of the disease, and subsequent modification of the dosage suggests a comeback for this drug. Until recently, the reasons for the few successes were due to the high complexity, cost, and risk of immunogenicity associated with the development of an ADC. More than twenty ADCs currently in clinical trials and recent approvals highlight the significant advances made in the field.

Seattle Genetic's SGN-35 is an antibody-directed prodrug therapeutic designed to target the CD30 receptor, a defining marker for Hodgkins lymphoma. The potent antimitotic agent monomethylauristatin E (MMAE) is attached via a self-immolative spacer to a novel cathepsin-cleavable linker (valine-citrulline), which in turn is attached to the sulfhydryl group of a cysteine residue on the antibody (Figure 28.32). Maleimide is the attachment group used to add selectively to one of the eight available sulfhydryl groups on the antibody. This yields a statistical mixture of between zero and eight drugs per antibody, with a DAR of 4 being the most prevalent [124]. Adcetris was approved by the FDA in 2011 to treat Hodgkin lymphoma and a rare lymphoma known as systemic anaplastic large cell lymphoma (ALCL).

Trastuzumab-DM1 (T-DM1) combines the approved anti-HER2 mAB (Herceptin) with the microtubule destabilizing agent emtansine (DM1) via a succinimidyl 4-[N-maleimidomethyl]cyclohexane-1-carboxylate (SMCC) linker attached to the ε-amino group of the lysine residue of the antibody (Figure 28.33). The impact of the linker on the efficacy and pharmacokinetics of an ADC is based on the empirical assessment of different linkers with the individual drug molecule and antibody in the context of the disease [127]. In this case, a thioether linkage provided the widest margin between the minimum efficacious dose and the best safety profile in an *in vivo* animal model. Following receptor internalization and lysosomal degradation of T-DM1 *in vitro*, the released lysine-SMCC-DM1 within the HER2-positive breast cancer cells caused mitotic arrest of cells in the G2/M stage. Lysine-SMCC-DM1 added exogenously to cells in a cell viability assay displayed poor cytotoxic effects, suggesting its limited cell permeability and the requirement of the antibody for internalization. More importantly, premature cleavage of the ADC will most likely not cause

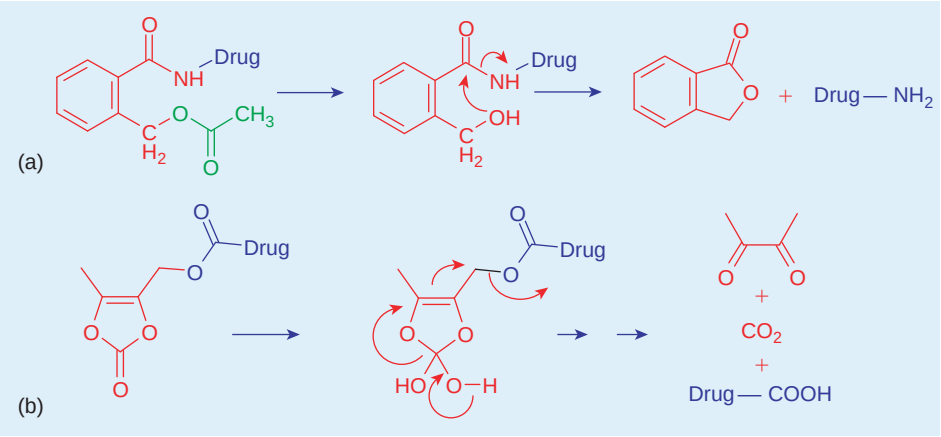


FIGURE 28.34 Cascade latentiation: (a) 2-Acyloxymethylbenzoic acids provide amides the lability of esters; [131,132]. (b) Substituted vinyl esters as lipophilic cascade carrier for carboxylic acid-containing drugs [61,133].

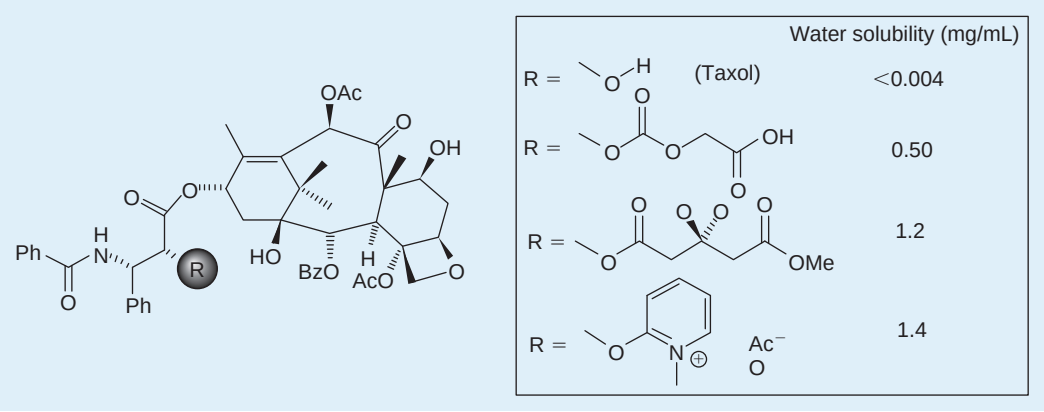


FIGURE 28.35 Water-soluble prodrugs of paclitaxel [136,137].

by-stander killing of nontumor cells, thereby improving the safety window. Metastatic breast cancer patients treated with Kadcyla had a median progression-free survival of 9.6 months compared to 6.4 months in patients treated with lapatinib plus capecitabine. The median overall survival was 30.9 months in the Kadcyla group and 25.1 months in the lapatinib plus capecitabine group [128]. The positive Phase III trial results led to the FDA approval of Kadcyla as a new therapy for patients with HER2-positive late stage breast cancer [129].

B. Use of Cascade Prodrugs

Most of the prodrugs described to this point are classical carrier-linked prodrugs. Classical carrier-linked prodrugs may sometimes be ineffective because the prodrug linkage is too stable (amides, nonactivated esters). In such cases, a β-assistance provided by an easily *in vivo* generated nucleophile can represent an interesting solution. The release of the active molecule from the prodrug proceeds through a two-step trigger mechanism, for which the name "cascade latentiation" was coined by Cain in 1975 [120,121].

The concept, also called distal hydrolysis [47] or the double prodrug concept [33,130], is illustrated by the use of 2-acyloxymethylbenzoic acids as amine-protective functions, providing amides with the lability of esters (Figure 28.34a), and by the use of substituted vinyl esters [= (2-oxo-1,3-dioxol-yl)methyl esters] as lipophilic cascade carriers for carboxylic acid-containing drugs such as ampicillin [133], α-methyldopa [61], or various cephalosporins (Figure 28.34b) [8,15,39,135].

1. Water-Soluble Paclitaxel Prodrugs

Paclitaxel (sold under the name Taxol) is a potent microtubule-stabilizing agent that has been approved for cancer treatment. Despite paclitaxel's therapeutic promise, its aqueous insolubility (<0.004 mg/mL) hampers its clinical application. Nicolaou et al [136,137] report the design, synthesis, and biological activity of prodrugs designed to improve water solubility that can also be considered as cascade prodrugs (Figure 28.35).

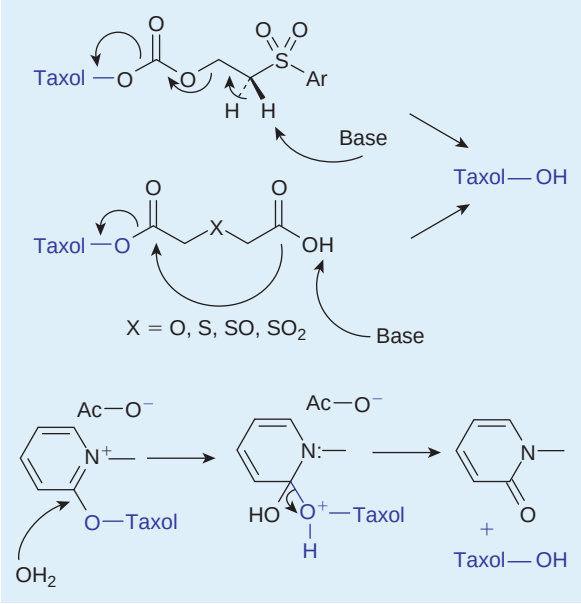


FIGURE 28.36 Paclitaxel release mechanisms from paclitaxel prodrugs [136,137].

The mechanistic rationale for the design of the first two paclitaxel prodrugs lies in the spontaneous decomposition of the carbonate ester after the abstraction of one of the activated protons or of an acidic proton (Figure 28.36). The release of paclitaxel from the pyridinium prodrug (paclitaxel -2′-methyl- pyridinium acetate; paclitaxel -2′-MPA) is presumed to be the result of a nucleophilic attack by water or another nucleophile at the 2′-position of the pyridinium moiety (Figure 28.36) [137].

2. *Bioactivation of an Antibacterial Prodrug*

Although amino acid **1** is a potent inhibitor of CMP−KDO synthase (Figure 28.37), a key enzyme in the biosynthesis of the lipopolysaccharide of Gram-negative bacteria, it is unable to reach its cytoplasmic target and is therefore inactive as an antibacterial agent. Simple lipophilic esters are not useful to enhance the delivery of amino acid **1** since they are not cleaved by the bacteria. On the other hand, double prodrug **3** has been found to solve the problem [134].

Upon entry into bacterial cells, the disulfide bond in compound **3** is reduced by sulfhydryl compounds present in the intracellular milieu, resulting in the formation of thiol **2**. This is highly unstable, and the active amino acid **1** is formed by a rapid, intramolecular displacement.

3. *Double Prodrugs for Peptides*

Amsberry and Borchardt [138,139] have applied Cain's cascade concept to prepare lipophilic polypeptide prodrugs. The amine functionality of the polypeptide is coupled to 2′-acylated derivatives of 3-(2′,5′-dihydroxy-4′, 6′-dimethylphenyl)-3,3-dimethylpropionic acid (Figure 28.38). Under simulated physiological conditions, the parent amine is regenerated in a two-step process: enzymatic hydrolysis of the phenolic ester, followed by a non-enzymatic intramolecular cyclization, leading to the release of the free amine (polypeptide) and a lactone.

The lactonization step is highly favored because of the steric pressure created by the three methyl groups (the "trimethyl lock" concept). An alternative to the hydrolytic first step involves a bioreductive generation of the intermediate phenolic amide (Figure 28.38).

C. Codrugs

Codrugs are also named "mutual prodrugs." Their design consists of the linking, in a unique molecule, of at least two different synergistic drugs that are released *in vivo* at the desired site of action. An example is found in the association of L-dopa to the catechol *O*-methyltransferase (COMT) inhibitor entacapone (Figure 28.39) [140].

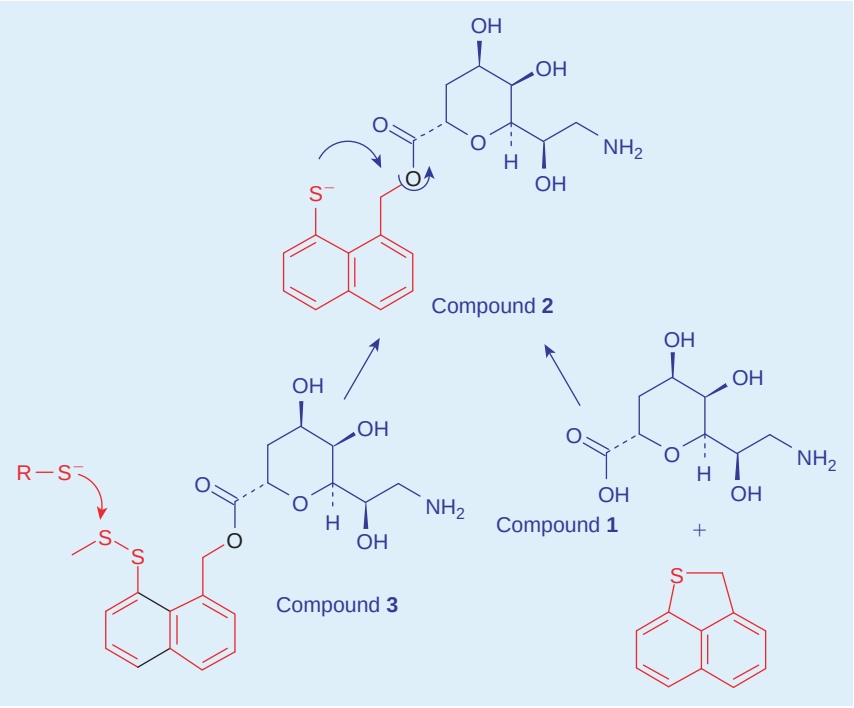


FIGURE 28.37 Bioactivation of the antibacterial prodrug of an impermeant inhibitor of 3-deoxy-D-manno-2-octulosonate cytidylyl- transferase [134].

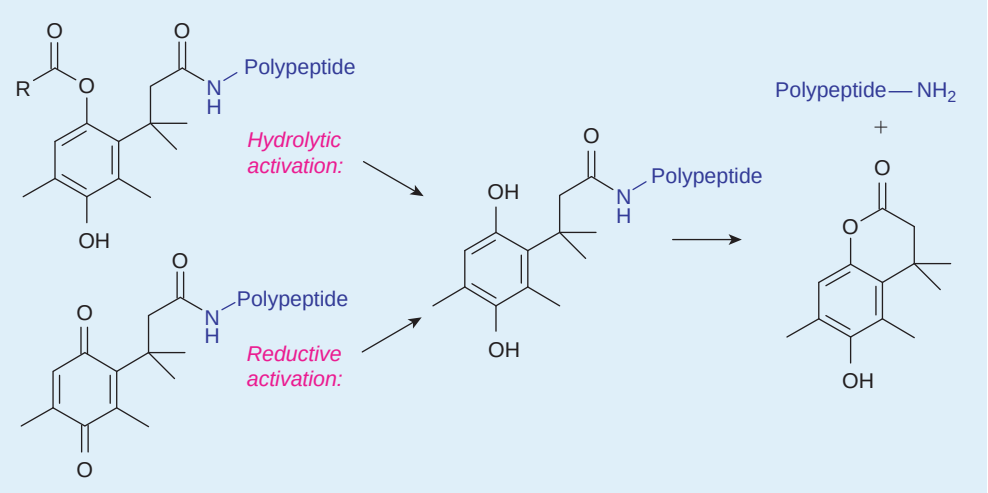


FIGURE 28.38 Proposed conversion of esterase-sensitive and redox-sensitive double prodrugs of peptides [138,139].

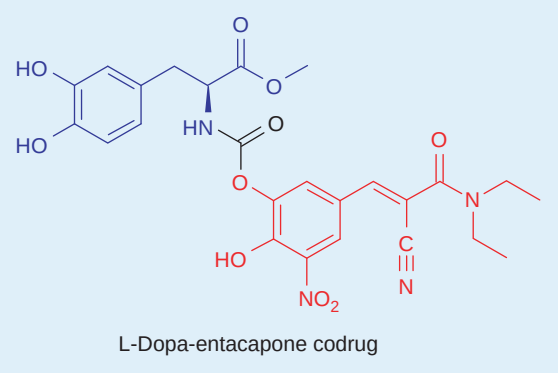


FIGURE 28.39 The amino functional group of L-dopa is linked to the phenolic function of entacapone by means of a carbonyl group; thus a carbamate function links the two active agents.

D. Carrier Prodrugs: Conclusion

The carrier prodrug approach is particularly successful in the antibiotics and oncology field, where there are clear improvements in some pharmacokinetic and toxicological parameters. The design of carrier prodrugs in medicinal chemistry represents the counterpart of the design of protective groups in organic chemistry. Both approaches have much in common; in both of them, imagination and creativity have no limits.

V. BIOPRECURSOR PRODRUG EXAMPLES

The following examples illustrate the bioprecursor prodrug approach. The intentional use of bioprecursor design is relatively recent, and—in some cases—there are doubts about the prospective or retrospective character of the design. The first examples relate to oxidative bioactivations. These are followed by examples of reductive bioactivations and mixed-type bioactivations. In general, the active species results from a cascade of metabolic reactions involving oxidative as well as reductive processes, complicated by hydrolytic reactions or hydration−dehydration sequences.

A. Oxidative Bioactivations

1. *Oxidative Bioactivation of Losartan*

A classic example of a bioprecursor prodrug is found in losartan, a nonpeptide angiotensin II receptor antagonist used as an antihypertensive medication [141]. It can also be considered a bioprecursor prodrug insofar as, *in vivo*, the primary alcohol is oxidized to a carboxylic acid (Figure 28.40), which represents the actual active principle [142].

2. *Conjugated Cyclohexeneones as Bioprecursors of Catecholamines*

Venhuis et al observed a particularly original oxidative bioactivation mechanism by which an α, β-unsaturated cyclic ketone is converted to the corresponding catechol and delivered enantioselectively into the CNS (Figure 28.41). This concept can be generalized and has the potential to lead to new anti-Parkinson's treatments [143].

3. *Site-Specific Delivery of Acetylcholine-Esterase Reactivator 2-PAM to the Brain*

N-methylpyridinium-2-carbaldoxime (2-PAM = **a**; Figure 28.42) constitutes the most potent reactivator of acetylcholinesterase poisoned by organophosphorus acylation. However, due to its quaternary nitrogen, 2-PAM penetrates the biological membranes poorly and does not appreciably cross the BBB. For this compound, Bodor et al [144] designed a novel dihydropyridine−pyridinium salt type of redox delivery system. The active drug is administered as its 5,6-dihydropyridine derivative (Pro-2- PAM = **b**), which exists as a stable immonium salt **c**. The lipoidal **b** (pKa = 6.32) easily penetrates the BBB, where it is oxidized to the active **a**. A dramatic increase in the brain delivery of 2-PAM by the use of Pro-2-PAM is thus achieved, resulting in a re-activation of phosphorylated brain acetyl-cholinesterase in vivo [145,146].

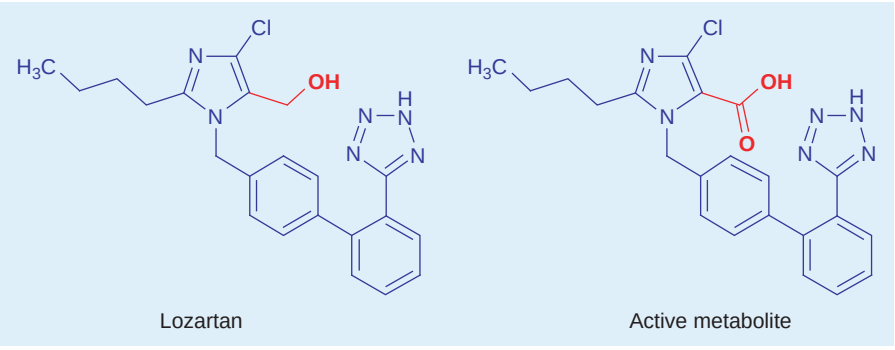


FIGURE 28.40 Oxidative bioactivation of losartan.

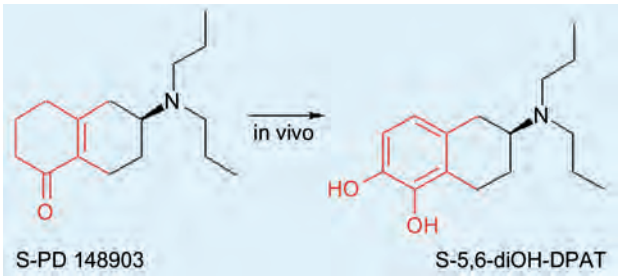


FIGURE 28.41 Oxidative bioactivations leading to catechols [143].

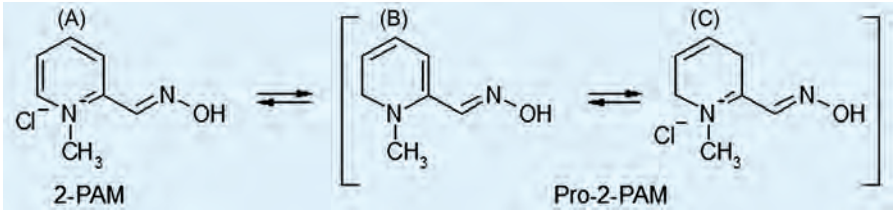


FIGURE 28.42 Metabolism of dihydro derivatives of 2-PAM [144].

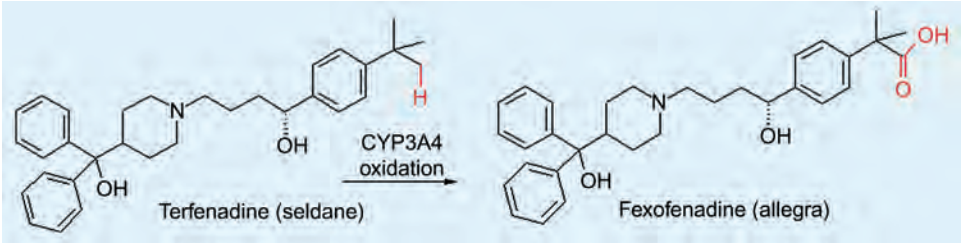


FIGURE 28.43 CYP3A4 oxidation of terfenadine.

4. *Oxidation of Terfenadine by CYP3A4*

Terfenadine, a bioprecursor prodrug, is an antihistamine used for the treatment of allergic conditions. It is completely metabolized in the liver by the enzyme cytochrome P450 3A4 isoform to give the active metabolite fexofenadine (Figure 28.43) [147]. Terfenadine was taken off the market due to the risk of cardiac arrhythmia caused by QT prolongation. The active metabolite is not cardiotoxic, however, and is now sold under the brand name Allegra. This is an example where the active component is better administered than a bioprecursor prodrug due to its pharmacologically undesirable activity.

5. *6-Deoxyacyclovir as a Bioprecursor of Acyclovir*

The antiherpetic agent acyclovir suffers from poor oral bioavailability; only 10–20 percent of an oral dose is absorbed in humans. This is ascribed to low water solubility due to strong interaction forces in the crystal lattice. The corresponding deoxo derivative (6-deoxyacyclovir) was shown by Krenitsky [148] to be eighteen times more water soluble and to be rapidly oxidized *in vivo* by xanthine oxidase to the parent drug (Figure 28.44). Studies in rats and in human volunteers showed that orally administered 6-deoxyacyclovir has a 5–6 times greater bioavailability than acyclovir [148,149].

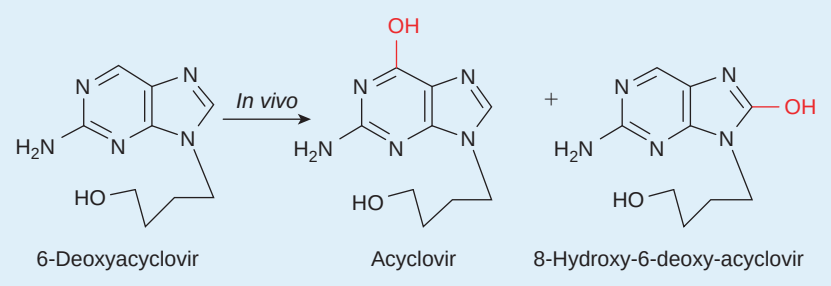


FIGURE 28.44 6-Deoxyacyclovir as a bioprecursor of acyclovir [148,149].

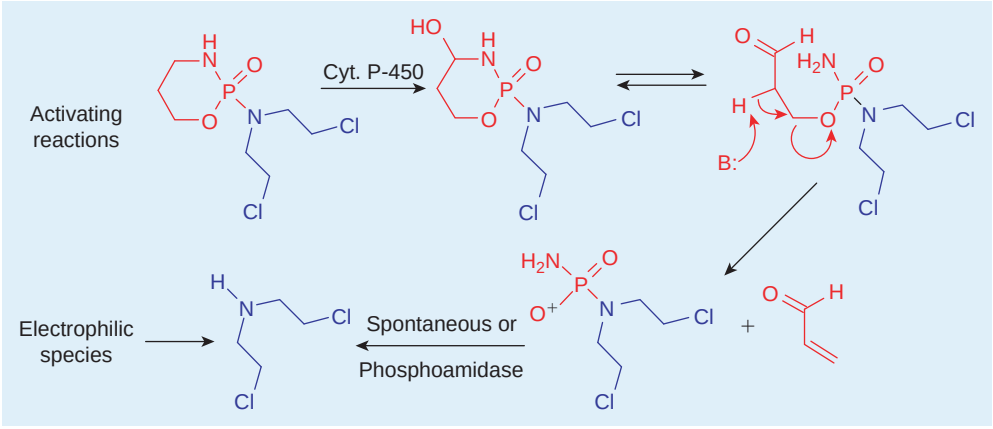


FIGURE 28.45 Bioactivation of cyclophosphamide [151,152].

6. *Bioactivation of Cyclophosphamide*

Cyclophosphamide is a cytotoxic (cytostatic), nonspecific cell cycle, antiproliferative agent, which is used in such diverse medical problems as neoplasia, tissue transplantation, and inflammatory diseases [150]. Chemically, it is an inert bioprecursor for a potent nitrogen mustard alkylation agent (Figure 28.45).

Cyclophosphamide was synthesized by Arnold et al [153–155] in the hope that it would be inert until activated by an enzyme present in the body, especially in a tumor. The activation mechanism is believed to require an initial oxidative dealkylation, followed by a spontaneous or phosphoramidase-catalyzed hydrolysis to the parent nitrogen mustard [151,152].

B. Reductive Bioactivations

1. *Reductive Bioactivation of Nitrogen Mustards*

Many conventional anticancer drugs display relatively poor selectivity for neoplastic cells, and solid tumors are particularly resistant both to radiation and to chemotherapy. However, solid tumors possess a few unique and important microenvironmental properties such as localized hypoxia, nutrient deprivation, and low pH [156]. On the other hand, as shown above for sulindac, sulfoxides can undergo two major biotransformations: reversible reduction to the sulfide and irreversible oxidation to the sulfone. The oxidation to the sulfone is the dominant process under normal physiological conditions, but the reduction to the sulfide becomes significant under anaerobiotic conditions [157]. Taking advantage of these findings, Kwon et al [158] devised a hypoxia-selective alkylating bioprecursor prodrug (Figure 28.46a).

2. *Reductive Bioactivation of Nitroimidazolylmethyluracils*

Thymidine phosphorylase (TP) is an angiogenic growth factor and a target for anticancer drug design (Figure 28.46b). Docking studies of the modeled TP predicted that the binding of aminoimidazolylmethyluracils was energically more favored than that of the corresponding nitro counterparts [159]. Effectively, the passage from the nitro to the amino analog was accompanied by a 1,000-fold increase in TP inhibition.

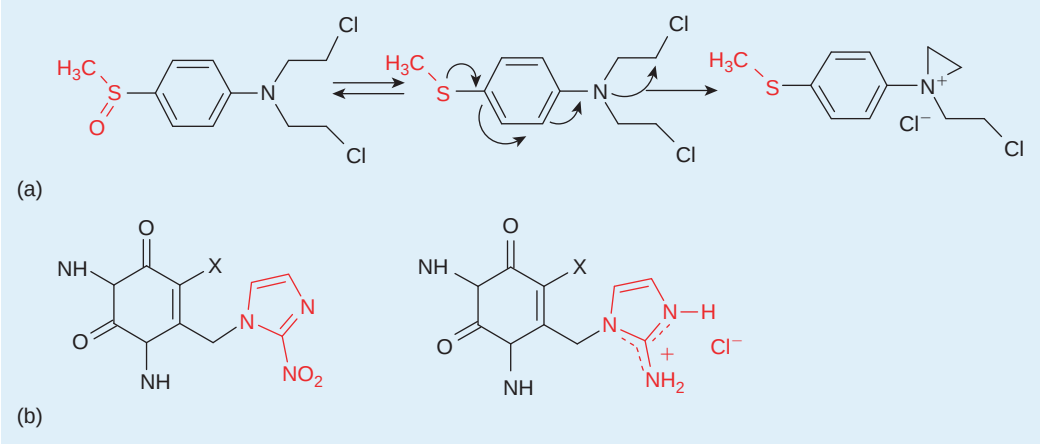


FIGURE 28.46 (a) and (b) Hypoxia-selective nitrogen mustard [158] and bioactivated TP inhibitors [159].

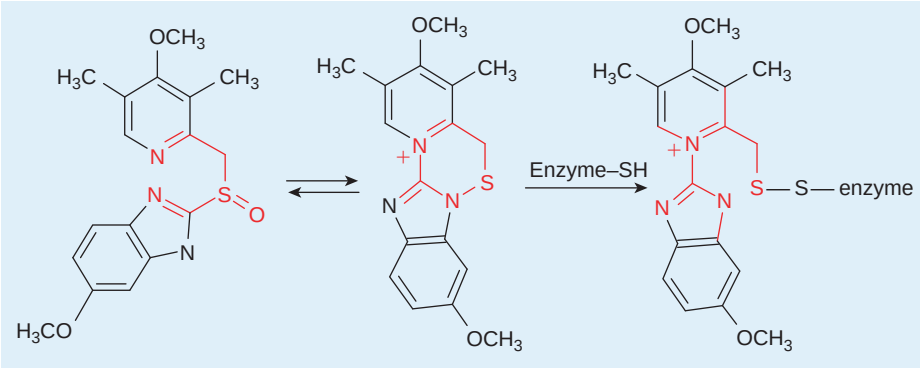


FIGURE 28.47 Reductive bioactivation of omeprazole.

3. *Reductive Bioactivation of Omeprazole*

Omeprazole effectively inhibits gastric secretion by inhibiting the gastric H^+, K^+-ATPase [160]. This enzyme is responsible for gastric acid production and is located in the secretory membranes of parietal cells. Omeprazole is an anti-ulcerative drug, used especially in the treatment of Zollinger–Ellison syndrome [161].

In vivo, omeprazole is transformed into the active inhibitor, a cyclic sulfenamide (Figure 28.47), which forms disulfide bridges with the thiol groups of the enzyme and thus inactivates it [162,163]. The high specificity in the action of omeprazole (pKa = 4.0) is due to its preferential concentration in the rather acidic parietal cells where it is activated. In neutral regions of the body, omeprazole is rather stable and is only partially converted to the active species.

C. Mixed Bioactivation Mechanisms

Certain bioactivation mechanisms involve several chemical sequences, some of them being oxidative and others being reductive. The prodrug SAH 51-641 (**1**) (Figure 28.48) illustrates an example of a mixed oxidative-reductive mechanism. SAH 51-641 is a potent hypoglycemic agent, which acts by inhibiting hepatic gluconeogenesis via inhibition of fatty acid oxidation [72]. This compound is metabolized by a sequential oxidation/reduction to the corresponding keto-acid (**2**) and the hydroxyl acid (**3**). Compound (**3**) is a substrate for medium-chain fatty acyl CoA ligase and represents the actual active agent [71].

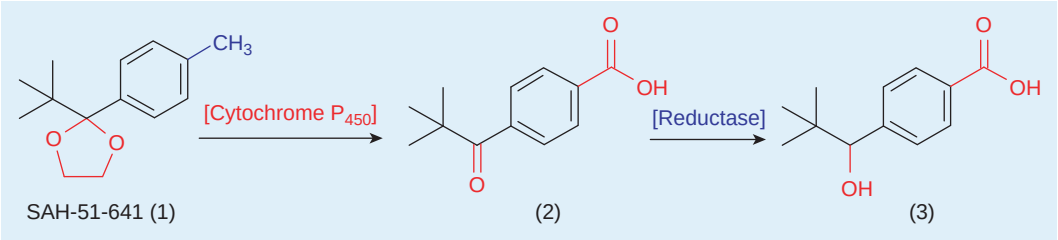


FIGURE 28.48 Mixed oxidative/reductive bioactivation of dioxolanes.

VI. DISCUSSION

Bioprecursors Versus Carrier Prodrugs

A comparative balancesheet established for the two prodrug approaches led us to the following conclusions (Table 28.2):

- The bioavailability of carrier prodrugs is modulated by using a transient transport moiety. Such a linkage is not implied for bioprecursors, which result from a molecular modification of the active principle itself.
- The lipophilicity is generally the subject of a profound alteration of the parent molecule in the case of carrier prodrugs, whereas it remains practically unchanged for bioprecursors.
- The bioactivation process is exclusively hydrolytic for carrier prodrugs. It involves mostly redox systems for bioprecursors.
- The catalysis leading to the active principle is hydrolytic (either through general catalysis or through extrahepatic enzymes) for carrier prodrugs. For bioprecursors, it seems largely restricted to Phase I metabolizing enzymes.

TABLE 28.2 Bioprecursors Versus Carrier Prodrugs

	Prodrugs	
	Carrier prodrugs	**Bioprecursors**
Constitution	Active principle + carrier group	No carrier group
Lipophilicity	Strongly modified	Usually slightly modified
Bioactivation	Hydrolytic	Mostly oxidative or reductive
Catalysis	Chemical or enzymic	Only enzymic

VII. DIFFICULTIES AND LIMITATIONS

The introduction of prodrugs in human therapy gave successful results in overcoming undesirable properties such as poor absorption, rapid biodegradation, or formulation problems. It can be expected that an increasing number of medicinal chemists will be tempted by this approach. However, they must keep in mind that prodrug design can also give rise to a large number of new difficulties, especially in the assessment of pharmacological, pharmacokinetic, toxicological, and clinical properties.

At the pharmacological level, for example, because bioactivation is necessary to create the active species, these compounds cannot be submitted to preliminary *in vitro* screening tests, namely, binding studies, neurotransmitter re-uptake, measurements of enzymatic inhibition, and activity on isolated organs.

The measurements of pharmacokinetic parameters can lead to numerous misinterpretations. Thus, pivampicillin has a half-life of 103 min in a buffered aqueous solution at 37°C, but it falls to less than one min after addition of only 1 percent of mouse or rat serum. In the presence of human serum (10 percent), however, the half-life is fifty min, whereas in whole human blood it is only five min. These results exemplify the care required to avoid incorrect conclusions. In addition, when a prodrug and the parent molecule are compared, one must take into

account the differences in their respective time courses of action. The maximum activity can appear later for the prodrug than for the parent compound, and often the comparison of the AUC could constitute a better criterion.

At the toxicological level, even when prodrugs derive from well-known active principles, they have to be regarded as new entities. Undesirable side effects can appear that are directly related to the prodrug (e.g., allergy to bucloxic acid) or derived from the bioactivation process (e.g., formation of unwanted or unexpected metabolites), or which can be attributed to the temporary transport moiety (e.g., digestive intolerance to pivampicillin, antivitamin-PP activity of nicafenine). This latter case is particularly illustrative. An apparently innocent carrier group such as *N*-hydroxyethylnicotinamide appeared as a promising candidate for improving the absorption of acidic anti-inflammatory drugs or clofibric acid [164–166]. However, during the clinical studies, side effects similar to vitamin PP deficiency appeared, suggesting that *N*-hydroxyethylnicotinamide could function as a nicotinamide antimetabolite. The compounds then had to be withdrawn (H. Cousse, Pierre Fabre & Co, personal communication).

In a review of potential hazards of the prodrug approach, Gorrod [167] cites four toxicity mechanisms:

1. Formation of a toxic metabolite of the total prodrug that is not produced by the parent drug.
2. Consumption of a vital constituent (e.g., glutathione) during the prodrug activation process. As L-cysteine is needed for the biosynthesis of glutathione, a supply with L-cysteine prodrugs can eventually confer some protection of the hepatic cells [168].
3. Generation of a toxic derivative from a transport moiety supposed to be "inert."
4. Release of a pharmacokinetic modifier (causing enzymatic induction, displacing protein-bound molecules, altering drug excretion, etc.).

Eventually, at the clinical stage, the predictive value of animal experiments is also questionable. Thus, for two prodrugs derived from α-methyldopa, the active doses in rats were identical, but they turned out to be very different during clinical investigations. One compound was just as potent as α-methyldopa, whereas the other one was 3–4 times more potent [169,170].

Finally, a patent application for a new prodrug should take into account all these aspects, and because the biological profile of each individual prodrug will differ depending on the linker/spacer and promoiety, it should in no way be regarded just as a complement to the main file. The question of patentability, however, is becoming more difficult to define as more prodrug examples appear in the literature, thus requiring inventors to meet the non-obvious criteria by demonstrating an unexpected result from the prodrug. Recent prodrugs designed for targeted delivery will certainly fulfill the requirements for patentability: novelty, utility, and non-obviousness.

VIII. CONCLUSION

For clarity, it is preferable to distinguish the carrier prodrug type from bioprecursor approaches. The first one, consisting in the attachment of a temporary carrier group to an active principle, largely proved its utility in the design of orally active antibiotics and, more generally, every time high bioavailability in plasma or peripheral organs is required. Novel carrier prodrugs targeting transporters to gain entry into the brain may offer a solution to the challenge of delivering drugs to the CNS. The opportunity for prodrugs to provide additional advantages—especially in the area of targeted delivery—is beginning to bear results as these newer prodrugs make their way into clinical trials and to patients.

The design of bioprecursors, which represents a creative application of the active metabolite concept in the forward-looking way, seems *a priori* more adequate for CNS delivery, but its clinical usefulness still needs to be proven.

In general, prodrug as a strategy should be considered relatively early during the drug-optimization phase as a complementary approach. Prodrug design is unique for each molecule, with each entity having its own particular issues; consequently optimization cannot be generalized. The examples provided highlight the level of complexity and sophistication in the type and targeting capabilities of recent prodrugs. Regardless of the class of prodrug, key optimization parameters are the pharmacokinetic profile, level of bioconversion, stability in the formulation, and safety of the released fragments (in the case of carrier prodrugs). Although *in vitro* tests (e.g., microsomal stability, pH and plasma stability, and stability in gastric/intestinal fluid in the case of oral drugs) may provide guidance and should be conducted, they can be misleading, and therefore care must be taken when interpreting the data. Prodrug selection should be based on comparative studies in preclinical *in vivo* models to establish the desired pharmacokinetics, biological activity, and toxicological profiles, but ultimately only clinical studies can prove its value.

References

[1] Stella V. Pro-drugs as a novel drug delivery system. In: Higuchi T, Stella V, editors. Pro-drugs: an overview and definition, vol. 14. Washington, DC: American Chemical Society; 1975. p. 1–115.
[2] Sinkula A. Prodrugs, protective groups, and the medicinal chemist. In: Mathieu J, editor. Medicinal chemistry, vol. 5. Amsterdam: Elsevier; 1977. p. 125–33.
[3] Wermuth CG. Designing prodrugs and bioprecursors. In: Wermuth CG, editor. The practice of medicinal chemistry. London: Academic Press; 2003. p. 561–85.
[4] Ettmayer P, Amidon GL, Clement B, Testa B. Lessons learned from marketed and investigational prodrugs. J Med Chem 2004;47: 2393–403.
[5] Testa B. Prodrug objectives and design. In: Taylor JB, Triggle DJ, editors. Comprehensive medicinal chemistry II, vol. 5. Amsterdam: Elsevier; 2007. p. 1009–41.
[6] Huttunen KM, Raunio H, Rautio J. Prodrugs: from serendipity to rational design. Pharmacol Rev 2011;63:750–71.
[7] Clas SD, Sanchez RI, Nofsinger R. Chemistry-enabled drug delivery (prodrugs): Recent progress and challenges. Drug Discovery Today 2014;19:79–87.
[8] Albert A. Chemical aspects of selective toxicity. Nature (London) 1958;182:421–3.
[9] Harper NJ. Drug latentiation. J Med Pharm Chem 1959;1:467–500.
[10] Wermuth CG. Bioprécurseurs contre prodrogues [Bioprecursors versus prodrugs]. In: Briot M, Cautreels W, Roncucci R, editors. Drug metabolism and drug design: Quo vadis? Montpellier: Sanofi-Clin-Midy; 1983. p. 253–71. [in French].
[11] Wermuth CG. Designing prodrugs and bioprecursors. In: Jollès G, Wooldrige KRH, editors. Drug design: fact or fantasy? London: Academic Press; 1984. p. 47–72.
[12] Wermuth CG. Modulation of natural substances in order to improve their pharmacokinetic properties. In: Beal JL, Reinhard E, editors. Natural products as medicinal agents. Stuttgart: Hippokrates Verlag; 1981. p. 185–216.
[13] Wermuth CG. Les prodrogues, des médicaments plus sûrs et plus maniables [Prodrugs: Safer and more manageable drugs]. Bull Soc Pharm Bordeaux 1980;119:107–29. [in French].
[14] Daehne WV, Frederiksen E, Gundersen E, Lund F, March P, Petersen HJ, et al. Acyloxymethyl esters of ampicillin. J Med Chem 1970; 13:607–12.
[15] Bodin NO, Ekström B, Forsgren U, Jalar LP, Magni L, Ramsey CH, et al. Bacampicillin: a new orally well-absorbed derivative of ampicillin. Antimicrob Agents Chemother 1975;8:518–25.
[16] Clayton JP, Cole M, Elson SW, Ferres H, Hanson JC, Mizen LW, et al. Preparation, hydrolysis, and oral absorption of lactonyl esters of penicillins. J Med Chem 1976;19:1385–91.
[17] Testa B, Jenner P. Drug metabolism: chemical and biochemical aspects. New York: Marcel Dekker; 1976.
[18] Shen TI, Witzel BE, Jones H, Linn BO, McPherson J, Greenwald R, et al. Synthesis of a new anti-inflammatory agent, cis-5-fluoro-2-methyl-1-[p-(methylsulfinyl) benzylidenyl]-indene-3-acetic acid. Fed Proc 1972;31:577.
[19] Hucker HB, Stauffer SC, White SD, Rhodes RA, Arrison BH, Umbenhauer ER, et al. Physiological disposition and metabolic fate of a new anti-inflammatory agent, cis -5-fluoro-2-methyl-1-[p-(methylsulfinyl) benzylidenyl]-indene-3-acetic acid in the rat, dog, rhesus monkey, and man. Drug Metab Dispos 1973;1:721–36.
[20] Duggan DE, Hare LE, Ditzler CA, Lei BW, Kwan KC. The disposition of sulindac. Clin Pharmaco Ther 1977;21:326–35.
[21] Duggan DE, Hooke KF, Noll RM, Hucker HB, Van Arman CG. Comparative biodisposition of sulindac and metabolites in five species. Biochem Pharmacol 1978;27:2311–20.
[22] Duggan DE, Hooke KF, Risley EA, Shen TY, Van Arman CG. Identification of the biologically active form of sulindac. J Pharmacol Exp Ther 1977;201:8–13.
[23] Higuchi T, Stella V. Pro-drugs as novel drug delivery systems, vol. 14. Washington, DC: American Chemical Society; 1975.
[24] Majumdar S, Duvvuri S, Mitra AK. Membrane transporter/receptor-targeted prodrug design: Stratiegies for human and veterinary drug development. Adv Drug Deliv Rev 2004;56:1437–52.
[25] Cao F, Gao Y, Ping Q. Advances in research of PepT1-targeted prodrug. Asian J Pharma Sci 2012;7:110–22.
[26] Kim RB. Transporters and drug discovery: why, when, and how? Mol Pharm 2006;3:26–32.
[27] Winiwarter S, Hilgendorf C. Modeling of drug-transporter interactions using structural information. Curr Opin Drug Dis Dev 2008;11:95–103.
[28] Newstead S, Drew D, Cameron AD, Postis VL, Xia X, Fowler PW, et al. Crystal structure of a prokaryotic homologue of the mammalian oliopeptide-proton symporters, PepT1 and PepT2. EMBO J 2011;30:417–26.
[29] Porras AG, Holland SD, Gertz BJ. Pharmacokinetics of alendronate. Clin Pharmacokinet 1999;36:315–28.
[30] Ezra A, Hoffman A, Breuer E, Alferiev IS, Monkkonen J, El-Hanany-Rozen N, et al. A peptide prodrug approach for improving bisphosphonate oral absorption. J Med Chem 2000;43:3641–52.
[31] Talluri RS, Samanta SK, Gaudana R, Mitra AK. Synthesis, metabolism, and cellular permeability of enzymatically stable dipeptide prodrugs of acyclovir. Int J Pharm 2008;361:118–24.
[32] Talluri RS, Gaudana R, Hariharan S, Mitra AK. Pharmacokinetics of stereoisomeric dipeptide prodrugs of acyclovir following intravenous and oral administrations in rats: a study involving in vivo corneal uptake of acyclovir following oral dosing. Ophthalmol Eye Dis 2009;1:21–31.
[33] Bundgaard H. Design and application of prodrugs. In: Krogsgaard-Larsen P, Bundgaard H, editors. A textbook of drug design and development. 1st ed. Chur: Harwood Academic Publishers; 1991. p. 113–91.
[34] McClure D. The effect of a pro-drug of epinephrine (dipivaloyl-epinephrine) in glaucoma-general pharmacology, toxicology, and clinical experience. In: Higuchi T, Stella V, editors. Pro-drugs as a novel drug delivery systems, vol. 14. Washington, DC: American Chemical Society; 1975. p. 224–35.
[35] Horn AS. Pro-drugs of dopaminergic agonists. Chem Ind 1980;:441–4.

[36] Westerink BHC, Dijkstra D, Feenstra MGP, Grol CJ, Horn AS, Rollema H, et al. Dopaminergic prodrugs: brain concentration and neurochemical effects of 5,6- and 6,7-ADTN after administration as dibenzoyl esters. Eur J Pharmacol 1980;61:7–15.
[37] Brass EP. Pivalate-generating prodrugs and carnitine homeostasis in man. Pharmacol Rev 2002;54:589–98.
[38] Tullar BF, Minatoya H, Lorenz RR. Esters of N-tert- butylarterenol: Long-acting new bronchodilators with reduced cardiac effects. J Med Chem 1976;19:834–8.
[39] Lin TS. Synthesis and in vitro antiviral activity of 3′-O-acyl derivatives of 5′-amino-2′-deoxy thymidine: Potential prodrugs for topical application. J Pharm Sci 1984;73:1568–71.
[40] Harnden MR, Jarvest RL, Boyd MR, Sutton D, Verehodge RA. Prodrugs of the selective antiherpes virus agent 9-[4-hydroxy-3-(hydroxymethyl)but-1-yl]guanine (BRL 39123) with improved gastrointestinal absorption properties. J Med Chem 1989;32:1738–43.
[41] Jones LA, Moorman AR, Chamberlain SD, Miranda P, Reynolds DJ, Burns CL, et al. Di- and triester prodrugs of the varicella-zoster antiviral agent 6-methoxypurine arabinoside. J Med Chem 1992;35:56–63.
[42] Vinsova J, Ferriz JJ. Prodrug design of phenolic drugs. Curr Pharma Design 2010;16:2033–52.
[43] Elger W, Schwarx S, Hedden A, Reddersen G, Schneider B. Sulfamates of various estrogens are prodrugs with increased systemic and reduced hepatic estrogenicity at oral application. J Steroid Biochem Molec Biol 1995;55:395–403.
[44] Patel JU, Prankerd RJ, Sloan KB. A prodrug approach to increasing the oral potency of a phenolic drug. Part 1. Synthesis, characterization, and stability of an O-(imidomethyl) derivative of 17-β-estradiol. J Pharm Sci 1994;83:1477–81.
[45] Patel J, Katovich MJ, Sloan KB, Curry SH, Pranker RJ. A prodrug approach to increasing the oral potency of a phenolic drug. Part 2. Pharmacodynamics and preliminary bioavailability of an orally administered O-(imidomethyl) derivative of 17-β-estradiol. J Pharm Sci 1995;84:174–8.
[46] Bodor N, Sloan KB, Little RJ, Selk SH, Caldwell L. Soft drugs 4. 3-Spirothiazolidines of hydrocortisone and its derivatives. Int J Pharm 1982;10:307–21.
[47] Collis AJ. Medicinal chemistry: The role of organic chemistry in drug research. In: Ganellin CR, Roberts SM, editors. Drug access and prodrugs. 2nd ed. London: Academic Press; 1993. p. 61–82.
[48] Anden NE, Corrodi H, Dahlström A, Fuxe K, Högfelt T. Effects of tyrosine hydroxylase inhibition on the amine levels of central monoamine neurons. Life Sci 1966;5:561–8.
[49] Etilevodopa. Drugs Fut 2001;26:219–23.
[50] Frey HH, Popp C, Löscher W. Influence of inhibitors of the high affinity GABA uptake on seizure threshold in mice. Neuropharmacology 1979;18:581–90.
[51] Ulm EH, Hichens M, Gomez HJ, Till AE, Hand E, Vassil TC, et al. Enalapril maleate and a lysine analogue (MK-521): disposition in man. Br Pharmaco 1982;14:357–62.
[52] Swanson BL, Vlasses PH, Ferguson RK, Berquist PA, Till AE, Irvin JD, et al. Influence of food on the bioavailability of enalapril. J Pharm Sci 1984;73:1655–7.
[53] Zannad F. Trandolapril: how does it differ from other angiotensin converting enzyme inhibitors? Drugs 1993;46:172–83.
[54] Tsybina N, Ostrovskaya RU, Protopova TV, Parin VV, Selezneva NI, Skolainov AP. Synthesis and pharmacological activity of gamma-aminobutyric acid derivatives. Khim Pharm Zh 1974;17:10–13.
[55] Ostrovskaya RU, Parin VV, Tsybina NM. The comparative neurotropic potency of gamma-aminobutyric acid and its cetyl ester. Byul Eks Biol Med 1972;73:51–5.
[56] Beardsley GP, Rosowsky A. Effect of methotrexate γ-monohexadecyl ester (γ-MHxMTX) on nucleoside uptake by human leukemic cells. Proc Am Assoc Cancer Res 1980;21:264. 21, 71 Meeting.
[57] Wermuth CG, Gaignault J-C, Marchandeau C. Designing pro-drugs and bioprecursors I: carrier prodrugs. In: Wermuth CG, editor. The practice of medicinal chemistry. London: Academic Press; 1996. p. 671–96.
[58] Jones G. Lipoidal pro-drug analogues of various anti-inflammatory agents. Chem Ind (London) 1980;7:452–6.
[59] Jansen ABA, Russell TJ. Some novel penicillin derivatives. J Chem Soc 1965;:2127–32.
[60] Binderup E, Godtfredsen WO, Roholt K. Orally active cephaloglycin esters. J Antibiot 1971;24:767–73.
[61] Saari WS, Halczenko W, Cochran DW, Dobrinska MR, Vincek WC, Titus DC, et al. 3-Hydroxy-α-methyltyrosine progenitors: Synthesis and evaluation of some (2-oxo- 1,3-dioxol-4-yl)methyl esters. J Med Chem 1984;27:713–17.
[62] Falch E, Krogsgaard-Larsen P, Christensen AV. Esters of isoguvacine as potential prodrugs. J Med Chem 1981;24:285–9.
[63] Svahn CM, Merenyi F, Karlson L, Widlund L, Grälls M. Tranexamic acid derivatives with enhanced absorption. J Med Chem 1986;29:448–53.
[64] Torriani H. Talniflumate. Drugs Fut 1979;4:448–50.
[65] Torriani H. Talmetacin. Drugs Fu 1982;7:823–4.
[66] Alexander J, Fromtling RA, Bland JA, Pelak BA, Gilfillan EC. Acyloxy)alkylcarbamate prodrugs of norfloxacin. J Med Chem 1991;34:78–81.
[67] Vepsalainen JJ. Bisphosphonate prodrugs. Curr Med Chem 2002;9:1201–8.
[68] Ahlmark M, Vepsaelaeinen J, Taipale H, Niemi R, Jaervinen T. Bisphophonate prodrugs: Synthesis and in vitro evaluation of novel clodronic acid dianhydrides as bioreversible prodrugs of clodronate. J Med Chem 1999;42:1473–6.
[69] Ebetino HF, Mazur AW, Dobson RLM. Bisphosphonate compounds. US2011/0098251 A1 (Patent); 2011.
[70] Johansson N.G., Zhou X.-X., Wahling H., Sund C., Wallberg H., Salvador L., et al. Prodrugs of phosphorous-containing pharmaceuticals. WO 99/41275 (Patent); 1999.
[71] Aicher TD, Bebernitz GR, Bell PA, Brand LJ, Dain JG, Deems R, et al. Hypoglycemic prodrugs of 4-(2,2-dimethyl-1-oxopropyl)benzoic acid. J Med Chem 1999;42:153–63.
[72] Young DA, Ho RS, Bell PA, Cohen DK, McIntosh RH, Nadelson J, et al. Inhibition of hepatic glucose production by SDZ 51641. Diabetes 1990;39:1408–13.
[73] Walsh DA, Moran HW, Shamblee DA, Welstead Jr WJ, Nolan JC, Sancilio LF, et al. Anti-inflammatory agents. 4. Syntheses and biological evaluation of potential prodrugs of 2-amino- 3-benzoylbenzeneacetic acid and 2-amino-3-(4-chlorobenzoyl)benzene acetic acid. J Med Chem 1990;33:2296–304.

[74] Galzinga L, Garbin L, Bianchi M, Marzotto A. Properties of two derivatives of γ-aminobutyric acid (GABA) capable of abolishing cardiazol- and bicuculline-induced convulsions in the rat. Arch Int Pharmacodyn 1978;235:73–85.
[75] Ahtee L, Halmekoski J, Heinonen H, Koskimies A. Comparison of the central nervous system actions of taurine and N-pivaloyltaurine. Br J Pharmacol 1979;66:480.
[76] Cundy KC, Annamalai T, Bu L, De Vera J, Estrela J, Luo W, et al. XP13512[(±)-1-([(α-Isobutanoyloxyethoxy)carbonyl]aminomethyl)-1-cyclohexand acetic acid], a novel gabapentin prodrug: II. Improved oral bioavailability, dose proportionality, and colonic absorption compared with Gabapentin in rats and monkeys. JPET 2004;311:324–33.
[77] Eriksson UG, Bredberg U, Hoffmann KJ, Thuresson A, Gabrielsson M, Ericsson H, et al. Absorption, distribution, and excretion of ximelagatran, an oral direct thrombin inhibitor in rats, dogs, and humans. Drug Metab Dispos 2003;31:294–305.
[78] Blech S, Ebner T, Ludwig-Schwellinger E, Stangier J, Roth W. The metabolism and disposition of the oral direct thrombin inhibitor, Dabigatran, in humans. Drug Metab Dispos 2008;36:386–99.
[79] Buur A, Bundgaard H, Falch E. Prodrugs of 5-fluorouracil. VII. Hydrolysis kinetics and physicochemical properties of N-ethoxy- and N-phenoxycarbonylmethyl derivatives of 5-fluorouracil. Acta Pharm Suec 1986;23:205–16.
[80] Larsen JD, Bundgaard H. Prodrug forms for the sulfonamide group. I. Evaluation of N-acyl derivatives, N-sulfonylamidines, N-sulfonyl-sulfilimines, and sulfonylureas as possible prodrug derivatives. Int J Pharm 1987;37:87–95.
[81] Bundgaard H, Buur A, Hansen KT, Larsen JD, Moss J, Olsen L. Prodrugs as drug delivery systems. 77. Phthalidyl derivatives as prodrug forms for amides, sulfonamides, carbamates and other NH-acidic compounds. Int J Pharm 1988;45:47–57.
[82] Leamon CP. Folate-targeted drug strategies for the treatment of cancer. Curr Opin Investig Drugs 2008;9:1277–86.
[83] Leamon CP, Reddy JA, Vlahov IR, Westrick E, Dawson A, Dorton R, et al. Preclinical antitumor activity of a novel folate-targeted dual drug conjugate. Mol Pharm 2007;4:659–67.
[84] Kim S-H, de Mas N, Parlanti L, Lyngberg OK, Strohlein G, Guo Z, et al. Synthesis, chromatographic purification, and isolation of epothilone-folic acid conjugate BMS-753493. Org Process Res Dev 2011;15:797–809.
[85] Leamon CP, Reddy JA. Folate-targeted chemotherapy. Adv Drug Deliv Rev 2004;56:1127–41.
[86] Yang J, Chen H, Vlahov IR, Cheng JX, Low PS. Evaluation of disulfide reduction during receptor-mediated endocytosis by using FRET imaging. Proc Natl Acad Sci 2006;103:13872–7.
[87] Rautio J. Prodrugs and targeted delivery: towards better ADME properties. Weinheim: Wiley-VCH; 2011.
[88] Chen K, Chen X. Integrin targeted delivery of chemotherapeutics. Theranostics 2011;1:189–200. Review.
[89] Arap W, Pasqualini R, Ruoslahti E. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. Science 1998;279:377–80.
[90] Kim JW, Lee HS. Tumor targeting by doxorubicin-RGD-4C peptide conjugate in an orthoopic mouse hepatoma model. Int J Mol Med 2004;14:529–35.
[91] Danhier F, Le Breton A, Preat V. RGD-based strategies to target alpha(v) beta(3) integrin in cancer therapy and diagnosis. Mol Pharmaceutics 2012;9:2961–73.
[92] Xiong JP, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, et al. Crystal structure of the extracellular segment of integrin αvβ3 in complex with an Arg-Gly-Asp ligand. Science 2002;296:151–5.
[93] Pilkington-Miksa M, Arosio D, Battistini L, Belvisi L, DeMatteo M, Vasile F, et al. Design, synthesis, and biological evaluation of novel cRGD paclitaxel conjugates for integrin-assisted drug delivery. Bioconjug Chem 2012;23:1610–22.
[94] Ai S, Duan J, Liu X, Bock S, Tian Y, Huang Z. Biological evaluation of a novel doxorubicin-peptide conjugate for targeted delivery to EGF receptor-overexpressing tumor cells. Mol Pharmaceutics 2011;9:375–86.
[95] Ai S, Jia T, Ai W, Duan J, Liu Y, Chen J, et al. Targeted delivery of doxorubicin through conjugation with EGF recptor-binding peptide overcomes drug resistance in human colon cancer cells. B J Pharma 2013;168:1719–35.
[96] Malakoutikhah M, Teixido M, Giralt E. Shuttle-mediated drug delivery to the brain. Angew Chem Int Ed 2011;50:7998–8014.
[97] Dalpiaz A, Filosa R, de Caprariis P, Conte G, Bortolotti F, Biondi C, et al. Molecular mechanism involved in the transport of a prodrug dopamine glycosyl conjugate. Inter J Pharm 2007;336(1):133–9.
[98] Gynther M, Ropponen J, Laine K, Leppanen J, Haapakoski P, Peura L, et al. Glucose promoiety enables glucose transporter mediated brain uptake of ketoprofen and indomethacin prodrugs in rats. J Med Chem 2009;52:3348–53.
[99] Gabathuler R. Approaches to transport therapeutic drugs across the blood–brain barrier to treat brain diseases. Neurobiol Dis 2010;37:48–57.
[100] Che C, Yang G, Thiot C, Lacoste M-C, Currie J-C, Demeule M, et al. New Angiopep-modified doxorubicin (ANG1007) and etoposide (ANG1009) chemotherapeutics with increased brain penetration. J Med Chem 2010;53:2814–24.
[101] Janoria KG, Gunda S, Boddu S, Mitra AK. Novel approaches to retinal drug delivery. Expert Opin Drug Deliv 2007;4:371–88.
[102] Dahlin A, Geier E, Stocker SL, Cropp CD, Grigorenko E, Bloomer M, et al. Gene expression profiling of transporters in the solute carrier and ATP-binding cassette superfamilies in human eye substructures. Mol Pharm 2013;10:650–63.
[103] Barot M, Bagui M, Gokulgandhi MR, Mitra AK. Prodrug strategies in ocular drug delivery. Med Chem 2012;8:753–68.
[104] Vooturi SK, Kadam RS, Kompella UB. Transporter targeted gatifloxacin prodrugs: Synthesis, permeability, and topical ocular delivery. Mol Pharm 2012;9:3136–46.
[105] Haas M, Moolenaar F, Meijer DKF, de Zeeuw D. Specific drug delivery to the kidney. Cardiovasc Drugs Ther 2002;16:489–96.
[106] Chae SY, Jang MK, Nah JW. Influence of molecular weight on oral absorption of water soluble chitosans. J Control Release 2005; 102:383–94.
[107] Park JH, Saravanakumar G, Kim K, Kwon IC. Targeted delivery of low molecular drugs using chitosan and its derivatives. Adv Drug Deliv Rev 2010;62:28–41.
[108] Yuan ZX, Sun X, Gong T, Ding H, Fu Y, Zhang ZR. Randomly 50% N-acetylated low molecular weight chitosan as a novel renal targeting carrier. J Drug Target 2007;15:269–78.
[109] Yuan ZX, Zhang ZR, Zhu D, Sun X, Gong T, Liu J, et al. Specific renal uptake of randomly 50% N-acetylated low molecular weight chitosan. Mol Pharm 2009;6:305–14.

[110] Wess G, Kramer W, Schubert G, Bickel M, Hoffmann A, Baringhaus KH. Coupling of drugs to modified bile acids for liver specific targeting. Abst Pap Am Chem Soc 1993;205(1–2):16 Pt. 1, MEDI 152.
[111] Huang J, Gao F, Tang X, Yu J, Wang D, Liu S, et al. Liver-targeting doxorubicin-conjugated polymeric prodrug with pH-triggered drug release profile. Poly Int 2010;59:1390–6.
[112] Liu J, Gray WD, Davis ME, Luo Y. Peptide- and saccharide-conjugated dendrimers for targeted drug delivery: a concise review. Interface Focus 2012;2:307–24.
[113] Aboul-Fadl T, Sheha MM, El-Azab AS, Abdel-Aziz HA. Amantadine amides prodrugs as hepatic delivery systems to enhance its activity against HCV. Dig J Nano Bio 2011;6:1675–83.
[114] Bagshawe KD, Sharma SK, Begent R. Antibody-directed enzyme prodrug therapy (ADEPT) for cancer. Expert Opin Biol Ther 2004;4:1777–89.
[115] NATAP. Valeant Seeks Partner for Pradefovir. Available from: <http://www.natap.org/2006/newsUpdates/041406_03.htm>; 2006.
[116] Erion MD, van Poelje PD, MacKenna DA, Colby TJ, Montag AC, Fujitaki JM, et al. Liver-targeted drug delivery using HepDirect prodrugs. JPET 2005;312:554–60.
[117] Orlowski M, Mizoguchi H, Wilk S. N-acyl-γ-glutamyl derivatives of sulfamethoxazole as models of kidney-selective prodrugs. J Pharmacol Exp Ther 1979;212:167–72.
[118] Cook AF, Holman MJ, Kramer MJ, Trown PW. Florinated pyrimidine nucleosides. 3. Synthesis and antitumor activity of a series of 5′-deoxy-5fluoropyrimidine nucleosides. J Med Chem 1979;22:1330–5.
[119] Au JL-S, Walker JS, Rustum Y. Pharmacokinetic studies of 5-fluorouracil and 5′-deoxy-5-fluorouridine in rats. J Pharmacol Exp Ther 1983;227:174–80.
[120] Stern L, Perry R, Ofek P, Many A, Shabat D, Satchi-Fainaro R. A novel antitumor prodrug platform designed to be cleaved by the endoprotease Legumain. Bioconjug Chem 2009;20:500–10.
[121] Schmidt F, Monneret C. Prodrug mono therapy: synthesis and biological evaluation of an etoposide glucuronide-prodrug. Bioorg Med Chem 2003;11:2277–83.
[122] Legian T, Clarhaut J, Tranoy-Opalinski I, Monvoisin A, Renoux B, Thomas M, et al. The first generation of β-galactosidase-responsive prodrugs designed for the selective treatment of solid tumors in prodrug monotherapy. Angew Chem Int Ed 2012;51:11606–10.
[123] Leriche G, Chisholm L, Wagner A. Cleavable linkers in chemical biology. Bioorg Med Chem 2012;20:571–82.
[124] Flygare JA, Pillow TH, Aristoff P. Antibody drug conjugates for the treatment of cancer. Chem Biol Drug Des 2013;81:113–21.
[125] Alley SC, Okeley NM, Senter PD. Antibody-drug conjugates: targeted drug delivery for cancer. Curr Opin Chem Bio 2010;14:529–37.
[126] Guardino E. Trastuzumab-DM1 (T-DM1) antibody-drug conjugate bridging targeted therapy and conventional chemotherapy. 2010 presentation.
[127] Erickson HK, Lambert JM. ADME of antibody-Maytansinoid conjugates. AAPS Journal 2012;14:799–805.
[128] Mathew J, Perez EA. Trastuzumab emtansine in human epidermal growth factor receptor 2-positive breast cancer: a review. Curr Opin in Oncol 2011;23:594–600.
[129] FDA. FDA approves new treatment for late-stage breast cancer. Available from: <http://www.fda.gov/newsevents/newsroom/pressannouncements/ucm340704.htm>; 2013.
[130] Bundgaard H. Novel chemical approaches in prodrug design. Drugs Fut 1991;16:443–58.
[131] Cain BF. The role of structure–activity studies in the design of antitumor agents. Cancer Chemother Rep 1975;59:679–83.
[132] Cain BF. 2-Acyloxymethylbenzoic acids: novel amine protective functions providing amides with the lability of esters. J Org Chem 1976;41:2029–31.
[133] Sakamoto F, Ikeda S, Tsukamoto G. Studies on prodrugs. II. Preparation and characterization of (5-substituted 2-oxo-1,3-dioxolen-4-yl)methyl esters of ampicillin. Chem Pharm Bull 1984;32:2241–8.
[134] Norbeck DW, Rosenbrook W, Kramer JB, Grampovnik DJ, Lartey PA. A novel prodrug of an impermeant inhibitor of 3-deoxy-d-manno-2-octulosonate cytidylyl-transferase has antibacterial activity. J Med Chem 1989;32:625–9.
[135] Borgman RJ, Smith RV, Keiser JE. The acetylation of apomorphine: An improved method for the selective preparation of diacetylapomorphine utilizing trifluoroacetic acid/acetyl bromide. Synthesis 1975;4:249–50.
[136] Nicolaou KC, Riemer C, Kerr MA, Rideout D, Wrasidlo W. Design, synthesis, and biological activity of protaxols. Nature (London) 1993;364:464–6.
[137] Nicolaou KC, Guy RK, Pitsinos EN, Wrasidlo W. A water-soluble prodrug of taxol with self-assembling properties. Angew Chem Int Ed 1994;33:1583–7.
[138] Amsberry KL, Borchardt RT. Amine prodrugs which utilize hydroxy-amide lactonization. I. A potential redox-sensitive amide prodrug. Pharm Res 1991;8:323–30.
[139] Amsberry KL, Gerstenberger AE, Borchardt RT. Amine prodrugs which utilize hydroxy-amide lactonization. II. A potential esterase-sensitive prodrug. Pharm Res 1991;8:455–61.
[140] Leppänen J, Huuskonen J, Nevalainen T, Gynther J, Taipale H, Järvinen T. Design and synthesis of a novel L-dopa-entacapone codrug. J Med Chem 2002;45:1379–82.
[141] Carini DJ, Duncia JV, Aldrich PE, Chiu AT, Johnson AU, Piera ME, et al. Non-peptide angiotensin II antagonists: the discovery of a series of N-(biphenylmethyl) imidazoles as potent, orally active antihypertensives. J Med Chem 1991;;34:2525–47.
[142] Wong PC, Price WAJ, Chiu AT, Duncia JV, Carini DJ, Wexler RR, et al. Non-peptide angiotensin II receptor antagonists. XI. Pharmacology of Exp 3174: an active metabolite of DuP 753, an orally active antihypertensive agent. J Pharmacol Exp Ther 1990;255:211–17.
[143] Venhuis BJ, Wikström HV, Rodenhuis N, Sundell S, Dijkstra D. A new type of prodrug of catecholamines: an opportunity to improve the treatment of Parkinson's disease. J Med Chem 2002;45:2349–51.
[144] Bodor N, Shek E, Higuchi T. Improved delivery through biological membranes. 1. Synthesis and properties of 1-methyl-1,6-dihydropyridine-2-carbaldoxime, a pro-drug of N-methylpyridinium- 2-carbaldoxime chloride. J Med Chem 1976;19:102–7.
[145] Shek E, Higuchi T, Bodor N. Improved delivery through biological membranes. 2. Distribution, excretion, and metabolism of N-methyl-1,6-dihydropyridine-2-carbaldoxime hydrochloride, a pro-drug of N-methylpyridinium-2-carbaldoxime chloride. J Med Chem 1976;19:108–12.

[146] Shek E, Higuchi T. Improved delivery through biological membranes. 3. Delivery of improved delivery through biological membranes. 3. Delivery of N-methylpyridinium-2-carbaldoxime chloride through the blood–brain-barrier in its dihydropyridine pro-drug form. J Med Chem 1976;19:113–17.
[147] Ling KH, Leeson GA, Burmaster SD, Hook RH, Reith MK, Cheng LK. Metabolism of terfenadine associated with CYP3A(4) activity in human hepatic microsomes. Drug Metab Dispos 1995;23:631–6.
[148] Krenitsky TA, Hall WW, de Miranda P, Beauchamp LM, Schaeffer HJ, Whiteman PD. Deoxyacyclovir: a xanthine oxidase-activated prodrug of acyclovir. Proc Natl Acad Sci USA 1984;31:3209–13.
[149] Whiteman PD, Bye A, Fowle ASE, Jeal S, Land G, Posner J. Tolerance and pharmacokinetics of A515U, an acyclovir analogue in healthy volunteers. Eur J Clin Pharmacol 1984;27:471–5.
[150] Gershwin ME, Goetzl EJ, Steinberg AD. Cyclophosphamide: use in practice. Ann Int Med 1974;80:531–40.
[151] Silverman RB. The organic chemistry of drug design and drug action, 1992. San Diego, CA: Academic Press; 1992. p. 380.
[152] Zon G. Cyclophosphamide analogues. In: Ellis GP, West GB, editors. Progress in medicinal chemistry. Amsterdam: Elsevier; 1982. p. 205–46.
[153] Arnold H, Bourseaux F, Brock N. NeuartigeKrebs- Chemotherapeutika aus der Gruppe der zyklischen N-Lost-Phosphamidester [Novel chemotherapeutic cancer agents from the group of cyclic nitrogen mustard phosphamide ester]. Naturwissenschaften 1958;45:64–6. [in German].
[154] Arnold H, Bourseaux F, Brock N. Chemotherapeutic action of a cyclic nitrogen mustard phosphamide ester (B518-ASTA) in experimental tumors of the rat. Nature 1958;181:931.
[155] Arnold H. Ueber die Chemie neuer zytostatisch wirksamer N- Chloroaethyl-phosphorsäureesterdiamide [About the chemistry of new cytostatic effective N- chloroaethyl phosphoric acid ester amides]. In: Spitzy K, Haschek H, editors. Proceedings of the fifth international congress on chemotherapy. Wien: Verlag der Wiener Medizinischen Akademie; 1967. p. 751–4. [in German].
[156] Kennedy KA, Teicher BA, Rockwell S, Sartorelli AC. The hypoxic tumor cell: a target for selective cancer chemotherapy. Biochem Pharmacol 1980;29:1–8.
[157] Davis PJ, Guenthner LE. Sulindac oxidation/reduction by microbial cultures: Microbial models for mammalian metabolism. Xenobiotica 1985;15:845–57.
[158] Kwon C-H, Blanco DR, Batura N. p-(Methylsulfinyl)phenyl nitrogen mustard as a novel bioreductive prodrug selective against hypoxic tumors. J Med Chem 1992;35:2137–9.
[159] Cole C, Reigan P, Gbaj A, Edwards PN, Douglas KT, Stratford IJ, et al. Potential tumor-selective nitroimidazolylmethyluracil prodrug derivatives: Inhibitors of the angiogenic enzyme thymidine phosphorylase. J Med Chem 2003;46:207–9.
[160] Wallmark B, Brändström A, Larsson H. Evidence for acid-induced transformation of omeprazole into an active inhibitor of ($H^+ + K^+$)-ATPase within the parietal cell. Biochim Biophys Acta 1984;778:549–58.
[161] Lamers CBHW, Lind T, Moberg S, Jansen JBM, OIbe L. Omeprazole in Zollinger-Ellison Syndrome. N Engl J Med 1984;310:758–61.
[162] Im WI, Shi JC, Blakeman DP, McGrath JP. Omeprazole, a specific inhibitor of gastric (H^+-K^+)-ATPase, is a H^+-activated oxidizing agent of sulfhydryl groups. J Biol Chem 1985;260:4591–7.
[163] Lindberg P, Nordberg P, Alminger T, Brändström A, Wallmark B. The mechanism of action of the gastric acid secretion inhibitor omeprazole. J Med Chem 1986;29:1327–9.
[164] Cousse H, Casadio S, Mouzin G. L'hydroxy éthyl nicotinamide vecteur d'acides thérapeutiquement actifs [Hydroxyethyl nicotinamide, a vector for therapeutically active acids]. Trav Soc Pharm Montpellier 1978;38:71–6. [in French].
[165] Casadio S, Cousse H, Mouzin G. Nouvelles formes modulées de médicaments utilisant les N-hydroxy alcoyl pyridine carboxamides comme vecteur [New forms of drugs modulated using N-hydroxy alkyl pyridine carboxamides as a vector]. 77.13478 (Patent) 1977 [in French].
[166] Vezin JC, Mouzin G, Cousse H, Casadio S. Nicafenine, a new analgesic. Arzneim Forsch 1979;29:1659–61.
[167] Gorrod JW. Potential hazards of the pro-drug approach. Chem Ind 1980;:457–61.
[168] Roberts JC, Nagasawa HT, Zera RT, Fricke RF, Goon DJW. Prodrugs of l-cysteine as protective agents against acetaminophen-induced hepatotoxicity. 2-(Polyhydroxyalkyl)- and 2- (polyacetoxyalkyl)thiazolidine-4(R)-carboxylic acids. J Med Chem 1987;30:1891–6.
[169] Saari WS, Freedman MB, Hartman RD, King SW, Raab AW, Randall WC, et al. Synthesis and antihypertensive activity of some ester progenitors of methyldopa. J Med Chem 1978;21:746–53.
[170] Vickers S, Duncan CA, White SD, Breault GO, Boyds RB, de Shepper PJ, et al. Evaluation of succinimidoethyl and pivaloyloxyethyl esters as progenitors of methyldopa in man, rhesus monkey, dog, and rat. Drug Metab Dispos 1978;6:640–6.

SECTION SIX

Chemical Modifications Influencing the Pharmacokinetic Properties

CHAPTER

29

Drug Delivery with Organic Solvents or Colloidal Dispersed Systems

Bernd U. Riebesehl

Technical Research & Development, Novartis Pharma AG, Basel, Switzerland

OUTLINE

I. INTRODUCTION

In general, only drug molecules dissolved in body fluids can permeate biological barriers and reach their site of action on a protein or cell surface or within certain intracellular structures. Therefore, the drug dose, its solubility in relevant body fluids, its release from delivery systems, and its permeability through biological barriers have a major impact on the therapeutic outcome. The selection of a suitable drug solubilization system depends on the physicochemical properties of the drug substance, the development phase, the dosing needs for the indication, the route of administration, and its related dosing and excipient constraints.

Physicochemical properties like high melting point or high heat of fusion limit the solubility of drugs in any solvent. Drug properties like lipophilicity, hydrophilicity, or amphiphilicity have to be taken into account when selecting the vehicle that solubilizes the drug most efficiently for its distribution in certain biologic fluids on the way toward its biologic target.

Each development phase has its specific priorities. In drug discovery, simple solvent systems are preferred for compound screening. For preclinical animal dosing where high doses shall be delivered, the tolerance of solubilizing vehicles constrains the choices, especially for the parenteral administration routes.

Each medical indication leads to specific patient and dosing needs that drive the selection of administration routes and solubilizing vehicles (e.g., emergency, short term, chronic, patient- and situation-dependent dosing).

Oral dosing is the preferred route of drug dosing, as it is convenient for most patients. However, only 803 of the 1,357 unique drugs acting on the known 337 human and pathogen drug targets have properties that they can be dosed orally, while 421 must be administered parenterally and 275 are delivered topically (including inhalation) [1]. The deliverable drug doses also differ widely by administration route (e.g., oral ≫ transdermal). This indicates that drug properties and dosage forms need to be tailored to the administration route to enable the

desired pharmacological action. This chapter is structured by administration route so that the drug-delivery system choices become more obvious.

Excipients are formulation components that support the administration, efficacy, safety, or stability of drug substances. However, their use is sometimes limited in concentration or dose and dependent on the route of administration. In practice this turns out to be a significant constraint for the feasibility of drug-delivery systems.

II. PHYSICOCHEMICAL DRUG PROPERTIES

Early in drug discovery, large libraries of substances are screened for their effects on biological targets. For this *in vitro* screening, drug candidate solutions in dimethyl sulfoxide are often used, as it dissolves a wide variety of compound types, including compounds that are practically insoluble in water. Further, it exhibits relatively low toxicity to most types of cells and has low chemical reactivity. For preclinical dosing to animals, more complex drug formulations are needed that provide for solubility, bioavailability, and safety, so that pharmacologic effects can be observed in dosing conditions relevant to later human use. Especially for the preclinical safety assessment, dosing needs to be escalated until the maximum tolerated dose is reached. A sequential investigation is applied for the selection for solubilizing media:

1. For ionizable drugs, a search for highly water-soluble salts or co-crystals should be undertaken to identify a suitable salt and solid form for the drug substance that is beneficial for stability and drug performance. Thus, salts and co-crystals may support supersaturating dosage forms, that is, if a drug has a higher solubility in the acidic stomach and a decreasing solubility during gastrointestinal passage as the pH increases in the intestine. Salt solubility is often 1000-fold higher than the solubility of the nonionic species [2].
2. For liquid dosage forms, the identification of solubilizing cosolvent−water mixtures that are acceptable for the route of administration and compatible with the dosage form is the next step. Such cosolvent solutions or cosolvents−water mixtures can be supersaturating dosage forms, as the drug solubility decreases in a log-linear way when diluted in water or body fluids. The addition of surfactants, lipids, or polymers may support the solution stability after such dilution.
3. Water-soluble complexes with cyclodextrins may be another option for liquid or solid dosage forms. This solubilization approach is covered in Chapter 32.
4. Synthetic modification of the drug structure by the covalent attachment of a water-solubilizing group (e.g., through a biodegradable linker) allows for the modification of solubility and even the plasma half-life (e.g., polyethylene glycol (PEG)) [3]. Likewise, the drug could be made more hydrophobic and less water-soluble by a covalent attachment of a lipophilic residue—again through a biodegradable linker—of a hydrophobic group [4]. This latter approach leads to a drug that may be more suitable for lipophilic solutes or nanoparticles that may improve drug targeting or enable injection depot formulations [5]. This approach is further discussed in Chapter 28.
5. Colloidal drug-delivery systems offer improved therapeutics but are also more difficult to design and manufacture. Liposomes, emulsions, nanoparticle, and solid-dispersion formulations enabling solubilization, membrane permeability, controlled release, or targeted delivery have been successfully marketed.

Which dosage form and drug-delivery approach is viable and preferable for an individual drug depends strongly on the anticipated route of administration. In the following sections, the possible drug-delivery approaches and dosage forms are discussed according to their administration route.

III. ORAL DRUG DELIVERY

When taken orally, small molecules must dissolve during their gastrointestinal passage before they pass the site of absorption. Further, they need to be stable in the gastrointestinal fluids and have properties that enable either their passive permeation following a concentration gradient or active transport into blood circulation.

After evaluating a large data set (2,245 compounds) of marketed oral drugs [6] Lipinski et al. reached a series of conclusions on the optimum properties of drugs for oral administration. In order to achieve efficient dissolution and passive oral absorption and cell permeability, drug candidates should have:

- fewer than five hydrogen bond donors (sum of OH and NH groups);
- a molecular weight less than 500 Da;

- a log P less than 5; and
- fewer than ten hydrogen bond acceptors (sum of O and N atoms in the structure).

This rapidly became known as "Lipinski's Rule of Five," as all the cut-off points are multiples of five. Some drugs and natural nutrients (e.g., amino acids, vitamins) cannot passively diffuse across biomembranes but are actively absorbed via membrane transporters.

Vieth investigated the factors influencing oral bioavailability by comparing marketed oral drugs and other marketed drugs. In contrast to oral drugs, injectables can have significantly higher mean molar weight, more oxygen or nitrogen atoms, more hydroxyl or NH groups, more rings, more rotatable bonds, and higher proton acceptor counts, while exhibiting lower mean calculated log P and halogen counts. Generally the differences are quite large and indicate that injectable drugs are significantly heavier, more polar, and more flexible than oral drugs. Interestingly, in four of these seven properties (MW, OHNH, rotatable bonds, and acceptor count), the oral and injectable sets are at or near the extremes for mean values across all the sets examined. We can therefore look at the means of these two groups as two extreme cases of physical properties acceptable for xenobiotics [7].

A. Dose, Solubility, Permeability, and Absorption

The amount of oral absorption and bioavailability is of great importance for drug efficacy and safety. At the same time, it is the result of sequential processes that are interrelated:

A high dose/solubility ratio may limit complete dissolution of a given dose during its time of gastrointestinal passage. On the other hand, drug permeability may also limit the fraction of a drug dose that is absorbed to become effective.

Amidon published a seminal paper on the biopharmaceutics classification system (BCS) that concerns the evaluation of the bioequivalence of immediate-release solid oral dosage forms. Based on their solubility and oral absorption characteristics, Amidon defined four different categories of drugs: [8]

1. Class I: High solubility, high permeability
2. Class II: Low solubility, high permeability
3. Class III: High solubility, low permeability
4. Class IV: Low solubility, low permeability

Basically, the US FDA has adopted this classification system as one of their "Guidance for Industry" series of documents [9].

In early development, another concept related to solubility and absorption gives useful information on the oral absorption potential of drug candidates. This is the maximum absorbable dose (MAD) as defined by Equation 29.1: [10]

$$\mathrm{MAD} = \mathrm{S} \times K_a \times \mathrm{SWWW} \times \mathrm{SITT} \tag{29.1}$$

where S = the drug solubility at pH 6.5 (reflects typical pH of small intestine), K_a = the absorption rate constant, which can be determined from perfusion experiments using rat small intestine, SWWW = the water volume in the small intestine that is available for drug dissolution (generally accepted to be 250 mL), and SITT = the small intestine transit (generally accepted to be approximately 270 min). While the equation was developed for K_a values experimentally determined using a portion of rat small intestine, it can be used to give an approximate range for the minimum acceptable solubility required for a given target dose. It has been reported [11] that the majority of drug substances have human K_a values in the range $0.001-0.05\ \mathrm{min}^{-1}$. If these values are used in Equation 29.2 together with a value for the MAD (or target dose), an approximate range for the target solubility required is calculated as:

$$\text{Minimum acceptable solubility} = \text{Target dose}/(\mathrm{S} \times K_a/\mathrm{MAD}) \tag{29.2}$$

Thus, for a target dose of 70 mg using the range of the human K_a minimum, solubilities between 20 μg/mL and 1 mg/mL are obtained.

Now more sophisticated *in silico* tools like Gastroplus® or SYMCIP® facilitate the modeling of plasma profiles and its underlying drug property-related and physiological factors [12]. For early formulation development for preclinical studies decision trees have been worked out. For oral drug candidates with very a limited fraction absorbed it is advisable to test multiple formulation approaches based on nanoparticles, solid dispersions or lipid-based drug-delivery systems.

Solubility of active pharmaceutical ingredients can be improved by employing metastable states that form supersaturating drug-delivery systems *in vivo* [13]. Promising formulation technologies used to overcome the problem of poor solubility include solid dispersions, lipid-based systems, or nanoparticle-based formulations that are manufactured as tablets, hard or soft gelatin capsules, or oral liquid formulations.

B. Dissolution Enhancement, Nanoparticles

Use of solubilizing prodrugs or salt formation are discussed in Chapter 30. Particle size reduction is used to enhance the dissolution rate so that more drug is in solution earlier during its transit through the intestine. Several marketed drugs with limited oral absorption were developed where the drug substance was milled to nanosize and achieved a higher bioavailability, allowing a reduction of the dose as compared to larger particle size-based formulations [14]. Table 29.1 shows examples of non-ionizable drugs that are marketed as nanoparticle-based oral solid dosage forms or as a nanosuspension. A nanosuspension is easier to swallow; moreover, the dose can be adjusted for pediatric use.

C. Supersaturating Delivery Systems

Supersaturating delivery systems can be utilized to raise the dissolved drug concentration above the drug equilibrium solubility of least soluble solid crystal form in the dissolving or diluting body fluid. This can be achieved by employing metastable states, salt forms, co-crystals, amorphous forms, solid dispersions, cosolvent solutions, or lipid-based systems [13]. If supersaturated concentrations are maintained during the gastrointestinal passage up to the site of absorption, this can have a beneficial effect on the bioavailability of a solubility constrained drug. On the other hand uncontrolled crystal nucleation and (re-)crystallization may lead to larger particle sizes of a less soluble solid form of the drug, with reduced oral absorption as a consequence. The extent to which a drug may oversaturate before nucleation leads to crystallization can be experimentally evaluated by a pH-titration method [15]. Addition of formulation components may reduce the undesired tendency for drug crystallization or precipitation, respectively.

Solid dispersions and lipid-based systems are examples of supersaturating delivery systems.

D. Solid Dispersions

Solid dispersions are typically supersaturating drug-delivery systems for poorly water-soluble crystalline drugs. These are typically composed of a drug dispersed in a polymer with surfactants and/or other excipients added. For achieving a solid dosage form that is physically stable during long-term storage and for forming an oversaturated drug solution after dosing and hydration, it is essential that the drug is molecularly dispersed in the excipient matrix and does not crystallize. The excipients need to be selected to be highly miscible with the drug and among one another in order to prevent demixing of dispersion components and the risk of drug crystal nucleation and crystallization [16]. The selection of excipients that are well miscible with amorphous drug substance can be achieved by casting polymer films from highly concentrated solvent solutions. Clear dried films indicate suitable mixtures that do not show phase separation of the components in the film. Alternatively, drug solubility tests in liquid monomers of pharmaceutically suitable polymers may identify the more miscible polymers. Attention needs to be paid to the glass transition temperature of the solid dispersions. The solid dispersion is more protected against phase separation when in a glassy state. Any hydration or solvation of the solid dispersion would reduce its glass transition temperature. If the resulting glass transition temperature was below the

TABLE 29.1 Marketed Oral Dosage Forms Based on Nanocrystals

Dosage Form	Brand	Active	Molar weight	Melting point	Log P	pKa	Single high dose (mg)
Nanoparticle tablet	Rapamune	Sirolimus	914	NA	4.3	NA	2
Nanoparticle capsule	Emend	Aprepitant	534	NA	4.5	NA	125
Nanoparticle tablet	TriCor	Fenofibrate	361	81	5.3	NA	145
Nanosuspension	Megace ES	Megestrol acetate	342	220	3.2	NA	625

NA = Not available.

storage temperature, this might over time lead to phase separation and crystal formation and is therefore to be avoided. Further, it is important to identify compositions that prevent drug crystallization after wetting of the dosage form so that a supersaturated solution is maintained in gastrointestinal fluids for sufficient time to benefit drug absorption. Solid dispersions can be manufactured by melt extrusion, where drug and excipients are mixed thoroughly at elevated temperatures and filled into capsules or formed to tablets. Alternatively, solid dispersions can be spray dried from solutions in organic solvents. Again, the spray product can be filled into capsules or be compressed to tablets [17]. Table 29.2 shows examples of marketed solid-dispersion dosage forms of weakly basic or nonionizable drugs.

E. Lipid-Based Systems

Lipid-based drug-delivery systems (LBDDS) have been successfully marketed filled into oral soft or hard gelatin capsules. Alternatively, volumetric dosed LBDDS can be used for children or patients with swallowing difficulties.

LBBDS are typically composed of oils, surfactants, or cosolvents solubilizing lipophilic drugs at high concentration and preventing precipitation during the gastrointestinal passage. Dilution, lipid digestion, and diffusion of water-soluble components after administration may change the composition and alter their solubilizing properties. Pouton proposed a lipid formulation classification system (LFCS) and related characteristics as described in Table 29.3 [18].

A recent review offers a formulation strategy for poorly water-soluble drugs [19]. Table 29.4 shows examples of marketed lipid-based products. Drugs are used in their nonionic form and often developed in both a soft gel capsule and an oral liquid dosage form. For highly potent, very low-dosed compounds like calcitriol, liquid filling offers accurate dose filling with less containment issues as compared to powder-based manufacturing processes [20].

TABLE 29.2 Marketed Oral Solid Dispersion Dosage Forms

Dosage Form	Brand	Active	Carrier	Molar weight	Melting point	Log P	pKa	Single high dose (mg)
Film coated tablet	Kaletra	Lopinavir, Ritonavir	PVPVA	628	127	1.7	NA	200
				720	125	1.2	2.8 base	50
Hard gelatin capsule	Sporanox	Itraconazole	HPMC	706	166	6.5	3.4 base	100
Tablet	Intelence	Etravirine	HPMC	435	260	>5	<3 base	200
Hard gelatin capsule	Prograf	Tacrolimus	HPMC	822	129	NA	NA	5

NA = Not available, PVPVA = Polyvinylpyrrolidonevinylacetate, HPMC = Hydroxypropylmethylcellulose.

TABLE 29.3 Characteristic Features of the Various Types of Lipid Formulation Systems

LFCS type	Characteristics	Advantages	Disadvantages
Type I	Nondispersing, requires digestion	GRAS status, simple, excellent capsule compatibility	Formulation has poor solvent capacity unless drug is highly lipophilic
Type II	SEDDS without water-soluble components	Unlikely to lose solvent capacity on dispersion	Turbid o/w dispersion (particle size 0.25–2.0 μm)
Type IIIA	SEDDS/SMEDDS with water-soluble components	Clear or almost clear dispersion, drug absorption without digestion	Possible loss of solvent capacity on dispersion, less easily digested
Type IIIB	SEDDS/SMEDDS with water-soluble components and low oil content	Clear dispersion, drug absorption without digestion	Likely loss of solvent capacity on dispersion
Type IV	Oil-free formulation based on surfactants and cosolvents	Good solvent capacity for many drugs, disperses to micellar solution	Loss of solvent capacity on dispersion, may not be digestible

TABLE 29.4 Marketed Oral Dosage forms Based on Lipid-Based Drug-Delivery Systems

Dosage Form	Brand	Active	Lipid formulation class	Molar weight	Melting point	cLog P	pKa	Single high dose (mg)
Soft gel capsule	Convulex	Valproic acid	I	144	130	2.8	4.6 acid	500
Soft gel capsule	Rocaltrol	Calcitriol	III	416	113	4.4	NA	0.0005
Soft gel capsule	Aptivus	Tipranavir	III-IV	584	89	7.8	NA	2x250
Oral suspension	Ciprobay	Ciprofloxacin	III	331	257	−0.8	6.1 base	500

NA = not available

A large number of lipids occur naturally, but not all are available commercially or established as pharmaceutical excipients. The Lipid Library [21] is a useful source of general information on most classes of lipids, fatty acids, and eicosanoids. In addition to this, the Lipid Metabolites and Pathways Strategy [22] provides a classification system for all known lipids and gives information on their structures.

IV. PARENTERAL DRUG DELIVERY

Parenteral articles are preparations intended for injection through the skin or other external boundary tissue rather than through the alimentary canal, so that the active substances they contain are administered using gravity or force directly into a blood vessel, organ, tissue, or lesion. Parenteral articles are prepared scrupulously by methods designed to ensure that they meet pharmacopoeial requirements for sterility, pyrogens, particulate matter, and other contaminants, and—where appropriate—contain inhibitors of the growth of microorganisms. An injection is a preparation intended for parenteral administration and/or for constituting or diluting a parenteral article prior to administration [23]. Parenteral dosing overcomes biological barriers for drug molecules that cannot passively be absorbed by other administration routes (e.g., large hydrophilic or poorly water-soluble molecules).

A. Selection of Drug-Delivery System

Many factors influence the feasibility of injectable dosage form design:

1. The clinical indication may be life-threatening and usually has an established standard of care that defines whether patients are hospitalized, require injections by health care professionals, or deserve self-medication for use at home. In general, intravenous and intramuscular injections are only administered by healthcare professionals, while subcutaneous dosing is viable for self-dosing by the patients. The standard of care also determines the acceptable cost.
2. The treatment may be chronic and be related to specific patient needs. Chronic treatment deserves convenient dosing, for example, through self-dosing with prefilled syringes or injection devices or less frequent depot injections by healthcare professionals.
3. The therapeutic index of a drug may be low so that the ratio of toxic dose and the therapeutic dose is low. In such cases, flat plasma profiles are desired so that constant dosing by infusion, infusion pumps, controlled-release formulations, or injection depots may be beneficial.
4. Pharmacologic activity depends on successful transport to biological targets and require passage of multiple biological barriers (e.g., highly charged siRNA-drugs require cellular uptake and endosomal release to reach the cytosol). Such drug candidates with unfavorable physicochemical properties require sophisticated drug-delivery strategies.

1. Route of Administration

A. INTRAVENOUS

The intravenous route is the most frequently used because it is most flexible regarding injection volume, rate of dosing, and excipient use. Infusions provide for controlled and constant plasma profiles and maximum infusion volumes. The fast dilution of injected solutions in the blood stream allows for a greater tolerance of the drug

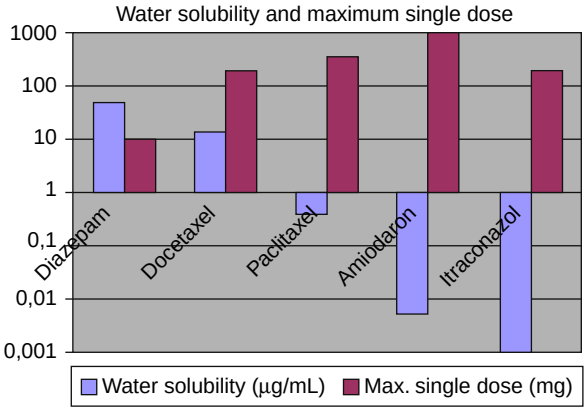


FIGURE 29.1 Water solubility and maximum single dose of marketed parenteral drugs.

product components with blood components or injection site tolerance as compared to intramuscular or subcutaneous dosing. Bolus injections are given in a short time, so that the plasma profile shows an initial concentration peak and decline.

B. INTRAMUSCULAR

Intramuscular injections require injection by healthcare professionals and are limited to volumes of up to 5 mL. This route can also be used for microparticle suspension depot injections, as particles pose no risk of blocking capillaries like the intravenous route. The released drug is transported by the lymph before it enters the circulation.

C. SUBCUTANEOUS

The subcutaneous route is normally limited to an injection volume of 1 mL. Safety considerations allows for self-medication by the patient. This route can also be used for microparticle or implant depot injections. Recently the injectable volume could also be increased by adding recombinant human hyaluronidase enzyme rHuPH20, which temporarily degrades hyaluronan, a structural protein in the interstitial space of the skin. This temporary alteration allows the injection of up to 5 mL of fluids. With this enzyme, pharmaceuticals that would normally be injected intravenously (i.v.) can be administered subcutaneously (s.c.). This change in route of delivery may improve patient convenience or enhance pharmacokinetics [24].

2. *Deliverable dose/Technology Type*

A. VOLUME OF DOSING

Often, parenteral small-molecule drugs require a high dosing volume, acidic or alkaline pH, or solubilizing excipients. In such cases, s.c. or i.m. dosing is not possible. Addition of solubilizing excipients like cosolvents frequently can bring about sufficient solubilization so that small injection volumes are achieved. Concentrates for infusion are particularly useful for highly dosed and poorly water-soluble drugs. Drug concentrates utilize the exponential increase in drug solubility by high cosolvent concentrations before infusion, and they are freshly diluted to safe cosolvent concentrations with standard infusion solutions such as glucose or saline solutions.

The drug dose / solubility ratio needs to be low for enabling small single-injection volumes per dose. Further, the drug and solution need to be chemically and physically stable over the shelf-life and in-use period of a drug product. Figure 29.1 shows the water solubility and maximum single dose of marketed parenteral drugs.

B. INJECTION DEVICES FOR PROGRAMMABLE, CONTINUOUS, AND SELF-DOSING

Patients with chronic diseases like diabetes that require continuous and individualized parenteral dosing benefit from new medical devices for home use. The standard of care evolved from subcutaneous prefilled syringes to elegant injection pens that enable dose adjustments and multiple injections from cartridge vials. Programmable subcutaneous infusion pumps can also be used by trained patients to enable convenient continuous dosing [25].

TABLE 29.5 Solvents and Cosolvents for Parenteral Use

Excipient	Frequency	Max. concentration	Example
Benzyl benzoate	3	44.7% w/v	Delestrogen 40 mg/mL (Bristol Myers)
Castor oil	2	11.50%	Delestrogen 40 mg/mL (Bristol Myers)
Cottonseed oil	1	87.4% w/v	Depo Testosterone (Pfizer)
N,N dimethylacetamide	2	33% w/v	Busulfex (Orphan Medical)
Ethanol/ethanol dehydrated	34	100%	Prograf (Fujisawa) 80% v/v, Alprostadil (Bedford Lab) 100%
Glycerin (glycerol)	17	70% w/v	Multitest CMI (Pasteur Merieux)
N-methyl-2-pyrrolidone	1[a]	[a]	Eligard 7.5 mg (Sanofi)
Peanut oil	1	[a]	Bal in Oil (Becton Dickinson)
PEG 300	4	65% w/v	VePesid (Bristol Myers)
PEG 400	4	67% v/v	Busulfex (Orphan Medical)
PEG 600	1	5% w/v	Persantine (Dupont-Merck)
PEG 3350	4	2.95% w/v	Depo-Medrol (Upjohn)
PEG 4000	1	3%	Invega Sustenna (Janssen)
Poppy seed oil	1	[a]	Ethiodol (Savage)
Propylene glycol	32	80%	Ativan (Wyeth-Ayerst)
Safflower oil	2	10%	Liposyn II (Abbott)
Sesame oil	7	100%	Solganal Injection (Schering)
Soybean oil	1	10% w/v	Diprivan Injection (Zeneca)
Vegetable oil	2	[a]	Virilon I.M. Injection (Star Pharmaceuticals)

[a]*Not applicable or no data available.*

B. Excipients

The use of excipients needs to follow a clear rationale, and their use may be limited by route, concentration, and total dose. In any case, the acceptance by health authorities of the excipient use in a new formulation will depend on the overall benefit/risk assessment. An FDA webpage lists concentrations of excipients used in approved drug products [26].

The use of excipients for injectable drugs is more limited than for the oral route. The use of novel excipients or their use at higher concentration or doses, respectively, may require additional safety studies. Regulatory approval for clinical studies or market access depends on the overall benefit/risk profile of the drug product including any excipient related risks.

Nema and Brendel reviewed excipient concentrations and doses in approved injectable drug products per administration route. Tables 29.5 and 29.6 show solvents, cosolvents, and solubilizing, wetting, suspending, emulsifying, or thickening agents, their frequency of use, and the maximum concentration in the drug product given as example [27]. The safety of excipients depends on the route and rate of administration, excipient concentration at the injection site, dilution by body fluids, and daily intake of the excipient. For example, intravenous infusion allows for faster dilution of higher excipient concentrations than subcutaneous injection where little interstitial fluid is available for dilution. Therefore the subcutaneous route will in general tolerate lower concentrations and doses of excipients than an intravenous infusion. Mottu reviewed the toxicity of organic solvents regarding hemolytic potential of cosolvents as follows: dimethyl acetamide < PEG 400 < ethanol < propylene glycol < dimethylsulfoxide [28,29].

Further excipients in parenteral dosage forms include pH-modifiers (acid, base, or buffer components), tonicity agents, viscosity or density modifiers, preservatives, or antioxidants. For pH-modification, the addition of a base or an acid is often preferred over the use of buffers. This approach minimizes buffer capacity that may be related to injection pain and minimizes ionic strength, which can also reduce drug solubility. The use of buffering components may be needed if the pH is unstable and needs tight control.

TABLE 29.6 Solubilizing, Wetting, Suspending, Emulsifying, or Thickening Agents

Excipient	Frequency	Max concentration	Example
Acacia	2	7%	Tuberculin Old Test(Lederle)
Aluminum monostearate	1	2%	Solganal Injection(Schering)
Carboxy methyl cellulose	4	0.55%	Bicillin (Wyeth-Ayerst)
Crosscarmellose sodium	21	3%	Nutropin Depot (Genentech)
Cremophor EL[a]	3	65% (w/v)	Sandimmune (Sandoz)
Cyclodextrin, gamma	1	5%	Cardiotec (BMS)
Cyclodextrin, alpha	1	0.14%	Edex (Schwartz)
Hydroxypropyl-beta-cyclodextrin	2	40%	Sporanox (Janssen)
Sulfobutylether cyclodextrin sodium	3	29.4%	Geodon (Pfizer)
Desoxycholate sodium	1	0.41% (w/v)	Fungizone (Bristol Myers)
Egg yolk phospholipid	3	1.2%	Cleviprex (The Medicines Co.)
Gelatin, hydrolyzed	1	16% (w/v)	Cortone (Merck)
Lecithin	8	1.2% (w/v)	Diprivan (Zeneca)
Polyoxyethylated fatty acid	1	7% (w/v)	AquaMephyton (Merck)
Polysorbate 80 (Tween 80)	72	100%	Taxotere (Aventis)
Polysorbate 20 (Tween 20)	22	0.4% (w/v)	Calcijex (Abbott)
PEG 40 castor oil[b]	1	11.5% (v/v)	Monistat (Janssen)
PEG 60 castor oil[c]	1	20% (w/v)	Prograf (Fujisawa)
Poloxamer 188 (Pluronic F68)	5	0.3%	Norditropin (NovoNordisk)
Povidone (Polyvinyl pyrrolidone, Crosspovidone)	7	0.6%	Bicillin (Wyeth-Ayerst)
Sodium dodecyl sulfate (sodium lauryl sulfate)	1	0.018% (w/v)	Proleukin (Cetus)
Sorbitol	3	50% (v/v)	Aristrospan (Fujisawa)
Triton X-100 (Octoxynol-9)	1	0.0085%	Fluarix (GSK)

[a]*Cremophor EL, Etocas 35, polyethoxylated castor oil, polyoxyethylene 35 castor oil.*
[b]*PEG 40 castor oil, polyoxyl 40 castor oil, castor oil POE-40, Croduret 40, polyoxyethylene 40 castoroil, Protachem CA-40.*
[c]*PEG 60 hydrogenated castor oil; Cremophor RH 60, hydrogenated castor oil POE-60, Protachem CAH-60.*

The use of microbiological preservatives is limited to multidose preparations or preparations for continuous subcutaneous infusion. The necessary type and level of preservatives needs to be determined through preservation effectiveness tests. The effective preservative concentration depends on pH, the vehicle, and packaging components, as the preservative may ionize and partition between hydrophilic and lipophilic phases as well as packaging components.

C. Solubilization

1. Cosolvent–Water Systems

When pH-modification, complexation, as discussed in other chapters, are insufficient to solubilize a poorly soluble drug, then organic solvents alone or in combination with surfactants are used. Solubility tests in cosolvent–water mixtures at varying weight fractions show which cosolvents are best suitable for solubilizing the drug for a ready to use injection. Mixtures of the best-solubilizing cosolvents may show synergies in solubilization.

In cases where the candidate is ionizable or zwitterionic, a combination of cosolvent(s) and pH adjustment of solution should be investigated. In any case, the selection of the qualitative and quantitative composition must ensure local tolerability at the injection site, not lead to drug precipitation on contact with plasma, and not provoke hemolytic effects. Pharmaceutically acceptable solvents are listed in Table 29.5. Of these, ethanol, propylene

glycol, glycerol, and PEG are used in approximately 66 percent of those products. Generally, the objective is to minimize the excipient dose and not to exceed established safe concentrations for the route of administration. Further the solution must not be oversaturated at the long-term storage temperature. High dosed poorly water-soluble drug solution concentrates for infusion allow dosing of high volumes over a longer time. The infusion concentrate solution is diluted into a carrier infusion. By this approach, the total dose of co-administered cosolvents can be minimized due to the log/linear solubility relationship. However, care must be taken that the final infusion solution is physically and chemically stable for the in-use stability period of several hours.

A. CHARACTERISTICS OF THE SOLVENT SYSTEM

The Hildebrand solubility parameter δ and its extension, the three-component (or three-dimensional) Hansen solubility parameters, have been successfully used for predicting solubility and identifying the best solvents for a given solute. These predictions take into account solution temperature, melting point, and heat of fusion of the solid solute and the solubility parameters of solute and solvent. A crystalline solute has a limited solubility, even in an ideal solvent (ideal solubility) and therefore in any solvent. The ideal solubility, expressed as mol fraction, can be calculated based on melting point and heat of fusion of a solid solute for a given solution temperature. The intermolecular interaction is quantified by the solubility parameters of solute and solvent. However, when solute and solvent interactions are dominated by hydrogen bonding, predictability of solubility is limited, so that simple solubility testing is more reliable and efficient.

Another useful physicochemical descriptor for solute and solvent properties is the partitioning coefficient P. It is more appropriate to regard log P as a measure of hydrophobicity, that is, the tendency to escape an aqueous environment. It is most commonly used in its logarithmic form, log P or Clog P (the calculated value).

For estimating the solubility of a given solute in cosolvent−water mixtures, a simple relationship between suitable parameters of the solute and the cosolvent would be helpful. A straightforward and reliable approach for selecting cosolvents and predicting their solubilization effects on drugs that requires little experimental data, and thus minimal time and drug, is the log-linear model proposed by Yalkowsky et al [30−32]. This model describes an exponential increase in the solubility of a nonionic drug with a linear increase in cosolvent concentration. The relationship is described by Equation 29.3:

$$\log S_{\text{tot}} = \log S_w + \sigma f_c \tag{29.3}$$

where S_{tot} is the total solute solubility in the cosolvent−water mixture, S_w is its water solubility, σ is the cosolvents solubilization power for the particular cosolvent−solute system, and f_c is the volume fraction of the cosolvent in the aqueous mixture. The σ term can be obtained from the slope of the $\log(S_{tot}/S_w)$ versus cosolvent volume fraction (f_c) profile of each selected drug and cosolvent. This implies the need for some experimental data for each cosolvent− solute system. However, it was demonstrated that a linear relationship also exists between σ and the logarithm of the solute's log P [33]. Therefore, σ can be related to the solute and the cosolvent by the following simple relationship Equation 29.4:

$$\sigma = \text{S} \log \text{P} + \text{T} \tag{29.4}$$

where S and T are constants dependent only on the particular cosolvent and log P is the solute's partition coefficient. In other words, the only *ab initio* data required to predict the solubility of a solute is the compound's octanol−water partition coefficient. Fortunately, this value can be determined experimentally or accurately predicted by a number of computational methods (Clog P, for example). A useful form of the log-linear equation is obtained by substituting from Equation 29.4 into Equation 29.3 to give:

$$\log \text{S}_{\text{tot}} = \log \text{S}_{\text{w}} + f_c(S \times \log \text{P} + \text{T}) \tag{29.5}$$

which expresses the total solubility in a mixed solvent system solely in terms of the pure components—water, cosolvent, and solute—obviating the need for any individual solute−cosolvent experiments. In Figure 29.2, the solubility is plotted against the composition (per cent of volume) of the water−cosolvent mixtures for three drug substances representative of different polarities: tioconazole 1 (log P = 4.52), caffeine 2 (log P = 0.57), and oxfenicine 3 (log P = −2.5) [34]. Several features can be derived from the graphs:

1. As the solute log P decreases, the solubilization by cosolvents decreases. The nonpolar tioconazole exhibits a logarithmic increase of solubility as the cosolvent volume fraction increases, whereas the low log P compound oxfenicine shows a log-linear decrease of solubility, following the model represented by Equation 29.5.

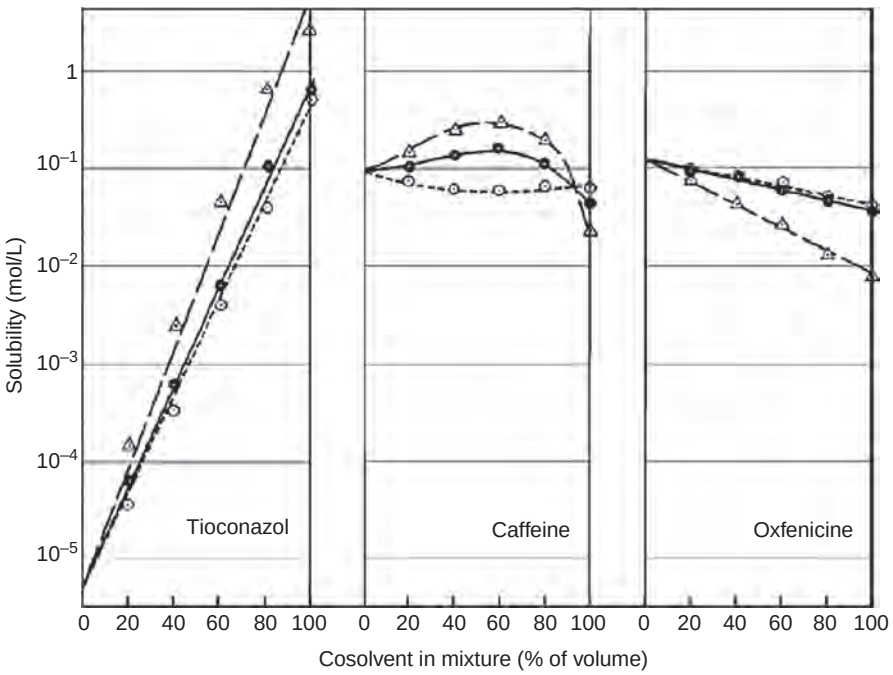


FIGURE 29.2 Solubilization of tioconazole, caffeine, oxfenicine.

2. For high log P solutes, the solubilization power of the cosolvents increases in the series glycerine < PEG 400 < propylene glycol < ethanol. As can be seen above, this correlates well with the log P of the cosolvents.
3. The intermediate case of caffeine does not follow this log-linear relationship. Instead, a parabolic course is generally observed with semipolar compounds. The deviation from linearity is related to the nonideal nature of water–cosolvent interactions. In consequence, the maximum solubility for semipolar substances must be sought experimentally. But at least the rank order of the degree of solubilization is reflected by the relative magnitude of σ values that may be calculated from the drug's log P value and the solvent coefficients in Table 29.4.

As exemplified by the case of oxfenicine, the solubilization approach by cosolvent mixtures does not appear to be useful for polar drug substances simply because they are less soluble in nonpolar solvents than they are in water. Electrolytes are at least partly ionized, and therefore they exhibit pH-dependent solubility in aqueous media, while the fraction of the less water-soluble, non-ionized species may benefit from cosolvent solubilization as shown in the example Figure 29.3.

Although the solubility of dexoxadrol hydrochloride in water exceeds that of the free base by more than two orders of magnitude, a significant increase can be achieved by additional cosolvent (PEG 300), although the effect is much less pronounced than with the free base. Finally, it should be noted that the log-linear model holds for mixed cosolvent systems in the following form (Equation 29.6):

$$\log S_{mix} = \log S_w + \sum(\sigma f_c) \tag{29.6}$$

where the individual solubilization powers and volume fractions of each cosolvent are linearly summed [35,36].

The formulation has to be devised such that the effect of dilution of the drug solution by the aqueous body fluids (e.g., blood, tissue interstitial fluids) during administration is considered in order to avoid precipitation upon injection. Solubilization needs to be maintained until the drug substance is diluted to its final concentration in the circulating bloodstream or in an infusion fluid, if the product is added to an i.v. infusion [37]. Otherwise, there is the risk that precipitated drug particles could mechanically or chemically irritate tissue, or even clog blood vessels.

A further part of the cosolvent selection process for a drug that is to be injected as a bolus involves the determination of the optimum concentration of solvent. This is a balance between the percentage that is required to maintain the drug in solution during manufacture and storage and the maximum concentration that can be safely injected. Similarly, for drugs that are infused i.v., the selection process should involve a dilution study using normal saline or 5 percent dextrose solution with the final concentration of the cosolvent at or close to its maximum accepted concentration to check for precipitation. Data on marketed parenteral products has been collated [38].

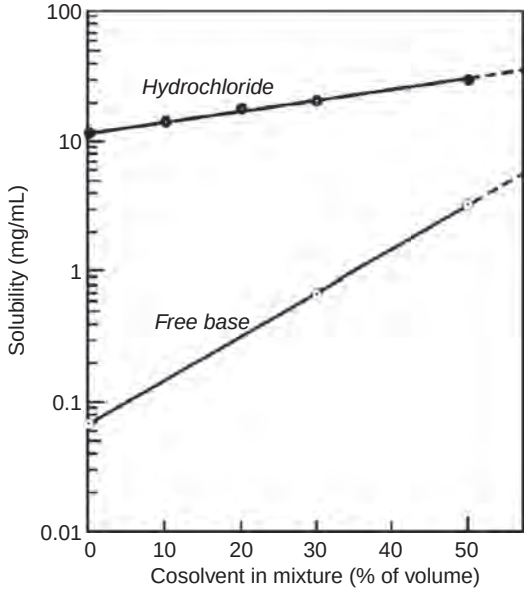


FIGURE 29.3 Solubility of the free base and the hydrochloride of dexoxadrol.

A decision-tree approach to parenteral formulation for use at the discovery stage of development has been published [39] covering compounds that are strong or weak acids, strong or weak bases, or neutral entities. In addition to this paper, several useful reviews have also been published that specifically cover the requirement for early formulations [40–42].

2. *Solubilization Mediated by Surfactants*

For drugs that have a very high dose–water-solubility ratio, solubilization by cosolvents alone may not be sufficient. A common strategy is to design concentrate solutions for infusion. These concentrates for infusion (e.g., Taxotere®, Busulfex®) maximize the drug solubility by high cosolvent concentrations or other solubilizing excipients. This approach allows for solubilizing a unit dose in a minimal volume, minimizing the total cosolvent dose given with the drug dose. The concentrate for infusion with its high excipient concentration is then diluted prior to use with a carrier infusion to excipient concentrations that are locally tolerated at the injection site. Additionally hemolysis, hypersensitivity, or other systemic adverse effects may be caused by excipients in solution or by particulate drug-delivery systems [43].

The preferred surfactants are nonionic. They are amphiphilic containing polar and nonpolar domains in their structures so that the molecule has a hydrophilic and a hydrophobic section. In aqueous solution, the polar groups are hydrated and oriented toward the water, while the hydrophobic part of the molecule that cannot be hydrated accumulates next to other hydrophobic molecules. At the air–water interface, their polar region is oriented toward the aqueous surface, while the nonpolar region is pointing outwards into the low-polar gas phase (Figure. 29.4a).

When hydrophobic drug substance particles are stirred into and suspended in the solution, the nonpolar region of the surfactant molecule attaches to the surface of these particles, while the polar group maintains intimate contact with the aqueous bulk phase. At low concentrations, surfactants are therefore used as wetting agents. As the concentration in aqueous solution is increased, it exceeds a critical value, the critical micelle concentration (CMC), and the surfactant molecules assemble cooperatively to form soluble, structured aggregates (Figure. 29.4b). The CMC value is a characteristic of a particular surfactant and depends on environmental factors such as temperature, pH, ionic strength, and others. Surfactant micelles are in a dynamic equilibrium with the surfactant monomers. Therefore, the number of molecules constituting a micelle is distributed around an average value over time and over a population of micelles. For example, the average aggregation number n in micelles of sodium dodecylsulfate is 64 at 25°C; for other surfactants this number varies over several orders of magnitude from 2–8 (bile salts) to many thousands. With a value of n as that of sodium dodecylsulfate, the shape of micelles is likely to be spherical; as n increases the shapes tend to be disks, rods, sheets, or even vesicles. Surfactant micelles can accommodate solute molecules in a number of orientations, arranged according to their

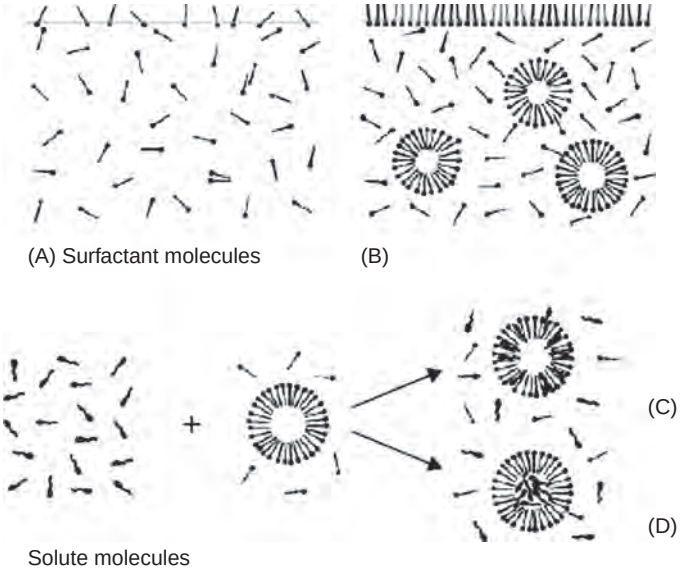


FIGURE 29.4 Formation of surfactant micelles and micellar solubilization.

polarity and hydrophobicity: (1) at the interface between the micelle surface and the aqueous bulk phase; (2) between the polar hydrophilic head groups; (3) within the hydrophobic part (the palisades layer) of the micelle shell (Figure. 29.4c); or (4) in the core (Figure. 29.4d). This is the mechanism of micellar solubilization. The totality of micelles represents a colloidal phase into which a substance is "dissolved" in the aqueous phase partitions. The capacity of the micellar phase to solubilize a solute can therefore be expressed as a partition coefficient, K_m. Hence, a linear relationship can be expected between the concentration S_{mic} of substance solubilized by micelles and the concentration of the surfactant C_{sfc} in the system. Because only micelles and not the monomeric surfactant molecules contribute to the solubilizing effect, the CMC C_{cmc} must be subtracted from the total of the surfactant concentration. The resulting total concentration of solute in the micellar solution is then:

$$S_{tot} = S_w + S_{mic} \tag{29.7}$$

Where

$$S_{mic} = K_m(C_{sfc} - C_{cmc}) \tag{29.9}$$

And

$$K_m = S_{mic}/S_w \tag{29.10}$$

The hydrophile/lipophile balance (HLB) system for the characterization of surfactants, devised by Griffin [88], was originally developed for selecting nonionic surfactants as emulsifiers for stabilizing oil-in-water (o/w) emulsions. According to the ratio of polar and nonpolar groups in the molecule, a surfactant is assigned a number, the HLB value, on a scale originally ranging from one to twenty with the most lipophilic surfactants having the lowest number,, and the more hydrophilic surfactants having higher numbers. Later inclusion of ionic surfactants in the system required an extension of the scale beyond twenty. Surfactants must be chosen with due care, as their strong interfacial activity can elicit adverse reactions in biological systems. Cell membrane constituents (e.g., phospholipids, cholesterol) can be solubilized at supermicellar concentrations, impairing the integrity of membranes. This disruption of the membrane enhances the permeability of drug substances and other substances present in the extracellular fluid. At low concentrations, such alterations are reversible and membranes recover rapidly, but higher concentrations can cause irreversible damage to mucosa and other tissues.

Some of the polyethoxylated surfactants used as solubilizers in parenteral and oral applications, in particular Cremophor EL®, are used in conjunction with ethanol to solubilize paclitaxel in Taxol® as a solution for dilution for infusion. Infusion of Cremophor EL®-containing drugs led to adverse events up to anaphylactic reactions, which led to the practice of premedication with corticosteroids and antihistamines prior to chemotherapy. Lecithin and other natural phospholipids, as one of the main constituents of all biological membranes, are the surfactants and solubilizers best tolerated in parenteral and oral applications. In order to keep the use concentrations low, the solubilizing power of one surfactant is rarely utilized alone. Thus, combinations of two or more

surfactants are used to optimize the solubilization effect by the formation of mixed micelles, especially with lecithin and bile acids. For injectable solutions several solubilization principles are usually combined, as can be seen in the following examples:

1. By employing pH adjustment and 5 percent sodium dodecyl sulfate, the solubility of the investigational cancer drug pyroxamide was increased by a factor of nearly half a million over its intrinsic solubility [44].
2. The presence of hydrophilic and hydrophobic regions in the same molecule is not a monopoly of surfactants; many drug molecules exhibit hydrophilic and hydrophobic regions as well, and this is indeed the structural requirement for their ability to permeate cell membranes and pass through absorption barriers. The amphiphilic properties are obvious in some drug classes (e.g., local anesthetics) [45]. On some individual occasions, a drug substance may exceed the CMC, whereupon micellar "self-solubilization" can occur. A recent example has been described by Hussain et al, who found that the hydrochloride of the basic analgesic DuP747 is surface active; the saturated solution (3 mg/mL at 22°C) lowered the surface tension of water to 50 dyne/cm. It was hoped that a different salt might be soluble enough to reach higher concentrations exceeding a CMC. This was indeed the case. The CMC of the methanesulfonate was found to be ca. 4 mg/mL, and a micellar solubility of 60 mg/mL could be achieved [46].

3. *Disperse Systems*

When simple ionization or cosolvent/surfactant-based solutions are not feasible due to a high dose−solubility ratio, the more complex disperse systems can be explored. Intravenous emulsions and liposomes have been successfully marketed. The drug release may occur quickly by partitioning from the oil droplet or liposome bilayer.

If drugs are injected as nanosized crystals, release occurs by dissolution. The nanocrystals may be eliminated from circulation by cellular uptake of the reticuolendothelial system (RES).

Figure 29.5 shows size ranges for commonly used colloidal drug-delivery systems. After injection, such colloidal systems interact with blood components, may fall apart, release components, may be eliminated if they are smaller than approximately 10 nm, or be opsonized and eliminated from circulation by cellular uptake of the reticuolendothelial system (RES).

4. *Emulsions*

Lipid-soluble drug substances, like fat-soluble vitamins and steroids, can be readily dissolved in lipid vehicles such as natural triglycerides (i.e., purified fatty oils) and semisynthetic triglycerides, such as the medium-chain triglycerides (e.g., Miglyol 812, consisting of the C8−C12 fatty acid triglycerides). Administration of oils by the parenteral route is restricted to subcutaneous and intramuscular injection. When given intravenously, oils would

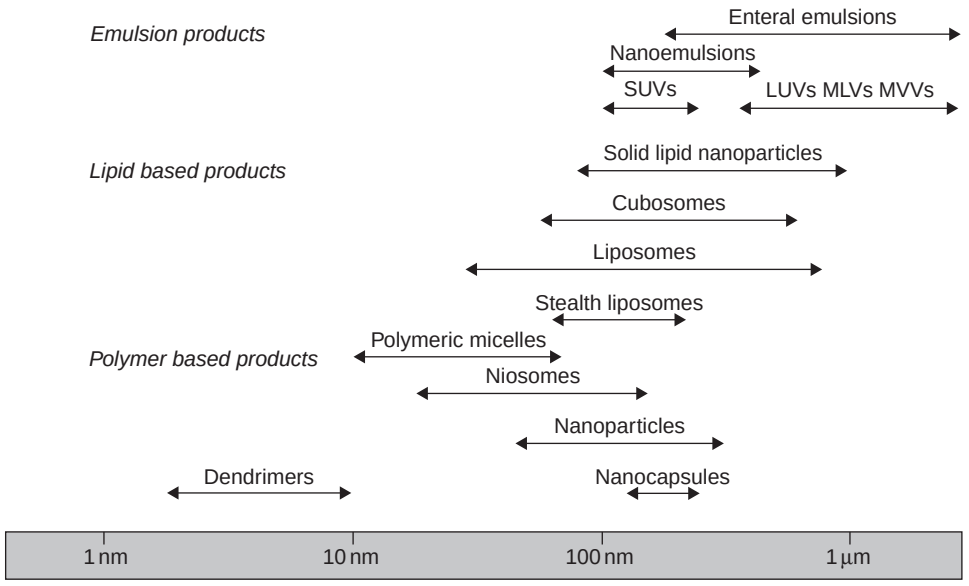


FIGURE 29.5 Approximate size ranges for the commonly used colloidal drug-delivery systems.

form droplets leading to fatal fat embolism. Therefore, homogeneous lipid phases must never be injected by the intravascular route. However, this problem can be circumvented if the oil is finely dispersed to a stable droplet size that can pass through narrow capillaries.

An emulsion is defined as a dispersion of two immiscible liquids, one of which is finely subdivided and uniformly distributed as droplets (the dispersed phase) throughout the other (the continuous phase). A third component (or multiple additional components), the emulsifying agent(s), is necessary to help stabilize the emulsion. The emulsifying agent(s) coats the droplets and prevents droplet coalescence by either reducing the interfacial tension or by creating a physical repulsion between the droplets [47]. The dispersed phase is occasionally also defined as the internal phase; the continuous phase is occasionally also defined as the external phase or dispersion medium. Virtually all emulsions are inherently physically unstable.

There are three main types of emulsion:

1. Oil-in-water emulsions (o/w), where the continuous phase is aqueous. These are normally used for oral or parenteral products.
2. Water-in-oil emulsions (w/o), where the continuous phase is an oil. These are normally used for topical products.
3. Multiple emulsions that can be either w/o/w or o/w/o and comprise droplets of one phase inside droplets of the second phase, with the continuous phase the same as phase one. These are sometimes used for delayed release products.

It is recommended that emulsions developed for the i.v. route of administration should have a submicron droplet size, although emulsions with a 10 mm droplet size have been used parenterally [48].

For parenteral nutrition, drug-free emulsions are used. A typical composition of a nutritional fat emulsion for i.v. administration is 200 g soybean oil or safflower oil, 12 g egg phospholipids, and 25 g glycerine per liter of emulsion with the pH adjusted to within the range 5.5−8.0. It is most important that the fat droplet particle size that does not exceed 0.5 μm. For experimental purposes, such fat emulsions (e.g., Abbolipid®, Intralipid®, Lipofundin®) are well suited and are used as early dosing vehicles for lipophilic drug candidates. Depending on the drug properties, emulsion formulations require significant optimization effort in tailoring the oily phase and emulsifiers. As an example, for the i.v. administration of the poorly soluble antimitotic agent paclitaxel in Taxol®, large volumes of a micellar and cosolvent-mediated dispersion consisting of Cremophor EL + ethanol + isotonic saline 5 + 5 + 90 have to be infused to reach the therapeutic dose of 30 mg, because the solubility of the drug is only 0.6 mg/mL in that vehicle. Intralipid was ruled out as an infusion vehicle (solubility of paclitaxel in soybean oil is 0.3 mg/mL). Emulsions may also be used for drug-targeting liver uptake. Formulations with a droplet size between 100 and 200 μm usually result in a high liver uptake after i.v. injection [49,50].

Often, drug toxicity may be reduced by formulation as an emulsion. The use of a w/o emulsion of amphotericin deoxycholate, as opposed to a simple solution, was shown to reduce the incidence and severity of renal impairment and chills in patients, while still maintaining the antifungal efficacy of the drug [51].

In contrast to ordinary emulsions, microemulsions are homogeneous, transparent, and thermodynamically stable. They can be considered swollen micelles. Microemulsions usually form spontaneously when the components are mixed at their optimized ratios. For optimum ease of formation and subsequent stability, it is essential that an interfacial tension as low as possible be achieved and maintained between the phases. Generally, this is not accomplished with just a single surfactant; a second amphiphile, typically a medium-chain-length alcohol, needs to be added as the cosurfactant. Microemulsions have several advantages over coarse emulsions, especially in respect of transparency and physical stability. However, they require significantly higher percentages of surfactants for formulation, thus restricting the choice of acceptable components [52].

The pharmaceutical significance of microemulsions lies in the fact that they are not merely solvents that solubilize drug substances that are difficult to dissolve in other solvents, but serve as systems that impart high penetration power to a drug product. For this reason, they have gained increasing interest. Limiting factors for their use on a broad scale include the need for well-tolerated components, particularly surfactants and cosurfactants, whose concentrations are higher than in ordinary emulsions. Furthermore, thermodynamic stability must be maintained over a temperature range between 4°C and 40°C. Drug carrying microemulsions so far have been used in the topical and transdermal field, but with the natural phospholipids the principle can be applied even to parenteral administered products. The breakthrough in oral drug delivery was achieved with the successful attempts of Kissel and Posanski to increase the bioavailability of cyclosporine-A by the microemulsion principle [53]. A solution formulation for an oral microemulsion with a drug content of 10 percent cyclosporine-A is composed of corn oil mono-, di-, and triglycerides, PEG-40 hydrogenated castor oil, propylene glycol, and anhydrous

ethanol. This product, replacing the former self-emulsifying drug-delivery system (SEDDS; see below), is actually a concentrate, which upon dilution with an aqueous phase generates the effective microemulsion. This principle of anhydrous concentrates was named SMEDDS (self-micro-emulsifying drug-delivery system). Such formulations lack the aqueous phase. On dilution, a SMEDDS spontaneously converts to an optically clear, thermodynamically stable microemulsion that contains the drug in molecular dispersion. The same principle of a water-free concentrate when leading to a macro-emulsion is called SEDDS. A review of self-dispersing lipid-based systems has been published recently [19].

As stated above, emulsions are essentially swollen micellar systems. The differences between a micelle containing solubilized oil and an emulsion comprising an oil droplet surrounded by an interfacial layer composed largely of surfactant are difficult to assess. The droplet sizes of the dispersed phase of an emulsion can be estimated by its appearance, as seen in Table 29.7. Table 29.8 shows marketed parenteral emulsions.

5. *The Choice and Classification of Emulsifying Agents*

Emulsification agents for parenteral emulsions are often chosen on the basis of their low toxicity and irritancy, in addition to their emulsification ability. While there is no list of emulsifying agents specifically approved for use in pharmaceutical products, a list of emulsifiers approved for use as food additives within the European Union is available. It is widely assumed that emulsifiers from this list are suitable for pharmaceutical products intended for internal use. The list comprises naturally occurring materials, together with some of their semisynthetic derivatives, such as polysaccharides, glycerol esters, cellulose ethers, sorbitan esters, and polysorbates. These are largely nonionic and are less irritant and less toxic than their anionic derivatives; cationic derivatives are usually significantly more toxic than the corresponding anionic derivatives.

Emulsions are usually used as a means of administering drugs that are poorly water-soluble the drug within the oil phase [54–56].

Intravenous infusions of emulsions with small droplet size can be safely dosed at high volumes for the purpose of parenteral nutrition [57]. Hypersensitivity reactions may occur in some patients, so the infusion rate needs to be limited. Initially, infusion rates are minimal in order to minimize hypersensitivity events.

Similarly, i.v. emulsions are useful for solubilizing lipophilic drugs and have the advantage of excellent injection site tolerance. High drug loads can be achieved, particularly for lipophilic liquids or low melting drug substances (e.g., Propofol®). After dosing, such drugs partition quickly into the blood and distribute to the site of action.

It may be necessary to add a buffer system to improve chemical stability, control tonicity, or to ensure physiological compatibility.

TABLE 29.7 Relationship Between Droplet Size and Appearance

Droplet size μm	Appearance
<0.05	Transparent
0.05–0.1	Grey, translucent
0.1–1	Blue-white (Tyndall effect)
>1	Milky

TABLE 29.8 Parenteral Emulsions

Brand	Active	Molar weight	Melting point	Log P	Single dose (mg)	Concentration (mg/mL)
Ropion	Flurbiprofen axetil	330	liquid	4.9	50	10
Diazepam Lipuro	Diazepam	285	126	2.9	10	5
Diprivan	Propofol	178	liquid	3.8	175	20
Etomidat Lipuro	Etomidat	244	67	2.9	28	2

6. Stability of Emulsions

A. PHYSICAL STABILITY

A stable emulsion is one where the globules retain their initial character and remain uniformly distributed throughout the continuous phase. The addition of emulsifying agents results in the formation of an interfacial film around each of the dispersed droplets. The physical nature of this film creates a barrier that eventually controls the coalescence of the droplets that approach one another. If this film is electrically charged through the use of charged surfactants, repulsion occurs before contact is made and improvements in stability are normally seen.

I. CRACKING OR BREAKING This is the separation of an emulsion back to its two constituting coherent phases, usually caused by agents that destroy the interfacial film. Some of the factors causing cracking are:

1. The addition of agents with charges opposite to those of the film; these can include surfactants, electrolytes, or iodized drugs with the opposite charge.
2. Bacterial growth.
3. Wide temperature changes, including freeze−thaw cycles.

II. FLOCCULATION This is an intermediate phenomenon, not to be confused with creaming, that is likely to occur in most pharmaceutical emulsions if there are weaknesses in the interfacial film. The droplets aggregate due to a shift in the interaction between attractive and repulsive forces, without fully coalescing. Shaking can normally re-disperse them.

III. CREAMING Creaming is primarily due to density differences between the two phases, and many emulsions cream on prolonged standing. The disperse phase (normally oily) is less dense and rises to the top to form a layer of more concentrated emulsion. Flocculation may occur in addition to creaming. Droplets of the creamed layer do not normally coalesce and can usually be re-dispersed by gentle shaking. Although it is not such a serious problem as cracking, it often leads to the formation of a product that is less elegant, with the added danger of incorrect dosing.

IV. PHASE INVERSION Normally, the relative volumes of each phase contribute to the type of emulsion formed. It is normal to have less than 50 percent of the disperse phase, but in numerous examples this volume exceeds 50 percent. Occasionally, attempts to significantly increase the amount of disperse phase result in cracking or phase inversion (an o/w emulsion switches over to a w/o emulsion).

B. CHEMICAL STABILITY

As stated earlier, ionic emulsifying agents are incompatible with materials or agents with the opposite charge. The presence of electrolytes can markedly affect the stability by either reducing the energies involved in the interactions between adjacent globules or by promoting salting out. High concentrations of electrolytes can strip the hydration layers from emulsifying agents, causing their precipitation. Changes in pH may also exert a significant effect by converting an ionized surfactant (e,g,, a sodium salt of an anionic surfactant) to its free acid or base form.

By their nature, emulsions usually contain high concentrations of oils or fats of natural (animal or vegetal) or semisynthetic origin. These may contain peroxide or be susceptible to the production of peroxides by the action of atmospheric oxygen or microorganisms unless due care is taken. Products can be formulated with antioxidants and/or microbiological preservatives, if considered necessary (see below). Obviously, if the active ingredient has been shown during the preformulation evaluation to be susceptible to peroxide oxidation, consideration should be given to the use of excipients with minimal propensity of peroxide formation or the formulation should be adequately protected by antioxidants.

C. MICROBIOLOGICAL CONTAMINATION AND STABILITY

The mixture of excipients that constitute an emulsion forms an excellent substrate for the growth of bacteria and fungi. Degradation due to microbial contamination can markedly affect both the physical and chemical stability, resulting in color and odor changes (rancidification), the production of gases (mainly carbon dioxide), hydrolysis of any ester groups in the fats and oils, pH changes in the aqueous phase (due to the production of organic acids), and a general break- down of the emulsion.

7. Drug Release from Emulsion Products

Emulsions are commonly used to deliver drugs by the oral, rectal and topical routes. Drug-delivery rates are usually high due the high surface area of the droplets of the dispersed phase [58–63]. They are also a convenient formulation for use in i.v. feeding, even though the choice of emulsifying agent is rather limited and the average globule size must be no greater than 4 μm. However, to be on the safe side it is generally recognized that for drug delivery by the i.v. route, the average droplet size should not be greater than 1 μm as particles greater than 5 μm cause capillary embolism.

It has been demonstrated that the gastrointestinal absorption of highly lipophilic substances can be enhanced impressively when given in lipid vehicles by entering the lymphatic route of fat absorption "piggy-back," where bile acids are the potent natural solubilizers. Charman and Stella have suggested that drugs require a log P octanol in excess of 5 and a triglyceride solubility of at least 50 mg/mL before mesenteric lymphatic transport is likely to become a major contributor to bioavailability [64,65]. Simple solutions of the antimalarial halofantrine hydrochloride (solubility in water: approximately 1 μg/mL, calculated log P = 8.5) in triglycerides increased the lymphatic uptake of the drug compared to an aqueous suspension. The long-chain glycerides (peanut oil) were most effective and surpassed by far the medium-chain triglycerides (fractionated coconut oil) and the still less effective tributyrin [66].

Emulsions can also be used as sustained release products by reversing the nature of the emulsion from o/w to w/o. The i.m. injection of w/o emulsions of some water-soluble vaccines results in a slow release of the antigen, a greater antibody response, and hence an immunity with longer duration. Emulsions may also be used to reduce the toxicity of the drug substance, and the use of a w/o emulsion rather than an aqueous solution can reduce the incidence and severity toxic effects.

Sodium metabisulfite and EDTA have been used individually as microbiological preservatives in different approved formulations of propofol injectable emulsion. The choice of antioxidant must be made with care, as some are restricted by use and/or concentration in different countries. BHA is widely used in fixed oils and fats at concentrations up to 0.02 percent but is permitted up to 0.1 percent in some essential oils. Its close relative, BHT, is recommended as an alternative to tocopherol at concentrations up to 10 ppm for the stabilization of liquid paraffin. Other anti- oxidants widely used in emulsion formulations include the propyl, octyl, and dodecyl esters of gallic acid at concentrations up to 0.001 percent (fixed oils and fats) and up to 0.1 percent (essential oils) [52].

8. Liposomes

Liposomes are formed by the self-assembly of phospholipid molecules in an aqueous environment. The amphiphilic phospholipid molecules form a closed spherical bilayer in an attempt to shield their hydrophobic groups from the aqueous environment, while still maintaining contact with the aqueous phase via the hydrophilic head group. The resulting closed sphere may encapsulate water-soluble drugs within the inner aqueous compartment or may encapsulate lipid-soluble drugs within the bilayer membrane. Often they are composed of anionic or zwitterionic phospholipids and a cholesterol-like molecule. The liposomes may vary in size and number of bilayers, depending on factors such as the manufacturing process, lipid composition, and surface charge. Like other nanoparticles, they generally have a relatively short lifetime in blood circulation. Liposome macrophage uptake by the liver and spleen (the reticuloendothelial system; RES) removes such liposomes from blood circulation. The lipid bilayer and/or the aqueous interior of liposomes can solubilize lipophilic and also amphiphilic drugs.

Visudyne® is a marketed liposome product that contains poorly water-soluble verteporphin solubilized in a liposome bilayer composed of dimyristoyl phosphatidylcholine (C14) and an unsaturated egg phosphatidylglycerol mixture (C14–C18). It has been shown that after intravenous injection, verteporphin is quickly transferred from its lipid-based formulation to serum proteins. This rapid transfer, particularly to lipoproteins, also provides a mechanism for its rapid delivery to cells [67].

D. Drug Targeting

Encapsulation within a nanoparticle can alter a drug's pharmacokinetics and distribution in ways that improve efficacy and reduce adverse side effects. These include liposomes, polymer conjugates, block copolymer micelles, nanoparticles, and nanosized crystals.

In 2011, seventy clinical cancer trials involved nanomedicines. Many of them were extensions to new clinical indications or new combinations [68,69] and [70] described key considerations relevant for successful tumor targeting:

1. Xenograft animal models that are often used as preclinical tumor models may have more leaky vasculature so that the EPR effect appears greater than in human studies,
2. The difficulty in evaluating biological potency, especially for complex formulations,
3. The importance of analytical methods that can determine the delivery system stability *in vivo*.
4. The appropriateness of current dose scaling techniques for estimation of starting doses in humans.

An excellent comprehensive review of the properties and use of colloidal systems, together with other similar technologies (collectively often referred to as nanomedicines) for use in medical imaging, has recently been published [68].

1. *Passive Targeting*

Passive drug targeting typically relies on an injected nanosized material that stays in the blood circulation long enough and is passively accumulated in the target tissue (e.g., tumors or inflamed tissue). The nanoparticle needs to be designed so that it releases the drug primarily after being accumulated in the target tissue. Further, it needs to be large enough that it is not filtrated by the kidney, small enough that it can distribute to target tissues, and have stealth properties that prevent it from eliminated by cells of the liver or spleen [71]. Another approach for passive targeting is based on local injections into tissue so that liposomes or other nanocarriers distribute in interstitial fluid toward local lymph nodes [72–74].

A. LIPOSOMES

The development history of Doxil details what was learned from a pioneering liposome product that may also apply to other nanosystems for passive tumor targeting. Major objectives for development of liposomes for passive targeting are the prolongation of liposome plasma circulation time, RES avoidance during vascular circulation, intra-tumor accumulation, and sufficient levels and retention of the drug substance inside the liposomes. The use of phospholipids with high melting point (Tm), cholesterol, and integration of DSPE-PEG into the liposome bilayer enables sterically-stabilized liposomes, reduces the insertion of plasma opsonins, and consequently prolongs the plasma circulation time of liposomes by reducing the uptake to the RES. A PEG chain length of 2000 Da was used for Doxil® liposomes [75]. Such liposomes escaping recognition and uptake by the liver and spleen are often referred to as "stealth liposomes," as this effect enables them to benefit of long circulation times. Liposome size has been shown to affect biodistribution. A size between 70 nm and 200 nm is necessary to achieve prolonged circulation times with stealth liposomes. When administered by i.v., stealth liposomes of about 100 nm in size passively target solid tumors by extravasation due to the disorganized tumor vasculature into their extracellular space. This effect is called the "enhanced permeation and retention" or EPR effect.

A sufficient active drug loading inside liposomes can be achieved by use of pH or ammonium ion gradients of amphiphilic weak bases or acids into long-circulating liposomes. The half-life of drug release from liposomes should be longer than the half-life of liposomes in circulation. In the case of Doxil®/Caelyx®, the active gradient loading technique leads to the formation of doxorubicin sulfate salt crystals inside the liposomes, so that 95 percent of the dose is crystalline. After injection, most of the plasma-circulating doxorubicin is measured as associated with liposomes as it reaches the tumor tissue. This liposome product is a concentrate for infusion and is approved for the treatment of Kaposi's sarcoma, recurrent ovarian cancer, metastatic breast cancer, and multiple myeloma.

Remote loading applies only to molecules that can accumulate in the internal aqueous phase of the liposome due to an ion or pH transmembrane gradient. Suitable candidates are amphipathic weak acids or weak bases having their log D (at pH 7) in the range of -2.5 to 2.0. Amphipathic weak bases should have a pKa ≤ 11, and weak acids should have pKa >3. Drug molecules that are too hydrophobic associate mainly with the lipid bilayer and will not be good candidates for remote loading. A computer-aided strategy for identifying drugs suitable for remote loading has been described [76]. The liposome dose has also been shown to affect the pharmacokinetics [77,78].

B. POLYMER CONJUGATES

Parenteral polymer therapeutics became successful as with PEG-protein conjugates [79]. Recently, small interfering ribonucleic acids (siRNAs) and dendrimers have also been developed. PEG was used as polymer attachment to antibodies or proteins due to its safety profile, its hydrophilicity, and its ability for

monofunctionalization. The pegylation prolongs the plasma half-life, and such drugs can often be administered subcutaneously. Autoinjectors or prefilled syringes are current devices for self-administration of such polymer conjugates for chronic dosing.

Other ligands, like styrene maleic acid-neocarzinostatin (SMANCS) or polyglutamic acid, have also been used [80].

Plasma proteins have also been suggested as drug carriers [81]. Albumin is the most abundant plasma protein (35–50 g/L human serum) with a molecular weight of 66.5 kDa and a plasma half-life of 19 days. It binds, solubilizes, and transports long chain, bilirubin, many therapeutic drugs—such as paclitaxel, copper(II), nickel(II), calcium(II), and zinc(II)—and is the major protein responsible for the colloid osmotic pressure of the blood. When human serum albumin (HAS) is broken down, the amino acids provide nutrition to peripheral tissue and also growing tumors. The tumor uptake, tumor blood flow, and the transport of molecules in the interstitium led Maeda and Matsumura to coin the expression "EPR" (i.e., enhanced permeability and retention of macromolecules) in relation to passive tumor targeting.

C. NANOSUSPENSIONS

Abraxane® is the first example for a marketed intravenous nanosuspension. The nanoparticles are generated in an emulsion-based process. The dosage form is a lyophilizate for reconstitution to a suspension for infusion [82].

D. STABILIZED MICELLES

Aqueous solutions for injection containing surfactants have been successfully used for solubilizing poorly water-soluble drugs. However, after intravenous injection and dilution in blood, these micelles may fall apart or release the drug quickly.

Stabilized micelles are micelles that are stable on dilution in plasma and are typically composed of cross-linked polymeric surfactants. Amphiphilic block copolymers, such as the Pluronics (polyoxyethylene–polyoxypropylene block copolymers), self-assemble into polymeric micelles, and hydrophobic drugs may be solubilized within the core of the micelle. Alternatively, drugs are conjugated to the micelle-forming polymer. Although micelles are rather dynamic systems that continuously exchange units between the micelle structure and the free units in solution, those composed of polyoxyethylene–poly(aspartic acid) have been found sufficiently stable.

2. *Active Targeting*

Ligand-targeted particulate nanomedicines have proven to be safe and efficacious in preclinical models and have mostly been investigated for the treatment of cancer. The targeting ligand typically does not cause localization in the target tissue but rather provides benefits to cell internalization and tissue retention. Future prospects may be to facilitate cell internalization of poorly permeable drugs, the treatment of drug-resistant tumors, targeting the tumor blood supply, crossing the blood–brain barrier, and generation of targeted vaccines. Lipid-based and polymer-based nanocarriers are the most advanced [83].

The development of stealth liposomes, which are not rapidly cleared by the liver and spleen, made the active and specific targeting of drugs a possibility. In an attempt to increase the tissue specificity, various antibodies have been conjugated to the surface of stealth liposomes to produce immunoliposomes for active targeting, as opposed to the passively targeted species discussed above. These antibodies can selectively bind to specific sites where there is a preponderance of antibody receptors and hence avoid healthy tissue. However, the effectiveness of the antibody in helping to target the liposome appears to be directly related to the concentration and molecular weight of the poly(oxyethylene) on the surface. A high poly(oxyethylene) density and poly(oxyethylene) with a molecular weight in excess of 2,000 Da reduces the directing ability of the antibody by shielding the antibody from its target.

A. ANTIBODY-DRUG CONJUGATES

Antibody–drug conjugates (ADC) have been the first approved ligand-targeted nanomedicines. All of them are treating oncology indications. They consist of an antibody as targeting moiety and 2 to 4 highly potent and covalently bonded toxin molecules per antibody. The toxin is released after internalization into specific antibody-binding cells and sequentially kills the cells. Adcetris® or brentuximab vedotin binds to and is internalized by lymphoma cells expressing CD33 marker and releases the toxin MMAE, after being subjected to intracellular lysosomal enzymes such as cathepsin B. However, the ADC is stable in plasma, which is

important for safety. The antibody brentuximab did not appear to have efficacy against Hodgkin's lymphoma, while the ADC brentuximab vedotin (Adcetris®) showed a 75 percent response rate in this indication. Trastuzumab emtansine (Kadcyla®) increased the response rate in metastatic breast cancer patients, extending the response rate of the parent monoclonal antibody from 15–26 percent to 26–64 percent [84].

B. LIGAND-TARGETED POLYMER NANOPARTICLES

Prostrate membrane specific antigen (PSMA) targeted docetaxel nanoparticles (BIND-014) have entered into Phase 2 clinical studies. The composition of the docetaxel containing nanoparticles was optimized in a multiparametric screening for settings of size, ligand density, PEG molecular weight, PEG content, zeta-potential, drug load, initial release rate, PL(G)A molecular weight, lactide/glycolide ratio, and type of encapsulation process.

3. *Intracellular siRNA and Gene Delivery*

RNA interference (RNAi) represents a breakthrough in understanding how genes are turned on and off in cells and the scientists Andrew Z. Fire and Craig C. Mello were awarded the 2006 Nobel Prize. Small interfering RNA (siRNA), the molecules that mediate RNAi, have now reached Phase 3 of clinical research. Nanoparticle-mediated siRNA delivery is a complex process that requires transport across numerous extracellular and intracellular barriers. The challenge is to deliver these highly anionic molecules parenterally to intracellular targets despite their vulnerability to metabolic elimination and lack of permeability into cells. Multiple parameters of the lipid nanoparticle (LNP) or degradable lipid particles (DLP) contribute to the success of drug delivery (e.g., siRNA entrapment, pka, LNP stability, cell uptake, and endosomal release). The use of cationic lipids enables the formation of lipid nanoparticles that are taken up by cells. Such liposomes are prepared from phospholipids with a hydrophilic amine head group (e.g., phosphatidylcholines, phosphatidylethanolamines or phosphatidylserines). The amines may be either quaternary ammonium, tertiary amines, secondary amines or primary amines, and liposomes prepared in this way are commonly called "cationic liposomes," because they possess a positive surface charge at physiological pH. These bind effectively to the negatively charged DNA. Cationic liposomes of varying description have been used to promote the cellular uptake of DNA with resultant therapeutic protein expression by various cell types *in vivo*. However, following i.v. injection, cationic liposomes are rapidly cleared from the blood. The knowledge gained in the development of anticancer liposomes, such as the prolongation of liposome circulation using poly-(oxyethylene) lipids, may be applied in the development of liposomes for gene delivery in order to produce efficient carriers. This strategy should aid liposome targeting in gene delivery. Thus far, targeting efforts with cationic liposomal gene formulations have been few and far between. Clinical trials with liposomal gene delivery have been initiated in cystic fibrosis patients, but they have concentrated on administration via nonparenteral routes for the obvious reason that the disease manifests itself in the alveolar epithelium. Although the experimental data indicate that cationic liposomes are able to facilitate the transfer of DNA into live mammalian cells, there are still major problems that need to be overcome in order to approach the ideal. In the future, it may be possible to treat this condition by the i.v. administration of the replacement gene–liposome complex [85].

E. Injectable Depots

Injectable depot drugs are particularly desirable in clinical indications where medication adherence needs to be improved (e.g., schizophrenia). Table 29.9 shows marketed injectable depot drugs. Such depot formulations are only effective and safe if daily doses are sufficiently low so that weekly or monthly doses can be safely delivered with a single subcutaneous (s.c.) or intramuscular (i.m.) injection. Furthermore, such depot formulations need to release the drug with a controlled rate to avoid a "burst release" (initially leading to high plasma levels) and to achieve limited fluctuations of plasma profiles that are safe and effective over the dosing interval [86].

1. *Nanoparticles*

Paliperidone palmitate (Invega sustenna®) was introduced in 2009 as the first nanocrystal injection product. Because of extreme low water solubility, paliperidone esters such as paliperidone palmitate dissolve slowly after an i.m. injection before being hydrolyzed to paliperidone and made available in the systemic circulation. It is administered intramuscularly as a one-month injection depot for the treatment of schizophrenia. The nanoparticle suspension is stable as a ready-to-use suspension and is administered from a ready-to-use syringe [5].

TABLE 29.9 Injectable Depot Systems

Brand	Active	Indication	Technology	Route: Dosing volume	Max Single dose (mg)	Dosing interval
Zyprexa Adhera	Olanzapine pamoate monohydrate	Schizophrenia	Powder in vial and aqueous vehicle in vial	Deep I.M.: 3 mL	405	4 weeks
Risperdal Consta	Risperidone	Schizophrenia	Powder in vial and aqueous vehicle in prefilled syringe	I.M. 2 mL	50	2 weeks
Invega Sustenna	Paliperidone palmitate	Schizophrenia	RTU aqueous nanosuspension in prefilled syringe	I.M. 0.75 mL	234	1 month
Haldol decanoate injection	Haldol decanoate	Schizophrenia	Oily solution in ampoule	I.M. 1 mL	100	1 month
Eligard	Leuprolide acetate	Prostate cancer	*In situ*-forming gel	S.C.		6 months

2. *Microparticles*

Microparticles based on PLGA copolymers have been marketed as s.c. or i.m. injectable depots for up to six-month dosing intervals.

3. *In Situ Depots*

Biodegradable injectable *in situ*-forming drug-delivery systems represent an attractive alternative to more established parenteral depot systems like microspheres or implants. These *in situ*-forming systems are typically cosolvent-dissolved polymer drug solutions that gel after subcutaneous injection as the cosolvents diffuses out of the remaining polymer matrix. These have advantages, such as ease of manufacturing and administration. An example is Atrigel™ [87].

References

[1] Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov 2006;5:993–6.
[2] Avdeef A. Solubility of sparingly-soluble ionizable drugs. Adv Drug Deliv Rev 2007;59:568–90.
[3] Greenwald RB, Choe YH, McGuire J, Conover CD. Effective drug delivery by PEGylated drug conjugates. Adv Drug Deliv Rev 2003;55 (2):217–50.
[4] Arouri A, Hansen AH, Rasmussen TE, Mouritsen OG. Lipases, liposomes, and lipid-prodrugs. Curr Opin Colloid Interface Sci 2013;18:419–31.
[5] Vermeulen A, Wouters A. Dosing regimen associated with long acting injectable paliperidone esters. US 2009/0163519Al (Patent); 2009.
[6] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 1997;1–3:3–25.
[7] Vieth M, Siegel MG, Higgs RE, Watson IA, Robertson DH, Savin KA, et al. Characteristic physical properties and structural fragments of marketed oral drugs. J Med Chem 2004;47:224–32.
[8] Amidon GL, Lennernäs H, Shah VP, Crison JR. A theoretical basis for a biopharmaceutic drug classification: the correlation of *in vitro* drug product dissolution and *in vivo* bioavailability. Pharm Res 1995;12:413–20.
[9] US Department of Health and Human Services, FDA Center for Drug Evaluation and Research. Guidance for industry: waiver of *in vivo* bioavailability and bioequivalence studies for immediate-release solid oral dosage forms based on a biopharmaceutics classification system, Available from: <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm070246.pdf>; 2000.
[10] Johnson KC, Swindell AC. Guidance in the setting of drug particle size specification to minimise variability in absorption. Pharm Res 1996;13(12):1795–8.
[11] Curatolo W. Physical chemical properties of oral drug candidates in the discovery and exploratory development settings. Pharm Sci Technol Today 1988;1(9):387–93.
[12] Xuan D, Rose JP, Van Gelder J. Developability assessment of clinical drug products with maximum absorbable. Int J Pharm 2012;427:260–9.
[13] Kawakami K. Modification of physicochemical characteristics of active pharmaceutical ingredients and application of supersaturatable dosage forms for improving bioavailability of poorly absorbed drugs. Adv Drug Deliv Rev 2012;64:480–95.
[14] Merisko-Liversidge E, Liversidge GG. Nanosizing for oral and parenteral drug delivery: a perspective on formulating poorly-water soluble compounds using wet media milling technology. Adv Drug Deliv Rev 2011;63(6):427–40.
[15] Box KJ, Volgyi G, Baka E, Stuart M, Takacs-Novak K, Comer JEA. Equilibrium versus kinetic measurements of aqueous solubility, and the ability of compounds to supersaturate in solution: a validation study. J Pharm Sci 2006;95(6):1298–307.
[16] Breitenbach J. Melt extrusion: from process to drug delivery technology. Eur J Pharm Biopharm 2002;54:107–17.

[17] Paudel A, Worku ZA, Meeus J, Guns S, van den Mooter G. Manufacturing of solid dispersions of poorly water soluble drugs by spray drying: formulation and process considerations. Int J Pharm 2013;453:253–84.
[18] Pouton CW. Formulation of poorly water-soluble drugs for oral administration: physicochemical and physiological issues and the lipid formulation classification system. Eur J Pharm Sci 2006;29:278–87.
[19] Muellertz A, Ogbonna A, Ren S, Rades T. New perspectives on lipid and surfactant based drug delivery systems for oral delivery of poorly soluble drugs. J Pharm Pharmacol 2010;62:1622–36.
[20] Mu H, Holm R, Müllertz A. Lipid-based formulations for oral administration of poorly water-soluble drugs. Int J Pharm 2013;453:215–24.
[21] The Lipid Library. Available from: <http://lipidlibrary.aocs.org/>.
[22] Lipid Maps. Lipid Metabolites and Pathways Strategy. Available from: <http://www.lipidmaps.org/>; 2015.
[23] US Pharmacopeial Convention. USP37-NF32. Rockville, Maryland: US Pharmacopeial Convention (USP); 2014.
[24] Bookbinder LH, Hofer A, Haller MF, Zepeda ML, Keller G-A, Lim JE. A recombinant human enzyme for enhanced interstitial transport of therapeutics. J Control Release 2006;114:230–41.
[25] Bode BW. Insulin pump use in type 2 diabetes. Diabetes Technol Ther 2010;12(Suppl 1):S17–21.
[26] US Food and Drug Administration. Inactive ingredient search for approved drugs. Available from: <http://www.accessdata.fda.gov/scripts/cder/iig/index.cfm>; 2013.
[27] Nema S, Brendel RJ. Excipients and their role in approved injectable products: current usage and future directions. PDA J Pharm Sci Technol 2011;65:287–332.
[28] Mottu F, Laurent A, Rüfenacht D, Doelker E. Organic solvents for pharmaceutical parenterals and embolic liquids: a review of toxicity data. PDA J Pharm Sci Technol 2000;54:456–69.
[29] Mottu F, Stelling M-J, Ruefenacht DA, Doelker E. Comparative hemolytic activity of undiluted organic water-miscible solvents for intravenous and intra-arterial injection. PDA J Pharm Sci Technol 2001;55:16–23.
[30] Yalkowsky SH, Roseman TJ. Solubilization of drugs by cosolvents. In: Yalkowsky SH, editor. Techniques of solubilization of drugs. New York: Marcel Dekker, Inc; 1981. p. 91–134.
[31] Yalkowsky SH, Flynn G, Amidon GL. Solubility of nonelectrolytes in polar solvents. J Pharm Sci 1972;61:983–4.
[32] Yalkowsky SH, Valvani SC, Amidon GL. Solubility of non-electrolytes in polar solvents IV: nonpolar drugs in mixed solvents. J Pharm Sci 1976;65:1488–94.
[33] Valvani SC, Yalkowsky SH, Roseman TJ. Solubility and partitioning IV: aqueous solubility and octanol–water partition coefficients of liquid nonelectrolytes. J Pharm Sci 1981;70:502–7.
[34] Gould PL, Goodman M, Hanson PA. Investigation of the solubility relationships of polar, semi-polar, and non-polar drugs in mixed co-solvent systems. Int J Pharm 1984;19:149–59.
[35] Alvarez-Nunez FA, Pinsuwan S, Lerkpulsawad S, Yalkowsky SH. Effect of the most common co-solvents upon the extent of solubilization of some drugs: log-linear model. AAPS Annual Meeting 1998;1(4):Poster 2433.
[36] Kramer SF, Flynn GL. Solubility of organic hydrochlorides. J Pharm Sci 1972;61:1896–904.
[37] Yalkowsky SH, Valvani SC. Precipitation of solubilized drugs due to injection or dilution. Drug Intel Clin Pharm 1977;11:417–19.
[38] Strickley RG. Review article: solubilizing excipients in oral and injectable formulations. Pharm Res 2004;21(2):201–30.
[39] Lee YC, Zocharski PD, Samas B. An intravenous formulation decision tree for discovery compound formulation development. Int J Pharm 2003;253:111–19.
[40] Bittner B, Mountfield RJ. Intravenous administration of poorly soluble new drug entities in early drug discovery: the potential impact of formulation on pharmacokinetic parameters. Curr Opin Drug Disc Dev 2002;5(1):59–71.
[41] Bittner B, Mountfield RJ. Formulation and related activities for the oral administration of poorly water-soluble compounds in early discovery animal studies: Part 2. Pharm Ind 2002;64(9):985–91.
[42] Neervannan S. Preclinical formulations for discovery and toxicology: physicochemical challenges. Expert Opin Drug Metab Toxicol 2006;2(5):715–31.
[43] Szebeni J, Muggia F, Gabizon A, Barenholz Y. Activation of complement by therapeutic liposomes and other lipid excipient-based therapeutic products: prediction and prevention. Adv Drug Deliv Rev 2011;63:1020–30.
[44] Jain N, Yang G, Yalkowsky SH. Solubilization of NSC-639829 by combined effects of pH with cosolvents, surfactants, and complex ants. Int J Pharm 2001;225:41–7.
[45] Thoma K, Albert K. Amphiphilic drugs. 2. Relation between colloidal properties and pharmaceutic-technological, pharmacokinetic, and pharmacodynamic behavior. Pharmazie 1983;38:807–17.
[46] Hussain MA, Wu LS, Koval C, Hurwitz AR. Parenteral formulation of the kappa agonist analgesic, DuP 747, via micellar solubilization. Pharm Res 1992;9:750–2.
[47] Florence AT, Attwood D. Physicochemical principles of pharmacy. 3rd ed. Hampshire: McMillan Press; 1998.
[48] Floyd AG. Top ten considerations in the development of parenteral emulsions. PSTT 1999;2:134–43.
[49] Lee MJ, Lee MH, Shim CK. Inverse targeting of drugs to the reticuloendothelial system: rich organs by microemulsions emulsified with Poloxamer 338. Int J Pharm 1995;113:175–87.
[50] Liu F, Liu D. Long circulating emulsions (oil-in-water) as carriers for lipophilic drugs. Pharm Res 1995;12:1060–4.
[51] Chavanet PY, Garry I, Charlier N, Caillot D, Kisterman JP, D'Athis M. Trial of glucose versus fat emulsion in preparation of amphotericin for use in HIV infected patients with candidiasis. BMJ 1992;305:921–5.
[52] Aulton ME, editor. Pharmaceutics: the science of dosage form design. 2nd ed. Churchill Livingstone; 2001. Chapter 6, p. 87–94.
[53] Drewe J, Meier R, Vonderscher JDK, Posanski U, Kissel T, Gyr K. Enhancement of the oral absorption of cyclosporin in man. Br J Clin Pharmacol 1992;34:60–4.
[54] Strickley RG, Anderson BD. Solubilisation and stabilisation of an anti-HIV thiocarbamate, NSC 629243, for parenteral delivery, using extemporaneous emulsions. Pharm Res 1993;10:1076–82.
[55] Lundberg B. Preparation of drug-carrier emulsions stabilised with phosphatidylcholine surfactant mixtures. J Pharm Sci 1994;83:72–5.

[56] Lundberg B. A submicron lipid emulsion coated with amphipathic polyethylene glycol for parenteral administration of paclitaxel (Taxol). J Pharm Pharmacol 1997;49:16–21.
[57] Vanek VW, Seidner DL, Penny A, Bistrian B, Collier S, Kathleen G, et al. A.S.P.E.N. position paper: clinical role for alternative intravenous fat emulsions. Nutr Clin Pract 2012;27:150–92.
[58] Prankerd RJ, Stella VJ. The use of oil-in-water emulsions as a vehicle for parenteral drug administration. J Parent Sci Technol 1990;44:139–49.
[59] Davis SS, Hadgraft J, Palin KJ. Medical and pharmaceutical applications of emulsions. In: Becher P, editor. Encyclopedia of emulsion technology. New York: Dekker; 1985. p. 159–238.
[60] Davis SS, Washington C, West P, Illum L, Liversedge G, Sternson L, et al. Lipid emulsions as drug delivery systems. Ann NY Acad Sci 1987;507:75–88.
[61] Singh M, Ravin JL. Parenteral emulsions as drug carrier systems. J Parenter Sci Technol 1986;40:34–41.
[62] Boyett JB, Davis CW. Injectable emulsions and suspensions. In: Liberman HA, Rieger MM, Banker GS, editors. Pharmaceutical dosage forms: disperse systems. New York: Dekker; 1989. p. 379–416.
[63] Collins-Gold L, Feichtinger N, Waernheim T. Are lipid emulsions the drug delivery solution? Mod Drug Disc 2000;3(3):44–8.
[64] Charman WN. Lipid vehicle and formulation effects on intestinal lym phatic drug transport. In: Charman WN, Stella VL, editors. Lymphatic transport of drugs. Boca Raton: CRC Press; 1992. p. 113–79.
[65] Charman WN. Lipids, lipophilic drugs, and oral drug delivery: some emerging concepts (mini-review). J Pharm Sci 2000;89:967–78.
[66] Charman WN. Lipids, lymph, and lipidic formulations. Bull Tech Gattefossé 1997;90:27–33.
[67] Chaudhary RK, Shariff I, Dolphin D. Drug release characteristics of lipid based benzoporphyrin derivative. J Pharm Pharmaceut Sci 2003;6(1):13–19.
[68] Duncan R, Gaspar R. Nanomedicine(s) under the microscope. Mol Pharm 2011;8(6):2101–41.
[69] Stern ST, Hall JB, Yu LL, Wood LJ, Paciotti GF, Tamarkin L, et al. Translational considerations for cancer nanomedicine. J Control Release 2010;146:164–74.
[70] Kwon IK, Lee SC, Han B, Park K. Analysis of the current status of targeted drug delivery to tumors. J Control Release 2012;164:108–14.
[71] Mann LL. Aberrant vascular architecture in tumours and its importance in drug-based therapies. Drug Discov Today 2003;8:396–403.
[72] Moghimi SM, Bonnemain B. Subcutaneous and intravenous delivery of diagnostic agents to the lymphatic system: applications in lymphoscintigraphy and indirect lymphography. Adv Drug Del Rev 1999;37:295–312.
[73] Philips WT, Klipper R, Goins B. Novel method of greatly enhanced delivery of liposomes to lymph nodes. J Pharmacol Exp Ther 2000;295:309–13.
[74] Moghimi SM. Modulation of lymphatic distribution of subcutaneously injected poloxamer 407 coated nanospheres: the effect of the ethylene oxide chain configuration. FEBS Lett 2003;540:241–4.
[75] Barenholz Y. Doxil—The first FDA-approved nano-drug: lessons learned. J Control Release 2012;160:117–34.
[76] Cern A, Barenholz Y, Tropsha A, Goldblum A. Computer-aided design of liposomal drugs: in silico prediction and experimental validation of drug candidates for liposomal remote loading. J Control Release 2014;173:125–31.
[77] Allen TM, Hansen C. Pharmacokinetics of stealth versus conventional liposomes: effect of dose. Biochim Biophys Acta 1991;1068:133–41.
[78] Allen TM, Stuart DD. Liposomal pharmacokinetics – classical, sterically-stabilised, cationic liposomes and immunoliposomes. In: Janoff AS, editor. Liposomes: rational design. New York: Marcel Dekker, Inc.; 2005. p. 63–87.
[79] Veronese FM, Pasut G. PEGylation: successful approach to drug delivery. Drug Discov Today 2005;10(21):1451–8.
[80] Vicent MJ, Ringsdorf H, Duncan R. Polymer therapeutics: clinical applications and challenges for development. Adv Drug Deliv Rev 2009;61:1117–20.
[81] Kratz F. Albumin as a drug carrier: design of prodrugs, drug conjugates, and nanoparticles. J Control Release 2008;132:171–83.
[82] Sun B, Yeo Y. Nanocrystals for the parenteral delivery of poorly water-soluble drugs. Curr Opin Solid State Mater Sci 2012;16:295–301.
[83] Van der Meel R, Vehmeijer LJC, Kok RJ, Storm G, van Gaal EVB. Ligand-targeted particulate nanomedicines undergoing clinical evaluation: current status. Adv Drug Deliv News 2013;65:1284–98.
[84] Ho RJY, Chien J. Trends in translational medicine and drug targeting and delivery: new insights on an old concept—targeted drug delivery with antibody-drug conjugates for cancers. J Pharm Sci 2014;103:71–7.
[85] Alabia CA, Lovea KT, Sahaya G, Yina H, Lulyc KM, Langer R. Multiparametric approach for the evaluation of lipid nanoparticles for siRNA delivery. Proc NAS 2013;110:12881–6.
[86] Kranz H, Bodmeier R. Structure formation and characterization of injectable drug loaded biodegradable devices: in situ implants versus in situ microparticles. Eur J Pharm Sci 2008;34:164–72.
[87] Packhaeuser CB, Schnieders J, Oster CG, Kissel T. In situ forming parenteral drug delivery systems: an overview. Eur J Pharm Biopharm 2004;58:445–55.
[88] Griffin WC. Classification of surface active agents by HLB. J Soc Cosmet Chem 1949;1:311.

CHAPTER

30

Preparation of Water-Soluble Compounds by Covalent Attachment of Solubilizing Moieties

Camille G. Wermuth[1] *and Dominique Lesuisse*[2]

[1]**Founder of Prestwick Chemical, Strasbourg, France;** [2]**Lead Generation Board, Sanofi,** AVENUE PIERRE BROSSOLETTE, CHILLY MAZARIN **- France**

OUTLINE

"Ajouter à sa queue, ôter à ses oreilles"
(Add to her tail, remove from her ears) ***Jean de LA FONTAINE, La Besace***

I. INTRODUCTION

The strategy described in this chapter aims to convert a water-insoluble drug into a water-soluble one by covalently attaching an appropriate solubilizing side chain. Surprisingly few reviews covering this subject are found in the literature [1–5]. Seen from the chemical side, the solubilizing moiety can be a neutral hydrophilic group or an ionizable organic base or acid. With the exception of possible crystallization problems, no major difficulties are expected in the synthesis of such compounds. One problematic aspect of the solubilization approach to be taken is to decide if the solubilizing moiety has to be fixed in a reversible manner, generating

TABLE 30.1 Successful Examples of Drug Solubilization by a Chemical Means

Solubilizing side chain	Therapeutic class	Compound
Phosphoric ester	Steroids	Betamethazone
Phosphoric ester	Vitamins	Menadione
Phosporic amide	Antibiotics	Ceftaroline fosamil [6]
Phosphoric amide	Antiemetic	Fosaprepitant [7]
Hemisuccinate	Cardiotonics	Benfurodil
Hemisuccinate	Antibiotics	Chloramphenicol
Hemisuccinate	Steroides	Prednisolone
Hemisuccinate	Benzodiazepines	Oxazepam
Acidic side chain	Theophylline	Etaphylline
Acidic side chain	Antisyphilitic	Solusalvarsan
Neutral side chain	Analgesic	Glafenine
Neutral side chain	Bronchodilator	Dyphylline
Neutral side chain	Antibacterial	Sulfapyridine N-glucoside
Basic side chain	Antibiotics	Rolitetracycline
Basic side chain	Flavonoids	Solurutine
Basic side chain	Morphine	Pholcodine
Basic side chain	Kinase inhibitor	Bosutinib [8]

a prodrug, or in an irreversible manner, yielding a new chemical entity. In the latter case, the solubilizing procedure may exact a cost in terms of the recognition mechanisms, the soluble analog being less potent or even showing a different pharmacological profile. In some instances, changes in one part of the molecule have to be compensated by changes in another part. As Jean de La Fontaine said of the elephant: "Add to her tail, remove from her ears." In addition, the solubilized analog of an already approved drug is considered by the governmental drug agencies as a totally new chemical entity, demanding a completely new development process. The financial investment that is then necessary can only be justified if enough sales of the solubilized form are expected. As a consequence, the attachment of solubilizing moieties has to be considered very early in the drug-discovery process or else limited to drugs with sizeable markets and undertaken when all other solubilizing stratagems fail. Despite these difficulties, many examples are found in therapy of successful drug solubilization by means of a chemical transformation of a parent drug (Table 30.1).

Chemically solubilized active principles render possible the preparation of parenteral and especially intravenous forms appreciated in the clinical practice. But even at the preclinical level, the use of water-soluble molecules is recommended as they are effectively much easier to study by *in vitro* tests, in cell or microorganism cultures, and on isolated organs. The inconveniences are that chemically modified structures may show modified pharmacological, pharmacokinetic, and toxicological properties.

II. SOLUBILIZATION STRATEGIES

Three points are decisive in terms of the solubilizing strategies: How will the solubilizing moiety be grafted? Where will it be grafted? What kind of side chain will be utilized?

A. Mode of Linking the Solubilizing Moiety

The solubilizing chain can be reversibly or irreversibly grafted to the parent molecule. In the case of reversible linkages, we are dealing in fact with prodrugs. Reversible linkages are usually provided by esters, peptides, or glucosides.

TABLE 30.2 Small and Large Solubilizing Moieties

Small groups or simple functionalities	Larger solubilizing moieties
—CO_2H	R—OH → R—O—CH_2—CH_2—CO_2H
—SO_3H, —OSO_3H	R—NH_2 → R—NH—CH_2—CH_2— CH_2—SO_3H
—PO_3H_2, —OPO_3H_2	$(R)_2C{=}O$ → $(R)_2C{=}N$—O—CH_2— CO_2H
—NH_2, —NHR, —NR_2	R—OH → O-morpholinylethyl
N-Oxides	R—OH → O-glucoside
S-Oxides	R—OH → O—CO—CH_2— CH_2—CO_2H
Sulfones	R—OH → *m*—O—C_6H_4—SO_3H

Irreversible attachment of side chains is achieved by *O*- and *N*-alkylation and creation of carbon−carbon bonds. The grafted side chains can be basic (dimethylaminoethyl or morpholinoethyl chains), acidic (carboxylic, sulfonic), or neutral (glyceryl).

Intermediate situations are found for enol and phenol phosphates as well as for some amides. For these compounds, only partial reversibility is observed *in vivo*.

B. Position of Linking the Solubilizing Moiety

First, a careful examination of the parent molecule must be undertaken in order to identify the parts of the molecule that present adequate chemical reactivity and are suitable as attachment points for the solubilizing chain. Functions such as OH, SH, NH, acidic CH, or CO_2H are reactive sites that furnish nucleophilic or basic entities. Conversely, aromatic double bonds are sensitive to electrophilic attack, whereas carbonyl groups and conjugated carbon−carbon double bonds are sensitive to nucleophilic attack. The second criterion that has to be considered is of a biological nature. The solubilizing chain can only be grafted to those parts of the molecule that are not involved in the drug−receptor interaction. Fixed at the wrong place, the solubilizing chain can completely inactivate the molecule.

C. Choice and Potential Impacts of the Solubilizing Chain

The size of the solubilizing chain is one of the selection parameters. The chains can be limited to the strict minimum and simply represent functional groups, or they can be made from larger residues containing several atoms (Table 30.2).

The nature of the side chains is the next selection parameter. It has to be decided if they can be ionizable (acidic or basic) or nonionizable. This choice will have a huge impact on several properties of the active principle, from pharmacology to pharmacokinetic [9] and toxicity [10], and not forgetting the galenic outcomes. This will be briefly reviewed in the following paragraphs.

Acidic ionizable moieties (e.g., carboxylic acids) yield readily crystalline compounds.. However, only a limited number of cations can be used to neutralize them. Traditional inorganic cations such as sodium, potassium, or magnesium can induce mineral surcharges and are no longer recommended. They are advantageously replaced by organic bases such as tromethamine, lysine, or *N*-methylglucamine. Carboxilic acids will represent a large hurdle for negatively charged lipid membranes and—in particular—CNS penetration [11]. Owing to their amphiphilic nature, they can show hemolytic properties. An additional potential liability of carboxylic acids is that their glucuroconjugation might give rise to reactive acyl glucuronides after rearrangement, provoking liver and other organs liabilities [12]. Another property of acids is that they are often linked to high-protein binding, as human albumin has several sites recognizing carboxylic acids [13]. This may positively impact several pharmacokinetics parameters, such as clearance and metabolism, but negatively impact *in vivo* activity of the molecule, in particular for CNS indications.

Basic ionizable moieties (e.g., substituted amines) can be neutralized by a very large number of organic and inorganic acids. The salts obtained are also readily crystallized and usually show less surface-active properties than salts from acidic chains. Drugs with acidic side chains should not be mixed with drugs having basic side chains, as it is likely that a salt formed between the two drugs might precipitate. One of the potential

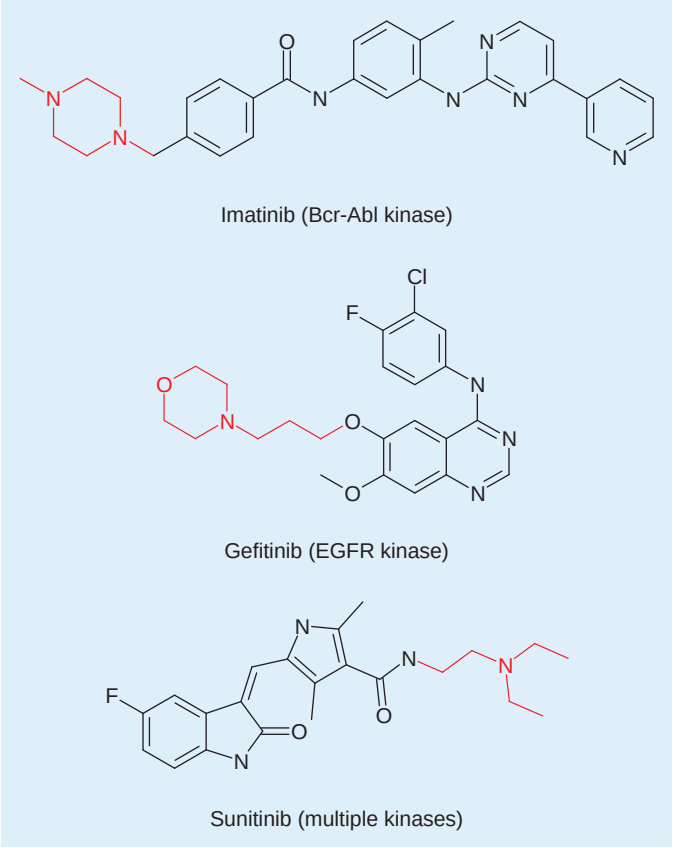


FIGURE 30.1 Kinase inhibitors on the market with basic solubilizing moieties.

disadvantages of this solubilizing moiety is its tendency to interfere with biogenic amines and neurotransmitters. In other words, attaching a basic amine functional group can seriously modify the pharmacological activity with regard to the parent drug. The strategy of solubilizing by means of a basic side chain has been extensively used in certain areas of medicinal chemistry, such as kinase inhibitors. More than twenty kinase inhibitors have come to the market in the last fifteen years. Several bear a basic side chain in order to bring solubility to these very flat and hydrophobic compounds (Figure 30.1). In many cases, however, the amine function has also brought some undesirable hErg channel inhibition, and the obtention of a compound devoid of cardiac liability has been a real challenge [14].

Nonionizable moieties (e.g., polyhydroxylated chains) do not present this disadvantage and are compatible with other drug preparations. As they can be delivered at pH values close to 7, they do not produce painful injections. Not always intuitive, they have the potential to bring more solubility than ionizable functions, as was recently pointed by a recent match pair analysis on various fragments [15]. The main problems encountered with nonionizable solubilizing moities is their lesser propensity to crystallize. In addition, increased cost can be expected from the necessity of added protection/deprotection steps during their synthesis.

III. ACIDIC SOLUBILIZING CHAINS

When planning the solubilization by means of a carboxylic acid side chain, one has to take into account the therapeutic properties peculiar to the carboxylic group. Thus, all arylacetic acids show more or less potent anti-inflammatory activities, and many α-functionalized carboxylic acids are chelating agents. Among them we find the chelating α-amino acids [10], antibacterial nalidixic acid-derived quinolones [16], and probably kynurenic acid analogs acting as antagonists at the glycine site [17].

A. Direct Introduction of Acidic Functions

Direct introduction of a solubilizing function can be achieved by carboxylation and by sulfonation. The historical example of carboxylation is the Kolbe synthesis of salicylic acid. Sulfonation was employed to solubilize guaiacol, camphor, and 7-chloro-8-hydroxyquinoline (Figure 30.2).

In a recent example, aryl-carboxylic solubilization was used to solubilize core-modified porphyrins [18]. Solubilizing aryl-carboxylic functions were also replaced by their tetrazole or 1,2,4-oxadiazolone bioisosteres in the design of second-generation, benzodiazepine-derived CCK-B antagonists (Figure 30.3) [19].

In a series of adenosine A_1 receptor antagonists [20], simple homologation of the carbon chain bearing the carboxylic group improved the bioavailability, the selectivity, and the water solubility (Figure 30.4).

B. Alkylation of OH and NH Functions with Acidic Chains

This procedure alkylates the hydroxy and the amino groups already present on the molecule with reactive intermediates bearing acidic functional residues (Table 30.3). These compounds are prepared starting from chloroacetic acid or its ethyl ester. For chains longer than acetic, cyanoethylation and hydrolysis of the nitrile obtained leads to the propionic chain, while alkylation with ethyl 4-bromobutyrate and saponification leads to the butyric chain. The propane-sulfonic chains are particularly accessible by means of ring opening of propane-sultone.

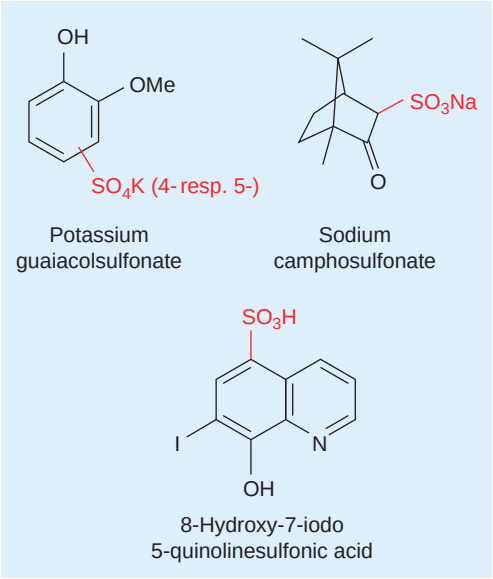


FIGURE 30.2 Sulfonic acid solubilization.

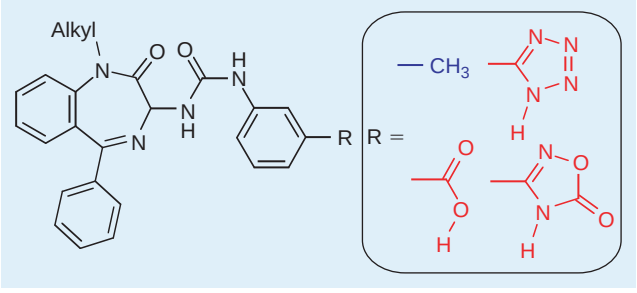


FIGURE 30.3 Carboxylic acid and heterocyclic bioisosteres as solubilizing groups.

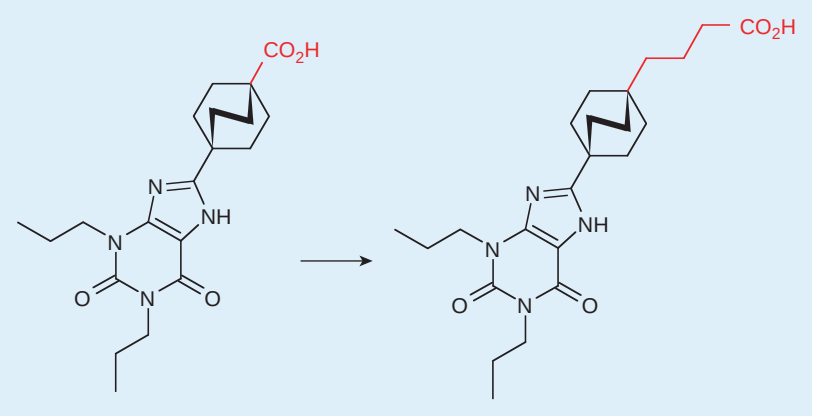


FIGURE 30.4 Homologation of the carboxylic solubilizing group provides improved bioavailability.

TABLE 30.3 Alkylation of OH and NH Functions with Acidic Chains

Starting derivative	Solubilized analog	Example	Reference
Ar—OH	Ar—O—CH_2—CO_2H	Solusalvarsan	[3]
Ar—OH	Ar—O—CH_2—CO_2H	Flavodic acid	[23]
Ar—NH_2	Ar—NH—CH_2—CO_2H	Acediasulfone	[24]
Ar—NH_2	Ar—NH—CH_2—CO_2H	Iodopyracet3	[25]
Ar—NH_2	Ar—NH—CH_2—SO_2H	Sulfoxone sodium	[26]

1. Dihydroartemisinin Ethers

A water-soluble derivative of artemisinin, the sodium salt of artesunic acid (the succinic half-ester derivative of dihydroartemisinin; Figure 30.5), can be administered by intravenous injection, a property that makes it especially useful in the treatment of advanced and potentially lethal cases of Plasmodium falciparum. Sodium artesunate is capable of rapidly reversing parasitaemia and causing the restoration to consciousness of the comatose cerebral malaria patient. The utility of sodium artesunate, however, is impaired by its poor stability due to the facile hydrolysis of the ester linkage. To overcome the ease of hydrolysis of the ester function in sodium artesunate, Lin et al [21] prepared a series of analogs in which the solubilizing moiety was joined to dihydroartemisinin by a ether rather than an ester linkage. One of the compounds prepared, artelinic acid (Figure 30.5), is both soluble and stable in 2.5 percent K_2CO_3 solution and possesses superior *in vivo* activity against Plasmodium berghei in comparison to artermisinin or artesunic acid [21].

Continuing the search for water-soluble dihydroartemisin derivatives with higher efficacy and longer plasma half-life than artesunic or artelinic acid, Lin et al prepared a series of dihydroartemisinoxy-butyric acids bearing an aryl substituent at the 4-position of the butyric side chain (Figure 30.5). The *p*-chlorophenyl and the *p*-bromophenyl derivatives showed a 5–8-fold increase in *in vitro* antimalarial activity against D-6 and W-2 clones of Plasmodium falciparum than artemisin or artelinic acid. They also showed *in vivo* oral antimalarial activity superior to that of artelinic acid [22]. Other ether-linked artemisinin-solubilizing chains containing asymmetric centers did not show activities superior to that of artelinic acid [27].

C. Acylation of OH and NH Functions with Acidic Chains

The acylation of OH and NH functions with acidic chains is probably the most popular mode of acidic solubilization. Alcohols and phenols are converted not only into half-esters such as hemisuccinates, hemiglutarates, hemiphtalates [28], and *meta*-benzenesulfonates [29] but also into phosphates or even sulfates (Figure 30.6). All these derivatives can give water-soluble sodium or amine salts. Similar acylation possibilities exist for amines, but peptide-like derivatives are often preferred because the enzymatic regeneration of the parent molecule *in vivo* is easier.

Carboxylic half-esters (e.g., hemisuccinates) of phenols are easily hydrolyzed in aqueous solution and are therefore not recommended for the solubilization of phenolic compounds. Even hemisuccinates of alcohols suffer

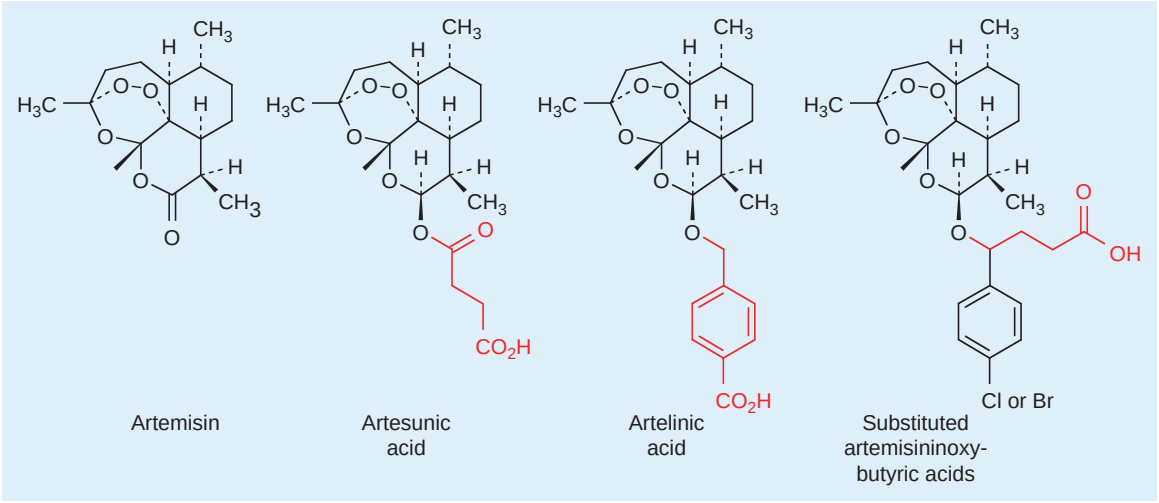


FIGURE 30.5 Solubilized forms of artemisin [21,22].

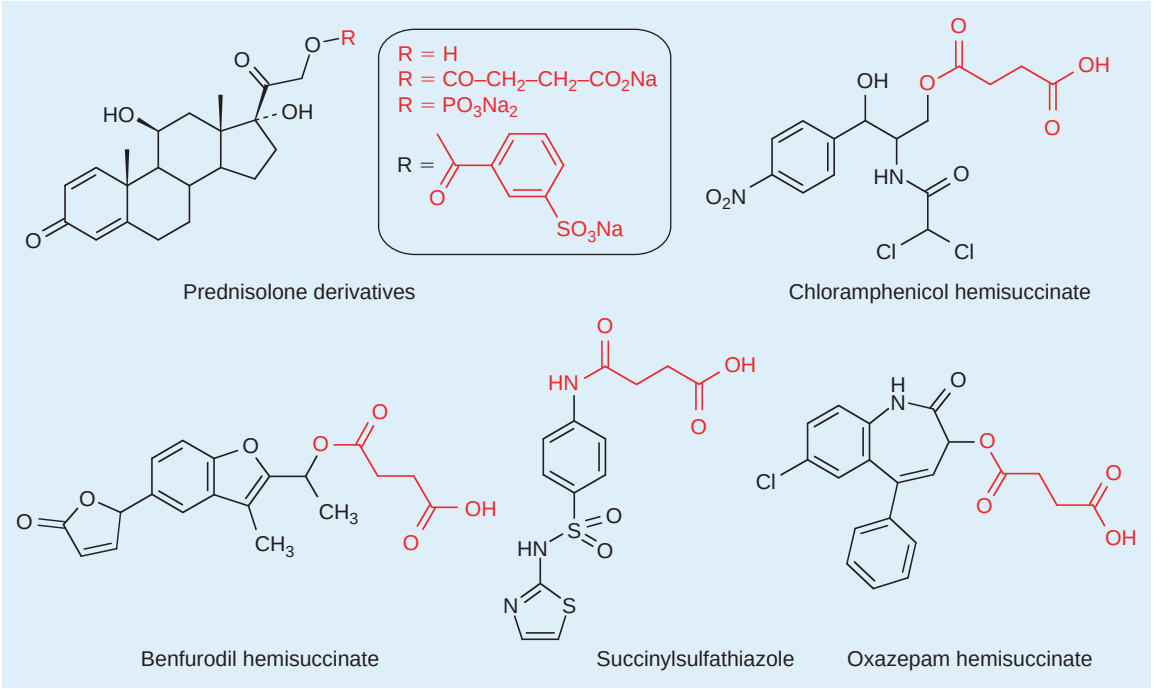


FIGURE 30.6 Acylation of OH and NH functions with acidic chains.

somewhat from stability problems and must be supplied as lyophilized (freeze-dried) powders for reconstitution in water and used within 48 hours (see, for example, the monographies chloramphenicol sodium succinate or hydrocortisone sodium succinate in *The Handbook on Injectable Drugs*; [31] see also Anderson et al [32,33]).

An additional difficulty occurring with hemisuccinates was discovered by Sandman et al [30]. These authors, in studying the stability of chloramphenicol succinate, found an unusual partial acyl transfer reaction of the succinyl group to give a cyclic hemi-orthoester (Figure 30.7).

Apparently hemiglutarates of phenolic drugs are more stable than hemisuccinates. An example is provided by the water-soluble diglutaryl-probucol, which prevents cell-induced low-density lipoproteins (LDL) oxidation (Figure 30.8) [34].

In the search for an improvement of solution stability (i.e., in minimizing the ester hydrolysis and decreasing the acyl migration), Anderson et al [35] synthesized a series of more-stable water-soluble methylprednisolone

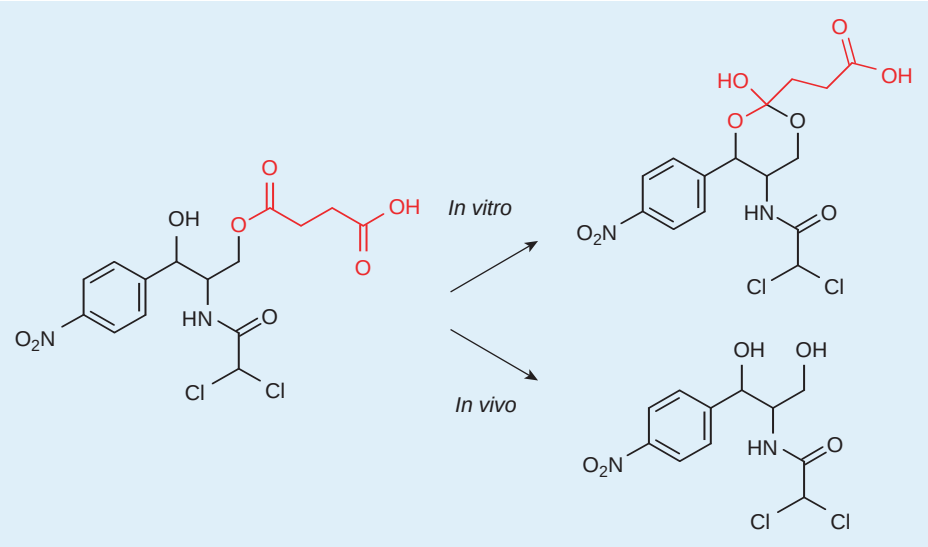


FIGURE 30.7 Formation of cyclic hemi-orthoesters from a hemisuccinate [4,30].

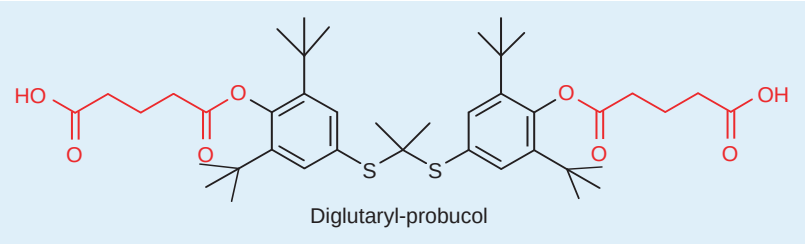


FIGURE 30.8 Structure of the diglutaryl ester of probucol [34].

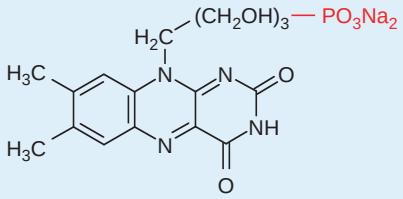
esters. Several of the analogs were shown to have shelf lives in solution of greater than two years at room temperature. Ester hydrolysis studies of these compounds in human and monkey serum indicated that derivatives having anionic solubilizing residues such as carboxylate or sulfonate are more slowly hydrolyzed by serum esterases than compounds with a cationic solubilizing moiety (tertiary amine) [35].

Phosphoric esters and amides (Figure 30.9) are generally more stable. They have been used in the steroid [45,46] and in the vitamin fields (vitamin C [47], vitamin B_1 [48], benfotiamine [49,50], riboflavine [51], dihydrovitamin K_1 [52]). Riboflavine-5′-phosphoric acid dihydrate, for example, has a solubility in water of 112 g/L at pH 6.9, compared to 0.06−0.33 g/L for riboflavine itself. Fosaprepitant, the phosphoramide derivative of prepitant, is freely water soluble and is used for injectable administration [53]. A large number of reported peptidomimetic compounds possess very low aqueous solubility at physiological pH owing to the high lipophilicity inherent in these structures. Phosphorylation can yield improved biological activities for such compounds. This is at least the case for the phosphorylated neurokinin-1 receptor antagonist [54] and the human immunodeficiency virus (HIV) protease inhibitor [36] of Figure 30.9, described by scientists from Merck and Upjohn respectively. Clean mono-phosphorylation methods are available. Some of them are shown in Figure 30.10.

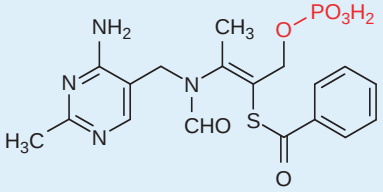
Formation of sulfate esters is one of the metabolic conjugation reactions (Phase II reactions; see Chapter 34). Sulfates of estradiol [44], glucose [55], menadiol [56], and hydroxyethyl-theophylline [57] have been prepared (Figure 30.11). Compared to their phosphoric analogs, sulfuric acid esters as a rule are resistant to enzymatic hydrolysis *in vivo* [58,59], and their conversion to the parent drug is questionable.

Sulfonic acids can be prepared by direct sulfonation (see above; Figure 30.2). Compounds containing conjugated double bonds have been solubilized by addition of sodium bisulfite. Treatment of menadione (vitamin K_3) with sodium bisulfite leads to two addition compounds (Figure 30.12). Mild warming of the reactants for a short time predominantly affords adduct (a), which arises from attack of bisulfite ion at carbon 2. Heating at reflux for an extended period yields adduct (b) from addition of bisulfite ion to carbon in position 3 [61].

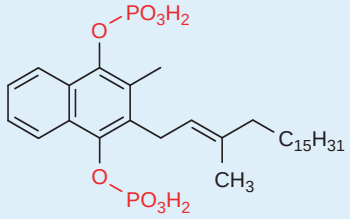
In a similar way, the treatment of N^4-cinnamylidenesulfanilamide (prepared from cinnamic aldehyde and sulfanylamide) with sodium bisulfite affords noprylsulfamide (Figure 30.12), according to the "soluseptazine principle" (noprylsulfamide is also called soluseptazine) [62]. Noprylsulfamide is freely soluble in water (200 g/L) and breaks down in the body with the liberation of free sulfanilamide.



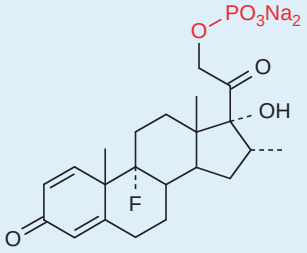
Riboflavine di-sodium phosphate



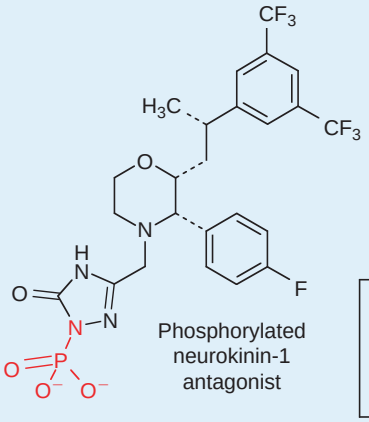
Benfotiamine



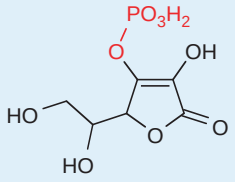
Dihydrovitamin K_1 diphosphate



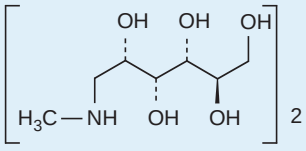
Sodium dexamethazone phosphate

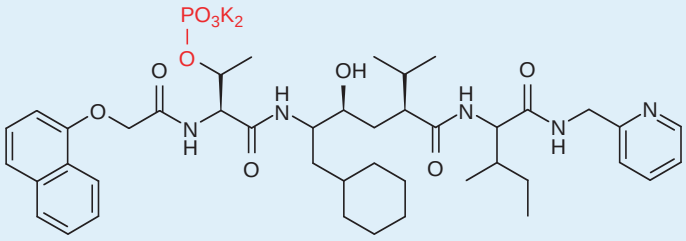


Phosphorylated neurokinin-1 antagonist

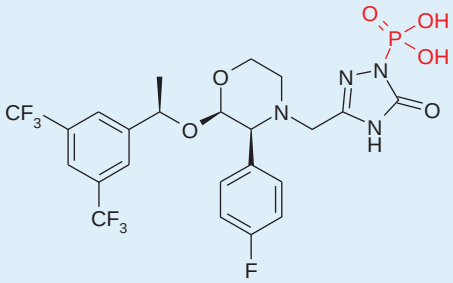


Ascorbic acid phosphate





Phosphorylated peptidomimetic HIV protease inhibitor



Fosaprepitant

FIGURE 30.9 Phosphate esters.

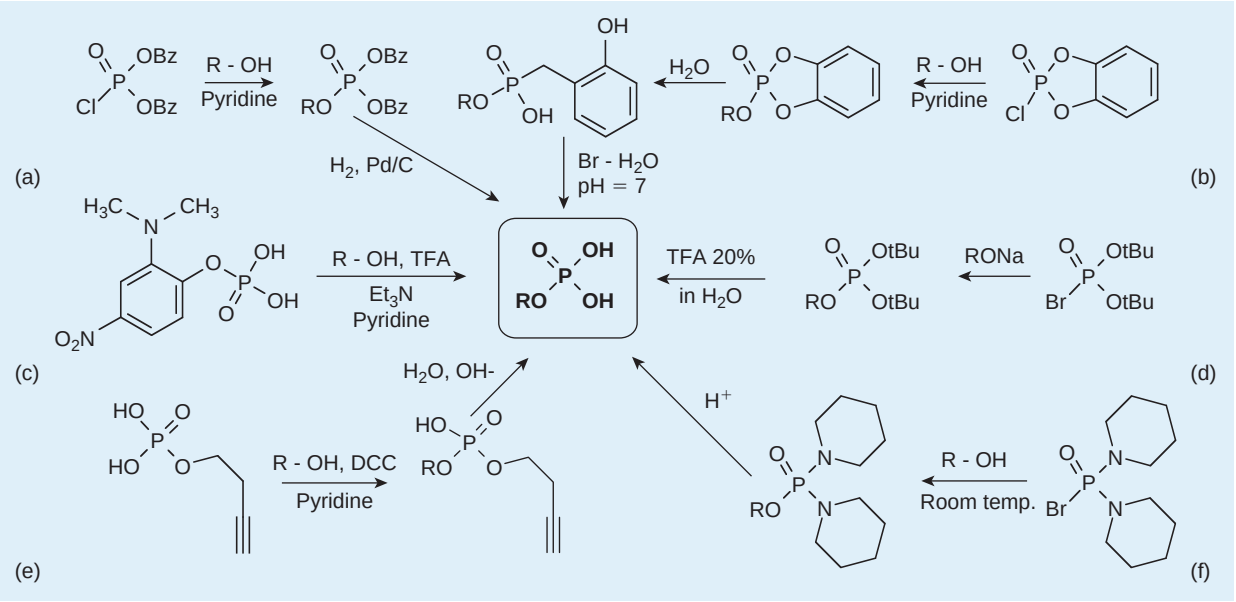


FIGURE 30.10 Useful syntheses of monophosphate esters (a) Ref. [36]. (b) Ref. [37,38], (c) Ref. [39], (d) Ref. [40,41], (e) Ref. [42–44].

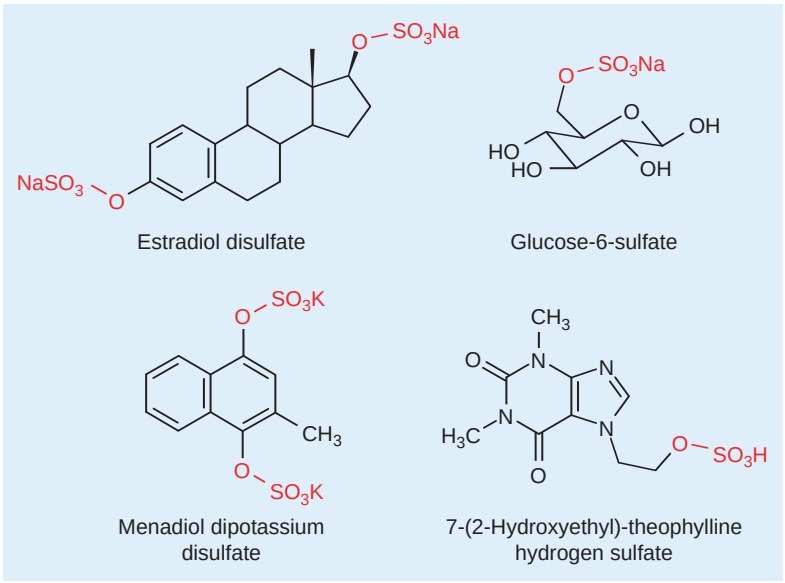


FIGURE 30.11 Sulfuric esters.

Treatment of 6-chloropurine riboside with *p*-aminobenzenesulfonic acid leads to the highly water-soluble N^6-(*p*-sulfophenyl)adenosine (solubility > 1.5 g/mL, <3 M). This compound (Figure 30.13) is a potent A_1 adenosine agonist in receptor binding, in inhibitory electrophysiological effects, in hippocampal slices, and in inhibition of lipolysis *in vivo* [60].

Treatment of primary and secondary amines with formaldehyde and sodium bisulfite generates stable methanesulfonates that can also act as solubilizing groups. The first example of this reaction found in the literature is the conversion of *p*-phenetidine into the corresponding methanesulfonate [63]. The compound obtained (Figure 30.14) still possesses antipyretic properties and is much less toxic than *p*-phenetidine. Nevertheless, it did not break into a new market [3].

Applied to noraminopyrine, the same solubilization strategy led to dipyrone [64,65], a water-soluble (1 g/1.5 mL) injectable form of aminopyrine (Pyramidon) used worldwide. Replacement of formaldehyde by acetaldehyde to yield ethanesulfonates has been claimed to lead to compounds with faster hydrolysis kinetics *in vivo* [66]. Replacement of formaldehyde by glucose afforded glucosulfone sodium, a soluble preparation of the leprostatic

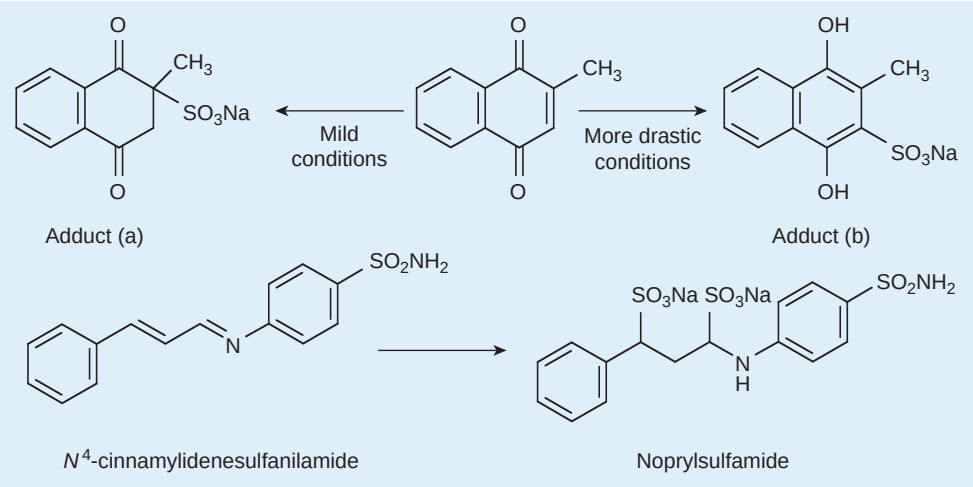


FIGURE 30.12 Bisulfite adducts.

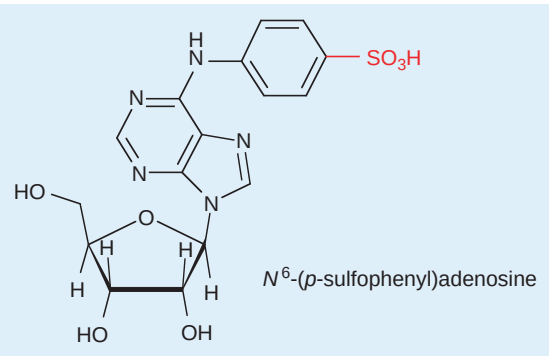


FIGURE 30.13 N^6-(*p*-sulfophenyl)adenosine, a freely water-soluble adenosine A1 receptor agonist [60].

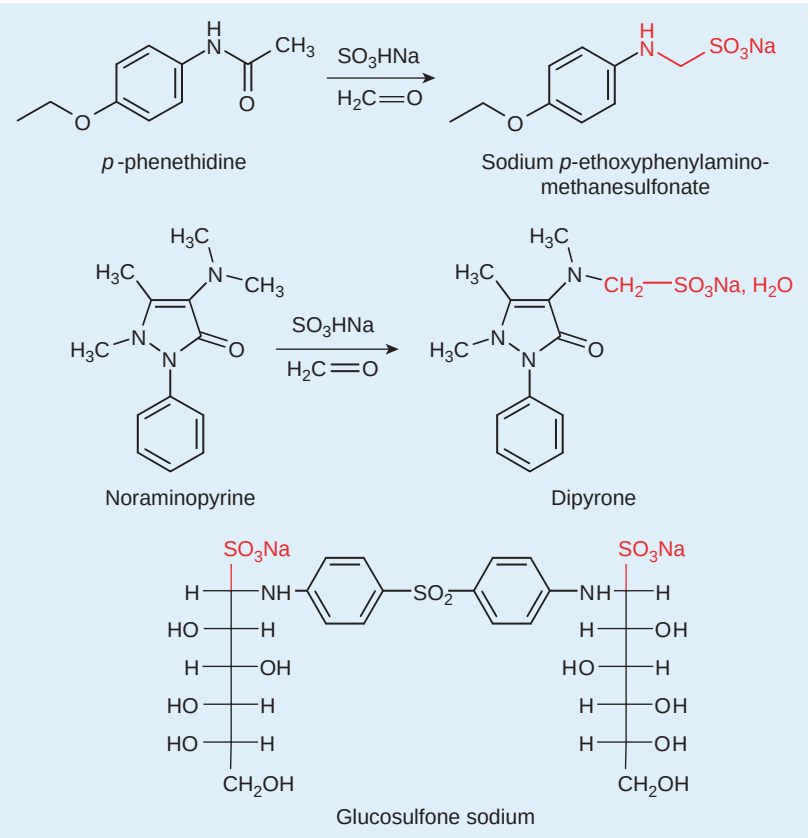


FIGURE 30.14 Sulfonates.

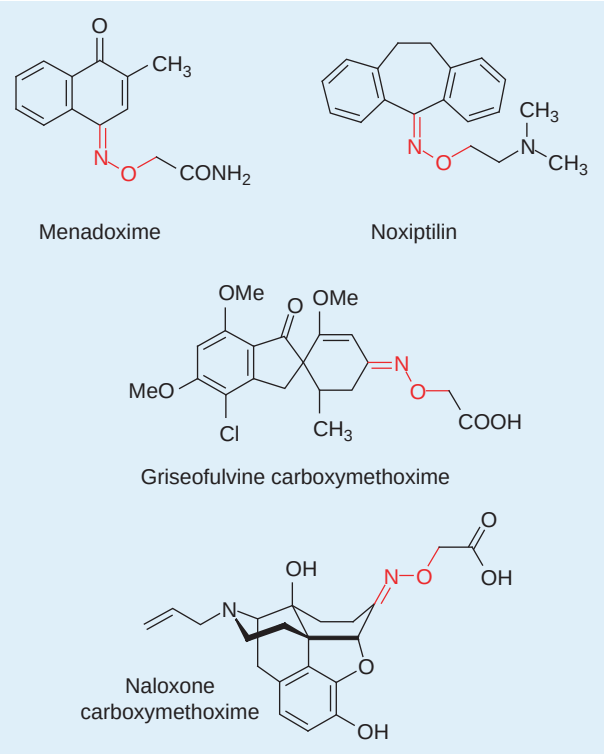


FIGURE 30.15 Carboxymethoximes as (reversible?) solubilizing chains for carbonyl-containing molecules.

4,4′-diaminodiphenylsulfone [67,68]. Replacement of formaldehyde-bisulfite by formaldehyde sulfoxylate is also claimed in some references [3,69].

Ketones can be solubilized as carboxymethoximes (Figure 30.15). Menadoxime is freely water soluble, can be sterilized by autoclaving, and like menadione itself shows high antihemorrhagic activity [66]. The carboxymethoxime of griseofulvin, on the other hand, is devoid of activity *in vitro* as well as *in vivo* [70]. This may be due either to an absence of conversion to the parent molecule or to rapid renal elimination. Unsubstituted oximes can be converted enzymatically to their corresponding ketones [71]. For carboxymethoximes, intramolecular assistance should even facilitate the hydrolysis to the initial carbonyl function. This was shown to be the case for the carboxymethoxime of naloxone. In addition, aside from solubilizing the active principle, the carboxy group has also been used to covalently attach the compound to a resin and proceed to characterize its membrane receptor, as was shown for corticosterone [72]. Other substituted oximes seem to be rather resistant, as apparent from the metabolic stability of noxiptylin [73].

IV. BASIC SOLUBILIZING CHAINS

Solubilization with basic side chains involves two essential strategies: either direct binding of the amine function on a carbon atom of the parent molecule or linking it to a function already present (e.g., alcoholic or phenolic hydroxyl, carboxylic acid, amine, amide).

A. Direct Attachment of a Basic Residue

Simple tertiary amines can be grafted to a carbon skeleton, either by exchange reactions or by Mannich reactions. The hydrochloride salt of the camptothecin derivative (a) (Figure 30.16) is soluble in water at concentrations up to 1 mg/mL. The comparable value for camptothecin itself is 0.0025 mg/mL [74]. Similar results were obtained in solubilizing the benzodiazepine (b) [75] and the quinazolinone (c) [76]. The adenosine A_1 antagonist KW-3902 (d) was solubilized in an original manner by converting it to an amidinic and cyclized bioisostere (e) [77].

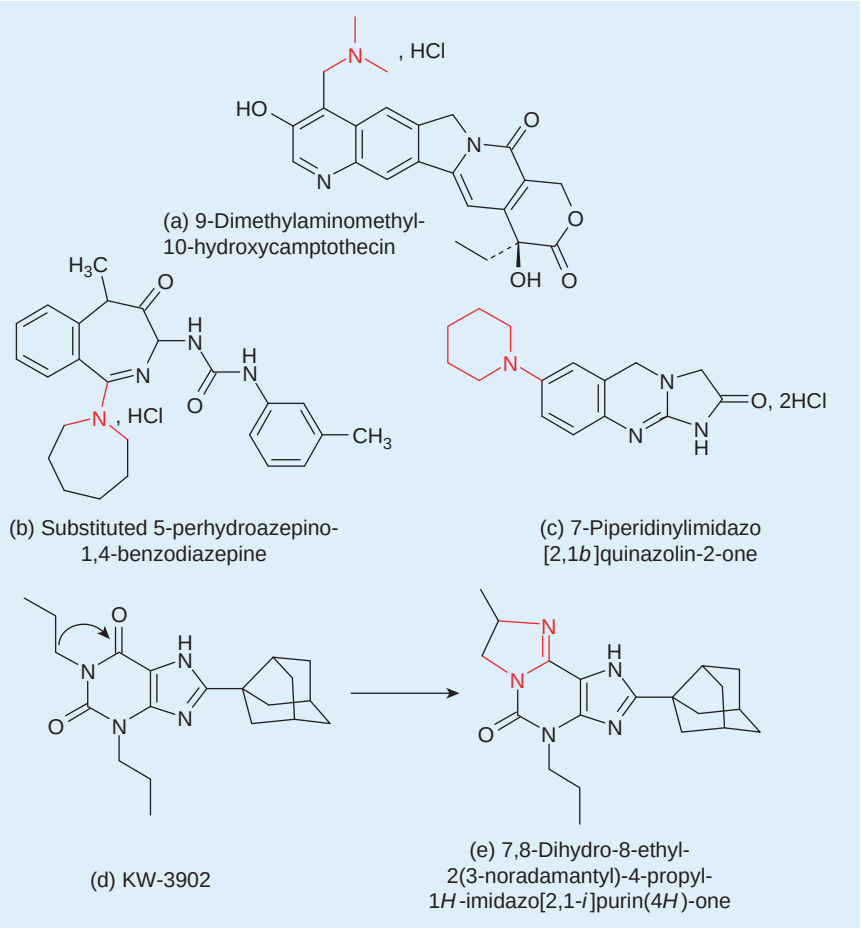


FIGURE 30.16 Solubilization by means of basic side chains.

B. Bioisosteric Exchange of a Basic Functionality

The K channel opener **8a** (Figure 30.17) suffers from poor aqueous solubility. Enhanced solubility was achieved by appending heterocycles to the **8a** scaffold (**14b** and **15a**), but the most surprising observation was that simple amino groups (**8j**) yielded excellent *in vitro* maxi-K ion channel opening activity and enhanced brain-to-plasma partitioning compared to the appended heterocycles [78].

C. Development of a Water-soluble Prodrug of Diazepam

In the search for water-soluble derivatives of the tranquillizer and anticonvulsant diazepam, scientists from Roche UK [79] imagined an original solubilization approach, based on the attachment of a basic side chain to the ring-opened diazepam molecule (Figure 30.18).

Avizafone, the peptide between the open form of diazepam (Valium®) and L-lysine, bears two primary amino functions that can be salified (by HCl) and thus render the compound freely water soluble. *In vivo* generation of free diazepam takes only a few minutes after the injection, thanks to an enzymatic cleavage of the peptidic bond followed by a spontaneous cyclization at the pH 7.4 of the bloodstream. The peptide with D-lysine is not a substrate for the esterases (Box 30.1).

D. Attachment of the Solubilizing Moiety to an Alcoholic Hydroxyl

Esterification of an alcoholic hydroxyl with dialkylglycine or its analogs is a very popular mode of solubilization of alcohols. It is illustrated (Figure 30.19) by soluble esters of forskolin [80], of allopurinol [81], and of metronidazole [82]. This mode of solubilization can also be applied to phenols such as paracetamol (see Chapter 43).

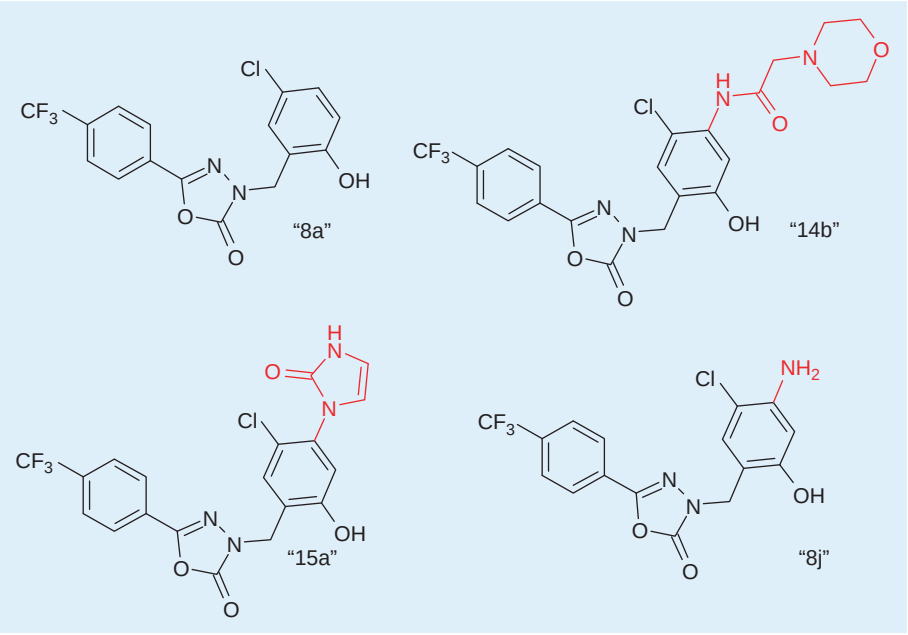


FIGURE 30.17 Solubility and bioavailability are enhanced by basic substituents on an aromatic ring.

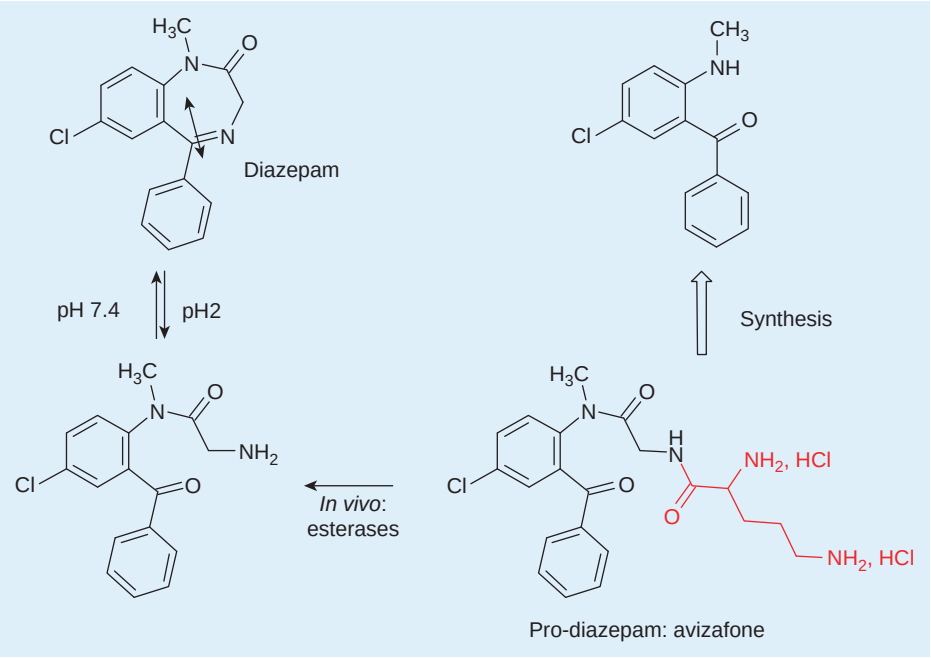


FIGURE 30.18 Avizafone, the peptide between diazepam and L-lysine, is an original water-soluble prodrug.

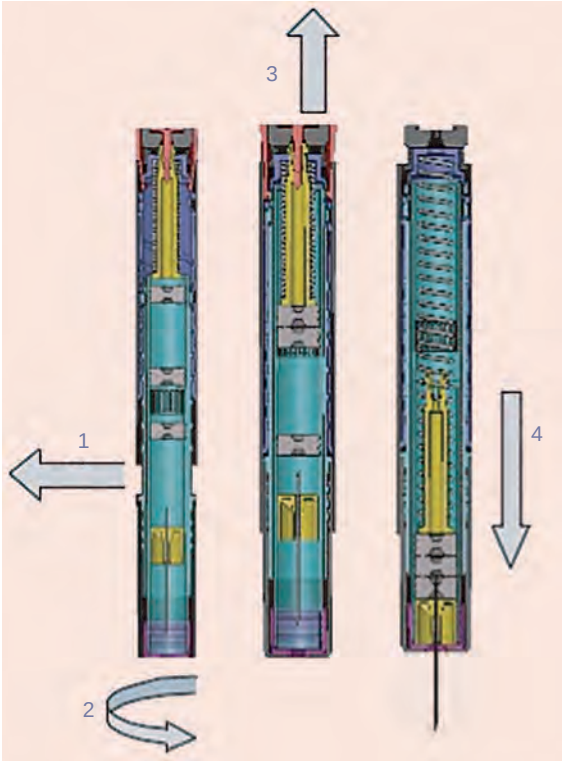
Many α-amino acid esters or related short-chained aliphatic amino esters show satisfying hydrolysis kinetics in plasma but exhibit poor stability in aqueous solution [82]. This poor stability is predominantly due to electron withdrawal by the positively charged amino group but may also involve intramolecular catalysis or assistance of the ester hydrolysis by the neighboring amino group. The replacement of the glycine unit by its benzologue, as shown for the metronidazole derivative of Figure 30.19, prevents the hydrolysis-facilitating effect of the amino group. Alternative solutions place the amino group more distant from the ester linkage in using 6-aminocaproic acid esters and sebacic acid-derived spacer groups [83] or render the ester function more resistant by replacing it by a carbonate or a carbamate function [84].

BOX 30.1

AVIZAFONE, AN ESSENTIAL COMPONENT OF THE NERVE AGENT ANTIDOTE SYSTEMS

Nerve agents such as tabun, sarin, and soman, are a class of organofluorophosphates that act in blocking the enzyme acetylcholinesterase. As a result, they produce an excessive accumulation of acetylcholine in the cholinergic synapses. Poisoning by a nerve agent leads to contraction of pupils, profuse salivation, convulsions, involuntary urination and defecation, and eventual death by asphyxiation as control is lost over respiratory muscles. Some nerve agents are readily vaporized or aerosolized, and the primary portal of entry into the body is the respiratory system. Nerve agents can also be absorbed through the skin, requiring that those likely to be subjected to such agents wear a full-body suit in addition to a respirator.

It has been known for many years that benzodiazepine compounds effectively antagonize seizures induced by organofluorophosphates. In the event of nerve gas poisoning, a combination of three drugs is commonly used: an anticholinergic drug (e.g., atropine), an acetylcholinesterase reactivator (e.g., pralidoxime 5 2-pyridine aldoxime), and an anticonvulsant (i.e., benzodiazepine). Atropine and most of the oximes are freely soluble in water, whereas many benzodiazepines are not. However, the diazepam prodrug avizafone is freely water-soluble and chemically stable. Thus, a typical self-administrable emergency treatment containing 2 mg of atropine, 350 mg of pralidoxime, and 20 mg of avizafone could be developed. The French Army Auto Injector, made by the Sedat Company, is displayed below.



Nerve agent antidote auto injector

INSTRUCTIONS

1 - Remove the red security collar
2 - Screw the cap in order to insure a perfect mix
3 - Remove the upper red security
4 - Press the injector against the arm or the leg, the jnjection will proceed automatically

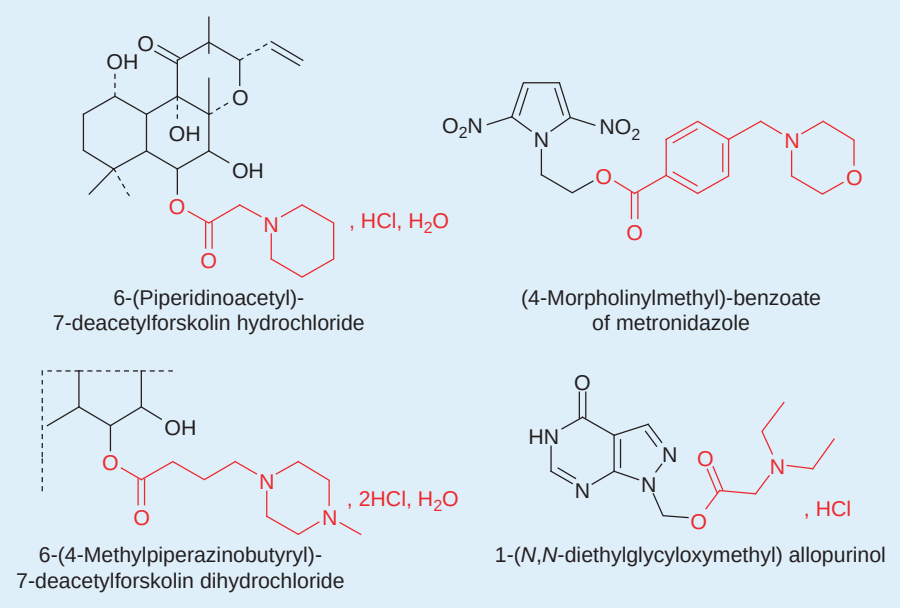


FIGURE 30.19 Dialkylglycines and related aminoesters.

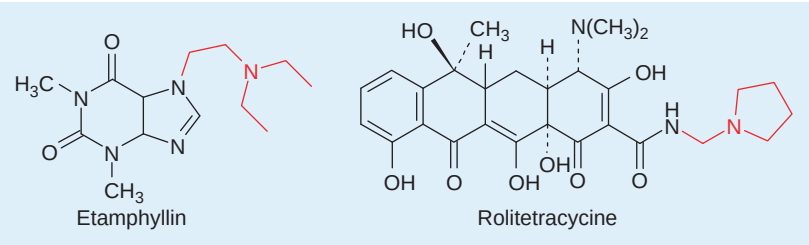


FIGURE 30.20 Basic chains on acidic NH functions.

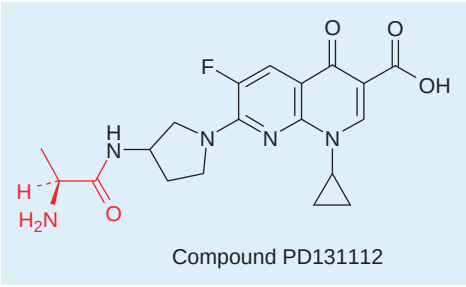


FIGURE 30.21 *S*-alanine-derived peptide of a quinolone antibacterial agent confers high water solubility and is sensitive to enzymatic cleavage *in vivo* [85].

E. Attachment of the Solubilizing Moiety to an Acidic NH Function

The NH group in theophylline is reactive and can easily be alkylated, yielding—among others—etamiphyllin (Figure 30.20), which is rendered water soluble as hydrochloride or as camphosulfonate [86].

Rolitetracycline, the Mannich base derived from the carboxamide function of tetracycline, formaldehyde, and pyrrolidine [87], is surprisingly stable and is used as an injectable form of tetracycline.

F. Attachment of the Solubilizing Moiety to a Basic NH_2 Function

An appropriate manner of solubilizing basic NH_2 functions is to form peptides with common amino acids, as shown above for the L-lysine peptide of the ring-opened form of diazepam. Similarly acylation of the 3-amino group of the pyrrolidine ring in a series of quinolone antibacterial agents yielded interesting compounds (Figure 30.21).

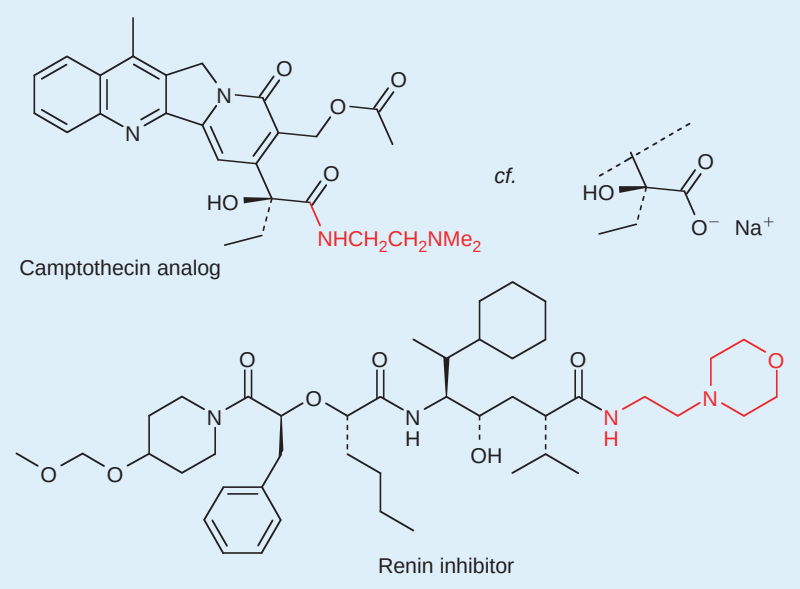


FIGURE 30.22 Basic amides of carboxylic acids.

The amino acid analogs were less active *in vitro* but had equal or increased efficacy *in vivo*. Indeed, it was shown that these compounds, which were stable to acid and base under the reaction conditions for their preparation, were rapidly cleaved in serum to give the parent quinolones. The amino acid derivatives showed a 3–70-times improved solubility when compared to the parent compounds [85].

G. Attachment of the Solubilizing Moiety to Carboxylic Acid Functionalities

As carboxylic esters of aminoalcohols are usually too sensitive to hydrolysis, amides with aminoalkylamines are preferred (Figure 30.22). The water-soluble E-lactone ring-modified 7-ethylcamptothecin analog bearing a dimethylaminoethyl amidic chain compares favorably with the sodium salt resulting simply from the lactone ring opening [88]. Introduction of basic substituents into modified hydroxyethylene dipeptide isosteres gave inhibitors with improved solubility as well as improved potency against human plasma renin [89].

V. NONIONIZABLE SIDE CHAINS

The most frequently employed solubilization approaches using nonionizable moieties involve hydroxylated and polyoxymethylenic side chains or glucosides and their analogs.

A. Glycolyl and Glyceryl Side Chains

These chains are present in some classical drugs, such as the muscle relaxant mephenesin [90] and the bronchodilator dyphylline (diprophylline; Figure 30.23) [91]. More recent applications are found in the venotropic troxerutin [92] and in the analgesic-anti-inflammatory drugs glafenin [29,93] and etofenamate [94].

B. Polyethylene Glycol Derivatives

Only a few examples of solubilization involving esters or ethers of poly(ethylene glycol) are found in the literature, and it is not always clear whether the main purpose of the polyoxyethylenic chain grafting was to increase the aqueous solubility or to produce another improvement such as sustained release. One of the most representative examples is the antitussive benzonatate (Figure 30.24) [95]. In a similar way, Nagakawa et al prepared soluble forms of vitamins A and E, as well as various steroids (prednisolone, testosterone, hydrocortisone, gitoxine) [96,97].

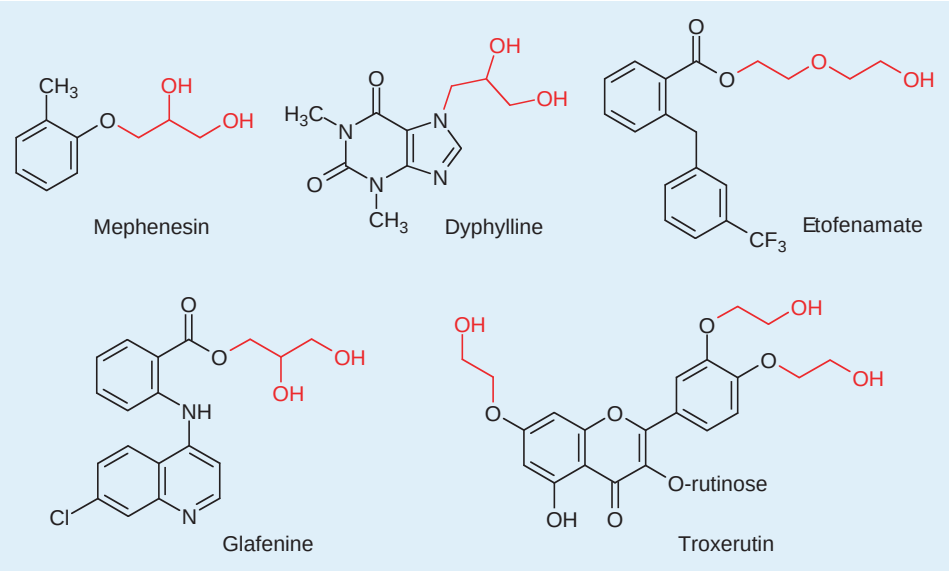


FIGURE 30.23 Glycolyl and glyceryl side chains.

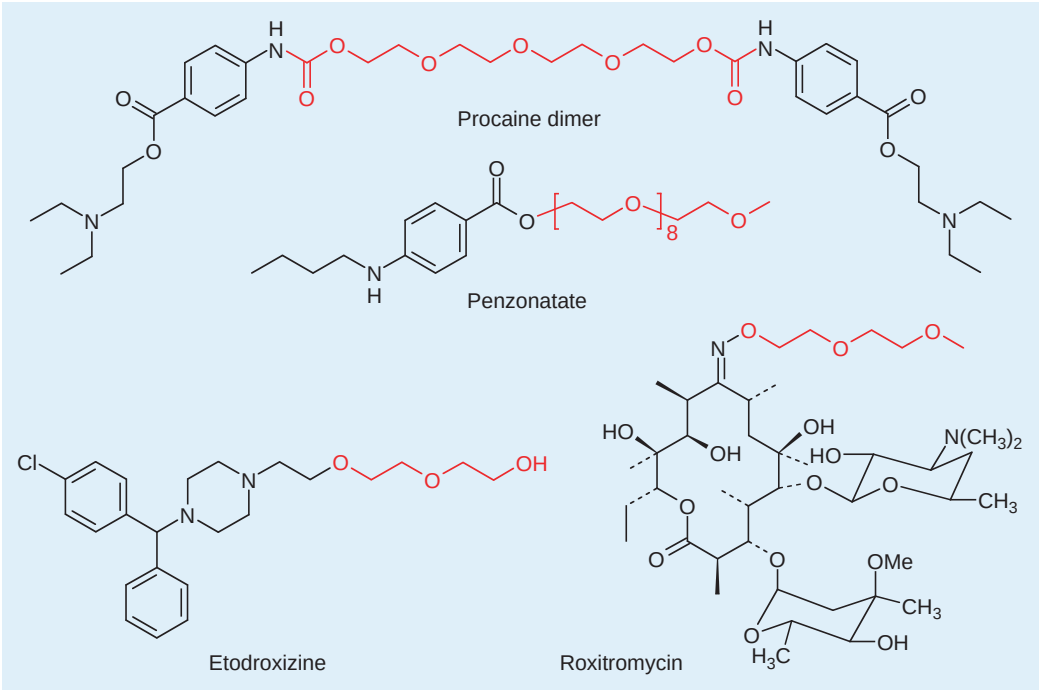


FIGURE 30.24 Polyethylene glycol derivatives.

The symmetrical attachment of the local anesthetic procaine to poly(ethylene glycol) increases the duration of action [98] and probably improves water solubility. The hypnotic etodroxizine bears a three-ethylene-oxy-unit chain but is nevertheless used as a dimaleate [99]. In roxithromycin, the oxygenated side chain is attached to the oxygen atom of the oxime of the antibiotic erythromycin [100].

C. Glucosides and Related Compounds

Despite their ubiquitous distribution in plants, human-made glucosidic derivatives of alcohols or phenols are rarely prepared in medicinal chemistry. A classic example is the sedative-hypnotic α-chloralose (Figure 30.25), which is presently used only as a surgical anesthetic for laboratory animals.

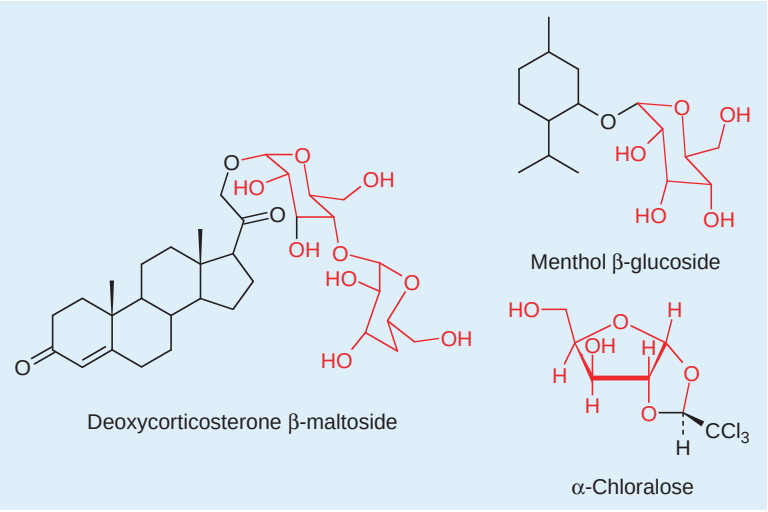


FIGURE 30.25 Glycosidic derivatives of alcoholic functions.

TABLE 30.4 Solubilities of Deoxycorticosterone Glycosides [98]

Deoxycorticosterone glycoside	Solubility in water
Glucoside	1.2‰
Galactoside	2.2‰
Lactoside	3.4‰
Lactosidoglucoside	Unlimited

Other examples are deoxycorticosterone β-maltoside [101,102] and menthol β-glucoside. Deoxycorticosterone glycosides show various solubilities depending on the sugar conjugate (Table 30.4) [103].

Menthol β-glucoside is a water-soluble, non-irritating prodrug of menthol that can be used—like glucovanillin, the β-D-glucoside of vanillin—as a pharmaceutic flavor adjuvant [104]. The use of sugar moieties as drug carriers has been reviewed by Chavis and Imbach [105].

Attachment of the sugar moiety to nitrogen atoms of amine, amide, or hydrazine functions is much more frequently encountered (Figure 30.26). Prontoglucal is the N^4-β-D-glucoside of sulfanylamide [106], the tuberculostatic glyconiazide is the isonicotinoylhydrazone of D-glucuronic acid lactone [107], and glucometacin results from the amidification of indomethacin with D-glucosamine [108]. Many highly water-soluble radiological contrast agents are solubilized as sugar conjugates. This is the case for metrizamide [109] and compound P297 [110].

VI. CONCLUDING REMARKS

The different chemical solutions to solubilizing problems discussed in this chapter reveal that in many cases the chemical transformation used also improves the activity profile of the parent molecule. This can be due to purely pharmacokinetic factors, such as a better resorption from the organism and faster transport and diffusion. These factors explain also why solubilized drugs are generally eliminated faster and therefore show fewer symptoms of toxicity. On the other hand, the pharmacological profile also can be affected. Chlorpromazine, for example (Figure 30.27), has neuroleptic properties, whereas the parent phenothiazine possesses anthelminthic properties. In this example, the attachment of the basic moiety has totally modified the pharmacological profile. However, the replacement of the basic moiety by its carboxylic counterpart yielded a compound totally inactive as neuroleptic (Wermuth, unpublished result). Conversely, in the tricyclic antidepressant series, the passage from the basic imipramine to the acidic amineptine conserved the antidepressant properties. In other words, there are no general rules available for the selection of the most appropriate solubilizing moiety. It is therefore recommended that for each new solubilization problem, acidic, basic, and neutral solubilized versions of the parent molecule be prepared.

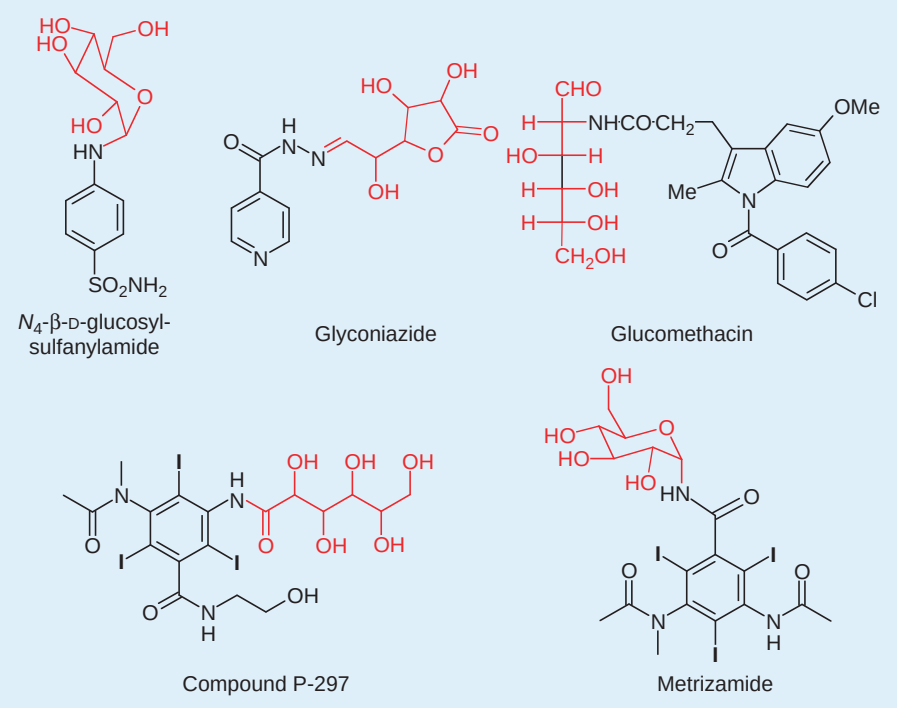


FIGURE 30.26 Glycosidic derivatives of amine, amide, and hydrazine functions.

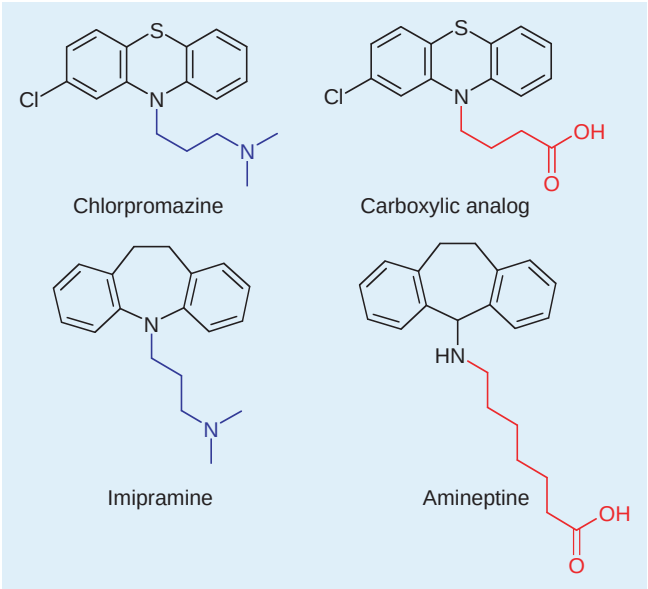


FIGURE 30.27 Acidic and basic side chains on tricyclic skeletons.

Finally, adding solubilizing moieties to active principles can also confer undesirable liabilities, such as the ones discussed earlier and those linked to basic or acidic moieties, or just unspecific toxicities or promiscuities linked to the increase of molecular weight and lipophilicity [111]. In many cases, rather than adding weight to a structure, it will be more appropriate to modulate the solubility by subtle scaffolds modifications [12].

References

[1] Li M, Wu RS, Tsai JSC. Water-soluble derivatives of lipophilic drugs. US 7,029,918 B2 (Patent); 2006.
[2] Marini-Bettolo GB. Metodos modernos de solubilizacion de medicamentos organicos [Modern methods of organic solubilization of drugs]. Ann Asoc Quim Pharm Urugay 1948;50:3–17 [in Spanish].

[3] Büchi J. Grundlagen der Arzneimittelforschung und der Synthetischen Arzneimittel [Fundamentals of drug discovery and synthetic drugs]. Basel: Birkhaüser Verlag; 1963. p. 220–35 [in German].
[4] Stella V. Prodrugs: an overview and definition. In: Higuchi T, Stella V, editors. Prodrugs as novel drug delivery systems, vol. 14. Washington, DC: American Chemical Society; 1975. p. 1–115.
[5] Silverman RB. The organic chemistry of drug design and drug action. San Diego, CA: Academic Press; 1992. p. 358–60.
[6] Ikeda Y, Ban J, Ishikawa T, Hashiguchi S, Urayama S, Horibe H. Stability and stabilization studies of TAK-599 (Ceftaroline Fosamil), a novel N-phosphono type prodrug of anti-methicillin resistant staphylococcus aureus cephalosporin T-91825. Chem Pharm Bull 2008;56 (10):1406–11.
[7] Navari RM. Fosaprepitant (MK-0517): a neurokinin-1 receptor antagonist for the prevention of chemotherapy-induced nausea and vomiting. Expert Opin Investig Drugs 2007;16(12):1977–85.
[8] Boschelli DH, Wang YD, Johnson S, Wu B, Ye F, Sosa ACB, et al. 7-Alkoxy-4-phenylamino-3-quinolinecarbonitriles as dual inhibitors of Src and Abl kinases. J Med Chem 2004;47(7):1599–601.
[9] Charifson PS, Walters WP. Acidic and basic drugs in medicinal chemistry: a perspective. J Med Chem 2014;57(23):9701–17. Available from: <http://dx.doi.org/10.1021/jm501000a>.
[10] Albert A. Selective toxicity. London: Chapman and Hall; 1979. p. 403–40.
[11] Mahar Dohan KM, Humphreys JE, Webster LO, Wring SA, Shamine LJ, Serabijt-Singh CJ, et al. Passive permeability and P-glycoprotein-mediated efflux differentiate central nervous system (CNS) and non-CNS marketed drugs. J Pharmacol Exp Ther 2001;303:1029–37.
[12] Regan SL, Maggs JL, Hammond TG, Lambert C, Williams DP, Park BK. Acyl glucuronides: the good, the bad, and the ugly. Biopharm Drug Dispos 2010;31(7):367–95.
[13] Petitpas I, Gruene T, Bhattacharya AA, Curry S. Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. J Mol Biol 2001;314(5):955–60.
[14] Sawa M. Strategies for the design of selective protein kinase inhibitors. Mini-Reviews in Medicinal Chemistry 2008;8:1291–7.
[15] Zhang L, Zhu H, Mathiowetz A, Gao H. Deep understanding of structure–solubility relationship for a diverse set of organic compounds using matched molecular pairs. Bioorganic and Medicinal Chemistry 2011;19:5763–70.
[16] Hammond ML. Recent advances in anti-infective agents. In: Bristol JA, editor. Annu Rep Med Chem, vol. 28. San Diego, CA: Academic Press; 1993. p. 119–30.
[17] Leeson PD, Carling RW, Kulagowski JJ, Mawer IM, Moore KW, Moseley AM, et al. Drugs interacting with the glycine binding site. In: Testa B, Kyburz E, Fuhrer W, Giger R, editors. Perspectives in medicinal chemistry. Basel: Verlag Helvetica Chemica Acta/VHC; 1993. p. 239–57.
[18] Stilts CE, Nelen MI, Hilmey DG, Davies S, Gollnick SO, Oseroff AR, et al. Water-soluble, core-modified porphyrins as novel, longer-wavelength-absorbing sensitizers for photodynamic therapy. J Med Chem 2000;43:2403–10.
[19] Bock MG, DiPardo RM, Mellin EC, Newton RC, Veber DF, Freedman SB. Second-generation benzodiazepine CCK-B antagonists: development of subnanomolar analogs with selectivity and water solubility. J Med Chem 1994;37:722–4.
[20] Kiesman WF, Zhao J, Conlon PR, Dowling JE, Petter RC, Lutterodt F, et al. Potent and orally bioavailable 8-bicyclo[2.2.2]octylxanthines as adenosine A1 receptor antagonists. J Med Chem 2006;49:7119–31.
[21] Lin AJ, Klayman DL, Milhous W. Antimalarial activity of new water-soluble dihydroartemisin derivatives. J Med Chem 1987;30:2147–50.
[22] Lin AJ, Sikry AB, Kyle DE. Antimalarial activity of new dihydroartemisin derivatives. 7. 4-(p-substituted phenyl)-4(R or S)-[10 (α or β)-dihydroartemisinoxy]butyric acids. J Med Chem 1997;40:1396–400.
[23] Mesanguy M. Procédé de préparation du flavone 5,7-dioxyacétate de sodium [Process for preparing the flavone 5,7-dioxyacetate of sodium]. 2.076.753 (Patent); 1970 [in French].
[24] Jackson EL. Certain N-alkyl, N-carboxyalkyl, and N-hydroxyalkyl derivatives of 4,49-diaminodiphenyl sulfone. J Am Chem Soc 1948;70:680–4.
[25] Reitmann J. Aliphatic amine salts of halogenated pyridones containing an acid group. US 1,993,039 (Patent); 1935.
[26] Bauer H. Organic compounds in chemotherapy. II. The preparation of formaldehyde sulfoxylate derivatives of sulfanilamide and of amino compounds. J Am Chem Soc 1939;61:617–18.
[27] Lin AJ, Lee M, Klayman DL. Antimalarial activity of new water-soluble dihydroartemisin derivatives. 2. Stereospecificity of the ether side chain. J Med Chem 1989;32:1249–52.
[28] Coker JD, Elks J, May PJ, Nice FA, Phillips GH, Wall WF. Action of some steroids on the central nervous system of the mouse. I. Synthetic methods. J Med Chem 1965;8:417–25.
[29] Allais A, Rousseau G, Girault P, Mathieu J, Peterfalvi M, Branceni D, et al. Sur l'activité analgésique et antiinflammatoire des 4-(29-alcoxycarbonyl phénylamino) quinoléïnes [The anti-inflammatory and analgesic activity of 4-(29-alkoxycarbonyl phenylamino) quinoleines]. Chimie Therapeutique 1966;1:65–70 [in French].
[30] Sandman B, Szulczewski D, Winheuser J, Higuchi T. Rearrangement of chloramphenicol-3-monosuccinate. J Pharm Sci 1970;59:427–9.
[31] Trissel LA. Handbook on injectable drugs. 14th ed. Bethesda, MD: American Society of Hospital Pharmacists; 2007.
[32] Anderson BD, Taphouse V. Initial rate studies of hydrolysis and acyl migration in methylprednisolone 21-hemisuccinate and 17-hemisuccinate. J Pharm Sci 1981;70:181–6.
[33] Anderson BD, Conradi RA, Lambert JW. Carboxyl group catalysis of acyl transfer reactions in corticosteroid 17- and 21-monoesters. J Pharm Sci 1984;73:604–11.
[34] Parthasarathy S. Evidence for an additional intracellular site of action of probucol in the prevention of oxidative modification of low density lipoprotein: use of a new water-soluble probucol derivative. J Clin Invest 1992;89:1618–21.
[35] Anderson BD, Conradi RA, Spilman CH, Forbes AD. Strategies in the design of solution-stable, water-soluble prodrugs III: influence of the pro-moiety on the bioconversion of 21-esters of corticosteroids. J Pharm Sci 1985;74:382–7.
[36] Chong K-T, Ruwart MJ, Hinshaw RR, Wilkinson KF, Rush BD, Yancey MF, et al. Peptidomimetic HIV protease inhibitors: phosphate prodrugs with improved biological activities. J Med Chem 1993;36:2575–7.
[37] Khawaja TA, Reese CB. o-Phenylene phosphochloridate. A convenient phosphorylating agent. J Am Chem Soc 1966;88:3446–7.
[38] Khawaja TA, Reese CB, Stewart JCM. A Convenient general procedure for the conversion of alcohols into their monophosphate esters. J Chem Soc (C) 1970;2090–100.

[39] Taguchi Y, Mushika Y. 2-(N,N-diethylamino)-4-nitrophenyl phosphate and its use in the selective phosphorylation of unprotected nucleosides. Tetrahedron Lett 1975;24:1913–16.
[40] Gajda T, Zwierzak A. Phase-transfer-catalysed halogenation of di-t-butyl phosphite: preparation of di-t-butyl phosphorohalidates. Synthesis 1976;243–4.
[41] Gajda T, Zwierzak A. Di-t-butyl phosphorobromidate: a new selective phosphorylating agent containing acid-labile protecting groups. Synthesis 1977;623–5.
[42] Tener GM. 2-Cyanoethyl phosphate and its use in the synthesis of phosphate esters. J Am Chem Soc 1961;83:159–68.
[43] Moffatt JG. The synthesis of orotidine-59 phosphate. J Am Chem Soc 1963;85:1118–23.
[44] Brownfield RB, Shultz WA. Direct method for the preparation of steroid-21-phosphates. Steroids 1963;2:597–603.
[45] Flynn GL, Lamb DJ. Factors influencing solvolysis of corticosteroid-21-phosphate esters. J Pharm Sci 1970;59:1433–8.
[46] Melby JC, St Cyr M. Comparative studies on absorption and metabolic disposal of water-soluble corticosteroid esters. Metabolism 1961;10:75–82.
[47] Cutolo E, Larizza A. Synthesis of 3-phosphoric ester of l-ascorbic acid. Gazz Chim Ital 1961;91:964–72.
[48] Wenz A, Göttmann G, Koop H. Verfahren zur Herstellung der freien Base und der Salze des Aneurin-orthophosphor-säureesters [Process for the preparation of the free base and salts of aneurin-ortho-phosphorous acid ester]. 1 110 646 (Patent); 1961 [in German].
[49] Ito A, Hamanaka W, Takagi H, Wada T, Kawada T. Verfahren zur Herstellung eines Acylierungsproduktes von Vitamin-B_1-orthophosphorsäureester und Salzen Davon [Process for the preparation of an acylation of vitamin-B1-ortho phosphoric acid esters and salts thereof]. 1,130,811 (Patent); 1960 [in German].
[50] Wada T, Tagaki H, Minakami H, Hamanaka W, Okamoto K, Ito A, et al. A new thiamine derivative, S-benzoylthiamine O-monophosphate. Science 1961;134:195–6.
[51] Viscontini M, Ebnoether C, Karrer P. Einfache Synthese kristallisierter Lactoflavin-59-phosphorsäure (Coferment des Flavinenzyms) [Simple synthesis of crystallized lactoflavin-59-phosphoric acid (coferment of flavinenzyms)]. Helv Chim Acta 1952;35:457–9 [in German].
[52] Fieser LF. Antihemorrhagic esters and methods for producing the same. US 2,407,823 (Patent); 1946.
[53] Tattersall FD, Rycroft W, Cumberbatch M, Mason G, Tye S, Williamson DJ, et al. The novel NK1 receptor antagonist MK-0869 (L-754,030) and its water soluble phosphoryl prodrug, L-758,298, inhibit acute and delayed cisplatin-induced emesis in ferrets. Neuropharmacology 2000;39(4):652–63.
[54] Hale JJ, Mills SG, MacCoss M, Dorn CP, Finke PE, Budhu RJ, et al. Phosphorylated morpholine acetal human neurokinin-1 receptor antagonists as water-soluble prodrugs. J Med Chem 2000;43:1234–41.
[55] Fex H, Lundvall KE, Olsson A. Hydrogen sulfates of natural estrogens. Acta Chem Scand 1968;22:254–64.
[56] Fieser LF, Fry EM. Water-soluble antihemorrhagic esters. J Am Chem Soc 1940;62:228–9.
[57] Stieglitz E, Matz M. Verfahren zur Herstellung von Xanthinalky-Schwefelsäuren oder deren Salze [A process for preparing xanthine alkyl-sulfonic acids or salts thereof]. 1,090,669 (Patent); 1958 [in German].
[58] Miyabo S, Nakamura T, Kuwazima S, Kishida S. A comparison of the bioavailability and potency of dexamethazone phosphate and sulphate in man. Eur J Clin Pharmacol 1981;20:277–82.
[59] Williams DB, Varia SA, Stella VJ, Pitman IH. Evaluation of the prodrug potential of the sulfate ester of acetaminophen and 3-hydroxymethyl-phenytoin. Int J Pharm 1983;14:113–20.
[60] Jacobson KA, Nikodijevic O, Ji X-D, Berkich DA, Eveleth D, Dean RL, et al. Synthesis and biological activity of N^6-(p-sulfophenyl)alkyl and N^6-sulfoalkyl derivatives of adenosine: water-soluble and peripherally selective adenosine agonists. J Med Chem 1992;35:4143–9.
[61] Greenberg FH, Leung KK, Leung M. The reaction of vitamin K_3 with sodium bisulfite. J Chem Ed 1971;48:632–4.
[62] Despois RL. Procédé de préparation de composés aminés aromatiques solubles possédant une valeur thérapeutique [Process for preparing soluble aromatic amine compounds having therapeutic value]. 831,366 (Patent); 1938 [in French].
[63] Verfahren zur Darstellung von p-äthoxyphenylaminomethylschwefligsauren Salzen [A process for the preparation of p-ethoxyphenylaminomethyl sulfuric acid salts]. 209,695 (Patent); 1907 [in German].
[64] Verfahren zur Darstellung von ω-methylschwefligsauren Salzen aminosubstituierter Arylpyrazolone. *Ger. Patent 259,577* (October 26, 1911; to Farbwerke vorm. Meister Lucius & Brüning in Hoechst am Main); 1911 [in German].
[65] Verfahren zur Darstellung von ω-methylschwefligsauren Salzen aminosubstituierter Arylpyrazolone. Ger. Patent 254,711 (July 21, 1911) [in German].
[66] Mutch N. A new sulphonamide (sulfonamide E.O.S.). Br Med J 1941;2:503–7.
[67] Jain BC, Iyer BH, Guha PC. The preparation of promin. Sci Cult 1946;11:568–9.
[68] Soluble diphenylsulfone. 234,108 (Patent); 1949.
[69] Bockmühl M, Krohs W, Racke F, Windisch K. N-methylsulphites and N-methanesulphinic acid salts of 1-aryl-2,3-dialkyl-4-alkylaminopyrazolones. US 2,193,788 (Patent); 1940.
[70] Fischer LJ, Riegelman S. Absorption and activity of some derivatives of griseofulvin. J Pharm Sci 1967;56:469–76.
[71] Tatsumi K, Ishigai M. Oxime-metabolizing activity of liver aldehyde oxidase. Arch Biochem Biophys 1987;253:413–18.
[72] Evans SJ, Moore FL, Murray FTM. Solubilization and pharmacological characterization of a glucocorticoid membrane receptor from an amphibian brain. J Steroid Biochem Molec Biol 1998;67:1–8.
[73] Aichinger G, Behner O, Hoffmeister F, Schütz S. Basische tricyclische Oximinoäther und ihre pharmakologischen Eigenschaften [Basic tricyclic oximino ethers and their pharmacological properties]. Arzneim-Forsch 1969;19:838–45 [in German].
[74] Kingsbury WD, Boehm JC, Jakas DR, Holden KG, Hecht SM, Gallagher G, et al. Synthesis of water-soluble (aminoalkyl)camptothecin analogues: inhibition of topoisomerase I and antitumor activity. J Med Chem 1991;34:98–107.
[75] Showell GA, Bourrain S, Neduvelil JG, Fletcher SR, Freedman SB, Kemp JA, et al. High-affinity and potent, water-soluble 5-amino-1,4-benzodiazepine CCKB/gastrin receptor antagonists containing a cationic solubilizing group. J Med Chem 1994;37:719–21.
[76] Ishikawa F, Saegusa J, Inamura K, Sakuma K, Ashida S-I. Cyclic guanidines. 17. Novel (N-substituted amino)imidazo[2,1-b] quinazolin-2-ones: water-soluble platelet aggregation inhibitors. J Med Chem 1985;28:1387–93.

[77] Suzuki F, Shimada J, Nonaka H, Ishii A, Shiozaki S, Ichikawa S, et al. 7,8-Dihydro-8-ethyl-2-(3-noradamantyl)-4-propyl-1H-imidazo[2,1-i]purin-5(4H)-one: a potent and water-soluble adenosine A1 antagonist. J Med Chem 1992;35:3572–81.
[78] Romine JL, Martin SW, Meanwell SA, Gribkoff VK, Boissard CG, Dworetzky SI, et al. 3-[(5-Chloro-2-hydroxyphenyl)methyl]-5-[4-(trifluoromethyl)phenyl]-1,3,4-oxadiazol-2(3H)-one, BMS-191011: opener of large-conductance Ca^{21}-activated potassium (maxi-K) channels, identification, solubility, and SAR. J Med Chem 2007;50:528–42.
[79] Hassall CH, Holmes SW, Johnson WH, Kröhn A, Smithen CE, Thomas WA. Peptido-aminobenzophenones-novel latentiated benzo-1,4-diazepines. Experientia 1977;33:1492–3.
[80] Khandelwal Y, Rajeshwari K, Rajagopalan R, Swamy L, Dohadwalla AN, de Souza NJ, et al. Cardiovascular effects of new water-soluble derivatives of forskolin. J Med Chem 1988;31:1872–9.
[81] Bundgaard H, Falch E. Improved rectal and parenteral delivery of allopurinol using the prodrug approach. Arch Pharm Chem Sci Ed 1985;13:39–48.
[82] Bundgaard H, Falch E, Jensen E. A novel solution-stable, water-soluble prodrug type for drugs containing a hydroxyl or an NH-acidic group. J Med Chem 1989;32:2503–7.
[83] Anderson BD, Conradi RA, Knuth KE. Strategies in the design of solution-stable, water-soluble prodrugs I: a physical-organic approach to pro-moiety selection for 21-esters of corticosteroids. J Pharm Sci 1985;74:365–74.
[84] Balkovec JM, Black RM, Hammond ML, Heck JV, Zambias RA, Abruzzo G, et al. Synthesis, stability, and biological evaluation of water-soluble prodrugs of a new echinocandin lipopeptide: discovery of a potential clinical agent for the treatment of systemic candidiasis and pneumocystis carinii pneumonia (PCP). J Med Chem 1992;35:194–8.
[85] Sanchez JP, Domagala JM, Heifetz CL, Priebe SR, Sesnie JA, Trehan AK. Quinolone antibacterial agents: synthesis and structure–activity relationships of a series of amino acid prodrugs of racemic and chiral 7-(3-amino-1-pyrrolidinyl)quinolones. Highly soluble quinolone prodrugs with in vivo pseudomonas activity. J Med Chem 1992;35:1764–73.
[86] Klosa J. Beitrag zur Reaktionsfähigheit der 7-Stellung des Theophyllins. 2. Mitt. Über Synthesen in der Theophyllinreihe [Contribution to response awareness of the 7-position of theophylline. 2nd communication. About syntheses in the theophylline series]. Arch Pharm 1955;288:301–3 [in German].
[87] Gottstein WJ, Minor WF, Cheney LC. Carboxamido derivatives of tetracyclines. J Am Chem Soc 1959;81:1198–201.
[88] Sawada S, Yaegashi T, Furuta T, Yokokura T, Miyasaka T. Chemical modification of an antitumor alkaloid, 20(S)-camptothecin: E-lactone ring modified water-soluble derivatives of 7-ethylcamptothecin. Chem Pharm Bull 1993;41:310–13.
[89] Boyd SA, Fung AKL, Baker WR, Mantei RA, Armiger Y-L, Stein HH, et al. C-terminal modifications of nonpeptide renin inhibitors: improved oral bioavailability via modification of physicochemical properties. J Med Chem 1992;35:1735–46.
[90] Morch P. Glykresin. Arch Pharm Chem 1947;54:327–32.
[91] Roth HJ. Zur Darstellung von β-Oxyalkyl-dimethyl-purinen; Umsetzung von Theophyllin und Theobromin mit 1,2-Epoxyden. Arch Pharm 1959;292:234–8 [in German].
[92] Courbat PJ. Qercetin and Quercetin Gycoside. US 3,420,815 (Patent); 1969.
[93] Mouzin G, Cousse H, Autin JM. A new, convenient synthesis of glafenine and floctafenine. Synthesis 1980;54–5.
[94] Boltze KH, Kreisfeld H. Zur Chemie von Etofenamat, einem Antiphlogistikum aus der Klasse der N-Arylanthranylsäurederivate [The chemistry of etofenamate, an anti-inflammatory drug from the class of N- aryl anthranilic derivatives]. Arzneimit Forsch Drug Des 1977;27:1300–12 [in German].
[95] Matter M. Polyethoxy ethers of isocyclic organic carboxylic acids. US 2,714,608 (Patent); 1955.
[96] Nagakawa T, Muneyuki T, Mori Y. Water solubilization by means of polyethylene glycol derivatives. 7455 CA (1961) 55, 5879b (Patent); 1961.
[97] Nagakawa T, Muneyuki T, Mori Y. Water solubilization of drugs. 17006 (Patent); 1960.
[98] Weiner B-Z, Zilkha A. Polyethylene glycol derivatives of procaine. J Med Chem 1973;16:573–4.
[99] Morren H. New derivatives of N-mono-benzhydril-piperazine and process for the preparation thereof. 817,231 (Patent); 1959.
[100] Gouin d'Ambrières S, Lutz A, Gasc JC. Novel erythromycin a derivatives. US 4,349,545 (Patent); 1982.
[101] Miescher K, Fischer WH, Meystre C. 6. Über Steroide. (33. Mitteilung). Über Glucoside des Desoxy-corticosterons [About steroids. 33rd communication. About glucosides of deoxycorticosterone]. Helv Chim Acta 1942;25:40–2 [in German].
[102] Miescher K, Meystre C. 26. Über Steroide. 34. Mitteilung. Über Sacharide des Desoxy-corticosterons II [About steroids. 34th communication. About saccharides of deoxycorticosterone II]. Helv Chim Acta 1943;26:224–33 [in German].
[103] Meystre C, Miescher K. Über Steroide. (35. Mitteilung). Zur Darstellung von Saccharidderivaten der Steroide [About steroids. 35th communication. For the presentation of saccharide derivatives of steroids]. Helv Chim Acta 1944;27:231–6 [in German].
[104] Higashiyama T, Sakata I. Mentholglykoside, Verfahren zur ihrer Herstellung, ihre Verwendung zur Entwicklung des Pfefferminzgeschmack sowie diese Verbindungen enthaltende Arzneimittel [Mentholglykosides: processes for their preparation, their use in the development of peppermint flavor, and pharmaceutical compositions containing them]. 2.242.237 (Patent); 1972 [in German].
[105] Chavis C, Imbach J-L. Sur une méthode de pharmacomodulation: l'utilisation du vecteur sucre [A method of pharmaco-modulation: the use of the sugar vector].. In: Combet-Farnoux C, editor. Actualités de chimie thérapeutique, Vol. 5. Châtenay-Malabry: Société de Chimie Thérapeutique; 1977. p. 3–28 [in French].
[106] Bognar R, Nanasi P. N-substituted glycosylamines derived from sulphanilamide and p-aminosalicylic acid. J Chem Soc 1953;:1703–8.
[107] Sah P. d-Glucuronolactone isonicotinyl hydrazone. J Am Chem Soc 1953;75:2512–13.
[108] Demetrio A, Ganzina F, Magi M, Serino E, Paroli E, Samueli F. Antiinflammatory di-(1)glucosamide of 1-(4-chlorobenzoyl)-2-methyl-5-methoxyindole-3-acetic acid. 2,223,051; CA 80, 83529e, (Patent); 1974.
[109] Almen THO, Haavaldsen J, Nordal V. N-(2,4,6-triiodobenzoyl)-sugar amines. US 3,701,771 (Patent); 1972.
[110] Hilae SK, Dauth GW, Hess KH, Gilman S. Development and evaluation of a new water-soluble iodinated myelographic contrast medium with markedly reduced convulsive effects. Radiology 1978;126:417–22.
[111] Hughes JD, Blagg J, Price DA, Bailey S, DeCrescenzo GA, Devraj RV, et al. Physiochemical drug properties associated with in vivo toxicological outcomes. Bioorg Med Chem Lett 2008;18:4872–5.

CHAPTER

31

Improving the Water-Solubility of Compounds by Molecular Modification to Disrupt Crystal Packing

Minoru Ishikawa and Yuichi Hashimoto

Institute of Molecular and Cellular Biosciences, The University of Tokyo, Japan

OUTLINE

I. INTRODUCTION

A. Importance of Improving Aqueous Solubility by Molecular Modification

Aqueous solubility is a key requirement for small-molecular drug candidates. Poor aqueous solubility often results in poor absorption, even if the permeation rate is high, since the flux of a drug across the intestinal membrane is proportional to the concentration gradient betwelen the intestinal lumen and the blood. In addition, risk assessment of poorly soluble compounds is challenging, because exposure may be difficult to define and the test sensitivity may be reduced. Further, high concentrations of poorly soluble drugs in organisms may result in crystallization and acute toxicity. Overall, poor solubility of drug candidates has been identified as the cause of numerous drug-development failures. Clinically acceptable solubility of drugs is determined by not only intrinsic solubility but also dose and gastric fluid volume. According to the current FDA guidance [1,2], the dose number (Do) is used as a parameter to evaluate the practical adequacy of drug solubility. It is calculated from the dose divided by an uptake volume of 250 ml and the drug's thermodynamic solubility: Do = (dose/250 mL)/ thermodynamic solubility. A candidate drug is considered highly soluble (=sufficiently soluble) when the highest dose to be administered is soluble in 250 ml of aqueous media (i.e., Do is <1).

According to the simplest definition, the thermodynamic solubility of a compound in a solvent is the maximum amount of the most stable crystalline form of the compound that can remain in solution under equilibrium conditions [3]. The most stable form is invariably the form that has the highest melting point. The importance of thermodynamic assessment is greater in late discovery/early development of candidate drugs. On the other hand, the kinetic solubility of a molecule depends on its crystal form, and crystal polymorphs of a molecule can show different kinetic solubility. Therefore, crystal modification can produce an increase in dissolution rate and a temporary or apparent increase of solubility, but it cannot produce a permanent alteration of solubility. Given sufficient time, the undissolved solute will revert to its most stable crystal form under the conditions, and the solubility will approach the true thermodynamic solubility [4]. (In this chapter, we define thermodynamic solubility as the solubility obtained after a solid compound has been agitated in the solvent for more than 24 h.) Therefore, the role of crystal modification is confined to increasing the dissolution rate (kinetic solubility) of drugs. Appropriate formulation can be helpful [5,6], but the extent of absorption and solubility enhancement that can realistically be achieved is very limited. Thus, it would be better to generate drug candidates with sufficient aqueous solubility at the drug discovery stage. In other words, it is much better to improve solubility by molecular modification.

B. Classical Modification to Improve Solubility: Decrease Hydrophobicity

In general, the aqueous solubility of small molecules depends on their hydrophobicity (LogP) [7]. Partition coefficient LogP is defined as follows: LogP = Log [(solute in *n*-octanol)/(solute in water)]. Increase of aqueous solubility leads to an increase of the denominator of the above equation and a decrease of LogP. Thus, decrease of LogP by molecular modification (i.e., introduction of hydrophilic group(s) into molecules) is a classic and general strategy for improving aqueous solubility. There are numerous examples of improvement in solubility by decreasing hydrophobicity. Lipophilic substituent constants (π), defined by Fujita et al. [7a], are useful guides in molecular design to modify the hydrophilicity of parent molecules.

This approach is not universally effective, however, because the introduced hydrophilic group(s) sometimes interferes with the target protein−drug interaction. In addition, this strategy is not effective when both solubility and hydrophobicity need to be increased, for example, to improve oral bioavailability or brain penetration of highly hydrophilic compounds with insufficient solubility. Furthermore, compounds with poor solubility in both octanol and water sometimes retain poor absolute values of aqueous solubility despite a decrease of LogP, because LogP values are just ratios. Therefore, an alternative strategy to increasing the aqueous solubility of lead compounds has been desired in drug discovery and medicinal chemistry. A prodrug approach through tethering the drug to a soluble moiety is one possibility to overcoming the problem of low aqueous solubility [8]. However, such a prodrug approach has several potential disadvantages, including manufacturing cost, species differences of prodrug cleavage, complexity of PK/PD prediction, and possible toxicity of metabolites.

II. RATIONALE FOR DISRUPTION OF CRYSTAL PACKING AS AN ALTERNATIVE METHOD TO IMPROVE SOLUBILITY

The three steps involved in dissolution of a drug are illustrated in Figure 31.1 [8]. Tight crystal packing requires a large amount of free energy to release a molecule from the crystal (step 1). A molecule with higher molecular weight requires a larger cavity in the polar solvent (step 2), and thus, for a solvent like water, more hydrogen bonds need to be broken. Finally, the free molecule occupies the cavity in the solvent (step 3) and solvation energy is released. Thus, the solubility of a solid solute in water is dependent on several factors: the crystallinity of the solute, its molecular size, and the ability of the solute to interact with water. Among the three steps, the solvation (hydration) step (step 3) is a good target to improve aqueous solubility (i.e., solubilizing modification to reduce hydrophilicity promotes hydration, as described above).

An alternative strategy for improving aqueous solubility would be the disruption of tight crystal packing of molecules (step 1). In other words, molecular modification to weaken the intermolecular interaction in the most stable crystal form of the compounds would improve thermodynamic solubility. The melting point itself is related to crystal lattice and crystal packing energies [9]. Therefore, the melting point can be utilized as a parameter of crystal packing. For molecules that do not have high crystal packing energy, such as low-melting-point solids and oils, water-solubility limitations are usually due to poor interactions with the solvent, water.

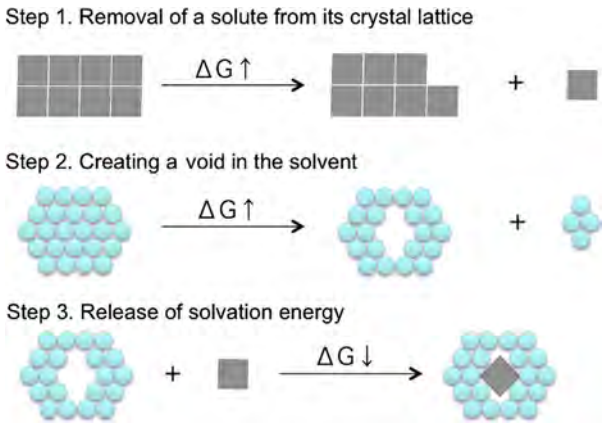
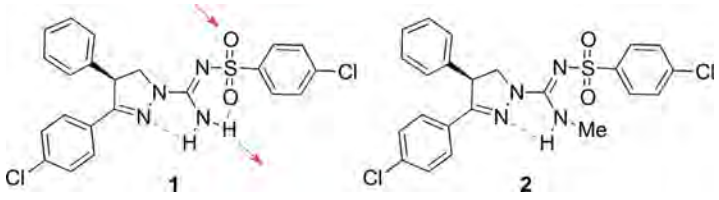


FIGURE 31.1 An illustration of the three steps needed for drug dissolution.

TABLE 31.1 Improvement in Solubility by Disruption of Intermolecular Hydrogen Bond



Compound	R	CB_1 Ki (nM)	Anti-hypertensive activity in rat EC_{50} (mg/kg)	Kinetic solubility[a]	Melting point (°C)	Crystal density (g/cm^3)	CLogP
1	H	16	>30	moderate	235	1.535	4.3
2	Me	7.8	5.5	easy	170	1.481	4.8

[a]*Solubility in EtOH at reflux.*
Hydrogen bonds are shown in red. Intermolecular hydrogen bonds are shown by arrows.

In 1980, Yalkowsky presented general solubility equations (GSE) derived on the basis of semi-empirical analysis. GSE includes not only LogP but also melting point; for example, Log[solubility (M)] = 0.5 − (LogP) − 0.01[[melting point(°C)] − 25] [10]. The equations were largely based upon examination of rigid, polycyclic, and halogenated aromatic compounds. Recently, several solubilizing modifications of complex pharmaceutical compounds to weaken crystal packing have been reported, including disruption of intermolecular hydrogen bonds, and disruption of molecular planarity and symmetry [11]. Some solubilizing modifications were aimed at both increasing hydrophobicity and disrupting intermolecular interaction simultaneously to clarify the contributions of the two mechanisms. If compounds soluble in both organic solvents and water are generated, aqueous solubility would be improved even if the LogP value is increased (i.e., the solubility ratio in octanol is increased). Examples of improvement in aqueous solubility by disruption of crystal packing are reviewed below.

III. IMPROVEMENT OF SOLUBILITY BY DISRUPTING INTERMOLECULAR HYDROGEN BONDS

Cannabinoid receptor (CB_1) antagonist [12]. CB_1 antagonist **1** showed no activity *in vivo* after oral administration, whereas **2** was active *in vivo* (Table 31.1). Compounds **1** and **2** were both very poorly soluble in water (<1 mg/mL at pH 7). However, the solubility of **1** in EtOH at reflux was moderate, whereas **2** readily dissolved under this condition. The lower aqueous solubility and dissolution rate of **1** might have resulted from a dissolution rate in the GI tract insufficient to achieve *in vivo* activity. X-ray structural analyses of **1** and **2** were

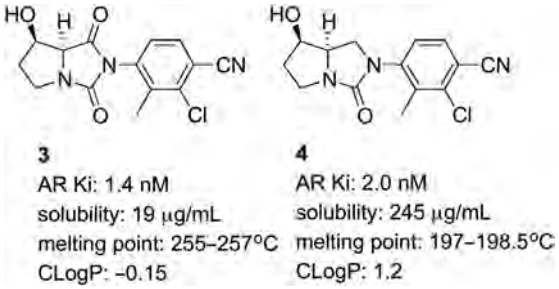
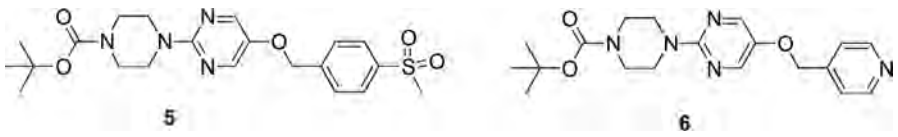


FIGURE 31.2 Improvement in aqueous solubility by disruption of intermolecular hydrogen bond.

TABLE 31.2 Improvement of Aqueous Solubility by Disruption of Intermolecular Interaction Involving Methylsulfonyl Group



Compound	Solubility (μM)	Melting point (°C)	Crystal density (g/cm^3)	log$D_{7.4}$	Human GPR119 EC_{50} (nM)
5	0.03	201–202	1.361	3.2	65
6	23	147–148	1.277	3.4	621

performed, revealing that the differences in crystal packing between **1** and **2** can be explained in terms of bridging ability. Both NH hydrogens of **1** form intramolecular hydrogen bonds with both the dihydropyrazole moiety and the sulfonyl group. These intramolecular hydrogen bonds of **1** maintain a planar orientation, and an intermolecular hydrogen bond exists between the SO_2 oxygen atom and hydrogen at the amidine moiety of another molecule of **1**. Introduction of a methyl group at the amidine moiety (**2**) disrupts the intermolecular hydrogen bond, resulting in a decrease of both crystal density and melting point.

Androgen receptor modulator [13]. Crystal structural analysis of compound **3** revealed an intermolecular hydrogen bond between the amide carbonyl of one molecule and the hydroxyl group of another molecule, resulting in tight crystal packing, which in turn leads to low aqueous solubility (Figure. 31.2). It was hypothesized that removal of the amide carbonyl group of the hydantoin moiety would eliminate the intermolecular hydrogen bond and thereby improve aqueous solubility. Removal of the amide carbonyl of the hydantoin ring (**4**) was indeed effective, and **4** was 13-fold more soluble than **3**. The melting point of **4** was also lower than that of **3**, whereas CLogP of **4** was higher than that of **3**. These results suggest that looser crystal packing is a likely reason for the improved aqueous solubility, despite the increased hydrophobicity. Several other reports have also examined the relationship between solubility and intermolecular hydrogen bonds in relation to crystal packing or melting point [14,15].

G Protein-coupled receptor 119 agonists [16]. Disruption of intermolecular interaction between oxygen of the sulfonyl group and hydrogen of a methyl group in the agonist **5** was reported. Compound **5** was highly insoluble (0.03 μM), and its melting point was high (201–202°C), although it does not contain any acidic hydrogens (e.g, OH, NH, and COOH) (Table 31.2). To identify interactions that contribute to the stability in the solid state, single-crystal X-ray structure analysis was performed. As shown in Figure 31.3a, compound **5** is rather flat, and the molecules stack very effectively. The strongest polar interactions are between oxygens of the sulfonyl groups and hydrogens of the methyl groups in other molecules (Figure. 31.3b). This head-to-head interaction aligns the molecules, generating a tight crystal lattice. A search of the Cambridge Crystallographic Database indicated that 18 percent of methyl sulfone moieties are associated with head-to-head molecular interaction, and 18 percent are involved in ladder-like interactions. A further 23 percent of methyl sulfones are involved in both types of interaction. This analysis led to replacement of the (methylsulfonyl)phenyl group with a pyridyl group (**6**). This substitution improved the solubility (23 μM), with a slight increase of log $D_{7.4}$ value. Single-crystal X-ray crystallographic analysis of **6** confirmed the absence of head-to-head interaction and the absence of the network of polar interactions associated with methyl sulfone. The resulting looser crystal packing led to decreased crystal density and melting point, and is presumably the reason for the solubility improvement.

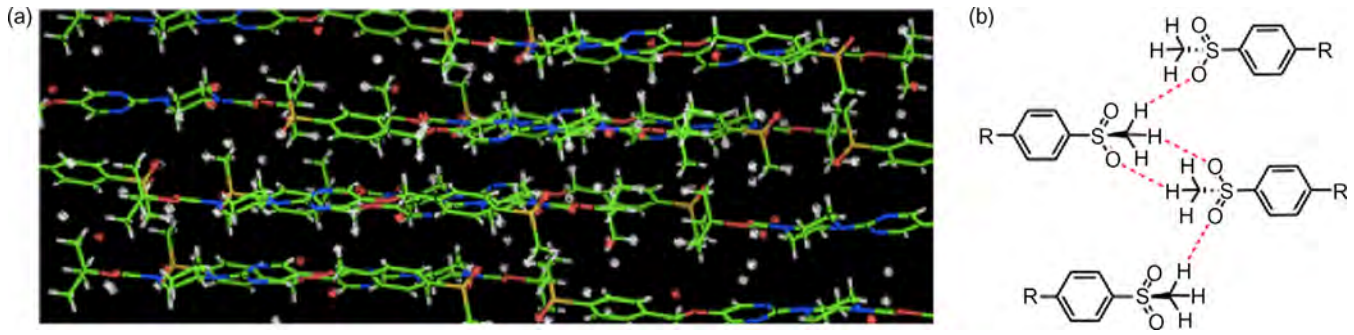


FIGURE 31.3 X-ray structure of compound **5**: (a) flat shape of **5** and its packing structure; (b) intermolecular head-to-head interaction of methyl sulfonyl groups (red dotted line).

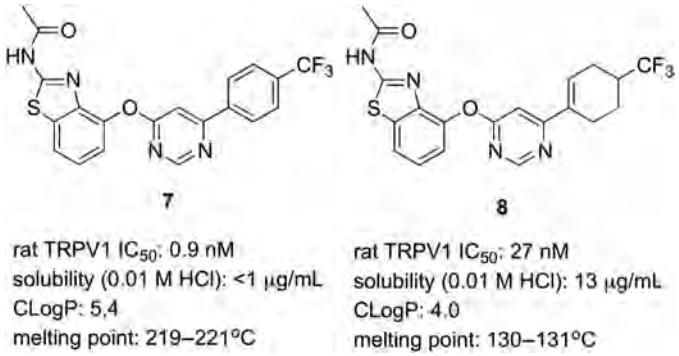


FIGURE 31.4 Improvement of thermodynamic aqueous solubility by increase of ratio of sp^3-hybridized carbons. Melting points are from the literature [17,18].

IV. IMPROVEMENT OF SOLUBILITY BY DISRUPTING MOLECULAR PLANARITY

A. Increase of Ratio of sp^3-hybridized Carbons

Vanilloid receptor-1 (TRPV1) antagonists [17]. Vanilloid receptor-1 (TRPV1) antagonist AMG 517 (**7**; Figure. 31.4) was found to have low thermodynamic aqueous solubility (<1 μg/mL in PBS or 0.01 M HCl), so a second-generation clinical candidate with increased aqueous solubility and a shorter half-life was desired. The strategy adopted was to introduce saturation into the 4-(trifluoromethyl)phenyl ring to reduce structural planarity, disrupt π−π stacking, and disrupt crystal packing. The partially saturated analog **8** showed improved thermodynamic solubility (13 μg/mL in 0.01 M HCl), being at least 13-fold more soluble than **7**, although this was accompanied with decreased potency. Compound **8** possessed a smaller CLogP value than **7**, suggesting that the solubility improvement might be due to both decrease of hydrophobicity and disruption of planarity. The fact that **7** had a higher melting point than **8** [17,18] also supports the view that disruption of planarity contributes at least in part to the improved solubility of **8**. Other reports suggest that removal of aromaticity can improve aqueous solubility, although the mechanism involved was not established [19,20].

Large-scale database study [21]. An analysis of the relationship between the ratio of sp^3-hybridized carbons (Fsp^3), where Fsp^3 = [(number of sp^3 hybridized carbons)/(total carbon count)], and stage of drug development indicated that Fsp^3 is correlated with success in compound transition from discovery through clinical testing and commercial introduction. In an attempt to explain these findings, solubility and melting point values were retrieved from another database of more than 1,000 compounds, although this included few complex pharmaceutical compounds. Fsp^3 was found to correlate positively with solubility and negatively with melting point. This analysis indicated that increasing the ratio of sp^3-hybridized carbons would be an effective strategy for improving solubility.

B. Increase of Dihedral Angle

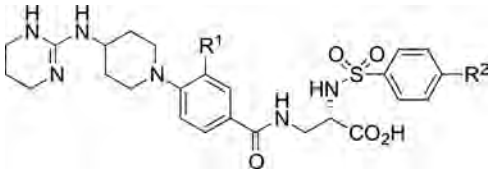
In the 1970s, the effects of polychlorinated biphenyls (PCBs) on the environment were widely discussed, and the aqueous solubility of PCBs—including the effect of reduced chlorine substitution—was examined in order to

understand the modes of transportation and attenuation of PCBs in the environment. It was found that *ortho*-chlorinated biphenyls, which would possess larger dihedral angles, showed higher solubility [22,23]. For example, 2,2′-dichlorobiphenyl (900 μg/mL) possessed greater aqueous solubility than 4-chlorobiphenyl (400 μg/mL) or 2,4′-dichlorobiphenyl (637 μg/mL). In the past decade, several researchers have investigated the relationship between solubility and dihedral angle of complex pharmaceutical compounds [11]. The reasons for focusing on the dihedral angle of bicyclic molecules were as follows: (1) a dihedral angle is a convenient numerical parameter that can be obtained by calculation or X-ray crystal structure analysis among various parameters of molecular planarity; and (2) numerous bioactive bicyclic molecules have been reported, because they can be easily synthesized by a Suzuki coupling reaction.

Integrin $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ dual antagonists [11,24–26]. The aqueous solubility of $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ dual antagonist **9** was very poor (<0.1 mg/mL). The first attempt to improve its solubility involved introduction of hydrophilic substituents (**10**–**12**). Briefly, these compounds showed decreased inhibitory activity in $\alpha_v\beta_3$-mediated cell adhesion assays using human vascular smooth muscle cells (VSMC) and/or $\alpha_{IIb}\beta_3$-mediated human platelet-rich plasma (hPRP) aggregation inhibition assays, as shown in Table 31.3. Therefore, introduction of hydrophilic substituents is unsuitable in this case. By contrast, the second approach (i.e., the introduction of hydrophobic substituents **14**–**16**) led to increased activity in receptor binding, VSMC, and hPRP assays. Furthermore, **14** and **16** showed increased aqueous solubility, being at least 6-fold and 13-fold more soluble than **9**, respectively (Table 31.4).

Next, the mechanism of the improved aqueous solubility of **14** and **16** was examined (Table 31.4). The hydrophobicity parameters (CLogP and retention time on reversed-phase HPLC [27]) indicate that **14** is more hydrophobic than **9**. In the case of the methoxy analog **16**, there was an apparent discrepancy, because the CLogP value was lower than that of **9**, whereas the retention time was larger than that of **9**. In addition, no clear relationship between aqueous solubility and hydrophobicity was observed among these compounds. On the other hand, a relationship was found between aqueous solubility and melting point: the rank order of aqueous solubility (**16** > **14** > **9**) showed a negative correlation with the order of melting points (**16** < **14** < **9**). Furthermore, the X-ray structure of **14** revealed a substantially increased dihedral angle between the piperidine ring and benzoyl group (Figure. 31.5) [24]. Therefore, it was suggested that the increase of aqueous solubility of **14** (and **16)** was caused by disruption of molecular planarity without a decrease of hydrophobicity, resulting in a decrease of intermolecular interaction. On the other hand, compound **13** showed decreased solubility despite its lower hydrophobicity compared with **14**. The melting point of **13** is higher than that of **14**, suggesting that the

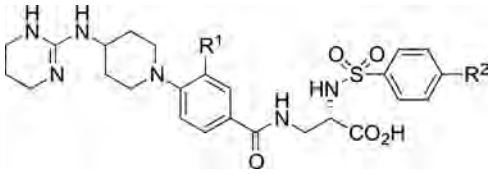
TABLE 31.3 Structure-Activity Relationships of Integrin $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ Dual Antagonists



Compound	R^1	R^2	IC_{50}			
			$\alpha_v\beta_3$	$\alpha_{IIb}\beta_3$	VSMC[a]	hPRP[b]
9	H	H	1.3	3.0	190	290
10	OH	H	0.44	0.98	530	170
11	F	CO_2H	0.77	1.2	660	930
12	OH	OH	0.30	0.94	390	510
13	F	OH	0.14	0.18	53	230
14	F	H	0.36	0.21	48	37
15	Cl	H	0.17	0.023	72	90
16	OMe	H	0.19	0.44	110	130

[a]*VSMC: $\alpha_v\beta_3$-mediated cell adhesion assay using human vascular smooth muscle cells (VSMC) and human vitronectin.*
[b]*hPRP: $\alpha_{IIb}\beta_3$-mediated human platelet aggregation inhibition assay.*

TABLE 31.4 Improvement in Aqueous Solubility of Integrin $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ Dual Antagonists by Increasing the Dihedral Angle in the Bicyclic Structures



Compound	R^1	R^2	Aqueous solubility[a] (mg/mL)	Melting point (°C)	HPLC retention time (min)[b]	CLogP
9	H	H	<0.1	252–254	8.25	1.1
14	F	H	0.6	182–184	9.73	1.7
16	OMe	H	1.3	162–164	8.72	0.79
13	F	OH	0.1	193–197	6.16	1.4

[a] *Solubility in water.*
[b] *Inertsil ODS-2 reversed-phase column.*

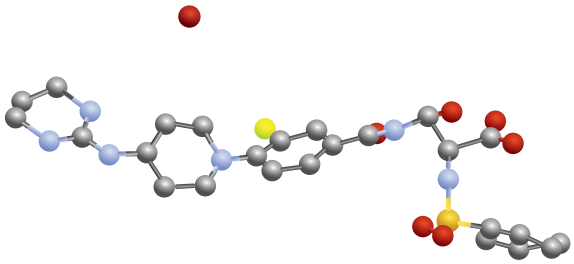


FIGURE 31.5 Single-crystal X-ray structure of compound **14**.

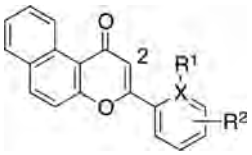
introduced hydroxyl group in **13** might form a new intermolecular hydrogen bond, which might lead to tighter crystal packing. It is noteworthy that disruption of molecular planarity (**14** and **16**) improved the solubility more than a decrease of hydrophobicity (**13**) in this case.

Aryl hydrocarbon receptor (AhR) agonists [11,26,28,29]. Improvement in the aqueous solubility of β-naphthoflavone **17**, an AhR agonist, was investigated. To decrease the planarity of **17**, several substituents (**18**–**26**) were introduced on the phenyl group of **17** (Table 31.5). To evaluate the effectiveness of this strategy, the 2-pyridyl analog **27** was also designed as a representative of the strategy of decreasing the hydrophobicity.

The thermodynamic aqueous solubility of **17**–**27** was evaluated. The aqueous solubility of **17** in a phosphate buffer (pH 7.4) was quite low (<0.15 μg/mL), so a mixture of an equal volume of phosphate buffer (pH 7.4) and EtOH was used as an aqueous medium for the evaluation of thermodynamic solubility. Even under this condition, the solubility of **17** was still poor (84.6 μg/mL). To clarify the reason for the better solubility of *ortho*-substituted naphthoflavones, various physicochemical parameters were measured or calculated (e.g., dihedral angle and maximum UV absorption (λmax) as parameters of molecular planarity, melting point for crystal packing, and retention time on reversed-phase HPLC and CLogP for hydrophobicity). The calculated dihedral angles of **17**–**27** for the optimized structures were obtained by means of density functional theory (DFT) calculations (gas phase, B3LYP/6-31G*). Extended conjugation of organic compounds results in a longer wavelength of λmax. Thus, compounds with similar conjugation should show a negative correlation of λmax and dihedral angle. Therefore, λmax in MeOH solution was evaluated as a parameter of molecular planarity.

First, the relationship between solubility and methyl group substitution was investigated. As expected, introduction of a methyl group increased hydrophobicity as judged from the CLogP values and retention times on reversed-phase HPLC. Despite the increased hydrophobicity, however, the solubility of *ortho*-methyl analog **18** (262 μg/mL) was 3-fold greater than that of **17**. In addition, the solubility depended on the position of the methyl group: the rank order of aqueous solubility was *ortho* (**18**) ≫ nonsubstituted (**16**) ≧ *meta* (**19**) > *para* (**20**). As for the mechanism of solubility improvement, the order of calculated dihedral angle was *ortho* (**18**) ≫ nonsubstituted

TABLE 31.5 Improvement in Thermodynamic Aqueous Solubility of AhR Agonists by Increasing the Dihedral Angle of the Bicyclic Structures



	R^1	R^2	X	Solubility[a] (μg/mL)	Melting point (°C)	Calculated dihedral angle[b] (°)	λmax (nm)	CLogP	HPLC retention time (min)[c]	EROD EC_{50} (μM)
17	H	H	C	84.6	165–167	17.8	273	4.8	7.24	1.4
18	H	*o*-Me	C	262	135–137	37.9	265	5.3	8.17	>10
19	H	*m*-Me	C	80.9	162	16.8	274	5.3	9.44	4.5
20	H	*p*-Me	C	35.4	194–195	16.5	285	5.3	9.43	2.8
21	Me	*o*-Me	C	1,270	92	70.0	261	5.7	9.13	>10
22	H	*o*-Cl	C	81.4	152	50.1	265	4.9	8.65	>10
23	Cl	*o*-Cl	C	150	160–161	79.7	261	5.1	10.7	>10
24	F	H	C	153	157	9.1	270	4.8	7.44	0.33
25	F	*o*-F	C	248	150	40.5	263	4.9	7.44	0.20
26	OMe	H	C	45.8	192–193	18.5	272	4.7	7.47	0.27
27	–	H	N	299	187–188	0.0	285	3.3	4.52	0.45

[a] *Solubility in an equal volume of EtOH and 1/15 M phosphate buffer (pH 7.4).*
[b] *Dihedral angle was estimated with Gaussian03.*
[c] *Inertsil ODS-4 reversed-phase column.*

(**17**) ≧ *meta* (**19**) ≧ *para* (**20**), and the order of λmax was *ortho* (**18**) ≪ nonsubstituted (**17**) ≦ *meta* (**19**) < *para* (**20**). Thus, the *ortho*-methyl analog **18** has the largest dihedral angle among these analogs. Furthermore, there was a relationship between aqueous solubility and melting point. The rank order of aqueous solubility corresponded to the order of melting points (*ortho* (**18**) ≪ *meta* (**19**) ≦ nonsubstituted (**17**) < *para* (**20**)). These results support the hypothesis that *ortho*-substitution disrupts molecular planarity by increasing the dihedral angle, in turn leading to a lower melting point and so increasing the solubility.

Next, the effect of the number of introduced methyl groups (**17**, **18**, **21**) was investigated. Naturally, the extent of hydrophobicity increase depends on the number of introduced methyl group(s). Despite increased hydrophobicity, *ortho*-dimethyl analog **21** was fifteen times more soluble (1,270 μg/mL) than **17**, and the solubility was dependent on the number of methyl groups. The rank order of aqueous solubility was dimethyl (**21**) > monomethyl (**18**) > nonsubstituted (**17**). This order is consistent with the rank order of lower melting point, larger calculated dihedral angle, and lower λmax. These results indicate that a larger number of *ortho*-methyl groups causes greater disruption of molecular planarity by increasing the dihedral angle, leading to a lower melting point and increasing the solubility.

The effect of chlorine substitution was similar to the case of the methyl groups. Dichloro analog **23** was much more soluble than **17**, in spite of having the highest hydrophobicity in this series (in terms of retention time on HPLC). However, the slight decrease in solubility of monochloro analog **22** and the higher melting point of dichloro analog **23** compared with monochloro analog **22** were unexpected results that were different from the case of the methyl analogs. The reasons for these phenomena may be that the highly hydrophobic character of chlorine favors a decrease of solubility, and the higher molecular weight of the chlorine atom contributes to increased melting point. Intermolecular halogen bonding involving the chlorine atom might lead to an increase of the melting point.

The solubility of *ortho*-fluoro analogs (**24**–**25**) also increased with increasing number of fluorine atoms, in association with a lower melting point, as expected. The physicochemical properties of fluorine analogs are basically consistent with theory. The only exception is a decreased calculated dihedral angle of **24** compared with **17**.

A possible explanation of the smaller dihedral angle of **24** would be interaction between the fluorine lone pair and hydrogen at the 2-position. However, several data—including the fact that the second-lowest energy conformation of **24** (ca. 1.5 kcal/mol less stable) has a larger calculated dihedral angle (38.4°), and **24** showed a lower melting point and smaller λmax—suggest that the actual dihedral angle might be larger than 9.1°.

Ortho-methoxy analog **26** was less soluble than **17**. Compound **26** also showed a higher melting point and almost the same calculated dihedral angle, compared with **17**. The reason for the relatively small dihedral angle may be similar to that in the case of monofluoro analog **24** (i.e., interaction between the oxygen lone pair and hydrogen at the 2-position). These results suggest that introduction of a methoxy group into **17** maintained the planarity by generating intramolecular interaction, leading in turn to maintained crystal packing energy and a higher melting point, and so decreasing the solubility.

Pyridine analog **27**, which lacks a hydrogen atom, showed a higher melting point, higher λmax, and decreased dihedral angle compared to **17**, though **27** showed better aqueous solubility (299 μg/mL), presumably because of its reduced hydrophobicity. Thus, this alternative strategy to improve solubility by increasing dihedral angle is quite distinct from the classic strategy based on decreasing the hydrophobicity of molecules. It is noteworthy that increase in dihedral angle (**21**) afforded a greater solubility improvement than decrease of hydrophobicity (**27**) in this case.

Single-crystal X-ray crystallographic analyses of representative compounds were performed [29]. Unfortunately, the crystal form of **17** was needle-shaped, which is not suitable for X-ray analysis. Therefore, methoxy analog **26** was selected as a representative compound with low solubility. Dimethyl analog **21** was also selected as a representative compound with high solubility. The X-ray crystal structures of **21** and **26** are shown in Figure 31.6. The dihedral angle measured from the X-ray crystal structure of **21** is larger than that of **26**. This order is consistent with the DFT calculations. Next, the packing structures were analyzed. The crystal density of **21** obtained from the X-ray crystal structure is lower than that of **26**, indicating that the crystal packing energy of **21** is lower than that of **26**. All these data indicate that the increased solubility of **21** can be explained in terms of the hypothesis that increase of dihedral angle results in disruption of molecular planarity, leading in turn to decreased intermolecular interaction (π−π stacking), thus increasing the solubility.

To evaluate the AhR-agonistic activity of the prepared compounds, CYP1A1-dependent 7-ethoxyresorufin *O*-deethylase (EROD) activity was measured (Table 31.5). Methyl analogs **18**−**21** and chloro analogs **22**−**23** showed decreased agonistic activity. On the other hand, the fluoro analogs **24** and **25**, methoxy analog **26**, and pyridine analog **27** showed increased activity. Among them, the difluoro analog **25** had the best overall profile, being seven

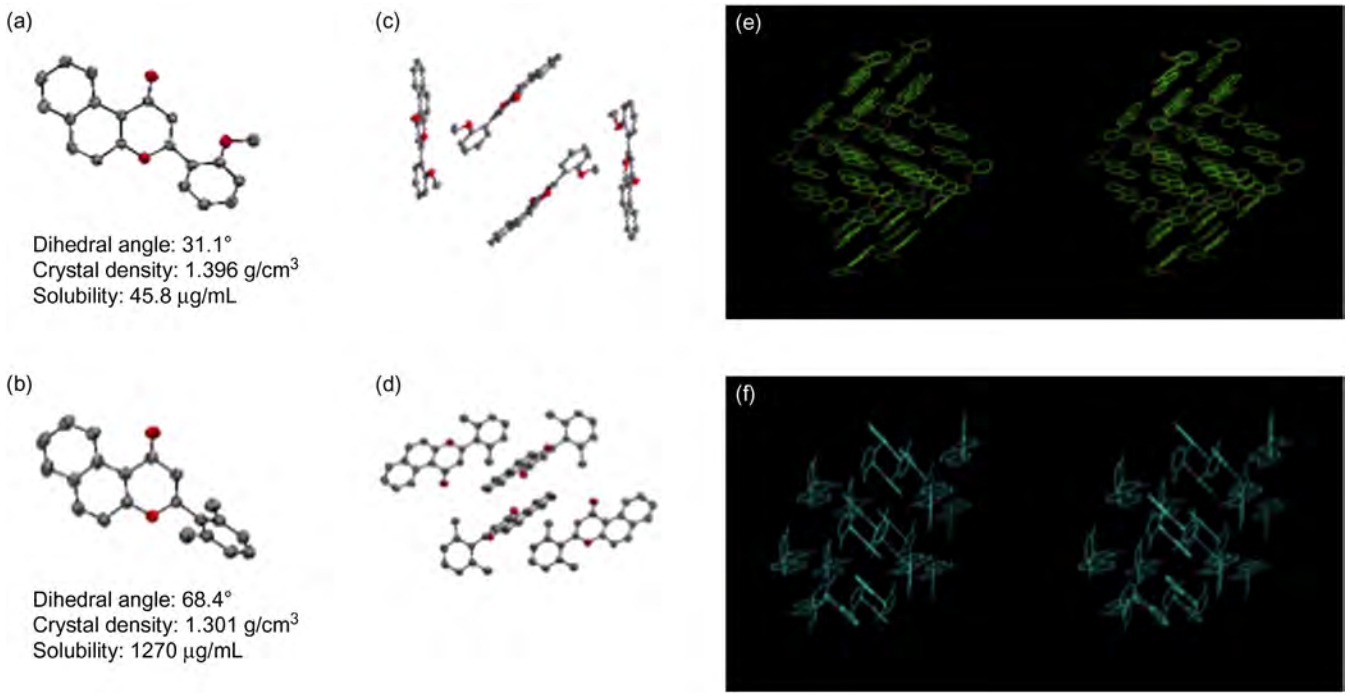
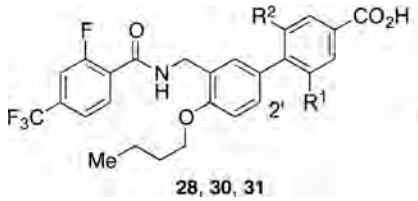
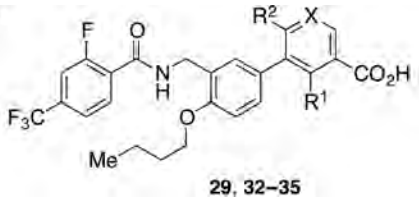


FIGURE 31.6 X-Ray crystal structures of **26** (a, c, and e) and **21** (b, d, and f). Single molecules are shown in a and b, packing structures in unit cells are shown in c and d, and stereo views of the packing structures are shown in e and f.

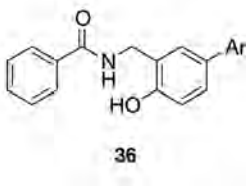
TABLE 31.6 Improvement in Thermodynamic Aqueous Solubility of PPARδ Partial Agonists by Increasing the Dihedral Angle of the Bicyclic Structures



28, 30, 31



29, 32–35



36

	R^1	R^2	X	Solubility (mg/mL) 50% EtOH[a]	Solubility (mg/mL) Phosphate buffer (pH 7.4)	Melting point (°C)	Calculated dihedral angle[b] (°)	CLogP	HPLC retention time (min)[c]	PPARδ EC_{50} (nM)
28	H	H	–	0.375	<0.001	259–262	43.5	6.3	7.98	170
30	Me	H	–	0.985	<0.001	241–243	52.5	6.8	9.46	11
31	F	H	–	3.22	<0.001	221–223	36.1	7.1	8.72	53
29	H	H	CH	1.35	<0.001	177–178	36.9	6.4	7.79	29
32	Me	H	CH	9.95	<0.001	146–149	57.5	6.9	8.65	1.6
33	F	F	CH	10.4	0.0217	177	46.2	7.5	7.42	5.7
34	H	H	N	9.03	0.00762	152	37.4	4.8	3.61	220
35	Me	Me	N	17.7	2.70	104–106	78.1	5.8	4.36	76

[a] *Solubility in a mixture of equal volumes of EtOH and 1/15 M phosphate buffer (pH 7.4).*
[b] *Calculated dihedral angles were estimated by using Gaussian03 with simplified models,* ***36****.*
[c] *Waters μBondapak reversed-phase column.*

times more potent and three times more soluble than **17**. The aqueous solubility of **17** in phosphate buffer (pH 7.4) was very low and undetectable by HPLC, whereas that of difluoro analog **25** was measurable under the same condition (0.021 μg/mL).

Peroxisome proliferator-activated receptor (PPAR) δ partial agonists [11,26,30]. To obtain crystals of PPARδ partial agonist complexes for X-ray analysis, the aqueous solubility of PPARδ partial agonists (biphenylcarboxylic acids **28** and **29**) needs to be increased. As the first approach for increasing the aqueous solubility of **28** and **29**, introduction of an oxygen atom into the *n*-butyl group and replacement of the phenyl ring with hetero rings were tried. Unfortunately, although these solubilizing modifications decreased the hydrophobicity of the molecules, they resulted in decreased PPARδ partial agonistic activity. Thus, an alternative approach to increasing aqueous solubility—that is, increase of dihedral angle—was investigated (Table 31.6). Briefly, introduction of methyl groups or fluorine atoms at the *ortho*-position of the biphenyl moiety led to increased PPARδ agonistic activity. Pyridyl analog **34** showed decreased activity, but the introduction of two methyl groups (**35**) resulted in three times stronger activity than **34** and comparable activity to the parent compound **29**.

The thermodynamic aqueous solubility of the partial agonists **28**–**35** was evaluated. The values of aqueous solubility of **28** and **29** in 1/15 M phosphate buffer (pH 7.4) were quite low (<0.001 mg/mL). A mixture of equal volumes of 1/15 M phosphate buffer (pH 7.4) and EtOH was also used as an aqueous medium for the evaluation of thermodynamic solubility. All the compounds shown in Table 31.6 had higher solubility than the parent compounds. In the *para*-carboxyl series **28**, **30**, and **31**, introduction of a methyl group (**30**) or fluorine atom (**31**) into the biphenyl moiety resulted in better solubility than that of **28**, as expected. The fluoro analog **31** was nine times more soluble (3.22 mg/mL) than **28** in aqueous EtOH. In the case of the *meta*-carboxyl series **29** and **32**–**35**, introduction of a methyl group (**32**) or two fluorine atoms (**33**) resulted in seven times greater solubility than that of **29** in aqueous EtOH. Furthermore, **33** showed moderate solubility in a phosphate buffer (0.0217 mg/mL); this was the first example among these PPAR ligands. Pyridyl analog **34** showed better solubility in both 50 percent EtOH and a phosphate buffer than did the parent compound **29**. Surprisingly, an increase of dihedral angle (**33**) resulted in greater solubility in the phosphate buffer than did a decrease of hydrophobicity (**34**) in this case. Dimethylpyridyl analog **35** showed two times greater solubility than **34** in aqueous EtOH. It is noteworthy that **35** also has greatly improved solubility in a phosphate buffer (2.70 mg/mL); it is 350 times more soluble than **34**.

The calculated dihedral angle, melting point, retention time on reversed-phase HPLC, and CLogP were evaluated to confirm the mechanism of the improved solubility of the biaryl analogs (Table 31.6). Dihedral angles in optimized structures of simplified models (**36**) were obtained by means of DFT calculations. All the compounds with methyl groups introduced at the *ortho*-position of the biaryl moiety (**30**, **32**, and **35**) showed increased hydrophobicity, larger dihedral angle, and lower melting point than those of the parent compounds **28**, **29**, and **34**, respectively. The most soluble analog, **35**, showed the lowest melting point and the largest dihedral angle in this series. These results suggest that introduction of methyl groups into biaryl molecules results in an increase of the dihedral angle, disrupting the molecular planarity and leading in turn to decreased crystal packing energy and a lower melting point, with a consequent increase of solubility.

When **35** was compared with **29**, it was found that the two kinds of modifications—replacement of the phenyl group with a pyridine ring and introduction of methyl groups—had resulted in more than 2,700 times higher solubility in phosphate buffer. This improvement of solubility could be ascribed to two factors: a decrease of hydrophobicity by introduction of the pyridine ring (**29** vs **34**) and an increase of dihedral angle by introduction of the methyl groups (**34** vs **35**). Thus, these results indicated that a combination of strategies for improving aqueous solubility is an effective approach.

In the case of compounds containing fluorine atoms, the reason for the improvement of aqueous solubility is not clear. Compound **31** possessed a smaller calculated dihedral angle, lower melting point, and higher hydrophobicity than **28**. A possible explanation of the small dihedral angle would be interaction between the fluorine lone pair and hydrogen at the 2′-position. Fluorinated β-naphthoflavone **24** showed improved solubility and a relatively small calculated dihedral angle (Table 31.5). However, **24** and **31** showed lower melting points (and smaller λmax), suggesting that the actual dihedral angle might be larger than the calculated value. On the other hand, **33** showed the same melting point as **29** but had a larger calculated dihedral angle and higher hydrophobicity. Taken together, these results indicate that introduction of fluorine atoms might improve aqueous solubility, not only by increasing the dihedral angle but also via other mechanisms.

Large-scale database study [31]. The AstraZeneca database of thermodynamic aqueous solubility measurements, rat plasma protein binding measurements, and rat oral exposure measurements was searched for occurrences of pair fitting. In this large-scale study, Leach et al. focused on matched pairs resembling those in Figure 31.7 (**37** and **38**). There were thirteen examples for Y = F with a mean *Z*-score (the number of standard deviations away from the mean difference) for the change in log(solubility) of + 0.54 (range = −0.44 to + 1.45), and there were five examples for Y = Cl with a mean *Z*-score of + 1.93 (range = +1.34 to + 2.42), although precise structures and solubility were not reported. These results indicate that *ortho*-substitution of biaryls tends to increase solubility more than expected or to decrease it less than expected.

Vascular endothelial growth factors receptor 2 (VEGFR2) kinase inhibitors [32]. Increase of the dihedral angle was also useful for improvement in solubility of benzamides, anilides, and phenylureas. Introduction of a methyl group at the 2-position of the anilide moiety of a VEGFR2 kinase inhibitor **40** led to increased kinetic solubility in JP2 (the second fluid of the Disintegration Test of the Japanese Pharmacopoeia, pH 6.8), and **40** was 44-fold more soluble than **39**. The melting point of **40** was lower than that of **39** (Table 31.7). It was also reported that *ortho*-methyl analog **40** was more soluble than *meta*-methyl analog **41** or *para*-methyl analog **42**. *Ortho*-fluoro analog **43** showed better solubility than **39** and **40**. The improved solubility of **43** could be due to both increased dihedral angle and lower hydrophobicity compared with **39**. It is interesting that the kinetic solubility of *ortho*-methyl/monomethyl pyrazole analog **44** was increased at least 67-fold over that of regioisomer **39**, even though they have exactly the same molecular formula. Fluoro analog **43** showed a high peak plasma concentration as well as high AUC_{0-8h} values, and the improved solubility might contribute to the improved pharmacokinetic properties. Compound **43** also showed potent antitumor efficacy in mice xenograft models.

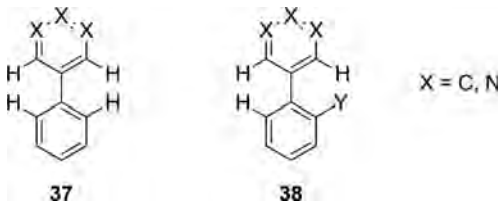
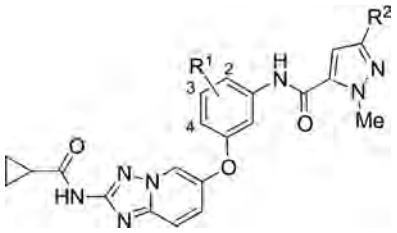


FIGURE 31.7 Matched pairs of compounds in the AstraZeneca database, used to examine the effect of *ortho*-substitution on thermodynamic aqueous solubility.

TABLE 31.7 Improvement in Kinetic Aqueous Solubility by *ortho*-Substitution of Anilide



Compound	R^1	R^2	VEGFR IC_{50} (μM)	Aqueous solubility[a] (μg/mL)	Melting point (°C)	CLogP
39	H	Me	1.1	1.2	244–245	2.4
40	2-Me	Me	1.9	53	217–220	2.2
41	3-Me	Me	2.1	3.2	233	2.9
42	4-Me	Me	3.7	3.7	222–225	2.9
43	2-F	Me	2.0	65	209	1.9
44	2-Me	H	1.6	>80	131–135	1.9

[a]*Kinetic solubility in JP2.*

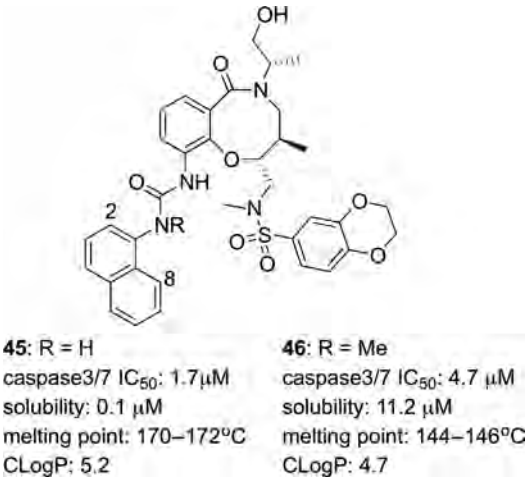
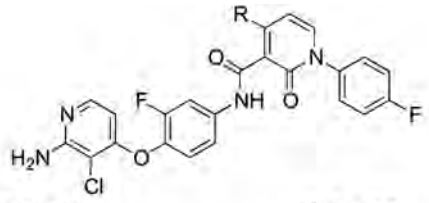


FIGURE 31.8 Improvement in thermodynamic aqueous solubility by increase in dihedral angle of urea.

STAT1 inhibitor [33]. The thermodynamic solubility of STAT1 inhibitor **45** bearing a urea moiety was determined to be 0.1 μM in PBS (pH7.4; Figure. 31.8). N-Methyl-N-1-naphthyl urea **46** was designed in an attempt to increase the aqueous solubility by increasing the dihedral angle. The addition of the methyl group to the urea nitrogen may prevent access to planar conformations, which should be hindered by a steric interaction between the N-methyl and C2- or C8-hydrogen of the quinolone moiety. In fact, **46** exhibited a 110-fold increase in solubility (11.2 μM). Decrease of crystal packing was supported by the change of melting point (i.e., the melting point of **46** was lower than that of **45**). Another possible explanation of the lower melting point for **46** might be that the introduced methyl group blocks intermolecular hydrogen bonding with the NH hydrogen of the urea.

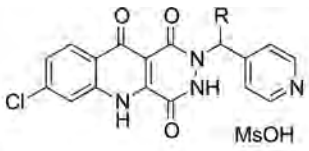
Met kinase inhibitors [34]. Highly planar Met kinase inhibitor **47** was modified by introducing an ethoxy group at the pyridone 4-position (Figure. 31.9). Compound **48** possessed comparable *in vitro* activity to **47** but showed >40-fold improvement in aqueous solubility at pH 1.0. Hydrophobicity parameters (CLogP and retention time on HLPC) of **47** and **48** were discrepant, so it was difficult to judge which compound is more hydrophobic. The pharmacokinetic parameters were examined in mouse (10 mg/kg), rat (10 mg/kg), and dog (5 mg/kg), and compound **48** was well absorbed after oral administration. The measured oral bioavailabilities (F_{po}) were



47: R = H
Met IC_{50}: 1.0 nM
solubility at pH 1: <10 μg/mL
CLogP: 4.3
HPLC retention time: 9.54 min

48: R = OEt
Met IC_{50}: 3.9 nM
solubility at pH 1: 400 μg/mL
CLogP: 4.9
HPLC retention time: 8.09 min

FIGURE 31.9 Improvement in aqueous solubility by *ortho*-substitution of pyridone carboxamide.



MsOH

49: R = H
^{3}H-MDL 105519 *Ki*: 115 nM
solubility at pH 7.4: 0.05 mg/mL
rat F: 5%
melting point: 277–278°C
CLogP: 1.4

50: R = Me
^{3}H-MDL 105519 *Ki*: 248 nM
solubility at pH 7.4: >0.29 mg/mL
rat F: 30%
melting point: 245–247°C
CLogP: 1.8

FIGURE 31.10 Improving aqueous solubility by α-substitution. Biological activity was evaluated in terms of displacement of radioligand binding at the NMDA glycine site ([^{3}H]MDL 105519) in rat brain membranes.

excellent (100 percent) in all three species. Compound **48** also showed potent antitumor efficacy in mouse xenograft models. Other reports suggest that *ortho*-substitution can improve aqueous solubility, although the mechanism involved was not established [35–37].

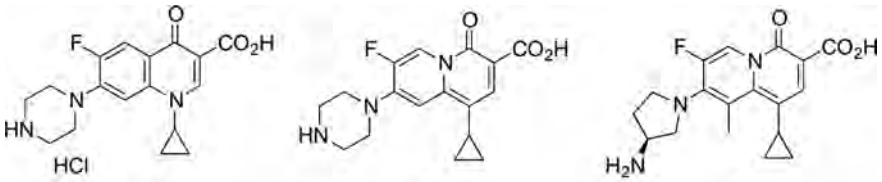
C. Introduction of Substituent at Benzylic Position

N-Methyl-D-aspartate (NMDA) antagonists [38]. A methyl group was introduced into the α-position of **49** to improve solubility by disrupting molecular planarity. The solubility of **50** was improved by at least 5.8-fold compared with that of **49** (Figure. 31.10). The melting point of α-methyl analog **50** was lower than that of **49**. The CLogP values indicated that introduction of a methyl group into **49** also increased the hydrophobicity. In accordance with the above results, the α-methyl analog **50** showed higher rat oral bioavailability (30 percent) than **49** (5 percent). An interesting extension of this approach would be stereoselective synthesis of compounds with asymmetric benzylic substitution. Several other reports suggest that introduction of substituents into benzylic positions improves aqueous solubility, although the mechanism of the solubility improvement is not clear [39,40].

D. Twisting of Fused Rings

DNA gyrase inhibitors [41]. When the quinolone ring of ciprofloxacin **51** was substituted with 2-pyridone, the resulting compound **52** showed about three times greater thermodynamic aqueous solubility and bioavailability, while retaining similar activity to **51** (Table 31.8). Compound **53** showed three times greater solubility than **51**, with eight times more potent activity. Apparently, transposition of the nitrogen of quinolone (**51**) to the bridgehead position (**52** and **53**) results in a change in the overall polarity of the molecule and perhaps in perturbation of the planarity of the fused bicyclic core structure. These changes are associated with favorable changes in the solubility and pharmacokinetics of the 2-pyridones. A single-crystal X-ray structure analysis of **53** was also reported. The steric congestion created by the interaction of the methyl group with the cyclopropane forces the

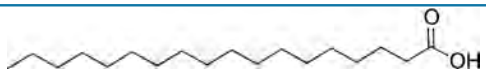
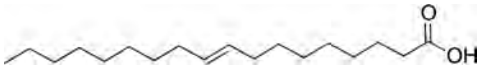
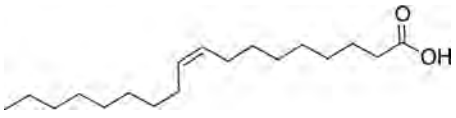
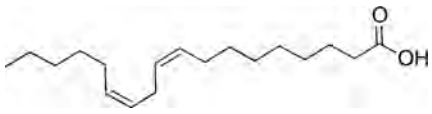
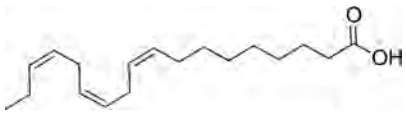
TABLE 31.8 Improvement of Thermodynamic Aqueous Solubility by Twisting of Fused Rings



Compound	CC_{50}[a] (μg/mL)	Aqueous solubility[b] (mg/mL)	Rat F (%)	CLogP
ciprofloxacin (**51**)	0.24	0.08	16	0.65
52	0.15	0.25	46	−0.15
53	0.03	0.25	32	0.04

[a]*CC_{50} is defined as the drug concentration that causes 50% inhibition of maximal gyrase (*E coli *H560)-mediated DNA cleavage.*
[b]*Thermodynamic solubility in phosphate buffer (0.05 M), pH 7.4, at 37 °C.*

TABLE 31.9 Melting Points of C18 Fatty Acids

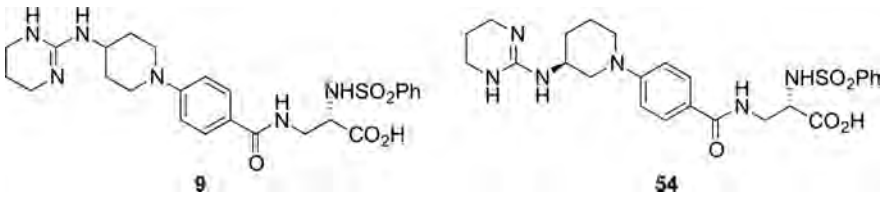
Fatty acid		Melting point (°C)
stearic acid		69
elaidic acid		46
oleic acid		13
linoleic acid		−5
linolenic acid		−11

methyl and cyclopropyl groups out of the plane of the pyridone ring, generating a twist that was estimated to be approximately 30°. This twisting conformation of **53** is considered to account at least in part for the enhanced solubility.

V. IMPROVEMENT OF SOLUBILITY BY BENDING THE MOLECULAR STRUCTURE

From the viewpoint of molecular symmetry, Gavezzotti reported in 1995 that *ortho*- and *meta*-disubstituted benzenes melt at lower temperatures than the *para*-isomers [42]. He noted that "a very old rule of thumb says that symmetrical molecules pack in a three-dimensional periodic lattice more easily than less symmetrical ones, and hence form more stable, higher-melting and less soluble crystals" [42]. The melting points of C18 fatty acids are shown in Table 31.9. The melting points of cis-fatty acids decrease with increasing number of cis-double bonds in the molecule. The rationale for this observation is that larger numbers of cis-double bonds cause the molecular structures to change from extended structures to more bent structures, which would decrease crystal packing. Therefore, molecular modification of extended compounds to generate bent structures—in other words, disruption of molecular symmetry—is expected to disrupt crystal packing and in turn to increase solubility.

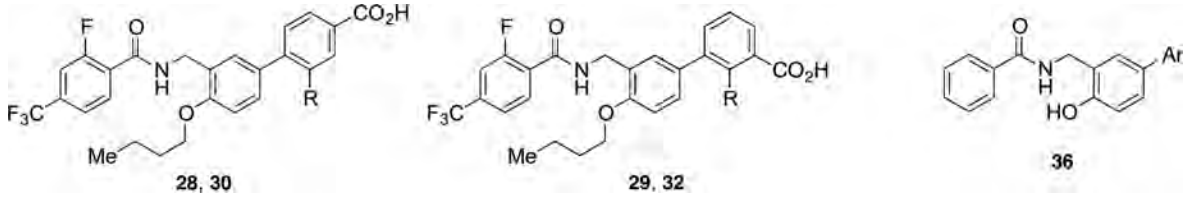
TABLE 31.10 Improvement in Aqueous Solubility of Integrin $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ Dual Antagonist by Disrupting Molecular Symmetry



	Aqueous solubility[a] (mg/mL)	CLogP	HPLC retention time (min)[b]	Melting point (°C)	IC_{50} (nM)			
					$\alpha_v\beta_3$	$\alpha_{IIb}\beta_3$	VSMC[c]	hPRP[d]
9	<0.1	1.1	8.25	252–254	1.3	3.0	190	290
54	3.5	1.1	12.2	181–184	0.48	0.56	31	83

[a] *Solubility in water.*
[b] *Inertsil ODS-2 reversed-phase column.*
[c] *VSMC: $\alpha_v\beta_3$-mediated cell adhesion assay using human vascular smooth muscle cells (VSMC) and human vitronectin.*
[d] *hPRP: $\alpha_{IIb}\beta_3$-mediated human platelet aggregation inhibition assay.*

TABLE 31.11 Improvement in Aqueous Solubility of PPARδ Partial Agonists by Disrupting Molecular Symmetry



	R	Solubility[a] (mg/mL)	Melting point (°C)	Calculated dihedral angle[b] (°)	CLogP	HPLC retention time (min)[c]
28	H	0.375	259–262	43.5	6.3	7.98
30	Me	0.985	241–243	52.5	6.8	9.46
29	H	1.35	177–178	36.9	6.4	7.79
32	Me	9.95	146–149	57.5	6.9	8.65

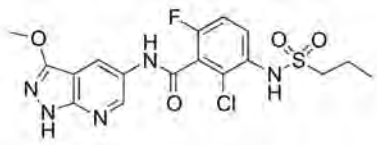
[a] *Solubility in a mixture of equal volumes of EtOH and 1/15 M phosphate buffer (pH 7.4).*
[b] *Calculated dihedral angle was estimated for the simplified models **36** by using Gaussian03.*
[c] *Waters μBondapak reversed-phase column.*

Integrin dual $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ antagonist [11,25,26,43]. Disruption of the molecular symmetry of $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ antagonist **9** was examined. As shown in Table 31.10, changing the position of the tetrahydropyrimidylamino group (**54**) resulted in at least 35-fold increase of aqueous solubility compared with that of the lead compound **9**. Compound **54** was more hydrophobic, as judged from retention time on reversed-phase HPLC. In addition, the melting point of **54** is lower than that of **9**. Thus, the increase of aqueous solubility of **54** would not have been caused by decreased hydrophobicity but rather by bending of the molecular structure, resulting in a decrease of intermolecular interaction. Interestingly, **54** was more soluble and more potent than **9**, apparently as a result of the disruption of molecular symmetry without any increase of molecular weight. Further chemical modification of 3-aminopiperidine derivative **54** was carried out, and the 3-aminopiperidine analogs were found to possess sufficient aqueous solubility.

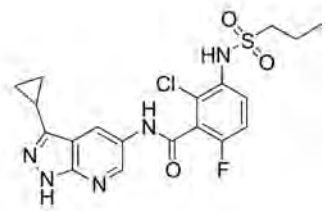
PPARδ partial agonists [11,26,30]. Focusing on the position of the carboxyl group in PPARδ partial agonists, the relationship between aqueous solubility and molecular symmetry was analyzed. When the position of the carboxyl group was changed from *para* to *meta*, the thermodynamic solubility increased (4-fold for **28** vs **29**; 10-fold for **30** vs **32**; Table 31.11). The melting points of *meta*-carboxylic acids **29** and **32** were lower than those of the corresponding *para*-analogs **28** and **30**, respectively. These results suggest that bending of the molecular structure results in decreased crystal packing energy and lower melting points, leading to increased solubility. Again, it is

TABLE 31.12 Improvement in Aqueous Solubility by Disrupting Molecular Planarity/Symmetry

55



56



57

	B-Raf IC_{50} (nM)	Solubility (μg/mL) pH6.5	pH7.4	Melting point (°C)	Crystal density (g/cm^3)	CLogP
55	4.8	4	9	229	1.559	1.6
56	2.2	<1	2	221		1.6
57	3.8	54	127	164	1.380	2.2

Intermolecular hydrogen bonds of **55** are shown as red dotted lines.

noteworthy that more soluble compounds (**29** and **32**) were generated by disruption of molecular symmetry without any increase of molecular weight.

B-Raf^{V600E} Inhibitors [44]. The B-Raf^{V600E} inhibitor **55** possessed a favorable ADME profile, but its aqueous solubility was low. Therefore, a solvent of 40 percent PEG400/10 percent EtOH/50 percent water was required for oral exposure in mice. Single-crystal X-ray structural analysis revealed that the monohydrate of **55** crystallizes in an extended conformation, an arrangement that leads to formation of head-to-tail dimers. These dimer pairs pack one on top of another, and they align to form hydrogen bonds both above and below with the amides from other dimer pairs (Table 31.12). These intermolecular interactions lead to a highly stable packing arrangement. To disrupt this tight crystal packing, two parts of **55** were modified. The first modification was the substitution of the fluorine atom with chlorine, which might lead to weakening of hydrogen bonds involving the central phenyl group bearing the bulky chlorine atom. Second was the substitution of a methoxy group with a bulky substituent, which should lead to disruption of planarity. As a result, the aqueous solubility of **57** was increased over 10-fold compared with that of **55**, and over 50-fold compared with that of **56**. Single-crystal X-ray structure analysis of **57** revealed that head-to-tail dimer formation did not occur, so that π-stacking interactions were reduced compared to **55**. In addition, the bond from amide carbonyl to the central phenyl ring is rotated approximately 180° compared to **55**, rendering dimer formation impossible. The bending conformation is expected to be unfavorable for crystal packing. These results indicated that molecular modification to disrupt planarity/symmetry of **55**, focusing on less effective crystal packing, resulted in a decrease of melting point and crystal density, with a consequent increase of aqueous solubility despite greater hydrophobicity.

VI. ADVANTAGES OF IMPROVING SOLUBILITY BY MOLECULAR MODIFICATION TO WEAKEN INTERMOLECULAR INTERACTION

As mentioned above, the strategy of decreasing crystal packing is quite distinct from the general/classic strategy to improve solubility by decreasing the hydrophobicity of molecules and offers several advantages.

First, the introduction of hydrophilic substituents often interferes with protein−drug interactions. For example, introduction of hydrophilic substituents into integrin antagonists and PPARδ partial agonists caused a decrease in their activities. By contrast, disruption of molecular planarity and symmetry of integrin antagonists and PPARδ partial agonists increased both biological activity and aqueous solubility. These results suggest that this alternative approach to improving aqueous solubility is particularly attractive when introduction of hydrophilic substituents is unsuccessful.

Second, disruption of molecular planarity coupled with an increase of hydrophobicity resulted in better aqueous solubility than a decrease of hydrophobicity in three cases (**14** vs **13**, **21** vs **27**, and **33** vs **34**). Introduction of hydrophilic groups may decrease aqueous solubility if the introduced hydrophilic groups form new intermolecular hydrogen bonds that induce tighter crystal packing. We can see an example of a less hydrophobic compound with worse aqueous solubility by comparing **14** to **13**.

Third, medicinal chemists sometimes find that both solubility and hydrophobicity need to be increased, for example, to improve the oral bioavailability of highly hydrophilic compounds with insufficient solubility or to improve brain penetration. In such cases, if compounds soluble in both organic solvents and water are generated, aqueous solubility would be improved even if the LogP value is increased (i.e., the solubility ratio in octanol is increased). Several of the compounds mentioned above showed improved solubility despite having higher hydrophobicity than their parent compounds. It is expected that the solubility of these compounds in organic solvents would also be increased.

Introduction of substituents often interferes with protein–drug interactions because of steric restrictions at binding sites. Fluorine has the smallest van der Waals radius of any substituent except for hydrogen. Thus, fluorine substitution should have the least impact on target protein–drug interaction, and thus fluorine might be the most versatile substituent to increase dihedral angle without loss of biological activity. In fact, fluorinated integrin antagonist **14**, AhR agonists **24** and **25**, PPARδ partial agonists **31** and **33**, and VEGFR2 kinase inhibitor **43** showed increased solubility. On the other hand, it appears that introduction of fluorine atoms might improve aqueous solubility, not only by increasing dihedral angle but also via other mechanisms.

Bending of molecular structures also provides a means to increase aqueous solubility without increasing the molecular weight, which could have adverse consequences for the pharmacokinetics. Examples of compounds whose aqueous solubility was increased without any increase of molecular weight are integrin antagonist **54** and PPARδ partial agonists **30** and **32**.

Finally, multiple strategies to improve aqueous solubility are likely to be superior to a single strategy. For example, two kinds of modifications (**35**; substitution of pyridine ring and introduction of methyl groups) of PPARδ partial agonist **29** led to more than 2,700 times higher solubility in a phosphate buffer, owing to the combination of reduction of hydrophobicity and increase of dihedral angle. As another example, two kinds of modifications of PPARδ partial agonist **28** (**32**; changing the position of the carboxyl group from *para* to *meta* and introduction of a methyl group at the *ortho*-position) led to twenty-seven times higher solubility in aqueous EtOH. Bending the molecular structure by changing the position of the carboxyl group contributes in part to the improved solubility (**28** vs **29**), as does the increase of dihedral angle by introduction of the methyl group (**29** vs **32**).

VII. CONCLUSION

Aqueous solubility is essential for drug candidates, and improvement of the aqueous solubility of small molecules is a major issue for medicinal chemists. The strategy of introducing hydrophilic groups into a molecule is generally used for improving aqueous solubility but is not universally effective. Thus, we have focused in this chapter on an alternative strategy for improving aqueous solubility, that is, weakening intermolecular interaction, which in turn leads to the disruption of crystal packing. The disruption of crystal packing is effective in increasing aqueous solubility. In one example, at least 350-fold improvement of aqueous solubility was obtained by increasing the dihedral angle, despite increased hydrophobicity. Strategies for molecular modification to weaken intermolecular interaction include disruption of intermolecular hydrogen bonds, removal of aromaticity, increase of dihedral angle, substitution of benzylic positions, twisting of fused rings, and bending the molecular structure. Melting point and crystal density derived from X-ray analysis can be considered parameters of intermolecular interaction and crystal packing.

It should be emphasized that the strategy of improving aqueous solubility by molecular modification to weaken intermolecular interaction—that is, blocking intermolecular hydrogen bonds, decreasing the number of sp^2-hybridized carbons, twisting, and bending—is quite distinct from the general/classic strategy based on decreasing the hydrophobicity of molecules. Possible advantages include:

- In cases where introduction of hydrophilic substituents is inappropriate, hydrophobic substituents can be used.
- Disruption of crystal packing is more effective than decreasing hydrophobicity in some cases.
- Both hydrophobicity and aqueous solubility can be increased simultaneously.

- Solubility can be improved with little or no increase of molecular weight.
- Combinations of different strategies may be superior to a single strategy.

For example, the combination of an increase in dihedral angle with a decrease of hydrophobicity improved solubility by at least 2,700-fold in one example.

References

[1] Food and Drug Administration Center for Drug Evaluation and Research. Guidance for industry: Waiver of in vivo bioavailability and bioequivalence studies for immediate release dosage forms based on a biopharmaceutical classification system; 2000.
[2] Yu LX, Amidon GL, Polli JE, Zhao H, Mehta MU, Conner DP, et al. Biopharmaceutics classification system: the scientific basis for biowaiver extensions. Pharm Res 2002;19:921–5.
[3] Bhattachar SN, Deschenes LA, Wesley JA. Solubility: it's not just for physical chemists. Drug Discov Today 2006;11:1012–18.
[4] Yalkowsky SH. Solubility and solubilization in aqueous media. Washington, DC: American Chemical Society; 1999.
[5] Palucki M, Higgins JD, Kwong E, Templeton AC. Strategies at the interface of drug discovery and development: early optimization of the solid state phase and preclinical toxicology formulation for potential drug candidates. J Med Chem 2010;53:5897–905.
[6] Babu NJ, Nangia A. Solubility advantage of amorphous drugs and pharmaceutical cocrystals. Cryst Growth Des 2011;11:2662–79.
[7] Leo A, Hansch C, Elkins D. Partition coefficients and their uses. Chem Rev 1971;71:525–616.
[7a] Fujita T, Iwasa J, Hansh C. A new substituent constant, π, derived from partition coefficients. J Am Chem Soc 1964;86:5175–80.
[8] Stella VJ, Nti-Addae KW. Prodrug strategies to overcome poor water-solubility. Adv Drug Deliv Rev 2007;59:677–94.
[9] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 2001;46:3–26.
[10] Jain N, Yalkowsky SH. Estimation of the aqueous solubility I: application to organic nonelectrolytes. J Pharm Sci 2001;90:234–52.
[11] Ishikawa M, Hashimoto Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. J Med Chem 2011;54:1539–54.
[12] Lange JHM, Coolen HKAC, van Stuivenberg HH, Dijksman JAR, Herremans AHJ, Ronken E, et al. Synthesis, biological properties, and molecular modeling investigations of novel 3,4-diarylpyrazolines as potent and selective CB1 cannabinoid receptor antagonists. J Med Chem 2004;47:627–43.
[13] Li JJ, Sutton JC, Nirschl A, Zou Y, Wang H, Sun C, et al. Discovery of potent and muscle selective androgen receptor modulators through scaffold modifications. J Med Chem 2007;50:3015–25.
[14] Kim I-H, Heirtzler FR, Morisseau C, Nishi K, Tsai H-J, Hammock BD. Optimization of amide-based inhibitors of soluble epoxide hydrolase with improved water-solubility. J Med Chem 2005;48:3621–9.
[15] Hodous BL, Geuns-Meyer SD, Hughes PE, Albrecht BK, Bellon S, Bready J, et al. Evolution of a highly selective and potent 2-(pyridin-2-yl)-1,3,5-triazine Tie-2 kinase inhibitor. J Med Chem 2007;50:611–26.
[16] Scott JS, Birch AM, Brocklehurst KJ, Broo A, Brown HS, Butlin RJ, et al. Use of small-molecule crystal structures to address solubility in a novel series of G protein coupled receptor 119 agonists: optimization of a lead and in vivo evaluation. J Med Chem 2012;55:5361–79.
[17] Wang HL, Katon J, Balan C, Bannon AW, Bernard C, Doherty EM, et al. Novel vanilloid receptor-1 antagonists. 3. The identification of a second-generation clinical candidate with improved physicochemical and pharmacokinetic properties. J Med Chem 2007;50:3528–39.
[18] Doherty EM, Fotsch C, Bannon AW, Bo Y, Chen N, Dominguez C, et al. Novel vanilloid receptor-1 antagonists: 2. Structure–activity relationships of 4-oxopyrimidines leading to the selection of a clinical candidate. J Med Chem 2007;50:3415–527.
[19] Brown A, Brown L, Brown TB, Calabrese A, Ellis D, Puhalo N, et al. Triazole oxytocin antagonists: identification of aryl ether replacements for a biaryl substituent. Bioorg Med Chem Lett 2008;18:5242–4.
[20] Bregman H, Chakka N, Guzman-Perez A, Gunaydin H, Gu Y, Huang X, et al. Discovery of novel, induced-pocket binding oxazolidinones as potent, selective, and orally bioavailable tankyrase inhibitors. J Med Chem 2013;56:4320–42.
[21] Lovering F, Bikker J, Humblet C. Escape from flatland: increasing saturation as an approach to improving clinical success. J Med Chem 2009;52:6752–6.
[22] Hoover TB. Water-solubility of PCB isomers. PCB Newslett 1971;3:4–5.
[23] Lee MC, Chian ESK, Griffin RA. Solubility of polychlorinated biphenyls and capacitor fluid water. Water Res 1979;13:1249–58.
[24] Ishikawa M, Kubota D, Yamamoto M, Kuroda C, Iguchi M, Koyanagi A, et al. Tricyclic pharmacophore-based molecules as novel integrin $\alpha_v\beta_3$ antagonists. Part 2: Synthesis of potent $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ dual antagonists. Bioorg Med Chem 2006;14:2109–30.
[25] Ishikawa M, Ajito K. Challenges toward acute ischemic diseases: preparation of integrin $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ dual antagonists. In: Torisawa Y, editor. Soyaku shien kenkyu no tenbo [Prospects for drug discovery support research]. Tokyo: CMC Publishing; 2008. p. 3–13.
[26] Ishikawa M, Hashimoto Y. Improvement in solubility in small molecule drug discovery programs focusing on dihedral angle and symmetry. J Synth Org Chem Jpn 2013;71:625–36.
[27] Valkó K. Application of high-performance liquid chromatography based measurements of lipophilicity to model biological distribution. J Chromatogr A 2004;1037:299–310.
[28] Fujita Y, Yonehara M, Tetsuhashi M, Noguchi-Yachide T, Hashimoto Y, Ishikawa M. β-Naphthoflavone analogs as potent and soluble aryl hydrocarbon receptor agonists: Improvement of solubility by disruption of molecular planarity. Bioorg Med Chem 2010;18:1194–203.
[29] Fujita Y, Yonehara M, Kitahara K, Shimokawa J, Hashimoto Y, Ishikawa M. Confirmation of molecular planarity disruption effect on aqueous solubility improvement of β-naphthoflavone analogs. Heterocycles 2011;83:2563–75.
[30] Kasuga J, Ishikawa M, Yonehara M, Makishima M, Hashimoto Y, Miyachi H. Improvement of water-solubility of biarylcarboxylic acid peroxisome proliferator-activated receptor (PPAR) δ-selective partial agonists by disruption of molecular planarity/symmetry. Bioorg Med Chem 2010;18:7164–73.

[31] Leach AG, Jones HD, Cosgrove DA, Kenny PW, Ruston L, MacFaul P, et al. Matched molecular pairs as a guide in the optimization of pharmaceutical properties: a study of aqueous solubility, plasma protein binding, and oral exposure. J Med Chem 2006;49:6672–82.
[32] Oguro Y, Cary DR, Miyamoto N, Tawada M, Iwata H, Miki H, et al. Design, synthesis, and evaluation of novel VEGFR2 kinase inhibitors: discovery of [1,2,4]triazolo[1,5-a]pyridine derivatives with slow dissociation kinetics. Bioorg Med Chem 2013;21:4714–29.
[33] Scully SS, Tang AJ, Lundh M, Mosher CM, Perkins KM, Wagner BK. Small-molecule inhibitors of cytokine-mediated STAT1 signal transduction in β-cells with improved aqueous solubility. J Med Chem 2013;56:4125–9.
[34] Schroeder GM, An Y, Cai Z-W, Chen X-T, Clark C, Cornelius LAM, et al. Discovery of N-(4-(2-amino-3-chloropyridin-4-yloxy)-3-fluorophenyl)-4-ethoxy-1-(4-fluorophenyl)-2-oxo-1,2-dihydropyridine-3-carboxamide (BMS-777607), a selective and orally efficacious inhibitor of the Met kinase superfamily. J Med Chem 2009;52:1251–4.
[35] Anderson DR, Meyers MJ, Vernier WF, Mahoney MW, Kurumbail RG, Caspers N, et al. Pyrrolopyridine inhibitors of mitogen-activated protein kinase-activated protein kinase 2 (MK-2). J Med Chem 2007;50:2647–54.
[36] Dublanchet AC, Ducrot P, Andrianjara C, O'Gara M, Morales R, Compere D, et al. Structure-based design and synthesis of novel non-zinc chelating MMP-12 inhibitors. Bioorg Med Chem Lett 2005;15:3787–90.
[37] Jones CD, Andrews DM, Barker AJ. Imidazole pyrimidine amides as potent, orally bioavailable cyclin-dependent kinase inhibitors. Bioorg Med Chem Lett 2008;18:6486–9.
[38] Brown DG, Urbanek RA, Bare TM, McLaren FM, Horchler CL, Murphy M, et al. Synthesis of 7-chloro-2,3-dihydro-2-[1-(pyridinyl)alkyl]-pyridazino[4,5-b]quinoline-1,4,10(5H)-triones as NMDA glycine-site antagonists. Bioorg Med Chem Lett 2003;13:3553–6.
[39] Fray JM, Bull DJ, Carr CL, Gautier ECL, Mowbray CE, Stobie A. Structure – activity relationships of 1,4-dihydro-(1H,4H)-quinoxaline-2,3-diones as N-methyl-d-aspartate (glycine site) receptor antagonists. 1. Heterocyclic substituted 5-alkyl derivatives. J Med Chem 2001;44:1951–62.
[40] Wang X, Chakrabarti PP, Ognyanov VI, Pettus LH, Tamir R, Tan H, et al. Trisubstituted pyrimidines as transient receptor potential vanilloid 1 (TRPV1) antagonists with improved solubility. Bioorg Med Chem Lett 2007;17:6539–45.
[41] Li Q, Chu DTW, Claiborne A, Cooper CS, Lee CM, Raye K, et al. Synthesis and structure – activity relationships of 2-pyridones: a novel series of potent DNA gyrase inhibitors as antibacterial agents. J Med Chem 1996;39:3070–88.
[42] Gavezzotti A. Molecular symmetry, melting temperatures, and melting enthalpies of substituted benzenes and naphthalenes. J Chem Soc Perkin Trans 1995;2:1399–404.
[43] Ishikawa M, Hiraiwa Y, Kubota D, Tsushima M, Watanabe T, Murakami S, et al. Tricyclic pharmacophore-based molecules as novel integrin $\alpha_v\beta_3$ antagonists. Part III: synthesis of potent antagonists with $\alpha_v\beta_3/\alpha_{IIb}\beta_3$ dual activity and improved water-solubility. Bioorg Med Chem 2006;14:2131–50.
[44] Wenglowsky S, Moreno D, Rudolph J, Ran Y, Ahrendt KA, Arrigo A, et al. Pyrazolopyridine inhibitors of B-RafV600E. Part 3: an increase in aqueous solubility via the disruption of crystal packing. Bioorg Med Chem Lett 2012;22:912–15.

CHAPTER

32

Chemical and Physicochemical Approaches to Solve Formulation Problems

Harvey Lieberman[1] *and N. Murti Vemuri*[2]

[1]Sanofi ,Waltham, MA, USA; [2]CMC Subject Matter Expert at BARDA/HHS, Conceptual MindWorks, Inc., Flemington, NJ, USA

OUTLINE

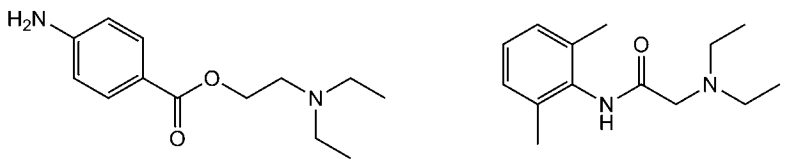
I. INTRODUCTION

The main goal of a medicinal chemist is to prepare active, selective, nontoxic compounds that can be developed into drugs. The chemistry designed into potent inhibitors can, however, have consequences for how they can be delivered, and it is generally the formulation scientist's task to develop a suitable formulation for animal and human use. This can often be challenging and may lead to huge expenses as a drug moves toward launch, especially if technologies need to be in-licensed or plants need to be adapted to the process. A better solution would be to fix potential problems at the molecular level, which would reduce long-term costs and offer more patent protection. As timelines for drug discovery and development are shortening, a stronger overlap between medicinal chemists and formulation scientists exists, allowing for early assessment of the drugability of candidate compounds.

II. STABILITY

A. Chemical Stability

The chemical stability of a drug can affect its performance in several manners. First, instability of a drug substance leading to degradation can lower its effectiveness, reducing the quantity of the active pharmaceutical ingredient that is administered to a patient. Second, degradation of a drug may lead to toxic products that can have adverse effects. Degradation of a drug substance may occur through exposure to its environment or



Procaine Lidocaine

FIGURE 32.1 Stereoelectronic shielding in lidocaine.

interactions with excipients, and care should be taken to avoid those that might accelerate the process. On the other hand, careful choice of excipients or manipulation of the physical form may retard a potential degradation.

1. *Decomposition of Drugs by Chemical Means*

A. HYDROLYSIS

Hydrolysis is the most common form of drug degradation due to the prevalence of water in isolation procedures, formulation strategies, and the large number of functional groups that can undergo hydrolysis. Hydrolysis generally occurs via an acid- or base-catalyzed mechanism but can also occur under neutral conditions where water can act as a base. The carbonyl function of esters, lactones, amides, lactams, carbamates, and imides are susceptible to hydrolysis, as are imines.

Chemical methods to deal with hydrolysis that may occur *in vivo* include the use of steric shields. For example, the use of a bulky alkyl group close to the functional group hinders the approach of a nucleophile or enzyme, and thus reduces the likelihood of hydrolysis. Electronic shielding can be used to replace a functional group with a more stable bioisostere. For example, if a labile ester function is replaced with a urethane or amide, the nitrogen can feed electrons into the carbonyl group, making it less reactive. This, in turn, increases chemical and metabolic stability. Stereoelectronic modification may also be used to increase stability without compromising activity. Changing the ester group in the readily hydrolyzed procaine to an amide and introducing aromatic methyl groups to hinder the attack of the carbonyl group significantly reduce the rate of hydrolysis. The resulting lidocaine is a longer-acting local anesthetic than procaine (Figure 32.1).

Several physicochemical options are available to deal with hydrolysis of a drug substance. If the dosage form is a liquid, then pH adjustment can improve the shelf life, although care should be taken that the drug's solubility and stability of excipients are not compromised. Alternative measures rely on protecting the drug from its aqueous environment. Use of surfactant-based systems such as micelles can reduce the hydrolysis of hydrophobic drugs. Alternatively, preparing a suspension reduces the amount of drug in solution and hence may increase stability. A lyophilized product will remove water and significantly reduce the rate of hydrolysis. An example of this approach can be found in the formulation of disodium 1-[2-(carboxylato)pyrrolidin-1-yl]diazen-1-ium-1,2-diolate (PROLI/NO), a nitric oxide donor prodrug which has a two second half-life for nitric oxide release at physiological pH [1]. PROLI/NO readily hydrolyzes but is stabilized by lyophilizing the prodrug and storing at -20°C. Stability studies show that no evidence of decomposition was observed over a 140-day period. A further approach to stabilizing a hydrolysable drug is to formulate it in a completely nonaqueous system. Dong and coworkers [2] demonstrated that N-epoxymethyl-1,8-naphthalimide could be stabilized in a combination of 70 percent Cremophor EL and 30 percent ethanol (v/v) for four months. Upon diluting with saline, less than 5 percent of the drug hydrolyzed over a four hour period. In contrast, this drug fully degrades in water within seventy-two hours.

For solid dosage forms, the choice of excipient and processing conditions can have a marked effect on the rate of hydrolysis. The choice of crystal or salt form may also affect the rate by altering surface groups and increasing lattice strength. Hydrolysis can also be used to a medicinal chemist's advantage. Targeting a pH in the GI-tract may make for an attractive prodrug. Bio-compatible polymers, which undergo hydrolysis, are used in slow-release implants.

B. OXIDATION

Oxidation may occur through autoxidation in which molecular oxygen propagates the reaction. Autoxidation proceeds without a catalyst and is generally slow. It is, however, extremely important, due to the difficulty in totally excluding oxygen from a pharmaceutical process and due to the permeability of oxygen through many materials. Oxidation may also occur through a catalyzed reaction with heavy metals, light, or heat. Groups prone to oxidation include alcohols, aldehydes, phenols, amines, olefins, ethers, thiols, and thioethers.

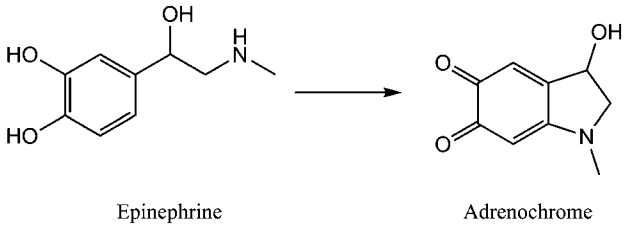


FIGURE 32.2 Oxidation of epinephrine to adrenochrome.

TABLE 32.1 Measured Levels of Hydroxyperoxides in a set of Common Pharmaceutical Excipients

Excipient	# Lots tested	Average HPO (nmole/g)	High HPO (nmole/g)	Low HPO (nmole/g)
PVP	5	7300	11000	3600
PEG 400	4	2200	3300	1000
PS80	8	1500	4600	180
HPX	21	300	890	50
Poloxamer	7	30	50	10
PEG solid	4	20	40	<10
MCG	3	<10	<10	<10
Microcrystalline cellulose	5	<10	10	<10
Mannitol	5	<10	<10	<10
Lactose	5	<10	10	<10
Sucrose	5	<10	20	<10

Reprinted with permission from [4].

Epinephrine, a drug used to treat cardiac arrest and anaphylaxis, undergoes oxidation to produce adrenochrome through a ring-closure mechanism (Figure 32.2). A review [3] of studies performed on injectable solutions of epinephrine suggests that temperature and pH have the greatest impact on stability.

Oxidation may occur in the solid state as well as in solution. For solid-state reactions, factors such as crystal packing, morphology, and surface groups contribute to the rate of oxidation, since oxygen has to diffuse through the crystal lattice. Oxidation may occur more readily in amorphous systems due to the increase in molecular mobility, and the presence of a small quantity of amorphous material, which may be generated during the formulation process, could lead to sufficient quantities of a degradant to cause a failure to meet a QC specification.

Since oxidation may occur in the dissolved or solid state, the presence of excipients in a formulation, as well as the manner in which the drug product is packaged and stored, may also contribute to the degradation of an active pharmaceutical ingredient. Care should be taken in formulating drugs that are prone to oxidation with buffers, since all buffers contain trace amounts of transition metals. Hydroperoxides (organic peroxide or H_2O_2) are present in many common excipients and can promote oxidative degradation if formulated with a susceptible drug. Table 32.1 [4] lists some common pharmaceutical excipients and their hydroperoxide levels. By way of example, Hartauer and coworkers [5] identified an N-oxide degradation product of raloxifene hydrochloride and managed to trace its origin back to the presence of peroxide impurities in povidone and crospovidone excipients.

If the risk of oxidation is related to impurities in excipients, it can be mitigated by increasing the drug concentration, as a dilute system, which contains proportionally more impurity per drug, undergoes a faster degradation. Metal impurities in the excipients may be reduced by using chelating agents such as EDTA or citric acid. Antioxidants may also be used to reduce the rate of oxidation. Antioxidants may act as oxygen scavengers such as chain terminators (molecules with weak X-H bonds that are attacked during the chain process and stop it prematurely; thiols and phenols are common chain terminators) or sacrificial molecules (those that are preferentially oxidized; ascorbic acid is an example). Another type of antioxidant is a radical trap or scavenger, such as butylated hydroxyanisole (BHA) and butylated hydroxytoluene (BHT), which work by stabilizing free radicals through their aromatic functions (Figure 32.3).

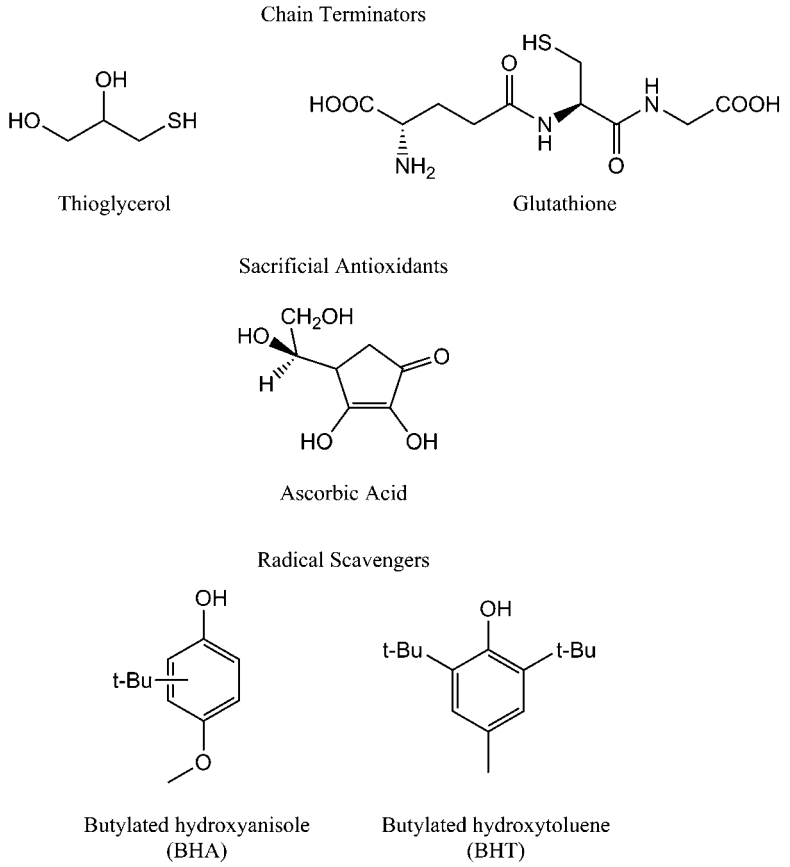


FIGURE 32.3 Common antioxidants.

C. PHOTOLYSIS

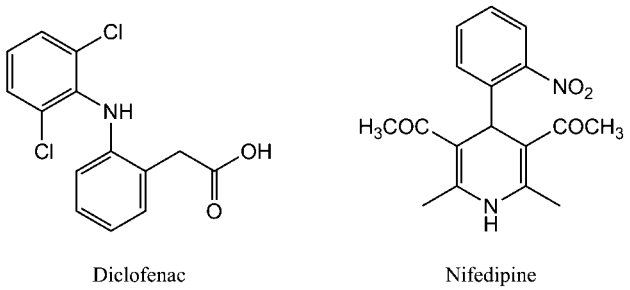
Energy absorbed in the UV or visible parts of the electromagnetic spectrum can be another source of pharmaceutical degradation. Groups that can undergo degradation by photolysis include ketones (via hydrogen abstraction), nitro compounds (via rearrangement), conjugated alkenes (via E-Z isomerization), aromatics, and heterocycles.

The simplest approach to stabilizing a photosensitive drug is to exclude it from light by placing the drug product in a protective package. Since photochemical reactions often lead to oxidative degradation, an inert atmosphere, chelating agents, and antioxidants can be used together. Complexation can also affect photoreactivity, and the use of cyclodextrins has been shown to promote, reduce, or have no effect on photostability [6].

Due to the photolabile nature of chlorine, chloroaromatic compounds can cause phototoxicity. Diclofenac is known to undergo photodegradation at the chlorine atoms, but this can be retarded by complexing the drug substance with 2-hydroxypropyl-β-cyclodextrin (Figure 32.4) [6]. Nifedipine (Figure 32.4) is highly photosensitive and has been shown to fully decompose in about five days under room light. Groof and coworkers demonstrated how the rate of decomposition is related to the physical form, with amorphous nifedipine degrading at about double the rate of its crystalline counterpart [7]. Complexing this drug substance with cyclodextrins reduced the rate of photodecomposition by up to 33-fold [8].

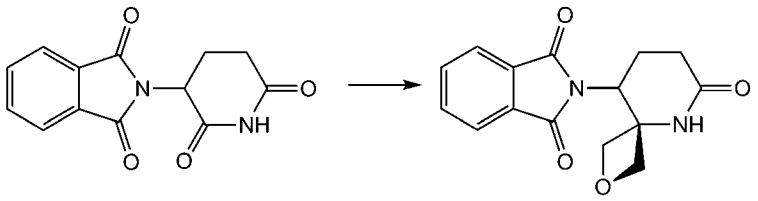
D. RACEMIZATION AND CHIRAL INVERSION

The trend for the past twenty years has shifted from production of pharmaceutical racemates to single-enantiomer drugs. Single enantiomers tend to offer reduced toxicity, lower doses, and lower development costs. Chemical control of racemization can be accomplished using steric bulk, but this can be difficult if the chiral center needs to be exposed for activity. Thalidomide is known to undergo conversion *in vivo*, leading to a racemic mixture of which the R-isomer is a sedative and the S-isomer is a teratogen. A recent study [9] in which the amide function was replaced with an oxetane suggests that the new oxetano-thalidomide (Figure 32.5) has a greater stability in plasma and could potentially limit racemization.



Diclofenac Nifedipine

FIGURE 32.4 Examples of molecules that readily undergo photolysis.



Thalidomide Oxetano-thalidomide

FIGURE 32.5 Stabilizing an amide function by replacement with an oxetane.

An alternative method of stabilization is to contain the molecular species and protect it from racemization. Wei and coworkers [10] demonstrated that the use of layered solids as a molecular container for L-tyrosine could inhibit its racemization under sunlight, UV, and thermal conditions.

In contrast to *in vivo* racemization (which leads to both enantiomers), chiral inversion (which accounts for the inversion of one chiral form to another) leads to a single entity. Ibuprofen, for example, is administered as a racemic mixture of the R- and S-enantiomers, but *in vivo* the R-enantiomer undergoes conversion to the more active S-isomer.

E. PREDICTION OF STABILITY

As a drug enters development it is useful to combine experimental degradation testing with *in silico* prediction in order to understand the mechanisms of degradation better and thus design a more robust formulation. Methods based on quantum mechanics [11] and expert systems [12] are now established within Sanofi. We have found the results to agree reasonably well with experiments despite the limitations of the computational methods, and the data are used to help interpret forced degradation studies.

2. *Influence of the Solid State on Chemical Stability*

Broadly speaking, solids can be classified into those with short- and long-range order (i.e., crystalline) or those with short-range but no long-range order (i.e., amorphous). Traditionally, active pharmaceutical ingredients are manufactured as crystalline solids due to their increased stability and relative ease of purification over the amorphous state. There is, however, an attraction in preparing pharmaceuticals in the amorphous state, particularly due to the fact that these high-energy materials can demonstrate a pronounced enhancement in solubility [13,14].

A. CHEMICAL STABILITY OF AMORPHOUS SOLIDS

Although a few small-molecule [15] and macromolecular exceptions [16] exist, amorphous materials tend to demonstrate a greater chemical instability compared to their crystalline counterparts. The presence of even a small quantity of water can increase this instability significantly [17,18].

Chemical reactivity of amorphous materials has been shown to be related to molecular mobility. Increased molecular mobility leads to enhanced chemical degradation [19]. Studies on the degradation of spirapril hydrochloride illustrate that the crystalline form is much more stable than the amorphous form and has far less molecular mobility as determined by ^{13}C solid-state NMR spectroscopy [20].

Δ^9-tetrahydrocannabinol (THC) is a highly unstable drug, susceptible to degradation by light, heat, and oxidation, with greater than 80 percent reported lost when stored for one month at 25°C and 0 percent relative

humidity [21]. The chemical stability of this amorphous drug was improved by blending it with polyethylene oxide (PEO) or hydroxypropylcellulose (HPC) along with processing aids using hot-melt casting methods [21,22]. The resulting PEO-vitamin E succinate (VES) patch, containing ascorbic acid as an antioxidant, underwent 5.8 percent degradation as opposed to 31.6 percent in the control after storage for two months at 40°C.

B. CHOICE OF POLYMORPHIC FORM

Polymorphs have different melting points, densities, and molecular interactions, and will therefore demonstrate different chemical stabilities. Tamoxifen citrate, for example, crystallizes in two polymorphic forms. Upon exposure to visible/UV light, one form is stable while the other one is unstable [23].

3. *Presence of Water*

The presence of water may affect the chemical stability of an active pharmaceutical ingredient (API) in a solid dosage form and may be present through the drug substance itself (via hydrates), excipients, processing procedures, or humidity. Water can lead to hydrolysis, such as in the case of aspirin. It can have a plasticizing effect, changing the mobility within the formulation, particularly when amorphous materials are present, and it can be adsorbed on the surface, leading to partial dissolution and potential degradation.

In order for water to affect the stability of the drug, it must be available. Hydrated APIs contain water that is bound in the crystal and is not free, but water may be released from the crystalline state through operations such as grinding. Ampicillin trihydrate, for example, demonstrates a degradation rate that is proportional to grinding time [24]. Moisture scavengers, such as silica, may be co-formulated in order to act as a desiccant and preferentially absorb any free water.

4. *Salt Form*

Salt forms are typically chosen in drug discovery and development in order to improve drug solubility characteristics, but salt formation may also have fortuitous implications on stability of a drug substance. For example, formation of a salt of an amine can protect it from N-oxide formation via oxidation in the solid state.

Salts, however, tend to be more hygroscopic than their free base or free acid counterpart due to the ionic nature of the solid material, and absorption or adsorption of water can lead to degradation of the API. The counterion will also affect the microenvironmental pH of the salt (the pH of the sorbed water layer at the surface of the solid) and can therefore impact its interaction with excipients. Procaine is known to undergo pH-dependent hydrolysis in solution. In a study of eleven crystalline salts, it was found that the rate of hydrolysis in the solid state was dependent on the pKa of the counterion [25].

5. *Use of Cyclodextrins to Improve Drug Stability*

Host–guest interactions will be covered later in this chapter, but it is worth highlighting that cyclodextrins have been used to improve chemical stability. Many examples illustrate how complexing with cyclodextrins can help protect pharmaceutical ingredients that are susceptible to photodegradation (see section II.A.1.c), but they can also improve stability with respect to hydrolysis and oxidation. Doxorubicin hydrochloride, for example, an antineoplastic agent, undergoes acid-catalyzed hydrolysis as well as photodegradation. Incorporation into γ-cyclodextrins reduced the rate of degradation by up to 10-fold [26]. Hydroxypropyl-β-cyclodextrin has been shown to improve the stability of both sulfamethazole and trimethoprim with respect to oxidation [27].

B. Physical Stability

1. *Influence of Polymorph on Physical Stability*

Polymorphism, namely the ability of a substance to exist in multiple crystal forms, is ubiquitous in the pharmaceutical industry. It has been estimated that as many as 80 percent of drug substances are polymorphic [28], echoing McCrone's famous claim that "It is at least this author's opinion that every compound has different polymorphs and that, in general, the number of forms known for a given compound is proportional to the time and money spent in research of that compound." [29] Polymorphs of a compound possess identical chemical behavior once in solution but can have widely different physical behavior in the solid state. Properties such as melting point, solubility, morphology, and surface free energy may differ [30].

At a given temperature and pressure, one polymorphic form will always be the most stable. This form may be the most stable form up to the melting point (monotropic) or it may undergo a reversible phase transition to

another form (enantiotropic) prior to the melt. It is common for the most stable polymorphic form to be chosen after extensive experimental screening due to development challenges associated with choosing a less stable form (e.g., polymorph conversion under typical processing conditions such as grinding, compression, or storage).

It is difficult to estimate the proportion of drugs marketed as metastable polymorphs. Some, such as chloramphenicol palmitate and norfloxacin, are intentionally marketed as metastable forms in order to increase a compound's solubility or bioavailablility [31,32]. Others, such as ritonavir, were thought to exist in a single form until a new, more stable form was discovered [33].

Several computational models exist to predict what potential polymorphs exist for a given molecular structure, but these methods cannot determine how to grow a particular form or even whether it can be grown under normal laboratory conditions. Prediction of crystal structure is a complex process [34,35] and, although improvements have been made, most models tend to suffer from overestimating too many low-energy possibilities and do not take entropic effects into consideration. Rather than a straight predictive tool, these methods are useful to help understand and direct crystallization experiments or aid in structure determination using X-ray powder data. For a more extensive appreciation of polymorphism and its role in the pharmaceutical industry the reader is directed to the excellent books authored by Bernstein [36], Brittain [37], and Hilfiker [38].

2. *Physical Stability of Salt Forms*

Disproportionation occurs when a salt converts back to a free acid and base via proton transfer. It can take place when small quantities of water are present in contact with the solid or if the solid is mixed with excipients that can affect the micro-environmental pH in the solid API [39]. Disproportionation can often be overlooked, as the methods commonly used to investigate stability—such as LC-MS—will be unable to distinguish between a salt form and its individual components. Studies on the maleate salt of a basic drug (dibenzo[b,f]oxepen-10-ylmethyl-methyl-prop-2-ynyl-amine hydrogen maleate) demonstrated that it degraded in a tablet formulation upon storage under accelerated stability testing and that this was triggered by the disproportionation of the salt [40]. The tablet was stabilized by adding citric acid to the formulation, which reduced the micro-environmental pH such that the conversion of salt to free base was not favored.

3. *Stabilizing Amorphous Solid Dispersions*

In recent years, more focus has been given to the amorphous state and how to ensure long-term stability [41,42]. Many new crystalline drug-development candidates possess less-than-desirable aqueous solubility, and the use of an amorphous form has the potential to increase the apparent solubility and dissolution rate. Amorphous systems are thermodynamically unstable and will convert to a crystalline form over time. In order to stabilize an amorphous phase, the thermodynamically-favorable crystallization process needs to be retarded.

One way to formulate an amorphous form and increase its physical stability is to form an amorphous solid dispersion (ASD). In this formulation the amorphous material is combined with a water-soluble polymer matrix along with other excipients, generally via a hot-melt extrusion or spray drying method. The polymer inhibits crystallization by raising the overall T_g (glass transition temperature) of the dispersion relative to that of the API, thus reducing molecular mobility. Polymers may also interact directly with the API through specific interactions that inhibit crystallization, such as hydrogen bonding.

Several of the factors that impact the physical stability of an ASD are discussed below.

Molecular mobility: The stability of amorphous materials will be impacted by molecular mobility, which can be characterized by global mobility (α relaxations) and local mobility (β relaxations). Global mobility is the cooperative motion of molecules as the temperature approaches T_g whereas local mobility involves movement of the whole or a part of a molecule in a uncooperative manner and tends to occur on a much faster timescale than global mobility. ASDs play a key role in reducing molecular mobility in two ways: by increasing the T_g of the dispersion and by engaging in preferential interactions with an API and hence reducing the possibility of forming stable nuclei. It has been suggested that an amorphous pharmaceutical solid should be stored at least 50°C below T_g in order to reduce molecular mobility, so that it remains stable for a sufficient length of time [43]. Nifedipine-PVP and phenobarbital-PVP solid dispersions have been investigated using solid-state NMR, and results suggest that hydrogen bond interactions between drug molecules and polymer contribute to the increase in physical stability [44].

Presence of water: The crystallization of an amorphous form can be accelerated by the absorption of water, which even at low levels can act as a plasticizer and lower T_g. Water may also reduce the drug−polymer miscibility, leading to a drug-rich phase that could promote crystallization [45].

TABLE 32.2 A Selection of Marketed Drugs Containing ASD

Drug	Polymer matrix	Indication
Afeditab (nifedipine)	Poloxamer/PVP	Anti-anginal, hypertensive
Certican/Zortress (everolimus)	HPMC	Anti-emetic, neuropathic pain
Cesamet (nabilone)	PVP	Anti-emetic
Gris-PEG (griseofulvin)	PEG	antifungal
Intelence (etravirine)	HPMC	HIV
Isoptin-SRE (verapamil)	HPC/HPMC	Hypertension, angina pectoris
Kaletra (ritonavir, lopinavir)	PVP/PVA	HIV
ProGraf (tacrolimus)	HPMC	immunosuppressive
Sporanox (itraconazole)	HPMC	antifungal

Temperature: Increasing temperature can reduce the stability of an ASD by increasing molecular mobility and weakening drug−polymer interactions. An investigation into the temperature effect on felodipine-PVP ASD [46] suggests that hydrogen-bonding interactions were weakened as temperature was increased, but the effect was found to be reversible.

Drug loading: The ratio of drug substance to polymer can have a pronounced effect on the stability of the dispersion. In ASDs of efavirenz containing 13−17 percent PVP and stored at 72°C / 35 percent RH, a 45-fold increase in crystallization rate was observed upon reducing the polymer concentration from 17 to 13 percent [47].

Although typically more challenging in their development than a crystalline form, ASDs may be found in a number of marketed drugs (Table 32.2).

III. BIOAVAILABILITY

A. Solubility

Poor aqueous solubility of a drug substance is a common consequence of modern drug-discovery programs. The adoption of high-throughput screening coupled with parallel medicinal chemistry approaches have led to the preparation and evaluation of large numbers of compounds. The downside to this approach is that, although potent compounds are identified, they tend to be of high molecular weight and lipophilicity. This, in turn, can lead to compounds with poor aqueous solubility. This is particularly true for specific classes of compounds, such as kinase inhibitors, whose chemistry favors molecules that are flat and aromatic with hydrogen-bonded hinge binders. This is same chemistry also reduces aqueous solubility. Many chemical and physicochemical approaches can be used to improve the solubility of a drug substance while maintaining its activity, several of which will be discussed below.

1. Covalent Attachment of Solubilizing Moieties

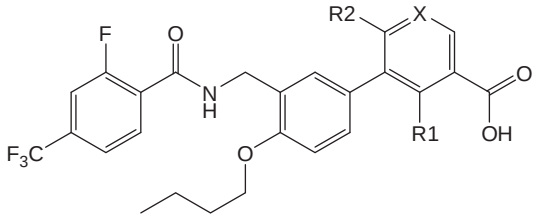
Increasing solubility by chemical means without sacrificing potency or grossly affecting PK parameters is generally the most cost-effective process and negates the time and efforts of having to develop a nonstandard formulation later. This is covered in detail in chapter 30.

2. Disruption of Crystal Packing by Changing the Molecule

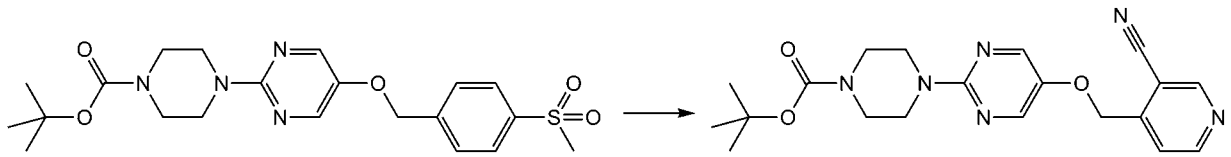
In order for a crystalline solid to dissolve, the interactions between neighboring molecules must first be disrupted. The two main interactions found within strong crystal lattices are intermolecular hydrogen bonds and short-range dispersion forces, and these are particularly prevalent in compounds containing large numbers of hydrogen-bonding donors and acceptors, planar conformations, and high degrees of symmetry.

Kasuga and coworkers [48] improved the aqueous solubility of two PPARδ selective partial agonists by disrupting the planarity and symmetry in a target molecule. Of the compounds investigated, structure **25** (Table 32.3) demonstrated 5-fold improved activity over its parent compound (structure **2**) with adequate solubility in phosphate buffer (22 μg/mL). Structure **27**, although slightly lower in activity as compared to the parent (3-fold less), was at least 2,700 times more soluble in phosphate buffer. These results are similar to those experienced during a Sanofi optimization program where two substituent positions on a planar kinase inhibitor were systematically altered, leading to a

TABLE 32.3 Activities and Solubilities of Selective Compounds Substituted at the 2-Position



Structure	R1	R2	X	PPARδ EC_{50} (nM)	Solubility in phosphate buffer pH 7.4 (mg/mL)	MPT (°C)
2	H	H	CH	29	< 0.001	177–8
25	F	F	CH	5.7	0.022	177
27	Me	Me	N	76	2.70	104–6



EC_{50} 65 nM
Solubility 0.03 μM

EC_{50} 21 nM
Solubility 32 μM

FIGURE 32.6 Increase in solubility by reduction of intermolecular interactions.

50,000-fold increase in aqueous solubility. In this case, the compound eventually chosen to move forward demonstrated thirteen times the solubility of the parent with minimal loss of potency and improved PK parameters.

In another example, Scott and coworkers [49] utilized small-molecule crystallography to identify areas of strong packing in the crystal structure of a BOC-protected potent lead agonist of GPR119. The resulting chemical switch from an aryl sulphone to a 3-cyanopyridyl motif led to an increase in potency coupled with an increase in solubility (Figure 32.6).

3. *Salt Forms*

Preparation of salt forms is a common first-choice approach to increase the solubility and dissolution rate of chemical species 50 (salts account for approximately 50% of marketed drugs [51]) and is the preferred method for increasing solubility for parenteral dosing [52]. Salts are formed when an ionizable drug is combined with a suitable counterion via proton transfer. The resulting solid material often has improved dissolution characteristics which can be attributed to the change in pH at the diffusion layer (the layer of solvent at the surface of the solid) of the dissolving solid. Even if the drug precipitates in its unionized form, it's most likely to so as microcrystalline or amorphous particles which have enhanced dissolution as compared to the crystalline free form. There are occasions when it is desirable to reduce the solubility of a drug, such as when it is formulated as a suspension. In this case more hydrophobic, larger counterions such as napsylate and embonate may be used. Care should be taken when assessing the physicochemical and biopharmaceutical properties of salts. Hydrochloride salts may experience a common-ion effect *in vivo* and preparation of salts from specific counterions can lead to toxicity issues, for example maleic acid is known to demonstrate toxicity in dogs [53] and mesylate esters may be generated in the formation of methylsulfonic acid salts. Salt screening is now a common strategy in late stage drug discovery or early in a development program [54].

4. *Co-Crystals*

Whereas the formation of salts is an established method for changing the physical nature of a substance while retaining its chemical activity, the interest in pharmaceutical co-crystals is much more recent. (In 2013 the FDA released a document on the regulatory classification of pharmaceutical co-crystals [55].) The definition of a

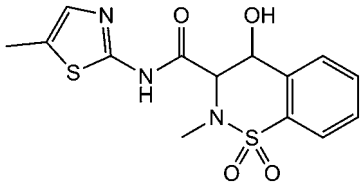


FIGURE 32.7 Meloxicam.

co-crystal is not as clear as that of a salt, but broadly speaking a co-crystal is a solid material consisting of two or more species that do not engage in proton transfer. The number of potential salts of an API that can be formed is somewhat limited by the list of salt formers and the acidic/basic nature of the active itself. A general rule of thumb is that the difference in pKa between the acid and base should be at least two units. By contrast, the number of potential co-crystals is much larger.

Weyna and coworkers [56] studied the solubility and dissolution of twelve co-crystals of meloxicam, a nonsteroidal anti-inflammatory drug prescribed for rheumatoid arthritis, osteoarthritis, postoperative pain, and fever (Figure 32.7). Meloxicam has a slow onset of action, and it is thought that an increase in solubility would enable the drug to act faster. *In vitro* dissolution experiments suggest that all twelve co-crystals exhibit faster early dissolution rates, with four of them (1-hydroxy-2-naphthoic acid, salicylic acid, succinic acid, and maleic acid) exhibiting a 10-fold or greater increase within the first two hours. *In vivo* PK experiments performed on rats demonstrated an increase in serum concentration as compared to meloxicam for most of the co-crystals within the first hour. In particular, the co-crystals of 1-hydroxy-2-naphthoic acid and salicylic acid achieved a concentration of 30.2 μg/mL within 21 and 11 minutes respectively—something that requires two hours for meloxicam alone. This demonstrates the potential for the development of co-crystals of meloxicam for mild- to medium-level acute pain.

5. Choice of Crystal Form

From the perspective of increasing bioavailability, the most attractive polymorph is the one with the lowest free energy, which would, in turn, possess the greatest solubility. The problem with this approach is that this same polymorphic form is also the least stable and will tend to convert to the most stable form where possible, often stepwise through other forms [57]. If a less stable polymorphic form could be stabilized, the gain in solubility is generally not great and up to a 2-fold increase is likely [13].

6. Use of the Amorphous Form

Amorphous materials can offer substantial improvement in apparent solubility and dissolution over crystalline counterparts. The determination of the exact enhancement is challenging due to the fact that an amorphous system is far from equilibrium [13,58]. In theory, several orders of magnitude are possible, but generally experimentally-measured enhancements of 1−50 fold have been reported. One challenge in using amorphous systems is their inherent instability and driving force to crystallize. A second is the possibility of the amorphous phase crystallizing once it has been administered *in vivo*.

A common method to stabilize an amorphous phase is to create an ASD (see section II.B.3). Vemurafenib is a practically insoluble drug with an aqueous solubility of 1 μg/mL. Preparation of an ASD of vermurafenib and hypromellose acetate succinate (HPMCAS) provided a faster dissolution, leading to a drug concentration of thirty times the crystalline drug solubility. Bioavailability was also improved by 4−5-fold [59].

7. Host/Guest Assemblies

Another method of altering the molecular environment involves forming a complex between the API and a host structure. Cyclodextrins (CD) have been used to increase the solubility of an API for many years. Cyclodextrins (Figure 32.8) are cyclic oligosaccharides that contain a hydrophobic core that can interact with an API and a hydrophilic surface that allows for solvation of the complex. Cyclodextrins can interact in a noncovalent manner with appropriately sized molecules and offer the potential of increased solubility and stability [60]. Furthermore, the chemistry of cyclodextrins may be modified to affect complexation.

Many cyclodextrins have been reported in the literature. The most commonly used are α-, β-, and γ-CDs containing 6, 7, and 8 glucopyranose units and possessing cavity sizes of 4.7−5.3, 6.0−6.5, and 7.5−8.3 Å respectively. Cyclodextrins have the potential of greatly enhancing solubility of poorly soluble drugs.

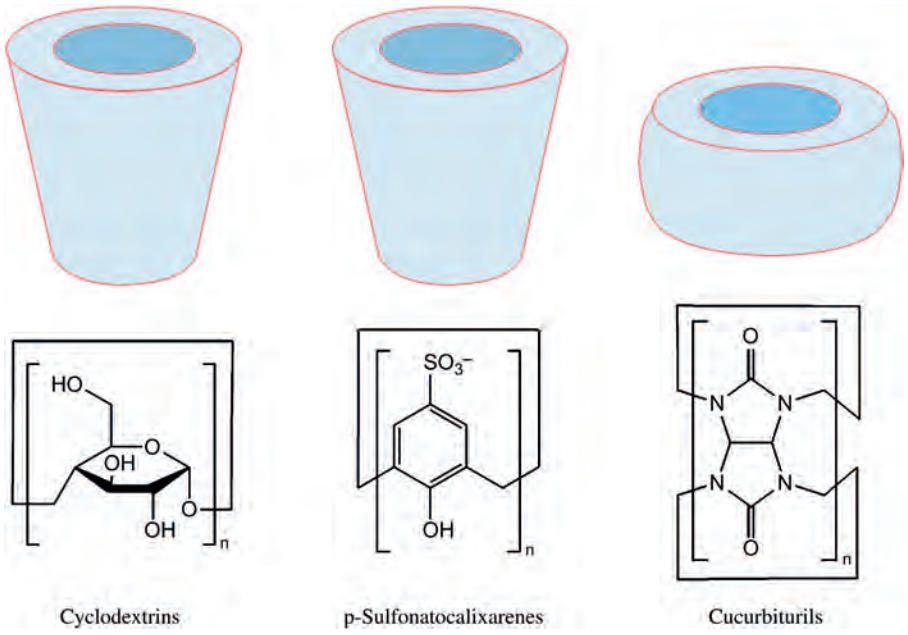


FIGURE 32.8 Examples of three classes of macrocycle hosts.

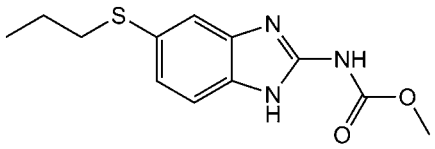


FIGURE 32.9 Albendazole.

The solubility of Paclitaxel, for example, can be increased 99,000 times when complexed with 50 percent dimethyl- β-cyclodextrin [61].

Cyclodextrins have been extensively studied in their use for drug delivery, but they are not the only class of macrocycles used in this manner. Calixarenes and cucurbiturils (Figure 32.8) are two other types of macrocycles which have been effectively used to increase the apparent solubility of poorly soluble drugs.

Calixarenes are basket-like structures formed from 4-substituted phenols, the sulfonated analogs being among those most studied [62]. Calixarenes, like cyclodextrins, can solubilize hydrophobic moieties but with additional ability to form hydrogen bonds through hydroxyl or, in the case of the sulfonated calixerines, sulfonato groups. The most commonly used calixarenes are calix[4]arenes, calix[6]arenes, and calix[8]arenes containing 4, 6, and 8 units and possessing cavity sizes of 3.0, 7.6, and 11.7 Å respectively.

Cucurbit[n]urils are a third class of macrocycles comprised of cyclic glycouril units bridged by methylene groups. They are barrel shaped and can bind neutral or protonated guest molecules. Cucurbit[n]urils have been studied in sizes from n = 5 to n = 10, leading to portal openings of 2.4−11.0 Å. Work carried out by Zhao and colleagues [63] illustrated the solubility enhancement that can be gained from using cucurbiturils, increasing the aqueous solubility of albendazole (Figure 32.9) from 3 μM to 6 mM at pH 6.6, a 2,000-fold gain.

8. *Other Molecular Interactions*

Hydrotropes are water-soluble, surface-active compounds which can significantly affect the solubility of poorly soluble drugs, most likely due to the formation of organized assemblies within solution. The issues encountered with traditional surfactant solubilization, such as emulsification, are not observed for hydrotropes due to the small hydrophobic part of the molecule. Some hydrotropes may aggregate in a stepwise manner, although many show no signs of self-aggregation even at high concentrations. Rasool and coworkers [64] studied the effect of nicotinamide (Figure 32.10) on various poorly soluble drugs and demonstrated a significant increase in solubility of up to 600-fold (Figure 32.11). Kim and coworkers [65] studied the solubilizing ability of N,N-diethylnicotinamide and N,N-dimethylbenzamide (Figure 32.10) on a series of structurally diverse drugs and

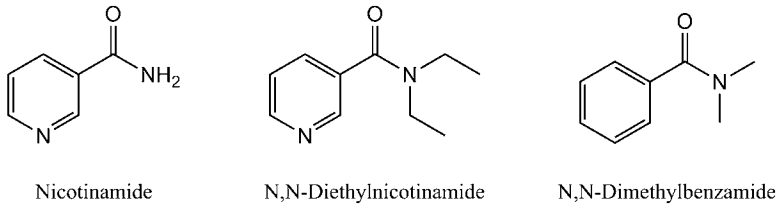


FIGURE 32.10 Examples of hydrotropes.

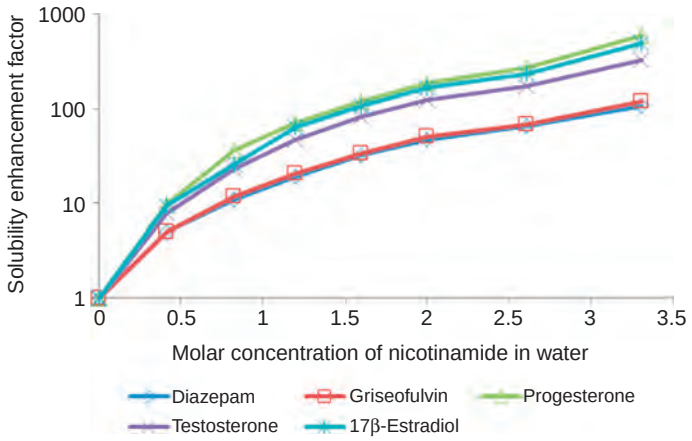


FIGURE 32.11 Solubility enhancement observed in using the hydrotrope nicotinamide. *adapted from [64].*

demonstrated solubility enhancements of 1,000−10,000-fold were achievable. One potential limitation of hydrotropes is the necessity of using high concentrations, which are often several moles in magnitude.

Considerable research has focused on attaching drugs to polymers in order to improve pharmacokinetic properties. Recently, dendrimers have emerged as a technology that can be used to improve the solubility of poorly soluble drugs. Dendrimers are highly branched macromolecular structures that provide many reactive surface groups, allowing for a high drug payload. They consist of a small core upon which multiple generations are built, and the resulting 3D structure is highly ordered. Niclosamide is a practically insoluble drug (aqueous solubility 230 ng/mL) used to treat tapeworms. Complexing with the PAMAM dendrimer increased its solubility by up to 6,178 times at the third generation [66], an increase greater than that observed by complexing with cyclodextrins, calixarenes, or mixtures of cyclodextrins and calixarenes [67].

B. Permeability

In order to increase absorption of a poorly permeable drug, either the drug or the membrane itself can be altered. Permeation enhancers generally work by affecting the properties of the membrane. The absorption of drugs by the transcellular route can be enhanced by altering the structure of the cell, and drugs absorbed via the paracellular route can be enhanced by opening the tight junctions. In this section, we will be more concerned with the properties of the drug and how its absorption can be improved without affecting the membrane structure.

1. Modifying Lipophilicity while Preserving Activity

The main routes for orally absorbed drugs to cross the intestinal membrane are passive diffusion or active transport, with passive diffusion being the most common mechanism. Passive diffusion may be accomplished by diffusion at the tight junctions between the intestinal enterocytes (paracellular) or through the lipid cell membrane itself (transcellular). Those drugs that undergo paracellular diffusion generally have low molecular weight (<250 Da) and are hydrophilic.

Relatively, the surface for transcellular diffusion is much greater than that for paracellular, and compounds that are absorbed via paracellular diffusion may demonstrate dose- and absorption-site-dependent differences [68]. It can, therefore, often be advantageous to increase transcellular uptake at the expense of paracellular.

One method is to increase the lipophilicity of a compound by replacing a hydrophilic group with a bioisosteric equivalent. Fluorine, a common isosteric replacement for hydrogen, can be used to increase lipophilicity; replacement

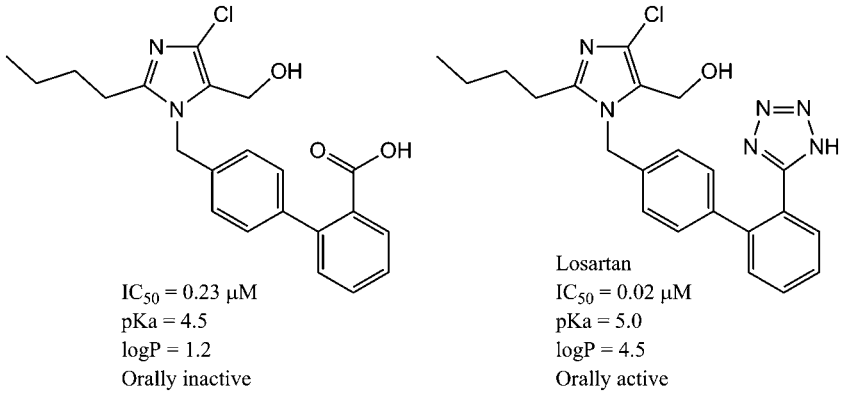


FIGURE 32.12 Losartan, an example of the use of a tetrazole bioisostere to increase lipophilicity.

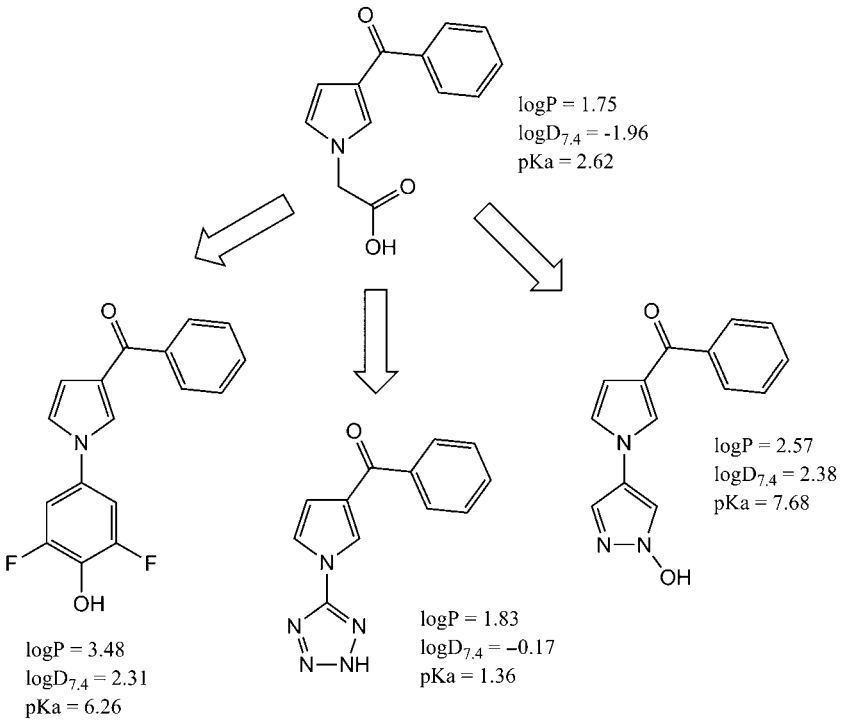


FIGURE 32.13 Effect of carboxylic acid group replacement on lipophilicity and pKa.

of a methyl with a trifluoromethyl group can significantly affect lipophilicity. A number of functional groups can be used to replace a carboxylic acid in order to increase its membrane permeability. Common groups include hydroxamic acids, sulfonates, and acyl-sulfonamides, as well as heterocyclic rings such as tetrazoles and oxo-thiazoles. Losartan is an example of the use of a tetrazole to increase lipophilicity [69]. Replacement of a carboxylic acid group by a tetrazole (Figure 32.12) increased potency and lipophilicity and allowed for the development of an oral drug [70].

Many aldose reductase inhibitors contain carboxylic acid groups that lead to poor bioavailability. Replacing the carboxylic acid group with 2,6-difluorophenol, tetrazole, or 1-hydroxypyrazole maintains or increases activity while allowing the lipophilicity and pKa to be modulated (Figure 32.13) [71–73].

2. *Use of Prodrugs*

A prodrug strategy can be used to increase the permeability of a drug. One method is to prepare a prodrug that contains a lipophilic portion in order to increase the overall lipophilicity of the molecule. Pivampicillin, for

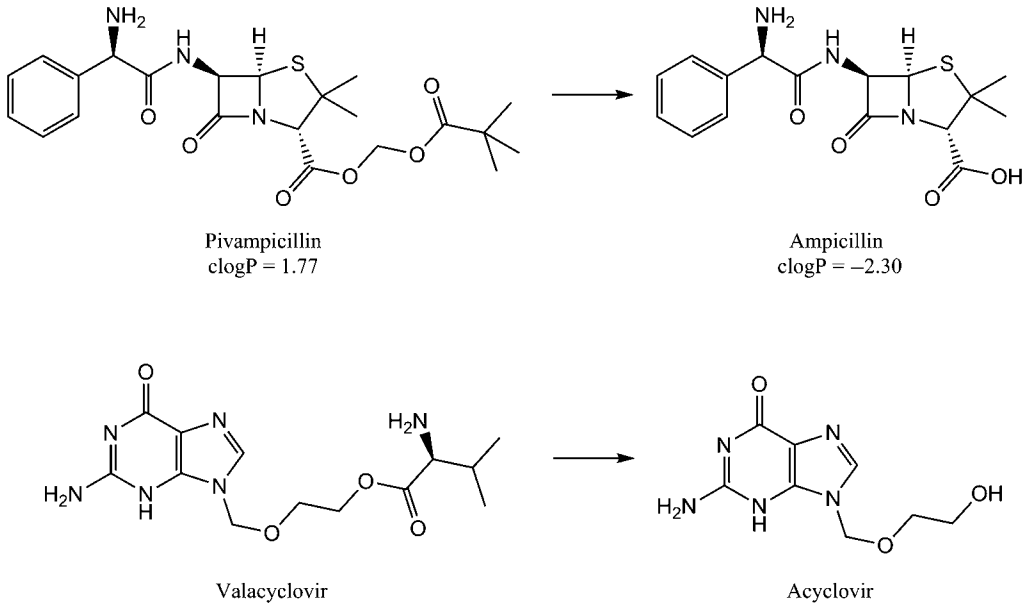


FIGURE 32.14 Examples of prodrug strategies to increase lipophilicity.

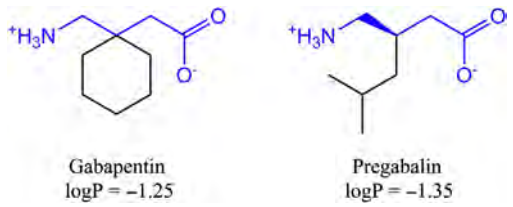


FIGURE 32.15 Targeting transporters: gabapentin and pregabalin.

example, is a pivaloyloxymethyl ester prodrug of ampicillin (Figure 32.14) that demonstrates a significant improvement in bioavailability.

Prodrugs that target specific transporters—in particular peptide transporters—are another route that can be exploited. Valacyclovir, the L-valyl ester prodrug of the antiviral acyclovir, improves the bioavailability over the parent 3–5-fold. Valacyclovir (Figure 32.14) is thought to target the PepT1 transporter, although it has also been found to interact with several other transporters [74].

3. *Ion Pair Interactions*

An ion pair consists of oppositely charged ions held together via Coulombic attraction that behave as a single unit. This has been exploited in drug delivery to enhance the permeability of hydrophilic drugs or drugs that are highly basic, acidic, or contain multiple pKas, such that they will be ionized throughout the physiological pH range. Ion pairing has been shown to enhance delivery via several routes, including transdermal, ocular, oral, and parenteral.

Piroxicam, a potent NSAID, is ionized at physiological pH. A series of ethanolamine salts have been prepared in order to improve its transdermal permeation properties in a number of vehicles. The results suggest that an improvement of up to 13-fold could be attained [75]. Formation of ion pairs is not always beneficial. Recent studies suggest that ion pair formation may be responsible for the decrease in bioavailability observed for poorly permeable drugs after food intake [76].

4. *Targeting Transporters*

Specific membrane proteins can be targeted in order to transport a hydrophilic drug. The general approach is to use a Trojan-Horse strategy of coupling a substance that does not cross a membrane barrier to one that does. Gabapentin and pregabalin are anticonvulsant GABA analogs that are believed to be transported via a system L amino acid transporter. Both molecules contain the common γ-aminobutyric acid substructure and exist as zwitterions at physiological pH (Figure 32.15). Since a transporter is being used, the process can become saturated.

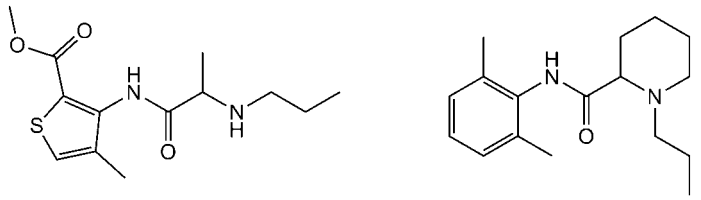


FIGURE 32.16 Half-life of local anesthetics.

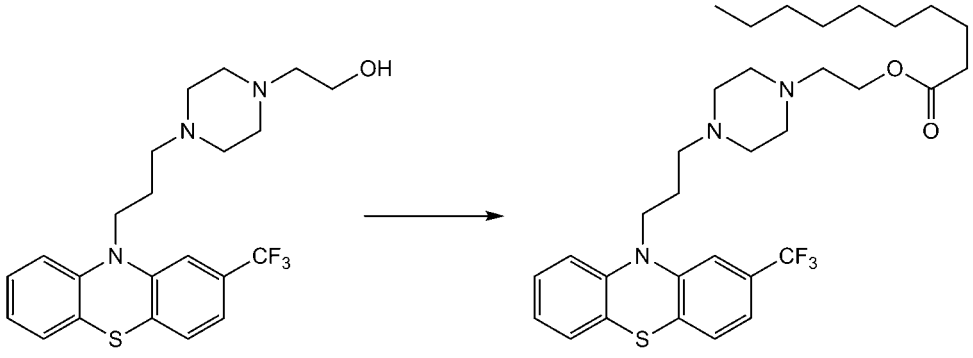


FIGURE 32.17 Fluphenazine and its decanoate prodrug.

The bioavailability of gabapentin, for example, varies inversely with dose. Use of membrane proteins to transport drugs is often coupled with preparation of a prodrug (see valacyclovir example in section III.B.2).

IV. MODIFYING THE DURATION OF ACTION

The duration of action of a drug depends on pharmacokinetic parameters such as half-life and the equilibration time between plasma and compartments such as tissue. Sometimes, the duration of action of a drug needs to be modified in order to ensure a suitable therapy. Among the chemical and physicochemical methods that can be used to adjust the duration of action are chemical modification, use of prodrugs, or conjugation of the drug to a polymer.

1. *Chemical Modification*

Introduction of hydrophobic groups will increase the concentration of a drug in fat tissue and slows the release into the blood supply. This strategy can be applied to drugs dosed intramuscularly. Articaine is a local anesthetic that is used in dentistry. It contains an ester group, which aids in rapid metabolism, leading to a half-life of twenty minutes. In contrast, bupivacaine, which is more lipophilic and more protein bound, has a half-life of 3.5 hours and is used post-surgery (Figure 32.16).

2. *Prodrugs*

Duration of action can be increased by using prodrugs to mask hydrophilic parts of a molecule and increase the lipophilicity of a drug. The residence time of insulin, for example, can be increased by attaching two 9-fluorenylmethoxycarbonyl (Fmoc) groups. The resulting prodrug has minimal activity itself and undergoes slow conversion to insulin under physiological conditions [77]. Fluphenazine, an antipsychotic, can be administered intramuscularly as a decanoate ester prodrug. The prodrug has a markedly extended duration of action, reducing the dosing frequency to once every 2–3 weeks from once every 6–8 hours for oral fluphenazine (Figure 32.17).

3. *Drug–Polymer Interactions*

One method to increase the blood residence of a drug is to conjugate it to a polymer where the increase in hydrodynamic volume translates into an increase in blood residence. Drugs can be conjugated directly to a polymer or may be conjugated via a spacer that can control the rate of release of the drug or target release under specific conditions. Among the polymers that have been used to adapt the properties of small-molecule and macromolecular drugs is poly(ethylene glycol) (PEG). Conjugation to PEG—or PEGylation—has been utilized for proteins for several decades as a way to improve their characteristics, but there are few small-molecule drug examples [78]. Binding a small-molecule drug to a large polymer can reduce its binding affinity but provides opportunities to increase bioavailability and half-life by reducing clearance.

An additional advantage in using this approach is the observed enhanced permeability and retention (EPR) of drugs in tissue that can be exploited in targeting tumors. Doxorubicin conjugated with different PEGs, for example, demonstrated a greater concentration in both blood and tumor as compared to free doxorubicin [79]. PEG linkages may be bound reversibly or irreversibly. Smaller PEGs tend to be used for irreversible linkages, as the PEG can interfere with the binding properties of the drug. Marcus and coworkers [80] demonstrated this in a study with the antibiotic gentamicin. Use of a reversible linkage between 5 and 40 kDa increased the half-life by up to 15-fold without loss of activity, whereas binding the drug directly to PEG led to a 100-fold loss of potency.

V. MANUFACTURING ISSUES

An API experiences different conditions of attrition, heat, and pressure during unit operations encountered in pharmaceutical processing. Drug materials interact with processing equipment, and some materials may undergo transformations as a result of processing [81]. Mechanical properties of materials can have an impact on processes, such as tableting or milling. In tableting operations, usually a higher drug-load formulation is more sensitive to properties of the API, and excipients have less of an impact in diminishing the impact of API properties. While the API has an impact on processing and hence the final drug product, it is potentially an option to anticipate effects and design or select material where undesirable effects could be minimized. In an idealized scenario material properties could be part of screening tree along with potency and ADME properties [82].

A. Material Characteristics and Tableting Behavior

During the tableting process, material deformation and bonding occur either through the plastic flow of materials into each other (ductile materials) or by attraction of new surfaces formed through elastic deformation and brittle fracture. Ideal material blends for effective tableting have a balance of plastic and elastic properties. Materials that are more plastic or exhibit viscoelastic behavior may demonstrate time-dependent flow and ineffective compression on higher speed machines, leading to tableting issues such as capping or lamination [83]. Inert excipients usually balance the properties of the API, and machine settings such as a precompression force can help manage such issues. In cases involving high doses and high drug loading, proactively predicting potential issues and designing the API with appropriate properties could help alleviate more complex problems.

Other tableting problems commonly encountered are picking and sticking, which may originate from the adhesive nature of powders. During a compression run, material slowly builds up on punches until it begins to deform the tablet surface resulting in picking, peeling or surface roughening. Sticking additionally involves adhesion to the die, resulting in needing increasing force to eject tablets. Mechanistically, these are the result of increased adhesion to punch / die surfaces relative to cohesive tablet bonding forces. This is managed during manufacturing by having highly polished or coated punch faces, lubrication of blends, or running the process at a slower speed. Smaller particles tend to adhere more, and decreasing the amount of fines in a blend can help. However, there is increasing interest in selecting materials that can avoid such problems in the first place.

In a study involving prism- and plate-shaped crystal habits of L-lysine monohydrochloride dihydrate, Sun and Grant [84] related the improved tableting of the plate shaped crystals to the favorable orientation of the slip planes, resulting in greater plasticity under load. Celik and coworkers [85] demonstrated the effect of crystal habit and particle size through compaction simulator studies. Improved compaction behavior has also been achieved using spherical crystallization techniques that actually result in spherical agglomerates of crystals [86,87]. Using different

crystal habits of ibuprofen and acetaminophen as models, Rasenack and Müller [88] proposed a method for assessing relative tableting behavior of crystal forms based on force-displacement curves.

Salt formation offers another lever for improving both physicochemical properties (discussed earlier in the chapter) and physicomechanical properties. Supuk and coworkers [89] demonstrated the effect of salt counter-ions on triboelectrification and compaction properties of flurbiprofen amine salts. Reduction in triboelectrification also resulted in reduced adhesion to processing surfaces. In studying different salt forms of L-lysine, Sun and Grant [90] related the tensile strength of compacts to melting points of the salts. Co-crystals also offer an interesting possibility, and Chow and coworkers [91] demonstrated improvements in product performance and processing of ibuprofen and flurbiprofen by forming 1:1 co-crystals with nicotinamide.

B. Corrosion and Abrasion

While wear and tear of processing equipment over the long run is expected, in some instances materials being processed may have a more significant impact. Weakly basic drugs are often converted into more soluble stable salts using strong acids. Under aqueous conditions (even with adsorbed moisture), these can generate a significantly acidic local pH, especially when the base is very weak. The resultant acidity can cause corrosion of metallic surfaces of equipment such as tablet punches. Hydrochloride salts, especially dihydrochlorides, are particularly notorious for affecting stainless steel surfaces.

Stahl and Nakano [92] relate the relative corrosion potential of two weakly basic drugs, a methanesulfonate and a hydrochloride salt, to the pKa of the base and the counter-ion used to form the salt. The weakly basic compound forming a hydrochloride salt was significantly more corrosive when applied to steel discs.

Some APIs can be quite hard and cause abrasion of unhardened/uncoated tablet punches. Calcium gluconate, for example, is known to be a particularly hard material that readily scratches tablet tooling. Similarly, some materials can be harder to micronize due to their ability to abrade mills. Measuring mechanical properties through techniques such as micro-indentation could help anticipate such issues.

VI. ADAPTING TO PATIENT'S NEEDS

A. Adapting Molecular and Material Properties to Choice of Route of Administration

The default choice for administering drugs is by the oral route, given the ease of self-administration, portability, and large scale manufacturability of such dosage forms. However, there are instances where the site of action, desired pharmacokinetic behavior, patient compliance concerns, or other benefit−risk assessment may favor an alternate route of administration. The molecular and material properties desirable for different routes of administration often have some specific nuances. These will be discussed in this section.

1. Oral Route

As already mentioned, oral delivery is most often the preferred route of administration for most drugs. Consequently, molecular and material properties of compounds resulting in adequate bioavailability have been the subject of much investigation and many resulting publications. Most drugs are transported from the gut into systemic circulation via diffusion across the large surface of the intestinal membrane. Following Fick's law, permeability of the drug through the membrane and solubility of the drug in the gut contents primarily determine the rate of drug diffusion. Further confounding factors that play a role in determining oral bioavailability include hepatic first-pass metabolism, gut-wall metabolism, and efflux transporters in the intestinal membranes. Molecular attributes (e.g., lipophilicity, ionization constants, hydrogen bonding) controlling solubility and permeability often work at cross-purposes with the need to be optimized to ensure maximal transport. One often used guide in this regard is what has come to be known as Lipinski's rule of five [93]. Lipinski's rule suggests poor oral absorption to occur when a molecule has two or more of the following:

- More than 5 H-bond donors (sum of −OH, -NH_x groups);
- Molecular weight greater than 500;
- Log P greater than 5; and
- More than 10 H-bond acceptors (sum of Ns and Os).

In the context of overall efficacy, one has to keep in mind the biological activity (potency) of the drug and the importance of dose by the oral route. A useful way [94] is the concept of calculating the maximum absorbable dose (MAD):

$$\mathrm{MAD} = \mathrm{S} \times \mathrm{Ka} \times \mathrm{SIWV} \times \mathrm{SITT}$$

where S is the solubility of drug at pH 6.5, Ka is the intestinal absorption rate constant (rat), SIWV is the small intestine water volume (~250 mL), and SITT is the small intestine transit time in minutes (~270 mins.). During discovery or early development (when a Ka value may not be available), this concept may still be used by making assumptions of the absorption rate constants (e.g., low = 0.003 min^{-1}, high = 0.03 min^{-1}) to estimate the range of doses likely to be absorbed. Permeability assessment through Caco-2 or PAMPA membranes can further refine these assumptions. MAD should not be taken in an absolute sense, but it does serve as a good communication tool to translate solubility into a clinical dose context.

In addition to properties that determine exposure, it is also important that compounds have adequate properties to withstand pharmaceutical processing and storage. Typically, a crystalline compound with a high enough melting point (>100°C) will be suitable for processing steps such as milling or compression. A compound with low solubility may need to be milled or micronized to improve dissolution rate from dosage forms. Often, a suitable salt might be chosen to optimize the ability to process without compromising solubility.

2. *Injectable Routes*

Injectable routes are best suited for situations requiring a fast onset of action, when oral delivery is not possible due to delivery challenges (biologics) or state of a patient (hospitalized/unconscious patient). The most common routes of parenteral administration for systemic exposure include intravenous (i.v.), subcutaneous (s.c.), and intramuscular (i.m.) routes. Molecular features and processing conditions are quite similar for each of these routes, but there may be some nuances specific to each. For instance, s.c. administration typically involves no more than 2 mL. An i.m. administration might involve 2–5 mL, while for i.v. administration it is possible to have a bag that delivers large volumes over a prolonged infusion period.

The preferred formulation for injectable drugs is an aqueous solution. Consequently, aqueous solubility and chemical stability are the most important features required of a molecule for parenteral formulations [95]. Ideally, it would be desirable for the molecule to have adequate stability to withstand terminal sterilization by autoclaving, but often that is not possible and the formulation is aseptically processed. If aqueous solutions are not stable, then chemical stability can sometimes be managed by developing a product that can be reconstituted (lyophilized or powder fill) with a diluent that can be added just prior to administration.

Consequently, solubility within an adequate pH range is a key attribute for formulation and delivery. For s.c. and i.m. administration, a pH range of 5.0–7.4 is preferred, although if needed pH 4.0–8.0 might be acceptable. An i.v. administration can more easily tolerate pH 4.0–8.0 due to the rapid dilution and buffering of blood. If pH adjustment alone does not achieve adequate solubility, other formulation approaches involving co-solvents, surfactant micelles, complexing agents, or even nanosuspensions can be utilized. The most robust solution is the one that is built into the molecule whenever it is feasible, and potential ways to modify solubility and stability without affecting efficacy have been discussed in some detail in earlier sections.

3. *Transdermal*

One of the largest organs of the body, the skin, is designed to be a barrier. However, delivering drugs through the skin is an attractive route, especially for chronic dosing indications where patient compliance could be a challenge. It has additional advantages of being painless and avoiding the challenges of oral absorption such as the first-pass effect, acidic pH of stomach, and gastric emptying.

Systemic delivery of drugs through the skin encounters barriers primarily in the stratum corneum, an external layer mainly comprising dead epidermal cells and the epidermis. The stratum corneum is more of a lipid barrier, while the epidermal and dermal layers are relatively hydrophilic. The main route of transport is passive diffusion through these layers although a small shunt might also exist via the transfollicular route [96].

Ideal drug candidates for transdermal delivery would have the following properties:

- Molecular weight <500;
- Log P 1–3;
- A low melting point; and
- Nonirritant, nonsensitizing, or photosensitizing.

Essentially small amphiphilic molecules that have some solubility in the lipidic stratum corneum and some solubility in the more aqueous epidermal and dermal layers can potentially be absorbed through the skin into the systemic circulation. A dose of 10–15 mg can potentially be delivered through this route.

4. *Inhaled and Nasal Delivery of Drugs*

Inhalation delivery has conventionally been intended for treatment of airway disorders (e.g., asthma, allergy, COPD), and in those instances absorption into systemic circulation is not intended, although good exposure to local tissues is critical. The first step of delivery to the lung involves deposition of particles or droplets. Particles have to be small enough to be inhaled and deposited within bronchioles and alveoli. Drug deposition within the lung is determined by impaction, sedimentation, or diffusion, which in turn is governed by the aerodynamic particle or droplet size (Figure 32.18) [97]. Particle shape also has an important effect on lung deposition. Elongated particles effectively have a smaller aerodynamic size and tend to be deposited deeper in the lung, whereas spherical particles might be deposited higher [98]. There has been interest in designing nature-inspired particle shapes such as pollen-shaped particles in order to modulate particle deposition patterns [99]. To act locally, deposited particles need to dissolve before they are removed by mucociliary clearance and / or alveolar macrophages.

The route does provide a possible route for delivery of drugs intended to act systemically, however. Unlike the case of oral, where the rule of five gives a precise definition for suitable drug candidates for absorption, with alternate routes there are wider possibilities [100] that may need to be explored experimentally. Small molecules that are mildly hydrophobic (log $P > 1$) are easily absorbed. Some examples would include nicotine, morphine, and fentanyl. Smaller peptides are easily absorbed, and proteins such as insulin have also been successfully delivered by this route in limited cases. Generally, deep lung (distal airways and alveoli) is best for systemic delivery, with the exception of instances when transporters located in the conducting airways are targeted (e.g., interferons, FSH, EPO). With unblocked peptides, one also needs to be concerned about enzymatic degradation by peptidases [97].

A device is integral to delivering drugs to the lung. Such devices might use dry powders (dry powder inhalers), solutions or suspensions in propellant gases (metered dose inhalers), or aqueous solutions or suspensions (nebulizers or soft mist inhalers). Key to product efficacy is lung deposition, which depends on aerodynamic particle size (2–7 μm). Drug properties and the device contribute equally to achieving this [101]. For solution formulation, good solubility and stability of the drug in the solvent system is important. For suspension formulations, ideally the API should be crystalline, be effectively reduced to a small size (~5 μm), and have a low solubility in the solvent system (hence decreased crystal growth potential). For dry powders, typically crystalline nonhygroscopic compounds that can be effectively size reduced are desirable. Formulations for such inhalers

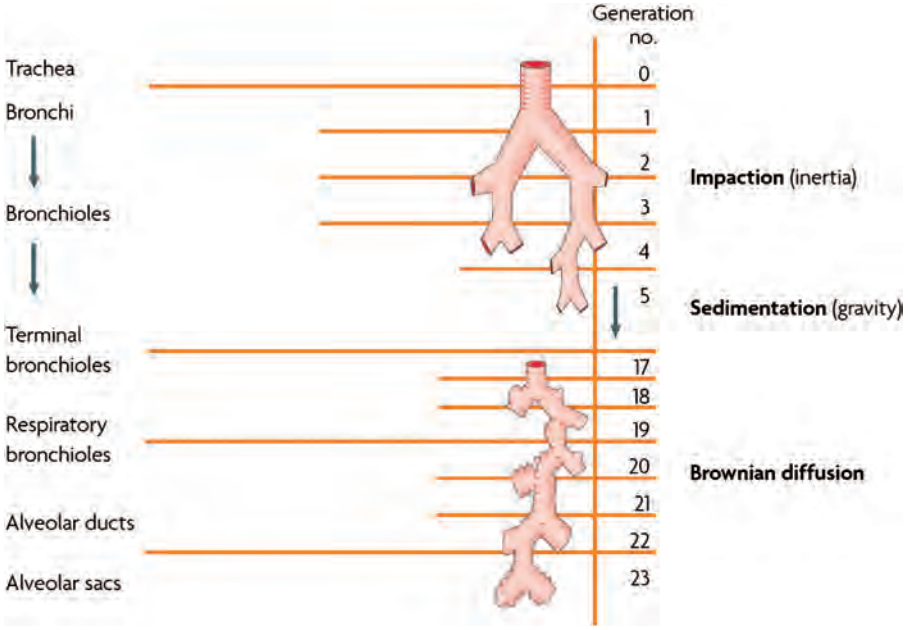


FIGURE 32.18 Factors that determine the deposition of inhaled particles. *Reprinted with permission from [97].*

often involve small drug particles dispersed on another carrier (e.g., lactose) with the design that the drug particle detaches and gets inhaled during the process of inhalation. Consequently, the impact of size reduction on surface properties (e.g., powder electrostatics, surface disorder, and surface energy) of drug particles has a significant impact on product performance and should be well characterized and understood.

In summary, when delivering drugs by the inhalation route it is important to optimize molecular and material properties, keeping in mind not just the biological barriers but also the formulation constraints of different device types required to achieve effective delivery.

Nasal delivery has some aspects in common with inhalation in that a delivery device is an integral aspect of drug delivery. However, the anatomy and physiology of the nose are quite different from the oral route and have an impact on delivery issues. Since the nasal epithelium is directly accessible from the outside, devices are simply designed to spray a metered amount of the drug in a solution or suspension form onto the nasal mucosa, from which they may act locally or be absorbed for potential systemic exposure. Locally acting drugs (e.g., antihistamines, corticosteroids, and nasal decongestants) are commonly used in the case of allergic rhinitis [102].

Systemically active drugs that are unstable in the gastric environment or undergo significant hepatic first-pass metabolism can be delivered by this route. Examples of such drugs would include peptides, proteins, and hormones [103]. Salmon Calcitonin, Desmopressin, and Estradiol are successfully delivered this way. CNS active drugs such as Sumitriptan and Fentanyl are delivered by this route with the advantage of a rapid onset of action as the absorbed drug is carried to the brain tissues via the carotid artery. There has also been a lot of recent interest in the area of nose-to-brain delivery, by delivering drug to the olfactory region of the nose not easily accessed by conventional spray-pump devices [104]. CNS active drugs and even compounds such as insulin have been shown to cross into the brain directly through intra-axonal pathways of the olfactory and trigeminal nerves [105].

Drugs that are unionized at the pH of nasal mucosa (5.5–6.5) are better absorbed. Generally, lipophilic drugs such as fentanyl or progesterone are absorbed through the nasal mucosa, but even polar drugs with molecular weights below 300 Da are absorbed by this route. As in the case of inhalation, mucociliary clearance and some potential for enzymatic degradation are considerations. Given the small surface area of the nasal mucosa, this route is best suited to potent low-dose compounds.

5. *Ocular Drug Delivery*

Eye diseases may affect the anterior segment (e.g., dry eye syndrome, conjunctivitis, glaucoma) and/or the posterior segment (e.g., age-related macular degeneration, diabetic retinopathy, diabetic macular edema, uveitis). The barriers and delivery challenges involved are quite different depending on the target tissue within the eye. Many ocular drugs on the market today aim to treat diseases affecting the anterior segment, where locally applied eye drops or ointments can be quite effective. Posterior eye diseases, which are the leading causes of vision loss in developed countries, have become in recent years an important therapeutic target with unmet medical needs [106].

In general, local (nonsystemic) routes for ocular drug administration and delivery technologies can be classified into invasive techniques and noninvasive drug delivery. Invasive delivery involves intravitreal or periocular injection of solutions, suspensions, or implants containing the drug. Lucentis is an example of such a solution used in AMD, and Vitrasert is a device implanted for posterior uveitis [107]. While these solutions achieve the therapeutic effects, they are less desirable from a patient-compliance perspective or may represent a high technological barrier. Eye drops or eye ointments are simpler options and have traditionally been quite effective for anterior eye therapy. More recently, a number of approaches are being tested in clinical trials using eye drop formulations for posterior eye diseases.

The barriers to delivering drugs to the eye are quite complex. Generally, drugs applied to the eye are cleared by naso-lachrymal flow (washed off by tear flow). As in the case of transdermal delivery, the penetration of ocular tissue is complicated by different tissue barriers within the organ. Physicochemical properties suitable for one layer may not be suitable for others. While specific studies will determine the suitability of a molecule for ocular delivery, some generalizations regarding suitable properties are identified below [108,109]:

- Sufficient solubility to accommodate required dose in 1–2 drops (~50–100 μL). Higher concentration at pH 7.4 will produce a greater driving force for permeation.
- Log P in the range of 1.5–4.0. If specific transporters are targeted then the log P is less relevant.
- Small molecular radius (≤10 Å).
- Stability in pH range of 5–8.
- Ability to bind to mucin in the eye by cationic drugs, which could help retain them longer.

- Should not absorb light in the range of 290–700 nm wavelength.
- Nonirritant, nonsensitizing, or photosensitizing.

Ocular formulations are generally processed similarly to injectable products and need to be sterilized by autoclaving, or involve aseptic processing and hence the compounds need to be able to withstand the rigors of appropriate manufacturing conditions.

B. Irritation Potential

The irritation potential of a drug compound or the formulation can become a significant issue for patient compliance and may become a source of pain or more serious toxicity. Many drug compounds are known to cause gastro-intestinal disorders [110], the cause being locally high concentrations and resultant pH/osmotic effects or more mechanism-based effects. Understanding the potential for irritation could eventually lead to designing out the particular feature. An example of this would be the gastric irritation caused by aspirin and other traditional nonsteroidal anti-inflammatory drugs such as indomethacin or naproxen. These drugs act by inhibiting the cyclo-oxygenase (COX) enzymes involved in prostaglandin synthesis. Along with reducing inflammation, these compounds have an effect on reducing the protective gastric lining, which is the cause of GI irritation with serious consequences in certain cases. Eventually, with the understanding of classification of COX-1 and COX-2 enzymes and their distribution, specific COX-2 inhibitors (e.g., Celecoxib) were developed with similar anti-inflammatory activity but reduced GI irritancy.

Similarly, injection site irritation and pain could be an impediment to patient comfort and compliance. More serious effects that accompany injection site irritation can include haemorrhage, oedema, inflammation, and tissue necrosis [111]. Factors contributing to these effects are the drug's solubility in aqueous medium, the pH and buffer capacity, the hypo- or hyperosmotic character of the injected drug solution, and the amount of the injected volume, as well as possible precipitation of the drug at the injection site and cell lysis. The anticonvulsant drug phenytoin, which is sparingly soluble in water, is formulated with high concentrations of propylene glycol and ethanol in order to solubilize the required dose. This results in a high risk of drug precipitation and pain with improper injection technique. Fosphenytoin sodium (Cerebyx) was developed as a prodrug of phenytoin intended for parenteral administration [112]. The drug's aqueous solubility is 142 mg/ml, over 4,400 times that of phenytoin. It is a phosphate ester derivative of phenytoin sodium and is enzymatically converted to phenytoin after parenteral administration by phosphatases located in human tissue (Figure 32.19).

C. Pediatric Development and Taste Masking

Government agencies have been increasingly emphasizing the need to include considerations of special populations during development of drug compounds. When it comes to pediatric drug development a carrot-and-stick approach has been implemented, which involves a 6-month extension in market exclusivity and a requirement to include a pediatric development plan (or approval of a waiver) as part of the initial submission dossier [113]. This has been necessitated by the fact that there are not enough medications on the market that are developed and tested for pediatric populations [114]. Thankfully, development of pediatric formulations has taken an important place in the overall development plans for new drugs.

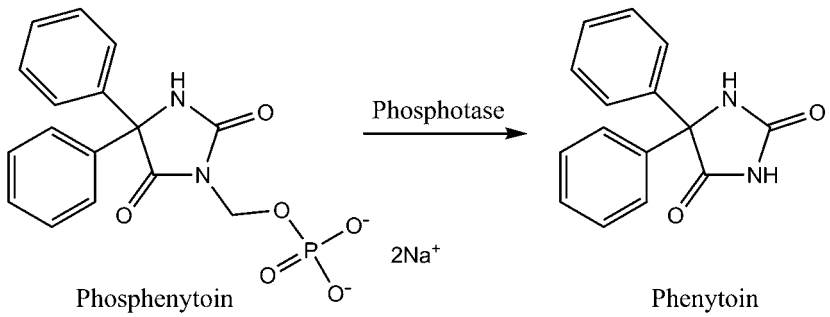


FIGURE 32.19 Enzymatic conversion of fosphenytoin to phenytoin.

One of the challenges of developing oral formulations for children is that conventional tablets and capsules cannot be swallowed by younger children. Suspensions, powders for reconstitution, rapidly disintegrating tablets, or fast dissolving strips are preferred dosage forms. In each of these cases, the taste of the drug and approaches to mask the taste become important steps in addition to conventional challenges of maintaining stability, uniformity, and other desired properties. Formulation approaches include addition of sweeteners and flavors or coating small particles with taste masking polymers. These approaches are often insufficient in masking the taste or managing the mouth feel (grittiness) of the product, and can sometimes be technically challenging.

Among the best approaches to solve the taste issue is identifying a form of the drug that has minimal solubility at the pH of the mouth but is still equally bioavailable when swallowed. Salts, prodrugs, and unionized or complexed forms are all potential states that could achieve the goal, and there are very good examples of successful product development. The palmitate ester prodrug of chloramphenicol is an old example of achieving effective taste masking without having to sacrifice exposure. Sinkula [115] described some general chemical rules to follow in designing out undesirable taste from potential prodrugs or other chemical modifications. The bitterness masking of propiverine hydrochloride was demonstrated by converting it to a free base form and simultaneously granulating it into small granules with a binder [116]. Pharmacokinetic assessment in dogs showed comparable plasma concentration–time profiles. A process involving complexation of chlorphenaramine maleate—a bitter drug—with tannic acid resulted in low solubility and effective taste masking at the pH of the mouth of 6.8, yet good dissolution and release at gastric pH [117]. Chrzan and coworkers [118] described a well-tolerated pediatric suspension formulation of fexofenadine hydrochloride—again a very bitter drug—that involved converting the drug to its zwitterionic dihydrate form. This formulation was shown to have very similar exposure (AUC and C_{max}) to a reference tablet formulation. In each of these examples, an understanding of the potential forms of the drug compound and their properties led to a very robust solution to the challenge of unacceptable taste.

References

[1] Waterhouse DJ, Saavedra JE, Davies KM, Citro ML, Xu X, Powell DA, et al. Injectable formulation of disodium 1-[2-(carboxylato)pyrrolidin-1-yl]diazen-1-ium-1,2-diolate (PROLI/NO), an ultrafast nitric oxide donor prodrug. J Pharm Sci 2006;95(1):108–15.

[2] Dong Y, Ng WK, Surana U, Tan RB. Solubilization and preformulation of poorly water soluble and hydrolysis susceptible N-epoxymethyl-1,8-naphthalimide (ENA) compound. Int J Pharm 2008;356(1–2):130–6.

[3] Hoellein L, Holzgrabe U. Ficts and facts of epinephrine and norepinephrine stability in injectable solutions. Int J Pharm 2012;434(1–2):468–80.

[4] Wasylaschuk WR, Harmon PA, Wagner G, Harman AB, Templeton AC, Xu H, et al. Evaluation of hydroperoxides in common pharmaceutical excipients. J Pharm Sci 2007;96(1):106–16.

[5] Hartauer KJ, Arbuthnot GN, Baertschi SW, Johnson RA, Luke WD, Pearson NG, et al. Influence of peroxide impurities in povidone and crospovidone on the stability of raloxifene hydrochloride in tablets: identification and control of an oxidative degradation product. Pharm Dev Technol 2000;5(3):303–10.

[6] Glass BD, Brown ME, Daya S, Worthington MS, Drummond P, Antunes E, et al. Influence of cyclodextrins on the photostability of selected drug molecules in solution and the solid-state. Int J Photoenergy 2001;3:205–11.

[7] Grooff D, Francis F, De VMM, Ferg E. Photostability of crystalline versus amorphous nifedipine and nimodipine. J Pharm Sci 2013;102(6):1883–94.

[8] Bayomi MA, Abanumay KA, Al-Angary AA. Effect of inclusion complexation with cyclodextrins on photostability of nifedipine in solid state. Int J Pharm 2002;243(1–2):107–17.

[9] Burkhard JA, Wuitschik G, Plancher JM, Rogers-Evans M, Carreira EM. Synthesis and stability of oxetane analogs of thalidomide and lenalidomide. Org Lett 2013;15(17):4312–15.

[10] Wei M, Yuan Q, Evans DG, Wang Z, Duan X. Layered solids as a "molecular container" for pharmaceutical agents: l-tyrosine-intercalated layered double hydroxides. J Mater Chem 2005;15(11):1197–203.

[11] Kieffer J, Brémond É, Lienard P, Boccardi G. In silico assessment of drug substances chemical stability. J Mol Struct THEOCHEM 2010;954(1–3):75–9.

[12] Parenty AD, Button WG, Ott MA. An expert system to predict the forced degradation of organic molecules. Mol Pharm 2013;10(8):2962–74.

[13] Hancock BC, Parks M. What is the true solubility advantage for amorphous pharmaceuticals? Pharm Res 2000;17:397–404.

[14] Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Solubility advantage of amorphous pharmaceuticals: I. A thermodynamic analysis. J Pharm Sci 2010;99(3):1254–64.

[15] Sukenik CN, Bonapace JAP, Mandel NS, Bergman RG, Lau P-Y, Wood G. Enhancement of a chemical reaction rate by proper orientation of reacting molecules in the solid state. J Am Chem Soc 1975;97:5290–1.

[16] Pikal MJ, Rigsbee DR. The stability of insulin in crystalline and amorphous solids: observation of greater stability for the amorphous form. Pharm Res 1997;14(10):1379–87.

[17] Luthra SA, Shalaev EY, Medek A, Hong J, Pikal MJ. Chemical stability of amorphous materials: specific and general media effects in the role of water in the degradation of freeze-dried zoniporide. J Pharm Sci 2012;101(9):3110–23.

[18] Ohtake S, Shalaev E. Effect of water on the chemical stability of amorphous pharmaceuticals. I. Small molecules. J Pharm Sci 2013;102 (4):1139–54.
[19] Yoshioka S, Aso Y. Correlations between molecular mobility and chemical stability during storage of amorphous pharmaceuticals. J Pharm Sci 2007;96(5):960–81.
[20] Byrn SR, Xu W, Newman AW. Chemical reactivity in solid-state pharmaceuticals: formulation implications. Adv Drug Deliv Rev 2001;48 (1):115–36.
[21] Munjal M, Stodghill SP, Elsohly MA, Repka MA. Polymeric systems for amorphous Delta 9-tetrahydrocannabinol produced by a hot-melt method. Part I: chemical and thermal stability during processing. J Pharm Sci 2006;95(8):1841–53.
[22] Munjal M, Elsohly MA, Repka MA. Polymeric systems for amorphous Delta 9-tetrahydrocannabinol produced by a hot-melt method. Part II: Effect of oxidation mechanisms and chemical interactions on stability. J Pharm Sci 2006;95(11):2473–85.
[23] Kojima T, Onoue S, Katoh F, Teraoka R, Matsuda Y, Kitagawa S, et al. Effect of spectroscopic properties on photostability of tamoxifen citrate polymorphs. Int J Pharm 2007;336(2):346–51.
[24] Takahashi Y, Nakashima K, Nakagawa H, Sugimoto I. Effects of grinding and drying on the solid-state stability of ampicillin trihydrate. Chem Pharm Bull (Tokyo) 1984;32(12):4963–70.
[25] Guerrieri P, Jarring K, Taylor LS. Impact of counterion on the chemical stability of crystalline salts of procaine. J Pharm Sci 2010;99 (9):3719–30.
[26] Brewster ME, Loftsson T, Estes KS, Lin JL, Fridriksdottir H, Bodor N. Effect of various cyclodextrins on solution stability and dissolution rate of doxorubicin hydrochloride. Int J Pharm 1992;79(2–3):289–99.
[27] Pourmokhtar M, Jacobson GA. Enhanced stability of sulfamethoxazole and trimethoprim against oxidation using hydroxypropyl-beta-cyclodextrin. Pharmazie 2005;60(11):837–9.
[28] Grunenberg A, Henck JO, Siesler HW. Theoretical derivation and practical application of energy/temperature diagrams as an instrument in preformulation studies of polymorphic drug substances. Int J Pharm 1996;129(1–2):147–58.
[29] McCrone WC. In: Fox D, Labes MM, Weissberger A, editors. Physics and chemistry of the organic solid state, vol. 2. New York: Wiley; 1965. p. 725–67.
[30] Brittain HG. Polymorphism and solvatomorphism 2010. J Pharm Sci 2012;101(2):464–84.
[31] Aguiar AJ, Zelmer JE. Dissolution behavior of polymorphs of chloramphenicol palmitate and mefenamic acid. J Pharm Sci 1969;58 (8):983–7.
[32] Barbas R, Marti F, Prohens R, Puigjaner C. Polymorphism of norfloxacin: evidence of the enantiotropic relationship between polymorphs A and B. Cryst Growth Des 2006;6(6):1463–7.
[33] Chemburkar SR, Bauer J, Deming K, Spiwek H, Patel K, Morris J, et al. Dealing with the impact of ritonavir polymorphs on the late stages of bulk drug process development. Org Process Res Dev 2000;4(5):413–17.
[34] Gavezzotti A. Are crystal structures predictable? Acc Chem Res 1994;27:309–14.
[35] Bardwell DA, Adjiman CS, Arnautova YA, Bartashevich E, Boerrigter SXM, Braun DE, et al. Towards crystal structure prediction of complex organic compounds: a report on the fifth blind test. Acta Crystallogr, Sect B: Struct Sci 2011;67(6):535–51.
[36] Bernstein J. Polymorphism in molecular crystals. Oxford: Oxford University Press; 2002.
[37] Brittain HG. editor. Polymorphism in pharmaceutical solids. 2nd ed. In J. Swarbrick, editor. Drugs and the pharmaceutical sciences. vol. 192. New Yorl: Informa Healthcare; 2009.
[38] Hilfiker R, editor. Polymorphism: in the pharmaceutical industry. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA; 2006.
[39] Guerrieri P, Taylor LS. Role of salt and excipient properties on disproportionation in the solid-state. Pharm Res 2009;26(8):2015–26.
[40] Zannou EA, Ji Q, Joshi YM. Serajuddin ATM. Stabilization of the maleate salt of a basic drug by adjustment of microenvironmental pH in solid dosage form. Int J Pharm 2007;337(1–2):210–18.
[41] Hancock BC, Zografi G. Characteristics and significance of the amorphous state in pharmaceutical systems. J Pharm Sci 1997;86(1):1–12.
[42] Newman A, Knipp G, Zografi G. Assessing the performance of amorphous solid dispersions. J Pharm Sci 2012;101(4):1355–77.
[43] Hancock BC, Shamblin SL, Zografi G. Molecular mobility of amorphous pharmaceutical solids below their glass transition temperatures. Pharm Res 1995;12(6):799–806.
[44] Aso Y, Yoshioka S. Molecular mobility of nifedipine-PVP and phenobarbital-PVP solid dispersions as measured by 13C-NMR spin-lattice relaxation time. J Pharm Sci 2006;95(2):318–25.
[45] Rumondor AC, Marsac PJ, Stanford LA, Taylor LS. Phase behavior of poly(vinylpyrrolidone) containing amorphous solid dispersions in the presence of moisture. Mol Pharm 2009;6(5):1492–505.
[46] Marsac PJ, Rumondor AC, Nivens DE, Kestur US, Stanciu L, Taylor LS. Effect of temperature and moisture on the miscibility of amorphous dispersions of felodipine and poly(vinyl pyrrolidone). J Pharm Sci 2010;99(1):169–85.
[47] Yang J, Grey K, Doney J. An improved kinetics approach to describe the physical stability of amorphous solid dispersions. Int J Pharm 2010;384(1–2):24–31.
[48] Kasuga J, Ishikawa M, Yonehara M, Makishima M, Hashimoto Y, Miyachi H. Improvement of water-solubility of biarylcarboxylic acid peroxisome proliferator-activated receptor (PPAR) delta-selective partial agonists by disruption of molecular planarity/symmetry. Bioorg Med Chem 2010;18(20):7164–73.
[49] Scott JS, Birch AM, Brocklehurst KJ, Broo A, Brown HS, Butlin RJ, et al. Use of small-molecule crystal structures to address solubility in a novel series of g protein coupled receptor 119 agonists: optimization of a lead and in vivo evaluation. J Med Chem 2012;55:5361–79.
[50] Serajuddin AT. Salt formation to improve drug solubility. Adv Drug Deliv Rev 2007;59(7):603–16.
[51] Paulekuhn GS, Dressman JB, Saal C. Trends in active pharmaceutical ingredient salt selection based on analysis of the Orange Book database. J Med Chem 2007;50(26):6665–72.
[52] Sweetana S, Akers MJ. Solubility principles and practices for parenteral drug dosage form development. PDA J Pharm Sci Technol 1996;50(5):330–42.
[53] Everett RM, Descotes G, Rollin M, Greener Y, Bradford JC, Benziger DP, et al. Nephrotoxicity of pravadoline maleate (WIN 48098-6) in dogs: evidence of maleic acid-induced acute tubular necrosis. Fundam Appl Toxicol 1993;21(1):59–65.

[54] Stahl PH, Wermuth CG, editors. Handbook of Pharmaceutical Salts: Properties, Selection, and Use. 2nd ed. Wiley-VCH; 2011.
[55] US Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for industry: Regulatory classification of pharmaceutical cocrystals, 2013.
[56] Weyna DR, Cheney ML, Shan N, Hanna M, Zaworotko MJ, Sava V, et al. Improving solubility and pharmacokinetics of Meloxicam via multiple-component crystal formation. Mol Pharm 2012;9:2094–102.
[57] Ostwald W. Studien über die Bildung und Umwandlung fester Körper [Studies on the formation and transformation of solid bodies]. Zeitschrift für Physikalische Chemie 1897;34:495–503 [in German].
[58] Murdande SB, Pikal MJ, Shanker RM, Bogner RH. Aqueous solubility of crystalline and amorphous drugs: challenges in measurement. Pharm Dev Technol 2011;16(3):187–200.
[59] Shah N, Iyer RM, Mair H-J, Choi DS, Tian H, Diodone R, et al. Improved human bioavailability of vemurafenib, a practically insoluble drug, using an amorphous polymer-stabilized solid dispersion prepared by a solvent-controlled coprecipitation process. J Pharm Sci 2013;102(3):967–81.
[60] Loftsson T, Brewster ME. Pharmaceutical applications of cyclodextrins. 1. Drug solubilization and stabilization. J Pharm Sci 1996;85 (10):1017–25.
[61] Sharma US, Balasubramanian SV, Straubinger RM. Pharmaceutical and physical properties of paclitaxel (Taxol) complexes with cyclodextrins. J Pharm Sci 1995;84(10):1223–30.
[62] Shinkai S, Mori S, Koreishi H, Tsubaki T, Manabe O. Hexasulfonated calix[6]arene derivatives: a new class of catalysts, surfactants, and host molecules. J Am Chem Soc 1986;108(9):2409–16.
[63] Zhao Y, Buck DP, Morris DL, Pourgholami MH, Day AI, Collins JG. Solubilisation and cytotoxicity of albendazole encapsulated in cucurbit[n]uril. Org Biomol Chem 2008;6(24):4509–15.
[64] Rasool AA, Hussain AA, Dittert LW. Solubility enhancement of some water-insoluble drugs in the presence of nicotinamide and related compounds. J Pharm Sci 1991;80(4):387–93.
[65] Kim JY, Kim S, Papp M, Park K, Pinal R. Hydrotropic solubilization of poorly water-soluble drugs. J Pharm Sci 2010;99(9):3953–65.
[66] Devarakonda B, Hill RA, Liebenberg W, Brits M, de Villiers MM. Comparison of the aqueous solubilization of practically insoluble niclosamide by polyamidoamine (PAMAM) dendrimers and cyclodextrins. Int J Pharm 2005;304(1–2):193–209.
[67] Yang W, de Villiers MM. Effect of 4-sulphonato-calix[n]arenes and cyclodextrins on the solubilization of niclosamide, a poorly water soluble anthelmintic. AAPS J 2005;7(1):E241–8.
[68] Piyapolrungroj N, Zhou YS, Li C, Liu G, Zimmermann E, Fleisher D. Cimetidine absorption and elimination in rat small intestine. Drug Metab Dispos 2000;28:65–72.
[69] Herr RJ. 5-Substituted-1H-tetrazoles as carboxylic acid isosteres: medicinal chemistry and synthetic methods. Bioorg Med Chem 2002;10 (11):3379–93.
[70] Duncia JV, Carini DJ, Chiu AT, Johnson AL, Price WA, Wong PC, et al. The discovery of DuP 753, a potent, orally active nonpeptide angiotensin II receptor antagonist. Med Res Rev 1992;12(2):149–91.
[71] Nicolaou I, Zika C, Demopoulos VJ. [1-(3,5-difluoro-4-hydroxyphenyl)-1H-pyrrol-3-yl]phenylmethanone as a bioisostere of a carboxylic acid aldose reductase inhibitor. J Med Chem 2004;47(10):2706–9.
[72] Pegklidou K, Koukoulitsa C, Nicolaou I, Demopoulos VJ. Design and synthesis of novel series of pyrrole based chemotypes and their evaluation as selective aldose reductase inhibitors: a case of bioisosterism between a carboxylic acid moiety and that of a tetrazole. Bioorg Med Chem 2010;18(6):2107–14.
[73] Papastavrou N, Chatzopoulou M, Pegklidou K, Nicolaou I. 1-Hydroxypyrazole as a bioisostere of the acetic acid moiety in a series of aldose reductase inhibitors. Bioorg Med Chem 2013;21(17):4951–7.
[74] Yang B, Hu Y, Smith DE. Impact of peptide transporter 1 on the intestinal absorption and pharmacokinetics of valacyclovir after oral dose escalation in wild-type and PepT1 knockout mice. Drug Metab Dispos 2013;41(10):1867–74.
[75] Cheong H-A, Choi H-K. Enhanced percutaneous absorption of piroxicam via salt formation with ethanolamines. Pharm Res 2002;19 (9):1375–80.
[76] Heinen CA, Reuss S, Amidon GL, Langguth P. Ion pairing with bile salts modulates intestinal permeability and contributes to food–drug interaction of BCS class III compound trospium chloride. Mol Pharm 2013;10(11):3989–96.
[77] Gershonov E, Shechter Y, Fridkin M. New concept for long-acting insulin: spontaneous conversion of an inactive modified insulin to the active hormone in circulation: 9-fluorenylmethoxycarbonyl derivative of insulin. Diabetes 1999;48(7):1437–42.
[78] Li W, Zhan P, De CE, Lou H, Liu X. Current drug research on PEGylation with small molecular agents. Prog Polym Sci 2013;38:421–44.
[79] Veronese FM, Schiavon O, Pasut G, Mendichi R, Andersson L, Tsirk A, et al. PEG – doxorubicin conjugates: influence of polymer structure on drug release, in vitro cytotoxicity, biodistribution, and antitumor activity. Bioconjug Chem 2005;16(4):775–84.
[80] Marcus Y, Sasson K, Fridkin M, Shechter Y. Turning low-molecular-weight drugs into prolonged acting prodrugs by reversible pegylation: a study with gentamicin. J Med Chem 2008;51(14):4300–5.
[81] Morris KR, Griesser UJ, Eckhardt CJ, Stowell JG. Theoretical approaches to physical transformations of active pharmaceutical ingredients during manufacturing processes. Adv Drug Deliv Rev 2001;48(1):91–114.
[82] Gardner CR, Walsh CT, Almarsson O. Drugs as materials: valuing physical form in drug discovery. Nat Rev Drug Discov 2004;3 (11):926–34.
[83] Patel S, Kaushal AM, Bansal AK. Compression physics in the formulation development of tablets. Crit Rev Ther Drug Carrier Syst 2006;23(1):1–65.
[84] Sun C, Grant DJ. Influence of crystal shape on the tableting performance of L-lysine monohydrochloride dihydrate. J Pharm Sci 2001;90 (5):569–79.
[85] Celik M, Ong JT, Chowhan ZT, Samuel GJ. Compaction simulator studies of a new drug substance: effect of particle size and shape, and its binary mixtures with microcrystalline cellulose. Pharm Dev Technol 1996;1(2):119–26.
[86] Kawashima Y, Cui F, Takeuchi H, Niwa T, Hino T, Kiuchi K. Improved static compression behaviors and tablettabilities of spherically agglomerated crystals produced by the spherical crystallization technique with a two-solvent system. Pharm Res 1995;12(7):1040–4.

[87] Di Martino P, Di Cristofaro R, Barthelemy C, Joiris E, Palmieri Filippo G, Sante M. Improved compression properties of propyphenazone spherical crystals. Int J Pharm 2000;197(1–2):95–106.
[88] Rasenack N, Muller BW. Crystal habit and tableting behavior. Int J Pharm 2002;244(1–2):45–57.
[89] Supuk E, Ghori MU, Asare-Addo K, Laity PR, Panchmatia PM, Conway BR. The influence of salt formation on electrostatic and compression properties of flurbiprofen salts. Int J Pharm 2013;458(1):118–27.
[90] Sun C, Grant DJ. Compaction properties of L-lysine salts. Pharm Res 2001;18(3):281–6.
[91] Chow SF, Chen M, Shi L, Chow AH, Sun CC. Simultaneously improving the mechanical properties, dissolution performance, and hygroscopicity of ibuprofen and flurbiprofen by cocrystallization with nicotinamide. Pharm Res 2012;29(7):1854–65.
[92] Stahl PH, Nakano M. Pharmaceutical aspects of the drug salt form. In: Stahl PH, Wermuth CG, editors. Handbook of pharmaceutical salts: properties, selection, and use. Weinheim: Wiley-VCH; 2002. p. 83–116.
[93] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 2001;46(1–3):3–26.
[94] Johnson KC, Swindell AC. Guidance in the setting of drug particle size specifications to minimize variability in absorption. Pharm Res 1996;13(12):1795–8.
[95] Vemuri NM. Preformulation. In: Nema S, Ludwig JD, editors. Pharmaceutical dosage forms: parenteral medications. 3rd ed. New York: Informa Healthcare; 2010;1:57–75.
[96] Hadgraft J, Lane ME. Skin permeation: the years of enlightenment. Int J Pharm 2005;305(1–2):2–12.
[97] Patton JS, Byron PR. Inhaling medicines: delivering drugs to the body through the lungs. Nat Rev Drug Discov 2007;6(1):67–74.
[98] Crowder TM, Rosati JA, Schroeter JD, Hickey AJ, Martonen TB. Fundamental effects of particle morphology on lung delivery: Predictions of Stokes' law and the particular relevance to dry powder inhaler formulation and development. Pharm Res 2002;19(3):239–45.
[99] Hassan MS, Lau R. Inhalation performance of pollen-shape carrier in dry powder formulation: effect of size and surface morphology. Int J Pharm 2011;413(1–2):93–102.
[100] Choy YB, Prausnitz MR. The rule of five for non-oral routes of drug delivery: ophthalmic, inhalation, and transdermal. Pharm Res 2011;28(5):943–8.
[101] Byron PR. Drug delivery devices: issues in drug development. Proc Am Thorac Soc 2004;1(4):321–8.
[102] Costantino HR, Illum L, Brandt G, Johnson PH, Quay SC. Intranasal delivery: physicochemical and therapeutic aspects. Int J Pharm 2007;337(1–2):1–24.
[103] Romeo VD, deMeireles J, Sileno AP, Pimplaskar HK, Behl CR. Effects of physicochemical properties and other factors on systemic nasal drug delivery. Adv Drug Deliv Rev 1998;29(1–2):89–116.
[104] Illum L. Nasal drug delivery: possibilities, problems and solutions. J Control Release 2003;87(1–3):187–98.
[105] Dhuria SV, Hanson LR, Frey II WH. Intranasal delivery to the central nervous system: mechanisms and experimental considerations. J Pharm Sci 2010;99(4):1654–73.
[106] Hughes PM, Olejnik O, Chang-Lin JE, Wilson CG. Topical and systemic drug delivery to the posterior segments. Adv Drug Deliv Rev 2005;57(14):2010–32.
[107] Del Amo EM, Urtti A. Current and future ophthalmic drug delivery systems: a shift to the posterior segment. Drug Discov Today 2008;13(3–4):135–43.
[108] Prausnitz MR, Noonan JS. Permeability of cornea, sclera, and conjunctiva: a literature analysis for drug delivery to the eye. J Pharm Sci 1998;87(12):1479–88.
[109] Ahmed I, Gokhale RD, Shah MV, Patton TF. Physicochemical determinants of drug diffusion across the conjunctiva, sclera, and cornea. J Pharm Sci 1987;76(8):583–6.
[110] Gore RM, Levine MS, Ghahremani GG. Drug-induced disorders of the stomach and duodenum. Abdom Imaging 1999;24(1):9–16.
[111] Sinkula AA, Yalkowsky SH. Rationale for design of biologically reversible drug derivatives: prodrugs. J Pharm Sci 1975;64(2):181–210.
[112] Boucher BA. Fosphenytoin: a novel phenytoin prodrug. Pharmacotherapy 1996;16(5):777–91.
[113] Milne CP, Bruss JB. The economics of pediatric formulation development for off-patent drugs. Clin Ther 2008;30(11):2133–45.
[114] Zajicek A, Fossler MJ, Barrett JS, Worthington JH, Ternik R, Charkoftaki G, et al. A report from the pediatric formulations task force: perspectives on the state of child-friendly oral dosage forms. AAPS J 2013;15(4):1072–81.
[115] Sinkula AA. Design of improved taste properties through structural modification. In: Roche EB, editor. Design of biopharmaceutical properties through prodrugs and analogs. Washington, DC: American Pharmaceutical Association; 1977. p. 442–5.
[116] Ogata T, Koide A, Kinoshita M, Ozeki T. Taste masking of propiverine hydrochloride by conversion to its free base. Chem Pharm Bull (Tokyo) 2012;60(8):976–84.
[117] Rahman Z, Zidan AS, Khan SR, Reddy IK, Khan MA. Cholorpheniramine tannate complexes: physicochemical, chemometric, and taste masking evaluation. Int J Pharm 2012;436(1–2):582–92.
[118] Chrzan K, Haribhakti R, Mermey M, Porcello CJ, Silvey GL, Tran V, et al. Fexofenadine suspension formulation. US2008/0299211 A1 (Patent) 2008.

CHAPTER

33

Discover a Drug Substance, Formulate, and Develop It to a Product

Pierre deMontigny[1], *David Harris*[1], *Chris Ho*[1], *Franz Weiberth*[1], *Bruno Galli*[2] *and Bernard Faller*[3]

[1]Sanofi, Waltham, MA, U.S.A; [2]Novartis Pharma AG, Lichtstrasse, Basel, Switzerland; [3]Novartis Pharma AG,Werk Klybeck Klybeckstrasse, Basel, Switzerland

OUTLINE

I. INTRODUCTION

Although we have made significant progress in the understanding of biology and the molecular basis of disease over the last four decades, the discovery and development of drugs remains a very challenging and costly enterprise. The average cost to develop a drug—including failures—rose from $140 million in the 1970s to $1.2 billion in the 2000s and has continued to rise [1,2]. In November 2014, the Tufts Center for the Study of Drug Development reported the cost to be $2.6 billion per approved compound based on the estimated average out-of-pocket cost of $1.4 billion and time costs (expected returns that investors forego while a drug is in development) of $1.2 billion [3]. The number of new molecular entities approved by health authorities declined from 1998 to 2008 and has remained roughly constant since. Attrition rates, development times, and expenditures have all continued to increase, however [4]. One explanation for this lower productivity is the industry's focus away from new drugs providing incremental improvement and on to difficult complex targets that may provide important breakthrough therapies. Another is the higher standards to which potential new drugs are subjected by health authorities [5]. Companies continue to look to innovation, the introduction of new technologies, and the development of new strategies to shorten cycle times and reduce costs.

This chapter presents information on the discovery and development process. It provides a brief description of aspects of discovery, drug-substance development, drug-product formulation, analytics, and health-authority

submission. It draws attention to some of the strategic technical points that are important for a project to progress effectively. Links are provided to essential regulatory documents of the European Regulatory Agency, the FDA, and the ICH.

II. THE DISCOVERY PHASE

The discovery phase includes several important steps leading to the selection of a new drug candidate for development. These steps include lead identification, lead optimization, and candidate selection.

A. Lead Identification

In the past, therapeutic breakthroughs often emerged from unexpected experimental results rather than rational drug discovery. Nowadays, molecular genetics and pathway screening are new technologies that are helping to identify new targets of interest. An analysis of the origins of the 113 first-in-class drugs approved by the FDA from 1999 to 2013 showed that seventy-eight were discovered through a hypothesis-driven molecular/pathway target based approach (forty-five small molecules and thirty-three biologics) [6]. Only eight were discovered by phenotypic screening, defined as testing a large number of compounds using a target agnostic assay that monitors a phenotypic change. Most of the remaining thirty-three were discovered by identification of an active substance from a plant or other source with known pharmacological activity (e.g., aspirin) or the derivatization of a pharmacologically active natural substance (e.g., steroids). Lead identification is a discovery phase that is expected to address important questions, such as:

1. How well do we know the disease we plan to treat?
2. Can we develop assays to follow the disease in animal models that are relevant to humans?
3. Are there models mimicking the human disease?
4. Can we study disease-specific perturbations from the nondisease state?
5. Did we identify biomarkers linked to the disease state that can provide a clear understanding of the changes of the clinical process?

Recent advances in biomarker discovery, proteomics, and metabolomics have improved our efforts in translational science to understand the disease state better, but it remains a challenging area [7]. Once the assay(s) are in place, the attempt is made to identify lead compounds by screening compound collections of the company or collaborating entities including, academia. The size, quality, and chemical diversity of the accessible compound collections obviously play a major role in the quality of the lead identification phase. A common problem in high-throughput screens is the potential instability and subsequent loss of purity of compounds while archived in solution phase. Compound solubility is also important. Poorly soluble compounds may not register as lead compounds due to their inability to achieve sufficient concentrations in the screen. In some cases re-synthesis is necessary.

Fragment-based drug discovery (FBDD) has emerged in recent years as an alternative to high-throughput screening for lead compounds [8]. FBDD identifies low molecular weight ligands that bind to biologically important molecules, and uses X-ray crystallography or NMR spectroscopy to understand their binding interactions to provide information that facilitates the optimization of their potency. The FBDD approach requires fewer compounds to be screened with a lower potency requirement than with the HTS approach. As these fragment ligands are of low molecular weight, they have space to add moieties that provide improvement in potency or other important properties such as solubility.

B. Lead Optimization to Candidate Selection

Lead compounds from the lead identification phase are then clustered based on specific criteria, such as structural similarities and desirable pharmacological profiles extracted by data mining the lead compounds. Selected lead compounds are subjected to additional screenings. Early structure–activity relationships (SAR) emerge. A chemistry program is initiated to synthesize analog to improve the molecular properties. The main goal of this phase of discovery is to optimize the structure of the lead series with respect to potency, selectivity, absorption, distribution, metabolism, elimination, and toxicity (ADMET) to select the best drug candidate. It is challenging to optimize multiple parameters simultaneously, because the objectives often clash. For protease inhibitors, for example, potency and permeability are often in conflict because of the

chemical constraints imposed by the pharmacophore. Binding is often increased through addition of hydrogen bond donor or acceptor atoms; the same features usually hinder permeability and thus oral absorption. Similarly, there is a balance between binding and solubility. Binding is usually gained during lead optimization by increasing lipophilicity (this is a simple way to gain binding), which in turn decreases aqueous solubility and hence drug disposition. In the past, one first optimized potency and selectivity. The ADMET properties were addressed at a later stage. The drawback of this approach was that, when one eventually focuses on the ADMET properties, it was usually too late to introduce changes in the molecule. Today one tries to optimize ADMET properties and potency in parallel. This new strategy calls for the development of tools that allow high-throughput profiling and have compound requirements compatible with drug development. It is usually helpful to develop a blueprint with the criteria for the selection of the best drug candidate. Because many drug candidates are sparingly soluble, it is often necessary to devise a dosage form even in the early phase of drug discovery. The early formulation development studies are important for getting good quality data on dose–response and exposure in animal models, particularly when the therapeutic dose is relatively high. These so-called PK/PD (pharmacokinetics/pharmacodynamics) experiments initiate the transition from the research to development stages. A reasonable, well-tolerated formulation at this stage allowing continuity in the subsequent stages significantly adds confidence to the results obtained. Early formulation ingredients should ideally not affect the animal model itself and should be tolerated with repeated and even chronic administration for toxicology evaluation. At this point in time, it is appropriate to develop a patent strategy with the legal group to ensure that the research efforts are properly protected in a timely manner.

C. Research–Development Interface

The research–development interface must be carefully managed so that the transition into full development can progress as quickly and effectively as possible because of two main drivers. First, there is fierce competition amongst pharma and biotech companies to advance their respective candidates. The company that first commercializes next-generation or novel therapies is rewarded with significant advantages in marketing and potential sales by virtue of having first access to patients. Second, the duration of marketing exclusivity of the active compound prior to patent expiry and resulting competition from generic manufacturers is maximized. Drug-substance (DS) supplies are frequently on the critical path at the research-development interface. This is especially true for complex molecules obtained only after long syntheses (e.g., >10 synthesis steps) that may take more than 6 months to execute. One strategy to facilitate project progression is to anticipate the DS requirements and initiate early synthesis as soon as practical to ensure DS supplies are available to support toxicological studies as a lead compound transitions to a potential Drug Candidate. Depending on the urgency, prioritization, confidence in early data and available budget and other resources within a company, a synthesis may even be initiated at risk, that is, prior to candidate selection. This investment in early chemical development will provide an improved synthesis with more scalable processes, for example, by eliminating chromatographic purification steps and highly hazardous reagents that are commonly employed in the original medicinal chemistry route as a means to expedite and ensure ample supplies of DS. Initial quantities of DS (10–50 g) are needed to supply early toxicological studies that support drug-candidate selection. This is followed by further up-scaling of the chemistry to provide DS to support IND-enabling toxicology studies and Phase 1 supplies (1–5 kg).

Ideally, the selected compound should be crystalline with desirable physical and chemical stability properties. Aqueous solubility may not be sufficient, however, requiring an enabling formulation to improve the dissolution characteristics necessary to achieve improved exposure *in vivo*.

Initial solubility determinations tend to be overly optimistic. For example, solubility and dissolution rates of actives may change—even by an order of magnitude—compared with earlier determinations performed with non-crystalline, less pure batches. Evaluation of different salts or other improved forms of the active compound including polymorphism screening, especially for poorly soluble compounds, can be initiated even with relatively small amounts of material (10 mg). A drug candidate molecule can exhibit polymorphic and pseudopolymorphic forms (e.g., hydrates and solvates), which need to be characterized. The preferred form of the drug candidate that will advance into development will have the optimal combination of drugability attributes with respect to solubility, dissolution rate, physical and chemical stability, hygroscopicity, and other physicochemical properties. Some of the more obscure properties start to become important, like adhesiveness and flow property, which can lead to substantial loss of the active compound through physical adsorption on the walls of vials, tubes, or pipettes. Particularly highly potent lipophilic compounds subject to such phenomena may lead to problems such as erratic dosing.

D. Learning Experiences

One first needs to stress that risk-taking is inherent to pharmaceutical discovery, and therefore one often needs to take conscious risks when one selects a development candidate. Although challenging compounds have been successfully developed (with massive investments in technical R&D) in the past (e.g., cyclosporine, tamoxifen, amiodarone), the question is how many of these difficult compounds can one afford in a drug-development pipeline? For each successful compound, how many failed despite significant (or sometimes desperate) efforts? From our experience in the past few years, a few points emerge:

1. Solubility alone is not necessarily meaningful; it must be considered with the intended dose, permeability, and the solubility/lipophilicity balance needs to be taken into account (less permeable, low soluble, low lipophilic compounds are the most challenging). In general, one should try to investigate early what limits the solubility of the compound (e.g., lipophilicity, wettability, crystal packing, other factors) as it translates into different types of additional investigations and risks.
2. Extreme lipophilicity (log D (7.4) > 5) continues to be a cause of compound termination and disappointment for research teams because pharmacological activity observed in early animal models cannot be translated to larger animal species.
3. It is difficult to have a suitable pharmaceutical salt formation, in particular with weak bases (p$K_a < 5$).

III. DEFINING EXPERIMENTAL FORMULATIONS, THE CREATIVE PHASE

A. Basic Thoughts on the Development Phase

An early development strategy is important in preparing for successful clinical trials and regulatory approvals. This strategy will guide the development scientists to plan the laboratory studies needed to provide a formulated product with the necessary quality and stability attributes that are critical for early development, including the judicious selection of formulation excipients that are fully compatible with the drug substance.

B. What is the Purpose of a Formulation?

A formulation is mainly intended to deliver an accurate dose of drug substance through a specific route of administration to achieve desirable a pharmacological effect. The physical and chemical properties of the compound in the formulation should be well-maintained.

Figure 33.1 shows the main interrelated factors that lead an individual substance to the appropriate formulation and subsequently to a dosage form. At early stages, the efforts will primarily consist in trying to understand and combine "basic factors" or the physicochemical properties of the DS, such as to readily set up a useful, versatile, and reliable formulation strategy. For an oral solid product, this translates into a maximization of solubility and—with it—possibly also its dissolution rate and *in-vivo* absorption. The mentioned "derived factors" (properties of the chosen dosage form-see figure) will predominantly have to be investigated, understood gradually, and validated later, possibly in parallel or later after optimizing the chemical and physical properties of the drug substance (see also Figure 33.4).

Most orally administrated compounds absorb through the gastrointestinal tract by passive diffusion. One of the potential tasks for a formulation is to improve in-vivo absorption. This is especially important for compounds with low solubility, slow dissolution rates, and low permeability. In some cases, enabling formulations are required to rescue compounds by compensating for their poor physicochemical characteristics. A good and integrated formulation approach should therefore always consider physicochemical properties of the compound first, relating them to the key information of any formulation activity, namely the envisaged dose.

However, biological factors such as metabolism (in, for example, the liver or the gastrointestinal tissues) and active molecule transporters like P-gP or MRP (in, for example, the gastrointestinal tract) influence the bioavailability and can mislead the prediction of the dose. The influence of such active transports is particularly noted if the main driving forces of the absorption process, such physicochemical parameters like solubility, dissolution rate, and permeability, are low.

The preferred route of administration is normally oral especially if systemic delivery of the drug is desired. However, due to the desire to target certain tissues such as the lung (by inhalation) or due to the properties of the

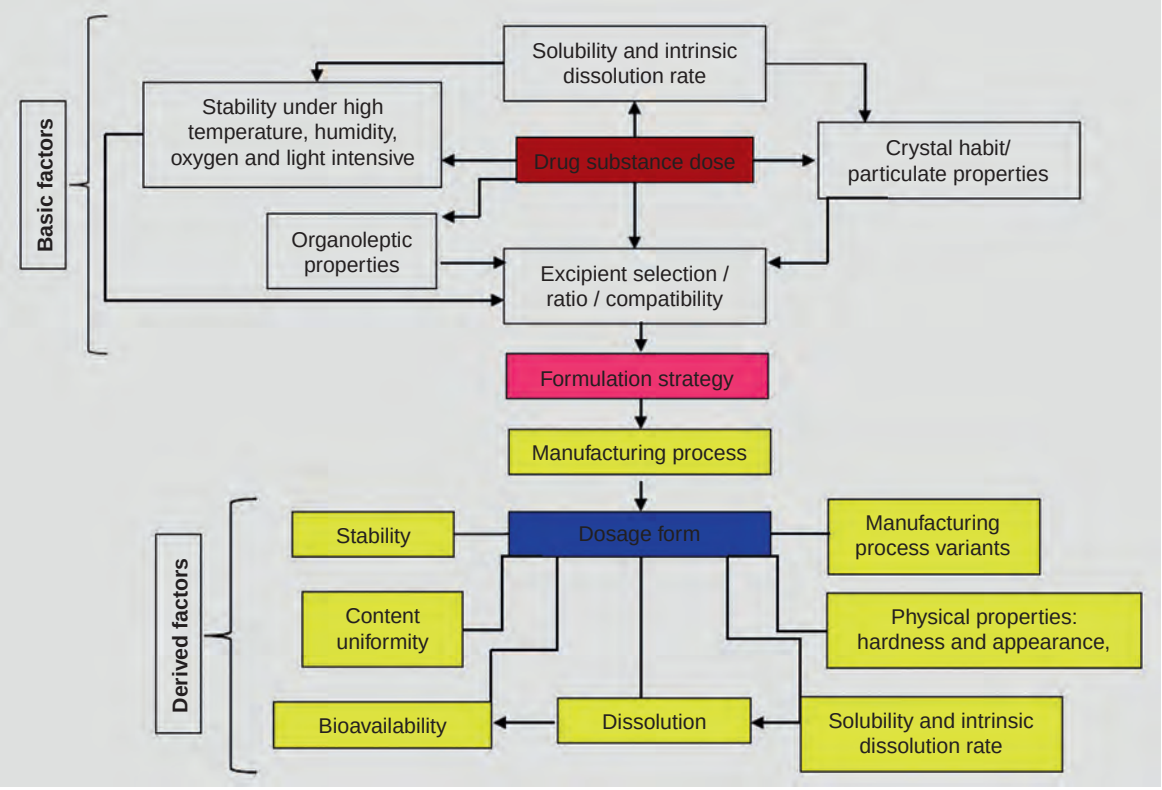


FIGURE 33.1 Main interrelated factors surrounding an oral formulation.

TABLE 33.1 Approximate Dosages in Human Depending on the Administration Route Chosen [14]

Administration route	"Realistic" dosage related to drug substance (human)
Oral	0.25–1,000 mg
Passive transdermal	0.05–5 mg
Inhalation	0.01–5 mg
Nasal, buccal, and sublingual	0.05–2 mg
Injection	Mainly depending on volume that can be administered and on physicochemical properties.

molecule such as instability in the GI tract (protein therapeutics are often administered by injection) other routes of administration are chosen. Table 33.1 provides a rough idea of the achievable doses related to the administration route.

C. Suggested Sequence of Activities Prior to Start Formulation

Prior to the start of multiple formulation efforts and intensive work for *in vivo* formulation optimization experiments, an assessment should be made on whether drug metabolism and transport are relevant pharmacokinetic driving factors for the substance to be formulated. This is done stepwise using following procedures:

1. *In silico* estimation of the physicochemical properties like calculated pK_a/Log *P* and polar surface area [9].
2. Experimental determination of physicochemical properties like pK_a/Log *P*, solubilities in biorelevant fluids and physiological pH, chemical stability, or even special amphiphilic properties [10].

3. Assessment of the ability of the compound to pass through membranes using *in vitro* models like PAMPA/CaCo-2/MDCK cell layers, together with a well-documented reference compound like mannitol [11].
4. Assessment *in vitro* of the metabolic pattern, biological stability, protein binding, and blood/plasma partitioning *in vitro* and on a need basis followed by verification *in vivo*.
5. *In silico* estimation of *in vivo* absorption characteristics, including the effects of dose and particle size of the drug substance on absorption.

D. Biopharmaceutical Classification of Compounds

Provided that metabolism and active transport systems are not the dominating factors for the drug absorption process, two factors fairly well characterize the oral absorption: solubility and permeability through membranes. Consequently, a Biopharmaceutics Classification System (BCS), which is a scientific framework for categorizing four classes of drug substance, has been defined by Amidon et al: [12]

Class 1. High Solubility—High Permeability (e.g., Metoprolol, L-Dopa)

Formulation principles chosen for Class I compounds are normally straightforward. Aqueous solutions and *in situ* suspensions cover the needs for early formulation.

Class 2. Low Solubility—High Permeability (e.g., Carbamazepine, Nifedipine)

Generations of pharmacists have struggled to overcome the difficulties related to this class of compounds. Scientific progress over the last two decades in the formulation field uncovered many interesting principles useful in the formulation of Class II compounds. Thus, handling of this class is now relatively well understood. The success is mainly due to improvements in controlling the solubility and dissolution rate by measures such as the selection of the appropriate salt, the use of the appropriate crystal form, micronization or nanonization, controlling of the wettability, the addition of surfactants, and the use of microemulsions. These principles have become the basis for adequate formulations. Highly dosed compounds (low potency) of this class of substances may face formulation limits regarding feasibility and attractiveness due to the massive use of special or new excipients needed to counterbalance the unfavorable properties.

Class 3. High Solubility—Low Permeability (e.g., Atenolol, Terbutaline)

The bioavailability of Class III compounds cannot practically be influenced by a formulation. Rarely, the bioavailability can be improved by increasing the dose or by modifying the permeability of the gastrointestinal membrane. Modification of the gastrointestinal membrane structure by penetration or absorption enhancers always bears a safety risk, because it may also improve the uptake of all kinds of toxins from the intestinal lumen and potentially cause severe irritations. What may be acceptable for an acute therapy may not be suitable for chronic treatment.

Class 4. Low Solubility—Low Permeability

Real Class IV compounds are not likely to be formulated. For highly potent drugs, there might be the opportunity to raise the dose, but in general one starts seeking suitable backups and switching to a different chemical class. Changing the route of administration might be viable (e.g., parenteral), but this has a strong impact on the attractiveness of the product or significant limitations.

A revised classification system, the Developability Classification System (DCS), was defined by Butler and Dressman [13] for oral drugs using BCS as a starting point. The revised system is designed to have a greater focus on drug developability. Intestinal solubility, the compensatory nature of solubility and permeability in the small intestine, and an estimate of the particle size needed to overcome dissolution rate limited absorption were all considered in the revised system. DCS can be used to identify factors limiting oral absorption and aid formulation choice for new drug candidates in early development.

For the DCS, the fasted state simulated intestinal fluid (FaSSIF) instead of buffers (pH 1.0−7.5) is used to estmate jejunal solubility. Assume 500 mL instead of 250 mL as a stomach volume to dissolve the drug substance. Classification of 2a and 2b are defined to distinugish between dissolution rate and solubulity-limited drug substance. Human jejunal permeability is predicted from a cell line measurement or from *in silico* estimation based upon its structure and log D.

E. How Do We Proceed at a Practical Level?

Solubility always has to be considered together with the envisaged dose (anticipated needed dose and reference volume of a stomach where it has to dissolve). This leads us to define:

$$Dose\ number = M_o/(V_0{}^*C_s)$$

M_o = dose to be administered
V_o = 250 mL (by definition, the reference stomach volume)
C_s = solubility (lowest solubility within physiological pH (1.0−8.0))

Dose numbers >20 are by definition related to "low soluble and dissolution rate limited" substances (see Amidon et al [12]). In such cases, special formulation approaches are to be considered from the scratch in the formulation strategy. A permeability assessment at an early stage can be done using, for example, the CaCo-2 cell model and a reference substance. Low permeable substances can be defined as those less permeable than, for example, mannitol.

Low permeable compounds always create serious formulation problems, whether it is due to the physico-chemistry or due to active transport systems. For both, the CaCo-2 model provides useful information. However, the crucial issue at this stage is to sort out in view of the evaluation of alternative routes whether the observed low bioavailability is due to metabolism, low permeability or both. Thus, permeability is always complex and not straightforward to assess.

In summary this leads to (see also Lipper [14]):

- Poor biopharmaceutical properties may sometimes be corrected by formulation, but at the expense of time and resources.
- Poor solubility and stability may be amenable to being fixed by the formulation.
- Poor permeability is very difficult to correct by formulation.
- First-pass metabolism problems cannot be fixed by oral formulations.

F. Which Formulation Principles are Used?

At early stages, stable solutions are preferred because of their defined physical state, which is assumed to be a molecularly dispersed system and due to ease of preparation and administration. However, this approach can be used only if the compound has reasonable solubility in buffer or in cosolvent to achieve the target concentration that is suitable for the study.

If the target drug concentration cannot be achieved in buffer or cosolvent, simple micellar-dispersed solutions are also directly usable by the experimenter. Concentration of surfactant in the solution should be high enough to form micelles that can solubilize the drug substance but not high enough to cause toxicity.

Complex micellar systems like microemulsions or special solid systems (nanosuspensions or solid dispersions) usually require a basic pharmaceutical development strategy and sufficient understanding of the technical and biological concept to ensure appropriate quality, safety, tolerability, and reproducibility of the formulation. Such sophisticated systems can usually be properly elaborated only at later development stages because of the lack of sufficient drug substance at early stages.

Microemulsion is a clear, transparent, thermodynamically stable dispersion of oil and water stabilized by an interfacial film of surfactant, frequently in combination with co-surfactants such as alcohols. Particle size of a microemulsion is generally $< 0.1\ \mu m$. It is an excellent solubilizing agent for hydrophobic drugs but requires a relatively large amount of surfactant (>10 percent) which must be nontoxic for use in pharmaceutical applications. The need for pharmaceutically acceptable ingredients limits the choice of components and leads to difficulties in formulation. Microemulsion stability can be influenced by environmental parameters such as temperature and pH [15].

Solid dispersions contains drug substance dispersed within an inert solid carrier matrix—typically water-soluble polymers—prepared by either melting (such as hot melt extrusion) or solvent-evaporating methods (such as spray drying). The drug is either suspended in the carrier as phase-separated crystalline material (undesirable) or a homogeneous molecular mixture of amorphous particles (desirable). The carrier can exist either in the amorphous or crystalline form. Solid dispersion that contains a drug in the amorphous form usually shows much faster dissolution rates as compared to a drug in crystalline form [16]. Stabilized amorphous solid dispersions can potentially accomplish the following objectives.

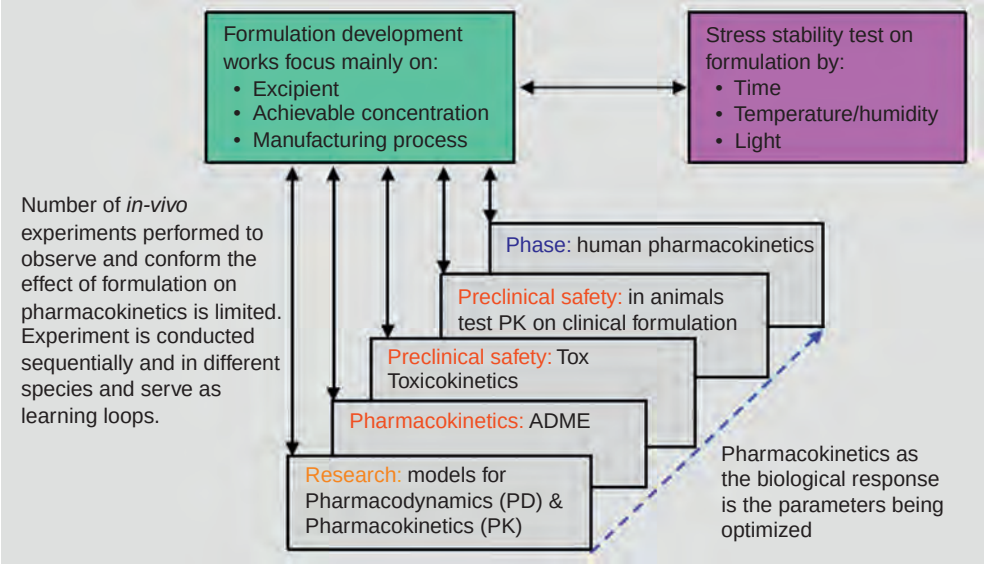


FIGURE 33.2 Iterative loops supporting the optimization of a formulation.

- Enhance the oral absorption of poorly water-soluble compounds by attaining and sustaining a supersaturated concentration of drug in gastrointestinal (GI) fluid.
- Provide a physically stable drug form (avoiding crystallization or phase separation of amorphous drug) that enables processing of the dispersion into solid dosage form.
- Apply to structurally diverse insoluble compounds across a wide range of physicohemical properties such as high melting point, weak base with low pKa, and log P.

In early preclinical drug testing, it is preferable to use excipients that can and will be used in later development phases. Whenever possible, the formulation principle and the excipients should be used consistently throughout the preclinical testing. This applies to all relevant studies with experimental animals, namely pharmacological, toxicological, and pharmacokinetic studies.

Toxicology studies often require high doses of poorly soluble compounds that open a permanent conflict between the maximal volume that can be administered and the safety/tolerability of the excipients needed. This situation will require serious biopharmaceutical problems to be resolved early and to have them resolved iteratively in subsequent pharmacokinetic experiments, applying the lessons learned from the human formulation studies. The iterative loops supporting formulation optimization are illustrated in Figure 33.2

Overall, the main challenges in predevelopment consist of distilling useful information as early as possible from the many experiments and developing an integrated formulation, dosage form, and drug-product concepts. In essence, this effort should lead to a good anticipation of the first human dose (even if only extrapolated from animals at this point) with a sufficient safety margin, the most appropriate route of administration, and a consistent formulation approach for the subsequent clinical trials (bullet point 1 in Figure 33.4).

IV. PREPARATION FOR A NEW DRUG-PRODUCT LAUNCH

With the anticipated increase in successful clinical trials to demonstrate the safety and efficacy of the new drug candidate, preparation for product launch will require stellar execution. Process scalability for the drug substance and the drug product with a good projection of the anticipated demand for the new product will require good planning using efficient, lean processes to ensure market readiness. Successful dossier submissions to the agencies will require documents that meet reviewers' expectations and successful pre-approval manufacturing site inspections. These activities will be facilitated by the proper anticipation of all the groups involved, including research, development, and production.

After the positive outcome of the clinical Phase IIa or proof-of-concept studies and commitment by the organization and stakeholders to continue with the development activities, selecting the final synthesis route and

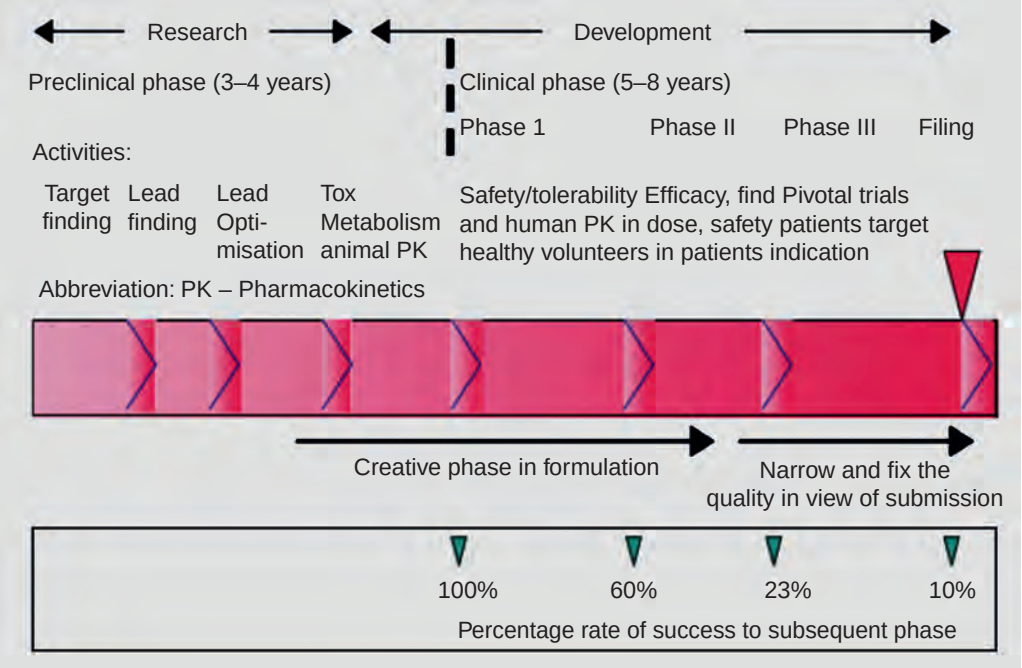


FIGURE 33.3 Main stages of drug development.

commercial formulation anticipated for Phase III studies and product launch will require significant preparation. Regulatory agencies require that manufacturing processes be designed and controlled to assure that in-process materials and finished products meet predetermined quality requirements aligned with the ICH and other regulatory guidelines [17–19]. Substantial resources will be engaged to validate the manufacturing processes for both drug substance and drug product, including identifying and controlling key process parameters, justifying and establishing specifications, developing and validating analytical methods, and manufacturing validation batches at defined launch sites. In addition, long-term stability studies need to be initiated early so that they do not become time-critical for filing.

V. CONCLUSION: DRUG DISCOVERY AND DEVELOPMENT IN INDUSTRY

The impact of pharmaceutical development in industry has been extraordinary in the past fifty years, as highlighted by Rowland et al [20]. The industry has made many important contributions to improve the health of patients. The discovery phase has improved with a better understanding of the disease biology map and with advances in translational sciences using proteomics, genomics, metabolomics, and bioinformatics. Personalized medicine will likely replace traditional medicine, with specific treatment regimens based on the patient's genetic makeup. Consequently, the pharmaceutical industry must explore new development opportunities differently, with external collaborations between large and small pharmaceutical and biotechnology companies.

The pharmaceutical development stages provided in Figure 33.3 and Figure 33.4 will change to diminish the attrition of new drug-candidate molecules with better translational data in research to address the disease state more effectively in the preclinical phases. The development phases will be expected to progress much more rapidly, with a higher rate of success, to product launch. However, it is very important to proceed with balanced risk/resource management, always considering the relative chance of success at each stage. In this situation, it is mandatory to constantly monitor the competitive situation as well.

At a later stage—from clinical Phase II onwards—it will be much more difficult to perform changes on the drug-product concept and its quality for regulatory and logistic reasons. Narrowing the specifications and the process parameters while fixing quality in preparation for the submission and forthcoming inspections by the health authorities will be the main drivers of the process from this point on. Normally at the beginning of the Phase IIb clinical studies, the dosage form should be defined to carefully anticipate the Phase III confirmatory

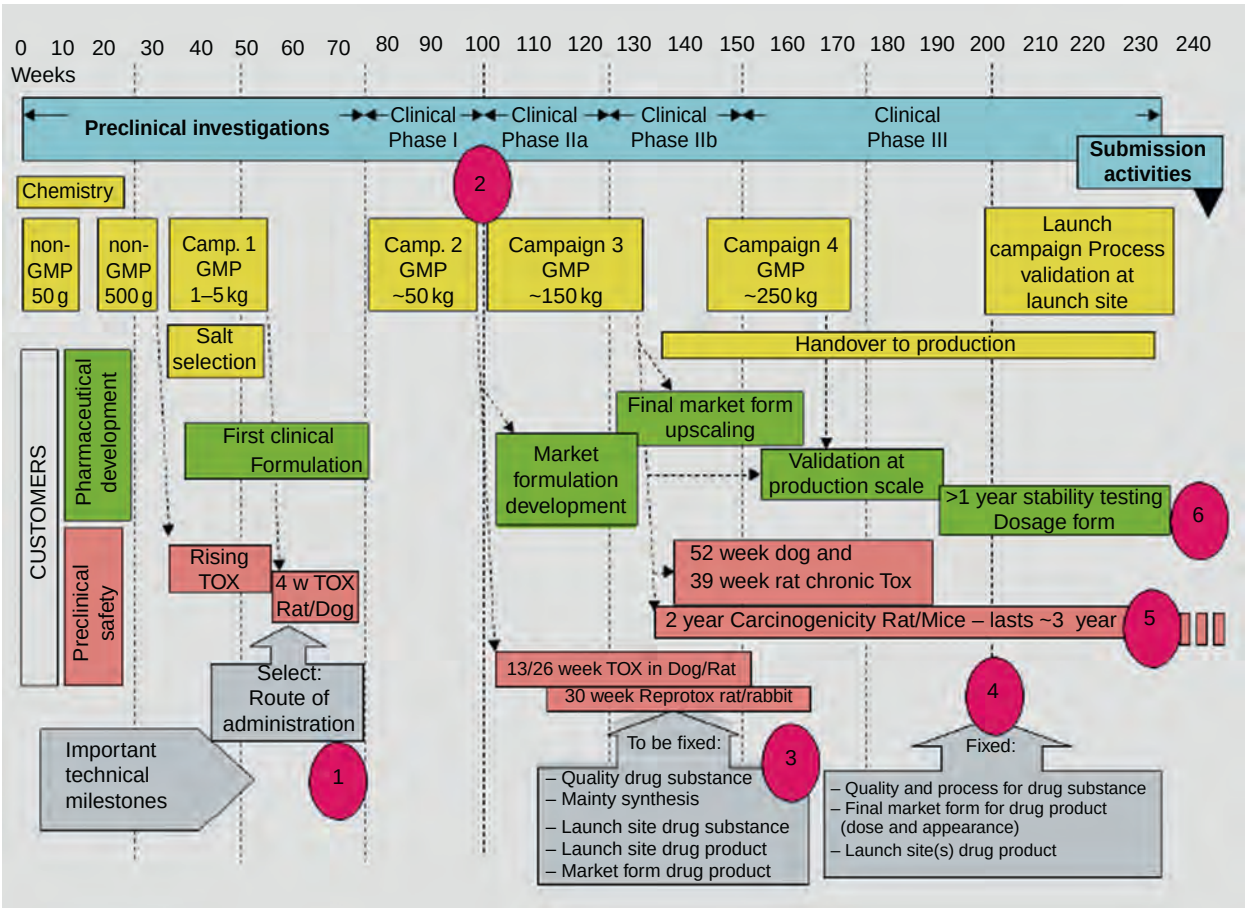


FIGURE 33.4 Scheme for a fast drug product development showing the interdependence of selected key activities.

clinical trials. It is very important to avoid significant formulation changes at this stage as they may require time-consuming, costly, and risky studies to prove that the formulations are bioequivalent.

In conclusion, it must be noted that the huge costs associated with large clinical Phase III trials have led many companies in the recent past to implement much more rigorous drug-development selection processes to identify the most promising projects as early as possible. An interesting review of the various current philosophies is provided by Shillingford et al [21].

Reference

[1] DiMasi JA, Grabowski HG. The cost of biopharmaceutical R&D: is biotech different? Managerial Decis Econ 2007;28(4–5):469–79.
[2] DiMasi JA, Hansen RW. The price of innovation: new estimates of drug development costs. J Health Econ 2003;22(2):151–85.
[3] Tufts Center for the Study of Drug Development. Cost to Develop and Win Marketing Approval for a New Drug Is $2.6 Billion. Available from: <http://csdd.tufts.edu/news/complete_story/pr_tufts_csdd_2014_cost_study/>; 2014.
[4] Pammolli F, Magazzini L, Riccaboni M. The productivity crisis in pharmaceutical R&D. Nat Rev Drug Discov 2011;10:428–38.
[5] Cook D, Brown D, Alexander A, March R, Morgan P, Satterhwaite G, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: a five-dimensional framework. Nat Rev Drug Discov 2014;13:419–31.
[6] Eder J, Sedrani R, Wiesmann C. The discovery of first-in-class drugs: origins and evolution. Nat Rev Drug Discov 2014;13:577–87.
[7] Anderson D, Kodukula K. Biomarkers in pharmacology and drug discovery. Biochem Pharmacol 2014;87:172–88.
[8] Murray C, Rees D. The rise of fragment-based drug discovery. Nat Chem 2009;1:187–92.
[9] Palm K, Stenberg P, Luthman K, Artursson P. Polar molecular surface properties predict intestinal absorption in humans. Pharm Res 1997;14:568–71.
[10] Kibbey C, Poole S, Robinson B, Jackson JD, Durham D. An integrated process for measuring the physicochemical properties of drug candidates in a preclinical discovery environment. J Pharm Sci 2001;90(8):1164–75.
[11] Arthusson P, Karlsson J. Correlation between oral absorption in humans and apparent drug permeability coefficient in human intestinal epithelial (CaCo-2) cells. Biophys Res Commun 1991;157:880–5.

[12] Amidon G, Lennernas H, Shah VP, Crison JR. Theoretical basis for a biopharmaceutical drug classification: correction of in vitro drug product dissolution and in vivo bioavailability. Pharm Res 1995;12(3):413–20.
[13] Bulter JM, Dressman JB. The development classification system: application of biopharmaceutics concept of formulation development. J Pharm Sci 2010;99:4940–54.
[14] Lipper R. How can we optimize selection of drug development candidates from many compounds at the discovery stage? Mod Drug Discov 1999;2(1):55–60.
[15] Constantinides PP. Lipid microemulsion for improving drug dissolution and oral absorption: physical and biopharmaceutical aspects. Pharm Res 1995;12(11):1561–72.
[16] Williams HD, Trevaskis NL, Charman SA, Shanker RM, Charman WN, Pouton CW, et al. Strategies to address low drug solubility in discovery and development. Pharmacol Rev 2013;65:325–499.
[17] European Regulatory Agency. Available from: <http://www.ema.europa.eu/ema>.
[18] US Food and Drug Agency. Guidance (Drugs). Available from: <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>; 2015.
[19] Tripartite International Committee of Harmonization (Europe, USA, Japan). ICH Guidelines. Available from: <http://www.ich.org/products/guidelines.html>.
[20] Rowland M, Noe CR, Smith DA, Tucker GT, Crommelin DJ, Peck CC, et al. Impact of the pharmaceutical sciences on health care: a reflection over the past 50 years. J Pharm Sci 2012;101(11):4075–99.
[21] Shillingford CA, Vose CW. Effective decision-making: progressing compounds through clinical development. Drug Discov Today 2001;6(18):941–6.

SECTION SEVEN

Pharmaceutical and Chemical Means to Solubility and Formulation Problems

CHAPTER

34

Drug Nomenclature

Raffaella G. Balocco Mattavelli[1], *Ji-Cui Dong*[1], *Sophie Lasseur*[1], *A. Romeo*[1] *and Sabine Kopp*[2]

[1]INN Programme, World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms; [2]World Health Organization, Geneva Essential Medicines and Health Products, Regulation of Medicines and Other Health, Technologies, Technologies Standards and Norms

OUTLINE

I. INTRODUCTION

The World Health Organization (WHO) has a constitutional mandate to "develop, establish, and promote international standards with respect to biological, pharmaceutical, and similar products." One way in which this mandate is discharged is through the Programme on International Nonproprietary Names (INNs) for pharmaceutical substances. This program was established in 1950 by resolution WHA3.11 of the World Health Assembly and became operational in 1953. At its heart is the concept that there should be a single name globally for a unique pharmaceutically active substance. A substance may have several designations, including an invented name, a systematic chemical name, a laboratory code, a trivial name, a Chemical Abstracts number. Assignment of an INN allows one name to be used in literature and in different circumstances so that there is a common basis for identification.

II. TRADE NAMES AND NONPROPRIETARY NAMES

Pharmaceutical preparations are usually marketed by industry under proprietary names or trademarks. In many countries, trade names—also called trademarks or brand names—are used when prescribing, dispensing, selling, promoting, or buying medicine. Trade names are usually selected by the owner of the product and

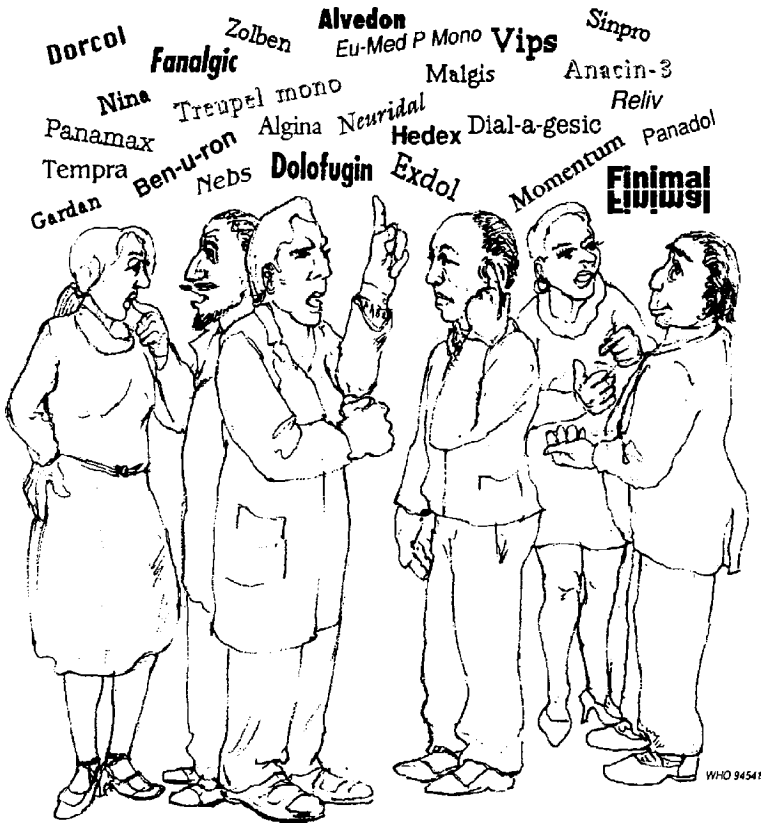


FIGURE 34.1 Various trade names for one substance; example *paracetamol*.

registered in national trademark or patent offices. They are private property and can be used only with the consent of the owner of the trademark [1,2].

In most cases, brand names are chosen for a finished pharmaceutical product, that is, for one or various active drug substances in a defined dosage form and formulation. Therefore, pharmaceutical preparations containing the same active drug substance are frequently sold under different brand/trade names, not only in different countries but even within the same country (see Figure 34.1).

In practice, this means that the number of trade names in one country is usually much higher than the number of active drug substances marketed and used. Nonproprietary names, also called generic or common names, are intended to be used as public property without restraint, that is, nobody should own any rights on their usage. These names are usually designated by national or international nomenclature commissions. Both trade names and nonproprietary names are normally published first in the form of proposals. Comments may be made and objections raised for a certain time period before final publication. Although nonproprietary names and trade names may appear similar in form to an outsider, there is, in fact, a big difference. First, nonproprietary names are designations to identify the active pharmaceutical drug substance rather than the final product. Secondly, the selection of a nonproprietary name follows established rules so that the name itself communicates to the medical and pharmaceutical health professional information about the therapeutic or pharmacological group to which the active drug substance belongs.

III. DRUG NOMENCLATURE

A. INNs for Pharmaceutical Substances

1. History

During the twentieth century, the rapid development of pharmaceutical chemistry has brought with it the need to identify large numbers of active drug substances by unique, universally available, and accepted names. The systematic chemical name, codified by international bodies including the International Union for Pure and Applied Chemistry (IUPAC) and International Union of Biochemistry (IUB), has the advantage of unambiguously

FIGURE 34.2 Various common names for one substance; example *paracetamol*.

defining a specific chemical substance, but it is often very long, difficult to memorize, and practically incomprehensible for the nonchemist. Moreover, it gives no indication as to the therapeutic action of the substance.

In order to avoid citation of difficult chemical names, generic names were created. In the beginning, however, different names were independently assigned to the same substance in different countries. For example, acetaminophen, *N*-(4-hydroxyphenyl) acetamide, 40-hydroxyacetanilide, *p*-acetamidophenol, *N*-acetyl-p-aminophenol, acetaminophen, and paracetamol are the same substance (see Figure 34.2).

When WHO started the Programme on INN, experts had to coordinate the activities of existing national nomenclature programs, which were especially active in France, the Nordic countries, the United Kingdom, and the United States. As a result of these national activities, many substances already had different, well-established national names. Members of the newly established INN program were faced with the difficulty of choosing a single name (paracetamol in the example given above; see Figure 34.3).

Since then, the activities of national commissions have been coordinated in order to achieve international standardization in nomenclature under the auspices of WHO, according to article 2a and 2u of its constitution [3]:

> In order to achieve its objective, the functions of the World Health Organization shall be: (a) to act as the directing and coordinating authority on international health work; . . . (u) to develop, establish and promote international standards with respect to food, biological, pharmaceutical and similar products . . . procedure (application, detailed steps).

Requests for recommended INNs are submitted on a form to the WHO Programme on INN (20 avenue Appia, 1211 Geneva 27, Switzerland). These requests are then submitted by the WHO Secretariat on behalf of the Director-General to the members of the WHO INN Expert Group. The following information has to be provided on the form:

- name and address of manufacturer and/or originator, including the name of the responsible person;
- suggested nonproprietary name(s) (various proposals possible);
- chemical name (following IUPAC rules) or description, and *ad hoc* information on biological products;
- molecular formula;
- graphic formula;
- stereochemical information;
- therapeutic use and pharmacological mode of action;
- code, trademark (known or contemplated);
- date of commencement of clinical trials; and
- letter from the Chemical Abstract Service (CAS) with CAS Registry Number and CA Index Name.

FIGURE 34.3 One international name for one substance; example *paracetamol*.

Each name suggested by the originator is then examined to determine whether it complies with the rules and guiding principles for the selection of an INN. When all members of the WHO INN Expert Group designated to select nonproprietary names agree on a name, it is first published as a proposed INN. During a four-month period, any person can forward comments or a formal objection (e.g., on the grounds of similarity to existing trademarks). When no objections are raised, the name is published a second time as a recommended INN and the Director-General of WHO gives notice to member states that the name has been selected by the WHO as a recommended INN.

2. *Selection Process and Selection Criteria*

General rules were established at the beginning of the INN program to allow health professionals to understand the rationale for a number of new names for pharmaceutical substances. At first, some countries used shortened chemical names as generic names, but this system was found to be very limited since many molecules contain similar elements and groups (such as phenol, chlorine, methyl, or benzyl-rings) in their chemical structures. In addition, a name that indicates a relationship to a group of pharmacologically similarly acting substances is more meaningful to users in most cases.

The following principles should generally be applied when selecting an INN. The name should: (1) be distinctive in sound and spelling, (2) not be too long, and (3) show a relationship to substances with the same pharmacological action. In addition, the new name should not conflict with any existing common names or trademarks, and patients should not be confronted with nonproprietary names that are likely to have anatomical, physiological, or pathological connotations. For example, a name starting *cancer-* would not be acceptable (see Figure 34.4).

In principle, INNs are given only to the active base or the active acid. Names for different salts or esters of the same active substance should differ only in respect of the name of the inactive moiety of the molecule. For example, *oxacillin* and *ibufenac* are INNs, and their salts are named *oxacillin sodium* and *ibufenac sodium*. The latter are also called "modified INNs" (INNM). Note that before the existence of this rule, some INNs were published for salts. The term "modified INN" may therefore sometimes be used for a base or acid. For example, levothyroxine sodium was published as INN and levothyroxine may thus be referred to as INNM.

To facilitate the transliteration and pronunciation of INNs for pharmaceutical substances certain letters, such as "h" and "k," should be avoided. Preference is given to "f" instead of "ph," "t" instead of "th," "e" instead of "ae" or "oe," and "i" instead of "y." The INN for amphetamine is therefore spelled amfetamine.

When devising an INN it is important to be aware of possible language problems. Since the name is used worldwide, not only should certain letters be avoided but experts need to be aware of unsuitable connotations in the major

General principles for guidance in devising international nonproprietary names for Pharmaceutical substances

1. International nonproprietary names (INN) should be distinctive in sound and spelling. They should not be inconveniently long and should not be liable to confusion with names in common use.
2. The INN for a substance belonging to a group of pharmacologically related substances should, where appropriate, show this relationship. Names that are likely to convey to a patient an anatomical, physiological, pathological, or therapeutic suggestion should be avoided.

These primary principles are to be implemented by using the following secondary principles:

3. In devising the INN of the first substance in a new pharmacological group, consideration should be given to the possibility of devising suitable INN for related substances, belonging to the new group.
4. In devising INN for acids, one-word names are preferred; their salts should be named without modifying the acid name, e.g. "oxacillin" and "oxacillin sodium," "ibufenac" and "ibufenac sodium."
5. INN for substances which are used as salts should in general apply to the active base or the active acid. Names for different salts or esters of the same active substance should differ only in respect of the name of the inactive acid or the inactive base.
   For quatemary ammonium substances, the cation and anion should be named appropriately as separate components of a quatemary substance and not in the amine-salt style.
6. The use of an isolated letter or number should be avoided; hyphenated construction is also undesirable.
7. To facilitate the translation and pronunciation of INN, "f" should be used instead of "ph", "t" instead of "th", "e" instead of "ae" or "oe," and "i" instead of "y;" the use of the letters "h" and "k" should be avoided.
8. Provided that the names suggested are in accordance with these principles, names proposed by the person discovering or first developing and marketing a pharmaceutical preparation, or names already officially in use in any country, should receive preferential consideration.
9. Group relationship in INN (see Guiding Principle 2) should if possible be shown by using a common stem. The following list contains examples of stems for groups of substances, particularly for new groups [list see text]. There are many other stems in active use. Where a stem is shown without any hyphens it may be used anywhere in the name.

FIGURE 34.4 General principles for guidance in devising INNs for pharmaceutical substances.

languages spoken in the world. A name may appear excellent for an English speaker but unacceptable in another language. For example, the name inglicretin could remind a French speaker of the term crétin anglais ("stupid Englishman") and might therefore not be the best choice for naming a pharmaceutical substance.

As INNs should show relationship to other substances of similar pharmacological action, common stems have been created. A large number of such common stems are in use, and new stems are created when necessary [4]. Some examples are given in Table 34.1.

Some examples of INNs ending *-entan*, the stem selected for designating endothelin receptor antagonists, are: ambrisentan, atrasentan, avosentan, bosentan, clazosentan, darusentan, edonentan, enrasentan, fandosentan, feloprentan, nebentan, sitaxsentan, tezosentan, and zibotentan.

When requesting selection of an INN, the manufacturer has often not yet finalized the precise indications for the therapeutic use of the compound. A name is usually requested during the development phase of a new compound, which means that the request is submitted to WHO during the clinical trials phase. A name is needed, however, as soon as an application for registration of a product is forwarded to the national authorities. This means that the naming process is close to all new scientific developments in the pharmaceutical field. External expertise is often needed for specific questions concerning new therapeutic groups and new types of products.

During the last few years, the selection process has become more complex. New receptors and pharmacological actions are discovered more and more frequently. This means in many cases that new stems have to be created. However, sometimes a structural relationship to existing molecules is present, and experts have to decide whether an existing stem may be used or whether a new one must be established. Fibrinogen receptor antagonists are an example. These substances act as platelet aggregation inhibitors for which the stem *-grel* existed for several years. The nomenclature experts have to decide whether the same stem should be used for the fibrinogen receptor antagonists or whether the group of new molecules is so important that a new stem needs to be established [4–6]. Guidelines have been set by the INN Expert Group on the establishment of new stems [7].

TABLE 34.1 Common Stems in INN Formation

Stem	Pharmacotherapeutic group
-ac	Anti-inflammatory agents; ibufenac derivatives
-adol/-adol-	Analgesics
-ast	Anti-asthmatic; anti-allergic substances not acting primarily as antihistaminics
-astine	Antihistaminics
-azepam	Diazepam derivatives
bol	Steroids, anabolic
-cain-	Class I antiarrhythmics; procainamide and lidocaine derivatives
-caine	Local anaesthetics
cef-	Antibiotics; cefalosporanic acid derivatives
-cillin	Antibiotics; derivatives of 6-aminopenicillanic acid
-conazole	Systemic antifungal agents; miconazole derivatives
cort	Corticosteroids, except prednisolone derivatives
-coxib	Selective cyclo-oxygenase inhibitors
-entan	Endothelin receptor antagonists
gab	Gabamimetic agents
gado-	Diagnostic agents; gadolinium derivatives
-gatran	Thrombin inhibitors; antithrombotic agent
gest	Steroids; progestogens
gli	Antihyperglycemics
io-	Iodine-containing contrast media
-metacin	Anti-inflammatory substances; indometacin derivatives
-mycin	Antibiotics, produced by *Streptomyces* strains
-nidazole	Antiprotozoals and radiosensitizers; metronidazole derivatives
-olol	Beta-adrenoreceptor antagonists
-oxacin	Antibacterial agents; nalidixic acid derivatives
-platin	Antineoplastic agents; platinum derivatives
-poetin	Erythropoietin-type blood factors
-pril(at)	Angiotensin-converting enzyme inhibitors
-profen	Anti-inflammatory agents; ibuprofen derivatives
prost	Prostaglandins
-relin	Pituitary hormone release-stimulating peptides
-sartan	Angiotensin II receptor antagonists; antihypertensive (nonpeptidic)
-vaptan	Vasopressin receptor antagonists
vin-/-vin-	Vinca alkaloids

On the other hand, a new mode of action is sometimes discovered for an existing substance. If further substances are developed with a similar mode of action, the question arises whether a new stem is needed, which would mean modifying the "old" name for the first compound in the series. For example, albifylline and pentoxifylline are *N*-methylxanthine derivatives, and the stem *-fylline* was therefore chosen for their names. These substances have also been found to suppress tumor necrosis factor-α [8]. Experts decided to retain the stem *-fylline* in this case, since the "new" action was nevertheless based on the typical xanthine-mediated inhibition of phosphodiesterase.

New approaches to naming pharmaceutical substances are being created because of increasing research using molecular design. "Simple" derivatives of known compounds are becoming more rare. Chemistry based on receptor structure and molecular design focuses more on synthesizing compounds to fit receptor binding sites. This means that nomenclature will move in the same direction. Chemical relationships will need to be looked at from a different standpoint, and the pharmacological activity might have to be considered in almost all cases as a basis for assigning a given substance to a group.

3. Procedure

The Annex reproduces the *Procedure for the Selection of Recommended INNs for Pharmaceutical Substances* as adopted by the WHO Executive Board in its resolution EB15.R7 and as amended in 2005 by resolution EB115.R4.

4. Publication

After acceptance of the selected name by the request originator, it is included in a list of proposed INNs, which is published in the *WHO Drug Information* [9,10]. For example:

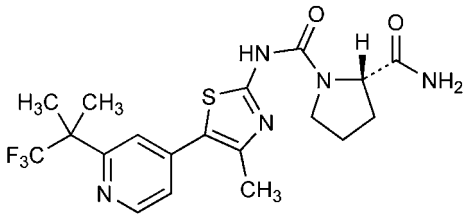
alpelisibum

alpelisib (2*S*)-N^1-{4-methyl-5-[1-(1,1,1-trifluoro-2-methylpropan-2-yl)pyridin-4-yl]-1,3-thiazol-2-yl}pyrrolidine-1,2-dicarboxamide *antineoplastic*

alpélisib (2*S*)-N^1-{4-méthyl-5-[1-(1,1,1-trifluoro-2-méthylpropan-2-yl)pyridin-4-yl]-1,3-thiazol-2-yl}pyrrolidine-1,2-dicarboxamide *antinéoplasique*

alpelisib (2*S*)-N^1-{4-metil-5-[1-(1,1,1-trifluoro-2-metilpropan-2-il)piridin-4-il]-1,3-tiazol-2-il}pirrolidina-1,2-dicarboxamida *antineoplásico*

$C_{19}H_{22}F_3N_5O_2S$ 1217486-61-7



Two lists of proposed INNs are published per year. If no objection has been raised during a four-month period following publication, the proposed name is published a second time as a recommended INN, as shown in the following example:

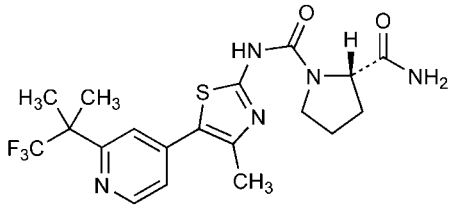
alpelisibum

alpelisib (2*S*)-N^1-{4-methyl-5-[1-(1,1,1-trifluoro-2-methylpropan-2-yl)pyridin-4-yl]-1,3-thiazol-2-yl}pyrrolidine-1,2-dicarboxamide

alpélisib (2*S*)-N^1-{4-méthyl-5-[1-(1,1,1-trifluoro-2-méthylpropan-2-yl)pyridin-4-yl]-1,3-thiazol-2-yl}pyrrolidine-1,2-dicarboxamide

alpelisib (2*S*)-N^1-{4-metil-5-[1-(1,1,1-trifluoro-2-metilpropan-2-il)piridin-4-il]-1,3-tiazol-2-il}pirrolidina-1,2-dicarboxamida

$C_{19}H_{22}F_3N_5O_2S$



The statements indicating action and use are largely based on the information supplied by the manufacturer. This information is merely meant to provide an indication of potential uses of new substances at the time they are accorded a proposed INN. Because of their provisional nature, these descriptors will neither be revised nor included in the recommended and cumulative lists of INN.

Lists of both proposed and recommended INNs are sent together with a circular letter to WHO member states (at present 194), to national pharmacopoeia commissions, and to other bodies designated by member states. In

this letter, the WHO requests that member states take such steps as are necessary to prevent the acquisition of proprietary rights in the published names, including prohibiting registration of these names as trade names.

Up to now, nearly 9,000 INNs have been selected. All names ever selected are published in a cumulative list of INNs, which is updated periodically [11]. The generic names are presented in alphabetical order by Latin name. Each entry includes:

- INNs in Latin, Arabic, Chinese, English, French, Russian, and Spanish;
- a reference to the INN list in which the name was originally published as a proposed or recommended INN;
- a reference to national nonproprietary names;
- a reference to pharmacopoeial monographs or similar official references;
- the molecular formula;
- its CAS number (at the time of publication);
- its chemical structure, amino acid, or DNA sequence (when available, electronic documents can be downloaded from Mednet [12]).

The cumulative list, published now only as a CD-ROM, contains the list of all published INN in the six UN languages (Arabic, Chinese, English, French, Russian, and Spanish) plus Latin, both in a PDF format and a searchable database that can be query by INN name (all or part of it), list number, CAS, alternate names, Anatomical Therapeutic Chemical (ATC) Classification, and other search parameters.

5. *Biologicals*

Biologicals and biotechnological medicine is one of the fastest growing sectors of the pharmaceutical market, and it has begun to play an increasingly important role in health care. With the scientific and technical developments currently taking place, many new biological products are being introduced for the prevention, diagnosis, or treatment of human disease. How to name these substances is therefore raising a novel challenge for the nomenclature committee [13]. New schemes and concepts need to be developed on a worldwide basis, and the established schemes need to be revised or updated when necessary. One example is the revised scheme for naming monoclonal antibodies [14] as given below:

INN for monoclonal antibodies (mAbs) are composed of a prefix, a substem A, a substem B, and a suffix. The common stem for mAbs is *-mab*, placed as a suffix. The stem *-mab* is to be used for all products containing an immunoglobulin variable domain that binds to a defined target. Substem B indicates the species on which the immunoglobulin sequence of the mAb is based.

Substem B for the species

a	rat
axo (pre-substem)	rat-mouse
e	hamster
i	primate
o	mouse
u	human
xi	chimeric
-xizu-	chimeric-humanized
zu	humanized

The distinction between chimeric and humanized antibodies is as follows:

Chimeric: A chimeric antibody is one in which both chain types are chimeric as a result of antibody engineering. A chimeric chain is a chain that contains a foreign variable domain (V-D-J-REGION) (originating from one species other than human or synthetic) linked to a constant region (C-REGION) of human origin.

Humanized: A humanized antibody is one where both chain types are humanized as a result of antibody engineering. A humanized chain is a chain in which the complementarity determining regions (CDR) of the variable domains are foreign (originating from one species other than human or synthetic), whereas the remaining chain is of human origin. By extension, an antibody is described as humanized if more recent protocols were used for the humanization.

The *-xizu-* infix is used for an antibody having both chimeric and humanized chains.
The *-axo-* infix is used for an antibody having both rat and mouse chains.
Substem A indicates the target (molecule, cell, organ) class.
Substem A for target class

-b(a)-	bacterial
-c(i)-	cardiovascular
-f(u)-	fungal
-k(i)-	interleukin
-l(i)-	immunomodulating
-n(e)-	neural
-s(o)-	bone
-tox(a)-	toxin
-t(u)-	tumor
-v(i)-	viral

In principle, a single letter (e.g., *-b-* for bacterial) is used as substem A. Whenever substem B starts with a consonant (e.g. *x* or *z*), an additional vowel indicated in the table (e.g., *-ba-*) is inserted to avoid problems in pronunciation

The prefix should be random. That is, the only requirement is to contribute to a euphonious and distinctive name. If the monoclonal antibody is conjugated to another protein or to a chemical (e.g., chelator), identification of this conjugate is accomplished by use of a separate second word or acceptable chemical designation. For instance, for mAbs conjugated to a toxin, the suffix *-tox* is used in the second word. If the monoclonal antibody is radiolabelled, the radioisotope is listed first in the INN (e.g., technetium (^{99m}Tc) nofetumomab merpentan (81)).

The prefix *peg-* can be used for pegylated mAbs, but this should be avoided if it leads to over-long INNs. In most cases, it is best to adopt a two-word INN for pegylated mAbs, with the first word describing the mAb and the second being *pegol* or a related designation.

Examples of INNs are abrilumab, actoxumab, amatuximab, begelomab, brentuximab vedotin, coltuximab ravtansine, lulizumab pegol, pasotuxizumab, visilizumab, and yttrium (^{90}Y) clivatuzumab tetraxetan.

In 2005, the two-word nomenclature scheme for gene therapy products was formally adopted by the members of the INN Expert Group designated to deal with the selection of nonproprietary names [15]. The 2013 updated scheme is shown as following [14]:

The first word describes the expression gene, and the second word the vector component:

1. Word one is for gene component.
   - **a.** The prefix is random and contributes to the distinctive name.
   - **b.** The infix identifies the gene using existing infixes for biological products (when available) or using similar infix as for the protein for which the gene codes.

-cima-	cytosine deaminase
-ermin-	growth factor
-kin-	interleukin
-lim-	immunomodulator
-lip-	human lipoprotein lipase
-mul-	multiple gene
-stim-	colony stimulating factor
-tima-	thymidine kinase
-tusu-	tumor suppression

   - **c.** The suffix is -(a vowel)gene (e.g., -(*o*)*gene*).

2. Word two is for the vector component.
   a. The prefix is random and contributes to the distinctive name.
   b. For viral vectors, the infix will be:

-adeno-	adenovirus
-cana-	canarypox virus
-foli-	fowlpox virus
-herpa-	herpes virus
-lenti-	lentivirus
-morbilli-	paramyxoviridae morbillivirus
-parvo-	adeno-associated virus (parvoviridae dependovirus)
-retro-	other retrovirus
-vaci-	vaccinia virus

   c. The suffix will be *-repvec* for replicating viral vector, *-vec* for nonreplicating viral vector, or *-plasmid* for plasmid vector.
3. In the case of nonplasmid naked DNA products, a second word in the name is not neded.

Examples of INNs are alferminogene tadenovec, beperminogene perplasmid, contusugene ladenovec, sitimagene ceradenovec,velimogene aliplasmid, and tipapkinogene sovacivec.

Since 2002, several meetings dealing with aspects of INNs for biologicals have been convened at WHO headquarters. Currently, the INN Expert Group is developing a Nomenclature Scheme for Celle Therapy Products [14,16,17].

The meeting held in April 2007 was dedicated to an in-depth review of the current INN policy for naming and defining biologicals to ensure consistency of approach and to identify and address anomalies. Specific areas for review included INN policies on posttranslational modifications of proteins, monoclonal antibodies, vaccines, gene, cell and tissue therapy products, blood products, biologicals derived through transgenic processes, and enzymes. Recommendations made in this meeting were finally adopted by the INN Expert Group [18]. One of these recommendations was that more information should be requested at the time of the application for an INN. This has been addressed by adding a special section to the INN application form as shown below:

1. For a protein substance, the complete mature amino acid sequence in a format that can be copied for analysis (Word or in the text of an e-mail), using the one-letter code with spaces between groups of ten characters, five groups per line and with a number indicating the position of the last amino acid at the end of each line; the positions of the disulfide bridges and all post-translational modifications listed after the sequence. For a glycoprotein/glycopeptides, the glycosylation pattern (site, the type of sugar, etc.). For a recombinant DNA protein, expression system; comparison with the native sequence. For a monoclonal antibody, the precursor nucleotide sequence with spaces between codons and translation and with numbers per line; CDR-IMGT; IG class and subclass; IG format; species or Taxonomy Related structure (chimeric, humanized, synthetic construct) (for each chain, if different); name/structure of the antigen against which the monoclonal antibody is directed; the complete mature amino acid sequence in a format that can be copied for analysis (Word or in the text of an e-mail), using the one-letter code with spaces between groups of ten characters, five groups per line and with a number indicating the position of the last amino acid at the end of each line; sites of disulfide bridges (intra-chain, inter-chains); the glycosylation pattern (site, the type of sugar, etc.); expression system; clone name(s); laboratory code name(s).
2. For a substance comprising nucleic acid (e.g., a DNA vaccine, oligonucleotide, gene therapy product), the full nucleotide sequence of the substance. Pertinent regions (e.g., coding regions, control regions) should be delineated. For a gene therapy product, a schematic map of the entire product should be provided in addition to an entire annotated sequence that delineates relevant parts of the sequence.
3. The details of pegylation: the end group; the polymer chain with the average number of repeat units (to 2 significant figures); details of the linker (not the reagent used); where the linker is attached to the active moiety, ideally, if multiple sites are involved in what proportion are modified.

As substances produced by biotechnology became more and more complex and challenging, the INN Expert Group had requested the WHO INN Secretariat to prepare a working document [14] intended to summarize and review the

past and present INN situation in this field. This document presents an inventory of the policy decisions taken by the INN Expert Group during all these years of change, and of the names assigned to biological and biotechnological substances. Considering the potential for further developments in the field of biologicals, this review is intended to be a living document which will be regularly updated to include new policies, and future INNs assigned.

6. INN Digital Data Access and Services

All INN and related information is now available in digital format on the Internet (as shown in Figure 34.5) and on searchable databases. Indeed the INN users can access INN MedNet [12] and search by name in English, French,

INN website: http://www.who.int/medicines/services/inn/en/
GUIDANCE
Guidelines on the use of INNs for pharmaceutical substances
INFORMATION
INN consultations and meetings
Dates, deadlines, and executive summaries
REQUESTING AN INN
Selection process of INNs
The procedure for the selection of recommended
INNs for pharmaceutical substances in the six official languages
INN online application: https://extranet.who.int/tools/inn_online_application/
To select an appropriate stem
INN stems
To select an appropriate stem
All about existing stems and pre-stems
WHO46.19: Nonproprietary Names for
Pharmaceutical Substances
PUBLISHED INN LISTS
Proposed and recommended lists
ALL INN PUBLICATIONS
Publishing INN lists
To order the cumulative list: http://apps.who.int/bookorders/
Lists of recommended and proposed INNs
WHO drug information series
Other publications and documents
stems
biologicals
modified INN
various
BIOSIMILARS
ONLINE ACCESS
To access the INN database on Mednet: http://mednet.who.int/
PROMOTION OF INN AND COLLABORATIONS
INN campaign: “La DCI, le vrai nom du médicament”
ISDB (International Society of Drug Bulletins)
"INN: an essential tool"
A Study on the use of INN in India

FIGURE 34.5 Web links to relevant information and documents on INN.

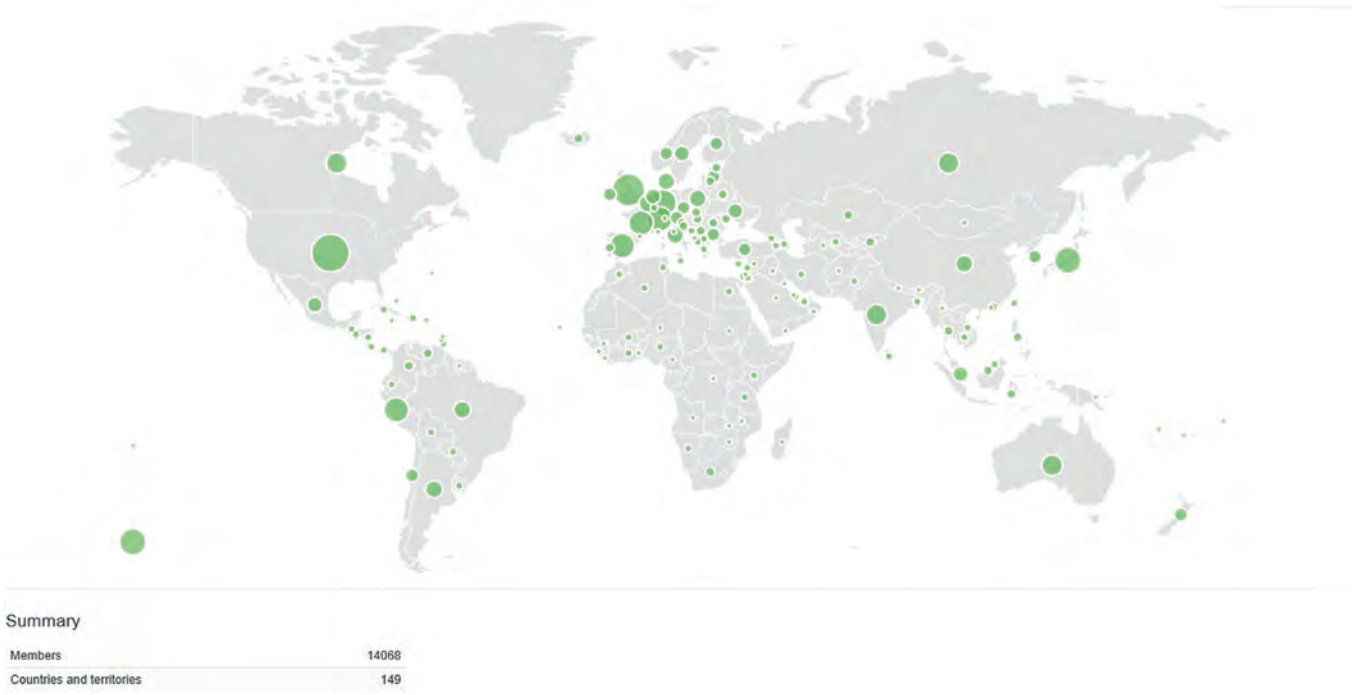


FIGURE 34.6 Mednet users around the world (updated 2014).

Chinese, Spanish, Arabic, Russian, and Latin or by list (proposed and recommended). The MedNet INN users' community is the largest WHO community, reaching more than 14,000 INN stakeholders all over the world (see Figure 34.6).

Moreover, the Programme on INN provides the INN Global Data Hub for institutions requiring it. This is a web service enabling an interoperable machine-to-machine interaction over the network. It has an interface described in a machine-readable format. Other systems interact with the web service in a manner prescribed by its description using the INN Hub API messages conveyed using HTTP protocol. The INN Global Data Hub allows transparent integration of the INN database on web site and/or applications. For off-line use, the CD-ROM of the cumulative list is published every two years.

B. Common Names Selected by the International Standards Organization and Nomenclature Bodies

1. *International Standards Organization*

The International Organization for Standardization (ISO) has laid down principles for selecting common names for pesticides and other agrochemicals [19]. These principles are comparable to the guiding principles for selecting INNs and have a similar purpose: to provide short, distinctive, and easily pronounced names for substances whose full chemical names are too complex for convenient use. The names chosen should not be permitted to become privately owned trademarks. ISO names are also given for salts and complex esters, as well as mixtures of isomers. The work of the INN and ISO committees sometimes overlaps, especially in the field of veterinary medicine. The two committees collaborate to avoid using different names for the same compound.

2. *National Nomenclature Bodies*

Since the INN Programme came into existence, WHO has coordinated the activities of national nomenclature commissions. Several INN experts are secretaries to national nomenclature commissions and, in most cases, the WHO Secretariat also acts as a corresponding member of these commissions. Differences between national and international nomenclature have become rare.

In most countries, national nomenclature commissions are part of or closely linked to the national pharmacopoeia. Some countries—such Germany and the Nordic countries—no longer have a commission and publish the INNs directly as national names in their legal publications. The latter countries previously published Nordic Pharmacopoeia Names (NFN). In others, the national nomenclature commission adopts INNs in the language of

the country as national names. All European directives include INNs as "usual terminology" (point 3 of Article 4 (2) of Directive 65/65/EEC). The European Pharmacopoeia uses INNs in the main titles of monographs.

National nomenclature commissions select and publish the following national nonproprietary names:

- British Approved Names (BAN): The Secretary, British Pharmacopoeia Commission, Market Towers, 1 Nine Elms Lane, London SW8 5NQ, UK.
- Dénominations Communes Françaises (DCF): Secretariat of the French Pharmacopoeia Commission at the Drug Agency, Direction des Laboratoires et des Contrôles, Unité Pharmacopée, 145–147, Boulevard Anatole France, 93200 Saint-Denis, France.
- Japanese Accepted Names (JAN): Japanese Ministry of Health and Welfare, New Drugs Division, Pharmaceuticals Affairs Bureau, 1-2-2, Kasumigaseki, Chiyoda-ku, Tokyo 100, Japan.
- United States Adopted Names (USAN): United States Adopted Names Council, American Medical Association, P.O. Box 10970, Chicago, Illinois 60610, USA.

IV. USE AND PROTECTION OF NONPROPRIETARY NAMES

The WHO INN Programme has been actively providing nonproprietary names since 1953. During this period, nearly 9,000 names have been published. New pharmaceutical substances are continually being developed, and some 100–150 new names are published every year. In order to avoid confusion, the WHO strongly recommends that new drug substances are identified by codes rather than arbitrary names until international nonproprietary names have been designated.

A. Use of Nonproprietary Names

International nonproprietary names are intended to be used in pharmacopoeias, labeling, advertising, drug regulation, and scientific literature, and as product names (e.g., for generics). Some countries have defined the minimum size of characters in which the generic name must be printed under the trademark labeling and advertising. In Canada, the USA, and Uruguay, the generic name must appear prominently in type at least half as large as that used for the proprietary or brand name. Certain countries (e.g., Mexico) have even gone as far as abolishing trademarks for the public sector.

B. Protection of Nonproprietary Names

Introducing INN common stems into trademarks, which seems to be increasingly popular, hampers the selection of new nonproprietary names within the established system. Given that all new INNs should be distinctive from existing INNs without similarity to trademarks, this practice can cause confusion to health professionals, may be the source of serious errors in prescribing and dispensing, and hinders the selection of future names for compounds in the same group of substances. Based on recommendations made by the WHO Expert Committee on the Use of Essential Drugs, a resolution [20,21] was adopted during the Forty-sixth World Health Assembly requesting member states to:

> enact rules or regulations, as necessary, to ensure that international nonproprietary names … are always displayed prominently; to encourage manufacturers to rely on their corporate name and the international nonproprietary names, rather than on trademarks, to promote and market multisource products introduced after patent expiration; to develop policy guidelines on the use and protection of international nonproprietary names, and to discourage the use of names derived from INNs, and particularly names including INN stems in trademarks.

V. SUMMARY

The existence of an international nomenclature for pharmaceutical substances, in the form of INNs, has proved since 1953 to be important for the safe prescription and dispensing of medicines to patients, and for communication and the exchange of information among health professionals worldwide. INNs identify pharmaceutical substances by unique names that are globally recognized and are public property. They are also known as generic names.

Common stems are developed for the selection of INNs to communicate to health professionals the type of pharmaceutical product in question. National and international nomenclature commissions collaborate closely to select a single name of worldwide acceptability for each active substance that is to be marketed as a pharmaceutical.

To avoid confusion, which could jeopardize the safety of patients, nonproprietary names and their common stems should not be used in trademarks. The selection of further names within a series should not be hindered by the use of a common stem in a brand name.

References

[1] Wehrli A. Pharmaceuticals: trademarks versus generic names. Trademark World 1986;16:31–5.
[2] WHO. Trademarks versus generic names for pharmaceuticals: A conflict that requires resolution. WHO Drug Information 1987;1:39–40.
[3] WHO. Basic Documents. 39th edition Geneva: WHO; 1992.
[4] WHO. The use of stems in the selection of INN for pharmaceutical substances. Geneva: WHO; 2013. Available from: <http://www.who.int/medicines/services/inn/StemBook_2013_Final.pdf>.
[5] Programme on International Nonproprietary Names (INN). Addendum to "The use of stems in the selection of International Nonproprietary names (INN) for pharmaceutical substances,". Available from: <http://www.who.int/entity/medicines/services/inn/Addendum _StemBook2013_201411.pdf>; 2014.
[6] Programme on International Nonproprietary Names (INN). Pre-stems: Suffixes used in the selection of INNs. Available from: < http://www.who.int/entity/medicines/services/inn/en>; 2014.
[7] Programme on International Nonproprietary Names (INN). Guidance on the establishment of new INN stems. Available from: <http://www.who.int/medicines/services/inn /stems_policy_explanation.pdf>; 2007.
[8] Semmler J, Gebert U, Eisenhut T, Moeller J, Schönharting M, Alléra A. Xanthine derivatives: Comparison between suppression of tumor necrosis factor-α production and inhibition of cAMP phosphodiesterase activity. Immunology 1993;78:520–5.
[9] List 110 of proposed INNs: WHO Drug Information 2013, 27(4). World Health Organization, Geneva.
[10] List 72 of recommended INNs: WHO Drug Information 2014, 28(3). World Health Organization, Geneva.
[11] WHO. Cumulative list no. 15 of international nonproprietary names (INN) for pharmaceutical substances. Geneva: WHO; 2013.
[12] WHO. Mednet INN services. Available from: <https://mednet-communities.net/inn/>.
[13] Wehrli A. Generic names for biotechnology-derived products. Drug News Perspect 1992;5:55–8.
[14] International nomenclature and gene therapy products. WHO Drug Information 2005;19(2):103.
[15] Programme on International Nonproprietary Names (INN). INN for biological and biotechnological substances (a review), INN working document 05.179. Available from: <http://www.who.int/entity/medicines/services/inn/BioRev2014.pdf>; 2014.
[16] Programme on International Nonproprietary Names (INN). WHO informal consultation on international nonproprietary names (INN) policy for biosimilar products. Available from: <http://www.who.int/medicines/services/inn/BiosimilarsINN_Report.pdf>; 2006.
[17] Programme on International Nonproprietary Names (INN). International nonproprietary names (INN) working group meeting on nomenclature for monoclonal antibodies (mAb). Available from: <http://www.who.int/medicines/services/inn/ApprovedFinalWHOINNWGroupMeetingNMAreport.pdf>; 2008.
[18] Forty-fourth INN consultation: Executive summary on international nonproprietary names (INNs) for pharmaceutical substances, INN working document 07.214. Available from: <http://www.who.int/medicines/publications/druginformation/innlists/en/>.
[19] ISO. Standard pesticides and other agrochemicals: Principles for the selection of common names (ISO/DIS 257, 1988). Geneva: International Organization for Standardization; 1988.
[20] Forty-sixth World Health Assembly resolution on nonproprietary names for pharmaceutical substances (WHA46.19). Geneva: WHO; 1993.
[21] Programme on International Nonproprietary Names (INN). WHO international nonproprietary names for pharmaceutical substances: Information leaflet for trademark departments. Available from: <http://www.who.int/medicines/services/inn/flyerINN.pdf>; 2004.

ANNEX

Procedure for the Selection of Recommended INNs for Pharmaceutical Substances

The following procedure shall be followed by the World Health Organization (hereinafter also referred to as "WHO") in the selection of recommended INNs for pharmaceutical substances, in accordance with resolution WHA3.11 of the World Health Assembly, and in the substitution of such names.

The authors alone are responsible for the views expressed in this article.

1. Proposals for recommended INNs and proposals for substitution of such names shall be submitted to WHO on the form provided therefor. The consideration of such proposals shall be subject to the payment of an administrative fee designed only to cover the corresponding costs of the Secretariat of WHO ("the Secretariat"). The amount of this fee shall be determined by the Secretariat and may, from time to time, be adjusted.

2. Such proposals shall be submitted by the Secretariat to the members of the Expert Advisory Panel on the International Pharmacopoeia and Pharmaceutical Preparations designated for this purpose, such designated members hereinafter referred to as "the INN Expert Group," for consideration in accordance with the "general principles for guidance in devising International Nonproprietary Names for Pharmaceutical Substances," annexed to this procedure. The name used by the person discovering or first developing and marketing a pharmaceutical substance shall be accepted, unless there are compelling reasons to the contrary.
3. Subsequent to the examination provided in article 2, the Secretariat shall give notice that a proposed international nonproprietary name is being considered.
   a. Such notice shall be given by publication in *WHO Drug Information* and by letter to Member States and to national and regional pharmacopoeia commissions or other bodies designated by Member States.
      i. Notice shall also be sent to the person who submitted the proposal ("the original applicant") and other persons known to be concerned with a name under consideration.
   b. Such notice shall:
      i. set forth the name under consideration;
      ii. identify the person who submitted the proposal for naming the substance, if so requested by such person;
      iii. identify the substance for which a name is being considered;
      iv. set forth the time within which comments and objections will be received and the person and place to which they should be directed;
      v. state the authority under which WHO is acting and refer to these rules of procedure.
   c. In forwarding the notice, the Secretariat shall request that Member States take such steps as are necessary to prevent the acquisition of proprietary rights in the proposed name during the period it is under consideration by WHO.
4. Comments on the proposed name may be forwarded by any person to WHO within 4 months of the date of publication, under article 3, of the name in *WHO Drug Information*.
5. A formal objection to a proposed name may be filed by any interested person within 4 months of the date of publication, under article 3, of the name in *WHO Drug Information*.
   a. Such objection shall:
      i. identify the person objecting;
      ii. state his or her interest in the name;
      iii. set forth the reasons for his or her objection to the name proposed.
6. Where there is a formal objection under article 5, WHO may either reconsider the proposed name or use its good offices to attempt to obtain withdrawal of the objection. Without prejudice to the consideration by WHO of a substitute name or names, a name shall not be selected by WHO as a recommended international nonproprietary name while there exists a formal objection thereto filed under article 5 that has not been withdrawn.
7. Where no objection has been filed under article 5, or all objections previously filed have been withdrawn, the Secretariat shall give notice in accordance with subsection (a) of article 3 that the name has been selected by WHO as a recommended international nonproprietary name.
8. In forwarding a recommended international nonproprietary name to Member States under article 7, the Secretariat shall:
   a. request that it be recognized as the nonproprietary name for the substance;
   b. request that Member States take such steps as are necessary to prevent the acquisition of proprietary rights in the name and to prohibit registration of the name as a trademark or trade name.
9. A. In the extraordinary circumstance that a previously recommended international nonproprietary name gives rise to errors in medication, prescription, or distribution or a demonstrable risk thereof because of similarity with another name in pharmaceutical and/or prescription practices, and it appears that such errors or potential errors cannot readily be resolved through other interventions than a possible substitution of a previously recommended international nonproprietary name, or in the event that a previously recommended international nonproprietary name differs substantially from the nonproprietary name approved in a significant number of Member States, or in other such extraordinary circumstances that justify a substitution of a recommended international nonproprietary name, proposals to that effect may be filed by any interested person. Such proposals shall be submitted on the form provided therefor and shall:
   i. identify the person making the proposal;
   ii. state his or her interest in the proposed substitution;

iii. set forth the reasons for the proposal;
iv. describe and provide documentary evidence regarding the other interventions undertaken in an effort to resolve the situation and the reasons why these other interventions were inadequate.

Such proposals may include a proposal for a new substitute international nonproprietary name, devised in accordance with the general principles, that takes into account the pharmaceutical substance for which the new substitute international nonproprietary name is being proposed.

The Secretariat shall forward a copy of the proposal for consideration in accordance with the procedure described in subsection (b) below to the INN Expert Group and the original applicant or its successor (if different from the person bringing the proposal for substitution and provided that the original applicant or its successor is known or can be found through diligent effort, including contacts with industry associations).

In addition, the Secretariat shall request comments on the proposal from:

i. Member States and national and regional pharmacopoeia commissions or other bodies designated by Member States (by including a notice to that effect in the letter referred to in article 3(a);
ii. any other persons known to be concerned by the proposed substitution.

The request for comments shall:

i. state the recommended international nonproprietary name that is being proposed for substitution (and the proposed substitute name, if provided);
ii. identify the person who submitted the proposal for substitution (if so requested by such person);
iii. identify the substance to which the proposed substitution relates and reasons put forward for substitution;
iv. set forth the time within which comments will be received and the person and place to whom they should be directed;
v. state the authority under which WHO is acting and refer to these rules of procedure.

Comments on the proposed substitution may be forwarded by any person to WHO within 4 months of the date of the request for comments.

B. After the time period for comments referred to above has elapsed, the Secretariat shall forward any comments received to the INN Expert Group, the original applicant or its successor, and the person bringing the proposal for substitution. If, after consideration of the proposal for substitution and the comments received, the INN Expert Group, the person bringing the proposal for substitution and the original applicant or its successor all agree that there is a need to substitute the previously recommended international nonproprietary name, the Secretariat shall submit the proposal for substitution to the INN Expert Group for further processing.

Notwithstanding the foregoing, the original applicant or its successor shall not be entitled to withhold agreement to a proposal for substitution in the event the original applicant or its successor has no demonstrable continuing interest in the recommended international nonproprietary name proposed for substitution.

In the event that a proposal for substitution shall be submitted to the INN Expert Group for further processing, the INN Expert Group will select a new international nonproprietary name in accordance with the general principles referred to in article 2 and the procedure set forth in articles 3–8 inclusive. The notices to be given by the Secretariat under article 3 and article 7, respectively, including to the original applicant or its successor (if not the same as the person proposing the substitution, and provided that the original applicant or its successor is known or can be found through diligent effort, including contacts with industry associations), shall in such event indicate that the new name is a substitute for a previously recommended international nonproprietary name and that Member States may wish to make transitional arrangements in order to accommodate existing products that use the previously recommended international nonproprietary name on their label in accordance with national legislation.

If, after consideration of the proposal for substitution and the comments received in accordance with the procedure described above, the INN Expert Group, the original applicant or its successor, and the person bringing the proposal for substitution do not agree that there are compelling reasons for substitution of a previously recommended international nonproprietary name, this name shall be retained (provided always that the original applicant or its successor shall not be entitled to withhold agreement to a proposal for substitution in the event that the original applicant or its successor has no demonstrable continuing interest in the recommended international nonproprietary name proposed to be substituted). In such an event, the Secretariat shall advise the

person having proposed the substitution, as well as the original applicant or its successor (if not the same as the person proposing the substitution, and provided that the original applicant or its successor is known or can be found through diligent effort, including contacts with industry associations), Member States, national, and regional pharmacopoeia commissions, other bodies designated by Member States, and any other persons known to be concerned by the proposed substitution that, despite a proposal for substitution, it has been decided to retain the previously recommended international nonproprietary name (with a description of the reason(s) why the proposal for substitution was not considered sufficiently compelling).

CHAPTER

35

Web Alert: Using the Internet for Medicinal Chemistry

David Cavalla

Numedicus Ltd, Cambridge, England

OUTLINE

Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. That's what we're doing. ***Jimmy Wales, founder of Wikipedia***

I. INTRODUCTION

The Internet has undergone substantial change since the 3rd edition of *Practice of Medicinal Chemistry*, both in the continued growth of the Internet and the availability of additional resources. This article can only describe the situation as it currently stands and give some predictions as to the future. It is in the nature of such reviews that they can never be complete and up to date, and they deteriorate rapidly. Furthermore, the resources on the

Internet are now too vast for a comprehensive article in the space available. A view has had to be taken to include only what can be considered to be the most significant and likely to last.

This review is intended to provide freely available resources for the tools that medicinal chemists generally use in the work they do, which necessarily involves a variety of tasks from drug design to chemical synthesis. Sites for prediction of physical activity are just as relevant as those for prediction of biological activity and patents.

While there has been a substantial increase in the number of commercial sites for chemistry since the last edition, the predicted shift in balance away from freely available resources has not occurred. If anything, a surge in open resources has taken place, such as in the open-access publishing and open-innovation movements. Further trend analysis is complicated by the observation of two fundamentally opposing developments. On the one hand there has been a huge growth in chaotic sources of information like Twitter, which were essentially non-existent four years ago. On the other hand, certain highly organized and well-linked sources such as PubChem have also grown.

II. BLOGS

The blog has rapidly risen to a position of importance in news and other forms of media. However, it has been introduced relatively slowly into science in general. There are no quality controls in the blog medium. However, there has been an aggregation around certain well-known sites, of which the following specialize on medicinal and pharmaceutical chemistry:

Title	URL	Comments
Medicinal Chemistry: In the Pipeline	http://pipeline.corante.com/	The author, Derek Lowe, has worked for several major pharmaceutical companies on drug-discovery projects against schizophrenia, Alzheimer's, diabetes, osteoporosis, and other diseases.
Sussex Drug Discovery	http://sussexdrugdiscovery.wordpress.com/	Medicinal, chemistry and biochemistry blog from the Translational Drug Discovery Group
MedChemBlog	http://www.medchemblog.blogspot.co.uk/	Blog directed to drug discovery-related information, especially medicinal chemistry and the pharmacological aspects of drugs.
In Vivo blog	http://invivoblog.blogspot.co.uk/	Commentary on recent developments in biopharmaceutical business development, R&D, financing, marketing, and policy.
MedChem Buzz	http://medchembuzz.wordpress.com/	Highlights a subset of the recent medicinal chemistry literature.
The Sceptical Chymist	http://blogs.nature.com/thescepticalchymist/	Blog from Nature Chemistry.
One in Ten Thousand	http://walkerma.wordpress.com/	General discussion on medicines and pharmaceutical chemistry.
Kinasepro	http://kinasepro.wordpress.com/	Interesting site on kinase medicinal chemistry, although a little old.

Further examples of blogs related specifically to synthetic chemistry are detailed below in the section on chemical synthesis.

Certain blogs are becoming established resources for highlighting articles of particular interest and importance, an increasingly necessary activity given the enormous amount of scientific literature that is published currently.

III. WIKIS

From software to encyclopedias, collaborative projects are one of the most evidently disruptive applications of the Internet, posing multiple challenges to commercial publishing organizations. Wikipedia (http://www.wikipedia.org/) is well established, authoritative, and increasingly comprehensive, even though its first page went online only in 2001. It is a generalist information source, but its scope and depth exceed many specialist alternatives. The word is a composite of wiki (the Hawai'ian word for quick) and encyclopedia (it is now the world's biggest). A interesting article on the origins, implementation, and phenomenal growth of Wikipedia was published in the *The Atlantic* in 2006 (http://www.theatlantic.com/doc/200609/wikipedia/).

It is the collaborative and indeed co-operative nature of wikis that has enabled the rapid growth of Wikipedia. By comparison with the first *Oxford English Dictionary*, which took seventy-eight years for the first product to be published in 1928 (http://www.oed.com/), Wikipedia is eleven years old and has more than 4.1 million articles in its English language edition, a number growing by nearly 2,000 a day. In 2005, a report in *Nature* (http://www.nature.com/) compared Encyclopedia Britannica and Wikipedia science articles and suggested that the former are usually only marginally more accurate than the latter. Between 2008 and 2010, articles in medical and scientific fields such as pathology, toxicology, oncology, and pharmaceuticals comparing Wikipedia to professional and peer-reviewed sources found that Wikipedia's depth and coverage were of a high standard (http://en.wikipedia.org/wiki/Reliability_of_Wikipedia/).

As an example, the entry for angiotensin II antagonists (http://en.wikipedia.org/wiki/Angiotensin_II_receptor_antagonist/) nicely leads to a list of seven members of the group. While incomplete as a list, each entry contains a graphical structure, a IUPAC name, a CAS number, a PubChem link, bioavailability, protein binding, metabolism, half-life, and other data. In the main section, the articles include information on regulatory status, dosing frequency, therapeutic indications, and side effects.

There are other wikis related to medicinal chemistry, but these specialist sources are becoming increasingly sparsely populated and superficial compared to the information provided by this generalist site. The Chemical Information Sources Wiki (http://en.wikibooks.org/wiki/Chemical_Information_Sources/) is a guide to the many sources of reference materials available for those with questions related to chemistry. The site includes information on primary, secondary, and tertiary publication sources, chemical information databases, physical property information, chemical patent searching, and molecular visualization tools and sites. The material is based on an undergraduate course offered for many years in the Indiana University Department of Chemistry by Gary Wiggins.

A. RSS Information and Twitter Feeds

Nowadays, web information is often delivered by means other than the simple (static) web browser. RSS (rich site summary) is a format for notifying users of new content at a website and a way of getting news in general. It has now been introduced for many journals. The user subscribes to a feed by entering into the reader the feed's URI (e.g., the address for the ACS Chem Eng news is http://feeds.feedburner.com/cen_latestnews/) or by clicking a feed icon in a web browser to initiate the subscription process.

Twitter (http://twitter.com/) is an increasingly large source of disorganized information. Some specific contributors are useful. A specific hashtag associated with content related to medicinal chemistry is #medchem.

IV. COMPOUND INFORMATION

There is a growing number of freely accessible compound databases challenging the dominance of Chemical Abstracts (http://www.cas.org), a commercial product. However the very diversity of these databases poses its own difficulty, and of course none is as comprehensive as CAS.

A. PubChem

Large, well-curated molecule databases such as PubChem (http://pubchem.ncbi.nlm.nih.gov/) have dramatically changed the landscape of publicly available cheminformatics resources for medicinal chemistry and pharmaceutical research. PubChem was launched by the NIH in 2004 to support the New Pathways to Discovery component of the Roadmap for Medical Research. PubChem archives and organizes information about the biological activities of chemical compounds into a comprehensive database and is the informatics backbone for the Molecular Libraries and Imaging Initiative, which is part of the NIH Roadmap. PubChem is also intended to empower the scientific community to use small molecule chemical compounds in their research as molecular probes to investigate important biological processes or gene functions. Nature Chemical Biology (http://www.nature.com/nchembio/) and NMRShiftDB (http://www.nmrshiftdb.org/; see below) data are available through PubChem, which also provides links to Medical Subject Annotations and PubMed biomedical literature citations. The database now has 100,000,000 substances and 200,000,000 bioactivities in its collection. PubChem provides a limited set of structure properties selected to be relevant for typical drug-design applications. Presently, it is possible to do chemical similarity searches based on SMILES, text name, molecular formula, and substructure.

PubChem also develops new bioassays and performs massive high-throughput screening experiments on a large number of compounds, resulting in a very large public store of biological activity data associated with chemical structures. The structure database contains full catalogs of major suppliers of screening compounds, as well as the structures from other public databases (NCI, NIAID, NIST), and provides extensive links to original data. Examples for productive queries in the PubChem system can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = pcsubstance. PubChem also claims it is already the largest freely accessible chemical structure store.

B. ChemSpider

ChemSpider (http://www.chemspider.com/) is a new chemistry search engine built with the intention of aggregating and indexing chemical structures and their associated information into a single searchable repository and making it available to everybody at no charge. Some properties have been added to each of the chemical structures within the database, such as structure identifiers like SMILES, InChI, IUPAC, and Index Names, as well as many physicochemical properties. In addition, ChemSpider provides access to a series of property prediction algorithms.

ChemSpider currently searches over 28 million compounds in multiple chemical structure databases. These include databases of curated literature data, chemical vendor catalogs, molecular properties, environmental data, toxicity data, and analytical data. ChemSpider intends to aggregate into a single database all chemical structures available within open access and commercial databases and to provide the necessary pointers from the ChemSpider search engine to the information of interest. This service will allow users to either access the data immediately via open-access links or have the information necessary to continue their searches into commercially available systems.

A blog supports the system for the science, politics, and vision behind ChemSpider, as well as for incremental changes in functionality (http://www.chemspider.com/blog/).

C. ChEBI Database

The Chemical Entities of Biological Interest (ChEBI) database (http://www.ebi.ac.uk/chebi/) is a highly curated database of molecular entities focused on small chemical compounds. The entities are either natural products or synthetic products used to intervene in the processes of living organisms. ChEBI includes an ontological classification (Fig. 34.1) whereby the relationships between molecular entities or classes of entities and their parents and/or children are specified. The database presently offers access to more than 440,000 compounds and is available for download by anonymous FTP (ftp://ftp.ebi.ac.uk/pub/databases/chebi/).

D. ChemBank

ChemBank (http://chembank.broad.harvard.edu/) is an initiative of Broad Institute Chemical Biology Program (BCB) and sponsored by the National Cancer Institute's Initiative for Chemical Genetics (ICG; http://www.broadinstitute.org/science/programs/chemical-biology/initiative-chemical-genetics). ChemBank was developed by the informatics group at the Harvard Institute of Chemistry and Cell Biology and utilizes software toolkits supplied by Daylight Chemical Information Systems (http://www.daylight.com/).

ChemBank is a freely available collection of data about small molecules and resources for studying their properties, especially their effects on biology. ChemBank contains structures for over 500,000 unique molecular entities. Each structure is stored with associated chemical information, including name(s), source information, molecular descriptors, assay results, and activity-related terms from the scientific literature. The database can be searched by chemical name or activity, by substructure (SMILES string input), or for structure similarity (SMILES string input). Searches can be limited to subsets of available compounds, defined as natural products, known drugs, FDA-approved drugs, commercially available compounds, orally available compounds, and primary metabolites, among other categories. ChemBank stores cell measurements derived from cell lines treated with small molecules, among other biological objects. It is possible to pick an assay and then view both the details of the screen and/or the data from the assay (e.g., a search on "Autophagy" yielded these hits: http://chembank.broadinstitute.org/chemistry/search/execute.htm?id = 5646885). An additional option enables viewing of the chemical structure employed in the assay, while another enables export of spreadsheet files into Microsoft Excel using comma-separated values.

ChemIDplus (http://chem.sis.nlm.nih.gov/chemidplus/) is a search engine that allows retrieval of about 395,000 chemical substance files. The structure-searchable database may include structure (302,000 structures available), official name, systematic name, other names, classification code (therapeutic use), molecular formula, STN locator code, and CAS registry number. Compounds are also searchable by toxicity data and physical properties.

ChemFinder (http://chembiofinder.cambridgesoft.com/) is a large and specific chemical substances search engine that provides basic information about chemicals such as physical property data and 2D chemical structures. Chemicals and pharmaceuticals can be searched by chemical name, CAS registry number, molecular formula, or molecular weight. The database contains links to information on compounds from the National Cancer Institute (NCI), the Merck Index, and Martindale, as well as sourcing information from ChemACX. Some of this information is available only with a subscription.

V. BIOLOGICAL PROPERTIES OF COMPOUNDS

In addition to the PubChem database referred to above, a range of resources provides information on biological activities of compounds.

ChEMBL (https://www.ebi.ac.uk/chembl/) is a database of bioactive drug-like small molecules. It contains 2D structures, calculated properties (e.g., LogP, molecular weight, Lipinski parameters) and abstracted bioactivities (e.g., binding constants, pharmacology and ADMET data). The data is abstracted and curated from the primary scientific literature and cover a significant fraction of the structure-activity relations (SAR) and discovery of modern drugs. Currently, the database contains over 1.2 million distinct compounds and over 10 million activities. The database has an informative blog that announces updates and news associated with the site (http://chembl.blogspot.co.uk/).

BindingDB (http://www.bindindb.org/) is a public web-accessible database of measured binding affinities, focusing chiefly on the interactions of proteins considered to be candidate drug targets with ligands that are small, drug-like molecules. It supports medicinal chemistry and drug discovery via literature awareness and development of SAR and QSAR. BindingDB contains nearly 1 million binding data points for over 6,000 protein targets and over 375,000 small molecules (http://www.bindingdb.org/bind/index.jsp).

The International Union of Basic and Clinical Pharmacology (IUPHAR) hosts its official database on receptor nomenclature and drug classification at http://www.iuphar-db.org/. It is both authoritative and increasingly comprehensive, containing detailed pharmacological, functional, and pathophysiological information on G protein-coupled receptors, voltage-gated ion channels, ligand-gated ion channels and nuclear hormone receptors.

OpenPHACTS (http://www.openphacts.org/index.php?option=com_content&view=article&id=116&Itemid=127) is a European project funded specifically to link information sources on compounds and their biological activities. It is currently in its early stages of development but aims to integrate multiple proprietary and public data sources using open standards, driving toward the creation of an open-innovation platform for data interchange.

The NCI DIS 3D database (http://dtp.nci.nih.gov/docs/3d_database/dis3d.html) is a collection of 3D structures for over 400,000 compounds that was built and maintained by the Developmental Therapeutics Program Division of Cancer Treatment, NCI (http://www.nci.nih.gov/). While the information stored therein is only a connection table of atomic linkages, it can be interpreted by computer software to provide a 3D structure for each entry. This can then be cross-checked against available biological pharmacophores representing the preferred 3D arrangement for certain biological activities. Drugs that match the pharmacophore could have similar biological activity but have very different patterns of atomic connections. This approach has been used to search for certain novel protein kinase C agonists (http://dtp.nci.nih.gov/docs/3d_database/pharms/pkcsearch.html), using a pharmacophore derived from phorbol. A similar approach has been used to find new ligands for HIV protease, HIV integrase, and HIV reverse transcriptase (http://dtp.nci.nih.gov/docs/3d_database/pharms/ncisearches.html).

Another of the NIH databases relates to human genes and genetic phenotypes. The Online Mendelian Inheritance in Man (OMIM) website (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db = OMIM) is a catalog of human genes and genetic disorders authored and edited by Victor A. McKusick and his colleagues at Johns Hopkins and elsewhere, and developed for the web by the National Center for Biotechnology Information (NCBA; http://www.ncbi.nih.gov/).

A. Prediction of Biochemical Properties

In addition to providing a database of known affinities, BindingDB (see above) also operates in predictive mode (http://www.bindingdb.org/bind/chemsearch/marvin/BatchStructures.jsp). The predictions are based on a similarity approach using the existing information in BindingDB's database of biochemical affinities, based on the principle that similar compounds tend to bind the same proteins.

Similarly, FlexX (http://www.biosolveit.de/flex/) claims it is a fast, robust, and highly configurable (FlexX-able) computer program for predicting protein–ligand interactions. Its main application is the prediction of binding. For instance, FlexX predicts the geometry of the protein–ligand complex for a protein with known 3D structure and a small ligand molecule and estimates the binding affinity. FlexX can operate in a virtual high-throughput screening (vHTS) mode, as it is capable of screening a database consisting of ~100.000 compounds in about eight hours on a thirty-node cluster. One of features is a module called PERMUTE (http://www.biosolveit.de/Permute/), which protonates molecules and generates tautomers. An evaluation license for FlexX is valid for approximately six weeks free of charge and provides access to the full functionality of the software. After the evaluation period the software must be purchased.

AutoDock (http://autodock.scripps.edu/) is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure. AutoDock has applications in X-ray crystallography, lead optimization, structure-based design, combinatorial chemistry, protein–protein docking, and chemical mechanism studies.

GRID (http://www.moldiscovery.com/soft_grid.php) is a computational procedure for determining energetically favorable binding sites on molecules of known structure. It may be used to study individual molecules such as drugs, molecular arrays such as membranes or crystals, and macromolecules such as proteins, nucleic acids, glycoproteins, or polysaccharides. Several different molecules can be processed one after the other.

B. Drug Repurposing

One of the applications of both knowledge of known biological activities and prediction of unknown ones is in the area of drug repurposing. A couple of resources are freely accessible in this field, such as the database at DrugRepurposing.info (http://drugrepurposing.info/), which contains over 9,000 compounds that have been either launched, registered, or investigated, covering over 375 indications with over 630 mechanisms. There are more than 1,750 links between mechanism and indication, over 165 links between compound and indication, and over 60 links between compound and mechanism—so-called off-target effects.

The database is populated with over 3,850 literature references, of which over 665 derive from prospective clinical studies or case reports, over 100 from retrospective or case control analyses, and over 1,125 from pharmacological *in vivo* studies. The rest are *in vitro* studies, patents, or genomic associations.

PROMISCUOUS (http://bioinformatics.charite.de/promiscuous/) is a similar resource of protein–protein and drug–protein interactions with the aim of providing a uniform data set for drug repositioning and further analysis. It contains three different types of entities—drugs, proteins, and side effects—as well as relations between them.

C. Molecular Datasets

A list of free molecular datasets is available at http://www.cheminformatics.org/. These can be used to correlate chemical structure and biological properties, incorporating information on QSAR, QSPR, toxicity, metabolism, and permeability. There are forty-four datasets available online. The Cheminformatics site includes the information in a tabular format, with links to the chemical datasets in structure–data format and to peer-reviewed articles accessible through a Document Object Identifier (DOI) linkup.

The site provides a lot of information in a ready-to-use format. For example, blood–brain barrier penetration data is available on a training set of fifty-seven compounds and a data set of thirteen more. Long-term animal carcinogenicity results are available for over 1,400 compounds drawn from the Carcinogenic Potency Database (CPDB), an initiative of the Lawrence Berkeley Laboratory (Berkeley, CA). Pharmacological data are available on a wide range of receptors.

D. Information on Metabolic Properties

The Human Metabolome Database (HMDB; http://www.hmdb.ca/) is a comprehensive curated collection of human metabolite and human metabolism data, containing records for more than 6,800 endogenous metabolites. It includes both literature-derived data and an extensive collection of experimental metabolite concentration data compiled from various mass spectra (MS) and nuclear magnetic resonance (NMR) metabolomic analyses performed on biological fluid samples. This is further supplemented with NMR and MS spectra on purified reference metabolites. Each metabolite entry in the HMDB contains data fields, including a comprehensive compound

description, names and synonyms, structural information, physicochemical data, reference NMR and MS spectra, biofluid concentrations, and disease associations, as well as extensive links to images, references, and other public databases. Recent improvements have included spectra and substructure searching.

Another useful database is the University of Minnesota Biocatalysis/Biodegradation Database (http://umbbd.ethz.ch/). The database includes search capability for compound, enzyme, or micro-organism name; chemical formula; CAS Registry Number; and EC (enzyme classification) code. It also has lists of reaction pathways, enzymes, micro-organism entries, and organic functional groups. It specifically includes a large number of reactions of naphthalene 1,2-dioxygenase and of toluene dioxygenase. A paper describing the database was published in *Nucleic Acids Research* in January 2010. It can be downloaded in full text or in PDF from the site.

PharmGKB (http://www.pharmgkb.org/) is an integrated resource about how variation in human genes leads to variation in our responses to drugs. The database contains genetic and clinical information about people who have participated in research studies at various medical centers. Genomic data, molecular and cellular phenotype data, and clinical phenotype data are accepted from the scientific community at large. These data are organized, and the relationships between genes and drugs are then categorized into clinical outcome, pharmacodynamics and drug responses, pharmacokinetics, molecular and cellular functional assays, and genotype. The database itself has been created by Stanford University in a nationwide effort funded by the US NIH. Data downloads are available at http://www.pharmgkb.org/downloads.jsp. It is linked to its own blog site at http://pharmgkb.blogspot.co.uk/.

The site refers to the interesting set of tools available at http://www.drug-interactions.com/, which is located in the Indiana University Department of Medicine. This site includes the Cytochrome P450 Drug Interaction Table, a text based list of drugs which are known to have interactions with cytochrome p450 (http://medicine.iupui.edu/clinpharm/DDIs/table.aspx). The drugs themselves are linked to entries in RxList (http://www.rxlist.com/) and to precomposed search routines on PubMed (http://www.ncbi.nlm.nih.gov/). The data are additionally categorized into compounds that are known substrates, inhibitors, and inducers of a particular p450 subtype. An abbreviated table used for clinical purposes is found at http://medicine.iupui.edu/flockhart/clinlist.htm.

VI. DRUG INFORMATION

A. Drug Bank

The University of Alberta, with the support of Genome Alberta and Genome Canada, has introduced the freely available online resource DrugBank, which contains detailed chemical, pharmaceutical, medical, and molecular biological information on more than over 6,700 approved or experimental drugs products and over 4,200 nonredundant protein (i.e., drug target/enzyme/transporter/carrier) sequences (http://www.drugbank.ca/). Specifically, the drug entities comprise around 1,450 FDA-approved small molecule drugs, 130 FDA-approved biotech (protein/peptide) drugs, 85 nutraceuticals, and more than 5,000 experimental drugs.

DrugBank brings the latest data from the Human Genome Project together with detailed chemical information about drugs and drug products. It provides more than eighty data fields for each drug, including brand names, chemical structures, protein and DNA sequences, links to relevant Internet sites, prescription information, and detailed patient information. Users may query DrugBank through a simple text query for general queries of the entire textual component of the database. They may browse for a tabular synopsis of database content, such as for compounds grouped by their indication. They may also draw the structure of a chemical compound (using a ChemSketch applet or SMILES string) to search for chemicals similar or identical to the query compound. Finally, there is also a facility to conduct BLASTP (protein) sequence searches of the 15,000 sequences contained in DrugBank. Both single and multiple sequence (i.e., whole proteome) BLAST queries are supported. A relational query search tool allows users to select or search various combinations of subfields.

While the FDA has a very good searchable websites of approved drugs at http://www.fda.gov/cder/ob/ and FDA-approved biologics and other biopharmaceutical products at http://www.biopharma.com/, these are not structure-searchable and do not contain information on compounds in development. More complete database products, like Informa's Pharmaprojects (http://www.pharmaprojects.com/), MedTRACK (http://oneview.medtrack.com/), Thompson Reuter's Drug News (http://drugnews.thomson-pharma.com/ddn/), and Integrity (http://thomsonreuters.com/products_services/science/science_products/a-z/integrity/), are only available for a substantial price.

Multiple other sources provide information on marketed compounds, similar to that which is conventionally available in pharmacopoeias. Indeed, the names of these sites often reflects that connection. The Internet Drug Index (http://www.rxlist.com/) is a prescription-drug database, which provides good basic information about products on the market, searchable by keyword, brand, or interaction. RxList is a trove of pharmaceutical knowledge with more than 4,500 medications on file, a pharmaceutical discussion board, and an online dictionary of medical jargon. It provides useful basic information about conventional drugs and a handful of herbal remedies as well in the form of drug FAQs (frequently asked questions) and patient monographs.

Another source is the electronic Medicines Compendium (eMC; http://www.medicines.org.uk/emc/), with electronic versions of data sheets and Summaries of Product Characteristics (SPCs, sometimes also called SmPCs to differentiate them from Supplementary Patent Certificates) for medicines. It provides the same information as that contained in the latest edition of the *Compendium of Data Sheets and Summaries of Product Characteristics*, which covers thousands of medicines licensed in the UK. As an ongoing process, the eMC is also incorporating the SPCs of several thousand other medicines approved by licensing authorities. The eMC aims to provide information on every licensed prescription, pharmacy, and general sale medicine in the UK, including generics. In addition to SPCs, the eMC aims to include all Patient Information Leaflets (PILs), with dynamic updating and online links to complementary sources of medicines information.

VII. PHYSICAL CHEMICAL INFORMATION

One of the most important sources of information is the NIST Chemistry webBook (http://webbook.nist.gov/chemistry/) from the National Institute of Standards and Technology (formerly the National Bureau of Standards), which lists up to forty-five thermochemical, thermophysical, and ion energetics properties that are available for over 40,000 compounds.

The MatWeb site (http://www.matweb.com/) is different, since it deals mostly with materials instead of individual chemical substances. The free sites, while offering significant amounts of data, do not compare with the information available from Beilstein's commercially-priced Crossfire product either in terms of the number of compounds or the number of properties for each hit.

Syracuse Research Corporation (SRC; http://www.syrres.com/esc/physdemo.htm) offers commercial online searches of a number of physical property databases, including online LogP measurements (octanol-water partition coefficient) and environmental fate for over 25,000 chemicals.

Information specifically on solvents is to be found at Solv-DB (http://solvdb.ncms.org/solvdb.htm). This site, sponsored by the National Center for Manufacturing Sciences (NCMS), gives information on eight different parameters including solvent name, CAS registry number, molecular formula, and chemical category for over 200 solvents. Nine different properties are range searchable, including flash point, vapor pressure, density, and surface tension. Up to thirty-three more properties can be displayed for each solvent. Results can be sorted by solvent name or any of the nine range-searchable properties. Extensive information is provided for each solvent with display of health, safety, regulatory, and environmental fate data.

The ChemExper Chemical Directory (http://www.chemexper.com/) is also listed below as a resource for searching available chemicals from various supplier catalogs. Links are provided to the supplier's website and to MSDS. Only the basic properties are directly provided: density, m.p., b.p., and flash point. However, links to the full text of the MSDS will usually provide some additional properties.

Finally, the Organic Compounds Database (http://www.colby.edu/chemistry/cmp/cmp.html), maintained at Colby College, features a database of 2,483 compounds compiled by Harry M. Bell of Virginia Tech. Though only a few common properties are provided, the search screen allows the selection of a wide variety of parameters including property values, element counts, and the presence or absence of certain broad structural entities such as amines or hydroxyl groups. Unfortunately, retrieval sets are limited to twenty compounds, though the search engine does report the total number of hit compounds.

VIII. PREDICTION AND CALCULATION OF MOLECULAR PROPERTIES

Molecular property prediction is becoming a useful tool in the generation of libraries enriched by drug-like molecules. Used in a more focused way, drug design and lead optimization benefit from an ability to predict

physical properties such as lipophilicity and solubility, as well as physical molecular properties such as polar molecular surface area (PSA). Methods for prediction of the latter are outlined in an old publication by David Clark at http://www.documentarea.com/qsar/davclark.pdf.

The Marvin Sketch and Calculator plugins, from ChemAxon, enable chemical drawing of structures and reactions and calculate various properties like pKa, LogP, logD, charge, topology analysis, and conformers (http://www.chemaxon.com/demosite/marvin/). A similar free and publicly accessible website has been created at Chemicalize.org (http://www.chemicalize.org/), based on ChemAxon's name-to-structure conversion software to identify chemical structures on webpages and other text. Structure-based predictions are then possible from the structure so generated, including partitioning parameters, (e.g., LogP), electronic characteristics (charge, polarizability), and geometry (PSA, topology).

Alternatively, a program called tpsa.c for the calculation of PSA directly from SMILES input is described at http://www.daylight.com/meetings/emug00/Ertl/. It claims to be two to three orders of magnitude faster than other methods,

In addition to experimental information, SRC (referenced above) have also developed software to predict physical properties, such as the Estimation Program Interface (EOI) Suite (http://www.epa.gov/oppt/exposure/docs/episuite.htm) which was developed for the US Environmental Protection Agency (EPA). By entering a single SMILES notation as the search key, results from ten separate programs are displayed. These are shown in the table below. The program contains a SMILES notation database searchable by CAS registry numbers. By entering a registry number, the SMILES notation is automatically retrieved and entered into the search box.

Aquatic toxicity (LD50, LC50)	Henry's law constant
Aqueous hydrolysis rates	M.P, B.P., and vapor pressure
Atmospheric oxidation rates	Octanol-water partition coefficient
Bioconcentration factor (BCF)	Soil sorption coefficient (Koc)
Biodegradation probability	Water solubility

Table: Properties available from the Estimation Program Interface (EOI) Suite at http://www.epa.gov/oppt/exposure/docs/episuite.htm

Another useful sites in this regard is ChemExper. In addition to resources for searching available chemicals and their physical properties (see above), ChemExper hosts the OSIRIS Property Explorer (http://www.chemexper.com/tools/propertyExplorer/) for calculation and prediction of a compound's physical parameters. OSIRIS calculates various drug-relevant properties using a user-drawn structure. Prediction results are valued and color coded. Properties with high risks of undesired effects—like mutagenicity or a poor intestinal absorption—are shown in red, whereas a green color indicates likelihood of conforming to drug-like behavior. As the user is building the molecule, the cLogP and solubility characteristics are being calculated. The kinds of toxicological and safety issues that are predicted include mutagenicity, tumorigenicity, reproductive effects, and irritancy. The algorithms used to calculate these properties are described in some detail. For instance the toxicity risk assessment is explained at http://www.chemexper.com/tools/propertyExplorer/toxicity.shtml. A substructure search process determines the occurrence frequency of any fragment (core and constructed fragments) within all compounds of that toxicity class. Similar explanations follow the fragment-based drug-likeness score (http://www.chemexper.com/tools/propertyExplorer/drugLikeness.shtml) and the overall drug-likeness score (http://www.chemexper.com/tools/propertyExplorer/drugLikenessScore.shtml). The OSIRIS Property Explorer is an integral part of Actelion's (http://www.actelion.com/) in-house substance registration system.

The Interactive Laboratory (I-Lab; http://www.acdlabs.com/ilab/) is a commercial product (but with a free demonstration version) available from Advanced Chemistry Development (ACD). It provides online computation of molecular physical properties for LogP, pKa, LogD, and aqueous solubility. I-Lab (for which 100 free credits per month are offered) also includes database searching of ACD's compilations of spectra and physical properties. The ACD/LogP calculator (http://www.acdlabs.com/resources/freeware/), now offered as freeware, has been compared with competitive products at http://www.acdlabs.com/products/phys_chem_lab/logp/competit.html. ACD claims it can calculate an accurate LogP derived from an internal ACD/LogP database containing over 5,000 experimental LogP values.

An interactive web service for the calculation of molecular properties relevant to drug design and QSAR has been established at the Molinspiration Cheminformatics website (http://www.molinspiration.com/cgi-bin/properties/). Properties calculated include LogP, total PSA, molecular volume, and Lipinski's rule of 5 parameters. A drug-likeness index will be available. Molinspiration is offering this as a free service to the Internet chemistry community for up to 100 determinations per month. Molinspiration works in conjunction with JME, a free Java applet that allows the generation and editing of molecules and reactions and the creation of molecule SMILES. JME was written by Peter Ertl from Novartis and has become a standard for molecule structure input on the Internet (http://www.molinspiration.com/jme/). Along similar lines, JMol is a free, open-source molecule viewer and editor initiated by Dan Gezelter at Columbia University. It is a collaboratively developed visualization and measurement tool for chemical scientists (http://jmol.sourceforge.net/).

The Reciprocal Net project, run by the Indiana University Molecular Structure Center (http://www.iumsc.indiana.edu/) is a distributed, open, extensible digital collection of molecular structures together with software tools for visualizing, interacting with, and rendering printable images of the contents. It provides automated conversion into standard formats that can be globally shared. The contents of the collection come principally from structures contributed by participating crystallography laboratories, which include universities from the US, the UK, and Australia. Reciprocal Net's common molecules include a section on therapeutic compounds (http://www.reciprocalnet.org/edumodules /commonmolecules/biochemical/list.html#therapeutic).

As far as visualization of chemical structure is concerned, the DEPICT service from Daylight (http://www.daylight.com/daycgi/depict/) accepts a SMILES string as input and returns an HTML page with an embedded image. Unfortunately, it provides no control over the output style and image size. Finally, OpenBabel is an open, collaborative project allowing anyone to search, convert, analyze, or store data from molecular modeling, chemistry, biochemistry, or related areas (http://sourceforge.net/projects/openbabel/). The project enables the interconversion of different chemical structure file formats.

IX. CHEMICAL SUPPLIERS

Currently, a very large amount of information on available chemicals is on the web. This information is relevant for both laboratory-scale synthesis and for larger-scale preparations, but it is more easily searched for laboratory synthesis.

Examples of sites for online searching of available chemicals are provided below:

Website address	Comments
http://www.buyersguidechem.de/	An excellent site with a wide variety of chemicals but no prices; useful for bulk and for MSDS; a directory of over 300,000 chemicals.
http://www.chemexper.be/	Excellent search capability on a wide variety of research chemicals; information includes the exact chemical name, formula, melting point, and other physical properties. Searching can be conducted by CAS number, molecular formula, substructure, name, and a range of other terms. Links allow the user to directly go to the individual supplier.
http://www.mdpi.org/molmall/	MolMall features the Rare Chemical Samples Exchange Center. Compounds are made available from small samples provided by individual researchers. Full-structure search or substructure searches are permitted on the website, as searches for the name of the submitter. Several other very useful searching functions are also available. Links to Molecules MolBank (http://www.mdpi.com/journal/molbank) papers are included if the compounds are published there.
http://www.chemindustry.com/apps/chemicals	ChemIndustry.com site enables the user to enter a product name to search a database of websites related to various chemical suppliers.
http://www.chemacx.com/	A commercial product through CambridgeSoft, Available Chemicals Xchange features the complete catalogs of over 200 vendors.

The site http://www.chemexper.com/ also allows access to Expereact™ WEB, a laboratory management program that helps the control of stock, ordering products, adding reactions (electronic laboratory journal), and exporting all the information to a word processor. Another site providing software for inventory management is

ChemSW (http://www.chemsw.com/). Products include the CIS Inventory System Pro 2000 and a digital MSDS digital filing cabinet (very useful for managing data sheets as they go out of date), as well as more conventional chemical drawing and molecular modeling programs.

Many suppliers offer their own database searching capabilities. Large companies such as Sigma-Aldrich have managed to offer a complete searchable database of their products by name, structure, and CAS number (http://www.sigma-aldrich.com/). They also feature online ordering via a secure interface. The smaller suppliers have been later in developing online databases with searching and secure ordering.

CHEMCATS (http://www.cas.org/content/chemical-suppliers/) is an online database accessible through Chemical Abstracts that contains over 13 million commercially available compounds, including pricing information when available from suppliers. Many compounds also include direct hyperlinks to suppliers' sites. CHEMCATS is routinely updated with new information provided by suppliers already in the database and with new suppliers and/or catalogs. This is another commercial product, but there is no up front fee. Pricing is pay-as-you−go. It can be accessed through STN Easy (http://stneasy.cas.org/).

X. CHEMICAL SYNTHESIS

WebReactions (http://www.WebReactions.net/) is a new, unique reaction search system offering direct retrieval of reaction precedents through the Internet. The WebReactions system is easy to learn and use; the user merely draws the reactant and product using a Java-based chemical drawing program. It is virtually instantaneous in displaying matches, not just for the input reaction itself but for as broad a range of analogs as desired.

The complete Organic Synthesis (OS) is now available free online at http://www.orgsyn.org/. Exact and substructure searches are supported following the download of a free ChemDraw plugin, as are chemical name, formula, OS reference, and keyword index searches. This site is available free of charge to all chemists and contains all of the eleven collective as well as annual volumes and indices. OS is a compilation of eighty-eight annual volumes containing selected and independently checked procedures and new reactions in the field of organic synthesis. Since the 1920s, volumes of OS consisting of synthetic procedures have been published annually. The first six collective volumes were published every ten years, while the last six covered five-year intervals.

For biotechnological synthesis, a superb database contains information on microbial bio-catalytic reactions and biodegradation pathways primarily for xenobiotic chemical compounds. It is called the University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) and can be found at http://umbbd.ethz.ch/. The goal of the UM-BBD is to provide information on microbial enzyme-catalyzed reactions that are important for biotechnology.

A number of blogs are oriented toward chemical synthesis:

Title	URL	Comments
Useful Chemistry	http://usefulchem.blogspot.com	A useful aggregative site that includes articles from a range of writers on various aspects of synthetic chemistry.
Org Prep Daily	http://orgprepdaily.wordpress.com/	Procedures for various simple reactions
Organic Chemistry Highlights	http://www.organic-chemistry.org	Stereoselective synthesis of natural products, new methods in synthetic organic chemistry, and computational organometallic chemistry in organic synthesis; 5−8 highlighted reactions per month and short reviews of organic, bioorganic, organometallic, and microwave chemistry; total synthesis of natural products and multicomponent reactions.
Curly Arrow	http://curlyarrow.blogspot.co.uk/	Discussion of all aspects of synthetic organic chemistry
Organometallic current	http://organometallics.blogspot.co.uk/	Organometallic chemistry-based blog
F-blog	http://www.fluorous.com/journal/	Fluorine-based synthetic chemistry
Totally Synthetic	http://www.totallysynthetic.com/blog/	Author is a synthetic chemist with S. Ley's group in Cambridge, England.

XI. CHEMOINFORMATICS SOFTWARE PROGRAMS

A few software programs capable of 3D structure, conformation generation, computer-aided drug design and/or molecular modeling are available with free licenses, at least for academic purposes:

Program name	URL	Description
CORINA	http://www2.chemie.uni-erlangen.de/services/telespec/corina/	3D structure generator; 1,000 structures can be generated for free.
FANTOM (Fast Newton-Raphson Torsion Angle Minimizer)	http://bose.utmb.edu/fantom/fm_home.html	Calculation of low-energy conformations of polypeptides and proteins, compatible with distance and dihedral angle constraints typically obtained from NMR experiments.
Moloc	http://www.moloc.ch/	Molecular modeling package, available as a free two month trial.
MMFF94	http://ccl.net/cca/data/MMFF94/	Set of validation molecules based on X-ray crystallographic data.
NEWLEAD	http://www.ccl.net/cca/software/SGI/newlead/README.shtml	Computer program for the automatic generation of candidate structures.
PADRE	ftp://ftp.CCL.net/pub/chemistry/software/UNIX/PADRE/	Analysis of the results of conformational searches and measurement of similarity and differences between molecules.
Pgchem::tigress	http://pgfoundry.org/projects/pgchem/	Chemoinformatics extension to the PostgreSQL database management system that enables PostgreSQL to handle molecules through SQL statements.
PyMOL	http://www.pymol.org/	Open source molecular visualization system
RasMol	http://www.umass.edu/microbio/rasmol/index2.htm	Molecular visualization software

A more extensive list of software generally available for pharmaceutical and biotechnological R&D is available from NetSci, a public information exchange (http://www.netsci.org/Resources/Software/Cheminfo/). This list includes chemical databases, reaction databases, QSAR, and other programs. The modeling section of NetSci is to be found at http://www.netsci.org/Resources/Software/Modeling/, and it includes both open-license and commercial software. A similar principle applies to the Computational Chemistry list at http://www.ccl.net/cca/software/. Most of the references in this section of the NetSci site are to software programs that are not free, even to academic licensees. Their exclusion from explicit mention in this review is not intended to imply any value judgment on their worth. Interested readers are encouraged to make their own enquiries if they wish to review the available offerings.

A. Datasets for Virtual Screening

ZINC is a free database of small molecules for docking that are commercially available (http://zinc.docking.org/). ZINC is a self-referential acronym for "ZINC is not commercial." It contains over 21 million compounds in ready-to-dock 3D formats. Downloads are available in sdf, mol2, and SMILES formats. Subsets of the libraries are available and can be browsed or searched by text, structure, or molecular properties. There is a subset of drug-like molecules assembled by searching the database according to Lipinski's rule of 5 (LogP < 5; mol wt < 500; number of H-bond donors < 5; number of H-bond acceptors < 10). It is also possible to create a bespoke subset by searching according to physical properties including structure, net charge, calculated LogP, rotatable bonds, number of H-donors, number of H-acceptors, polar desolvation, apolar desolvation, and molecular weight (http://zinc.docking.org/search/structure.php).

XII. CHEMICAL ANALYSIS

Analytik (http://www.analytik.de/) is a comprehensive German information site for analytical chemists. It relates discussions of analytical problems, contains a small but excellent link collection to chemical databases and literature (with an application database), and press releases from the German Chemical Society.

An open source NMR database, NMRShiftDB, is available at http://nmrshiftdb.nmr.uni-koeln.de/. The software and database content can be downloaded via http://www.sourceforge.net/projects/nmrshiftdb/, although online searching also works without any special plugins. NMRShiftDB is a web database for organic structures and their nuclear magnetic resonance (NMR) spectra. It allows spectrum prediction (currently only for 13C, 1H, and other nuclei) and searches by spectra, structures, and other properties. Currently, the database contains over 41,000 structures and over 49,000 measured spectra (as well as about 500 calculated spectra).

The Java applet that comes with NMRShiftDB includes an array of features for molecular display (e.g., ball-and-stick, wireframe, space-fill), translation into SMILES nomenclature, and the possibility of structure editing. Searching of chemicals can be based on chemical name, keyword, CAS number, literature title/author, and chemical formula among others.

XIII. CHEMICAL PUBLICATIONS

A. Journals

Nearly all journals have a web presence, and an increasing majority have electronic versions of their publications (including archives) available through the website. A convenient listing of them is available in the chemistry section of the WWW Virtual Library at http://www.liv.ac.uk/Chemistry/Links/journals.html. Salient journals related to medicinal chemistry include those shown in the table below:

Publisher	Journal title	URL
American Chemical Society	*Bioconjugate Chemistry*	http://pubs.acs.org/journals/bcches/
	Journal of Natural Products	http://pubs.acs.org/journals/jnprdf/
	Journal of Medicinal Chemistry	http://pubs.acs.org/journals/jmcmar/
	Organic Process Research & Development	http://pubs.acs.org/journals/oprdfk/
Bentham Scientific Publishers	*Current Medicinal Chemistry*	http://benthamscience.com/journal/index.php?journalID = cmc
	Current Pharmaceutical Design	http://benthamscience.com/journal/index.php?journalID = cpd
	Current Drug Discovery Technologies	http://benthamscience.com/journal/index.php?journalID = cddt
Current Drugs	*Current Opinion in Biotechnology*	http://www.current-opinion.com/journals/current-opinion-in-biotechnology/
Elsevier	*Bioorganic & Medicinal Chemistry*	http://www.journals.elsevier.com/bioorganic-and-medicinal-chemistry/
	Bioorganic & Medicinal Chemistry Letters	http://www.journals.elsevier.com/bioorganic-and-medicinal-chemistry-letters/
	Drug Discovery Today	http://www.drugdiscoverytoday.com/
	European Journal of Medicinal Chemistry	http://www.journals.elsevier.com/european-journal-of-medicinal-chemistry/
Nature	*Nature Biotechnology*	http://www.nature.com/nbt/
	Nature Medicine	http://www.nature.com/nm/
	Nature Reviews Drug Discovery	http://www.nature.com/nrd/
	Nature Chemical Biology	http://www.nature.com/nchembio/

The Experimental Data Checker and OSCAR toolkit is software to extract data from literature. It is available from http://www.rsc.org/Publishing/Journals/guidelines/AuthorGuidelines/AuthoringTools/ExperimentalDataChecker/. Experimental data on new molecules in organic and inorganic chemistry are presented in a standard form that varies little from journal to journal. Typically, the appearance of the compound is described, followed by its melting

points (if applicable); R_f, infra-red, and NMR data; and mass spectral information. OSCAR will extract this information from either a paragraph of experimental data or a full paper, and then run some checks to test the data for consistency. After the user pastes the experimental data into a form, the program returns both the analytical information and a critical assessment of the same. It can also plot the 1H NMR spectrum from the analyzed information.

B. Theses

Many universities are installing searchable and accessible thesis archives. At least theoretically, this is a welcome addition to the web-searchable pantheon of scientific literature. The practical difficulty associated with this task is the sheer diversity of information sources, which are not all archived in a central location. Lists such as these must be assembled by hand. They are not only vast but also constantly changing. A partial solution to this problem is addressed through the Networked Digital Library of Theses and Dissertations (NDLTD; http://www.ndltd.org/). Established in 1996, its members include hundreds of universities around the world, as well as partner organizations, including Adobe, the American Library Association, the Association of Research Libraries, the Coalition for Networked Information, the Joint Information Services Committee, OCLC Online Computer Library Center, Proquest/UMI, and Theses Canada. The Union Catalog Project (http://thumper.vtls.com:6090/) is an attempt to make these individual collections appear as one seamless digital library of ETDs to students and researchers seeking out theses and dissertations. ETDs are owned and maintained by the institutions at which they were produced or archived, while the metadata (e.g., title, author) have been gathered into a central search engine.

The Massachusetts Institute of Technology (MIT) Libraries' Document Services department is found at http://dspace.mit.edu/handle/1721.1/7582/. MIT is one of the institutions at the forefront of this effort, and its site offers the full text of selected master's and doctoral theses from all MIT departments. These include theses that have been previously requested and scanned by Document Services as well as theses from the university's pilot project in electronically-submitted theses. Users can search the database by keyword, perform an advanced search with separate fields, or browse by author or year. All theses can be viewed as low-resolution (100dpi) greyscale inline gif images. The theses of some of their Nobel Prize-winning alumni are available at http://libraries.mit.edu/docs/nobeltheses.html/.

XIV. PATENT INFORMATION

The major world patent databases are online and searchable, along with a plethora of tools for the desk scientific researcher.

Esp@cenet at the European Patent Office (EPO) Databases (http://www.epo.org/searching/free/espacenet.html?hp = stages) allows free online patent searching of over 30 million documents (in EPO member states and worldwide) by entering information such as keywords, patent numbers, and institute names. This website also supports translation to and from the major European languages.

The US Patent and Trademark Office (USPTO) web patent database (http://www.uspto.gov/patft/) provides access to the US Patent Bibliographic Database, which includes bibliographic data from 1976 to the present. There is a patent number search page, as well as boolean and advanced search pages for text field searching. Both cited and citing patents are hyperlinked to each patent. Hyperlinks connect the classification numbers and their definitions, and good help pages are available for each search type.

Although the USPTO site provides images of the actual hard copies, the user currently has to combine singly downloaded tiff or pdf files in order to generate a single-file document. This tedious process has been obviated by commercial patent engines such as MicroPatent (http://www.micropatent.com/) and Delphion (http://www.delphion.com/), but now there are free alternatives to address this issue. FreePatentsOnline (http://www.freepatentsonline.com/) is, as the name suggests, a freely accessible database that contains all patents published by the USPTO since number 4,000,000. It is automatically updated weekly, is searchable, and can retrieve images of the results from the patent text pages. The search methods are similar to those available at the USPTO site. Search terms can be entered in certain fields, such as title, abstract, and assignee (owner), to locate patents or published patent applications having the entered terms in the specified fields (in the specified sections of the patents or applications). Search strings can also be connected with Boolean terms such as AND, OR, and

ANDNOT, and parentheses can be used to order the connected terms. The ends of search terms can be truncated, and the wildcard symbol "$" can be used. The most comprehensive search method is based on searching the specification field. To identify US classes for particular fields of technology, users can access the Manual of US Patent Classification at http://www.uspto.gov/web/classification/.

SureChem (http://www.surechem.com) offers something extra in addition to freely available patent information, because it is able to mine public information patent documents for chemical structure information through chemical name and image recognition. Its dataset includes 20 million annotated US, EP, and PCT full-text documents and Japio abstracts. SureChem has extended its capability to include mining of Medline abstracts.

Google has established its own patent service (http://www.google.com/patents/) that covers the entire collection of granted patents and published patent applications from the USPTO and the EPO. US patent documents date back to 1790, while EPO documents since 1978 are available. In addition to a simple Google search box, it also offer an advanced mode (http://www.google.com/advanced_patent_search/) with criteria such as patent number, inventor, classification, and filing date.

A repository of general interest in patent literature and intellectual property-related news and decisions is available through IP News Flash (http://www.ipnewsflash.com/). It is a meta-information portal that browses your information channels and presents only relevant, recent, and customizable IP information on a single page, updated hourly with information on patents and other intellectual property-related matters. The site offers an e-mail news feed with content delivered daily or monthly.

Various other sites provide general information on patents. One example is the USPTO site about patents and patenting procedures (http://www.uspto.gov/web/offices/pac/doc/general/). Other useful bits of information about patent terms and procedures in other countries are available from Derwent (http://www.derwent.com/). A comparison table of web patent databases from Duke University is presented by the university library to help users compare the resources available and assess which is best for each individual's needs (http://guides.library.duke.edu/patent/).

A. Japanese Patents

The Japanese Patent Office (JPO) website (http://www.jpo.go.jp/) now provides certain information free in English. It provides more information in Japanese, including free legal status information from the JPO's intellectual property digital library (IPDL) pages. There are five methods of searching the IPDL patent database (http://www.ipdl.inpit.go.jp /homepg_e.ipdl). The form for retrieval of patent images based on patent number is somewhat difficult to navigate (http://www4.ipdl.inpit.go.jp/Tokujitu/tjsogodben.ipdl?N0000 = 115) but is backed up by a useful help area at http://www.ipdl.inpit.go.jp/HELP/tokujitu/db_en/help_index.html.

Further commercial resources available for English translations of Japanese patent documents include Paterra, Inc. (http://www.paterra.com/). The InstantMT™ service retrieves the requested patent by number and rapidly provides a translated version, which is rendered for download in a two-column formatted PDF file. The system covers all Japanese Kokai (A documents) published after January 1, 1993, and all granted Japanese patents (Toroku) published since May 27, 1996. New documents are entered into the system within two weeks of being published in Japan.

XV. TOXICOLOGY

There are a number of toxicology databases available on the Internet, and recently there has been an amalgamation of the best in the form of TOXNET (http://toxnet.nlm.nih.gov/), a cluster of databases on toxicology, hazardous chemicals, and related areas. The website provides access to an impressive array of files containing information related to the toxicity and other hazards of chemicals. Users can readily extract toxicology data and literature references, as well as toxic release information on particular chemicals. Alternately, one can perform a search with subject terms to identify chemicals that cause certain effects. A variety of display and sorting options are available.

A summary of further resources in this area (including subsets of the TOXNET database array) is provided in the table below:

Type of database	Database name	URL	Description
Toxicology Data Files	HSDB (Hazardous Substances Data Bank)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB	Databank of over 4,700 potentially hazardous chemicals. In addition to toxicity data, the files cover emergency handling procedures, environmental fate, human exposure, detection methods, and regulatory requirements. The data are fully referenced and peer-reviewed by expert toxicologists and other scientists.
	International Toxicity Estimates for Risk (ITER) Database	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?iter	A new database within the TOXNET site that contains human health-risk values from major organizations worldwide for over 600 chemicals of environmental concern. It is a product of Toxicology Excellence for Risk Assessment (TERA), a non-profit group whose mission is to protect public health by developing and communicating risk assessment values, improving risk methods through research, and educating the public on risk assessment issues.
	IRIS (Integrated Risk Information System)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?IRIS.htm	Online database of the EPA; http://www.epa.gov/) containing carcinogenic and noncarcinogenic health-risk information on over 500 chemicals. Data have been scientifically reviewed and represent EPA consensus.
	CCRIS (Chemical Carcinogenesis Research Information System)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?CCRIS	Sponsored by the NCI (http://www.nci.nih.gov/), CCRIS contains scientifically evaluated data derived from carcinogenicity, mutagenicity, tumor promotion, and tumor inhibition tests on about 8,000 chemicals.
	GENE-TOX (Genetic Toxicology)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX	Another EPA database. It contains genetic toxicology test results on over 3,000 chemicals selectively reviewed for each of the test systems under evaluation. The GENE-TOX data bank is the product of these data review activities.
	Columbia Environmental Research Center (CERC)	http://www.cerc.usgs.gov/	Acute toxicity of over 400 chemicals and sixty aquatic animals. The results have been provided from aquatic acute toxicity tests conducted by the USGS CERC. The acute toxicity test provides a relative starting point for hazard assessment of contaminants and is required for federal chemical registration programs for fungicides, rodenticides, and pesticides. Data are organized and searchable by combinations of compound and species data (e.g., LC-50 data for various chemicals and exposure times).
Toxicology Literature Files	TOXLINE	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE	Bibliographic database covering the biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals. Contains over 4 million citations, almost all with abstracts and/or index terms and CAS Registry Numbers.
	DART/ETIC (Development and Reproductive Toxicology/Environmental Teratology Information Center)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC.htm	Bibliographic database covering literature on teratology and other aspects of developmental toxicology, containing over 200,000 references to teratology literature published since 1965.

(*Continued*)

(Continued)

Type of database	Database name	URL	Description
Toxic Releases Files	TRI (Toxic Chemical Release Inventory Files)	http://www.epa.gov/tri/	Contains information on the annual estimated releases of toxic chemicals into the environment. It is based upon data submitted to the EPA from industrial facilities throughout the US and includes the amounts of certain toxic chemicals released into the environment on over 650 chemicals and chemical categories. Pollution prevention data are also reported.
Carcinogenicity	Carcinogenic Potency Project	http://potency.berkeley.edu/listofpubs.topic.html	A useful resource related to carcinogenicity, it includes a wide array of publications from the Carcinogenic Potency Project. The references include papers on methodological analysis of the relevance of carcinogenicity prediction from bioassays, species comparisons, target organs, mechanism of carcinogenesis, risk assessment techniques, possible cancer hazards of natural and synthetic chemicals and causes, and prevention of cancer.

PubMed links to chemicals found in TOXNET's HSDB (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB) through a LinkOut feature that appears when a user clicks the links part of any PubMed reference, shown on the far right-hand side of the screen. The links now appear as specific chemical names. LinkOut provides PubMed users with connections to full-text articles, consumer health information, and supplementary data related to a PubMed citation.

The NCMS has a free and fairly extensive database of solvents that allows rather extensive input of physical property ranges and various user-specified limits such as "not a carcinogen" or "not listed on the Montreal protocol (ozone)" (http://solvdb.ncms.org/solvdb.htm).

Fee-based resources include the updated US EPA Toxic Substance Control Act (TSCA) Chemical Inventory of over 64,000 chemicals, which is available cross-referenced with Superfund Amendments and Reauthorization Act (SARA) Title III RCRA reporting requirements on CD-ROM. It features SARA III fields integrated with TSCA information and PDF for instant search/retrieval. For details, see http://www.ntis.gov/products/ots.aspx.

Finally, the OSIRIS Property Explorer (http://www.chemexper.com/tools/propertyExplorer/), listed above in the section Prediction and Calculation of Molecular Properties, has useful capabilities for prediction of toxicological properties.

XVI. META-SITES AND TECHNOLOGY SERVICE PROVIDER DATABASES

Meta-sites providing access to a range of resources devoted to chemistry that have not already been referenced are listed in the following table:

Title	URL	Comments
Chemistry section of the WWW Virtual Library	http://www.liv.ac.uk/Chemistry/Links/links.html	Thorough, up-to-date, and accurate listings of a large number of chemistry sites. The chemical database section at http://www.liv.ac.uk/Chemistry/Links/refdatabases.html gives details of a collection of about 110 chemical databases (among them: Analytical Abstracts, Beilstein, CCDC, CA Selects Plus, ChemFinder, DrugDB, FT-IR Library, and STN).
Organic Chemistry Resources Worldwide	http://www.organicworldwide.net/	Created by Koen Van Aken, a Belgian chemist. A well-organized and highly useful site for all engaged in synthetic organic chemistry research.

(*Continued*)

(Continued)

Title	URL	Comments
Caltech	http://libguides.caltech.edu/chemistry/	An indispensable point of call for databases and search engines for chemistry.
The Chemical Database Service (CDS)	http://cds.dl.ac.uk/	The CDS provides online access to numerous chemical databases, which are available free of charge to academics at UK universities. The chemistry links cover a large variety of topics (among them general information sites, reference databases, chemical sources, chemical websites, UK universities, and chemistry FTP sites).
University of Cincinnati Online Database Collection	http://www.engrlib.uc.edu/selfhelp/alphlist.htm	Links to engineering, biology, and chemistry databases are listed on this important site.
Chemiedatenbanken	http://www.chemie-datenbanken.de/	An excellent collection of German and international chemical databases (e.g., free resources, general collections, and commercial database providers).

XVII. CONCLUSION

Medicinal Chemistry, like all other sciences, has been transformed by the Internet. The salient changes involve much easier access to information, both in terms of documents and as searchable databases; increasing access to free information, since paid-for resources have to clearly evince value in order to gain subscribers; and increasing possibilities to conduct calculations of (for instance) molecular properties online. Unfortunately, it is the nature of the Internet that this information changes on a regular basis, making some of the entries in this chapter likely to become redundant or outdated over time. Nevertheless, we have passed through the most revolutionary period in the Internet's formation and are entering a period of maturation during which a certain stability will ensue.

CHAPTER

36

Protection of Inventions in Medicinal Chemistry

Richard Luthi[1] *and Christopher Brice*[2]

[1]Founding partner, Luthi & Co., Intellectual Property Law Offices, of Tel Aviv, Israel; [2]Consultant, Luthi & Co., Intellectual Property Law Offices, of Tel Aviv, Israel

OUTLINE

The Congress shall have power… to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries. ***US Constitution Article One, Section 8, Clause 8***

I. PATENTS AND THE MEDICINAL CHEMIST

A patent is a time-limited monopoly over an invention, granted by the state. It provides the owner with the right to stop others from using the invention without permission. Once the patent monopoly period has elapsed, anyone is free to use the invention as they please.

While most governments see it as their duty to provide their citizens with a certain level of medical care and to fund academic institutions, generally they do not directly engage in or fund the development of new drugs,

relying instead on the private sector. Patents are a means for governments to create the necessary incentive for the private sector to take the huge risk of investing in pharmaceutical research and development. The idea is simple: the patent system affords a period of exclusivity to those who have taken the financial risk and successfully developed an innovative medicine, so that they may seek to recoup their investment and make a profit before facing competition from others who have taken no such risks. Of course, in those countries where governments pay or contribute towards the price of drugs for patients, the government does indirectly fund investment in those drugs that reach the marketplace. In other words, they reward success after the event but do not involve the taxpayer in the high risk of failure.

Patents are crucial for the innovative pharmaceutical industry. Without a patent monopoly or some other form of product exclusivity, it would be impossible to recover the enormous costs of drug design and discovery, preclinical testing, and (especially) clinical trials. Patents are not only important to large multinational companies; they are also very important to universities and to small start-up companies that are engaged in early stage medical research but do not have the resources to develop drugs and take them all the way to the market. They generally expect to sell the company or to license the technology to large pharmaceutical companies. Most pharmaceutical companies and investors will not invest in an unproven drug or buy a start-up company unless they are assured of adequate patent protection.

In addition to the potential to reward innovation, patents have other important roles. In order to receive a patent, an inventor is obliged to disclose his invention in sufficient detail to allow a person of average skill in the art to practice it. Patents are therefore an invaluable source of technical information and competitor intelligence. The disclosed information becomes rapidly available because most patent applications are published within eighteen months after the application is filed.

Chemists tackling a particular problem can gain valuable information from third-party patents and applications that address similar problems. We frequently see that the publication of initial data on the first member of a new and useful chemical class is followed in short order by the testing of other similar but better molecules addressing the same medical problem or biological target.

By monitoring the patent filings of rival companies, pharmaceutical companies can ascertain competitors' directions of research. Particularly promising areas of investigation can usually be identified by the intensity of patenting activity. By monitoring the filings of others, companies can also ensure that their own products do not infringe competitors' patents. If a pharmaceutical company considers that a competitor's patent or application unjustifiably hinders the use of technology that it wants to employ, steps can be taken in advance to challenge the patent or try to obtain a licence from the patentee. In the USA, the first generic company to succeed in challenging the validity of the patent, achieves six months' exclusivity from generic competition. When the patent expires, the disclosed information allows generic companies to offer the same molecule at significantly lower prices.

Having said all that, this chapter is not directed only to those involved in the development of novel drugs. Generic companies must evaluate all relevant patents covering the product they are developing in order to ensure that their generic product is indeed free from any claim of patent infringement and, to the extent necessary and possible, to develop technological solutions that do not fall foul of the patents.

Accordingly, there is good reason for the medicinal chemist to know something about patents for medical inventions, whether he or she is engaged with innovative research or generic drug development. But this chapter is not a condensed legal textbook and our aim is not to turn chemists into lawyers. Rather, we aim to help the chemist understand the rudiments of the patent system and how best to secure and enforce patents, to challenge unjustified monopolies, and to know when to turn to a patent attorney for advice on patentability and how to assist a lawyer in doing a good job of enforcing patents.

Any patent attorney who takes on the job of drafting patent applications relating to medicinal compounds will almost certainly have formal training in chemistry, pharmacy, or biology, but even an in-house patent attorney with experience in handling inventions from a particular research group will not have the detailed knowledge and expertise of the inventors themselves. Lawyers handling patent litigation may have no formal training in chemistry at all, though they will generally be experienced in coming to grips with all kinds of science and technology in considerable detail in a short space of time. Patent work is necessarily teamwork, made more effective (and more interesting) when each member of the team has some understanding of the framework in which the others are operating.

The development of a new drug, whether a synthetic or a biological molecule, is a lengthy and expensive stepwise process in which, typically, many inventions and technological innovations are made by persons with various scientific skills: biologists, physicians, chemists, pharmacologists, pharmacists, and others. A biologist may have discovered a new biological pathway that may be stimulated or inhibited to achieve a beneficial clinical outcome. Chemists and biologists may have discovered novel, safe, and efficacious molecules that stimulate or

antagonize that biological pathway and designed innovative processes for their industrial manufacture. They may synthesize unique and advantageous salts, hydrates, or crystal forms of molecules. Pharmacists may formulate unusual delivery systems. Chemists and biologists may develop novel diagnostic tools. All of these and many more such discoveries and technologies may be inventive and qualify for a patent. The medicinal chemist—and particularly the team leader—needs to be alert to the need to secure patent protection for these frequently very valuable technological advances that required the investment of substantial resources and expertise.

The development of new medicines is unique among paths of commercial endeavours, not only in the high cost and risks involved but also in the lengthy period, typically twelve to fifteen years, after a new molecule is first synthesized before it reaches the market. This means that the twenty-year monopoly period of patents filed in the early days of research that covers, for example, the new chemical entity or its medicinal use, will be largely used up during the development phase, so that by the time the product is launched only a very limited term of patent protection will remain. Generally, the term is not long enough to recoup the investment. Evolving patent law has sought to compensate for this by extending the period of protection for medicinal patents by a few years, but this is frequently still insufficient. However, later patents, covering technology that naturally only emerges during the advanced stages of development—such as industrial synthesis or commercial formulations—may offer additional layers of protection to ensure marketing exclusivity and adequate returns from the cumulative investment.

Critics of the innovative pharmaceutical industry, in particular companies who want to sell generic equivalents as soon as a medicinal compound comes off patent, frequently complain that patents beyond the initial compound patent are an unfair "evergreening" of the originator's monopoly. However, there is no good reason why later inventions—made at considerable expense and with the input of specialist knowledge and skills—that offer technological advances that genuinely meet the requirements of patentability should not be duly protected.

Patents are territorial. No worldwide patents exist. Separate patents need to be secured for each country, and the national patent can only be enforced to prevent infringement in that particular country. There is a regional patent offices in Europe (and in Eurasia and part of Africa), which examines patent applications for the entire region but the European Application results in the grant of a bundle of national patents. The law and the court systems vary from jurisdiction to jurisdiction. This means that patent applications for the same invention may be granted in one country and rejected in another. Not infrequently, a patentee may succeed in enforcing his patent in one country but not in another, although the same allegedly infringing product is involved.

Naturally, a patent can only be enforced in those territories where it is granted and in force. Some of the poorer developing countries, as well as countries that have a relatively developed generic drug manufacturing industry, have been reluctant to provide patent protection at the level customary in developed countries. Until a few years ago, many countries would not grant patent protection for novel molecules for medicinal use. The situation improved following the TRIPS Agreement,[1] signed by almost every country in the world, and most countries have now implemented patent systems for medicinal products. They may not all be perfect, but they do provide a degree of protection. Nonetheless, some countries, (e.g., India) that are motivated by local economic interests still offer a lesser degree of protection for drugs, much to the frustration of large innovative companies.

Companies with a global presence must be versed in the particularities of patent law in many countries and most certainly in the larger markets. We do not intend to deal here with detailed local differences. Rather, we shall try to explain, in general terms, principles that are common to most legal systems, with some reference to the peculiarities of the more important jurisdictions, especially the USA and the EPO.

The medicinal chemist should be aware that in addition to patents, other types of intellectual property protection are available, including various forms of exclusivity that are particular to medicinal products. We shall discuss these briefly in Section VI below.

II. WHAT KINDS OF MEDICAL INVENTIONS CAN BE PATENTED?

Patents are granted for inventions capable of practical application and not (as the common phrase has it) for "mere ideas." Subject to the general requirements and statutory exclusions discussed in Section III below, it is possible to secure patents for the invention of new and useful products (meaning things that can be made), methods, and processes.

[1]Agreement on Trade-Related Aspects of Intellectual Property Rights. Annex 1C of the Marrakesh Agreement Establishing the World Trade Organization, signed in Marrakesh, Morocco, on 15 April 1994.

Obviously, new and useful medicinal molecules are eligible. A new medical (or other) use for a molecule could be patentable, whether or not the molecule itself is new. So could the use of a new drug-discovery target or screening method. Inventions relating to the mode of administration or dosing regimen of drugs, including combining two or more known molecules into a single fixed-dose combination, may also be eligible for patent protection.

Novel manufacturing processes can be the subject of a patent, whether the compound thereby produced is new or old. Examples include processes with a higher yield or a purer product; processes employing fewer, safer, or simpler manufacturing steps; processes using fewer or cheaper reagents; and processes requiring the consumption of less energy. Any special equipment that facilitates an advantageous manufacturing process might also be patented. The discovery of a chemical intermediate could qualify as a useful invention. Patents relating to the manufacturing process are inherently less useful than a patent covering the final medicinal compound, because they do not prevent third parties from making the drug but only prevent them from doing so using the particular process, intermediate, or equipment that is the subject of the patent. In some cases, it may be difficult to devise a cost-effective alternative process, or the commercial advantage of the invention may be hard for generic manufacturers to resist. In such instances, these patents have real value.

Chemists are well aware that the salts and physical forms of many compounds, as well as other modifications, can have a radical effect on their stability, handling, dissolution and bioavailability, and pharmaceutical companies expend considerable efforts developing the most advantageous forms of their molecules. To give a few examples that may be appropriate subject matter for a patent:

a. salts, esters, amides, and other derivatives;
b. crystals or amorphous forms;
c. hydrates or solvates;
d. other physical characteristics, such as particle size or shape;
e. resolved isomers; and
f. pro-drugs designed to metabolise to the desired active molecule.

Patents for different forms of a medicinally active compound can be important, because they offer distinct technological advantages. It may also be difficult for a company seeking to manufacture the original molecule after the basic patent covering it has expired to achieve bioequivalence and the necessary stability and purity required by the regulatory authorities without employing the same form of the compound. These patents are more difficult to secure, since in many cases the relevant technology is known. Only in circumstances where there is something unexpected is the particular novel form (or the process needed to obtain it) would a patent be granted. However, where the invention is, for example, a new synthetic process, intermediate product or crystal form then it may as a practical matter be very difficult to demonstrate the use of the invention in a competitor's finished product.

Possibly less important for the medicinal chemist—but very important for pharmaceutical companies—are patents relating to delivery systems (formulations) and (again) processes for their production. Particularly where the administration of a problematic molecule is challenging, formulation patents may be difficult to work around because of regulatory requirements. A competitor cannot market its generic drug in an inferior formulation; the generic product must be bioequivalent to and at least as stable as the original drug. Even inventions relating to the product packaging of sensitive molecules (e.g., photosensitive or hygroscopic molecules) may be patentable subject matter.

Many other types of inventions may have commercial significance. It may be important to patent diagnostic tools and methods, especially as clinical practice moves towards individualised therapy where drugs are administered to patients only when specific test results (such as the presence or absence of certain gene mutations) show that the patient is likely to respond to therapy.

III. THE BASICS OF PATENT LAW

A. The Requirements for Patentability

From late medieval times, state sponsored monopolies ("letters patent") have been used to promote trade and manufacture, but also to reward political favorites and to raise revenue. During the modern era, a strong patent system has become one of the cornerstones of policy for promoting scientific and industrial development, and

the implementation of such a system was deemed a sufficiently important part of government function to merit mention in the US Constitution, as can be seen from the quotation that introduced this chapter.

The patent system may usefully be regarded as a social contract, the basis of which is succinctly and elegantly described in the following textbook passage, still pertinent some forty years on: [2]

> "[I]n order to encourage improvement, and to encourage also the disclosure of improvements in preference to their use in secret, any person devising an improvement in a manufactured article, or in ... methods for making it, may upon disclosure of his improvement at the Patent Office demand to be given a monopoly in the use of it for a period of [20] years. After that period it passes into the public domain; and the temporary monopoly is not objectionable, for if it had not been for the inventor who devised and disclosed the improvement nobody would have been able to use it at that or any other time, since nobody would have known about it."

Accordingly, the fact that the subject matter of a patent application is eligible in principle for a patent does not mean that it is assured of receiving patent protection. The invention must satisfy four basic conditions defined by the elements of the social contract between the inventor and the public, namely novelty, inventive step, utility, and sufficiency of description. The corresponding questions are:

a. Was the claimed invention new?
b. Did it involve an inventive step?
c. Is it technologically useful?
d. Does the description in the patent sufficiently disclose how to put the claimed invention into practice?

If the answer to any of these questions is "no," then the patent application should be refused. If the application is nonetheless accepted by the examiners in the patent office, then the granted patent will be invalid and may be attacked by others whose business is threatened by the unjustified monopoly.

Although the patent laws of most countries adopt these four conditions of patentability in one form or another, there is significant variability in how the requirements are applied. Not infrequently, this means that a patent may survive an attack on its validity in one country, and yet be revoked elsewhere, simply because the conditions are interpreted differently.

Apart from these four substantive requirements, a patent may be refused or held invalid on many other, more formal grounds, such as ambiguity of the claims, lack of support for the claims in the specification, the addition of new material during patent prosecution, lack of entitlement to the claimed priority date, or failure to meet various deadlines. Patents may also be rejected on grounds relating to the patentee's conduct, such as failure to make certain disclosures required by law or other actions considered to be evidence of bad faith. Furthermore, if the applicant has misappropriated the invention or if he is not for any other reason the rightful owner of the patent, he may be stripped of the patent, which can also be transferred to the rightful owner. These objections vary greatly from jurisdiction to jurisdiction and involve legal technicalities that go beyond the scope of this chapter.

We will now discuss the four principal conditions for patentability, which provide the major substantive grounds on which applications may be refused or patents may be attacked for invalidity:

1. Novelty

Novelty is the most self-evident and easily understood requirement. The grant of a monopoly is only justified because the inventor has developed and disclosed in the patent application something new that has never been disclosed before. If the invention claimed by the applicant has already been disclosed in a written or oral description or by use—anywhere in the world—in such a was as to make the invention available to the public before the date of the patent application, the inventor has contributed nothing new that would warrant granting a period of exclusivity for that invention. Though the requirement for novelty is apparently simple, its application raises some complicated issues.

When an invention lacks novelty, the patent is said to be anticipated by the prior art and invalid for anticipation. The relevant date for considering novelty is the priority date of the patent, which is the filing date of the earliest patent application by the inventor that describes the invention (see section III.E below).

Novelty is only assessed against public disclosure (usually in a published document) or public use. The question is not whether anyone actually read the prior art document or inspected the product or apparatus disclosing or implementing the invention, but rather whether a member of the public could have obtained access as a matter of right if they so desired. If it were kept secret, an earlier document in which someone else described

[2]Blanco White, T.A. Patents for Inventions and the Protection of Industrial Design. 4th ed. London: Stevens & Sons; 1974. p.1.

exactly the same invention would not be considered as prior art that could nullify the novelty of a later-filed patent. The same applies if an identical invention had previously been demonstrated by another in a private showing that was confidential and not open to the public. In such a case, the invention is not anticipated.

This exclusion from consideration of secret prior knowledge or use reflects the social contract; the patent is a reward for making the invention available to the public for the first time (even if the act of invention was a matter of pure serendipity) and not a prize for intellectual ability or practical ingenuity. For the same reason, where two applications for the same invention are filed close together in time, the patent is granted to the person who filed the first application (i.e., to the inventor with the earlier priority date) and not to the person who first made the invention. Until recently, the USA was the exception in this regard: priority was given to the first person who made the invention even if someone else filed his patent application first. Following the passage of the America Invents Act in 2011, which made many sweeping amendments to US patent law, the priority of all patents filed after 16 March 2013 is to be judged by the "first to file" rule.

However, patent applications are only published eighteen months after filing. During that period, an applicant cannot be certain whether the invention has been pre-empted or anticipated by another. The difficult question of whether an unpublished patent application constitutes novelty-destroying prior art for a later application for the same invention is resolved by provisions that differ from country to country.

The rules governing the requirement for novelty lead pharmaceutical companies to file a patent application as soon as possible after an invention is made, because if delay enables a competitor to beat them to the patent office or to publication, an earlier date of invention would be of no use in obtaining a patent or in challenging the competitor's earlier-filed patent. The company may find itself unable to exploit commercially the invention it has made, even though it had no prior knowledge of the competitor's parallel invention.

The disadvantage of early filing is that it starts two clocks running: (1) the term of the patent monopoly; and (2) the time at which the invention will be published. In most cases, neither of these is likely to justify any deferral.

The requirement of absolute novelty also means that the inventor should ensure that no disclosure of his invention is made before filing the patent. Scientists like to publish, to present their work at conferences, and to discuss their achievements with colleagues at scientific meetings, but this should never be done before a patent application has been filed. Although under US law[3] (and to a lesser degree in some other jurisdictions), the inventor has a one-year grace period in which to file a patent after the date that he himself publishes the invention, it is generally unwise to rely on this grace period unless the inventor is only interested in obtaining a US patent, as the inventor will be unable to obtain a patent in many other important jurisdictions.

In order to deprive an invention of novelty, the prior publication or prior public use must disclose the entire invention. A partial disclosure, which makes available to the public only some components of the claimed invention, would not destroy novelty, even if the missing components could be supplemented by reference to another publication. It is frequently said that to establish lack of novelty one cannot make a mosaic of prior publications. To put it another way, only something that would infringe the patent if done after the patent were granted would anticipate the patent (i.e., would destroy novelty and validity) if made public before the filing date.

Because only a complete prior disclosure will anticipate an invention, a patent can incorporate and build on earlier inventions. The novelty comes from the additional elements or the combination of elements. It follows from this that patents are not mutually exclusive. To use a simple mechanical example, neither a prior patent for the bicycle nor a prior patent for the internal combustion engine would have destroyed the novelty of a later patent for the motorcycle, because neither discloses the entire motorized vehicle. A third-party motorcycle could potentially have infringed all three patents: the engine patent because of the use of an engine, the bicycle patent for the use of a two wheeled vehicle and of course the motorcycle patent itself. (In practice, infringement would depend on the scope of the specific patent claims, as compared to the specific motorcycle.)

For the same reason, the discovery of a new medical use of a known medicinal compound can be patented if all other conditions and limitations of the law are met: a prior patent for (or other publication of) the compound as such would not destroy novelty, provided the publication did not also disclose the specific use claimed in the later patent.

Another rule, which also differs somewhat from jurisdiction to jurisdiction, is that a prior publication, prior use, or unpublished patent application will only destroy the novelty of a patent if the invention was disclosed in such a way that it could have been put into practice by those with the relevant skills and understanding. (This reflects the requirement that a patent must disclose the claimed invention in such a way that it can be performed

[3]35 U.S.C. § 102(b) (2014).

by others working in the relevant field; see "sufficient description," section III.A.4 below.) For example, a scientist may present a poster at a conference, describing early *in vitro* results for a new medicinal compound, giving only the company product code without actually disclosing the formula of the molecule. She could even hold up a sample, without describing the active ingredient or allowing third-party analysis. Such a limited disclosure would not destroy the novelty of a patent relating to the medicinal compound, because it would not enable anyone to practice the invention. Nonetheless, even such limited disclosures are unwise, because of the risk of some unintentional disclosure that may undermine patent validity.

One novelty issue is of particular relevance to patents for new chemical entities, namely, whether it is possible to patent a specific product that was included in some broad disclosure in an earlier patent or publication. This question arises most frequently in chemical cases because of the need for the inventor to claim more than the handful of active compounds that she has actually synthesized in order to fend off competition from someone who makes unimportant changes to the molecule. This problem (which arises also in sections III.A.3 and IV.B below) has resulted in chemical patents using a kind of algebraic structural formula, in which various substituents are represented by variables (traditionally designated as R) in place of specific elements (C, H, N etc) or groups. The variable R is then defined to be any one of a list of possible atoms or groups (e.g., to quote the sildenafil patent, "wherein R^1 is H, C_1 -C_3 alkyl, C_3 -C_5 cycloalkyl, or C_1 -C_3 perfluoroalkyl"). This chemical shorthand is known as a Markush formula (In honour of Dr Eugene A Markush, a prolific inventor of dyestuffs in the first half of the twentieth century, who fought a US Patent Office rejection of his generic patent claims; but this kind of generic formula may have been used as early as the 1850s[4]) and allows vast numbers of compounds to be defined in a generic claim with relatively little effort. When multiple variables are used (R^1 to R^N) and some of the variable groups are in turn defined only in general terms (such as "substituted phenol, aryl, or carbocyclic") or even by reference to further variable groups, the possible permutations quickly multiply and can cover literally millions of different compounds.

This means that the patentee will have tested only a few of the compounds covered by his claims, and he may certainly fail to discern that a particular subclass of compounds encompassed by the formula possesses a uniquely advantageous property limited to this subclass. If the subclass and its particular advantageous property are discovered later—whether by the same inventor or by another—then it may be possible to obtain another patent that selects out those individual compounds or that specific subclass of compounds. These selection patents, as they are called, are regarded in some jurisdictions as meeting the requirement for novelty, even though the compounds were notionally disclosed in the original patent. As one senior UK judge recently said, disclosing a forest does not make available a particular and special leaf of a tree in that forest.[5]

The limitations of the later selection patent have to be appreciated. It could very well be that the selected compounds infringe the broader patent, so that the later inventor will have to wait until that earlier patent expires or secure a licence under it. On the other hand, the earlier inventor would be prohibited from using the selected subset of compounds in the later patent, although she notionally invented those compounds and disclosed them in the earlier patent.

Novelty issues can also arise when patents are sought for single enantiomers of racemic compounds or other stereoisomers, where the specific isomer is difficult to separate from the racemate or in the event that it possesses unexpected properties and uses. It has been held in some jurisdictions that the publication of the mixture will not necessarily destroy novelty, because it is not a disclosure of the single isomer or of the new use or properties.

2. *Inventive Step*

The fact that a product or process is new (has not previously been made available to the public) does not mean that it is necessarily entitled to a patent. A patent monopoly is justified only by the disclosure of something new that would not readily have occurred to a skilled man without invention or, in patent law terms, that was not obvious.

The test is whether, without the benefit of hindsight, the invention disclosed in the patent would have been obvious to persons skilled in the field and familiar with the state of the art. The inventive step need not be very high. To continue the metaphor, the test presents quite a low threshold to patentability. Few patents would be granted if only breakthrough inventions were worthy of a patent. Certainly, the simplicity of an invention is no bar to satisfying this requirement. An invention may even be brilliant in its simplicity.

[4]Cooke, *Org Biomol Chem* (2004), 2, 3179-3191, at 3182.

[5]*Dr Reddy's v. Eli Lilly* [2010] RPC (9) 222 at 236.

While the obviousness test is difficult to define accurately or to quantify, it simply means that the claimed improvement should not stand out as the natural thing to do. "Obvious" has been said to mean that the improvement or solution to a problem is "very plain" [6] or that it would "at once occur" [7] to any skilled person who was addressing the problem addressed by the alleged invention.

As with novelty, the relevant date is the priority date of the patent, and again the background is all the relevant information in the public domain (unpublished patent applications are not relevant here). Unlike the requirement for novelty, in which any objection must be based on a single prior art document (or prior use) that discloses the claimed invention in its entirety, one may show that the invention was obvious by combining disclosures from a number of prior art references and the common general knowledge of competent workers in the relevant field. In other words, one may make a mosaic of the published prior art references, provided it would have been obvious for people in the field to assemble that combination of documents.

The requirement for an inventive step does not involve the logical or philosophical complications discussed in the section on novelty, but as a practical matter, the question of what would be obvious in any given case is far more difficult than the factual question of whether an invention is novel.

First, as mentioned, no precise and uniformly applicable yardstick by which to measure an inventive step is available. Second, in considering the inventive step of any patent or application while knowing all the time what the invention is, one must try to determine in retrospect—but without using hindsight—the mindset of a hypothetical person skilled in the art who had no knowledge of the invention.

The great difficulty in determining inventive step is transporting oneself back in time to see how things appeared at the priority date, before the idea embodied in the patent application was known. Anyone who has seen a magician's trick and then been told how it was performed is familiar with the way the feeling of wonder and joy subsides into a feeling that one has been taken in by nothing more than a tawdry and simple deceit. We are often a little disappointed in ourselves: How could we have failed to see the obvious? So, too, the magic of an invention falls away when it is observed and dissected *ex post facto*. It is safe to say that with the advantage of hindsight, the majority of inventions seem obvious, even when they are not.

A third difficulty stems from the need to decide which prior art publications this hypothetical skilled artisan would have considered. Here it is very difficult (but still necessary) to avoid hindsight. Once the invention has been made, it is much easier to point to various publications that include components of the invention. The relevance of these documents—and more importantly the idea of bringing them together—may not have been apparent at the time the invention was made but has only become obvious in retrospect.

A fourth difficulty stems from identifying the skilled person for whom the invention would or would not have been obvious. If the theoretical skilled person is a very highly trained, clever, and creative person, he or she may be capable of making inventions, in which case anything may be obvious to him or her. The average or unimaginative person working in the field would not readily appreciate the same solution. The degree of skill and imagination of the skilled person is generally a major controversy in patent litigation.

The skilled person is considered an average or ordinary person of the art with a good grasp of the relevant field and a capacity to unearth the relevant publications, but rather unimaginative, devoid of a creative spark, and not prone to making creative leaps. This test, like the test of obviousness itself, is easy to formulate but difficult to apply. For example, it may be almost impossible to work at all in the field of medicinal chemistry without a considerable ability to find creative solutions to problems as they arise in the laboratory or production plant.

Even the exact field of expertise in which the ordinary person is skilled can be hotly disputed in particular cases. Modern inventions are frequently made by teams of scientists having various skills. In such cases, the notional skilled addressee of the patent may also be a team of scientists operating in the field. Then again, it might have been part of the inventive step" to put together the right skill set in the first place in order to come up with the invention. (We have sometimes used the term "inventive journey" to describe a multistep inventive process.)

Although the obviousness test is the standard for patentability in almost all patent law systems in the world, great diversity exists in the application of the test, which leads to different and often unpredictable results in various jurisdictions.

[6] *General Tire v. Firestone* [1972] 89 RPC 457 at 497.

[7] *Vickers v. Siddell* (1980) 7 RPC 292 at 304 (HL).

In the EPO, a very structured approach has been devised. It is generally called the "problem/solution approach." First, one must determine the closest prior art to the claimed invention, which is defined as the most promising starting point—in the eyes of the skilled man of the art—for developing the invention. It is the piece of prior art that achieves the most similar result in the most similar way. Then, one must identify the objective technical problem, that is, the problem solved by the patent that was not solved by the closest prior art. Finally, starting from that closest prior art reference and bearing in mind the objective technical problem, one must determine whether the solution of the patent would have been obvious.

The major difficulty with this formulaic approach is that the choice of the key publication representing the closest prior art can only be made with hindsight, based on a knowledge of the invention of the patent. Because the problem to be solved is defined by the closest prior art, the definition of the problem is also affected by hindsight.

The US courts have grappled for many years with trying to establish a single clear legal test for obviousness. One frequently used test called "teaching-suggestion-motivation" (TSM) laid down the principle that if the various features of the invention are only made obvious by combining various prior art publications into a mosaic, then the prior art must also supply some reason or motivation to combine those disparate pieces of prior art.[8]

In 2006, the US Supreme Court determined that the TSM test had previously been applied in an excessively rigid manner and concluded that an improvement may be obvious, even though the need to combine the prior art references is not itself specifically taught in the prior art.[9] This has increased the uncertainty of US decisions on obviousness.

Sometimes, the best way to assess inventive step is to consider real-life evidence as to what other scientists did when faced with a comparable problem. Rather than concentrating on a particular prior publication chosen in hindsight with knowledge of the invention under consideration, a survey of the relevant literature may show that the natural and obvious way forward "in real time" was quite different from the approach that led to the claimed invention. Other real-life evidence can also be of help. For example, proof that there was a well-known and long-standing need to solve the problem solved by the inventor is strong evidence that the solution was not obvious, particularly when the invention becomes a commercial success soon after its creation.

Similarly, the difficulties encountered by the inventor in reaching the invention or the failures of others in the field when attempting to do so may be indicative of invention, although some courts, notably the UK Patents Court, have criticized this approach as too subjective. Certainly, the fact that an inventor may have reached the invention quickly and without undue difficulty is no proof that the invention was obvious. It may simply be proof of her ingenuity and originality.

Other considerations of this nature may be the reaction of the scientific and commercial community to the invention. Did other scientists react with disbelief on hearing of the invention? Were they filled with admiration? Either of these reactions would indicate that the invention was not obvious. Did other scientists or commercial entities immediately shift the focus of their attention to the new invention or try to copy it? Had there been warnings against attempting to perform a key element of the invention or predictions of failure? All these may shed important light and offer objective evidence on the question of whether the invention was obvious or not in real time.

In prosecuting a patent application—and particularly in patent litigation—the issue of inventive step generally requires more time and effort than any other issue. Lawyers rely heavily on expert support from the medicinal chemists or other scientists involved to build a picture of the expertise and information that the notional skilled person would have brought to the problem under consideration and how he or she would have understood the prior publications.

3. *Utility/Industrial Application*

The requirement for utility in patent law encompasses two separate conditions. First, in order to be patentable, an invention must be industrially useful, since the purpose of the patent system is to encourage practically useful inventions. Patent law does not, however, set a high standard of usefulness. Of mechanical inventions, it is said to be enough that "the wheels go round," and the same principle would apparently apply in other fields. It is in any event hard to mount an attack of inutility in this sense, at least in respect of claims relevant to the attacker, because it is virtually impossible for the objector to answer the question, "If the invention is useless, why do you want to attack the patent?"

[8]Harmon, R.L. Harmon on Patents: Black Letter Law & Commentary. Arlington, VA: BNA Books; 2007. p.488.

[9]KSR Int Co v. Teleflex Inc, 82 U.S.P.Q.2d 1385, (2007).

The second requirement can be important in the field of medicinal chemistry. Where a patent expressly promises certain advantages, the claimed invention must in fact provide those advantages or else the patent will be invalid. Obviously, the patent office examining the application is not called upon to test whether the invention does offer the promised advantages. If this is not apparent from the data in the application itself, in most cases it is a third party challenging the patent who tries to provide proof that the invention does not fulfil its promise.

A refinement of these rules is that the invention should be useful and fulfil its promise throughout the full scope of its claims, although the cases indicate that if the invention is inoperable for only a small proportion of the claimed scope that may not deny it validity.

The rule is potentially of enormous importance in patents for new chemical entities, especially for the first examples of a new class of medicinal compounds. As we have mentioned, when claims using broad generic (Markush) formulae are filed, it is likely that many or most of the compounds within the formula have not been tested and so may not in fact possess the claimed medicinal advantages. These will have been established only for a relatively small number of representative compounds. There is some debate as to whether the medicinal advantages must extend over the entire range of compounds for the patent to be valid. In principle, all the members of the claimed class should be useful (albeit not necessarily to the same extent), and here the applicant for a patent faces a dilemma. The class must be defined in general terms, since it would be simple to circumvent a list of individual compounds, but the generalization will be vulnerable to an objection of inutility.

For this reason, it is important to include more specific and narrower claims in a patent as well as broad generic claims. Claims that nest inside one another like Russian dolls increase the chances that at least some of the narrower claims will survive an attack. It is common practice to include separate claims limited to the five or six specific compounds tested—or even to the single most promising compound (see section III.D below).

The question of what degree of proof of utility is required in a patent is also highly pertinent to medicinal chemistry. Usually, only minimal evidence of utility need be present in the patent itself, enough to make the invention plausible to the skilled man. Many patents disclose only *in vitro* tests (or at most, tests in animals) and extrapolate from these to claim a broad pharmacological utility that includes therapeutic use in humans. This is considered acceptable, despite the fact that experimental models are sometimes poor at predicting response in humans. Safety and efficacy in humans need not be demonstrated in the patent as long as the claimed utility is credible.

The burden of proving the lack of utility falls on the party challenging the patent. A factual scientific dispute is likely over whether and how the purported benefit should be evaluated and tested. Since clinical trials for the purposes of litigation would be impossible on ethical grounds, decisions concerning what type of tests and what results would demonstrate success or failure may be controversial. It will generally be difficult and expensive for an infringer to challenge the utility of the invention, and even if he knocks out the broader claims, he may still infringe narrower claims to the specific product that he wishes to make.

Unlike attacks on novelty and inventive step, which can only be based on the prior art, an attack against the patent's utility can be based on what are colloquially called "post art" publications, in other words, material not in the public domain at the priority date. Except in the rare case where published experiments undermine the claimed advantages of the invention, the only way to attack utility in these circumstances is for the attacking party to carry out its own experiments for the purposes of the litigation, as discussed above. In practice, if the patent itself passes the minimal plausibility standard and the patentee can show clinical utility for one or more compounds by reference to trials carried out after the patent was filed, that will generally be the end of the matter.

4. Enablement/Sufficient Description

Because the social contract implicit in the patent system requires that the public has free access to an invention after expiry of the guaranteed period of exclusivity, a patent must disclose the claimed invention in such a way that it can be performed by skilled persons seeking to work the invention.

The test of enablement is often said to be whether undue experimentation is required to perform the invention. We would put the emphasis here on "undue." Minor tests and adjustments may have to be made to operate the invention as described, but this will not disqualify the patent. However, sufficient information must be given so that the suitably qualified reader is not left with a research project to perform in order to make the claimed product or to work the claimed process.

For example, if a novel class of medicinal compounds (or some of them) can only be made using a synthetic approach or specific reaction conditions that would not be readily apparent to a competent chemist, then these

things must be made clear in the description of the invention. Patent claims that do not satisfy this requirement will be held invalid. On the other hand, the patent need not teach expressly what those working in the field already know.

B. Exclusions from Patentability

Certain things that would be considered by laymen to be inventions are excluded from patentability for policy reasons. These exclusions vary between jurisdictions and with time. The most important exclusions for the pharmaceutical industry are the exclusions of scientific discoveries and of products of nature. However, it is the discoveries as such that are excluded, and most often the practical (and commercial) application can be patented.

A recent and well-publicized but controversial example is found in the field of biotechnology. In 2013, the US Supreme Court held that purified, naturally occurring DNA sequences that have been found to be relevant to certain diseases are not patentable because they are products of nature.[10] Prior to this decision, such sequences had been the subject of numerous patents (on the grounds that a particular gene is not found in its isolated form in nature but only as part of a long sequence of genomic DNA). However, the court held that this exclusion does not apply to synthetic forms of DNA, such as cDNA, which does not exist in nature and is of much greater commercial utility.

This exclusion would also seem to exclude patents for other naturally occurring molecules (e.g., peptides, receptors), even if previously unknown or never previously isolated. Again, however, synthetic forms are more likely to be used commercially and could probably be patented. Their medicinal use may also be patentable, as discussed below.

In contrast to the US Supreme Court, the EPO considers isolated and purified genes to be patentable, provided only that the isolated gene can be put to some useful purpose. However, another relevant exclusion exists in European law (but not in the US). European law prohibits patenting methods for treatment of the human body. At first blush, this would seem like an exclusion that seriously curtails the ability to secure patents for many of the types of inventions discussed above. In reality, however, this limitation has been construed narrowly. It is designed to prevent patents that would restrict doctors in their clinical practice. While a patent cannot be secured for the method by which a doctor carries out surgery or the way she examines and diagnoses a patient, a patent can be secured for special equipment designed to carry out these operations. More importantly for the medicinal chemist and the pharmaceutical industry, while one cannot obtain a patent that would prevent a doctor from treating a patient in a certain way, a patent can be obtained for a new drug used in that treatment. It can even be possible to secure patents for particular dosage forms and thus indirectly for dosage regimens of certain drugs.

An extreme example in the field of medicinal chemistry is the not uncommon discovery that an existing drug is clinically useful in a second and unrelated medical indication. It might seem that such a discovery would fall foul of both the US and European exclusions, but in Europe and many jurisdictions, claiming the drug "for use in the manufacture of medicaments for the treatment of" the second or subsequent indication is permitted. In the US, which allows an inventor to patent methods of treatment, such circumlocutions are unnecessary, and the preferred wording for the claim to a second, novel, medical indication for a known drug is simply "a method of treating disease X with drug Y."

Other exclusions include abstract inventions such as scientific theories, mathematical methods, methods for performing mental acts, and computer programs that do not in themselves create a physical change. These may be relevant to research or diagnostic methods, but further discussion is beyond the scope of this chapter.

Some countries, such as India, have imposed additional exclusions aimed at limiting the scope of pharmaceutical patents. For example, Section III(d) of the Indian Patents Act declares that "the mere discovery of a new form of a known substance which does not result in the enhancement of the known efficacy of that substance" is not considered an invention; this would seem to exclude the patenting of a new form that is useful industrially (for example in formulation) but makes no clinical difference once administered. Neither is "the mere discovery of any new property or new use for a known substance" patentable in India.

[10] *Association for Molecular Pathology v. Myriad Genetics, Inc*, 689 F 3d 1303, Section 3: "The following are not inventions within the meaning of this Act . . .(d) the mere discovery of a new form of a known substance which does not result in the enhancement of the known efficacy of that substance or the mere discovery of any new property or new use for a known substance or of the mere use of a known process, machine or apparatus unless such known process results in a new product or employs at least one new reactant."

C. The Patent Specification

The specification of a patent application or a granted patent includes:

a. a description of the invention;
b. the claims (discussed in the next section);
c. any drawings necessary to make the description comprehensible or referred to in the claims; and
d. an abstract (usually).

The published specification will generally also include a bibliography, comprising a list of the prior art to be considered by the relevant patent office.

The description and any figures are where the inventor describes the invention and its advantages over the state of the art and provides instructions on how to make and use the patented product, process, or method. It is here that the patentee must provide sufficient description, so that the reader is enabled to perform the invention. Applications for biological inventions that cannot be fully described in words must be accompanied by a deposit of the relevant material (usually in a culture of modified microorganisms) with an authorized repository.

The description frequently includes working examples of the invention, perhaps with test results, which show how to carry out the invention. These can also be used to demonstrate its advantageous properties.

D. Patent Claims

The claims are probably the most important part of the patent specification. They are where the inventor defines, in words of his own choice, his invention and the scope of the monopoly he is seeking. It is where the inventor draws an imaginary fence around what he claims to have invented and stakes out the area he believes should be reserved for his exclusive use until the patent expires. It is the invention as defined by the claims that must satisfy the conditions of novelty, inventive step, and utility, and that must be enabled by the description.

Obviously, the patentee would like to secure the broadest claims possible to prevent others from circumventing the patent by making minor changes that would still provide the benefits of the invention but might escape the wording of a narrow claim. On the other hand, if the claims are cast too broadly, they may cover the prior art, subject matter that is obvious over the prior art, or unworkable ("inutile") subject matter, so that the patent would be vulnerable to attack.

To meet these problems, applicants generally draft patent claims as one or more sets of nesting claims, where each claim is narrower than the one before. The idea is that if one or more of the broad claims turns out to be invalid, the narrower claims will survive.

Thus, in the case of a patent for a new chemical entity, there may be:

a. a set of claims covering generically defined classes of compounds, with each successive claim defining a narrower class than the one before;
b. further generic claims limited to those classes of substituted molecules that have (or are expected to have) higher activity, greater selectivity, or other advantageous properties, according to the experimental data available;
c. much narrower claims, limited to those compounds described in examples, which have actually been synthesized and tested by the inventor;
d. claims listing preferred compounds; and
e. the narrowest claim or claims (and strongest against attack), limited to the lead compound or compounds.

E. First Filing, Priority Date, and International Arrangements

As we noted at the beginning of this chapter, global patents do not exist. Patents are territorial, and separate patent applications have to be filed and secured in each territory where protection is sought. However, international conventions have been put in place in order to reduce the financial and administrative burden of filing many concurrent patent applications throughout the world, and a few regional patent offices have been established serving several countries, most notably the EPO.

The initial application will generally be filed in the local national patent office. This first filing date becomes the international priority date for the invention as described and defined in the application. The priority date is important for several reasons discussed above, so that the first application in respect of a commercially

significant invention should be filed as soon as this is justified by the available data. If necessary, the initial application can be supplemented with additional experimental information by filing further applications within twelve months of the priority date, although the priority date of additional subject matter will be the filing date of the application in which it is first described. The US offers a cheaper approach of filing-a provisional patent application, with less stringent requirements and much lower fees. A provisional application requires an enabling description but it may be filed without claims and it establishes the priority date of the invention without starting the clock running on the life of the patent, which commences only when the full patent application is filed. A provisional application is not examined or published and will automatically expire 12 months after it is filed, unless a corresponding full application is filed within that period.

Under the Paris Convention of 1883, which has been ratified by almost every country in the world, the initial filing in one country secures a global priority date. This will then be the relevant date for determining the validity and entitlement of the corresponding patents in each of the other signatory countries, provided the inventor files his foreign applications within twelve months of the priority date.

The Patent Cooperation Treaty (PCT) comes into play at the end of this one-year priority period, so that the applicant does not need to file patents in each individual territory. Instead she can file a single PCT application, which is administered by the World Intellectual Property Organization (WIPO) acting through national or regional patent offices to conduct preliminary formal processing and a search for relevant prior art.

The PCT application effectively allows the applicant to postpone deciding the specific countries in which to file her patent for an additional eighteen months, following which the application is split into separate national and regional patent applications to be examined by the relevant patent offices.

The combined effect of the Paris Convention and the PCT is to postpone commitment to the considerable costs and administrative burden of national filing by about thirty months from the priority date, so that the inventor has some time to decide the value of the invention and how much to spend protecting it.

Some regional patent offices, most importantly the European Patent Office (EPO), examine and grant patents centrally for members of the European Patent Organization (currently comprising the twenty-eight member states of the EU plus ten other countries) and for a few extension states. This significantly reduces the cost and complexity of patent prosecution in Europe, but a European Patent Application still results in a bundle of separate national patents that need to be enforced separately in each member state.

In an attempt to address this enforcement problem, most (but not all) the member states of the EU are now working on a pan-European patent (the Unitary Patent), and agreement has been reached for a Unified Patent Court that would enable patentees to enforce their Unitary Patent in a single court with jurisdiction extending throughout the signatory countries.

F. Patent Prosecution

After dealing with preliminary formal matters, the first substantive task of the patent office examiner is to search for relevant prior art. Nowadays, this usually involves a computer search of the relevant databases—including patent databases—that might uncover similar inventions.

In the US and some other countries, the patentee must assist the patent office in its examination by disclosing all relevant prior art of which she is aware. Failure to do so may be grounds for invalidating the patent later with no possibility of remedying the situation. The applicant is not obliged to search for prior art, but if she should do so, she may not withhold any relevant results.

Based on the search results, the examiner forms a view of the novelty and inventiveness of the claimed invention, and considers whether the description is sufficient and meets all formal requirements. There generally follows an exchange of correspondence in which the examiner raises objections to the application and the applicant responds to these objections, frequently offering amendments designed to meet the objections. If the examiner persists in the objections, the deadlock may be broken by requesting an interview (frequently held by conference call). It is often easier to explain the invention to the examiner orally and then come to some understanding of what claims should be granted. If the examiner finally refuses a patent application, the applicant generally has recourse to a system for appeal in appropriate cases.

The process of back-and-forth interaction or negotiation with a patent office is called "patent prosecution." This is not to be confused with litigation before the courts. The documents generated in the course of patent prosecution are called the "file wrapper." These documents, which are publicly available (often online), can be helpful in understanding the weaknesses and strengths of a given patent. Importantly, even after a patent is granted, a patentee may be bound by admissions or concessions he is forced to make during patent prosecution in order to get the patent issued.

In the US, the EPO, and some other countries, it is possible to file additional later applications called divisionals—and, in the US, continuations and continuations in part—which are based on the original disclosure in the patent application and enjoy the original priority date. These later applications are designed to enable the patentee to modify and change the scope of his claims, provided the original specification contains support for the new claims. If the patentee finds that the patent examiner is willing to grant some of claims but not others, it is a useful way for securing early patent protection for the undisputed claims, which courts can then enforce, while maintaining the discussion with the patent examiner over the claims the examiner is initially unwilling to grant.

G. Challenges to Patent Validity

Since no search of the prior art, however thorough, can definitively unearth all relevant prior art publications to enable the examiner to determine unequivocally whether the invention is worthy of a patent, the grant of a patent is not conclusive of patentability. Examiners cannot have the depth of knowledge and experience of expert practitioners in the field of every patent application put before them and because even patent examiners can make mistakes.

Throughout the patent term, the validity or the ownership of the patent can be challenged. This can be done through various procedures that vary widely from country to country. Challenges are made either in the relevant patent office or in court. Lack of validity is also a defense or counterclaim available to a defendant accused of infringing the patent, as is a determination that the plaintiff is not the rightful owner of the patent.

In most countries, third parties may lodge objections to the grant of the patent after a patent application is accepted. This procedure, known as an "Opposition" or—in the US—as "Post Grant Review," results in quasi-judicial proceedings over patent validity that take place in the patent office. An opponent, who is usually a competitor of the patentee, often has more time and resources than the examiners, as well as a greater specialized knowledge of the field concerned. Opponents will frequently uncover and cite new prior art of which the examiner was unaware.

H. The Period of Protection

A patent remains in force for twenty years from the filing date of the application on which it was based (as long as the appropriate maintenance fees are paid). As stated above, an application relying on the priority of an earlier application must be filed within twelve months of the priority date, so that the patent term is limited to twenty-one years from the priority date.

The most valuable months and years of a successful pharmaceutical product are usually the last months and years of patent life. Nonetheless, the potential benefits of an early priority date generally outweigh the cost of reducing the patent term. If certain conditions are met, the disadvantage can be mitigated by an extension of protection by up to 5.5 years (as explained in the following section) and by other forms of exclusivity (as discussed in section VI.B below).

I. Patent Extension/Restoration and Supplementary Protection Certificates

Pharmaceuticals are highly regulated. In the case of pharmaceutical inventions, a substantial part of the twenty-year patent term (typically twelve to fifteen years) can be taken up with preclinical and clinical trials and by regulatory review. All of this takes place before the product is approved for marketing and can begin to generate a return on the huge research and development costs that have been incurred. This problem has been recognized, and to some extent mitigated, in the provision of patent extensions (known as Supplementary Protection Certificates (SPCs) in Europe) and the grant of data and marketing exclusivity. Since patent extensions prolong the years of peak profitability, when a pharmaceutical company achieves the greatest return on its sunk costs, they tend to be of great economic importance.

In order to benefit from a patent extension, a product must have undergone regulatory review before it was commercially marketed. In the US, half of the period of time during which a product was being tested and all the time during which it was under regulatory review at the Food and Drug Administration (FDA) can be restored to a patent, up to a maximum of 5 years. This is on the condition that the total patent life for the product—including the patent extension—cannot exceed fourteen years from the date the product was first approved.

The only patents that can be extended are those that claim a drug product, a method of using the product, or a method of manufacturing the product. The FDA determines a product's eligibility for patent term restoration

and provides information to the US Patent and Trademark Office (PTO) regarding a product's testing and regulatory review periods. The PTO then determines the period of patent extension. The application for an extension must be submitted within sixty days of product approval.

In many respects, the SPC scheme operated by the EU is quite similar in form and function. Although important differences exist, these are outside the scope of this chapter. Suffice it to say that in the EU, an SPC may be available to extend the term of a patent for an active ingredient (or a combination) by a maximum of five years (or 5.5 years for some pediatric products). Exclusivity may be extended to a maximum of 15 (or 15.5) years.

J. Patent Infringement and Patent Enforcement

Generally, anyone who uses a patented process or makes, imports, sells, offers, or uses a patented product—including a product made by a patented process)—for commercial purposes will infringe the patent, though the basic laws of infringement vary from jurisdiction to jurisdiction.

An infringing product or process is one that falls within the scope of the monopoly defined by the claims of a valid patent. Naturally, the parties to litigation frequently disagree over the proper interpretation of the claims and the proper scope of the patent. The approach of the courts to claim interpretation (claim construction) varies greatly between countries. For example, the US courts take a very formal and generally narrow linguistic approach to determining the scope of the claims, and it is becoming increasingly tough to prove infringement. In other countries, including most of Europe, the courts take a more liberal approach to claim language, which involves trying to understand the real inventive contribution embraced by the patent.

It may be that the accused infringer has modified his product so that, at least linguistically, what he has done does not fall within the claims, but it is generally accepted that patent claims should be read to cover subject matter beyond their narrow linguistic scope. Complex and often imprecise rules govern the scope of protection afforded in such cases of nonliteral infringement and, again, the rules vary greatly from country to country. In the US, the doctrine of equivalents is intended to prevent an infringer simply replacing an element of the invention with an obviously equivalent substitute that performs substantially the same function in substantially the same way to reach substantially the same result. The rule is somewhat formalistic, however, so that proving nonliteral infringement in the US can be difficult. In Europe, giving the language of the claims a meaning that seeks to encompass the inventive contribution of the patent produces a more flexible—and correspondingly less predictable—result.

A patent is a purely negative right. It only gives the patentee the right to prevent others from using the claimed invention and does not give the patentee any right to use the invention. Thus, a patentee may find that he cannot use his patented invention because doing so would infringe someone else's patent. To give an example, a company developing an improved formulation for a compound covered by a third-party patent cannot make the formulation until the patent covering the compound expires. Similarly, the owner of the patent covering the compound cannot use the newly patented formulation for the compound without securing the consent of the later patentee.

There are exemptions from infringement, one of which is particularly important in the pharmaceutical industry. This is the experimental use exception. Patent law does not prohibit the making and using of a patented invention if the aim is to carry out research related to that invention. Thus, in the case of a patent for a new chemical entity, the exemption would allow research into, say, a new medicinal use of the compound. As long as such use is truly experimental and is not a guise for commercial exploitation, then such use is permitted.

Some countries have extended this exemption to allow a generic pharmaceutical company to take steps that are on their face commercial rather than experimental, even while the patent is in force, in order to secure regulatory approval. By allowing these activities, the public is assured that cheap generic drugs will become available promptly after the patent expires. Many patent systems offer limited defences to someone who practiced an invention in secret before the date of a third-party patent filing.

Proving the facts required to show infringement is not always simple. While it is usually not difficult to protect specific compounds (the name of the compound will be printed on the packaging), proving that a medicinal product embodies inventions such as synthetic processes or crystalline forms may involve complex forensic analysis that requires chemists with special talents. For example, proving the route of synthesis of the active ingredient may involve detecting minute traces of distinctive intermediate compounds and debating the probabilities of chance contamination. In the case of formulation patents, it may be necessary to determine by chemical and

physical analyses the ingredients or structure of the formulation or, in the case of patents for crystal structures, the crystal structure of the active ingredient within a formulation.

A defendant accused of infringing a patent can defend himself not only by arguing that his product or process does not infringe the patent, but also by challenging the validity of the patent. In fact, defendants almost invariably do this. The procedural options available depend on the country in question.

One peculiarity of US law is that patent infringement cases involving a claim for damages are heard by a jury, and it is not uncommon for juries to have to decide on both infringement and validity, which may involve complex scientific issues. This is less frequent in pharmaceutical cases because of the Hatch–Waxman provisions, special legislation put in place to ensure that generic companies challenge the validity or infringement of the patent during the process of securing regulatory approval, well before the generic product is due to be launched and damages to the patentee accrue. These cases are heard by a federal judge instead of a jury.

Other countries—such as the UK and Germany—employ specialist patent courts and judges to hear patent cases. In the UK, the same judge will rule on both infringement and validity, whereas in Germany, infringement and validity proceedings are conducted in separate courts. (In our experience, the separation of issues can lead to unsatisfactory results.) As mentioned above, the EU is in the process of establishing a centralized court that will hear patent infringement cases based on unitary patents.

The inherent complexities of patent litigation mean that it may take some time for a court to reach a conclusion. If certain conditions are satisfied, most countries provide the patentee with the ability to secure interim measures (injunctions) to stop infringement until final judgment is reached. If the patentee then loses the case, she will need to compensate the defendant for losses incurred as a result of the unjustified injunction.

Because of the complexity and consequent cost and uncertainty of patent litigation and the possible effects on the market, some countries (e.g., the UK) prohibit patentees from threatening patent proceedings against anyone other than the original manufacturer or an importer.

K. Employees' Inventions

The general rule is that inventions made as a result of and in the course of employment belong to the employer. The rule is eminently reasonable as applied to medicinal chemists and other scientists who are effectively employed to make inventions. In many countries, employees are obligated to inform their employer of any inventions they made.

Nonetheless, in some countries, the law makes provision to reward inventors personally for inventions that turn out to be of special commercial value to their employers. In Germany, inventor employees are entitled as of right to such a reward. In the UK, this will only happen in exceptional circumstances, when the invention is of outstanding benefit to the employer. In the US, no such right exists unless the employer and employee contractually stipulated to such an arrangement.

Some academic research institutes have schemes to share royalty income with the scientists working at the institute who contribute to patents that are successfully licensed to industry. In Israel for example, the royalty share can be quite substantial (in some cases up to 40 percent of the income received by the university), leading many scientists to concentrate on research that may yield income to their university.

IV. THE ROLE OF THE MEDICINAL CHEMIST IN THE PATENT ARENA

A medicinal chemist working in an established pharmaceutical company will certainly be working in a controlled environment, with work and documentation policies and practices designed to optimize the protection of intellectual property, in cooperation with the patent department. This will inevitably leave considerable scope for personal judgment and initiative.

A. When to Notify the Patent Department

Since an early priority date can bring important advantages (as discussed in section III.E above), it is in our view better to err on the side of providing too much rather than too little information to the patent department and to give notice of anything that might be an invention.

The creative medicinal chemist should be very wary of discounting as obvious any surprising or encouraging new result. Frequently, the medicinal chemist who made the invention sees the steps taken in the research as having flowed logically from what was known toward the target. In that sense, the invention may seem obvious to the inventor. An inventor, by definition, brings inventive skills to the task, but invention should be measured by the eyes of the unimaginative person skilled in the art, to whom the invention may not have been at all obvious.

One of the problems faced by large organizations is ensuring that management is made aware of potentially valuable new technology. Some companies have procedures in place involving formal invention disclosures. Along with invention committees and reward programs, these are designed to encourage scientists to make invention disclosures to their superiors.

B. Exploring the Breadth of an Invention

It is important to establish as clearly as possible the boundaries of an invention, in order to obtain broad and robust patent protection. In order to protect a medicinal compound from competition by a similar drug, patent claims must extend as broadly as possible to cover the class of compounds that embody the invention. Needless to say, it would not be possible to test or even synthesize all of them, but the data generated in a well-designed screening program and an exploration of structure/activity relationships may make it possible to reach tentative conclusions about what is likely to work and what is not, which may be enough to justify the broad claims. Negative as well as positive results should be discussed with the patent attorney preparing the application and prosecuting it through the patent office.

As also discussed above, the patent application must describe how to put the invention into practice. This condition, too, should be satisfied across the full breadth of the claims. If any compound within the claimed class requires nonstandard synthetic strategies or processing conditions, these should be specified. Similarly, where the invention is a process, it is important to investigate the effective ranges for the novel processing conditions, as well as the extent of the class of products that can benefit from the use of the inventive methods.

C. Documenting the Invention

The work protocols of an established pharmaceutical research department will include detailed regulations for keeping laboratory notebooks and/or computerized records. We wish to emphasize that in our long experience of litigating pharmaceutical patents, well-kept laboratory notebooks and records have on many occasions been invaluable, whether in rebutting unfounded attacks on patent validity or in exposing unjustified monopolies.

Records of invention or disclosure statements prepared for the patent department or for other administrative purposes are also valuable sources of information, especially if cross-referenced thoroughly to the laboratory notebooks. Written after the event, however, they may provide an *ex post facto* rationalization, that can give a false impression. Such documents (like the laboratory notebooks) may well have to be disclosed to the opponent in any subsequent litigation and the inventor should be wary of making an invention look obvious by mis- or over-interpreting earlier work (published or unpublished) with the benefit of hindsight, making buried and apparently irrelevant data look like a clear signpost toward the invention.

What we say in relation to disclosure statements naturally applies with even more force to scientific publications and any other documents put out in the public domain. Another potential danger is that the brief introduction to the new work described in the paper may inadvertently belittle prior work (perhaps in a different department of which the author knows little) that represented a substantial leap forward at the time and is the basis of a broad and fundamental patent supporting the whole project. For example, a paper relating to industrial scale synthesis may give a spurious rationale for a drug-discovery program, written in hindsight and dismissing in a sentence research that took years of inspired experimentation. Such summaries should be written with care or avoided altogether.

Of course, information relating to any invention should not be made available to the public before a patent application has been filed in respect of that information (or until a decision is made not to seek patent protection). For this reason, the patent department should generally vet proposed publications by a chemist employed in the pharmaceutical industry.

D. Drafting and Prosecuting the Patent Application

The chemist's role in securing effective patent protection is by no means completed by handing over the relevant laboratory notebooks and a record of invention. Bearing in mind everything that we have said above, the inventor can assist in many ways, such as in generating a claim set that is a reasonable generalization of the invention, supported so far as possible by experimental results. She should point out any surprises or peculiarities in the results achieved or in the methods used, to ensure that the invention is properly described. The inventor should also disclose to the patent department any relevant prior publications of which she is aware, as well as unpublished information derived from others in the company or from third parties that assisted or inspired the research leading to the invention.

During the prosecution of the application, the patent attorney will receive prior art search results and objections to patentability from patent offices around the world. The inventor can assist the patent attorney to understand how the prior art would have been viewed by those working in the field and in meeting the objections by refuting misconceived arguments, by correcting misunderstandings of the invention and of the prior art, and by helping to formulate amendments to the patent application.

It must be emphasized that the inventor should not be tempted to withhold relevant prior art just because it seems to weaken his claim to a patent. Not only is this against the law in some jurisdictions (as discussed above), but doing so will actually handicap the patent attorney (who may have been able to draft the claims to avoid the prior art, if it had been disclosed) and weaken the resulting patent, which will be susceptible to attack by any third party who subsequently uncovers that prior art.

E. Opposition and Court Proceedings

A patent may be revoked in opposition proceedings in the relevant patent office (either before or after a formal grant, depending on the jurisdiction) or in court. Such opposition is brought by competitors who want to see the patent refused or revoked as invalid on the grounds that it does not meet the requirements for patentability discussed in section III.A above.

Once granted, the mere existence of a patent can operate effectively as a "please keep off the grass" notice, but if it relates to a commercially important pharmaceutical product, legal proceedings may have to be initiated by the patent owner to enforce the monopoly by obtaining an injunction restraining infringement and by claiming damages for past infringement.

A medicinal chemist could be involved in any of these proceedings, both in supporting the legal team behind the scenes and in giving evidence, either as a witness of fact (e.g., the inventor may have to relate the history of the invention) or in giving an expert opinion (e.g., on issues of obviousness or on experiments proving infringement).

In many countries, litigation involves the disclosure (called "discovery") of all relevant documents in the possession or control of the parties, including laboratory notebooks, files of publications and other papers held by the inventor, and the like.

V. PATENTS AS A SOURCE OF SCIENTIFIC INFORMATION

The social contract embodied in the patent system requires medicinal chemists to disclose their inventions and how to make them work, in order to secure a patent monopoly. We have discussed above the importance of this knowledge to competing pharmaceutical companies. Patents also contain information that may be useful scientifically. The medicinal chemist may benefit from the information published by other inventors working in the same field of chemistry or by using techniques borrowed or adapted from another field.

While the claims of valid patents should be respected, not all of the information in a patent forms part of the monopoly. For example, some of the information may help researchers avoid wasting time exploring blind alleys. Of course, the entire contents of expired patents are in the public domain. That availability is a major aim of the patent system.

It should be born in mind when designing a literature search that patents may in some cases be more useful than journal articles, because the requirements for disclosure in patent applications may mean that they contain data—particularly negative data—that the authors would not consider worth publishing in a scientific journal.

VI. OTHER FORMS OF PROTECTION

A. Confidentiality

A medicinal chemist employed to conduct research will naturally be bound to keep secret all of the employer's confidential information. Even after a patent application has been filed, the information it contains will not be available to the public until publication some eighteen months after the priority date. The fact and contents of the application remain confidential during that time. After publication, the chemist will likely be in possession of trade secrets that are not included in the specification, and whose confidentiality should be respected indefinitely. If a competitor obtains confidential information illicitly, its use may be restrained by injunction and damages claimed.

B. Marketing Exclusivity and Data Exclusivity

Unlike the Patent Extension/Restoration and SPC systems discussed in section III.I above, marketing and data exclusivity are not dependent upon the existence of a patent but only on regulatory approval of the drug.

Whereas the innovator of a drug must spend hundreds of millions of dollars on preclinical experimentation and clinical trials, a generic manufacturer who proves that its drug is sufficiently pure and bioequivalent to the originator's product can obtain regulatory approval based on the information gathered at great expense by the innovating company. In the US, this is called an Abbreviated New Drug Application (ANDA).

The governments of the EU, the US, Japan and certain other countries recognized that—irrespective of patent protection—the originator should be given a period of exclusivity during which the health authority should not be allowed to rely on the originator's preclinical and clinical data on in order to register a competing product. *De facto* exclusivity is assured because the regulatory authority either refuses to accept an application by a generic competitor or refuses to grant final approval to the generic product, thereby keeping it off the market.

The EU offers pharmaceutical companies a total of eleven years of protection for a new medicinal product. This comprises eight years of data exclusivity, during which a competitor cannot submit a generic application that relies on data already provided to the regulator by the originator. This is followed by two years of marketing exclusivity, which may be further extended by one year.

During the two-year period of marketing exclusivity, it is possible to submit a generic application, and this application can be processed by the regulator. However, the application cannot be approved by the regulator until the two-year exclusivity has expired.

In the US, several types of marketing exclusivity exist. The details of these (including the length and extent of protection) depend on factors such as the regulatory status of the active ingredient, the prevalence of the indication for which they were approved (see orphan drug exclusivity, discussed in section VI.B.4 below), and whether they had been studied in children (see pediatric exclusivity, discussed in in section VI.B.3 below).

1. *New Chemical Entity (NCE) Exclusivity*

A drug product containing a NCE (i.e., an active moiety never before approved by the FDA) is entitled to five years of exclusivity from its date of approval.

2. *Clinical Trials Exclusivity*

Sometimes manufacturers conduct additional clinical trials in order to modify an existing product (e.g., adding a new indication or new dosage form). In such cases, the FDA may not approve an application in respect of the same modification based on that manufacturer's clinical data for a period of three years following regulatory approval (contrast this with the five-year NCE exclusivity described above). This limited exclusivity protects only the modification but no other forms or uses of the product. Bioequivalence studies are not covered by this practice.

3. *Pediatric Exclusivity*

Children are not simply miniature adults. They react differently to drugs. In the past, so few clinical trials were carried out on children that it was rarely possible to assess the efficacy and safety of a drug in that population. Accordingly, the EU and US have offered drug manufacturers an incentive to conduct clinical research into the way children react to drugs and thus to broaden the therapeutic options for young patients.

In the EU, conducting clinical trials in children prolongs the SPC (i.e., extends the life of the patent) for an additional six months. In the US, six months of additional exclusivity is awarded to a company that, at the request of the FDA, conducts clinical trials in children.

4. Orphan Drug Exclusivity

Many diseases are comparatively rare. Paradoxically, the sheer number of rare diseases means that a sizeable proportion of the population either suffers from or is at risk of developing one or another rare disease. Nonetheless, the very limited number of patients affected by each individual disease provides little incentive to develop drugs for these populations, as such drugs would not be economically viable. The fact that so few drugs were being developed for these indications led to them being termed "orphan diseases." In order to provide an incentive for drug companies to develop effective treatments for these diseases, the US and the EU provided a period of exclusivity for drugs for orphan diseases—so-called "orphan drugs"—irrespective of patentability.

In the US, an orphan drug is defined as a drug intended to treat a disease that effects fewer than 200,000 Americans, or one for which US sales are not expected to recoup the costs of development. Orphan drug exclusivity is granted for seven years and postpones or limits both generic and innovator competition.

In the EU, orphan drugs are those that have substantial benefits in serious, previously untreatable, conditions that afflict no more than five people in 10,000 in the EU, or drugs that would not otherwise be developed for lack of economic incentives. Depending on the circumstances, the period of protection lasts from six to twelve years.

C. Trade Marks

In countries where prescription-only drugs are funded by national or other insurers, a trademark is unlikely to help maintain the originator's market beyond the life of the patent, as the funding entity will almost certainly insist on generic substitution. However, if the drug is off patent and is accepted for over-the-counter sales—as in the case of Voltaren (diclofenac sodium), Zovirax (acyclovir), and aspirin (acetylsalicylic acid)—the consumer can choose which brand to buy. The trademark may then have substantial commercial value. A very successful and notorious trademark, such as Viagra (sildenafil citrate), will be valuable even if the drug can only be obtained by prescription.

VII. CONCLUSION

It will doubtless be appreciated that the future of pharmaceutical research depends on—among other things—an effective patent system. Unfortunately, in recent years the patent system in many countries has come under attack, and legislators and courts are seeking to limit the scope of patent protection, making it more difficult to secure patents and enforce them. Some of the criticism voiced against the patent system stems from abuse of the patent system, but some criticism is more populist and is sometimes based on a misunderstanding of the patent system and of science. Surely measures must be taken to prevent abuse, but in the absence of a better system to stimulate research, it is important that an effective patent system be maintained.

We hope that those reading this chapter will appreciate that medicinal chemists should possess at least a rudimentary understanding of the basics of patent law. We have tried to provide a sufficiently broad overview without unduly burdening the reader with detail. Because we sought to cover and explain in simple terms a complex area of law, which differs in different countries, we have at times had to sacrifice the linguistic accuracy expected of lawyers in their daily work. For this we apologize. Naturally, this chapter is not to be relied upon nor substituted for proper legal advice. We thank Bruce Genderson of Williams & Connolly, Washington DC, for reading and making helpful comments.

Index

Note: Page numbers followed by "*f*," "*t*," and "*b*" refer to figures, tables, and boxes, respectively.

B

C

E

F

G

H

I

O

P