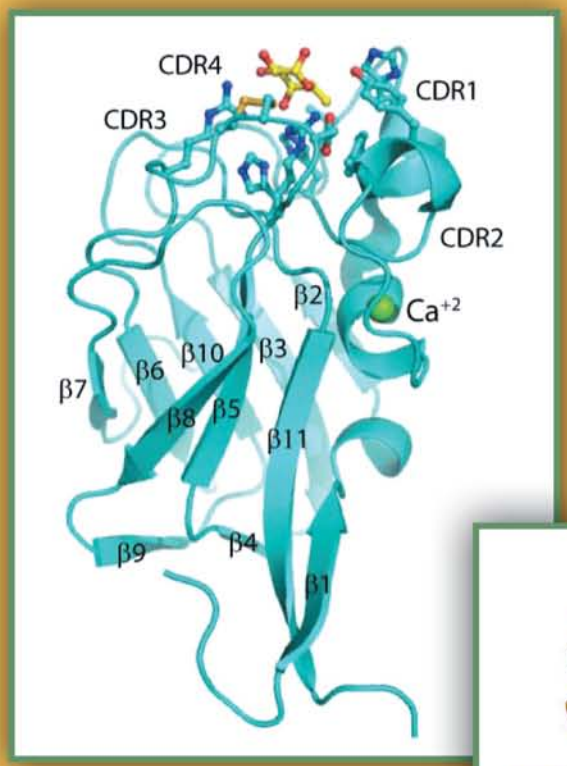
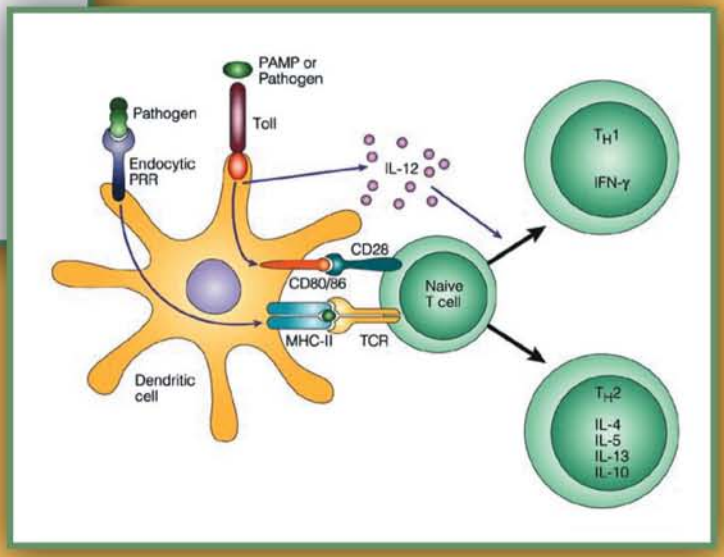
ANIMAL LECTINS

A Functional View





Edited by

Gerardo R. Vasta
Hafiz Ahmed

CRC Press
Taylor & Francis Group

ANIMAL LECTINS

A Functional View

ANIMAL LECTINS

A Functional View

Edited by
Gerardo R. Vasta
Hafiz Ahmed

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-0-8493-7269-8 (Hardcover)

Library of Congress Cataloging-in-Publication Data

Animal lectins : a functional view / editors, Gerardo R. Vasta and Hafiz Ahmed.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-8493-7269-8 (hardcover : alk. paper)
1. Lectins. I. Vasta, Gerardo R. II. Ahmed, Hafiz, 1959-
[DNLM: 1. Lectins. QW 640 A598 2009]

QP552.L42A55 2009
572'.69--dc22 2008017811

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Dedication

Dedicated to our families and friends

Contents

PART III Glycoprotein Folding, Sorting and Secretion, Targeting, Degradation, and Clearance

PART IV Cell Adhesion and Cell Surface Lattice Formation

PART V Cell–Cell Interactions, Signaling, and Transport

PART VI Recognition and Effector Functions in Innate Immunity

Foreword

It is now 120 years since the first report on a lectin appeared and more than 60 years since the term "lectin" was coined. For a long time, attention was focused almost exclusively on plant lectins, and it was intensified by the demonstration in the 1960s that these sugar-binding proteins are invaluable tools for the study of carbohydrates in solution and on cell surfaces. Broad interest in animal lectins began only two decades ago and is currently growing at an explosive rate, fueled by the realization that they act as cell recognition molecules, a feature originally demonstrated for bacterial surface lectins. A key reason for the accelerated progress in this field is the application of recombinant DNA techniques, which have also permitted the construction of animals deficient in particular lectins. Another factor is the vast improvement in the methods of structural analysis of the carbohydrates of glycoconjugates. Last but not least is the availability of glycoarrays with hundreds of different saccharides, which greatly facilitates the detection of lectins and identification of lectin specificity. By now, numerous animal lectins, classified into nearly 20 families, have been isolated and characterized, and the functions of a number of these lectins have been clarified. Among others, these lectins have been shown to control the biosynthesis of glycoproteins and their targeting to subcellular organelles, to participate in cell signaling, monitor the migration of leukocytes in blood vessels, serve as innate immunity agents against microbial pathogens, and contribute to tumor progression.

The National Institutes of Health recognized the importance of animal lectins and in 2001 awarded a megagrant of $75 million for a 10-year period to the Consortium for Functional Glycomics, the aim of which was to study the role of carbohydrate–lectin interactions at the cell surface and in cell–cell communication. At present, over 350 glycobiology research groups are affiliated to the consortium.

The knowledge accrued about animal lectins is certain to have a considerable impact on medical practice, by improving the pharmacokinetics of glycoprotein drugs and blocking the undesirable migration of leukocytes to sites of inflammation, as well as for disease diagnosis, in the treatment of cancer, and in lectin replacement therapy. There is no doubt that in the future the list of animal lectins will continue to grow and that novel functions and applications will be found for these fascinating proteins.

The editors should be commended for assembling a book that contains authoritative reviews on various aspects of the present state of knowledge of animal lectins. In such a rapidly expanding field, no review is likely to be exhaustive or up-to-date in a few years time. However, I am confident the book will be helpful to many readers and will stimulate them to further work in this exciting field.

Nathan Sharon
The Weizmann Institute of Science
Rehovot, Israel

Preface

It is now well established that protein–carbohydrate interactions play significant roles in modulating cell–cell and cell–extracellular matrix (ECM) interactions, which, in turn, mediate various biological processes such as cell activation, growth regulation, cancer metastasis, and apoptosis. Thus, the identification of carbohydrate-binding proteins (lectins) and their ligands, and the detailed understanding of the molecular mechanisms and downstream effects of these protein–carbohydrate interactions are subjects of intense research at present. During the preparation of this book, a PubMed search of "animal lectins" at the National Center of Biotechnology Information Web site (www.ncbi.nlm.nih.gov/) yielded almost 73,000 entries, reflecting the exponential growth of this area of research.

Why a "functional view of animal lectins" now? It is perhaps the growing interest on the rapidly expanding glycobiology field, together with the advent of novel technologies for disruption of gene expression and knockouts in genetically tractable animal models and the availability of their genomes in public databases that have generated an explosive rate of productivity in research related to the biological roles of lectins in both invertebrate and vertebrate animal models. Thus, from the contents of the chapters that follow, we hope the reader will agree with our assessment about the timing of this volume.

Part II is designed to provide the reader with an overview on the current approaches used for assessing lectin functions. These include structural and thermodynamic approaches to study lectin–ligand interactions; the use of glycan arrays for the identification of lectin ligands and the qualitative and quantitative analyses of their interactions; whole-genome analysis to characterize the molecular, structural, and evolutionary diversity of lectins; and finally, the animal models available for assessing their biological roles.

Part III through VI describe lectin functions within the cell, followed by those lectins that function at the cell surface, and finally, with those that mediate interactions between cells and self/nonself ligands related to innate and adaptive immunity. Accordingly, Part III is dedicated to lectins that participate in glycoprotein folding, sorting and secretion, targeting, degradation, and clearance. Part IV focuses on lectin-mediated cell adhesion and cell surface lattice formation. Part V deals with the lectins that are involved in cell–cell interactions, signaling, and transport. Part VI is dedicated to lectins that are involved in the recognition and effector functions in innate and adaptive immunity. Some new lectins and lectin families are also included in the last part.

This book would not have been possible without the contributions of many people. First of all, we are very grateful to all contributors for their outstanding and very exciting reviews. Unfortunately, not all color figures could be included in the color insert printed within this volume. An accompanying CD provided in the back cover of this book contains all the color illustrations, which can be expanded on a computer screen for further detailed examination. We are deeply indebted to Professor Nathan Sharon (The Weizmann Institute of Science, Rehovot, Israel) for his gracious foreword to this book. Our sincere thanks to Dr. Judith E. Spiegel (former editor, Taylor & Francis) and Barbara E. Norwitz (current editor, Taylor & Francis) for finding our proposal worthwhile for publication. We would like to thank the entire team at Taylor & Francis, especially Patricia Roberson, Joette Lynch, and Suryakala Arulprakasam at SPi for turning the manuscripts into a book. Finally, we are very grateful to our family members for their patience and support.

Gerardo R. Vasta
Hafiz Ahmed

Editors

Gerardo R. Vasta earned a PhD in biochemistry from the University of La Plata, Buenos Aires, Argentina, in 1978 and a PhD in zoology in 1980. He undertook his postdoctoral studies at the Roswell Park Cancer Institute, Buffalo, New York, and later collaborated with John J. Marchalonis at the Department of Biochemistry and Molecular Biology, Medical University of South Carolina. Dr. Vasta then became a faculty member of the department. He is currently serving as a professor at the University of Maryland Biotechnology Institute, Baltimore, Maryland.

Dr. Vasta's primary scientific interests include molecular aspects and evolution of self/nonself recognition mechanisms in innate immunity, host–parasite interactions, and developmental processes of vertebrates and invertebrates, with particular emphasis on protein–carbohydrate interactions. During the early 1980s, based on the structural relationships of lectins from invertebrates and some acute phase reactants from vertebrates, he proposed that invertebrate C-type lectins may function as nonself recognition factors that mediate pathogen recognition and phagocytosis. His research team under his guidance later identified a homologue of the mammalian mannose-binding proteins in a protochordate species. Recently, Dr. Vasta along with his research associates discovered a new lectin family with a novel sequence motif and structural fold (named F-type lectins) present in a range of organisms from prokaryotes to invertebrates and ectothermic vertebrates. His laboratory is currently investigating the roles of galectins and F-type lectins in development and innate immunity using the zebrafish as a model organism.

Dr. Vasta has received a number of national and prestigious international awards, including the J. William Fulbright Scholar International Award, U.S. Department of State; John E. Fogarty Award, National Institutes of Health; International Scholarship for Scientific Research, National Council of Scientific and Technological Research of Argentina, among others. He served as a division editor of the *Journal of Experimental Zoology*, and is currently a member of review panels at several federal funding agencies. He has authored over 130 publications, including peer-reviewed articles, invited contributions, and book chapters, and has coedited a book, *New Directions in Invertebrate Immunology,* published by SOS Publications in 1996.

Hafiz Ahmed received a PhD in biochemistry from Jadavpur University, Calcutta (now called Kolkata), India, in 1986 for his work on the purification and characterization of carbohydrate-binding proteins (lectins) from plants, invertebrates, and bacteria. As a postdoctoral fellow at the Max Planck Institute for Experimental Medicine (Gottingen, Germany), Roswell Park Cancer Institute (Buffalo, New York), and Center of Marine Biotechnology, University of Maryland Biotechnology Institute (Baltimore, Maryland), he conducted studies on animal lectins, especially on the structure and functions of galectins and C-type lectins. He has published more than 60 papers in peer-reviewed journals. He is the author of *Principles and Reactions of Protein Extraction, Purification, and Characterization,* published by CRC Press in 2004. Dr. Ahmed is the current editor-in-chief of *Organic Chemistry Insights*, published by Libertas Academica. He is currently serving as an assistant professor at the Center of Marine Biotechnology. His recent research interest is on the roles of galectins in development and cancer metastasis.

Contributors

Coen M. Adema
The Center for Evolutionary and Theoretical Immunology
Department of Biology
University of New Mexico
Albuquerque, New Mexico

Hafiz Ahmed
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

L. Mario Amzel
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University School of Medicine
Baltimore, Maryland

Vitaly Balan
Karmanos Cancer Institute
Wayne State University
Detroit, Michigan

Linda G. Baum
Department of Pathology and Laboratory Medicine
UCLA School of Medicine
Los Angeles, California

Mario A. Bianchet
Department of Biophysics and Biophysical Chemistry
Johns Hopkins University School of Medicine
Baltimore, Maryland

German Bianco
Laboratorio de Inmunopatología
Instituto de Biología y Medicina Experimental
Buenos Aires, Argentina

C. Fred Brewer
Departments of Molecular Pharmacology, and Microbiology and Immunology
Albert Einstein College of Medicine
Bronx, New York

Gordon D. Brown
Institute of Infectious Disease and Molecular Medicine
University of Cape Town
Cape Town, South Africa

Julio J. Caramelo
Laboratory of Structural Cell Biology
Fundación Instituto Leloir
Buenos Aires, Argentina

Paul R. Crocker
Division of Cell Biology and Immunology
College of Life Sciences
University of Dundee
Dundee, United Kingdom

Richard D. Cummings
Department of Biochemistry
Emory University School of Medicine
Atlanta, Georgia

Nancy M. Dahms
Department of Biochemistry
Medical College of Wisconsin
Milwaukee, Wisconsin

Tarun K. Dam
Departments of Molecular Pharmacology, and Microbiology and Immunology
Albert Einstein College of Medicine
Bronx, New York

Michael Demetriou
Department of Microbiology
and Molecular Genetics
University of California
Irvine, California

Kevin M. Dennehy
Institute of Infectious Disease
and Molecular Medicine
University of Cape Town
Cape Town, South Africa

James W. Dennis
Samuel Lunenfeld Research Institute
Mount Sinai Hospital
Department of Molecular Genetics
University of Toronto
Toronto, Ontario, Canada

Nazzareno Dimasi
Department of Antibody Discovery
and Protein Engineering
MedImmune, Inc.
Gaithersburg, Maryland

Shao-J. Du
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

Yuichi Endo
Department of Immunology
Fukushima Medical University
Fukushima, Japan

M. Espeli
Centre d'Immunologie de Marseille Luminy
CNRS, INSERM
Université de la Méditerranée
Marseille, France

Franco H. Falcone
School of Pharmacy
The University of Nottingham
Nottingham, United Kingdom

Teizo Fujita
Department of Immunology
Fukushima Medical University
Fukushima, Japan

L. Gauthier
Innate-Pharma SA
Marseille, France

Jill E. Gready
John Curtin School of Medical
Research
Australian National University
Canberra, Australia

Edward N. Harris
Department of Biochemistry and Molecular
Biology
The University of Oklahoma Health Sciences
Center
Oklahoma City, Oklahoma

Tomomitsu Hatakeyama
Department of Applied Chemistry
Nagasaki University
Nagasaki, Japan

Hans-Peter Hauri
Biozentrum
University of Basel
Basel, Switzerland

Jun Hirabayashi
Research Center for Medical Glycoscience
National Institute of Advanced Industrial
Science and Technology
Tsukuba, Ibaraki, Japan

Uffe Holmskov
Medical Biotechnology Center
University of Southern Denmark
Odense, Denmark

Nobuko Hosokawa
Department of Molecular
and Cellular Biology
Institute for Frontier Medical Sciences
Kyoto University
Kyoto, Japan

and

Core Research for Evolutional Science
and Technology
Japan Science and Technology Agency
Saitama, Japan

Daniel K. Hsu
Department of Dermatology
Davis School of Medicine
University of California
Sacramento, California

Juan M. Ilarregui
Laboratorio de Inmunopatología
Instituto de Biología y Medicina Experimental
Buenos Aires, Argentina

Shawn S. Jackson
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

Jens Chr. Jensenius
Department of Medical Microbiology and Immunology
University of Aarhus
Aarhus, Denmark

Mitsuru Jimbo
Department of Marine Bioscience
School of Fisheries Sciences
Kitasato University
Ofunato, Japan

Shaden Kamhawi
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Rockville, Maryland

Hisao Kamiya
Department of Marine Bioscience
School of Fisheries Sciences
Kitasato University
Ofunato, Japan

Michael R. Kanost
Department of Biochemistry
Kansas State University
Manhattan, Kansas

Ann M. Kerrigan
Institute of Infectious Disease and Molecular Medicine
University of Cape Town
Cape Town, South Africa

Jung-Ja P. Kim
Department of Biochemistry
Medical College of Wisconsin
Milwaukee, Wisconsin

Kazuhiko Koike
Faculty of Biological Science
Hiroshima University
Hiroshima, Japan

Yvette van Kooyk
Department of Molecular Cell Biology and Immunology
VU University Medical Centre
Amsterdam, the Netherlands

Ken S. Lau
Department of Pathology
Harvard University
Boston, Massachusetts

Jin Kyu Lee
Complex Carbohydrate Research Center
University of Georgia
Athens, Georgia

Fu-Tong Liu
Department of Dermatology
Davis School of Medicine
University of California
Sacramento, California

Eric S. Loker
The Center for Evolutionary and Theoretical Immunology
Department of Biology
University of New Mexico
Albuquerque, New Mexico

S.J.C. Mancini
Centre d'Immunologie de Marseille Luminy
CNRS, INSERM
Université de la Méditerranée
Marseille, France

Misao Matsushita
Department of Applied Biochemistry
Tokai University
Kanagawa, Japan

Niraj R. Mehta
Department of Pharmacology and Molecular Sciences
Johns Hopkins University School of Medicine
Baltimore, Maryland

Mette Møller-Kristensen
Medical Research Laboratories
Aarhus University Hospital
Aarhus, Denmark

Shaun Morroll
School of Pharmacy
The University of Nottingham
Nottingham, United Kingdom

F. Mourcin
Centre d'Immunologie de Marseille Luminy
CNRS, INSERM
Université de la Méditerranée
Marseille, France

Ivan R. Nabi
Department of Cell Biology and Physiological Sciences
University of British Columbia
Vancouver, British Columbia, Canada

Kazuhiro Nagata
Department of Molecular and Cellular Biology
Institute for Frontier Medical Sciences
Kyoto University
Kyoto, Japan

and

Core Research for Evolutional Science and Technology
Japan Science and Technology Agency
Saitama, Japan

Beat Nyfeler
Biozentrum
University of Basel
Basel, Switzerland

Eric W. Odom
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

Cornelia Oetke
Division of Cell Biology and Immunology
College of Life Sciences
University of Dundee
Dundee, United Kingdom

Linda J. Olson
Department of Biochemistry
Medical College of Wisconsin
Milwaukee, Wisconsin

Nades Palaniyar
The Hospital for Sick Children
University of Toronto
Toronto, Ontario, Canada

Armando J. Parodi
Laboratory of Glycobiology
Fundación Instituto Leloir
Buenos Aires, Argentina

Michael Pierce
Department of Biochemistry and Molecular Biology
Complex Carbohydrate Research Center
University of Georgia
Athens, Georgia

Gabriel A. Rabinovich
Laboratorio de Inmunopatología
Instituto de Biología y Medicina Experimental
Buenos Aires, Argentina

Avraham Raz
Karmanos Cancer Institute
Wayne State University
Detroit, Michigan

B. Rossi
Innate-Pharma SA
Marseille, France

Keiko Saito
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

Ryuichi Sakai
Faculty of Fisheries
Hokkaido University
Hakodate, Japan

Mariana Salatino
Laboratorio de Inmunopatología
Instituto de Biología y Medicina Experimental
Buenos Aires, Argentina

C. Schiff
Centre d'Immunologie de Marseille Luminy
CNRS, INSERM
Université de la Méditerranée
Marseille, France

Ronald L. Schnaar
Departments of Pharmacology and Neuroscience
Johns Hopkins University School of Medicine
Baltimore, Maryland

David F. Smith
Department of Biochemistry
Emory University School of Medicine
Atlanta, Georgia

Susan H. Smith
Department of Immunology
Duke University Medical Center
Durham, North Carolina

Grith L. Sorensen
Medical Biotechnology Center
University of Southern Denmark
Odense, Denmark

Barbara A. Stout
Department of Biology
The Center for Evolutionary and Theoretical Immunology
University of New Mexico
Albuquerque, New Mexico

Sean R. Stowell
Department of Biochemistry
Emory University School of Medicine
Atlanta, Georgia

Satoshi Tasumi
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

Thomas F. Tedder
Department of Immunology
Duke University Medical Center
Durham, North Carolina

Steffen Thiel
Department of Medical Microbiology and Immunology
University of Aarhus
Aarhus, Denmark

Marta A. Toscano
Laboratorio de Inmunopatología
Instituto de Biología y Medicina Experimental
Buenos Aires, Argentina

Sharon Turner
School of Pharmacy
The University of Nottingham
Nottingham, United Kingdom

Jesus G. Valenzuela
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Rockville, Maryland

Gerardo R. Vasta
Center of Marine Biotechnology
University of Maryland Biotechnology Institute
Baltimore, Maryland

Yi Wang
Karmanos Cancer Institute
Wayne State University
Detroit, Michigan

Paul H. Weigel
Department of Biochemistry and Molecular Biology
The University of Oklahoma Health Sciences Center
Oklahoma City, Oklahoma

Markus W. Wendeler
Biozentrum
University of Basel
Basel, Switzerland

Xiao-Qiang Yu
Division of Cell Biology
and Biophysics
School of Biological Sciences
University of Missouri, Kansas City
Kansas City, Missouri

Alex N. Zelensky
Department of Genetics
Erasmus Medical Center
Rotterdam, the Netherlands

Si-Ming Zhang
Department of Biology
The Center for Evolutionary and Theoretical
Immunology
University of New Mexico
Albuquerque, New Mexico

Part I

Introduction

1 Introduction to Animal Lectins

Gerardo R. Vasta and Hafiz Ahmed

CONTENTS

1.1 EARLY RESEARCH ON LECTINS: HISTORIC PERSPECTIVES

Lectins are generally defined as proteins or glycoproteins (other than antibodies and enzymes) that bind carbohydrates [1] (see Table 1.1 for major milestones in lectin research). As carbohydrates are abundant on the surface of virtually all cells, lectins can mediate cell–cell and cell–extracellular matrix interactions. Following the discovery of the first animal lectins from the albumin gland of the snail *Helix pomatia* [2], the Charcot Leyden crystals protein (now galectin-10) [3–5], snake venom [6,7] described from the mid- to the late nineteenth century, and the agglutinins from crustaceans and horseshoe crabs [8,9] identified in the early twentieth century, the landmark discovery of the carbohydrate nature of the H blood group substances by Walter J.T. Morgan and Winifred M. Watkins at the Lister Institute in London in the early 1950s, by use of the eel agglutinin [10], was perhaps the milestone that initiated the transition from serological approaches to a period of intense biochemical research during the 1960s and 1970s. During these two decades, several animal lectins from slime molds to horseshoe crabs [11], snails [12,13] and vertebrates including mammals [14,15] were identified and biochemically characterized. Since then, the pace of lectin research increased dramatically, greatly facilitated by the introduction of affinity chromatography by Agarwal and Goldstein [16] and probably due to the realization that lectins could serve as useful tools for practical applications as blood group typing reagents [17] and the identification and characterization of carbohydrate structures on the surface of normal and malignant cells [18,19]. During the preparation of this volume, a PubMed search of "animal lectins" at National Center of Biotechnology Information Web site (www.ncbi.nlm.nih.gov/) yielded almost 73,000 entries, reflecting the exponential growth of this field.

1.2 CURRENT CLASSIFICATION OF ANIMAL LECTINS

With the availability of amino acid sequences and crystal structures of increasing number of animal lectins, it is now possible to classify them into several groups mostly based on conserved sequence motifs for sugar binding and cation coordination, and structural folds, in addition to additional properties such as their requirements for divalent cation or reducing environment, or their nominal carbohydrate specificities (see Table 1.2). In 1988, Kurt Drickamer pioneered this initiative by identifying characteristic sequence motifs in the carbohydrate-binding domains of lectins that either

TABLE 1.1
Major Milestones of Lectin Research

Year	Discovery or Breakthrough
1888	Description of first phytoagglutinin (ricin) by Peter Herman Stillmark in his doctoral thesis to the University of Dorpat (now Tartu, Estonia)
1891	Paul Ehrlich applied ricin and abrin (toxic hemagglutinin from jequirity bean, *Abrus precatorius*) as model antigens for immunological studies
1899	Description of first animal lectin from the albumin gland of snail *Helix pomatia* (M.L. Camus)
1902	Description of first bacterial agglutinin (R. Kraus)
1907	Description of nontoxic plant agglutinins (K. Landsteiner and H. Raubitschek)
1919	J.B. Sumner crystallized concanavalin A from the jack bean *Canavalia ensiformis*, thereby obtaining pure hemagglutinin for the first time
1936	First suggestion that carbohydrates on the surface of red blood cells may be the ligands for concanavalin A-induced hemagglutination
1941	First description of viral agglutinin (G.K. Hirst)
1947	Description of blood group specificity of certain hemagglutinins (W.C. Boyd and K.O. Renkonen)
1952	Discovery of the carbohydrate nature of the H blood group substance by Walter J.T. Morgan and Winifred M. Watkins by use of the eel agglutinin
1954	W.C. Boyd and E. Shapleigh introduced the term "lectin" (from the Latin *legere*, to pick out or choose)
1960	P.C. Nowell described that the lectin from the red kidney bean *Phaseolus vulgaris*, known as phytohemagglutinin, possesses the ability to stimulate lymphocytes
1963	I.J. Goldstein introduced affinity chromatography for the isolation of lectins
1972	First lectin concanavalin A, whose primary structure and 3D structure were elucidated (G.M. Edelman, K.D. Hardman, and C.F. Ainsworth)
1974	Isolation of first mammalian lectin (asialoglycoprotein receptor) from liver (G. Ashwell)
1980	I.J. Goldstein et al. defined lectin as a carbohydrate-binding protein of nonimmune origin that agglutinates cells
1988	S.H. Barondes defined lectin as a carbohydrate-binding protein other than an antibody or an enzyme

TABLE 1.2
Classification of Animal Lectins

Lectin Groups	Ca^{2+}	Specificity	Defining Features
C-type	Yes	Variable	C-type sequence motif
Galectins	No	β-Galactosides	S-type sequence motif
P-type (M-6-P R)	Variable	Mannose-6-P	P-type sequence motif
I-type	No	Variable	Ig-like domains
Pentraxins	Most	PC/Galactosides	Multimeric binding motif
Heparin-binding type	No	Heparin/heparan-SO_4^{2-}	Basic amino acid clusters
F-type	Variable	L-fucose	F-type sequence motif
Calnexin	Yes	Glc_1Man_9	Calnexin sequence motif
M-type	Yes	Man_8	M-type sequence motif
L-type	Yes	Variable	L-type sequence motif
R-type	No	Variable	R-type sequence motif
F-box	No	$GlcNAc_2$	F-box sequence motif
Ficolin	Yes	GlcNAc, GalNAc	Ficolin sequence motif
Chitinase-like (Chilectins)	No	Chito-ologosaccharides	Triose-phosphate isomerase (TIM) barrel-like structure
Intelectins (X-type)	Yes	Gal, galactofuranose, pentoses	Intelectin sequence motif

required calcium or a reducing environment for binding their ligands, leading to propose that these lectins be classified as C- and S-types, respectively, on the basis of their known amino acid sequences [20]. Since then, the identification of conserved sequence motifs in the CRDs of numerous lectin groups has resulted in the establishment of multiple lectin families, and the search for such motifs constitutes the most frequently used diagnostic tool for the family assignment of any newly identified lectin. For lectin families like the C- and the F-type, in which the CRDs can be found in combination of a great variety of apparently unrelated domains to form mosaic or chimeric molecules, it is more appropriate to describe them as domain families rather than lectin families. Some lectin sequence motifs such as that typical of the C-type lectin domain (CTLD), however, can also be found in nonlectin proteins such as the type II [21] antifreeze protein [22,23] and galectin-like domain in the hematopoietic stem cell precursor, HSPC159 [24]. Similarly, the finding of typical lectin structural fold in apparently unrelated, nonlectin proteins, in which the lectin signature motif is poorly conserved or absent, has raised interesting questions about their evolutionary history, either cooption or convergence [25–27]. Finally, lectins, for which no extensive *bona fide* families have been identified so far, constitute a significant group of proteins that remain to be investigated and likely to yield novel insight into structure–function relationships.

1.3 NEW INSIGHT INTO THE BIOLOGICAL ROLES OF LECTINS: A ROADMAP TO THE BOOK'S CONTENTS

Part II of this volume is designed to provide the reader with an overview of the current approaches used for assessing lectin functions. These include structural (Bianchet et al., Chapter 2) and thermodynamic (Dam and Brewer, Chapter 3) approaches to study lectin–ligand interactions, the use of glycan arrays for the identification of lectin ligands (Smith and Cummings, Chapter 4), and the qualitative and quantitative analyses of their interactions (Hirabayashi, Chapter 5), a whole-genome analysis to characterize the molecular, structural, and evolutionary diversity of lectins (Zelensky and Gready, Chapter 6), and finally, the animal models available for assessing their biological roles (Ahmed et al., Chapter 7). In Chapter 2, Bianchet et al. provide a general description of the structural aspects of the lectin–ligand recognition and discuss the mechanism by which lectin discriminates among different sugars. In Chapter 3, Dam and Brewer discuss general mechanisms for the increased binding avidity and specificity of multivalent lectin molecules for ligands and counterreceptors and extended site binding of a receptor for a multivalent ligand. Smith and Cummings (Chapter 4) describe a new approach (glycan microarray) for high-throughput screening of a variety of glycan-binding proteins (GBPs) to identify glycans (candidate ligands) that are recognized with the highest affinity by the GBPs. Using these microarrays, these investigators explore the functional recognition of the vast repertoire of glycans present in nature. Hirabayashi, in Chapter 5, discusses the development of the frontal affinity chromatography (FAC), which allows quantification of lectin–carbohydrate interactions. The author also describes "lectin microarray," which can be used to identify glycan-related biomarkers as well as high-throughput quality control of various cells and glycoprotein drugs. By analysis of whole-genome and other data resources, Zelensky and Gready (Chapter 6) characterize the molecular, structural, and evolutionary diversity of CTLDs and add three new groups (XV–XVII) to the already characterized 14 CTLDs (I–XIV) by Drickamer and his associates. At the end of this part, Ahmed et al. (Chapter 7) address the progression and development of structural–functional studies on lectins using a variety of animal species and illustrate the current use of animal models (mammalian and alternative nonmammalian systems) for the characterization of galectin function.

The following parts are organized in a progression that starts with lectin functions within the cell, followed by lectins that function at the cell surface, and finally, with those that mediate interactions between cells and "self"/"nonself" ligands related to innate and adaptive immunity. Accordingly, Part III is dedicated to lectins that participate in glycoprotein folding, sorting and secretion, targeting, degradation, and clearance. In Chapter 8, Caramelo and Parodi discuss how endoplasmic

reticulum (ER)-resident lectins, calreticulin (a soluble protein) and calnexin (a membrane-bound homologue of calreticulin), in conjunction with two ER enzymes, glucosidase II and UDP-Glc:glycoprotein glucosyltransferase (UGGT), participate in a quality control system responsible for protein folding. Both calreticulin and calnexin act as nonclassical chaperones and enhance folding efficiency by preventing nonspecific protein–protein interactions and favor correct disulfide bridge pairing by facilitating the action of ERp57, a member of the protein disulfide isomerase family. In most cases, UGGT is the sole element in the system in charge of sensing the conformational and assembly status of protein monomers or oligomers. Nyfeler et al. (Chapter 9) describe the molecular features of the four animal L-type lectins (ERGIC-53, VIP36, ERGL, and VIPL) and their roles in glycoprotein sorting and trafficking. Nancy Dahms and her associates (Chapter 10) describe two members of the P-type lectin family, the 46 kDa cation-dependent mannose-6-phosphate receptor (CD-MPR) and the 300 kDa cation-independent mannose-6-phosphate receptor (CI-MPR), and their abilities to direct the delivery of approximately 50 different newly synthesized soluble lysosomal enzymes bearing mannose-6-phosphate (Man-6-P) on their N-linked oligosaccharides to the lysosome. From the crystallographic studies, these investigators demonstrate that the two MPRs use related, but yet distinct, approaches in the recognition of phosphomannosyl residues. In Chapter 11, Hosokawa and Nagata describe the role of M-type lectins in the sorting of newly synthesized glycoproteins to the secretory pathway. ER-associated protein degradation (ERAD) is a mechanism to eliminate proteins that fail to adopt correct conformations in the ER after retrotranslocating the proteins to the cytosol. For the disposal of glycoproteins by ERAD, mannose is trimmed from the N-glycans by ER α-mannosidase I and the misfolded glycoproteins in the form of Man8B ($Man_8GlcNAc_2$ isomer B) is recognized by a lectin and sorted for the disposal pathway. These authors also discuss the cloning and function of mammalian ER-degradation enhancing α-mannosidase-like protein (EDEM1) and yeast homologue Htm1p/Mnl1p molecules, which enhance glycoprotein ERAD. In the last chapter (Chapter 12) of this part, Harris and Weigel explore the biological diversity of the hyaluronan and chondroitin sulfate receptors such as TSG-6, RHAMM, CD44, LYVE-1, and HARE with regard to their functional activities related to glycosaminoglycan-binding and cellular responses.

Part IV focuses on lectin-mediated cell adhesion and cell surface lattice formation. In Chapter 13, Wang et al. describe the role of galectin-3 in apoptosis, proliferation, and metastasis of tumor cells and summarize the clinical significance of galectin-3 expression in tumors. Schnaar and Mehta (Chapter 14) describe a member of the Siglec family of sialic acid–binding lectins, Siglec-4, which is exclusively expressed in the nervous system and regulates axon–myelin stability and axon regeneration possibly through interactions with nonreducing glycan terminus NeuAcα2-3Galβ1-3GalNAc of gangliosides GD1a and GT1b present in axons. Harris and Weigel (Chapter 15) describe four lecticans (such as aggrecan, neurocan, brevican, and versican) that bind hyaluronan and their possible roles in matrix stabilization, cellular differentiation, and tissue morphogenesis by keeping the cells together. Kamiya et al. (Chapter 16) describe the roles of coral lectins in symbiosis, one of the most intriguing biological phenomena in shallow areas of the tropical sea. They particularly focus on morphological changes in zooxanthellae and discuss mechanisms of interaction between corals and their algal symbionts that are mediated by protein–carbohydrate recognition.

Part V deals with the lectins that are involved in cell–cell interactions, signaling, and transport. In Chapter 17, Oetke and Crocker describe sialoadhesin and the CD33-related siglecs and their binding to sialylated pathogens, and their function in fine-tuning cells of the immune system. Kerrigan et al. (Chapter 18) focus on the signaling pathways mediating the various cellular responses induced by a pattern recognition receptor, Dectin-1, and how these responses are tailored to initiate antifungal immunity. Smith and Tedder (Chapter 19) describe the role of CD22 (another member of Siglec family) in regulation of normal B cell signal transduction and function, activation-induced cell death, and the homeostatic survival of B cells in the periphery. In Chapter 20, Espeli et al. report on the early steps of B cell development in the bone marrow by focusing on the specific interactions between early B cell precursors and the bone marrow microenvironment. They focus on the mode of

activation of the receptor expressed by early pre-B cells, called the pre-BCR, which depends on the secretion of galectin-1 by bone marrow stromal cells and on specific interactions between the pre-BCR, galectin-1, and integrin members. In the last chapter of this part (Chapter 21), Lau et al. discuss about Golgi N-glycan processing and galectin function. They demonstrate that the N-glycan multiplicity cooperates with the physical and kinetic properties of the Golgi pathway to determine relative levels of surface glycoproteins and suggest that galectins, the N-glycan multiplicity, and the Golgi N-glycan branching pathway have coevolved to balance cellular sensitivities to multiple extracellular cues controlling growth and arrest in mammals.

Part VI is dedicated to lectins that are involved in recognition and effector functions in innate and adaptive immunity. In Chapter 22, Møller-Kristensen et al. describe mannan-binding lectin (MBL) polymorphisms and infectious diseases. They consult several published studies to investigate a possible connection between MBL deficiency and susceptibility to infections by bacteria, viruses, fungi, and parasites, and conclude that MBL deficiency appears to be linked to increased susceptibility to infections. Palaniyar et al. (Chapter 23) discuss on immunoregulatory roles of two collectins, lung surfactant proteins A and D. These two collectins serve as endogenous antibiotics and mediators of inflammation and thus have emerged as effector molecules of the innate immune system. These collectins also involve in the crosstalk between innate and adaptive immunity through interactions with complement system via antibody molecules and dendritic cells (DC) via Toll-like receptors (TLRs). van Kooyk (Chapter 24) describes C-type lectin receptors (CLR) on DC, the crucial immune cells located at peripheral tissues throughout the human body. It has been shown that the recognition of pathogens by DC through CLRs and TLRs can differentiate DC in various directions, leading to specific production of cytokines that can either activate T cells or inhibit T cell function. The investigator demonstrates that both CLRs and TLRs fine-tune immune responses through the concerted action of pattern recognition receptors (PRR) occupancy by the signature of the pathogen. Dimasi (Chapter 25) describes structural and functional roles of CLRs on natural killer (NK) cells. NK cells are large granular lymphocytes, which, unlike cytotoxic T cells, do not require a previous challenge and preactivation for their actions. NK cells are equipped with two different sets of surface receptors, inhibitory and activating, and a fine balance between these inhibitory and activating stimuli determines whether NK cells will destroy the target cells. These inhibitory and activating receptors on NK cells belong to either the immunoglobulin-like receptor (IgSF) superfamily such as killer cell immunoglobulin-like receptors (KIRs) and leukocyte Ig-like receptors (LIRs) or the CTLD superfamily, such as Ly49s, CD69, NKG2D, and CD94/NKG2. The author describes in a functional context the structure of the NK cell surface receptors of the C-type lectin superfamily and their complex with the major histocompatibility complex (MHC) Class I or Class I-like molecules that have been resolved through high-resolution x-ray crystallography. In Chapter 26, Yu and Kanost discuss the roles of a family of C-type lectins (immulectins) in innate immunity of tobacco hornworm. Immulectins serve as PRRs and bind to structural molecular patterns, generally termed pathogen-associated molecular patterns, present on the surface of most microorganisms but not on the host cells. Several chapters are devoted on the regulatory roles of galectins in immune responses. Rabinovich et al. (Chapter 27) summarize recent developments in understanding of the role of galectins within different immune cell compartments, and in the broader context of the inflammatory microenvironments. The authors provide solid evidence to claim that galectins can serve as immunosuppressive agents or targets for anti-inflammatory drugs. Stowell and Cummings (Chapter 28) evaluate the effects of galectins on leukocytes in aspects of innate immunity. The authors propose that galectins may serve as key immune regulatory elements, which can induce key changes in leukocytes, including apoptotic death, removal and turnover, or immunosuppressive alterations in cytokine production. Galectins are believed to play a key role to signal leukocyte elimination or shift toward a more tolerable Th2 outcome, following their release during excessive immune-mediated tissue damage. Hsu and Liu (Chapter 29) very effectively describe in detail the regulation of immune responses by galectin-3.

Some new lectins/lectin families are also included in this part. Lee and Pierce (Chapter 30) describe the X-lectins, a new family of lectins with homology to the *Xenopus laevis* oocyte lectin

XL-35, which lack a CTLD but carry a fibrinogen-like motif. X-lectins participate in the formation of the fertilization envelope that blocks sperm entry, and in the embryo, in cell–cell interactions and cell–matrix adhesion. Vasta et al. (Chapter 31) describe F-type (L-fucose-binding) lectins, a new family of recognition factors. The F-type lectin from European eel revealed a novel lectin fold (the "F-type" fold) with unique fucose- and calcium-binding sequence motifs. The F-type lectin family comprises a large number of proteins present from prokaryotes to vertebrates, exhibiting single, double, or greater multiples of the F-type motif, either arrayed in tandem or in mosaic combinations with other domains, yielding subunits of variable sizes even within a single species. This widespread, heterogeneous distribution of the F-type domain suggests an extensive structural–functional diversification of this lectin family. In Chapter 32, Stout et al. discuss the biology of fibrinogen-related proteins (FREPs) from the freshwater snail *Biomphalaria glabrata*. A brief overview of the role of FREPs in internal defense across the animal phyla is provided. The discovery and basic structural features of FREPs and their roles in defense responses against trematodes are discussed. Kamhawi and Valenzuela (Chapter 33) describe lectins in sand fly–*Leishmania* interactions and hypothesize that the lectin–*Leishmania* interactions may be necessary for a successful infection in the fly. In Chapter 34, Matsushita et al. discuss the structural basis for recognition plasticity of ficolins, a family of proteins that have both collagen-like and fibrinogen-like domains with a carbohydrate specificity for *N*-acetylglucosamine. Serum ficolins play an important role as pattern recognition molecules in innate immunity by complexing with serine proteases termed MASPs. Upon binding to carbohydrates on the surfaces of microbial pathogens, serum ficolins eliminate them by direct opsonization. In Chapter 35, Hatakeyama describes characteristic features of a Ca^{2+}-dependent, Gal/GalNAc-specific lectin (CEL-III) from the sea cucumber (*Cucumaria echinata*) and its putative mechanism of hemolytic activity on the basis of the structural information. This lectin exhibits hemolytic and cytotoxic activities through formation of ion-permeable pores composed of its oligomers in the target cell membranes. In the last chapter, Morroll et al. describe the structural features of chitinases and chilectins, which may have arisen from the gene duplication events. Chilectins have lost the ability to cleave chitin due to mutations in the active site, but retain the ability to bind carbohydrates.

REFERENCES

1. Barondes, S.H. 1988 Bifunctional properties of lectins: Lectins redefined. *Trends Biochem. Sci.* 13, 480–482.
2. Camus, M.L. 1899 Recherches experimentales sur une agglutinine produite par la glande de l'albumen chez *l'Helix pomatia. C.R. Acad. Sci.* 129, 233.
3. Charcot, J.M. and Robin, C. 1853 Observation de leocythemie. *C. R. Mem. Soc. Biol.* 5, 44–50.
4. Leyden, E. 1872 Zur Kenntniss des bronchial-asthma. *Arch. Pathol. Anat.* 54, 324–344.
5. Swaminathan, G.J., Leonidas, D.D., Savage, M.P., Ackerman, S.J., and Acharya, K.R. 1999 Selective recognition of mannose by the human eosinophil Charcot–Leyden crystal protein (Galectin 10): A crystallographic study at 1.8 A° resolution. *Biochemistry* 38, 13837–13843.
6. Flexner, S. and Noguchi, H. 1902 Snake venom in relation to haemolysis, bacteriolysis, and toxicity. *J. Exp. Med.* 6, 277–301.
7. Kilpatrick, D.C. and Green, C. 1992 Lectins as blood typing reagents. In: Franz, H. (Ed.), *Advances in Lectin Research,* Vol. 5. Ullstein Mosby, Berlin, pp. 51–94.
8. Noguchi, H. 1903 On the multiplicity of the serum haemagglutinins of cold-blooded animals. *Zentralbl. Bakteriol. Abt. l Orig.* 34, 286.
9. Cantacuzene, J. 1912 Sur certains anticorps naturels observes chez *Eupagurus prideauxii. C.R. Soc. Biol.* 73, 663.
10. Watkins, W.M. and Morgan, W.T.J. 1952 Neutralization of the anti-H agglutinin in eel serum by simple sugars. *Nature* 169, 825–826.
11. Marchalonis, J.J. and Edelman, G.M. 1968 Isolation and characterization of a hemagglutinin from *Limulus polyphemus. J. Mol. Biol.* 32, 453–465.
12. Hammarstrom, S. and Kabat, E.A. 1969 Purification and characterization of a blood-group A reactive hemagglutinin from the snail *Helix pomatia* and a study of its combining site. *Biochemistry* 8, 2696–2705.

13. Springer, G.F. and Desai, P.R. 1971 Monosaccharides as specific precipitinogens of eel anti-human blood group H (O) antibody. *Biochemistry* 10, 3749–3760.
14. Hudgin, R.L., Pricer, W.E. Jr., Ashwell, G., Stockert, R.J., and Morell, A.G. 1974 The isolation and properties of a rabbit liver binding protein specific for asialoglycoproteins. *J. Biol. Chem*. 249, 5536–5543.
15. Teichberg, V.I., Silman, I., Beitsch, D.D., and Resheff, G. 1975 A β-Dgalactoside binding protein from electric organ tissue of *Electrophorus electricus. Proc. Natl Acad. Sci. USA* 72, 1383–1387.
16. Agrawal, B.B.I. and Goldstein, I.J. 1965 Specific binding of concanvalin A to cross-linked dextran gels. *Biochem. J.* 96, 23C–25C.
17. Watkins, W.M. 2001 The ABO blood group system: Historical background. *Transfus. Med*. 11, 243–265.
18. Kocourek, J. 1986 Historical background. In: Liener, I.E., Sharon, N., and Goldstein, I.J. (Eds.), *The Lectins: Properties, Functions and Applications in Biology and Medicine*. Academic Press, Orlando, FL, pp. 1–32.
19. Sharon, N. and Lis, H. 2004 History of lectins: From hemagglutinins to biological recognition molecules. *Glycobiology* 14, 53R–62R.
20. Drickamer, K. 1988 Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem*. 263, 9557–9560.
21. Zelensky, A.N. and Gready, J.E. 2005 The C-type lectin-like domain superfamily. *FEBS J.* 272, 6179–6217.
22. Ewart, K.V., Li, Z., Yang, D.S., Fletcher, G.L., and Hew, C.L. 1998 The ice-binding site of Atlantic herring antifreeze protein corresponds to the carbohydrate-binding site of C-type lectins. *Biochemistry* 37, 4080–4085.
23. Drickamer, K. and Dodd, R.B. 1999 C-Type lectin-like domains in *Caenorhabditis elegans*: Predictions from the complete genome sequence. *Glycobiology* 9, 1357–1369.
24. Cooper, D.N. 2002 Galectinomics: Finding themes in complexity. *Biochim. Biophys. Acta*. 1572, 209–231.
25. Bianchet, M.A., Odom, E.W., Vasta, G.R., and Amzel, L.M. 2002 Identification of a fold common to a family of innate immunity recognition factors: Structure of a fucolectin from the European eel (*Anguilla anguilla*). *Nature (Struct. Biol.)* 9, 628–634.
26. Vasta, G.R., Ahmed, H., and Odom, E.W. 2004 Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. *Curr. Opin. Struct. Biol*. 14, 617–630.
27. Odom, E.W. and Vasta, G.R. 2006 Characterization of a binary tandem domain F-type lectin from striped bass (*Morone saxatilis*). *J. Biol. Chem*. Jan 20; 281(3), 1698–1713.

Part II

Modern Approaches for Assessing Lectin Function

2 Structural Aspects of Lectin–Ligand Interactions

Mario A. Bianchet, Hafiz Ahmed, Gerardo R. Vasta, and L. Mario Amzel

CONTENTS

2.1 INTRODUCTION

Animal lectins are carbohydrate-binding proteins expressed in a variety of tissues. They are excreted or membrane-bound proteins and mediate many biological processes including cell-to-cell adhesion, protein trafficking, and innate responses to pathogens (Drickamer and Taylor, 1993; Gabius, 1997). Lectins recognize carbohydrate-containing structures of the host itself and of the host's pathogens. Their most common ligands are carbohydrates structures such as glycoconjugates of endogenous glycoproteins or oligosaccharide patterns on external membranes of pathogens (Buz et al., 2006).

Recognition of such structures by the lectin may be traced back primarily to binding by the lectin carbohydrate recognition domain (CRD) of a few monosaccharide moieties of the oligosaccharide, frequently terminal or capping monosaccharides such as sialic acid (Sia), galactose (Gal), fucose (Fuc), or sulfated sugars (Rudd et al., 2001). Although, this monosaccharide specificity characterizes the lectin function, monosaccharides bind to a single-lectin CRD with high specificity but with low affinity (dissociation constants K_D in the millimolar range (Rini, 1995; Weis and Drickamer, 1996; Weis, 1997)). Lectins have the ability, required by their biologic function, to differentiate among a vast variety of complex carbohydrate structures sharing the same monosaccharide. In particular, innate immunity functions require discrimination between host and nonhost carbohydrate structures. The required enhanced avidity for a particular glycan is achieved by secondary recognition of other carbohydrate moieties of the glycan in an extended binding site or by combined recognition by multiple CRDs (multiplicity) of several glycans. Clustering of CRDs, built either in a single polypeptide chain or as a result of quaternary association, endows lectins with specificity for arrangements of glycans present in the host or the nonhost structures, enabling their discrimination. For example, mannose-binding proteins (MBPs) bind with high avidity the cell surface of pathogens that present a dense arrangement of high mannose structures. In this chapter, we focus our discussion on structural aspects of carbohydrate recognition by lectins.

Animal lectins can be classified into two classes: lectins that contain an evolutionarily conserved CRD and lectins that do not contain this conserved unit and appear to have converged by evolving equivalent carbohydrate binding properties. Based on sequence homologies and evolutionary relatedness, most of the lectins have been classified into several families: C-type, S-type or galectins, P-type or Man-6-P lectins, Ig-type or Siglecs, R-lectins, and F-type lectins (Table 2.1). The other class of animal lectins, which includes pentraxins (pentameric subunit arrangements), glycoaminoglycan-binding proteins (basic aminoacids cluster), sulfoglucoronosyl lipid-binding proteins, ganglioside-binding proteins, resists classification because members of this class do not show a detectable sequence homology or evolutionary relation. (We are going to limit our discussion to first class because a comprehensive discussion of the other class is outside of the scope of this chapter.) The most consistent structural characteristic responsible for carbohydrate recognition by lectins is the presence of a shallow pocket containing a specific pattern of hydrogen-bonding groups that interact with sugar hydroxyls and aromatic groups or hydrophobic patches that stack or interact with the carbohydrate apolar face or with hydrophobic substituents.

In the following sections we will discuss the structural characteristics of the different lectin families.

TABLE 2.1
Lectin Families

Lectin Family	Cation	Carbohydrate Specificity	Defining Features
C-type	Yes	Man, Gal, Fuc (Variable)	C-type sequence motif
S-type (Galectins)	No	β-galactosides	S-type sequence motif
F-type	Yes	Fucose	F-type sequence motif
I-type (singlecs)	No	Sia (Variable)	Ig-like domains
P-type (Man-6-P)	Variable	Mannose-6-P	P-type sequence motif
R-type	No	β-Galactosides	R-type sequence motif
Others Families			
Heparin-binding type	No	Heparin/heparan-$(SO_4)^{2-}$	Basic amino acid clusters
Pentraxins	Most	PC/galactosides	Pentameric repeats

PC, phosphoryl-choline.

2.2 C-TYPE LECTINS

C-type lectins (or C-lectins) are a family of animal proteins that exhibit Ca^{2+}-dependent carbohydrate-binding properties. C-lectins mediate biological processes including endocytosis of ligands, cell-to-cell adhesion, and serum glycoproteins turnover (Drickamer and Taylor, 1993). The earliest C-type lectins studied were MBPs (Drickamer et al., 1986; Taylor et al., 1992). Members of this family present a variety of modular architectures; all share one or more modules, the C-type lectin domain (CLD), with a signature sequence motif (Figure 2.1A). In this motif, four conserved cysteines residues, which form two structurally important disulfide bridges, are strictly conserved. Other proteins that lack carbohydrate-binding properties also contain C-type lectin-like domains (CTLDs), suggesting that this domain is a characteristic of a larger protein superfamily (Drickamer, 1993). Nevertheless, we are going to focus this presentation on a subset of these proteins with carbohydrate-binding properties.

2.2.1 Three-Dimensional Structure of the C-Type Lectin Domain

Many structures of CLD-containing proteins have been determined (Table 2.2); in those, the regular elements of secondary structure are conserved with changes concentrated in the loops connecting these elements. The CLD is a globular α/β structure between 115 and 130 amino acids

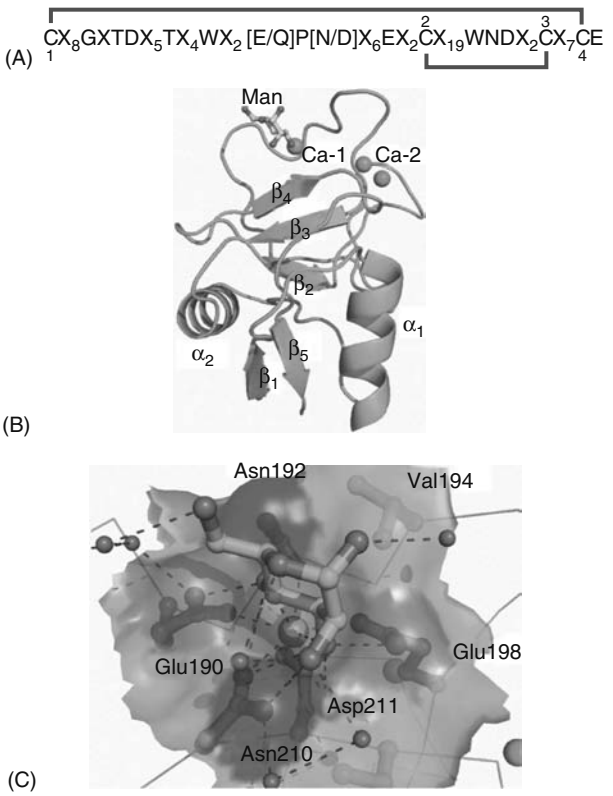


FIGURE 2.1 (See color insert following blank page 170. Also see CD for color figure.) Sequence motif and structure of C-lectins. (A) Sequence motif of C-lectins. The long sequence motif characteristic of C-lectins shows the conserved residues in blue, the cysteines in green, and the conserved proline in red. The two disulfide bridges that connect the cysteines are shown by the thick magenta lines. (B) Overall structure. The structure of the C-lectin CRD is the small α/β module shown in the figure using the PDB 1RLD. The three Ca^{2+} are shown as green spheres. At the top of the CRD, the binding site for ligand is shown occupied by mannose. (C) Mannose bound to a C-lectin. The binding site is shown as a transparent surface. The groups interacting with the ligand are visible through the surface. Residues that contact the ligand are visible through the surface. H-bond interactions are indicated by the dashed line.

in length (Figure 2.1B). The carbohydrate binding site is located at one end of the domain. Between one and four bound Ca^{2+} atoms are observed in the crystal structures of different CLDs (Table 2.2); only one of Ca^{2+} sites is strictly conserved among these lectins. Man-binding proteins have two calcium atoms, Selectins have only one. The N- and C-termini of the CLD form a two strands antiparallel β-sheet (β_1 and β_7 at the opposite end of the carbohydrate-binding site. Two consecutives helices (α_1 and α_2) surround these strands. Helix α_2 and the β-sheet (β_1, β_7) participate in oligomer formation (Drickamer, 1999). These secondary structure elements encompass roughly one-half of the CLD residues and constitute a scaffold for the other half, which is commonly described as a double loop or "loop in the loop" structure (Zelensky and Gready, 2005). The N-terminal half of the double loop is a long loop (Figure 2.1B) that displays the greatest variation among C-lectins. The size of this long loop allows sorting of CLDs in two types: canonical CLDs and compact CLDs (Drickamer, 1999; Zelensky and Gready, 2005). This region contains most of the hydrophobic core residues of the domain and the less conserved first calcium-binding site (Ca-1).

TABLE 2.2
Crystal Structures of Lectins

Lectin Type	PDBid	Ligand	Species	Comments
Galectin-1	1A78	TDG	Amphibian	
	1HLC	Lac	Human	
	1SLT	LacNAc	*B. taurus*	
Galectin-3	2NN8	Lac		
Galectin-4	1X5O		Human	NMR
Galectin-7	1BKZ		Human	
Galectin-9	2D6M		Mouse	
C-lectin	1BYF	D-Gal	Tunicate	
	1G1S	S-Le^x, peptide	Human	P-selectin
	1G1T	S-Le^x	Human	P-selectin
	1H8U		Human	MBP
	1IXX		Snake	
	1SL6	Le^x	Human	
	1SZB		Human	MAP19
	1TN3		Human	Tetranectin
	1WMZ	GalNAc	Echinoderm	
	2AFP		Fish	NMR
C-lectin (CD23)	2M2T			
	2KMB	3′NeAu Le^x		
	1C3A			
	1E9L			
	1JC9			
F-lectin	1K12	Fuc	Fish	*A. anguilla*
	3CQ0	Fuc	Fish	*Striped Bass*
	2J22	Fuc	*S. pneumoniae*	
IG-lectin	1NKO			
R-lectin	2AO3	OMe-GalNAc	Earthworm	
Tackylectin-5	1TL2	GlcNac	Crab	
P-lectin	1M6P	Man-6-P	*B. taurus*	
	1SZ0	Man-6-P	*B. taurus*	

The C-terminal region of the double loop binds carbohydrate in a shallow depression made by three stretches of residues S1, S2, and S3 in a S1–S3–S2 sequence, surrounded by five loops. S2 and S3 are β-strands in an antiparallel β-sheet (β_5, β_6; Figure 2.4B). S1 contains the motif [E/Q] P[N/D]X_6E, S2 contains the conserved Cys, and S3 a WNDX_2C motif. This highly conserved and more hydrophilic portion of the fold is stabilized by one conserved disulfide bridge that joins the onset of S2 and the end of S3. A conserved Ca^{2+} site (Ca2) bridges S1 and S3 (strand β_6). The structural and carbohydrate-binding roles of the calcium sites correlate with the cation dependence of these lectins (Ng et al., 1998). An extra segment that folds as a hairpin, stabilized by an additional pair of bridged cysteines separated by 10 residues, is observed in several structures including tetranectin (Nielsen et al., 1997) and are referred as long-form CLDs.

2.2.2 Carbohydrate-Binding Site and Specificity of C-Type Lectins

The Ca^{2+} and its equatorial ligands bind two contiguous hydroxyl groups of the carbohydrate, such as the 3- and 4-OH groups in Man and GlcNac or the equivalent 2- and 3-OH groups of Fuc. Five residues from S1 and β_6 coordinate the cation at its binding site. Strand S1 provides three oxygen atoms as Ca^{2+} equatorial ligands: a Glu (or Gln see below), an Asn one residue away and a Glu further down (in blue in the motif of Figure 2.1A). The *cis*-proline residue in S1, between the first two Ca^{2+} ligands, plays an essential role in forming correct structure of the binding site. This residue has been shown to go through a *cis–trans* isomerization reaction when Ca^{2+} binds the protein (Ng et al., 1998). Strand S3 provides two oxygen atoms as cation ligands, one equatorial from an Asn and one axial from a nearby Asp (Figure 2.4B). Early in the discovery of C-type lectins, a distinction was made between C-lectins that preferentially bind Man-containing carbohydrates and those that prefer Gal-containing structures. The monosaccharide preference was correlated with a short sequence in S1: EPN for Man-binding and QPD for Gal-binding. Iobst and Drickamer (1994) switched the specificity of a mannose-binding C-lectin to Gal by substituting the Glu to Gln and the Asn to Asp (EPN → QPD). Their structural analysis of the Gal-specific MBP-A mutant showed that the substitutions do not produce changes in the binding site and that the switch in the specificity was induced by swapping the hydrogen-bond donor (N of the amides) and acceptor (O, of the carboxylates) across the sugar-binding plane that results in changes in the hydrogen-bonding pattern from the Man-type asymmetrical (around the Ca^{2+}) to Gal-type symmetrical. The same distribution of H-bonding partners was observed in the Gal-binding lectin TC-14 from the tunicate *Polyandrocarpa misakiensis* (Poget et al., 1999). The TC-14 CLD contains an unusual EPS motif in S1, which shows similarities to the motifs of the Man-binding proteins but contains a Ser as a hydrogen-bond donor instead of the Asp in MBP-A. In addition, the crystal structure revealed that this change in specificity is due to a compensatory modification on the opposite side of the ligand-binding site (the S2 WND-motif is changed to LDD), and a 180° rotation of the Gal residue compared with the orientation observed in the Gal-binding MBP-A mutant.

2.2.3 Extended Carbohydrate Site

Other moieties of complex oligosaccharides show additional interactions with groups of the protein outside of the binding site that result in enhanced specificity. For example, a shallow hydrophobic pocket close to the Ca-2 site provides a binding site for the N-acetyl group of GalNac in the asialoglycoprotein receptor and results in a 50-fold enhancement of the affinity for GalNac compared to the nonsubstituted Gal (Kolatkar et al., 1998).

2.2.4 Quaternary Association and C-Lectin Subfamilies

Collectins and selectins are oligomeric molecules composed of CLDs attached to collagen regions or α-coiled neck regions.

Collectins are a class of C-lectins containing a collagen-like domain. Mannose-binding lectin (MBL), conglutinin, and collectin-43 (CL-43) are serum proteins produced by the liver. MBLs have affinity for high-mannose oligosaccharides on cell surfaces of pathogens (Vasta et al., 1999; Ng et al., 2002). They assemble in large oligomeric complexes containing up to 27 units. MBP-A forms, with its collagen like tails, a trimeric helical structure stabilized with disulfides bridges in its cysteine-rich N-terminal region.

Selectins are integral membrane proteins that mediate adhesion in leukocyte migration. The three types, E-, L-, and P-selectins, bind Fuc and the negatively charged sialic acid. As a consequence, all three proteins bind Lewis x (Le^x; Gal α(1,4)-[Fuc α(1–3)]-GlcNac) and sialyl-derivatized Le^x (sLe^x; NeuNAc α(2,3)–Galα(1,4)-[Fuc α(1–3)]-GlcNac). All selectins share the domain architecture of an N-terminal CLD followed with a primary sequence similar to the elongation growth factor (EGF) precursor, and followed for a variable number, of tandem repeats of 62 residues each (CR-repeats, complement-regulatory or consensus repeat) with a transmembrane segment and a short cytosolic domain (Lasky et al., 1989). The number of CR-repeats is type-dependant: E-selectins have four repeats, P-selectins have nine repeats, and L-selectins have two (CD62L protein), six (CD62E protein), and nine repeats (CD62P protein) (Lasky, 1995).

2.3 GALECTINS

Galectins are a family of β-galactoside-binding proteins that are believed to participate in a variety of biological processes, such as early development, tissue organization, immune functions, tumor evasion, and cancer metastasis (Leffler et al., 2004; Vasta et al., 2004; Liu and Rabinovich, 2005). To date, 15 members of the galectin family (Ahmed et al., 1990) have been identified in higher vertebrates displaying either one (proto and chimera types) or two CRDs (proto and tandem-repeat types) (Cooper, 2002). Galectins-1, 2, 5, 7, 10, 11, 13, 14, and 15 are examples of the proto type galectins, of which galectin-5 and 7 are monomeric, whereas all others are homodimeric. Galectin-4, 6, 8, 9, and 12 are tandem-repeat type galectins. Galectin-3 is the only chimera-type galectin, with three distinct structural domains: a 12-amino acid N-terminal domain, a collagen-like sequence rich in proline and glycine, and a carboxy terminal domain containing a CRD (Gong et al., 1999).

Galectins have a high affinity for small β-galactosides, but binding specificity for complex glycoconjugates varies considerably within the family. Structural analysis of lectins has provided insights into the molecular basis of protein–carbohydrate interactions. Moreover, the quaternary structure of the lectins provides information about multivalency in protein–carbohydrate interactions.

2.3.1 Three-Dimensional Structure of Galectins

The structures of five mammalian galectins has been determined so far (Lobsanov et al., 1993; Bourne et al., 1994; Liao et al., 1994; Leonidas et al., 1995, 1998; Seetharaman et al., 1998): one avian (Varela et al., 1999), one amphibian (Bianchet et al., 2000), two fish (Shirai et al., 1999, 2002), and one fungal (Walser et al., 2004). They all share a similar structure consisting of a jelly roll topology typical of legume lectins. The crystal structure of the bovine galectin-1 complexed with the disaccharide *N*-acetyllactosamine [LacNAc, Galβ(1,4)GlcNAc] determined at 1.9 Å resolution (Liao et al., 1994) (Figure 2.2A) defined the fold of these lectins. Each subunit of the galectin-1 dimer is composed of a β-sandwich of two antiparallel β-sheets formed by six (β_1, β_1, β_{3-6}) and five (β_{11}, β_2, β_{7-9}) strands respectively. Galectin-1 and -2 are dimeric and the N- and C-terminals of each subunit are located at the dimer interface. The two monomers of the galectin-1 are related by a twofold rotation axis perpendicular to the plane of the β-sheets. The structure reveals that there is one carbohydrate-binding site per monomer. These carbohydrate-binding sites, located on the same side of the β-sandwich and at opposite ends of the dimer 46 Å apart, are defined by clefts formed by strands β_4–β_6 of the six-stranded β-sheet (Figure 2.2A and C). Interestingly, despite the

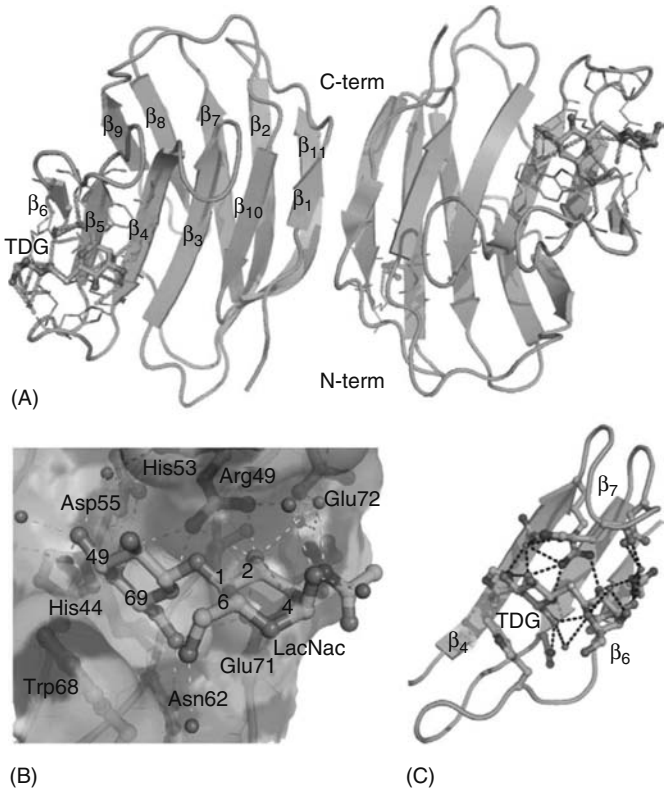


FIGURE 2.2 (See color insert following blank page 170. Also see CD for color figure.) Structure of galectins. (A) Overall structure of a galectin dimer. The dimer is shown with the twofold axis perpendicular to the plane of the figure. The strands forming the two β-sheets are labeled. TDG is shown bound to the lectin. (B) Details of LacNac binding to *B. taurus* galectin-1. A transparent rendition of the surface of the binding cavity is shown with the amino acids visible behind the surface. (C) Detail of TDG binding to the *B. arenarum* galectin-1.

lack of significant sequence homology, galectins fold with the same jelly roll topology as that of the legume lectins such as pea lectin ConA, and *Erythrina corallodendron* lectin (Liao et al., 1994). (The mammalian pentraxins, serum amyloid protein (SAP) and the C-reactive protein (CRP), also share the legume lectin fold (Emsley et al., 1994).) In contrast, there is no resemblance between the fold of galectins and that of C-type lectin and no sequence homology, indicating these two families of lectins are not related.

2.3.2 Quaternary Structures

Although galectins share the same fold, their quaternary association may differ (Lobsanov et al., 1993; Bourne et al., 1994; Liao et al., 1994; Leonidas et al., 1995, 1998; Seetharaman et al., 1998; Shirai et al., 1999, 2002; Varela et al., 1999; Bianchet et al., 2000; Walser et al., 2004). Furthermore, some unique structural features can be identified in the most divergent family members. For example, Congerin I, a galectin from the conger eel *Conger myriaster*, has conformationally extended N- and C-termini that are believed to participate in intersubunit "strand swapping" when dimerized (Shirai et al., 1999). This "strand-swapinp" contributes to stabilize the dimer by increasing intersubunit interactions, and perhaps explains the high thermostability of the protein (Shirai et al., 1999). CGL2, a galectin from the fungus *Coprinus cinereus*, forms a tetramer characterized by two perpendicular twofold axes of rotation, with the C-terminal amino acids of the four monomers meeting at the center of the tetramer interface (Walser et al., 2004).

The Galectin-7 structure was determined in the native form and in complexes with Gal, GalNac, lactose (Lac), and *N*-acetyllactosamine (LacNac) (Leonidas et al., 1998; Saussez and Kiss, 2006). Although binding of carbohydrate by the Galectin-7 protomer is quite similar to that other galectins, the crystal structure (Table 2.2) of galectin-7 shows that dimerization of this lectin is very different from those of dimeric galectins. A hydrophobic patch opposite to the carbohydrate-binding site mediates the dimerization burying 1484 $Å^2$ (Leonidas et al., 1998; Rini and Lobsanov, 1999).

2.3.3 Carbohydrate-Binding Specificity of Galectins and Their Carbohydrate-Binding Sites

The ability of galectins to discriminate among carbohydrate structures is striking. For example, the binding affinity of bovine galectin-1 for either D-Gal or its α/β methyl derivative is almost 200 times lower than for the β-Gal-containing disaccharide Lac (Ahmed et al., 1990, 1996, 2004). For most galectins, LacNAc and thiodigalactoside (TDG) are 5–10 times more active than Lac. In general, the 4′-OH, 6′-OH groups of the Gal residue of Lac/LacNAc are critical for binding, and thus, any changes (epimers or substitutions) of these hydroxyls reduces binding efficiency. The equatorial 3-OH of the GlcNAc residue of LacNAc is also important for binding because Galβ1,3GalNAc (T-disaccharide) is a very poor inhibitor for most galectins. Substitution of the 2′-OH, the 3′-OH of the Gal residue or both, however, does not affect the binding of galectin-1 to the carbohydrate ligand.

Close examination of the carbohydrate-binding specificities of galectins, however, reveals diversity in their binding properties (Ahmed et al., 1994, 2002). For example, the T-disaccharide is a good ligand for galectin-3 and other chimera galectins and for the 16 kDa galectin from *Caenorhabditis elegans* (Leffler and Barondes, 1986; Ahmed et al., 2002). Moreover, GalNAcα(1,3)[Fucα(1,2)] Galβ(1,4)Glc (A-tetrasaccharide with substitutions at 2′ and 3′-OH of the Gal residue) has almost 100-fold higher affinity for galectin-3 than for galectin-1 (Leffler and Barondes, 1986). The crystal structures of galectins complexed with their carbohydrate ligands allowed the identification of not only the amino acids involved, but also of the hydroxyl groups of the ligands that participate in the interactions. Moreover, the molecular basis of the variable binding properties of certain galectins can be explained in terms of their 3-D structure.

2.3.3.1 Conserved Carbohydrate-Binding Site

The carbohydrate-binding site of the bovine galectin-1 is formed by three continuous concave strands β_4–β_6 containing all residues involved in direct interactions with the LacNac (Figure 2.2B). The Gal residue of the bound LacNac is deeply buried in the carbohydrate-binding cleft. The 4′-OH group of the Gal makes hydrogen bonding interactions with the highly conserved residues His44, Asn46, and Arg48. The 6′-OH makes similar interactions with Asn61 and Glu71. Trp68 participates in a stacking interaction with the Gal ring carbons and restricts the orientation of the 4′-OH to the axial form. In the GlcNAc moiety of the LacNAc, the 3-OH makes hydrogen bonds with Arg48, Glu71, and Arg73. Additional interactions involve a water molecule that bridges the nitrogen of the NAc group with His52, Asp54, and Arg73. The latter interactions, present in most galectins-1, may explain a five- to eightfold increase in affinity of LacNAc over lactose (Leffler et al., 1989; Ahmed and Vasta, 1994).

The prototype galectin from *B. arenarum* resembles the mammalian galectin-1 with respect to its carbohydrate-binding profiles and its carbohydrate-binding sequence motif. All nine residues responsible for carbohydrate binding, His45, Asn47, Arg49, His53, Asp55, Asn62, Trp69, Glu72, and Arg74 are present (Ahmed et al., 1996). The structures of the *B. arenarum* galectin complexed with LacNAc or TDG further support this close similarity (Bianchet et al., 2000). The two galectin complexes, one with TDG and one with LacNac, show that the topologically related hydroxyl groups of the disaccharides exhibit similar interaction patterns with the protein.

2.3.3.2 Extended Carbohydrate-Binding Site

Unlike galectin-1, galectin-3 has an extended carbohydrate-binding site. As shown in Figure 2.2A, a galectin carbohydrate-binding site is formed by a cleft open at both ends. In this cleft, the LacNAc is positioned in such a way that the reducing end of the LacNAc (GlcNAc) is open to solvent, but the nonreducing moiety (Gal) is in close proximity to residues in the strand β_3. In galectin-3, which favors higher affinity interactions with the GalNAcα(1,3)[Fucα(1,2)] Galβ(1,4)Glc, it was expected that residues on β_3 interact with the α-linked GalNAc residue. The crystal structure of galectin-3 (Table 2.2) reveals that indeed, Arg29 in β_3 makes a direct hydrogen bond interaction with the bound oligosaccharide (Seetharaman et al., 1998). The extended binding site beyond the nonreducing end of bound galactosides has also been observed in the congerin II crystal structure (Shirai et al., 2002; Table 2.2).

2.3.3.3 Variable Carbohydrate-Binding Site

As in most galectins, the carbohydrate-binding cleft is usually formed by three strands β_4–β_6; the length of the loops connecting strands can be a critical factor for determining the carbohydrate-binding properties of many galectins. For example, the carbohydrate-binding specificity of the 16 kDa galectin from *C. elegans* is unique because it interacts with most blood group precursor oligosaccharides [Tα, Galβ(1,3)GalNAcα; Tβ, Galβ(1,3)GalNAcβ; type 1, Galβ(1,3)GlcNAc, and type 2, Galβ1,4GlcNAc] (Ahmed et al., 2002). A homology model of the *C. elegans* 16 kDa galectin reveals a shorter loop region between strand β_4 and β_5 of the CRD, which may be responsible for its unique binding profile. Particularly relevant are the interactions of Glu67 (Asp54 in bovine galectin-1) with the 3(4)-OH (equatorial) of GlcNAc [in Gal β(1,3[4])GlcNAc] and that of Glu67 with the 4-OH (axial) of GalNAc (in Gal β(1,3) GalNAc) due to shortening of the loop (Ahmed et al., 2002). Interaction of the equivalent residue, Asp54, with the axial 4-OH of Gal β(1,3)GalNAc is not possible for typical galectins-1 (i.e., those from bovine and *B. arenarum*) because of steric hindrance (Liao et al., 1994; Bianchet et al., 2000).

2.4 F-TYPE LECTINS

Members of this family of fucose-binding proteins have been characterized as immune recognition molecules in invertebrates such as the horseshoe crab (*Tachypleus tridentatus*) (Saito et al., 1997) and vertebrates such as the Japanese eel (*Anguilla japonica*) (Honda et al., 2000) and *Xenopus laevis* (Seery et al., 1993). Homologous domains have been also detected in bacteria, such as *Streptococcus pneumoniae*, probably as consequence of lateral transference of genes (Bianchet et al., 2002; Odom and Vasta, 2006). F-lectins exhibit a 140-residue module (F-type lectin domain; FLD) showing conservation of two motifs: h_2DGx, and HX_{24}RXDX_4 [R/K] (Figure 2.3), where h stands for a small hydrophobic residue (i.e., Val, Ala, or Ile), "x" for a small hydrophilic residue (i.e., Asn, Asp, or Ser) and X for any residue.

As in C-type lectins, similar structural domains have been observed in other proteins such as bacterial sialidase (PDB code 1EUT) and fungal galactose oxidase (PDB code 1GOF). In addition, FLD folds have been observed in a large family of adhesion proteins such as the C1 and C2 domains of the coagulation factors Va (FVa C1/2) and VIII (FVIII) and others from unrelated pathways (APC10/DOC1(Wendt et al., 2001) and XRC11), suggesting either convergent evolution or a high tolerance for mutations in this fold. Interestingly, the common function of these protein domains is to bind specific ligands, such as DNA nucleotides, in the case of XRC11, or phospholipids of the mammalian cell membrane, in the case of the coagulation factors. Interestingly, FVa and FVIII have sequence similarities to members of the discoidin domain family (Baumgartner et al., 1998), most of which have been implicated in cell adhesion or developmental processes (Vogel, 1999). Discoidin I and II, proteins from the slime mold *Dictyostelium discoideum*, were initially described as lectins with a high affinity for Gal (Poole et al., 1981). Thus, it is plausible that carbohydrate recognition is the ancestral function of this domain, which may have evolved into a phospholipid-binding domain, as in the case of the coagulation factors V and VIII.

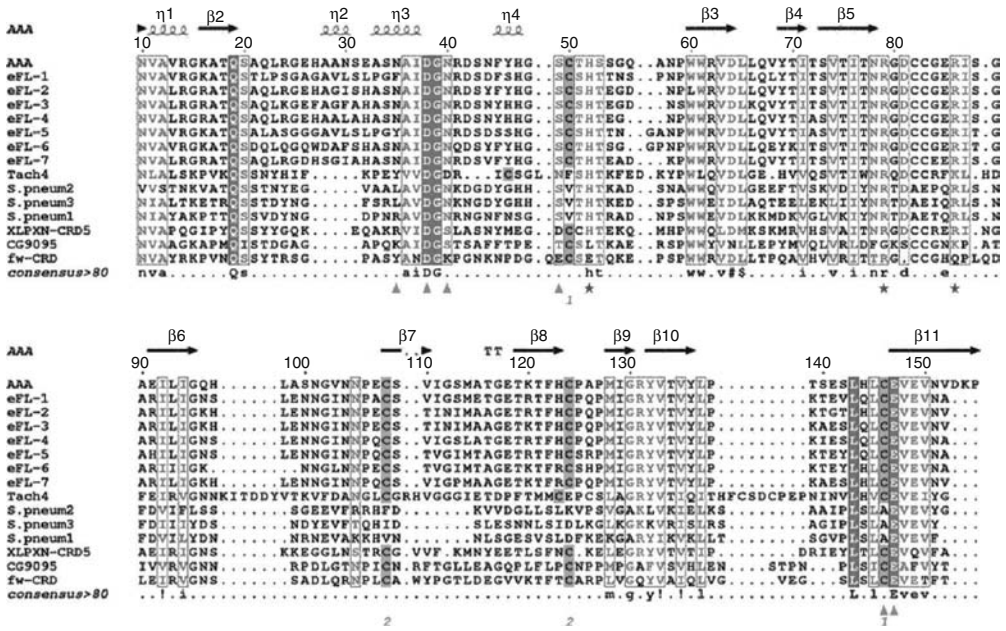


FIGURE 2.3 (See CD for color figure.) Sequence alignment of selected F-lectins. Fully conserved residues are shown on a magenta background; partially conserved residues are shown in red and cysteine residues are shown in green. Arrows indicate residues participating in the metal binding site and stars indicate residues of the conserved F-lectin motif. AAA, *anguilla anguilla* aglutinin (american eel); eFL-1 to 7, isoforms of european eel fucolectin; Tach4, tachylectin 4; S.pneum 1-3, *Streptococus Pneumoniae*; XLPXN-CRD5, *Xenopus Laevis* pentraxin; CG9095 and fw-CRD, *D. Megaloganster* furrowed-like and furrowed receptor respectively.

2.4.1 Three-Dimensional Structure of F-Type Lectin

The crystal structure of the complex of *Anguilla Anguilla* agglutinin (AAA) with α-L-Fuc (Bianchet et al., 2002) revealed the fold of this lectin family: an eight-strand β-barrel with jellyroll topology (Figure 2.4A). Two short antiparallel strands (β_4 and β_9) close the barrel at the end opposite to the binding site. The N- and C-terminal strands (β_1 and β_{11}) protrude 15Å from this end of the barrel. Five loops (CDR1–5) connect the main beta-sheets encircling a positively charged depression where Fuc binds (Figure 2.4B). Highly conserved disulfide bridges (two) and salt bridges (two), together with the bound cation clamp the structure together. On the other side of the hollow, CDR4 (Figure 2.4B) has two solvent-exposed Cys residues (Cys82 and Cys83) that form a rare disulfide bridge (between consecutive residues). At one side of the β-barrel, a 3_{10} helix-rich substructure, which includes CDR1 and CDR2, tightly binds a cation. A number of structural elements associated with the cation site are involved in shaping loops CDR1 and CDR2 of the binding site, stressing the importance of the conformation of these two loops for the extended carbohydrate binding. A disulfide bridge between Cys50 and Cys146 links the N-terminus of β_{11} with the C-terminus of CDR2, and a salt bridge between Arg41 and Glu149 attaches the loop between η_2 and η_3 to β_{11}. Both of these interactions hold the cation-binding substructure against the bulk of the fold. A disulfide bridge between Cys108 and Cys124 links strands β_7 and β_8, connecting the two main β-sheets. An intrasheet salt bridge between Asp64 and Arg131 links the N-terminus of β_3 with the C-terminus of β_{10} (in the front sheet, Figure 2.3A).

The cation is coordinated with seven oxygen atoms from the main chain or side chain of six residues: Asn35 (O), Asp38 (Oδ1), Asn40 (O), Ser49 (O, Oγ1), Cys146 (O), and Glu147. Even though the Ca^{2+} site is distant from the carbohydrate-binding site, an enhancement of carbohydrate affinity by calcium addition has been observed in these proteins (Saito et al., 1997). Tachylectin-4,

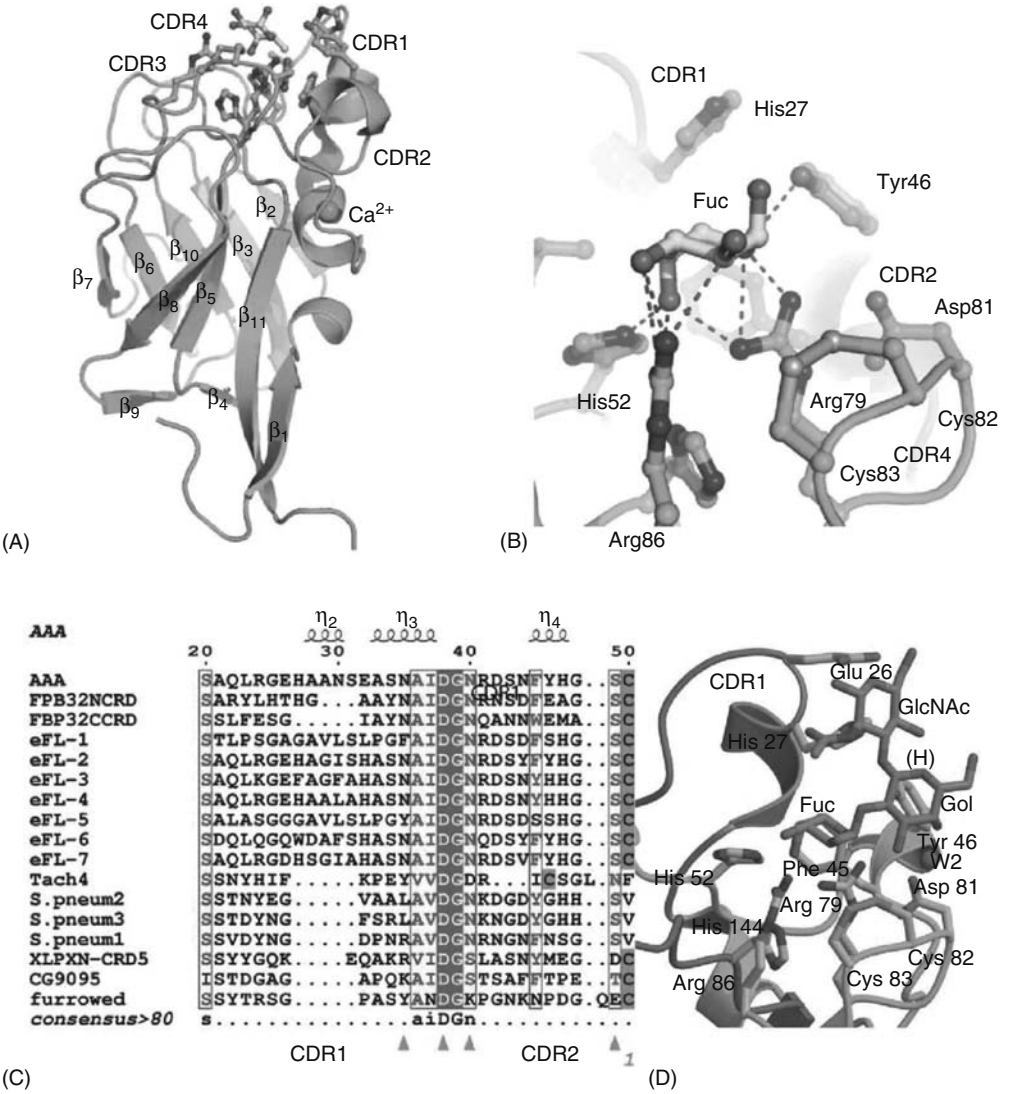


FIGURE 2.4 (See color insert following blank page 170. Also see CD for color figure.) Structure of F-lectins. (A) Overall structure. The two β-sheets that form the core of the molecule are clearly visible. The carbohydrate complementarity determining regions (CDRs) are indicated. (B) Fucose recognition. Residues that interact with the fucose ligand are shown and the H-bonds to the sugar indicated. The disulfide bridge formed by two consecutive residues (Cys82 and Cys83) is shown interacting with the hydrophobic portion of the ligand. (C) Alignment of the sequences of CDR1 and CDR2 of several F-lectins. Completely conserved residues are shown on a magenta background; partially conserved residues are shown in red and cysteines are shown in green. (D) Computer model of H type 1 trisaccharide bound to the AAA binding site. The model predicts interactions between the third carbohydrate moiety and the CDR1.

horseshoe-crab F-lectin, shows a fourfold enhancement of its hemagglutination activity for A-type erythrocytes upon the addition of Ca^{2+}. This enhancement, which is abolished by EDTA (Saito et al., 1997), may result from possible changes induced by cation binding of the conformation of the contiguous loops CDR1 and CDR2, which are involved in binding of the ligand.

2.4.2 Carbohydrate-Binding Site

AAA recognizes the ring O5 and the equatorial 3- and axial 4-OH groups of the α-Fuc using the nitrogen atoms of three conserved residues: Nε of a histidine and the guanidinium groups of two arginines (Figure 2.4B). A network of hydrogen bonds maintains this triad of residues in optimal positions at the center of the hollow to recognize the axial 4-OH group of the fucose with hydrogen bonds in a perfect tetrahedral geometry. In addition, the arginine residues of the triad recognize O5 and the equatorial 3-OH group of the Fuc. Two hydrophobic patches provide van der Waals contacts to the monosaccharide: its ring atoms C1 and C2 rest over a rare disulfide bridge between consecutive cysteine residues and its C6 docks loosely in a hydrophobic pocket, stacking against the rings of His27 and Phe45. These residues, together with Leu23 and Tyr46, form the binding pocket.

2.4.2.1 Extended Binding Site

Except for the Drosophila proteins, most of the sequences of Figure 2.3 conserve a fucose-binding sequence motif (Figure 2.4) and the size and hydropathic profile of CRD2 (Figure 2.4C), which forms part of the hydrophobic pocket for the Fuc methyl group. The second cation-binding motif (h_2DGx) is also conserved in these proteins. This motif provides three of the seven oxygen atoms that bind the cation.

Non-eel proteins (in the alignment in Figure 2.3) contain a shorter CDR1, lacking five of the residues that form 3_{10}-helix η_2. This erases the loop protruding-feature observed in AAA, reducing or eliminating interactions between CDR1 and the putative oligosaccharide antennae and perhaps broadening the specificity of these fucolectins for Le^x oligosaccharides (Bianchet et al., 2002). Organisms other than bacteria conserve the cysteines involved in the two interstrand disulfide bridges of AAA (50–146 and 104–128).

European eel serum presents several isoforms of AAA (Horejsi and Kocourek, 1978; Kelly, 1984) suggesting a tuning of the FLD to different pathogens. All fucolectins from the *Anguilla* genus shown in Figure 2.3 conserve the binding site triad showing conservation of the histidine and the CDR4 sequence. These F-lectins conserve the size of CDR1 and CDR2 (Figure 2.4D), although the loops present sequence variations in residues associated with the C6 pocket. The loop CDR1 that is in position to interact with other moieties of Fuc-containing oligosaccharides (Figure 2.4D) shows the greatest sequence variability. Most eel F-Lectins conserve polar residues at the apex of this loop (like Glu 26 and His27 in AAA), probably to make polar interactions with other carbohydrate moieties of the ligand. Also, most eel sequences conserve two aromatic CDR2 residues in the N-terminus of η_4 (Phe45 and Tyr46 in AAA) that form the pocket for the Fuc methyl group.

2.4.3 Architecture and Quaternary Structure of F-Type Lectins

From bacteria to vertebrates, several proteins contain one or more FLDs (Honda et al., 2000; Bianchet et al., 2002; Odom and Vasta, 2006). The N-terminal domains of the "furrowed" receptor and of the protein CG9095 of *Drosophila melanogaster* (Adams et al., 2000), the horseshoe crab tachylectin-4 (Saito et al., 1997), and the N-terminal domain of the *Xenopus laevis* (Seery et al., 1993) pentraxin 1 fusion protein are proteins involved in innate immunity that contain FTLD domains (Honda et al., 2000). FLDs are observed in chimeras with others domains such as the pentraxin-1 of *X. laevis*, the C-lectin type CRDs and the short CR (Meindl et al., 1995), which form complex toll receptors in the *Drosophila* proteins. Interestingly, in modern fishes (i.e., zebra fish, steelhead trout, stickleback, and pufferfish), the arrangement is either a duplicate or a quadruplicate tandem of FLDs (Odom and Vasta, 2006). Similarly, tandem arrangements are present in *Xenopus* spp., including either triplicate or quintuple concatemers. Clearly, the placement of the termini in the FLD favors the formation of concatenated CRD modules in numbers that appear lineage related. These tandem arrays may yield chimeric proteins by including pentraxin (*X. laevis*) or C-type domains (*D. melanogaster*).

AAA is a physiological trimer (Kelly, 1984). Although the crystal structure of AAA shows a protomer per asymmetric unit, the crystal belongs to a space group (R32) that includes a threefold symmetry. This fact and the area buried by the trimer generated by application of the threefold suggest that the trimer present in the crystal is physiological. Since the function of lectins is to bind surface glycans in a multivalent fashion, the proposed trimer of the AAA–fucose complex provides the basis for understanding the function of this protein in innate immunity. Similar arrangement of protomers has been observed in the crystal structure F-lectin from Stripped Bass (MsFPB2; Bianchet, Odom, Amzel, and Vasta, private communication). Fuc-binding sites placed in each FLD of the trimer can crosslink Fuc-carrying groups on the surface of a pathogen separated at least 23 Å, about half of the distance observed in the MBP trimer (45 Å) (Kogan et al., 1995). Furthermore, the high positive charge of the binding site (Figure 2.4B) correlates with the activity against negatively charged liposaccharides present in bacterial membranes.

2.5 IG-TYPE LECTINS (SIGLECS)

One unusual, and may be surprising, family of carbohydrate binding proteins are the Ig-type lectins (Amzel and Poljak, 1979). The carbohydrate-binding module of Ig-type lectins is a single immunoglobulin fold (Ig) domain. Ig domains are the building blocks of the members of the immunoglobulin superfamily, a large family of proteins that includes antibodies, cellular receptors, and cell adhesion molecules. The Ig-fold is a 100–110-residue module comprising 7 (Ig-C domain) or 9 (Ig-V domain) β-strands arranged in two sheets, connected by loops with a consistent topology. A large proportion of Ig-type lectins recognize structures ending in sialic acid and are classified as a subfamily called siglectins (sialic acid Ig-superfamily lectins). We will concentrate our discussion of Ig-type lectins on the siglec subfamily. (For a comprehensive review of this family, see Varki and Angata, 2006.)

Although carbohydrate-recognizing antibodies have been known for a long time (Rovis et al., 1972; Kabat, 1978; Rao et al., 1984; Borden and Kabat, 1988), the mode of carbohydrate recognition by Siglecs does not resemble that of antibodies: in antibodies the binding site is between two Ig-domains, and in Siglecs is formed by amino acids from a single domain (Alphey et al., 2003; Dimasi et al., 2004; Zaccai et al., 2007). Furthermore, while antibodies use the loops connecting the strands to interact with the hapten, Siglecs use side chains of residues present on the outside face of one of the sheets (Figure 2.5).

All Siglecs are extracellular receptors consisting of one Ig-V domain, followed by a variable number (1–16) of Ig-C domains, a transmembrane region consisting of a single transmembrane helix, a short cytoplasmic domain, in some cases two or more immunoreceptor tyrosine-based inhibitory motifs (ITIMs), [I/V]XYXXL). The Ig-V domain contains the carbohydrate-binding site and the ensuing Ig-C domains appear to function as spacers separating the carbohydrate binding site from the membrane. The first domain contains several residues that are well-conserved in Siglecs: an arginine residue that forms a salt bridge with the carboxyl group of sialic acid, two aromatic residues (one near the N-terminal and the other after the conserved arginine), and three cysteines, one of which forms an unusual interdomain disulfide bridge with a cysteine in the second Ig domain.

The Ig-V domain of Siglecs has three features that are less common in other Ig-V domains: an intrasheet disulfide bridge formed by two of the three conserved cysteines that links two adjacent strands, the interdomain disulfide mentioned above, and the H-bonding pattern of the last β-strand—it is split into two halves in such a way that makes H-bonds with strands in both sheets (Figure 2.5).

The binding site is formed mainly by four side chains (Figure 2.5A). In all cases, the conserved arginine makes a doubly H-bonded salt bridge with the carboxylate of the first sialic acid of the carbohydrate (Figure 2.5B) and interacts with the two conserved aromatic residues described above (Figure 2.5A). In the case of Siglec-7 bound to disialic acid these residues are Arg124, Tyr26, and Trp132. Also in Siglec-7 Asn133 is close (3.76 Å) to the first hydroxyl of the glyceryl side chain of the second sialic acid.

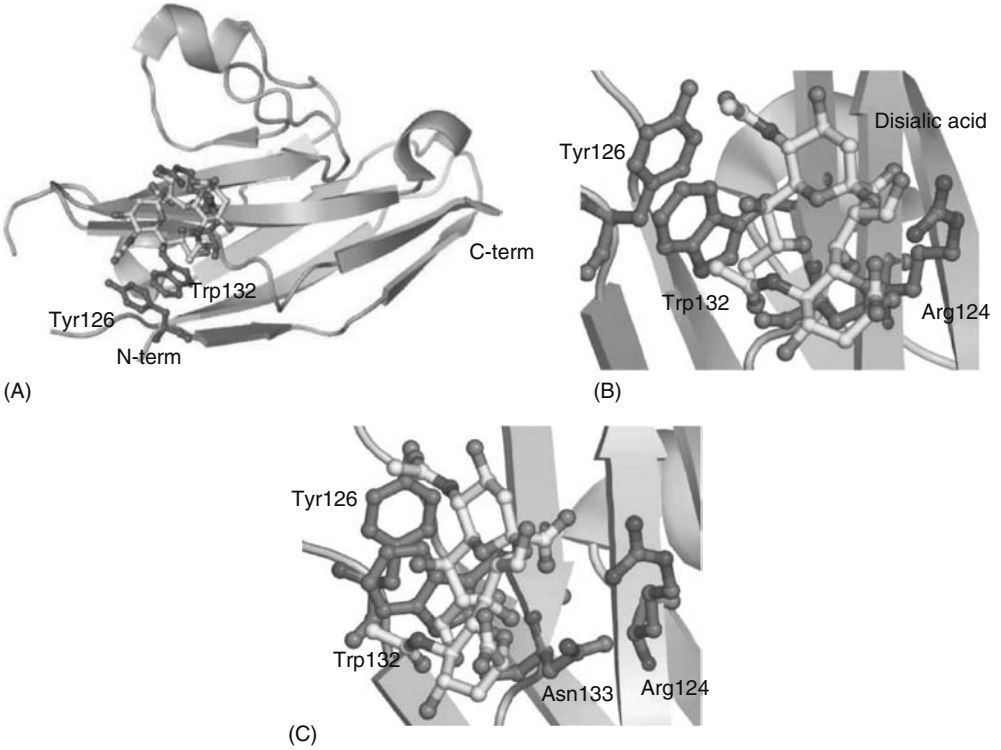


FIGURE 2.5 (See color insert following blank page 170. Also see CD for color figure.) Structure of Siglec-7 (PDB Accession 2HRL). (A) Overall structure. Structure of Selectin-7 showing the two β-sheets and the location of the carbohydrate binding site. Only the disialic portion of the ligand is shown. The bottom strand forms H-bonds with strands in both sheets. The long loop, containing the short helix, shown on the right is not a general feature present in all Siglecs. (B) Disialic acid binding by Siglec-7. Residues that interact with the ligand, Arg124, Tyr26, Trp132, and Asn133, are shown in purple. (C) Carboxylate recognition by Siglec-7. The figure is similar to (B), but it is rotated clockwise to show the recognition of the carboxylate of the first sialic acid by the lectin: Arg124 makes a double H-bonded salt bridge to the negatively charged group of the sugar.

Siglecs are expressed in many cell types including macrophages (spleen, lymph nodes, and bone marrow), B-cells, myeloid cells, Schwan cells, granulocytes, and monocytes (Brewer and Dam, 1999). Specific types of Siglecs are expressed in each of the tissues. In all cases in which it was studied, Siglecs are part of a signaling system that reports to the cells involved in a structure containing the appropriate carbohydrate has been recognized by the lectin. In many cases, the signal may be mediated by cytosolic structures of the Siglecs that carry the ITIM. The ITIM is found not only in Siglecs but also in several receptors that modulate activation signals in lymphocytes. The motif functions by recruiting the inhibitory phosphatases SHP-1, SHP-2, and SHIP that carry one or more SH2 domains that preferentially bind the phosphorylated tyrosines in the ITIM.

2.6 P-TYPE LECTINS

Lysosomal enzymes (LE) carry out the final degradation of most cellular macromolecules. They are synthesized on membrane-bound ribosomes in the endoplasmic reticulum (ER). As many other newly synthesized proteins, LEs travel the ER-Golgi pathway. Obviously, these enzymes must be segregated away from all other glycoproteins and selectively delivered to the lysosomes. Mannose 6-phosphate (Man-6-P) residues in N-linked oligosaccharides of the newly synthesized LE give the signal to direct them to the right organelle. Recognition of these residues is carried out by two receptors in the trans-Golgi network: the cation-dependent Man-6-P receptor (MW: 45 kDa) and the insulin-like

growth factor II (MW: 275 kDa; also called cation-independent Man-6-P receptor) (von Figura and Hasilik, 1986; Hasilik, 1992). Together these receptors constitute the P-lectin family (Drickamer and Taylor, 1993). The elucidation of the pathway mediated by these lectins was the first connection between glycoprotein biosynthesis and human diseases.

2.6.1 Three-Dimensional Structure of P-Type Lectins

The crystal structure of the complex with Man-6-P and Mn^{2+} 155 residue-N-terminal region of CD-MPR was used to define the folds of the P-lectin CRD (Roberts et al., 1998). Recently, the N-terminal 432-residue region of the CI-MPR, encompassing domains 1–3, was determined in the presence of bound mannose 6-phosphate (Olson et al., 2004). The CI-MPR receptor contains two high-affinity carbohydrate recognition sites within its 15-domain extracytoplasmic region, with essential residues for carbohydrate recognition located in domains 3 and 9. Expression of truncated forms of the receptor demonstrated domain 9 alone maintain nanomolar affinity for Man-6-P, whereas domain 3 alone resulted in a protein with three orders of magnitude lower affinity toward a LE.

2.6.2 Man-6-P Binding Site

Similar interactions between protein and Man-6-P are observed in the CD- and in the IGF-II/CI-MPR. The binding site is located in the cleft between the two β-sheets (Figure 2.6A). Centered in cleft, the Man residue is recognized by several interactions with the protein. Arg111 and Tyr143 (CD-MPR numbering) from the five-strand sheet recognize the important 2-OH group of the Man-6-P. Glu133 and Arg135 form a hydrogen bond with the equatorial 4-OH group. Glu133, Tyr143, and Gln66 form hydrogen bonds with the equatorial 3-OH group (Figure 2.6B). The axial 1-OH of the α-Man residue forms a hydrogen bond with Tyr45. No apolar interactions with the carbohydrate are observed. The phosphate is tightly recognized, mainly by the loop between β_6 and β_7 of the wider sheet. Five nitrogen atoms from residues 103–105, three from main chain atoms and two from side-chain atoms are at hydrogen-bonding distance from a phosphate oxygen atom. Nδ of His105 together with Asn104 Nδ2 one each makes hydrogen bonds to two oxygen atoms of the phosphate. His105 is the candidate to provide pH-dependence, required to release the LE in the lower pH of the prelysosomal compartment.

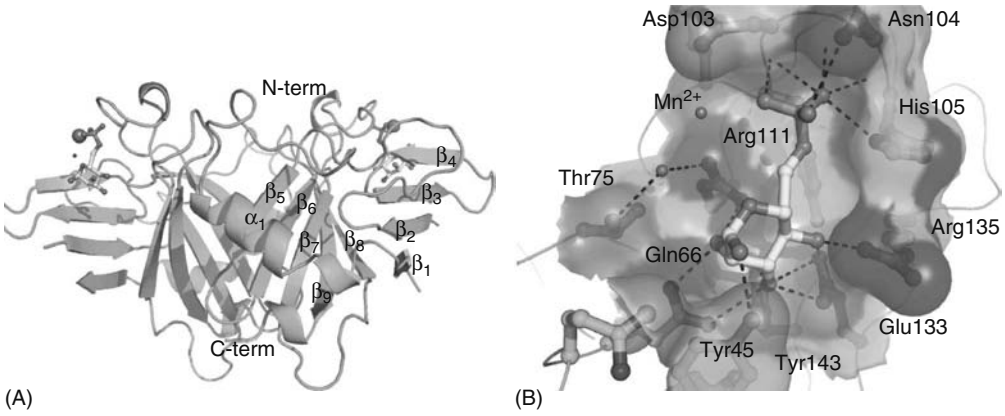


FIGURE 2.6 (See CD for color figure.) Structure of P-lectins. (A) Overall structure. The P-lectin dimer is shown with the twofold axis approximately parallel to the plane of the drawing. The ligand binding site is found between the two sheets of each monomer. (B) Binding of mannose-5-phosphate by P-lectin. The binding site is shown as a transparent surface. The groups interacting with the ligand are visible through the surface. It is interesting to note that the groups interacting with the phosphate are all part of the main chain.

The Mn^{2+} (Figure 2.6A) binds near the N-terminus of the loop between β_6 and β_7 (103–105) close to the phosphate residue. The manganese atom location suggests that the shorter loop in IGF-II/CI-MPR may explain its cation independence (Rini and Lobsanov, 1999) although the observed distances in the crystal structure seems to be longer than the typical Mn–O distance of 2.2Å (Harding, 2006).

2.6.3 Architecture and Quaternary Structure of P-Type Lectins

The CD- and IGF-II/CI-MPR are membrane-bound proteins. CD-MPR has only one N-terminal CRD and a transmembrane domain. The CRD used in the crystallographic studies is a dimer in solution and in the crystal a large apolar interface area with 1550 $Å^2$ per monomer is buried (Rini et al., 1999), suggesting a dimer. The architecture of the extracytoplasmic region of the CI-M6P receptor is a 15-domain structure. Two of these domains, 3 and 9, have the essential residues for carbohydrate recognition (Olson et al., 2004).

2.7 CONCLUSION AND FUTURE DIRECTIONS

The diversity of protein glycosylation, early thought to be paradoxical, is now known to be consistent with the increasing number of functions being discovered for these protein modifications. Structural studies of the interactions between lectins and oligosaccharides are providing keys for the deciphering of 1.44×10^{15} glycome code (compare this number with 67×10^6 peptides or 5×10^3 nucleotides that may be generated with six or fewer monomers) (Laine 1994). The ongoing structural work in animal lectins has blossomed as a dynamic field with the identification of new lectin-encoding genes and subsequent structural studies, which filled gaps in the old lectin families and contributed to the identification of new ones. Several lectin families have members displaying mutimeric tandem arrangements of domains. However understanding the function of this gene-encoded multiplicity is still in an early stage of development and represents one of the many challenges in the field.

REFERENCES

Adams, M. D., S. E. Celniker, et al. 2000. The genome sequence of *Drosophila melanogaster. Science* 287(5461): 2185–2195.

Ahmed, H., H. J. Allen, et al. 1990. Human splenic galaptin: Carbohydrate-binding specificity and characterization of the combining site. *Biochemistry* 29(22): 5315–5319.

Ahmed, H., M. A. Bianchet, et al. 2002. Novel carbohydrate specificity of the 16-kDa galectin from *Caenorhabditis elegans*: Binding to blood group precursor oligosaccharides (type 1, type 2, Talpha, and Tbeta) and gangliosides. *Glycobiology* 12(8): 451–461.

Ahmed, H., S. J. Du, et al. 2004. Biochemical and molecular characterization of galectins from zebrafish (*Danio rerio*): Notochord-specific expression of a prototype galectin during early embryogenesis. *Glycobiology* 14(3): 219–232.

Ahmed, H., N. E. Fink, et al. 1996. Galectin-1 from bovine spleen: Biochemical characterization, carbohydrate specificity and tissue-specific isoform profiles. *J Biochem (Tokyo)* 120(5): 1007–1019.

Ahmed, H. and G. R. Vasta 1994. Galectins: Conservation of functionally and structurally relevant amino acid residues defines two types of carbohydrate recognition domains. *Glycobiology* 4(5): 545–548.

Amzel, L. M. and R. J. Poljak 1979. Three-dimensional structure of immunoglobulins. *Annu Rev Biochem* 48: 961–997.

Baumgartner, S., K. Hofmann, et al. 1998. The discoidin domain family revisited: New members from prokaryotes and a homology-based fold prediction. *Protein Sci* 7(7): 1626–1631.

Bianchet, M. A., H. Ahmed, et al. 2000. Soluble beta-galactosyl-binding lectin (galectin) from toad ovary: Crystallographic studies of two protein–sugar complexes. *Proteins* 40(3): 378–388.

Bianchet, M. A., E. W. Odom, et al. 2002. A novel fucose recognition fold involved in innate immunity. *Nat Struct Biol* 9(8): 628–634.

Borden, P. and E. A. Kabat 1988. The specificities of polyclonal and monoclonal anti-idiotypes to anti-alpha(1—6) dextrans; possible correlations of idiotype with amino acid sequence. *Mol Immunol* 25(3): 251–262.

Bourne, Y., B. Bolgiano, et al. 1994. Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides. *Nat Struct Biol* 1(12): 863–870.

Brewer, C. F. and T. K. Dam 1999. *Essentials of Glycobiology*, Edited by A. Varki, R. Cummins, J. Esko, H. Freeze, G. Hart and J. Marth, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Buz, E.I. et al. 2006. Carbohydrate recognition systems in autoimmunity. *Autoimmunity* 39(8): 691–704.

Cooper, D. N. 2002. Galectinomics: Finding themes in complexity. *Biochim Biophys Acta* 1572(2–3): 209–231.

Dimasi, N., A. Moretta, L. Moretta, R. Biassoni, and R. A. Mariuzza 2004. Structure of the saccharide-binding domain of the human natural killer cell inhibitory receptor p75/AIRMI. *Acta Crystallogr D Biol Crystallogr* 60(Pt 2): 401–403.

Drickamer, K. 1993. Evolution of Ca^{2+}-dependent animal lectins. *Prog Nucleic Acid Res Mol Biol* 45: 207–232.

Drickamer, K. 1999. C-type lectin-like domains. *Curr Opin Struct Biol* 9(5): 585–590.

Drickamer, K., M. S. Dordal, et al. 1986. Mannose-binding proteins isolated from rat liver contain carbohydrate-recognition domains linked to collagenous tails. Complete primary structures and homology with pulmonary surfactant apoprotein. *J Biol Chem* 261(15): 6878–6887.

Drickamer, K. and M. E. Taylor 1993. Biology of animal lectins. *Annu Rev Cell Biol* 9: 237–264.

Emsley, J., H. E. White, et al. 1994. Structure of pentameric human serum amyloid P component. *Nature* 367(6461): 338–345.

Gabius, H. J. 1997. Animal lectins. *Eur J Biochem* 243(3): 543–576.

Gong, H. C., Y. Honjo, et al. 1999. The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells. *Cancer Res* 59(24): 6239–6245.

Harding, M. M. 2006. Small revisions to predicted distances around metal sites in proteins. *Acta Crystallogr D Biol Crystallogr* 62(Pt 6): 678–682.

Hasilik, A. 1992. The early and late processing of lysosomal enzymes: proteolysis and compartmentation. *Experientia* 48(2): 130–151.

Honda, S., M. Kashiwagi, et al. 2000. Multiplicity, structures, and endocrine and exocrine natures of eel fucose-binding lectins. *J Biol Chem* 275(42): 33151–33157.

Horejsi, V. and J. Kocourek 1978. Studies on lectins. XXXVI. Properties of some lectins prepared by affinity chromatography on O-glycosyl polyacrylamide gels. *Biochim Biophys Acta* 538(2): 299–315.

Iobst, S. T. and K. Drickamer 1994. Binding of sugar ligands to Ca^{2+}-dependent animal lectins. II. Generation of high-affinity galactose binding by site-directed mutagenesis. *J Biol Chem* 269(22): 15512–15519.

Kabat, E. A. 1978. Dimensions and specifications of recognition sites on lectins and antibodies. *J Supramol Struct* 8(1): 79–88.

Kelly, C. 1984. Physicochemical properties and N-terminal sequence of eel lectin. *Biochem J* 220(1): 221–226.

Kogan, T. P., B. M. Revelle, et al. 1995. A single amino acid residue can determine the ligand specificity of E-selectin. *J Biol Chem* 270(23): 14047–14055.

Kolatkar, A. R., A. K. Leung, et al. 1998. Mechanism of N-acetylgalactosamine binding to a C-type animal lectin carbohydrate-recognition domain. *J Biol Chem* 273(31): 19502–19508.

Laine, R. A. 1994. A calculation of all possible oligosaccharide isomers both branched and linear yields $1.05 \times 10(12)$ structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4(6): 759–767.

Lasky, L. A. 1995. Selectin–carbohydrate interactions and the initiation of the inflammatory response. *Annu Rev Biochem* 64: 113–139.

Lasky, L. A., M. S. Singer, et al. 1989. Cloning of a lymphocyte homing receptor reveals a lectin domain. *Cell* 56(6): 1045–1055.

Leffler, H. and S. H. Barondes 1986. Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. *J Biol Chem* 261(22): 10119–10126.

Leffler, H., S., Carlsson, et al. 2004. Introduction to galectins. *Glycoconj J* 19(7–9): 433–440.

Leffler, H., F. R. Masiarz, et al. 1989. Soluble lactose-binding vertebrate lectins: A growing family. *Biochemistry* 28(23): 9222–9229.

Leonidas, D. D., B. L. Elbert, et al. 1995. Crystal structure of human Charcot–Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins. *Structure* 3(12): 1379–1393.

Leonidas, D. D., E. H. Vatzaki, et al. 1998. Structural basis for the recognition of carbohydrates by human galectin-7. *Biochemistry* 37(40): 13930–13940.

Liao, D. I., G. Kapadia, et al. 1994. Structure of S-lectin, a developmentally regulated vertebrate beta-galactoside-binding protein. *Proc Natl Acad Sci USA* 91(4): 1428–1432.

Liu, F. T. and G. A. Rabinovich 2005. Galectins as modulators of tumour progression. *Nat Rev Cancer* 5(1): 29–41.

Lobsanov, Y. D., M. A. Gitt, et al. 1993. X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-A resolution. *J Biol Chem* 268(36): 27034–27038.

Meindl, A., M. R. Carvalho, et al. 1995. A gene (SRPX) encoding a sushi-repeat-containing protein is deleted in patients with X-linked retinitis pigmentosa. *Hum Mol Genet* 4(12): 2339–2346.

Ng, K. K., S. Park-Snyder, et al. 1998. Ca^{2+}-dependent structural changes in C-type mannose-binding proteins. *Biochemistry* 37(51): 17965–17976.

Ng, K. K. S., A. R. Kolatkar, et al. 2002. Orientation of bound ligands in mannose-binding proteins. Implications for multivalent ligand recognition. *J. Biol. Chem.* 277(18): 16088–16095.

Nielsen, B. B., J. S. Kastrup, et al. 1997. Crystal structure of tetranectin, a trimeric plasminogen-binding protein with an alpha-helical coiled coil. *FEBS Lett* 412(2): 388–396.

Odom, E. W. and G. R. Vasta 2006. Characterization of a binary tandem domain F-type lectin from striped bass (*Morone saxatilis*). *J Biol Chem* 281(3): 1698–1713.

Olson, L. J., R. D. Yammani, et al. 2004. Structure of uPAR, plasminogen, and sugar-binding sites of the 300 kDa mannose 6-phosphate receptor. *Embo J* 23(10): 2019–2028.

Poget, S. F., G. B. Legge, et al. 1999. The structure of a tunicate C-type lectin from *Polyandrocarpa misakiensis* complexed with D-galactose. *J Mol Biol* 290(4): 867–879.

Poole, S., R. A. Firtel, et al. 1981. Sequence and expression of the discoidin I gene family in *Dictyostelium discoideum*. *J Mol Biol* 153(2): 273–289.

Rao, A. S., J. Liao, E. A. Kabat, E. F. Osserman, M. Harboe, and W. Nimmich 1984. Immunochemical studies on human monoclonal macroglobulins with specificities for 3,4-pyruvylated D-galactose and 4,6-pyruvylated D-glucose. *J Biol Chem* 259(2): 1018–1026.

Rini, J. M. 1995. Lectin structure. *Annu Rev Biophys Biomol Struct* 24: 551–577.

Rini, J. M. and Y. D. Lobsanov 1999. New animal lectin structures. *Curr Opin Struct Biol* 9(5): 578–584.

Roberts, D. L., D. J. Weix, et al. 1998. Molecular basis of lysosomal enzyme recognition: Three-dimensional structure of the cation-dependent mannose 6-phosphate receptor. *Cell* 93(4): 639–648.

Rovis, L., E. A. Kabat, and M. Potter 1972. Immunochemical studies on a mouse myeloma protein having specific binding affinity for 2-acetamido-2-deoxy-D-mannose. *Carbohydr Res* 23(2): 223–227.

Rudd, P. M., T. Elliott, et al. 2001. Glycosylation and the immune system. *Science* 291(5512): 2370–2376.

Saito, T., M. Hatada, et al. 1997. A newly identified horseshoe crab lectin with binding specificity to O-antigen of bacterial lipopolysaccharides. *J Biol Chem* 272(49): 30703–30708.

Saussez, S. and R. Kiss 2006. Galectin-7. *Cell Mol Life Sci* 63(6): 686–697.

Seery, L. T., D. R. Schoenberg, et al. 1993. Identification of a novel member of the pentraxin family in *Xenopus laevis*. *Proc R Soc Lond B Biol Sci* 253(1338): 263–270.

Seetharaman, J., A. Kanigsberg, et al. 1998. X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1-Å resolution. *J Biol Chem* 273(21): 13047–13052.

Shirai, T., Y. Matsui, et al. 2002. Crystal structure of a conger eel galectin (congerin II) at 1.45 Å resolution: Implication for the accelerated evolution of a new ligand-binding site following gene duplication. *J Mol Biol* 321(5): 879–889.

Shirai, T., C. Mitsuyama, et al. 1999. High-resolution structure of the conger eel galectin, congerin I, in lactose-liganded and ligand-free forms: Emergence of a new structure class by accelerated evolution. *Structure* 7(10): 1223–1233.

Taylor, M. E., K. Bezouska, et al. 1992. Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor. *J Biol Chem* 267(3): 1719–1726.

Varela, P. F., D. Solis, et al. 1999. The 2.15 Å crystal structure of CG-16, the developmentally regulated homodimeric chicken galectin. *J Mol Biol* 294(2): 537–549.

Varki, A. and T. Angata 2006. Siglecs–the major subfamily of I-type lectins. *Glycobiology* 16(1): 1R–27R.

Vasta, G. R., H. Ahmed, et al. 2004. Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. *Curr Opin Struct Biol* 14(5): 617–630.

Vasta, G. R., M. Quesenberry, et al. 1999. C-type lectins and galectins mediate innate and adaptive immune functions: their roles in the complement activation pathway. *Dev Comp Immunol* 23(4–5): 401–420.

Vogel, W. 1999. Discoidin domain receptors: Structural relations and functional implications. *FASEB J* 13(Suppl): S77–S82.

von Figura, K. and A. Hasilik 1986. Lysosomal enzymes and their receptors. *Annu Rev Biochem* 55: 167–193.

Walser, P. J., P. W. Haebel, et al. 2004. Structure and functional analysis of the fungal galectin CGL2. *Structure* 12(4): 689–702.

Weis, W. I. 1997. Cell-surface carbohydrate recognition by animal and viral lectins. *Curr Opin Struct Biol* 7(5): 624–630.

Weis, W. I. and K. Drickamer 1996. Structural basis of lectin–carbohydrate recognition. *Annu Rev Biochem* 65: 441–473.

Wendt, K. S., H. C. Vodermaier, et al. 2001. Crystal structure of the APC10/DOC1 subunit of the human anaphase-promoting complex. *Nat Struct Biol* 8(9): 784–788.

Zaccai, N. R., A. P. May, R. C. Robinson, L. D. Burtnick, P. R. Crocker, R. Brossmer, S. Kelm, and E. Y. Jones 2007. Crystallographic and in silico analysis of the sialoside-binding characteristics of the Siglec sialoadhesin. *J Mol Biol* 365(5): 1469–1479.

Zelensky, A. N. and J. E. Gready 2005. The C-type lectin-like domain superfamily. *Febs J* 272(24): 6179–6217.

3 Thermodynamic Approaches to the Study of Affinity of Clustered Carbohydrate Epitopes in Galectin–Glycoconjugate Interactions

Tarun K. Dam and C. Fred Brewer

CONTENTS

3.1 INTRODUCTION

Many biological ligands possess clustered epitopes that bind to specific receptor molecules. However, the effects of clustered epitopes on ligand–receptor interactions are not well understood. For example, Equation 3.1 illustrates binding of monovalent ligand L to a receptor R. Their interactions can be described by affinity constant K_a in Equation 3.2:

$$[L] + [R] \rightleftarrows [LR] \tag{3.1}$$

$$K_a = [LR]/[L][R] \tag{3.2}$$

Binding of tetravalent ligand L4 to receptor R is more complicated. In this case, L is covalently clustered into four epitopes (L4), with each epitope of L4 binding an individual R molecule (no steric effects). Historically, the binding of L4 to R is often described by a single inhibition or binding constant. However, binding of L4 to R cannot be described by Equation 3.1 since four molecules of R bind to one molecule of L4. Therefore, a single binding constant such as that in Equation 3.2 is not correct. Instead, binding of L4 to R must be minimally written as the four microequilibria below:

$$[L4] + [R] \rightleftarrows [L4R] \tag{3.3}$$

$$[L4R] + [R] \rightleftarrows [L4R2] \tag{3.4}$$

$$[L4R2] + [R] \rightleftarrows [L4R3] \tag{3.5}$$

$$[L4R3] + [R] \rightleftarrows [L4R4] \tag{3.6}$$

It follows that there are four microequilibrium binding constants, K_{a1}, K_{a2}, K_{a3}, and K_{a4}, associated with microequilibria Equations 3.3, 3.4, 3.5, and 3.6, respectively. Until recently, the presence of these microequilibria for multivalent ligand–receptor interactions have not been widely described (cf. [1]), nor has the range of values for the microaffinity constants, K_{a1}, K_{a2}, K_{a3}, and K_{a4}, in the present example, been estimated or determined.

This chapter will review recent findings on the binding of multivalent (clustered) carbohydrates and glycoproteins to lectins including galectins in the field of glycobiology. We will show how such microequilibria are detected and what is the range of microaffinity constants for a given multivalent ligand. The results have important implications for not only the binding of clustered carbohydrates to lectins, but also for clustered ligand–receptor interactions throughout biology.

3.2 DETECTION OF MICROEQUILIBRIA IN THE BINDING OF MULTIVALENT CARBOHYDRATES TO LECTINS

Hill plots have been recently used to detect the presence of microequilibria in the binding of multivalent carbohydrates [2] and a glycoprotein [3] to lectins. Traditionally, the Hill plot, in which log {concentration of free ligand} is plotted versus log {fraction of bound protein}/{fraction of free protein} has been used to investigate positive or negative cooperativity in the binding of monovalent ligands to multisubunit proteins in which subunit–subunit interactions in the latter are responsible for the cooperativity effects (cf. [4,5]). A linear Hill plot with a slope of 1.0 indicates no binding cooperativity. Hill plots with slopes greater than 1.0 indicate positive cooperativity and slopes less than 1.0 negative cooperativity. The Hill plot also has the advantage of assigning numerical values to the degree of cooperativity and a logarithmic representation that allows plotting of all obtainable data unlike double reciprocal or half-reciprocal plots that may have open upper limits on the abscissa and ordinate [6].

Recently, Hill plots have been used to investigate the binding of synthetic multivalent analogs **2**, **3** and **4** in Figure 3.1 to the Man-specific lectins concanavalin A (ConA) and *Dioclea grandiflora* lectin (DGL) [2]. Trimannoside **1** in Figure 3.1 binds to both lectins as a monovalent ligand [7]. **2**, **3**, and **4** are di-, tri- and tetravalent structural analogs of **1**, respectively, and thus examples of clustered epitopes in the glycosides. Using isothermal titration calorimetry (ITC), **2**, **3**, and **4** were shown to have 6-, 11- and 35-fold greater observed affinity constants, respectively, for

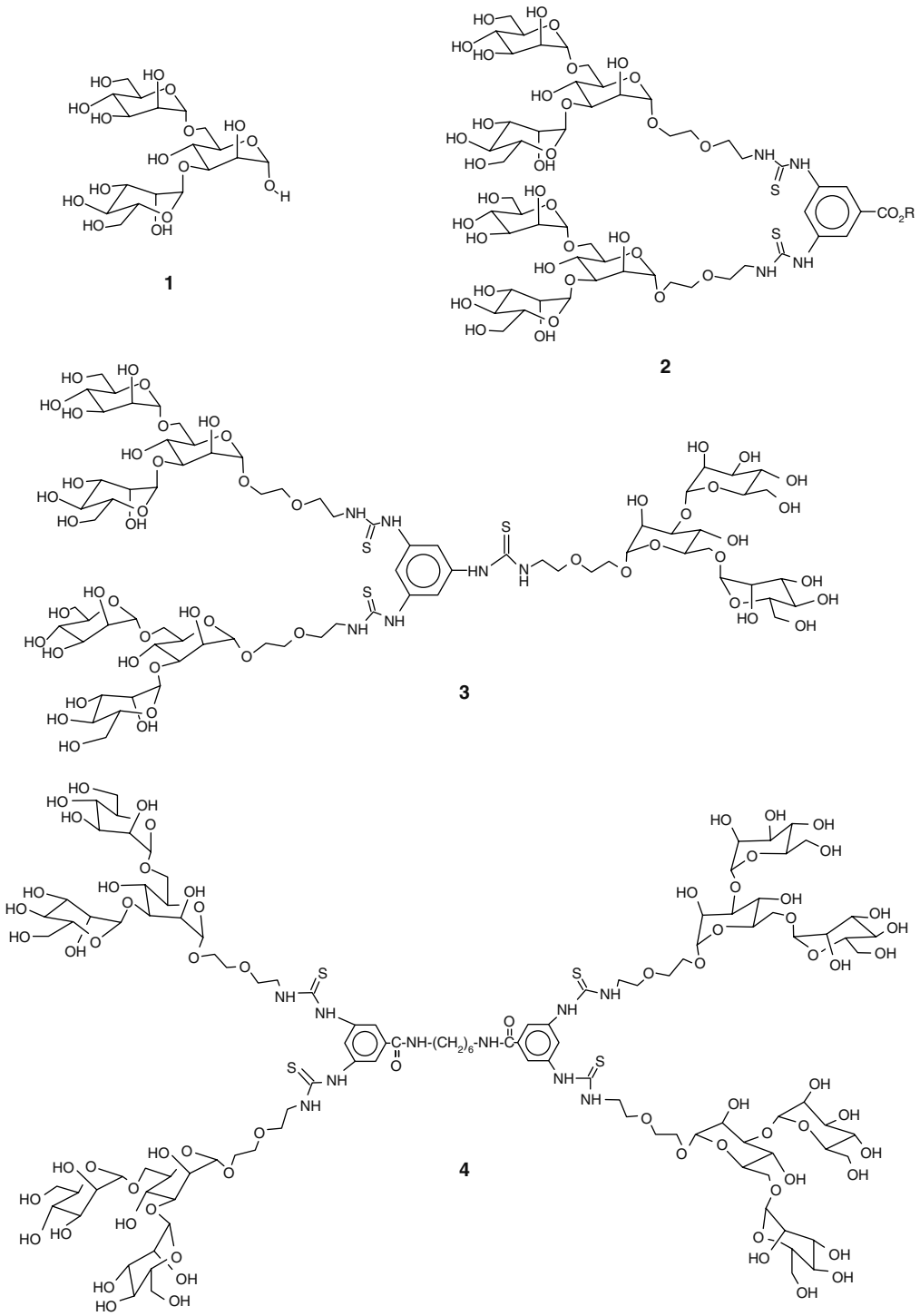


FIGURE 3.1 Structures of trimannoside **1**, and di-, tri-, and tetraantennary analogs **2**, **3**, and **4**, respectively.

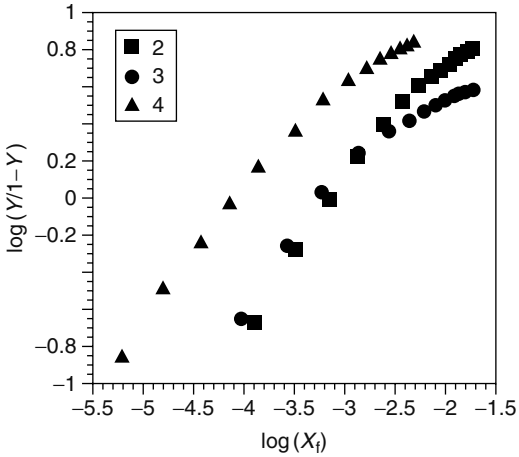


FIGURE 3.2 Hill plots of the ITC data for analogs **2**–**4** binding to ConA (20 mM). The concentration of **2** was 650 mM, **3** was 670 mM, and **4** was 240 mM. Hill plots were constructed by plotting log $\{Y(i)/1-Y(i)\}$ versus log $\{X_f(i)\}$ where $Y(i)$ is $\{X_b(i)\} \times$ (functional valency of ligand)/$Mt(i)$, which is a modified version of the Hill plot that takes into account the functional valency of the ligand. $X_b(i)$ and $X_f(i)$ are the fraction of ligand bound and free, respectively, after the ith injection. (From Di Cera, E., *Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules*, Cambridge University Press, New York, 1995.) The functional valencies of **2** and **3** are two, and the functional valency of **4** is four, as determined from ITC derived n values. (From Dam, T.K., Roy, R., Das, S.K., Oscarson, S., and Brewer, C.F., *J. Biol. Chem*., 275, 14223, 2000.)

binding with ConA, and 5-, 8- and 53-fold greater observed affinities constants, respectively, for binding with DGL. Details of the thermodynamic binding data can be found in the original publication [7].

Hill plots of the raw ITC data for the binding of **2**, **3**, and **4** to ConA are shown in Figure 3.2 [2]. Hill plots for the binding of the analogs to DGL were similar to those observed for ConA [2]. Curvilinear plots are observed in Figure 3.2, indicating increasing negative cooperativity for all three analogs binding to both ConA as well as DGL. The tangent slopes of progressive three-point intervals of the Hill plot for analog **4** binding to ConA is shown in Figure 3.3, and a bar graph of the three-point tangent slopes of the Hill plots of analogs **2**–**4** binding to ConA are shown in Figure 3.4. Importantly, Hill plots of the binding of monovalent analog **1** to both lectins

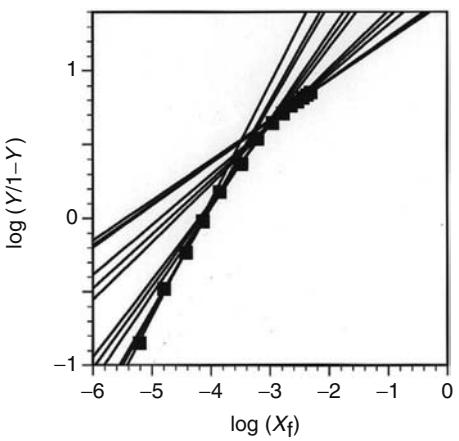


FIGURE 3.3 Tangent slopes of progressive three point intervals of the Hill plot for analog **4** with ConA.

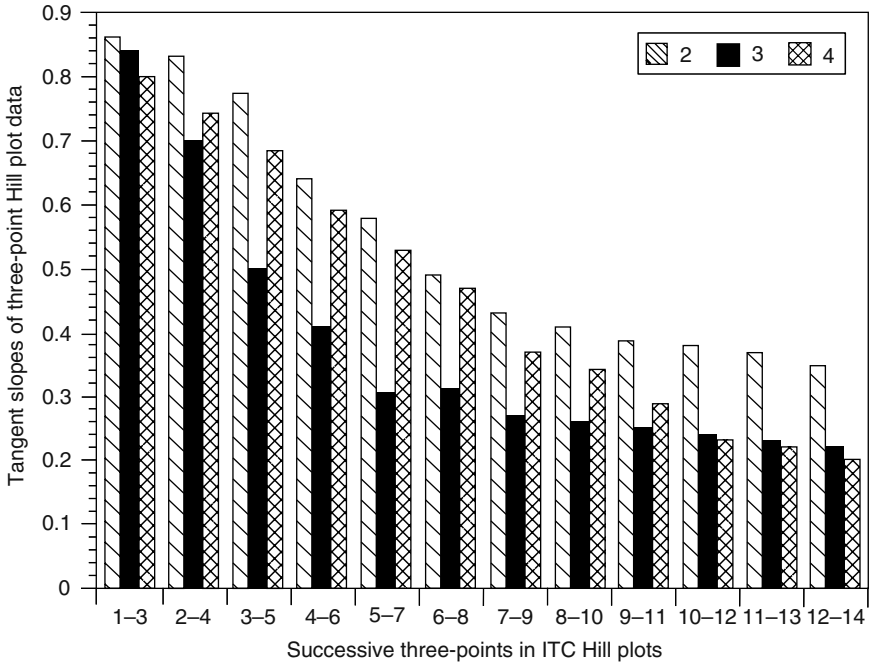
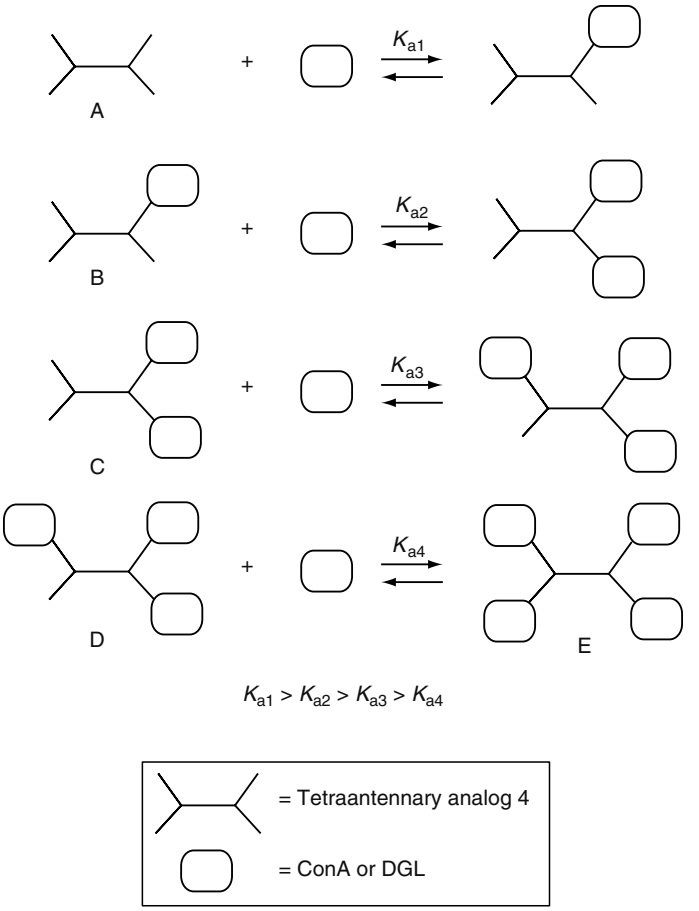


FIGURE 3.4 Bar graphs of the three-point tangent slopes of the ITC data Hill plots of analogs **2–4** binding to ConA.

show linear plots with slopes close to 1.0 (not shown). Thus, the increasing negative cooperativity observed in the Hill plots is due to the multivalent carbohydrates and not lectins.

3.3 PHYSICAL BASIS FOR THE INCREASING NEGATIVE COOPERATIVITY OF ANALOGS 2–4 BINDING TO ConA AND DGL

The physical basis for the increasing negative cooperativity of multivalent carbohydrates **2–4** binding to ConA and DGL is attributed to the decreasing functional valence of the carbohydrates upon progressive binding of their multiple epitopes [2]. This is shown in Scheme 3.1 for binding of tetraantennary analog **4** to ConA. Scheme 3.1 shows the microequilibria constants for **4** as it sequentially binds one, two, three, and four molecules of ConA (or DGL). (Scheme 3.1 is equivalent to Equations 3.3 through 3.6.) The functional valence of unbound **4** (species A) is tetravalent, the functional valence of **4** with one bound lectin molecule (species B) is trivalent, the functional valence of **4** with two bound lectin molecules (species C) is divalent, and the functional valence of **4** with three bound lectin molecules (species D) is monovalent. The increasingly curvilinear Hill plots for **4** binding to ConA (Figure 3.2) and DGL are consistent with the decreasing valence and increasing negative binding cooperativity of **4** with increasing sequential occupancy of the four epitopes of the analog. The same is true for analogs **2** and **3** with ConA and DGL, respectively. The decreasing observed (macroaffinity) constants of **4**, **3**, and **2** for ConA and DGL obtained by ITC measurements support these conclusions [7]. Hence, it is expected that the microaffinity constants in Scheme 3.1 for **4** binding to ConA and [7] DGL have relative values such that $K_{a1} > K_{a2} > K_{a3} > K_{a4}$. Thus, the presence of this decreasing range of microaffinity constants associated with the four epitopes of **4** is consistent with the increasing negative cooperativity observed in the Hill plots of **4** with both lectins.



SCHEME 3.1

3.4 RANGE OF MICROSCOPIC AFFINITY CONSTANTS FOR MULTIVALENT CARBOHYDRATES BINDING TO ConA AND DGL

Based on Scheme 3.1, Equation 3.7 was derived to describe the relationship between the observed macroscopic free energy of binding and the microscopic free energies of binding of the various epitopes of a multivalent carbohydrate binding to a lectin [2]:

$$\Delta G(\text{obs}) = \{\Delta G_1 + \cdots + \Delta G_n\}/n \tag{3.7}$$

Equation 3.7 states that the observed macroscopic ΔG value (ΔG(obs)) of a multivalent carbohydrate is the average of the microscopic $\Delta G(\Delta G_n)$ values of the individual epitopes, where n is the number of epitopes [2]. This equation correctly estimates the difference in microscopic ΔG values of the two epitopes of divalent analog **2** binding to ConA [8]. In this case, Equation 3.8 for the divalent carbohydrate

$$\Delta G(\text{obs}) = \{\Delta G_1 + \Delta G_2\}/2 \tag{3.8}$$

shows that ΔG(obs) from a normal ITC experiment allows calculation of ΔG_1, the first epitope of the divalent carbohydrate, assuming that ΔG_2 for the second epitope is the same as that of a monovalent ligand. This latter assumption was shown to be true from a reverse ITC experiment that

allows direct determination of ΔG_1 and ΔG_2 [8]. The difference between ΔG_1 and ΔG_2 calculated from Equation 3.8 using ΔG(obs) from a normal ITC experiment agreed well with that determined from the reverse ITC [8].

Equation 3.7 can also be used to estimate the spread in microscopic ΔG values for the tri- and tetraantennary carbohydrates that bind to ConA and DGL. Equation 3.9 describes the relationship between the macroscopic ΔG(obs) and four microscopic ΔG values for binding of the tetraantennary analog **4** to DGL:

$$\Delta G(\text{obs}) = \{\Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_4\}/4 \tag{3.9}$$

ΔG_1 in Equation 3.9 is associated with the binding of the first carbohydrate epitope of tetraantennary analog **4**, ΔG_2 with the second carbohydrate epitope, ΔG_3 with the third carbohydrate epitope, and ΔG_4 with the fourth epitope. The macroscopic ΔG(obs) for binding of the **4** to DGL is −10.6 kcal/mol [7], while ΔG_4 in Equation 3.9 can be taken as the ΔG(obs) for binding monovalent analog **1** (Figure 3.1) to DGL, which is −8.3 kcal/mol [7]. Since ΔG(obs) is the average of the four microscopic ΔG values, then

$$\Delta G_1 - \Delta G(\text{obs}) = \Delta G(\text{obs}) - \Delta G_4 \tag{3.10}$$

assuming that $\Delta G(\text{obs}) - \Delta G_2 \sim \Delta G_3 - \Delta G(\text{obs})$ (i.e., there is a symmetrical distribution of microscopic ΔG values on either side of ΔG(obs)). The numerical value of ΔG_1 calculated from Equation 3.10 is −12.9 kcal/mol, which is 4.6 kcal/mol greater than ΔG_4. This difference between ΔG_1 and ΔG_4 translates to a difference in microscopic K_a values, K_{a1} and K_{a4}, of approximately 2800. In absolute terms, K_{a1} is approximately 0.3 nM while K_{a4} is approximately 0.8 μM. Thus, the microscopic K_{a1} of the first unbound epitope of tetraantennary analog **4** binding to DGL is 2800-fold greater than K_{a4} for binding of the fourth epitope. For **4** binding to ConA, this difference between K_{a1} and K_{a4} is nearly 1200-fold [7]. This indicates a decreasing gradient of microscopic binding constants of the four epitopes of **4** binding to ConA and DGL. These differences have been postulated to be due to kinetic effects on the off-rates of the various fractionally bound complexes of the multivalent carbohydrates [7].

3.5 DETECTION OF MICROEQUILIBRIA IN THE BINDING OF ASIALOFETUIN, A MULTIVALENT GLYCOPROTEIN, TO GALECTINS

Negative cooperativity has also recently been shown in the ITC data of asialofetuin (ASF), a naturally occurring 48 kDa glycoprotein that possesses nine *N*-acetyl-D-lactosamine (LacNAc) epitopes (Figure 3.5), binding to galectins-1, -2, -3, -4, -5, and -7, and truncated, monomer versions of galectins-3 and -5, which are members of a family of β-Gal-specific animal lectins [3]. The observed K_a values for ASF binding to the galectins and two truncated forms are 50- to 80-fold greater than that of LacNAc, a monovalent ligand. Since galectins-1 and -7 have been shown by sedimentation velocity and equilibrium data to be dimers in solution [9], and galectin-3 and its truncated version to be predominantly monomers in solution [9], the similarity of K_a values for galectins-1, -3, and -7 together with truncated galectins-3 and -5 indicate that the observed K_a values for ASF are independent of the quaternary structures of the galectins. Sedimentation data for galectins-4 and -5 have not been reported. However, galectin-5 is observed to agglutinate rat erythrocytes [10], suggesting that it may oligomerize in the presence of certain carbohydrate ligands. However, it is a monomer as determined by mass spectrometry [11,12] and gel filtration [10].

Bovine galectin-1 [13] as well as avian liver galectin-1 [14] bind and precipitate with ASF in solution. The stoichiometry of the precipitation complexes is 9:1 galectin/ASF under conditions of excess galectin-1, and 3:1 galectin/ASF with increasing ASF concentrations. The 9:1 galectin/ASF cross-linked complex is consistent with binding of galectin-1 to all nine LacNAc epitopes of ASF in

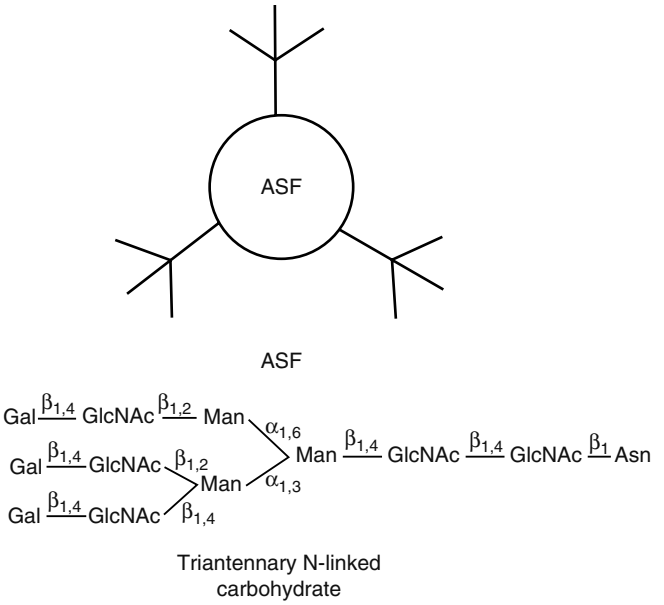


FIGURE 3.5 Schematic representation of ASF and its asparagine (N-) linked triantennary carbohydrate chain found at three positions on ASF.

the cross-linked complexes [13,14]. The ITC-derived *n* values of ASF binding to galectins-1, -2, -3, -4, -5, and -7, as well as truncated galectins-3 and -5, in the ITC experiments are all close to 0.1 [3], which are consistent with binding of all nine LacNAc epitopes of ASF to the galectins in the ITC experiments.

Hill plots of the raw ITC data for LacNAc binding to galectins-1, -2, -3, -4, -5, and -7 as well as truncated galectins-3 and -5 are straight lines with slopes close to 0.93 [3]. These values are close to a value of 1.0 for noncooperative binding interactions [4,5], indicating that the galectins do not undergo cooperative binding interactions with the monovalent disaccharide.

The Hill plot of the ITC data for ASF binding to galectin-3 is shown in Figure 3.6 (open squares) [3]. The plot is curvilinear rather than linear and disposed around the zero point on the ordinate, as observed for monovalent LacNAc, after correction for the functional valence of ASF. The tangent slopes of progressive three-point intervals of the *x*-axis of the Hill plot for ASF is shown in Figure 3.7, which shows decreasing tangent slopes along the binding curve. Figure 3.8 shows a bar graph comparison of the three-point tangent slopes of ASF binding to galectin-3. The initial tangent slope value is close to 1.0; the final tangent slope value is approximately 0.15. These results show increasing negative cooperativity with increasing binding of ASF to galectin-3.

The Hill plot of the ITC data for ASF binding to truncated galectin-3 is shown in Figure 3.6 (closed circles) [3]. The plot is also curvilinear and disposed around the zero point on the ordinate. Figure 3.8 shows a bar graph comparison of the three-point tangent slopes of ASF binding to truncated galectin-3. The initial tangent slope value is close to 1.0; the final tangent slope value is approximately 0.20. These results show increasing negative cooperativity with increasing binding of ASF to truncated galectin-3.

Hill plots similar to those for binding of ASF to galectin-3 and truncated galectin-3 were observed for ASF binding to galectins-1, -2, -4, -5, and -7 as well as truncated galectin-5 were also observed (not shown) [3]. These results indicate that increasing negative cooperativity occurs for ASF binding to all of the above galectins that represent all three subfamilies of galectins. Furthermore, the negative cooperativity is not due to the quaternary structures of the galectins since dimeric

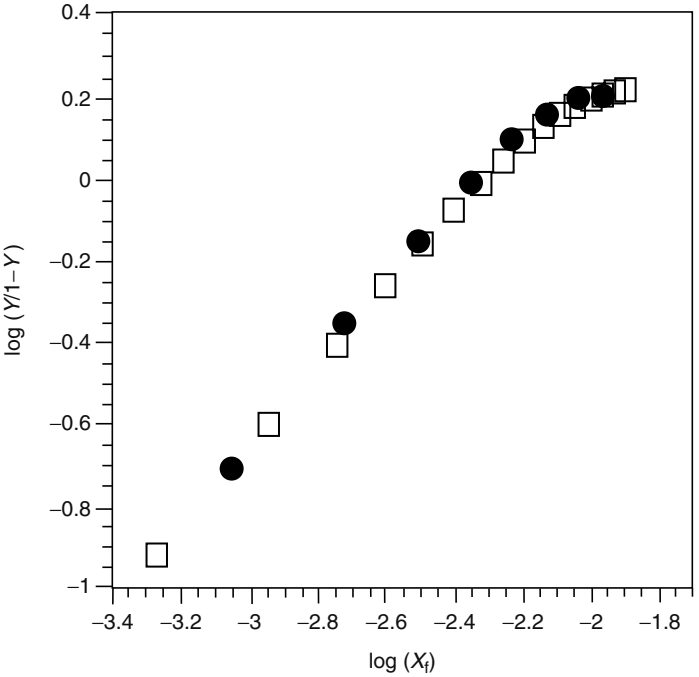


FIGURE 3.6 Hill plots of ITC data for ASF (330 μM) binding to human galectin-3 (34 μM) and ASF (640 μM) binding to truncated galectin-3 (45 μM). Plot definitions are given in the legend of Figure 3.2. The functional valency of ASF for the galectins is nine as determined by ITC. (From Dam, T.K., Gabius, H.-J., Andre, S., Kaltner, H., Lensch, M., and Brewer, C.F., *Biochemistry*, 44, 13564, 2005.)

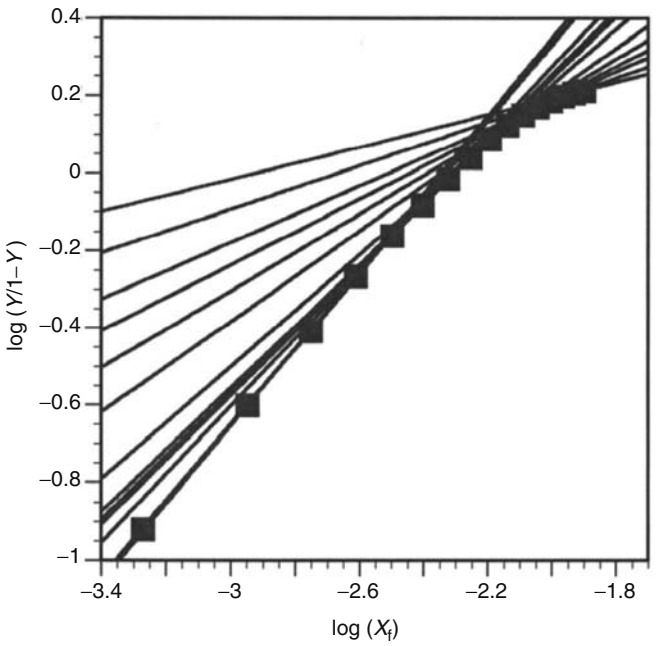


FIGURE 3.7 Tangent slopes of progressive three-point intervals of the Hill plot for ASF with human galectin-3.

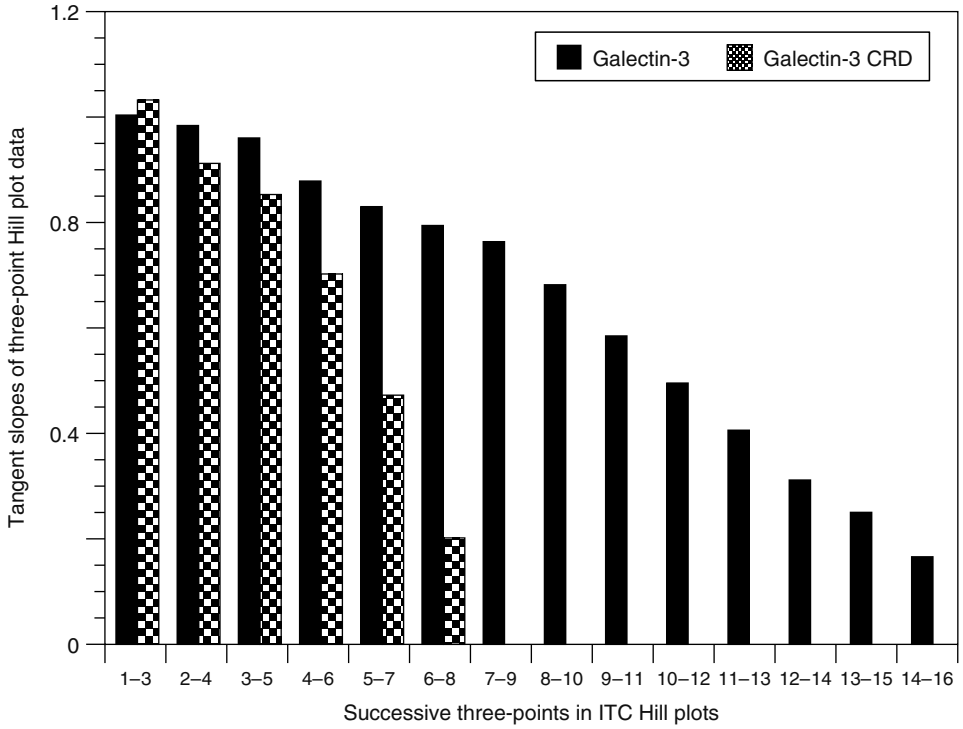


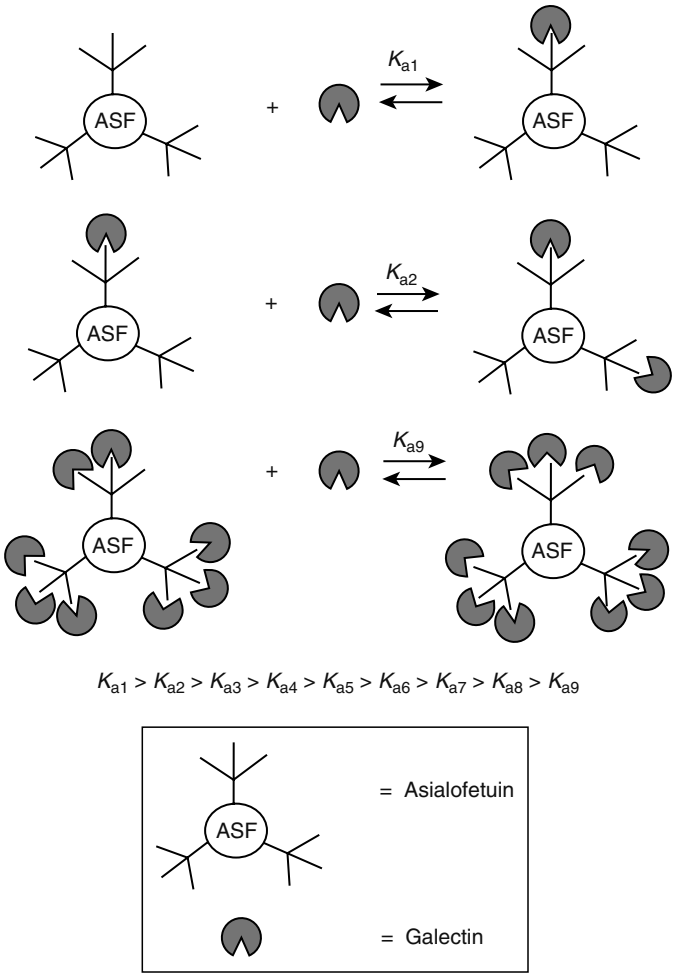
FIGURE 3.8 Bar graphs of the three-point tangent slopes of the ITC data Hill plots of ASF binding to galectin-3 and truncated galectin-3.

galectins such as galectins-1, -2, and -7 show similar results as found for monomeric galectins including galectin-3 and truncated galectins-3 and -5.

3.6 PHYSICAL BASIS OF THE NEGATIVITY COOPERATIVITY OF ASF BINDING TO THE GALECTINS

The physical basis of the negativity binding cooperativity of ASF to the galectins appears to be similar to that observed for the binding of the di-, tri-, and tetraantennary carbohydrates (**2**, **3**, and **4)** in Figure 3.1 to ConA and DGL [2]. The increasing negative cooperativity observed in the present study is due to the reduction in functional valence of ASF as it binds an increasing number of galectin molecules. Scheme 3.2 shows the various microequilibria constants for ASF as it sequentially binds one, two, and up to nine molecules of a galectin. Unbound ASF is functionally nonavalent, with subsequent reductions in its valence with further binding of galectin molecules until the ninth galectin molecule binds to a functionally monovalent ASF molecule. The increasingly curvilinear Hill plots in Figure 3.6 for ASF binding to galectin-3 and its truncated form, respectively, are consistent with the decreasing functional valence of ASF with increasing binding of galectin molecules. The same is true for ASF binding to the other galectins.

Another physical factor that was suggested to play a role in the curvilinear Hill plots of **2**, **3**, and **4** binding to ConA and DGL [7] was the formation of noncovalent cross-linked complexes between lectin molecules and the carbohydrates. However, the present study shows that ASF possesses essentially the same degree of negative binding cooperativity with truncated galectins-3 and -5 as the full-length molecules. Since truncated galectin-3 has been shown to be a monomer in solution [9], and truncated galectin-5 is expected to be a monomer, their negative binding cooperativity with



SCHEME 3.2

ASF is consistent with the reduction in functional valence of ASF upon binding of the carbohydrate recognition domains (CRDs) of these galectins. Hence, the formation of noncovalent cross-linked complexes between ASF and multisubunit galectins such as galectins-1, -2, and -7 does not appear to significantly contribute to their observed negative binding cooperativity.

3.7 RANGE OF MICROSCOPIC K_a VALUES FOR ASF BINDING TO THE GALECTINS

The observed K_a values for ASF binding to the galectins are the average of the microscopic binding free energy terms ($-\Delta G$) of the different epitopes of ASF. Hence, Equation 3.3 can be written as Equation 3.7 for ASF binding to the galectins, in which there are nine microscopic ΔG values representing the nine microequilibria as shown in Scheme 3.2.

$$\Delta G(\text{obs}) = \{\Delta G_1 + \Delta G_2 + \cdots + \Delta G_9\}/9 \tag{3.7}$$

The nine microequilibrium constants of ASF are represented by K_{a1}, K_{a2}, …. and K_{a9} for binding of the first LacNAc epitope of ASF, the second epitope of ASF, and the last unbound LacNAc epitope

of ASF to a galectin molecule, respectively. Hence, the observed ΔG values (ΔG(obs)) of ASF for the galectins are the average of the nine microscopic ΔG terms as shown in Equation 3.7 for each galectin. The relative values of ΔG_1, ΔG_2, …, ΔG_9 decrease based on the decreasing microscopic ΔG values of the epitopes of multivalent analogs with decreasing functional valencies, as observed for multivalent carbohydrates binding to ConA and DGL [7]. In kinetic terms, the microscopic off-rate (k_{-1}) for K_{a1} in Scheme 3.2 ($K_{a1} = k_1/k_{-1}$) would be expected to be slower than the microscopic off-rate for K_{a2}, etc. due to binding and recapture of the first bound galectin molecule by the remaining unbound LacNAc residues of ASF before full dissociation of the complex. Since the ITC-derived ΔG(obs) value is the average of the nine microscopic ΔG values in Equation 3.7, and assuming that there is a symmetrical distribution of decreasing microscopic ΔG values on either side of the average ΔG value (that is, ΔG_8–$\Delta G_7 \sim \Delta G_3$–$\Delta G_2$, etc.), then the value of ΔG(obs) is nearly equal to ΔG_5 in Equation 3.7. It follows then that if ΔG_9, which represents binding of the last free epitope of ASF to a galectin, is nearly equal to ΔG for LacNAc binding to a galectin, then the difference between ΔG(obs) and ΔG_9 is half the difference between ΔG_1 and ΔG_9. Thus, the 2.8 kcal/mol increase in ΔG(obs) for ASF binding to galectin-3 versus ΔG for LacNAc binding [3] indicates that the difference between ΔG_1 and ΔG_9 for ASF binding to galectin-3 is ~5.6 kcal/mol. Since 2.8 kcal/mol is a 78-fold increase in affinity of ASF for galectin-3 over LacNAc, the difference in affinity of the first unbound LacNAc epitope of ASF for galectin-3 is 78 × 78 or ~6000-fold increase in affinity over LacNAc, with a decreasing gradient of affinities down to that of LacNAc for the last (9th) unbound epitope of ASF.

Using the same data analysis, the estimated increase in affinity of the first unbound LacNAc epitope of ASF for truncated galectin-3 versus LacNAc is nearly 3000-fold, demonstrating that the truncated monomeric form of the galectin also shows similar enhanced affinities for the first unbound epitope of ASF. In fact, all of the galectins including truncated galectin-5 show K_a(obs) values that are 50 to 78-fold greater than that of LacNAc for the respective galectin [3]. This indicates that the first unbound epitope of ASF binds to all of the galectins with between 3000- and 6000-fold higher affinity than LacNAc and the last unbound epitope of ASF, and therefore, that there is a gradient of decreasing affinities for the remaining epitopes of ASF for all of the galectins. In terms of numerical affinity constants, a 6000-fold increase in affinity of galectin-3 for the first unbound epitope of ASF is equivalent to a 10 nM affinity constant using the K_a value for LacNAc as an estimate for K_{a9} in Scheme 2. This estimated range of nM affinity constants for K_{a1} is typical for all of the galectins binding to the first unbound epitope of ASF.

It is important to point out that the above estimates of the range of affinity constants for ASF binding to the galectins depends, in part, on the assumed affinity of the galectins binding to the last unbound epitope of ASF. In the above calculations, that affinity constant was estimated to be similar to that of the corresponding monovalent ligand, LacNAc. However, if the affinity of binding the galectins to the last unbound epitope of ASF is lower than LacNAc due to steric crowding or other mechanisms, then the estimated range of enhanced affinities sites on ASF (the first unbound LacNAc epitopes) would be even greater. Likewise, a nonsymmetrical distribution of decreasing binding constants associated with the nine LacNAc epitopes of ASF would also affect the range of estimated affinity constants for the epitopes. Nevertheless, a large range of decreasing microaffinity constants exists for the nine epitopes of ASF binding to the galectins, and the estimates for this range in the present study are reasonable given the assumptions for such calculations.

3.8 IMPLICATIONS OF LARGE GRADIENTS OF MICROSCOPIC AFFINITY CONSTANTS OF MULTIVALENT CARBOHYDRATES AND GLYCOPROTEINS FOR LECTINS INCLUDING GALECTINS

The observation that tetravalent carbohydrate analog **4** and nonavalent ASF bind to different lectins with a similar range of microaffinity constants has important implications in glycobiology. Tetravalent analog **4** binds to DGL with an estimated microaffinity constant for its first epitope that is nearly

3000-fold greater than its fourth epitope. This can be compared to the estimated microaffinity constant of the first epitope of nonavalent ASF binding to the galectins that is 3000- to 6000-fold greater than the ninth epitope of ASF [3]. Hence, it appears that a maximum increase in microaffinity constants is approached with approximately four epitopes of a multivalent carbohydrate binding to lectins. This will have to be tested systematically with higher carbohydrate valence analogs to DGL and lower valence analogs to the galectins. However, these findings may have general implications for the design of multivalent carbohydrate inhibitors for lectins.

The higher valency of ASF compared to tetravalent analog **4** also has important implications for the cross-linking activities of lectins to multivalent glycoprotein receptors [3]. Lectin-mediated cross-linking of glycoprotein receptors on the surface of cells is often associated with signal transduction effects including apoptosis [15]. For example, many galectins including galectin-1 are associated with proapoptotic activities i n cells [16]. The presence of a gradient of microaffinity constants on glycoprotein receptors such as ASF would result in relatively low concentrations of galectins binding to only a few high-affinity sites on the receptors. In fact, as few as three galectin-1 molecules bound to ASF lead to homogeneous cross-linking of the molecules into large insoluble aggregates [17]. This contrasts with the requirement of binding to nearly all of the epitopes of lower valence molecules such as bi-, tri-, and tetravalent carbohydrates (**2**–**4** in Figure 3.1) for cross-linking, which requires much higher concentrations of lectins to occupy the higher and lower affinity epitopes in such molecules (cf. [8]).

In this regard, CD43 is one of several galectin-1 counter receptors on the surface of T cells [18], and possesses approximately 80 serine/threonine (O-) linked chains with terminal LacNAc epitopes [19]. Thus, CD43 is expected to possess a large gradient of microscopic binding constants for galectin-1 as well as other galectins. Only a fraction off high-affinity epitopes on CD43 may need to be bound by galectin-1 in order to cross-link the receptor and induce apoptosis in susceptible T cells [18]. Furthermore, binding and cross-linking of glycoprotein receptors like CD43, which may possess individual affinity sites on each glycoprotein of ~10 nM on the surface of a cell by a dimeric galectin such as galectin-1 would result in an overall avidity of the galectin-1 of ~10^{16} M. Hence, cross-linking by a dimeric galectin would be essentially irreversible under these conditions. Glycoprotein receptors with relatively few carbohydrate chains and hence lower valency would be expected to form less stable cross-linked complexes with galectins. Regulating the total carbohydrate valence of a glycoprotein receptor may influence its cross-linking activities with lectins and hence its biological activities including receptor cycling kinetics and cell surface receptor density.

The present observations also apply to other types of lectins interacting with clustered glycan receptors. Fractional high-affinity binding of lectins to multivalent receptors could lead to supramolecular assemblies of homogeneous cross-linked receptors [20] or heterogeneous cross-linked receptors [21]. Such assemblies, in turn, can trigger cell surface signal transduction mechanisms similar to those observed for galectin-1 on susceptible T cells [16].

3.9 IMPLICATIONS FOR OTHER MULTIVALENT BINDING SYSTEMS

In general, the exclusive use of macroscopic binding parameters such as observed affinity or inhibition constants to describe binding of a multivalent ligand to a receptor fails to reveal the underlying binding cooperativity and range of microscopic affinity constants of the interactions. This is important since at low concentrations a multivalent ligand is expected to bind more tightly to a receptor than predicted by its macroscopic affinity constant. For example, tetravalent analog **4** binds to DGL with a macroscopic affinity constant that is 50-fold greater than that of the corresponding monovalent analog (cf. [7]). However, the microaffinity constant of the first epitope of **4** is estimated to be nearly 60-fold greater than its macroscopic affinity constant. These observations suggest that multivalent analogs of drugs with specific target receptors including enzymes may possess enhanced activities at lower concentrations than predicted by their macroscopic affinity or inhibition constants.

The present findings are also relevant to other types of biological molecules with clustered epitopes and binding domains. These include cell adhesion molecules such as the integrins and cadherens (cf. [22]), and molecules that function in immunity including the so-called immunological synapse of the T cell receptor [23] and antigen–antibody interactions [24]. Multifunctional proteins including extracellular matrix proteins with repeat domains are also candidates for enhanced microaffinity gradients with specific ligands or counter-receptors [25]. Glycosaminoglycans with clustered carbohydrate epitopes may also possess enhanced microaffinity gradients toward lectin-like molecules including growth factors [26]. Interestingly, mucins like CD43, a counter receptor for galectin-1, would also be included in these effects.

The so-called "spare receptors" as described in the pharmacological literature [27] are also candidates for microaffinity gradient interactions with ligands. Maximum dose–activity responses are often observed at relatively low fractional occupancy of such receptors. "Spare receptors" such as those that occur in neural synapses may exhibit enhanced affinity for a specific ligand through clustering in a manner similar to that observed for ASF binding to galectins. The enhancement in affinity of a ligand could be as much as 1000- to 10,000-fold for clustered receptors. Occupancy of a portion of the clustered or "spare receptors" by antagonist would diminish the total number of unbound receptors, but may not reduce the "avidity" of the remaining clustered receptors [27]. The "efficacy" of the agonist effect may be related to cross-linking or activating in some manner a fraction of the receptors that is necessary for full pharmacology effect. Indeed, CD43 can be considered an example of a "spare receptor" in that molecules of CD43 on the surface of T cells contain many more carbohydrate epitopes than required for binding and cross-linking by galectin-1 [18].

3.10 CONCLUSION AND FUTURE DIRECTIONS

The results discussed in this review represent one general mechanism for the increased avidity of multivalent molecules for ligands and counter-receptors. An important second mechanism involves extended site binding of a receptor for a multivalent ligand, such as occurs between branched chain carbohydrates and the asialoglycoprotein receptor [28]. Together, both mechanisms represent important means of increasing the binding affinity and specificity of multivalent molecules in biological systems.

ACKNOWLEDGMENTS

This work was supported by Grant CA-16054 from the National Cancer Institute, Department of Health, Education and Welfare, and Core Grant P30 CA-13330 from the same agency C. Fred Brewer (C.F.B).

REFERENCES

1. Cera, E. D. 1998. Site-specific thermodynamics: Understanding cooperativity in molecular recognition. *Chemical Reviews 98*, 1563–1591.
2. Dam, T. K., Roy, R., Pagé, D., and Brewer, C. F. 2002. Negative cooperativity associated with binding of multivalent carbohydrates to lectins. Thermodynamic analysis of the "multivalency effect." *Biochemistry 41*, 1351–1358.
3. Dam, T. K., Gabius, H.-J., Andre, S., Kaltner, H., Lensch, M., and Brewer, C. F. 2005. Galectins bind to the multivalent glycoprotein asialofetuin with enhanced affinities and a gradient of decreasing binding constants. *Biochemistry 44*, 13564–12571.
4. Stryer, L. 1988. *Biochemistry*, 3rd ed., W. H. Freedman and Company, New York.
5. Di Cera, E. 1995. *Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules*, Cambridge University Press, New York.
6. Weber, G. and Anderson, S. 1965. Multiplicity of binding. Range of validity and practical test of Adair's equation. *Biochemistry 4*, 1942–1947.

7. Dam, T. K., Roy, R., Das, S. K., Oscarson, S., and Brewer, C. F. 2000. Binding of multivalent carbohydrates to concanavalin A and *Dioclea grandiflora* lectin. Thermodynamic analysis of the "multivalency effect." *The Journal of Biological Chemistry 275*, 14223–14230.
8. Dam, T. K., Roy, R., Pagé, D., and Brewer, C. F. 2002. Thermodynamic binding parameters of individual epitopes of multivalent carbohydrates to concanavalin A as determined by "reverse" isothermal titration microcalorimetry. *Biochemistry 41*, 1359–1363.
9. Morris, S., Ahmad, N., Andre, S., Kaltner, H., Gabius, H.-J., Brenowitz, M., and Brewer, F. 2004. Quaternary solution structures of galectins-1, -3 and -7. *Glycobiology 14*, 293–300.
10. Gitt, M. A., Wiser, M. F., Leffler, H., Herrmann, J., Xia, Y.-R., Massa, S. M., Cooper, D. N. W., Lusis, A. J., and Barondes, S. H. 1995. Sequence and mapping of galactin-5, a β-galactoside-binding lectin, found in rat erythrocytes. *The Journal of Biological Chemistry 270*, 5032–5038.
11. Andre, S., Kaltner, H., Lensch, M., Russwurm, R., Siebert, H. C., Fallsehr, C., Tajkhorshid, E., Heck, A. J. R., von Knebel-Doberitz, M., Gabius, H.-J., and Kopitz, J. 2005. Determination of structural and functional overlap/divergence of five proto-type galectins by analysis of the growth-regulatory interaction with ganglioside GM1 *in silico* and in vitro on human neuroblastoma cells. *International Journal of Cancer 114*, 46–57.
12. Kopitz, J., Andre, S., von Reitzenstein, C., Versluis, K., Kaltner, H., Pieters, R. J., Wasano, K., Kuwabara, I., Liu, F.-T., Cantz, M., Heck, A. J. R., and Gabius, H.-J. 2003. Homodimeric galectin-7 (p53-induced gene 1) is a negative growth regulator for human neuroblastoma cells. *Oncogene 22*, 6277–6288.
13. Mandal, D. K. and Brewer, C. F. 1992. Cross-linking activity of the 14-kilodalton β-galactose-specific vertebrate lectin with asialofetuin: Comparison with several galactose-specific plant lectins. *Biochemistry 31*, 8465–8472.
14. Gupta, D., Kaltner, H., Dong, X., Gabius, H.-J., and Brewer, C. F. 1996. Comparative cross-linking activities of lactose-specific plant and animal lectins and a natural lactose-binding immunoglobulin G fraction from human serum with asialofetuin. *Glycobiology 6*, 843–849.
15. Perillo, N. L., Pace, K. E., Seilhamer, J. J., and Baum, L. G. 1995. Apoptosis of T cells mediated by galectin-1. *Nature 378*, 736–739.
16. Perillo, N. L., Uittenbogaart, C. H., Nguyen, J. T., and Baum, L. G. 1997. Galectin-1, an endogenous lectin produced by thymic epithelial cells, induces apoptosis of human thymocytes. *Journal of Experimental Medicine 185*, 1851–1858.
17. Gupta, D. and Brewer, C. F. 1994. Homogeneous aggregation of the 14-kDa β-galactoside specific vertebrate lectin complex with asialofetuin in mixed systems. *Biochemistry 33*, 5526–5530.
18. Pace, K. E., Lee, C., Stewart, P. L., and Baum, L. G. 1999. Restricted receptor segregation into membrane microdomains occurs on human T cells during apoptosis induced by galectin-1. *The Journal of Immunology 163*, 3801–3811.
19. Daniels, M. A., Hogquist, K. A., and Jameson, S. C. 2002. Sweet "n" sour: The impact of differential glycosylation on T cell responses. *Nature Immunology 3*, 903–910.
20. Brewer, C. F. 1997. Cross-linking activities of galectins and other multivalent lectins. *Trends in Glycoscience and Glycotechnology 9*, 155–165.
21. Ahmad, N., Gabius, H.-J., André, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F., and Brewer, C. F. 2004. Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *Journal of Biological Chemistry 279*, 10841–10847.
22. van Kooyk, Y. and Figdor, C. G. 2000. Avidity regulation of integrins: The driving force in leukocyte adhesion. *Current Opinion in Cell Biology 12*, 542–547.
23. Krummel, M. F., Sjaastad, M. D., Wulfing, C., and Davis, M. M. 2000. Differential clustering of CD4 and CD3d during T cell recognition. *Science 289*, 1349–1352.
24. Nussbaum, G., Cleare, W., Casadevall, A., Scharff, M. D., and Valadon, P. 1997. Epitope location in the *Cryptococcus neoformans* capsule is a determinant of antibody efficacy. *Journal of Experimental Medicine 185*, 685–694.
25. Engel, J. 1991. Common structural motifs in proteins of the extracellular matrix. *Current Opinion in Structural Biology 3*, 779–785.
26. Krufka, A., Guimond, S., and Papraeger, A. C. 1996. Two hierarchies of FGF-2 signaling in heparin: Mitogenic stimulation and high-affinity binding/receptor transphosphorylation. *Biochemistry 35*, 11131–11141.
27. Goldstein, A., Aronow, L., and Kalman, S. M. 1974. *Principles of Drug Action*. John Wiley & Sons, New York, pp. 101–104.
28. Lee, Y. C. 1993. Biochemistry of carbohydrate-protein interaction. *The FASEB Journal 6*, 3193–3200.

4 Deciphering Lectin Ligands through Glycan Arrays

David F. Smith and Richard D. Cummings

CONTENTS

4.1 INTRODUCTION TO GLYCOMICS—FUNCTIONAL AND OTHERWISE

A major emphasis of modern molecular biology and biochemistry has been on the genome and its protein products. Genes and proteins are linear macromolecules and with evolution of powerful methods including DNA synthesis, polymerase chain reaction, and peptide synthesis, these molecules became relatively easy to study. This situation is not the case with complex carbohydrates, which have multiple sites for substitution, are often branched, and are difficult to be synthesized. As major biological functions of complex carbohydrates are being recognized, modern studies on glycobiology and the emerging field of glycomics have begun to define the structural interactions between glycans and glycan-binding proteins (GBPs) (also called carbohydrate-binding proteins or lectins). Modern glycomics is a term coined in the late 1990s by Professor Vern Reinhold following the neologisms proteome/proteomics [1] and may be defined as the constellation of glycan structures (oligo- and polysaccharides) synthesized by cells and found in glycoproteins, glycolipids, and proteoglycans and in free glycans, e.g., milk and other fluids. Functional glycomics is a term employed by the consortium for Functional Glycomics to denote the exploration of the recognition and interactions of the glycome in biological systems http://www.functionalglycomics.org/static/index.shtml.

Among the major questions that functional glycomics must address are: How many glycans are synthesized by a single type of cell? How many different GBPs exist in humans or other animals? A lot of evidence is emerging to show that many GBPs are involved in innate immunity and regulation of the adaptive immune response. How are self and nonself glycans distinguished? What is the affinity of interaction between glycans and GBPs?

To address these types of questions, new methods and approaches are required. The remarkable advances that have been made in mass spectroscopy of glycoconjugates over the past 10–15 years have provided direct and sensitive methods for sequencing and sizing microquantities of glycans. These methods are beginning to reveal the vastness of the glycomes of human and other organisms. Glycan microarrays are a more recently developed novel tool for functional glycomics. Glycan microarrays represent multiple glycan probes that are either noncovalently or covalently attached to a microchip, a glass slide, or microsphere-sized beads, and allow simultaneous and rapid evaluation of glycan recognition by a candidate receptor or GBP. This chapter focuses on the development of glycan microarrays and how they are being used to decipher the recognition determinants of various GBPs in pure form or as present in viruses and bacteria, as well as for evaluating the specificity of antibody binding to glycans. Following a historical overview, we will mainly focus on printed glycan microarrays, in which glycans are covalently linked to glass slides analogous to DNA and protein microarrays.

4.2 HISTORICAL BACKGROUND ON GLYCAN MICROARRAYS

Glycan microarrays evolved from many early studies on lectins, antibodies, and organisms that interact with glycoconjugates. Plant lectins, which initiated studies of glycan recognition [2], were discovered over 100 years ago by Stillmark (1889) and Hellin (1891). Antibodies in serum to different blood group antigens were discovered also over 100 years ago by Landsteiner [3], who coined the term hapten. Sumner and Howell in their studies of plant seed components including enzymes (urease) discovered the well-known agglutinin termed concanavalin A (meaning that it occurred with another protein called canavalin) that could precipitate glycogen and agglutinate erythrocytes [4]. The word lectin was coined in 1954 to denote those proteins from plants with specific precipitating activity [5]. Obviously, investigations in this area required new approaches to evaluate the binding of one substance by another.

Almost all of these early studies on lectins and blood group antibodies were based on visual measurements, and agglutination of erythrocytes was the most common method for these studies. The types of glycans recognized by the lectin were explored indirectly by hapten inhibition studies or the use of exoglycosidases to alter the structure of either the target cell or the inhibitory haptens. These approaches are fundamentally sound in many ways and actually led to the discovery of the blood group antigens by Winifred Watkins and Walter Morgan [6,7]. However, inhibition studies required large amounts of glycans and to be thorough they also require a large repertoire of defined glycans and lectin itself; in addition, the analyses were relatively labor-intensive and not readily amenable to high throughput.

4.3 NONCOVALENT COUPLING STRATEGIES FOR MICROARRAYS

Solid-phase binding assays represented a better approach that allowed direct binding to glycan targets with less material and led the way to modern high-throughput techniques. One of the first solid-phase assays for detecting protein binding to glycans was developed over 30 years ago and was an enzyme-linked immunosorbent assay (ELISA)-type approach to detect antibodies to bacterial polysaccharides that were directly adsorbed on plastic [8].

Another key early development in solid-phase assays for glycan binding proteins was the direct separation of naturally occurring glycolipids by thin layer chromatography and the use of a glycolipid-overlay technique, in which the separated glycolipids were directly assayed for binding with ^{125}I-labeled toxins or antibodies [9,10]. Multiple glycolipid targets could be simultaneously assayed for binding using extremely small quantities of labeled protein or antibody. This approach was extended by the synthesis of neoglycolipids (NGLs) using reductive amination of reducing glycans with dipalmitoylphosphatidyl ethanolamine [11]. These NGLs can be resolved by thin layer chromatography, incorporated into liposomes, or placed in microtiter plates, and probed with antibodies

or lectins to define the coupling glycan moieties [12]. This approach subsequently evolved into modern glycan microarrays in which NGLs are immobilized noncovalently by printing them on nitrocellulose-based glass slides [13,14].

The early solid-phase approaches with bacterial polysaccharides were extended by biotinylating the bacterial polysaccharides and capturing them on immobilized avidin-coated microtiter plates for use in ELISA [15]. Such noncovalent approaches for attaching glycans to a solid-phase surface are illustrated in Figure 4.1. Over the years, a number of investigators have utilized biotinylated glycans in various formats with immobilized streptavidin and demonstrated the utility of these approaches [16–19]. Glycan microarrays employing streptavidin/biotin capture of glycans will be discussed in more detail below. One of the early types of glycan microarrays was termed a "GlycoChip" and was based on the capture of biotinylated oligosaccharides linked to 30 kDa polyacrylamide on a microarray platform with Streptavidin XNA GOLD [20]. This array was successfully utilized to explore the specificity of antibodies to carbohydrate antigens.

One of the first large-scale noncovalent glycan microarrays, which used immobilized bacterial polysaccharides and lipopolysaccharides, was developed by Wang et al. [21]. In this approach, robotic printing was used to adsorb 48 glycans on nitrocellulose-coated glass slides using a relatively wide variety of glycans from glycosaminoglycans, polysaccharides, glycoproteins, and semisynthetic material. This glycan microarray was usefully explored by antibodies to glycans and visualization relied on using a biotinylated antihuman IgG that was detected with cyanine 3-streptavidin using a standard fluorescence scanner.

A variety of such noncovalent methods for generating glycan arrays were developed by several groups. Some of these arrays involved linkage of activated sugars to self-assembled monolayers through covalent methods [22], while others relied on adsorption of glycan or glycan conjugates to a

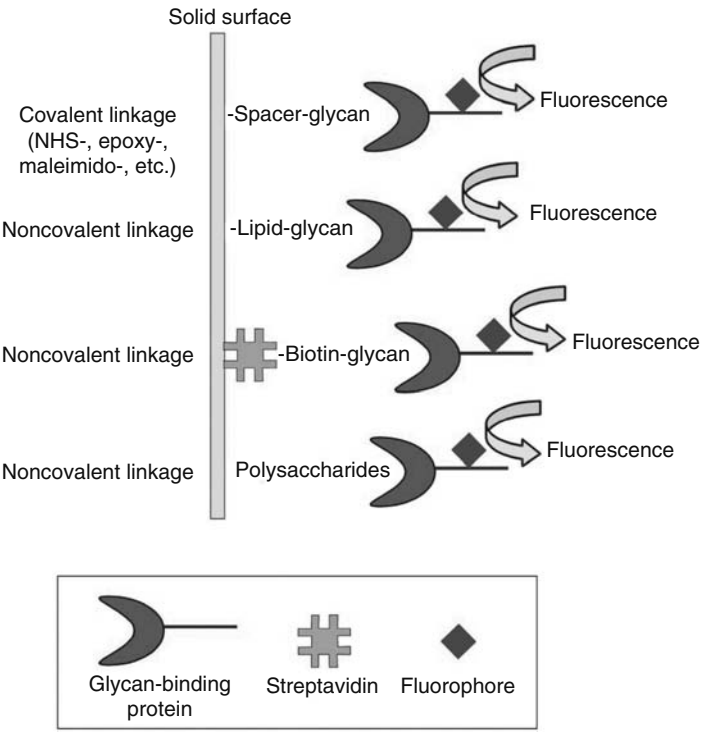


FIGURE 4.1 (See CD for color figure.) Multiple strategies for probing the binding of GBPs to specific glycans in a microarray approach. Both covalent and noncovalent methods of immobilizing glycans have been used.

nitrocellulose solid surface [14,21,23,24]. Schwarz et al. from Glycominds, Ltd. developed a novel glycan array using the strategy of coupling the glycans via a glycan-*p*-aminophenol linker at the reducing end to a novel linker (an oligomer of 1,8-diamino-3,6-dioxaoctan) on the solid support [25].

4.4 COVALENT COUPLING STRATEGIES

One of the earliest covalent coupling approaches to glycan microarrays was developed by Park and Shin, who synthesized maleimide-containing glycans and immobilized them covalently to thiol-derivatized glass slides with printing of 100–120 um diameter spots [26]. Another approach developed by the Seeberger Laboratory was to derivatize thiol-terminated ethylene glycol-derivatized glycans directly to glass slides, which had been precoated with bovine serum albumin that was preactivated with maleimide functional groups for covalent coupling [27]. This group also developed a number of other strategies for covalent coupling of glycans to activated glass slides, including the use of sugars with an ethanolamine linker at their reducing ends and coupling to an immobilized homobifunctional disuccinimidyl carbonate linker [28]. In addition, Wong and coworkers developed covalent microtiter plate-based glycan microarrays using alkyne chemistry and cleavable linkers [29], which was an extension of their original studies on noncovalent capture of glycolipids directly in microtiter plates [30].

4.5 LARGE-SCALE ELISA-TYPE GLYCAN MICROARRAYS

The use of biotinylated glycans in an ELISA-type format (ELISA glycan microarray) was widely developed as the first publicly available array through the Consortium for Functional Glycomics funded by the NIH/NIGMS and established in 2001. The ELISA glycan microarray initially contained several dozen biotinylated glycans. Some of the glycans were chemically synthesized with aliphatic spacers containing an azido moiety that was reduced to a primary amine and was subsequently biotinylated. This array was later expanded to well over 100 glycans and was used in many studies on plant and animal glycan binding proteins [31–42].

An example of the type of data obtained with the ELISA glycan microarray is shown in Figure 4.2, where the binding of murine galectin-4 and recombinant domains of the protein were tested [39]. Galectin-4 is a tandem repeat galectin and contains two carbohydrate recognition domains (CRD1 and CRD2) at the N- and C-termini, respectively. The ELISA glycan microarray was used to demonstrate that each CRD has unique carbohydrate specificity and that the specificity of the intact galectin-4 is a composite of the contribution of the individual CRDs. The binding studies in these

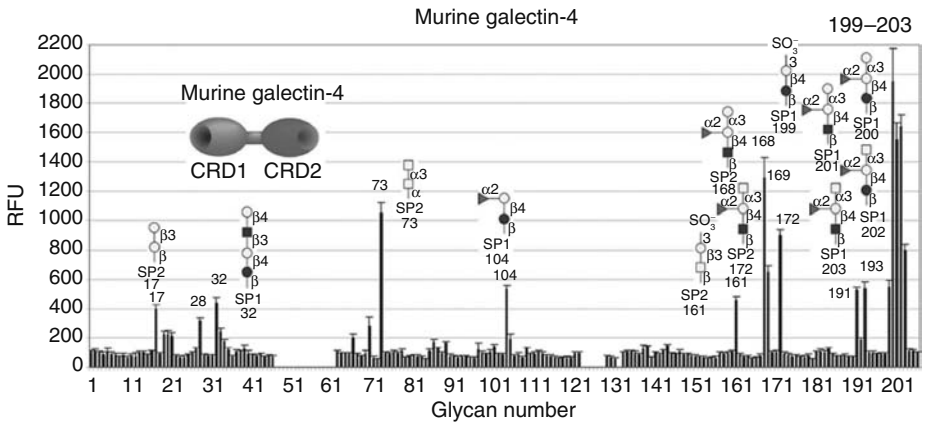


FIGURE 4.2 (See CD for color figure.) Example of the binding of murine galectin-4, a tandem-repeat galectin, to the glycan microarray containing biotinylated glycans captured on streptavidin. In this experiment, a recombinant form of murine galectin-4, a tandem repeat galectin, was fluorescently tagged with Alexa Fluor 488.

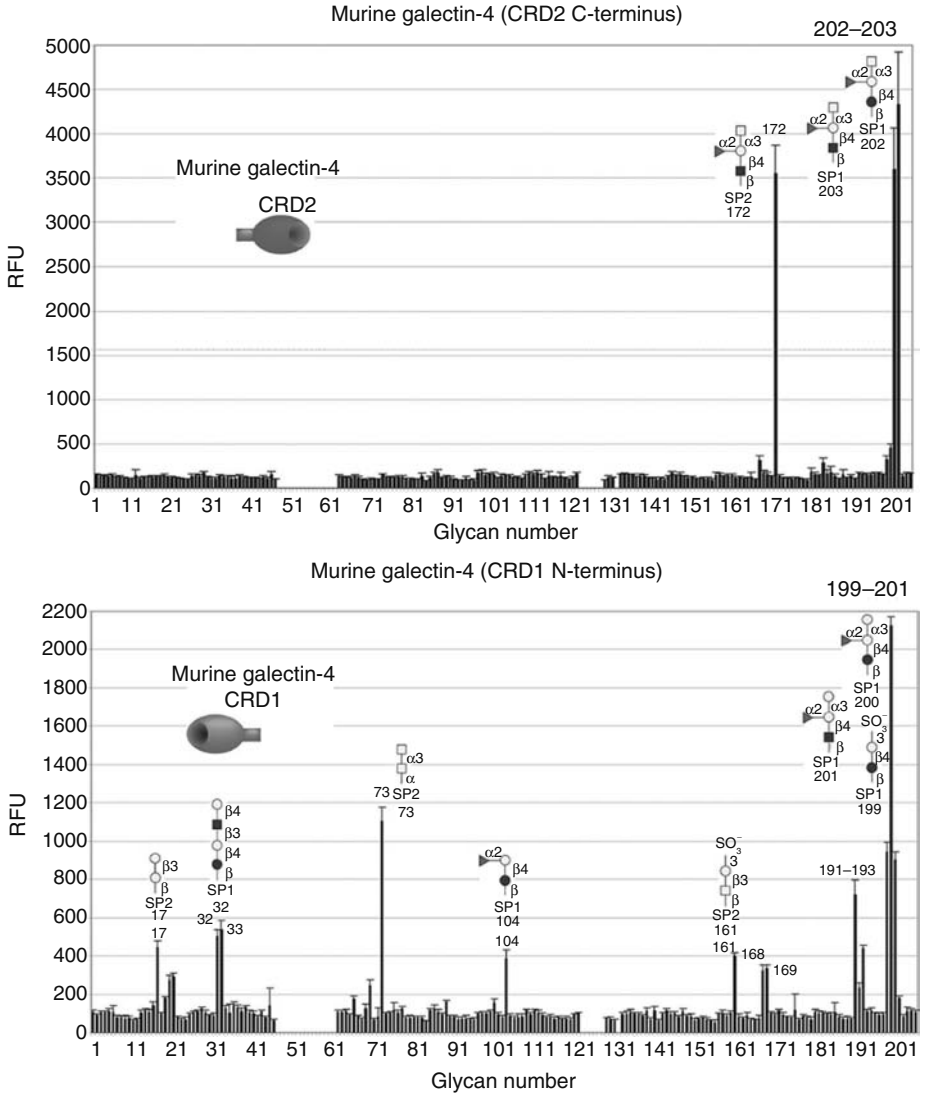


FIGURE 4.2 (continued) In addition, the C-terminal domain (CRD2) and the N-terminal domain (CRD1) were separately expressed and also fluorescently tagged. Each of these proteins were added to microtiter wells (mGal-4 2.8 µg/mL, CRD1 17.7 µg/mL, CRD2 12.3 µg/mL), and incubated for 1 h at room temperature. The plates were washed and bound galectin was directly measured in a Victor-2TM 1420 Multilabel Counter (PerkinElmer Life Sciences) at 485 nm excitation and 535 nm emission. The top panel shows the binding of the holoprotein, the middle and bottom panels show binding of recombinant forms of either the C-terminal or N-terminal domain containing the carbohydrate recognition domains 2 or 1, respectively. Each biotinylated glycan on the microarray is numbered. Some of the pertinent glycans bound by galectin-4 are shown. The heights of each peak in this bar graph represent the fluorescent intensity measured in a fluorescent microplate reader. (From Markova, V. et al., *Int. J. Mol. Med.* 18, 65, 2006.)

ELISA-type assays are basically to exploit common ELISA detection systems, which include fluorescence- or colorimetric-based reagents. In the typical experiments shown in Figure 4.2, the lectins are directly fluorescently labeled. Alternatively, fluorescently labeled secondary antibodies have been used to detect the primary binding protein. For example, studies on binding of human Siglec-8 employed a recombinant form of Siglec-8 as a chimera with the Fc portion of human IgG1 (Siglec-8-Ig) and binding was detected by incubation with goat antihuman IgG-Alexa 488 [32].

4.6 CURRENT METHODOLOGY FOR LARGE-SCALE GLYCAN MICROARRAYS

A key contribution from the consortium for functional glycomics (CFG) was the preparation of covalent glycan microarrays by coupling glycans with an aliphatic amine spacer directly to commercial glass slides containing polyethylene glycol surface with *N*-hydroxysuccinimide (NHS) coupling chemistry [43]. The great advantage of this glycan microarray is that it is printed by conventional contact printing approaches used for DNA and protein microarrays and that it utilizes direct covalent coupling with NHS chemistry, which is robust and highly efficient. The glycan microarray provided by the CFG is publicly available upon request and hundreds of different samples have been analyzed in Cores D and H of the consortium. Because this array is widely used, it will be the primary focus of the remainder of this chapter.

4.6.1 Source of Glycans

The CFG glycan microarray is periodically expanded by the addition of new glycans to the array. Initially, glycans used for populating the array were chemically synthesized as glycosides or glycosylamines with an amino-functionalized spacer at the reducing end for attachment to the NHS-derivatized glass slide [43]. Some of these glycans were modified by specific glycosyltransferases in the presence of appropriate sugar nucleotides to produce chemoenzymatically synthesized structures [44]. In addition, purified glycopeptides and glycoproteins were coupled to the array via their available free amino groups. A schematic showing the steps in producing and interrogating the CFG glycan microarray is found in Figure 4.3. The content of current and past versions of the microarray can be found at the Web site of the consortium http://www.functionalglycomics.org/static/consortium/resources/.

Ideally the glycan microarray should represent all known glycan structures, but this is not possible by current chemical or chemoenzymatic methods due to the complexity and the large number of potential structures. Such determinants are not well-defined, but historically they have been considered to be largely at the nonreducing termini of glycans. However, there can be recognition of internal glycan determinants along with clustered determinants. To aid in expanding the array, recent approaches to derivatization of naturally occurring free glycans or glycans released from glycoconjugates or polysaccharides by chemical or enzymatic methods have been successfully developed. For example, reducing glycans can be coupled to 1,2-dihexadecyl-*sn*-glycero-3-phosphoethanolamine to form NGLs that can be printed on nitrocellulose as a noncovalent microarray [14] or derivatized with a bifunctional spacer containing both a methyl-*N*,*O*-hydroxylamine for coupling to reducing sugars and a primary amine functionality suitable for coupling to NHS-activated surfaces [45]. Another useful approach has been the use of fluorescent bifunctional diamino spacers such as 2,6-diaminopyridine [46,47] that can be coupled with reducing glycans from natural sources. Fluorescent conjugates have an advantage in detection and permit the separation and purification of nanomolar amounts of material that can then be identified or structurally characterized and printed as arrays.

4.6.2 Printing of the Glycan Microarrays

The glycan microarrays are printed from stock solutions at concentrations of 10 to 100 μM by robotic pin printers that deposit approximately 0.6 nL of six replicates of glycan derivatives upon contact to the surface of NHS-derivatized slides as described for the glycan array produced for the

FIGURE 4.3 (See color insert following blank page 170. Also see CD for color figure.) Preparing covalent glycan microarrays printed on NHS-activated glass slides using glycans with a spacer at the reducing that contains a primary amine for coupling. The immobilized glycans are printed as small spots (50–150 μm) and then interrogated with material that may bind glycans, including GBPs, cells, viruses, bacteria, and serum. Binding is detected by fluorescence.

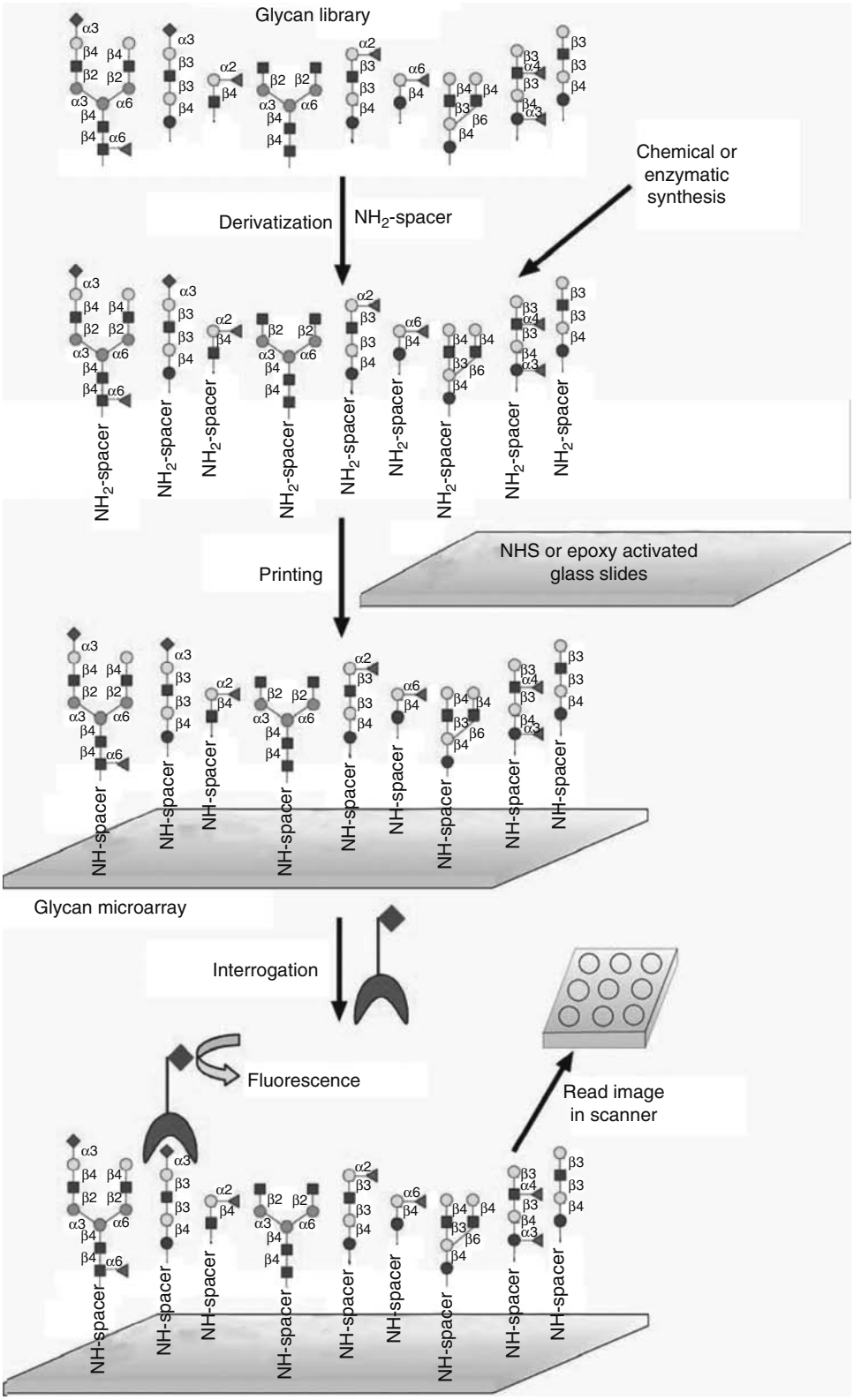


FIGURE 4.3 (See caption on page 54.)

consortium [43]. Alternatively, amino-functionalized fluorescent glycans can also be printed at high precision (<5% CV) at volumes of 0.3 nL on either NHS- or epoxy-derivatized slides using noncontact printing methods such as piezoelectric printing [46]. Fewer replicates due to more precise printing and smaller volumes decrease the amount of glycan derivative utilized for noncontact printing. In either case, fluorescent glycan derivatives can be visualized in the fluorescence scanner to permit location of the subarrays for quantitation and provide a convenient method for quality control of microarray fabrication. Regardless of the type of printing method or linker, the printed slides must be blocked to eliminate any residual reactive groups and prevent nonspecific binding of protein. The printed slides are washed with appropriate buffer components [43,46,47], rinsed with water, dried, and stored desiccated at room temperature until used.

4.6.3 Binding of Samples to the Glycan Microarray

Several major considerations for glycan analyses are the signal–noise ratio for detecting binding over a background noise, the concentration of immobilized glycans, the concentration of the GBP or receptor source (e.g., virus, bacteria), the effect of derivatization on the activity and stability of the GBP (e.g., fluorescent labeling or biotinylation), and the affinity of the GBP or receptor. All of these parameters are typically considered when a new binding study is undertaken and where there is little available information about the binding affinity of the GBP or receptor or even whether they bind carbohydrate. A trial and error approach is often taken where different strategies of derivatization, detection or cross-linking of the GBP are explored until a proper condition is found to conduct the experiments. Some GBPs may be inactivated by derivatization and care should be taken to confirm that the labeled GBP retains activity. Some GBPs appear to have low affinity that is enhanced by cross-linking them. For example, recombinant influenza virus hemagglutinins (HAs) are often detected by using His-tagged HAs that are made to be highly multivalent by cross-linking with anti-His mAb along with antimouse-IgG-Alexa Fluor488 in a proper ratio [48]. However, many GBPs bind with high affinity directly, such as galectins, and no special cross-linking is required to detect their binding [49].

Analyses of GBPs are carried out on slides after rehydration in buffer. The GBP is prepared in binding buffer that contains an appropriately buffered isotonic saline solution to support any cofactors such as divalent cations that may be required for binding and 1%–3% bovine serum albumin (BSA) or other blocking protein and 0.05% Tween 20 to prevent nonspecific binding. The GBP concentration usually varies in the range of 1–200 μg/mL depending upon the affinity for its coupling glycan. The GBP is added to the array in 50–70 μL and placed under a cover slip. Alternatively, if individual subarrays on a slide are contained by a barrier drawn with a hydrophobic pen, approximately 1 mL of GBP solution may be added to the surface of the slide. GBPs that are directly labeled with a fluorescent tag can be analyzed after a one-step process of binding followed by dipping the slide gently into solutions of a series of buffers containing no BSA, followed by buffer containing no detergent, and finally distilled water. The slide is then dried by spinning in a slide centrifuge or with a gentle stream of nitrogen to remove excess water.

Detection of GBPs bound to the array that are not directly labeled can be accomplished by most indirect immunochemical technique as described previously [43]. Modifications of these simple protocols are being developed to detect binding of whole viruses, bacteria, and other microorganisms and cells to glycans on the microarray. An example of a 12×16 subarray of the CFG glycan microarray that has been interrogated with a cocktail of 17 different fluorescently labeled lectins to demonstrate the distribution of glycans is shown in Figure 4.4. The replicated pattern is produced as a result of the six replicates of each glycan being distributed over the entire array that is comprised of a matrix of 4×8 subarrays. Thus the complete array is capable of supporting over 1000 glycan targets represented in replicates of six.

4.6.4 Analyzing the Binding Data

Fluorescence of four replicates (following the elimination of the highest and lowest values of each glycan) represented on the array are reported as the average of relative fluorescence units (RFU) detected using a

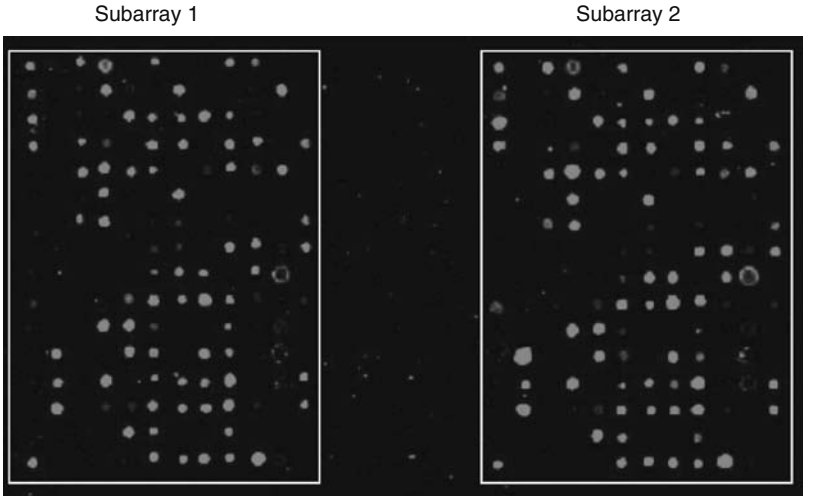


FIGURE 4.4 (See CD for color figure.) Example of the fluorescent-labeled lectin detected on a printed glycan microarray. Multiple duplicate subarrays are on the glass slide and this shows the reproducible binding of the lectin to two adjacent subarrays. Glycans are regularly spotted in a geometric grid on the microarray. Thus, in this example, blank spaces between green fluorescent spots indicates glycans that are not bound by the lectins tested. Data are typically scanned in a ScanArray 5000 (PerkinElmer Inc.) confocal scanner and the images are analyzed using the IMAGENE image analysis software (BioDiscovery, El Segundo, California).

ProScanArray fluorescence scanner (PerkinElmer). Analysis of the fluorescence image is carried out using IMAGENE image analysis software (BioDiscovery, El Segundo, California) for the Consortium for Glycan Array [43] or software included with the ProScanArray Scanner for noncontact printed arrays [46]. Data are recorded using Microsoft Excel software in tables listing all of the glycan structures, their corresponding average RFU values, the standard error of the mean (SEM) and the coefficient of variation reported in percent (%CV) and a histogram as shown in Figure 4.5.

Thus, the methodology for printing and the instruments for "reading" slides, through fluorescence detection, are widely available and access to the glycan microarray is publicly available from the consortium, where investigators can make an application to a Steering Committee that evaluates the scientific basis and available resources to decide whether to undertake the proposed study. In most cases, investigators with projects approved by the consortium submit samples to a core facility. These samples are typically fluorescently labeled GBPs, viruses, or even bacteria for analyses. For serum samples, where antibodies to glycan antigens may be evaluated, the serum samples are unlabeled, and bound IgG or IgM is identified by fluorescently labeled secondary antibody binding. Results of all experiments, when finalized, are made publicly available at the consortium Web site.

4.7 INSIGHTS INTO LECTIN LIGANDS USING GLYCAN ARRAYS

The glycan microarray has been used to screen hundreds of different samples, including lectins, antibodies, viruses, and bacteria. To illustrate the power of the glycan microarray, we will highlight the recent studies on viruses in this section. The initiation of most animal virus infections may be the binding of a viral coat protein to glycan cell surface receptors. This pathway is well documented for influenza viral infection where the major coat protein, the HA, binds specifically to sialic acid-containing cell surface glycans. Influenza virus that infects birds prefers to bind glycans that terminate in α2,3-linked sialic acids on intestinal epithelial cells, whereas human viruses specifically bind α2,6-linked sialic acids on lung and upper respiratory tract epithelial cells [50,51]. Adaptation of the avian virus to humans is presumed to be due to a switch in receptor specificity, which has been shown to occur with a single amino acid substitution in an influenza virus HA. Since the HA is not only responsible for the this binding specificity, but is also the primary viral protein recognized by

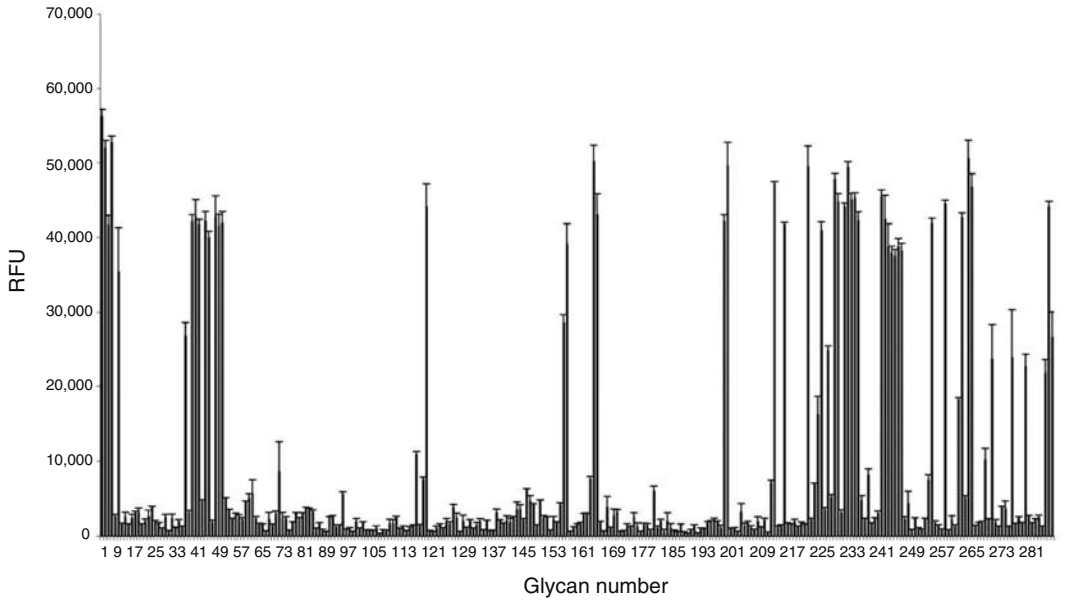


FIGURE 4.5 Glycan microarray analysis of the N-terminal domain of human galectin-8. The lectin domain was analyzed at 200 μg/mL (3.5–6 μM) and the microarray slides were comprised of 285 glycan targets represented on each slide at $n = 6$. The data shown are the average RFUs of four replicates after removal of the highest and lowest values, and error bars indicate the SEM.

neutralizing antibodies in vaccines, rapid methods for correlating mutations and structural changes with virus receptor specificities would be useful in predicting potential epidemics [48,52,53].

Prior to the advent of the printed glycan microarray, most of our understanding of influenza HA specificity was limited to differentiating α2,3- and α2,6-sialic acid linkages using cell-based assays. Hemagglutination or hemadsorption of erythrocytes that are enzymatically modified to express only α2,3- or α2,6-linked sialic acids was the first method of choice [54]. It is now possible, however, using glycan microarrays to rapidly analyze recombinant HAs in a cell-independent assay that permits screening for binding to sialoglycans whose number is only limited by what structures can be synthesized or isolated from natural sources. The current version (v3.0) of the glycan array from the consortium is comprised of 320 glycan targets distributed over the glass surface where approximately 30% contain sialic acid with 66 having α2,3-linked Neu5Ac, 26 having α2,6-linked Neu5Ac, 10 containing α2,8-linked Neu5Ac, and eight containing combinations of the various sialic linkages and six that are further substituted with sulfate. In addition, there are seven sialoglycans with α2,3- or α2,6-linked NeuGc. Thus, it is now possible in a single analysis to interrogate all of the printed glycans to reveal not only sialic acid linkage preference, but also subtle specificity differences associated with other features of the glycans including the size or length of the oligosaccharide, additional fucosylation, or additional charges associated with additional sialic acids or sulfates.

Stevens et al. [48] have recently demonstrated that the glycan array analysis not only revealed clear α2–3 and α2–6 sialic preferences, but also detected subtle differences in HA specificity for fucosylation and sulfation. Furthermore, they showed that the species barrier from birds to humans could be circumvented by changes at only two positions in the influenza HA. Thus the glycan microarray is useful in identifying mutations that may enable adaptation of influenza serotypes into the human population [53].

Adeno-associated viruses have shown promise as vectors for gene transfer. Understanding the tropism of these viruses at the molecular level could ultimately lead to a better understanding of their utility in gene therapy applications. Classical approaches to defining the specificity of different

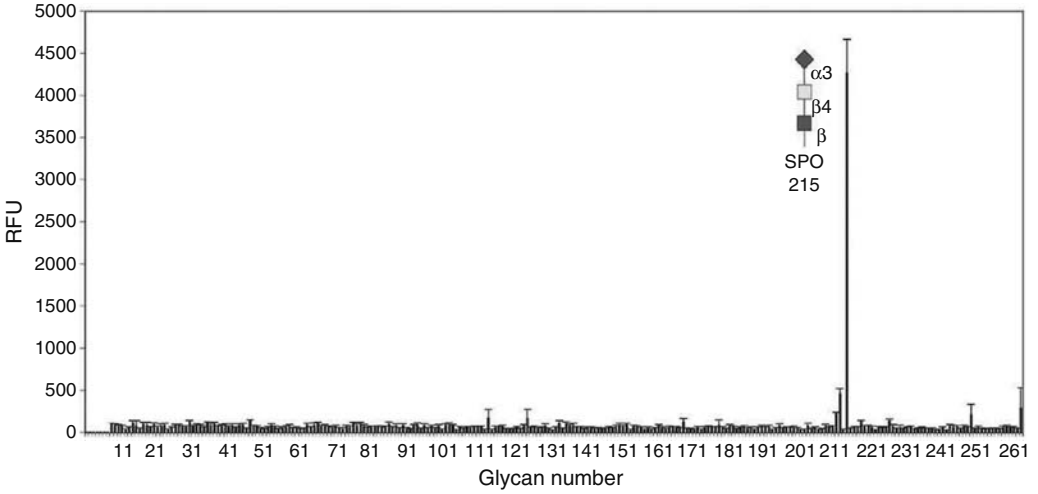


FIGURE 4.6 Binding of AAV1 capsid to the printed glycan array. A total of 264 glycans were screened for binding to the capsids and the plot shows the average relative fluorescence (RFU) for the six addresses for each glycan (thick line) versus glycan number. The standard error measurement is indicated for each glycan. The glycan measurably bound by AAV1 was glycan #215, as indicated. SPO is a space group. (From Wu, Z., Miller, E., Agbandje-McKenna, M., and Samulski, R.J., *J. Virol*., 80, 9093, 2006.)

adeno-associated virus strains involved hemagglutination, enzymatic treatments, and resialylation of erythrocytes, as well as competition binding assays using known glycans. Such studies provided strong evidence that cell-surface, sialic-containing glycoproteins are involved in binding and transduction of target cells [55]. Recent studies indicate that the glycan microarray will be a major tool to understand the subtle differences in specificity of viral capsid proteins for cell surface glycans. For example, a recent analysis of the specificity of adeno associated virus1 (AAV1) using the glycan microarray indicated that among the 264 glycans printed on the microarray, fluorescently labeled capsid AAV1 bound to a single glycan NeuAcα2–3GalNAcβ1–4GlcNAc (glycan 215) [56] (Figure 4.6). In related studies on parvovirus minute virus of mice (MVM), data from the glycan microarray showed that the prototype virus (MVMp), three virulent mutants, and an immunosuppressive strain (MVMi) all bound to a terminal sialic acid linked α2–3 to a common Galβ1–4GlcNAc motif of 3′SiaLN-LN, 3′SiaLN-LN-LN, and sLe^x- sLe^x- sLe^x. while one of the virulent strains and the immunosuppressive strain also recognized multisialylated glycans terminating in NeuAcα2–8 linkages characteristic of gangliosides GD3, GT3, and GD2 [57]. Thus, the glycan microarray provides a high level of structural definition of structures recognized by MVM and provides a rationale for the tropism of MVM for malignantly transformed cells that possesses the sLe^x motif, as well as the neurotropism of MVMi for interactions with ganglioside-type structures.

4.8 CONCLUSION AND FUTURE DIRECTIONS

Glycan microarrays offer a new approach for high-throughput screening of a variety of GBPs and glycan-binding organisms. Previous studies without such microarrays were limited by hapten inhibition approaches or direct affinity chromatography approaches. Through the use of glycan microarrays hundreds of samples including purified proteins, antibodies, serum samples, viruses, and bacteria have been studied in the past few years. The results are beginning to show that each GBP has unique recognition and affinity. Some GBPs appear to recognize a highly restricted set of glycans (see Figure 4.6), while others see more of a pattern of glycan structures within a restricted range (see Figures 4.4 and 4.5). Such studies are highlighting the need to take a fresh look at the

nature of protein–glycan interactions at both the atomic level with more complex glycans and at the cellular level. In addition, the interpretation of data from glycan microarrays should be interpreted cautiously, since these microarrays do not have the protein or lipid components to which glycans are linked; thus, such microarrays may miss key determinants contributed by the protein components. For example, P-selectin binds well to the N-terminal glycosulfopeptides domain of P-selectin glycoprotein ligand-1 (PSGL-1), but binds poorly to the glycan alone [58,59]. New developments in the formation of arrays using new linker chemistries and new glycan derivatization approaches are underway and attest to the robust interest in this new format. Also new developments are underway to directly capture natural glycans from human, animal, and microbial sources. Future studies are likely to expand the glycan microarrays and provide a richer matrix to help us explore the complex nature of glycan recognition and the functional roles of glycans in host and pathogen biology.

ACKNOWLEDGMENTS

The work of the authors was supported by NIH Grant P01 HL085607 to R.D.C. and in part by the Consortium for Functional Glycomics under NIH Grant GM62116.

REFERENCES

1. Wilkins, M. R. et al. Progress with proteome projects: Why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* 13, 19–50, 1996.
2. Sharon, N. and Lis, H. History of lectins: From hemagglutinins to biological recognition molecules. *Glycobiology* 14, 53R–62R, 2004.
3. Schwarz, H. P. and Dorner, F. Karl Landsteiner and his major contributions to haematology. *Br J Haematol* 121, 556–565, 2003.
4. Sumner, J. B. and Howell, S. F. Identification of hemagglutinin of jack bean with concanavalin A. *J Bacteriol* 32, 227–237, 1936.
5. Boyd, W. C. and Shapleigh, E. Specific precipitating activity of plant agglutinins (lectins). *Science* 119, 419, 1954.
6. Watkins, W. M. Blood-group substances. *Science* 152, 172–181, 1966.
7. Morgan, W. T. and Watkins, W. M. Unravelling the biochemical basis of blood group ABO and Lewis antigenic specificity. *Glycoconj J* 17, 501–530, 2000.
8. Carlsson, H. E., Lindberg, A. A., and Hammarstrom, S. Titration of antibodies to salmonella O antigens by enzyme-linked immunosorbent assay. *Infect Immun* 6, 703–708, 1972.
9. Magnani, J. L., Brockhaus, M., Smith, D. F., and Ginsburg, V. Detection of glycolipid ligands by direct binding of carbohydrate-binding proteins to thin-layer chromatograms. *Methods Enzymol* 83, 235–241, 1982.
10. Magnani, J. L., Smith, D. F., and Ginsburg, V. Detection of gangliosides that bind cholera toxin: Direct binding of 125I-labeled toxin to thin-layer chromatograms. *Anal Biochem* 109, 399–402, 1980.
11. Tang, P. W., Gool, H. C., Hardy, M., Lee, Y. C., and Feizi, T. Novel approach to the study of the antigenicities and receptor functions of carbohydrate chains of glycoproteins. *Biochem Biophys Res Commun* 132, 474–480, 1985.
12. Feizi, T. and Childs, R. A. Neoglycolipids: Probes in structure/function assignments to oligosaccharides. *Methods Enzymol* 242, 205–217, 1994.
13. Liu, Y., Chai, W., Childs, R. A., and Feizi, T. Preparation of neoglycolipids with ring-closed cores via chemoselective oxime-ligation for microarray analysis of carbohydrate–protein interactions. *Methods Enzymol* 415, 326–340, 2006.
14. Fukui, S., Feizi, T., Galustian, C., Lawson, A. M., and Chai, W. Oligosaccharide microarrays for high-throughput detection and specificity assignments of carbohydrate–protein interactions. *Nat Biotechnol* 20, 1011–1017, 2002.
15. Sutton, A., Vann, W. F., Karpas, A. B., Stein, K. E., and Schneerson, R. An avidin–biotin based ELISA for quantitation of antibody to bacterial polysaccharides. *J Immunol Methods* 82, 215–224, 1985.
16. Bovin, N. V. Neoglycoconjugates: Trade and art. *Biochem Soc Symp*, 143–160, 2002.
17. Bovin, N. V. et al. Synthesis of polymeric neoglycoconjugates based on N-substituted polyacrylamides. *Glycoconj J* 10, 142–151, 1993.

18. Leteux, C. et al. Biotinyl-l-3-(2-naphthyl)-alanine hydrazide derivatives of N-glycans: Versatile solid-phase probes for carbohydrate-recognition studies. *Glycobiology* 8, 227–236, 1998.
19. Rothenberg, B. E., Hayes, B. K., Toomre, D., Manzi, A. E., and Varki, A. Biotinylated diaminopyridine: An approach to tagging oligosaccharides and exploring their biology. *Proc Natl Acad Sci USA* 90, 11939–11943, 1993.
20. Galanina, O. E., Mecklenburg, M., Nifantiev, N. E., Pazynina, G. V., and Bovin, N. V. GlycoChip: Multiarray for the study of carbohydrate-binding proteins. *Lab Chip* 3, 260–265, 2003.
21. Wang, D., Liu, S., Trummer, B. J., Deng, C., and Wang, A. Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nat Biotechnol* 20, 275–281, 2002.
22. Houseman, B. T. and Mrksich, M. Carbohydrate arrays for the evaluation of protein binding and enzymatic modification. *Chem Biol* 9, 443–454, 2002.
23. Bryan, M. C., Lee, L. V., and Wong, C. H. High-throughput identification of fucosyltransferase inhibitors using carbohydrate microarrays. *Bioorg Med Chem Lett* 14, 3185–3188, 2004.
24. Willats, W. G., Rasmussen, S. E., Kristensen, T., Mikkelsen, J. D., and Knox, J. P. Sugar-coated microarrays: A novel slide surface for the high-throughput analysis of glycans. *Proteomics* 2, 1666–1671, 2002.
25. Schwarz, M. et al. A new kind of carbohydrate array, its use for profiling antiglycan antibodies, and the discovery of a novel human cellulose-binding antibody. *Glycobiology* 13, 749–754, 2003.
26. Park, S. and Shin, I. Fabrication of carbohydrate chips for studying protein–carbohydrate interactions. *Angew Chem Int Ed Engl* 41, 3180–3182, 2002.
27. Ratner, D. M. et al. Probing protein–carbohydrate interactions with microarrays of synthetic oligosaccharides. *Chembiochem* 5, 379–382, 2004.
28. Disney, M. D. and Seeberger, P. H. The use of carbohydrate microarrays to study carbohydrate–cell interactions and to detect pathogens. *Chem Biol* 11, 1701–1707, 2004.
29. Bryan, M. C. et al. Covalent display of oligosaccharide arrays in microtiter plates. *J Am Chem Soc* 126, 8640–8641, 2004.
30. Bryan, M. C. et al. Saccharide display on microtiter plates. *Chem Biol* 9, 713–720, 2002.
31. Avril, T., North, S. J., Haslam, S. M., Willison, H. J., and Crocker, P. R. Probing the *cis* interactions of the inhibitory receptor Siglec-7 with alpha2,8-disialylated ligands on natural killer cells and other leukocytes using glycan-specific antibodies and by analysis of alpha2,8-sialyltransferase gene expression. *J Leukoc Biol* 80, 787–796, 2006.
32. Bochner, B. S. et al. Glycan array screening reveals a candidate ligand for Siglec-8. *J Biol Chem* 280, 4307–4312, 2005.
33. Kim, Y. M. et al. Lectin from the Manila clam *Ruditapes philippinarum* is induced upon infection with the protozoan parasite *Perkinsus olseni*. *J Biol Chem* 281, 26854–26864, 2006.
34. Singh, T. et al. Carbohydrate specificity of an insecticidal lectin isolated from the leaves of *Glechoma hederacea* (ground ivy) towards mammalian glycoconjugates. *Biochem J* 393, 331–341, 2006.
35. Tateno, H., Crocker, P. R., and Paulson, J. C. Mouse Siglec-F and human Siglec-8 are functionally convergent paralogs that are selectively expressed on eosinophils and recognize 6'-sulfo-sialyl Lewis X as a preferred glycan ligand. *Glycobiology* 15, 1125–1135, 2005.
36. Coombs, P. J., Graham, S. A., Drickamer, K., and Taylor, M. E. Selective binding of the scavenger receptor C-type lectin to Lewisx trisaccharide and related glycan ligands. *J Biol Chem* 280, 22993–22999, 2005.
37. Coombs, P. J., Taylor, M. E., and Drickamer, K. Two categories of mammalian galactose-binding receptors distinguished by glycan array profiling. *Glycobiology* 16, 1C–7C, 2006.
38. Guo, Y. et al. Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* 11, 591–598, 2004.
39. Markova, V. et al. Role of the carbohydrate recognition domains of mouse galectin-4 in oligosaccharide binding and epitope recognition and expression of galectin-4 and galectin-6 in mouse cells and tissues. *Int J Mol Med* 18, 65–76, 2006.
40. McGreal, E. P. et al. The carbohydrate-recognition domain of Dectin-2 is a C-type lectin with specificity for high mannose. *Glycobiology* 16, 422–430, 2006.
41. Powlesland, A. S. et al. Widely divergent biochemical properties of the complete set of mouse DC-SIGN-related proteins. *J Biol Chem* 281, 20440–20449, 2006.
42. van Liempt, E. et al. Specificity of DC-SIGN for mannose- and fucose-containing glycans. *FEBS Lett* 580, 6123–6131, 2006.
43. Blixt, O. et al. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci USA* 101, 17033–17038, 2004.

44. Blixt, O. and Razi, N. Chemoenzymatic synthesis of glycan libraries. *Methods Enzymol* 415, 137–153, 2006.
45. Bohorov, O., Andersson-Sand, H., Hoffmann, J., and Blixt, O. Arraying glycomics: A novel bi-functional spacer for one-step microscale derivatization of free reducing glycans. *Glycobiology* 16, 21C–27C, 2006.
46. Song, X., Xia, B., Lasanajak, Y., Smith, D. F., and Cummings, R. D. Quantifiable fluorescent glycan microarrays. *Glycoconj J* 25, 15–25, 2008.
47. Xia, B. et al. Versatile fluorescent derivatization of glycans for glycomic analysis. *Nat Methods* 2, 845–850, 2005.
48. Stevens, J. et al. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *J Mol Biol* 355, 1143–1155, 2006.
49. Carlsson, S. et al. Affinity of galectin-8 and its carbohydrate recognition domains for ligands in solution and at the cell surface. *Glycobiology* 17, 663–676, 2007.
50. Connor, R. J., Kawaoka, Y., Webster, R. G., and Paulson, J. C. Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. *Virology* 205, 17–23, 1994.
51. Rogers, G. N. and D'Souza, B. L. Receptor binding properties of human and animal H1 influenza virus isolates. *Virology* 173, 317–322, 1989.
52. Stevens, J., Blixt, O., Paulson, J. C., and Wilson, I. A. Glycan microarray technologies: Tools to survey host specificity of influenza viruses. *Nat Rev Microbiol* 4, 857–864, 2006.
53. Stevens, J. et al. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312, 404–410, 2006.
54. Glaser, L. et al. A single amino acid substitution in 1918 influenza virus hemagglutinin changes receptor binding specificity. *J Virol* 79, 11533–11536, 2005.
55. Kaludov, N., Brown, K. E., Walters, R. W., Zabner, J., and Chiorini, J. A. Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J Virol* 75, 6884–6893, 2001.
56. Wu, Z., Miller, E., Agbandje-McKenna, M., and Samulski, R. J. Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J Virol* 80, 9093–9103, 2006.
57. Nam, H. J. et al. Identification of the sialic acid structures recognized by minute virus of mice and the role of binding affinity in virulence adaptation. *J Biol Chem* 281, 25670–25677, 2006.
58. Leppanen, A. et al. A novel glycosulfopeptide binds to P-selectin and inhibits leukocyte adhesion to P-selectin. *J Biol Chem* 274, 24838–24848, 1999.
59. Leppanen, A., White, S. P., Helin, J., McEver, R. P., and Cummings, R. D. Binding of glycosulfopeptides to P-selectin requires stereospecific contributions of individual tyrosine sulfate and sugar residues. *J Biol Chem* 275, 39569–39578, 2000.

5 Chromatography and Related Approaches for Qualitative and Quantitative Analyses of Lectin Specificity

Jun Hirabayashi

CONTENTS

5.1 FRONTAL AFFINITY CHROMATOGRAPHY

5.1.1 History of Improvement

Frontal affinity chromatography (FAC) is a quantitative technique of affinity chromatography originally developed by Kasai and Ishii [1] (for a detailed review, see Ref. [2]). Although it was first investigated as a tool for biomolecular interactions between enzymes and their inhibitors, its high potential and versatility have been demonstrated for many other biomolecules, for example, for quantitative specificity analysis of a representative plant lectin, concanavalin A, using a series of radiolabeled asparagine-linked oligosaccharides [3]. Originally, FAC was operated manually in a cold room using an open column and a fraction collector. Such procedures are laborious, time-consuming, and not sufficiently reproducible. In 2000, the system was greatly improved by the incorporation of a high-performance liquid chromatography (HPLC) system to assure high reproducibility and by the introduction of fluorescently labeled oligosaccharides to achieve high sensitivity [4]. The first such system introduced was quite simple, comprising a single isocratic pump, a manual injector, a fluorescence detector, and a personal computer. A specialized data analysis program was also developed [5]. With this specialized system, several dozens of interaction

analyses per day were performed for the determination of dissociation and association constants (K_d and K_a, respectively) between galectins and pyridylaminated (PA) oligosaccharides [6,7].

Two years prior to the technical advance mentioned above based on detection of fluorescent oligosaccharides, a different type of improvement was introduced. D.C. Schriemer and O. Hindsgaul applied the FAC principle to a modern analytical technique, electrospray ionization-mass spectrometry (ESI-MS). With this innovative coupling, online MS detection became possible for a series of carbohydrate derivatives synthesized by combinatorial chemistry [8]. This is an excellent combination, as various compounds can be subjected to FAC interaction analysis as a mixture, provided they can be differentiated on the basis of molecular mass [9–13]. FAC-ESI-MS has several disadvantages: These include difficulty in precise determination of elution volume and difficulty in ionization of relatively large glycans. The FAC user thus has the option, depending on the application, of MS or fluorescence detection. More recently, automated instruments for fluorescence-detection FAC (designated hereafter FAC-FD; Figure 5.1) have been developed in the course of the New Energy and Industrial Technology Development Organization (NEDO) project in Japan [14]. As described in previous reviews, the FAC-FD system has a number of advantages for the systematic comparison of the sugar-binding specificities of related lectins [15–17]. Perhaps the most important attribute of FAC-FD is its ability to provide precise values for K_a or K_d using minimal amounts of counterpart molecules (oligosaccharides). Most other analytical methods (e.g., equilibrium dialysis, surface plasmon resonance, enzyme-linked solid-phase assays, hemagglutination inhibition assays, glycan arrays) require greater amounts of the glycans that possess weak lectin affinities.

FIGURE 5.1 Outlook of an automated FAC instrument, FAC-T. The instrument was developed in the course of the NEDO project designated "SG" (Structural Glycomics) by collaboration with Shimadzu Corporation (Kyoto, Japan). The FAC-T instrument consists of two isocratic pumps, an autosampling system, a column oven, a pair of miniature capsule-type columns, and a PC workstation. In this system, FAC-T is equipped with a fluorescence detector. For details, see text. (From Nakamura-Tsuruta, S., Uchiyama, N., and Hirabayashi, J., *Methods Enzymol*., 415, 311, 2006; Nakamura-Tsuruta, S. et al. In: Nilsson, C.L., Ed., *Lectins: Analytical Technologies*, Chapter 10, Elsevier, Amsterdam, 2007.)

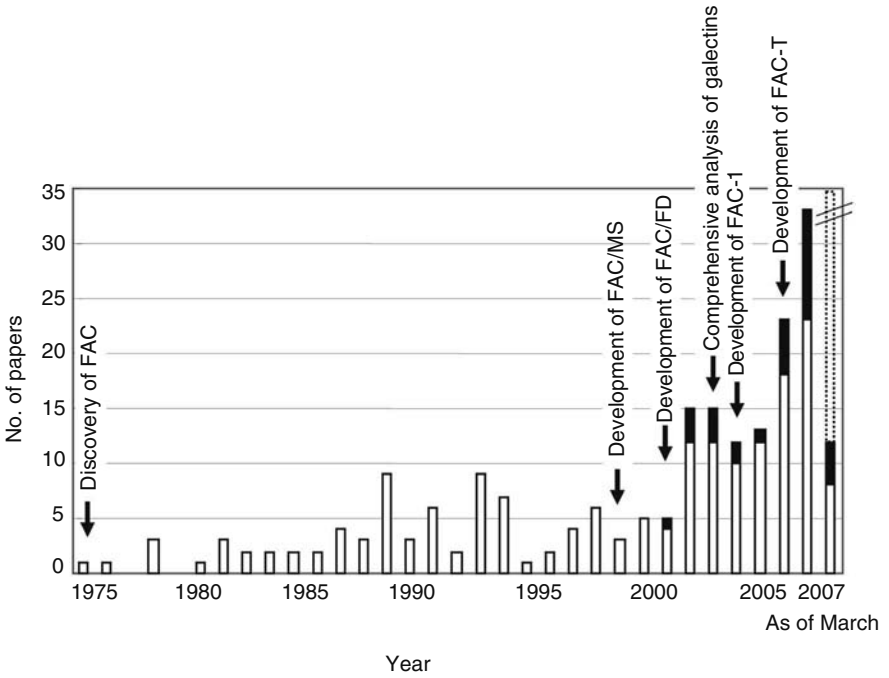


FIGURE 5.2 Increase in the publication of FAC-related papers. Closed bars represent numbers of papers per year describing FAC-FD (fluorescence detection), while open bars represent those describing other FAC technologies (conventional FAC and FAC-MS). Remarks of epoch-making events are noted above arrows.

The rapid uptake of FAC technology is clearly demonstrated by the increase in rate of publication of FAC-related papers following improvements published in 2000 [14]. Papers published in 2006 (10) account for 30% of the total number (33). (Figure 5.2).

5.1.2 Principle and Procedures

Details of the separation principle and mathematical aspects of FAC have been published [2]. More recently, the practical operation of the improved systems has been described [15–17]. This section describes the essence of the FAC procedure and the operation of the automated FAC-1 and FAC-T instruments.

A lectin-immobilized column must be prepared for analysis, together with fluorescently labeled oligosaccharides. Lectins are immobilized on NHS-activated Sepharose 4 Fast Flow (Amersham) and the resultant resin is packed into a capsule-type miniature column.

Appropriately diluted oligosaccharide solution is prepared and an excess volume is applied to the lectin-immobilized column. In the case of FAC-1 and FAC-T, at least 0.5 mL of diluted (e.g., 2.5 nM for *N*-glycans and 5 nM for others) PA-saccharide is applied by an autoinjection system. Since the miniature capsule-type column is sufficiently small (diameter, 2 mm × 10 mm; bed volume, 31.4 μL), an excess volume of the saccharide solution relative to the column (>15.9-fold) is assured. Under conditions that ensure dynamic equilibrium between the immobilized lectin and mobile-phase oligosaccharides (flow rate, 0.125 mL/min; temperature, 25°C), the latter molecules undergo repeated association and dissociation processes with the immobilized lectin molecules. This results in generation of a "leakage" of the elution front when the accumulation of the saccharide exceeds the capacity of the column. This leakage is actually observed as a delay (retardation) of the elution front (V) compared with that of an appropriate control saccharide (V_0), which shows no affinity to

the lectin. This observed delay, i.e., $V-V_0$ (expressed in mL), increases when either the affinity constant (K_a) to the lectin or the immobilized lectin content (B_t) becomes larger. This relationship is defined by the basic equation of FAC (Equation 5.1), where $[A]_0$ is the initial concentration of saccharide (expressed in M) and B_t is the column capacity (mol). K_a and K_d are related as shown in Equation 5.2.

$$K_d = B_t(V-V_0)-[A]_0 \tag{5.1}$$

$$K_d = 1/K_a \tag{5.2}$$

When K_d is much larger (e.g., >20-fold) than $[A]_0$ (note that both are expressed in M), Equation 5.2 is further simplified to

$$K_d = B_t(V-V_0) \tag{5.3}$$

Since lectin–carbohydrate interactions are generally weak, and in almost all cases not more than 10^7 M^{-1} in terms of K_a (i.e., in K_d, 0.1 μM), Equation 5.3 applies to almost all cases in FAC. With the use of fluorescent PA–oligosaccharides, concentrations of 1–10 nM are suitable for detection. Discussion of results is straightforward. "Twice-stronger retardation is observed for saccharide A compared with saccharide B." means twice-stronger affinity of A over B. Among other labeling reagents, 2-amino-benzamide (2-AB) also shows satisfactory performance as far as examined. Each V value may be obtained approximately by calculating the elution volume (ml) corresponding to a half maximum value of the "plateau" (i.e., $1/2[A]_0$). In most cases, a more precise determination of V is necessary. This is achieved by a mathematical procedure developed recently [5].

5.1.3 Advantages of FAC

As described above, the current system for FAC-FD has numerous advantages from both basic and practical viewpoints.

1. *Clear principle*: Essentially the same as that of enzyme kinetics, both of which are based on the classic Langmuir adsorption equation.
2. *Simple operation*: Utilization of an isocratic elution system guarantees high-throughput and reproducible analysis. Moreover, once the B_t value of the lectin column (described later) has been determined, one K_d value can be determined by a single interaction analysis (i.e., one injection).
3. *Full applicability to weak interaction analysis*: This is particularly important for analysis of lectin–carbohydrate interactions because they are relatively weak compared with those of antigen–antibody. Only a few methods are available for the determination of K_d (dissociation constant) >10^{-4} M, while FAC usually deals with K_d values between 10^{-3} and 10^{-7} M.
4. *Sensitivity*: When using PA-saccharides in FAC-FD, no more than 1 pmol is required for each analysis.
5. *Accuracy and reproducibility*: In accordance with Equation 5.3 described above, extents of retardation of the elution front ($V-V_0$) no longer depend on the initial concentration of an analyte saccharide ($[A]_0$). Therefore, in most cases ($K_d >> [A]_0$), highly accurate data on $V-V_0$ are obtained, which allow precise determination of K_d.
6. *Usefulness of PA-saccharides*: From a practical viewpoint, it is important that fluorescently labeled standard glycans be readily available. As the PA-labeling method [18] is now widely utilized, a number of (>100) PA-oligosaccharides with defined structures have

become commercially available; e.g., from Takara Bio Inc. (Kyoto, Japan) and Masudsa Chemical Industry (Takamatsu, Japan). PA-oligosaccharides do not show any undesirable nonspecific adsorption, as far as examined, to agarose and other HPLC support materials. These features are of great assistance in the construction of a glycan library for a lectin–carbohydrate interaction database.

5.1.4 Development of Automated Instruments FAC-1 and FAC-T

The NEDO project for structural glycomics was intended to develop high-throughput instruments, which include those for mass spectrometry and lectin-based glycan profiling [14]. FAC is a unique method, whose principle is simple and clear, for the provision of valuable information on lectin–glycan interactions in terms of affinity constants (K_a). Recent improvements in automatic injection and data analysis will allow the rapid collection of new research data, which are essential for the elucidation of lectin functions *in vivo*.

A prototype instrument for automated FAC, FAC-1, has been developed in collaboration with Shimadzu [16,17] (Kyoto, Japan; Figure 5.1). FAC-1 consists of two isocratic pumps (pump A for analysis and pump B for column washing), an autosampling system (enabling 210 injections), a column oven (usually set at 25°C), and a pair of miniature capsule-type columns. The system can be equipped with either fluorescence (Shimadzu, RF10AXL; e.g., for detection of PA-saccharides in specificity analysis) or UV detectors (Shimadzu, SPD-10A VP; e.g., for detection of *p*-nitrophenyl-saccharides in concentration analysis), and a PC workstation loaded with a conventional "LC solution" software. The FAC-1 system is primarily designed for efficient analysis with minimal dead time: This was attained by using a parallel column-switching system. During analysis with one column, the other column is subjected to washing, greatly reducing cycle analysis time to 5–6 min (Figure 5.3).

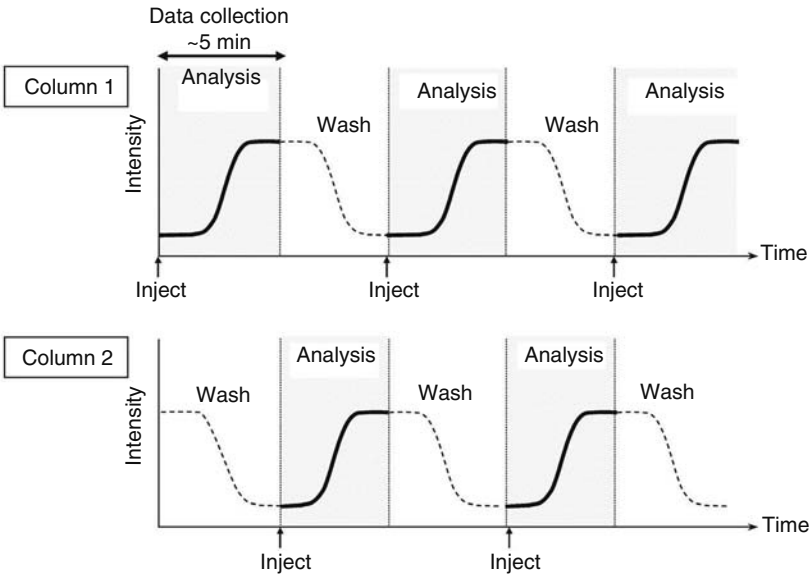


FIGURE 5.3 A scheme of pairwise analysis (data collection) and washing procedures in FAC-1. The instrument consists of two columns (columns 1 and 2) and two pumps for efficient analysis with minimal dead time: This was attained by using a parallel column-switching system. Hereby, a relatively short cycle analysis time (5–6 min) is achieved.

A more advanced instrument, FAC-T was developed, which, while still utilizing two pumps, is equipped with four columns in a "twin-parallel" (2 × 2) arrangement. (For details, see Refs. [16,17]). Both FAC-1 and FAC-T enable more than 200 analyses per day and require only 0.25–0.5 pmol/0.5 mL of PA-saccharide solution for each analysis. When stronger retardation is expected, application of a larger volume (e.g., 0.8 mL) is necessary to determine a precise V value. UV-sensitive saccharide derivatives are generally used for concentration-dependence analysis to determine B_t. These include *p*-nitrophenyl, *p*-aminophenyl, *p*-methoxyphenyl, and methotrexate derivatives. The concentration range used depends largely on lectin affinity to the saccharide used for analysis. In practice, if the concentration corresponding to K_d of the saccharide is applied to a lectin column, its V–V_0 value becomes $(V_{max}–V_0)/2$ in theory [2], where V_{max} is maximum V, which is obtained when the minimal concentration of PA-saccharide ($[A]_0 \ll K_d$) is used.

5.1.5 Publication of Results on the Internet (Lectin frontier DataBase)

It is suggested that results of lectin–carbohydrate interactions be published on the Internet in the form of a comprehensive database, as lectins have proved to be extremely useful tools and their functions *in vivo* have attracted intense attention from a range of scientific fields. Since initial investigations almost 120 years ago, many lectins have been isolated from plants, fungi, microbes, viruses and animals, including humans, and their distinctive biological properties have been characterized. Although there are various lectin databases, none containing experimental data on lectin–carbohydrate interactions have been published. In this context, we recently constructed a novel database, designated Lectin frontier DataBase (LfDB), which was designed along with the integrated database designated Carbohydrate sequence DataBase (CabosDB) by collaboration with Mitsui Knowledge Industry Co. Ltd. (Tokyo, Japan). The databases are planned for publication on the Internet in the near future [19].

LfDB provides molecular information on about 220 lectins, with common and formal nomenclature of the source organisms and tissues, lectin family, number of carbohydrate recognition domains (CRDs), amino acid sequences (deduced or directly determined), types of polypeptide fold (e.g., β-sandwich), references, and links to other relevant databases (e.g., GenBank, Protein Data Bank, Pfam). However, the most important feature is that LfDB gives comprehensive, experimental data on lectin affinity determined by FAC. The lectin entries can be retrieved through the database using a variety of search options (e.g., keywords, molecular properties of lectins, structural features of glycans, affinity strength of lectin–oligosaccharide interactions). Users can browse the derived affinity patterns, characteristics of the target lectins, and structures of the relevant glycans.

5.2 LECTIN MICROARRAY

The lectin microarray is an emerging technique in the field of structural glycomics. The idea originated from the concept of glycan profiling [14], by which complex features of glycans and glycocoproteins, and even their mixtures (e.g., cell extracts, body fluids containing various glycoproteins of multiple glycoforms) can be characterized by a simple procedure. Such an approach is urgently required for quality control of glycoprotein drugs and cells for medical use (regenerative medicine) and for investigation of glycan-related biomarkers. The same approach has been applied to other profiling methods, for example, a mass spectrometry technique called spectral matching (MS^n) and multidimensional HPLC mapping (LC^n) using PA or 2-AB-labeled oligosaccharides. Unlike these methods based on physical or chemical separation principles, the lectin microarray is based on the biological recognition properties of lectins. In comparison with antibodies, another group of biomolecules having specific affinity to particular molecules, lectins generally show weaker affinity and broader specificity. This is in clear contrast to antibody arrays, in that lectin microarrays works well for glycan profiling but not for glycan identification.

Considering the rapid accumulation of fundamental data on lectin–carbohydrate interactions, as well as the pressing requirement for a rapid and sensitive profiling method other than MS^n (expensive) and LC^n (low throughput), it was a natural progression to use multiple lectins as deciphering

molecules on a microarray platform, in analogy to DNA and protein microarrays. Academic papers reporting lectin microarrays emerged almost simultaneously from 2005 [20–25]. Microarray applications were also reported, which include the use of crude samples derived from cultured cells, and clinical samples [26–29]. As lectin–carbohydrate interactions are relatively weak compared with other bio-molecular interactions, development of a lectin microarray requires a special technique uncommon in conventional DNA and antibody microarrays. The same problem is also true for carbohydrate microarrays, unless high-density immobilization is achieved for glycan immobilization [30]. In the following sections, reasons why the evanescent-field fluorescence-activated detection principle was adopted for the lectin microarray are explained [31], followed by a description of its potential for differential profiling in the special context of biomarker investigation.

5.2.1 Evanescent-Field Fluorescence-Activated Detection Principle

It has been long known that an evanescent-field wave is generated between two phases (e.g., glass slide and probing buffer) significantly differing in reflection index when excitation light is injected to the solid phase (glass) with an injection angle appropriate to cause total reflection (Figure 5.4). The wave thus generated is restricted to an extremely narrow distance from the glass surface (substantially <200 nm). This "near optic field" is utilized for activation of fluorescently labeled

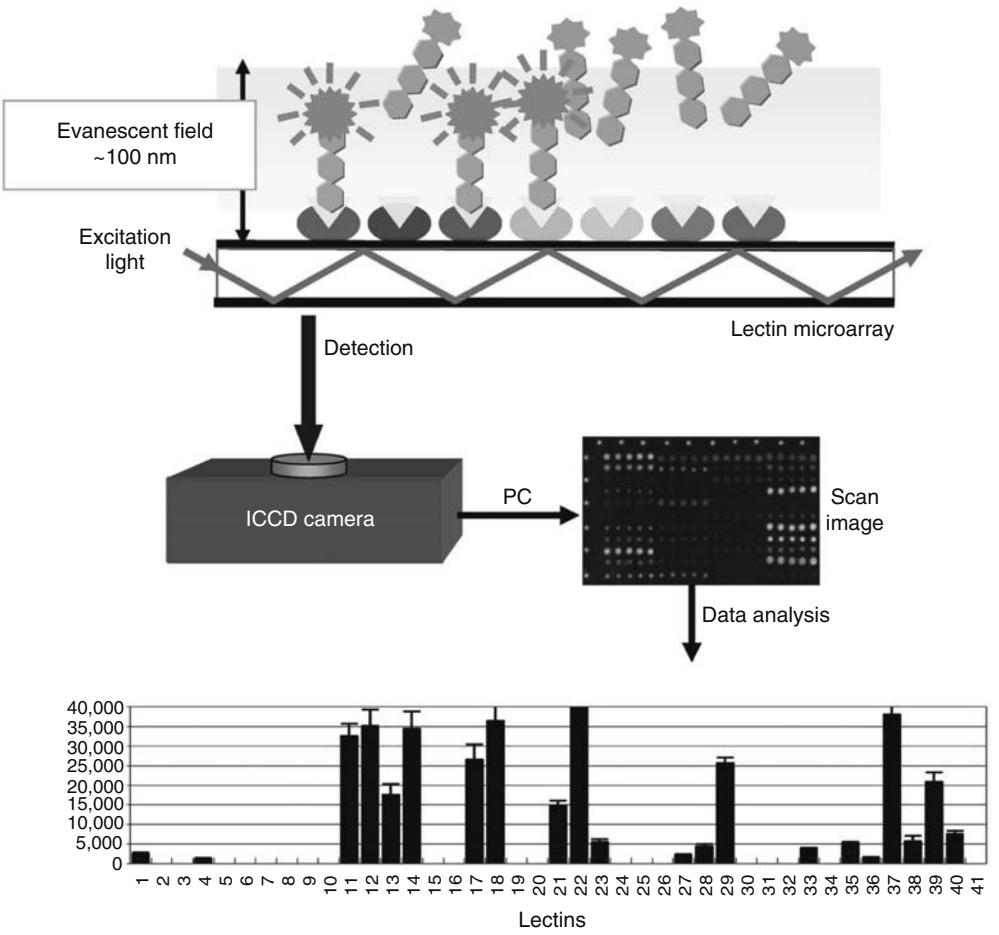


FIGURE 5.4 (See CD for color figure.) A schematic drawing of the evanescent-field fluorescence-activated detection principle adopted for the development of lectin microarray. With this principle, relatively weak lectin–carbohydrate interaction is observed without washing procedures in a real-time manner. For detailed explanation, see text.

FIGURE 5.5 Outlook of an evanescent-field activated-fluorescence detection-type instrument (SC-Profiler) for lectin microarray. The instrument was developed in the course of the NEDO project designated "SG" (Structural Glycomics) by collaboration with Moritex (Tokyo, Japan).

glycans, which are added to the lectin microarray [23,31]. In this context, a specialized instrument for lectin microarray, SC-Profiler, was developed in the course of NEDO project by collaboration with Moritex (Tokyo, Japan; Figure 5.5). The adopted detection principle, i.e., evanescent-field fluorescence-activated method, is considered superior to others in the following ways:

1. *No requirement for washing procedures*: Because only the limited detection area is excited by the evanescent wave, no washing process to remove unbound fluorescent probes is necessary before detection.
2. *Real-time observation under equilibrium conditions*: This is not achievable by other detection methods that use washing procedures, while it is possible to detect in situ lectin–carbohydrate interactions at equilibrium in an aqueous phase.
3. *Applicability to crude and mixed samples*: Unlike FAC-FD, the described lectin microarray enables specific detection of target molecules among mixed glycans/glycoproteins, which include clinical samples.
4. *Versatility and applicability*: The method is applicable to glycans, glycopeptides, and glycoproteins as well as cell extracts differing in glycosylation machineries. For detection of a target glycoprotein, a sandwich method using specific antibody is adopted [28,29].

It must be noted that the evanescent-field fluorescence-detection system has some inherent properties, which may be disadvantageous in some cases. These include:

1. The method requires prior labeling of target glycans/glycoproteins, unlike the surface plasmon resonance method, though this is common to all of the reported lectin microarray systems. For labeling, commercially available fluorescent reagents, e.g., Cy3, Cy5, and tetramethylrhodamine (TAMRA), can be used. In the case of a fixed target molecule (glycoprotein), the use of a specific antibody for probing the protein moiety of the glycoprotein will be effective for high-throughput analysis [28,29].

2. The evanescent wave, as an excitation light, will activate a small fraction of nonbound fluorescent probes occurring in the near-optic field, which will generate a significant level of background in some cases. There are various methods for improvement: these include reduction of the probe amount applied to the array and addition of a minimal washing procedure to remove the bulk of unbound fluorescent probes. A more fundamental solution to this potential problem may be achieved by either addition of some colloidal particles to the probe solution (Ebe et al., unpublished results). The solution largely depends on which options are chosen for each experimental condition and on the level of requirements (e.g., purpose of experiments, required quality of results, etc.).

5.2.2 Inherent Properties of the Lectin Microarray

When using this special technique, some inherent properties different from those of other profiling methods (e.g., MS and HPLC) should be considered. Firstly, all lectin microarray systems reported so far do not necessarily give equal signal intensities on different lectin spots, i.e., each lectin spot shows a different maximal value. Signal intensity depends on accessibility and availability of immobilized lectins, affinity between lectins and glycans/glycoproteins, and the amount of target glycans/glycoproteins. Thus, the maximum signal value on each spot cannot be adequately controlled. It is recommended that the lectin microarray in its present form is best used for "differential profiling," which compares glycan profiles under the same experimental conditions.

Secondly, distinct from other profiling methods (e.g., MS and LC), lectin microarray results do not provide direct information on covalent structures of glycans. They do give useful clues for glycan structures in terms of epitope and other characteristic features. For interpretation, accumulated information on lectin–carbohydrate interactions by FAC is essential. From a practical viewpoint, combined use of the lectin microarray with other analytical methods is encouraged.

5.2.3 Difficulties and Complexity

The major application of the lectin microarray is in the field of differential profiling. However, it cannot be assumed that compared samples will show significantly distinct profiles. There is a tendency that the higher the complexity of glycan structures, the greater the difficulty in differentiating between samples. For example, relatively simple profiles are obtained for TMR-labeled glycans, which show clearly distinct patterns to one another (e.g., high-mannose type and biantennary complex type N-glycans) [23]. Such a feature is maintained for glycopeptides and glycoproteins if the number of glycosylation attachment sites and types of glycans (e.g., N-glycans and O-glycans) are few. However, in relatively large glycoproteins (e.g., laminin) with multiple glycosylation sites in a mixture of N-glycans and O-glycans, lectin array profiles become much more complex. Nevertheless, different glycoproteins including those from different species (e.g., transferrin from mouse, chick, and pig) can be clearly differentiated. Differentiation between different types of cells is also possible. These include chinese hamster ovary (CHO) and its mutant lectin (LEC) cells [26]. Apparent homogeneity of the "glycome" is maintained in each type of cell. In other words, the glycome is basically defined by outputs of glycosylation-related enzymes expressed in each cell. On the other hand, tissues and organs are heterogeneous in terms of cell type, and thus, in terms of glycome. It is therefore sometimes difficult to differentiate different types of tissue. Difficulty in differential profiling is probably greatest in blood or serum, which are fairly complex mixtures of secreted (glyco)proteins, as well as their degradation products, originating from multiple tissues and cells. In the latter samples, some enrichment procedures are necessary to differentiate individuals and normal patients. Serum also contains an extremely abundant component, albumin. This is a critical issue in the investigation of useful biomarkers.

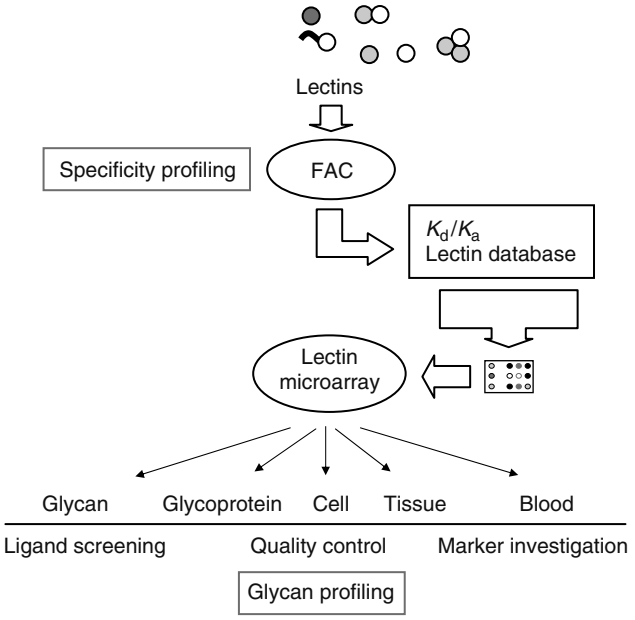


FIGURE 5.6 A possible strategy for lectin-based structural and functional glycomics. FAC provides essential information of diverse lectins from both plants and animals in a high-throughput manner, in terms of K_a and K_d, whereby fundamental knowledge is accumulated in the Lectin Affinity DataBase. With these bioresources (lectins and specificity information), lectin microarray will be most effectively utilized for various purposes in various fields, which include investigation of ligand molecules (glycans/glycoproteins) for animal lectins, quality control of glycoprotein drugs in industrial companies, and searching for differentiation markers in immune cells as well as disease-related biomarkers.

5.3 CONCLUSION AND FUTURE DIRECTIONS

A new NEDO project termed "Medical Glycomics" was commenced in April 2006, with the aim of finding disease-related biomarkers by means of various biotools, including the evanescent-type lectin microarray described in this chapter. Various other applications of this emerging technique are greatly hopeful (Figure 5.6). Results obtained with the lectin microarray described in this chapter will be described elsewhere.

ACKNOWLEDGMENTS

The author thanks all members of Glycostructure Analysis Team (Research Center for Glycoscience, AIST; November 2006) and Lectin Application and Analysis Team (Research Center for Medical Glycoscience, AIST; December 2006~) involved in the studies of FAC and lectin microarray described in this chapter. Industrial members from Shimadzu and Moritex are also acknowledged for the development of the instruments (FAC-1, FAC-T, and SC-Profiler). Financial support was obtained from New Energy and Industrial Technology Development (NEDO), Japan. Special thanks to Prof. Dr. Ken-ichi Kasai (Teikyo University) for helpful discussions.

REFERENCES

1. Kasai, K. and Ishii, S. Quantitative analysis of affinity chromatography of trypsin: A new technique for investigation of protein–ligand interaction, *J. Biochem*., 177, 261, 1975.
2. Kasai, K. et al. Frontal affinity chromatography: Theory for its application to studies on specific interaction of biomolecules, *J. Chromatogr*., 376, 323, 1986.

3. Ohyama, Y. et al. Frontal affinity chromatography of ovalbumin glycoasparagines on a concanavalin A-Sepharose column: A quantitative study of the binding specificity of the lectin, *J. Biol. Chem*., 260, 6882, 1985.
4. Hirabayashi, J., Arata, Y., and Kasai, K. Reinforcement of frontal affinity chromatography for effective analysis of lectin–oligosaccharide interactions, *J. Chromatogr.*, 890, 261, 2000.
5. Arata, Y., Hirabayashi, J., and Kasai, K. Application of reinforced frontal affinity chromatography and advanced processing procedure to the study of the binding property of a *Caenorhabditis elegans* galectin, *J. Chromatogr.*, 905, 337, 2001.
6. Arata, Y., Hirabayashi, J., and Kasai, K. Sugar binding properties of the two lectin domains of the tandem repeat-type galectin LEC-1 (N32) of *Caenorhabditis elegans*: Detailed analysis by an improved frontal affinity chromatography method, *J. Biol. Chem*., 276, 3068, 2001.
7. Hirabayashi, J. et al. Oligosaccharide specificity of galectins: A search by frontal affinity chromatography, *Biochim. Biophys. Acta*, 1572, 232, 2002.
8. Schriemer, D.C. et al. Micro-scale frontal affinity chromatography with mass spectrometric detection: A new method for the screening of compound libraries, *Angew. Chem. Int. Ed*., 37, 3383, 1998.
9. Zhang, B. et al. Frontal affinity chromatography coupled to mass spectrometry for screening mixtures of enzyme inhibitors, *Anal. Biochem*., 299, 173, 2001.
10. Palcic, M.M. et al. Evaluating carbohydrate–protein binding interactions using frontal affinity chromatography coupled to mass spectrometry, *Methods Enzymol*., 362, 369, 2003.
11. Schriemer, D.C. Biosensor alternative: Frontal affinity chromatography, *Anal. Chem*., 76, 440A, 2004.
12. Ng, E. and Schriemer, D.C. Emerging challenges in ligand discovery: New opportunities for chromatographic assay, *Expert Rev. Proteomics*, 2, 891, 2005.
13. Ng, E.S. et al. High-throughput screening for enzyme inhibitors using frontal affinity chromatography with liquid chromatography and mass spectrometry, *Anal. Chem*., 77, 6125, 2005.
14. Hirabayashi, J. Lectin-based structural glycomics: Glycoproteomics and glycan profiling, *Glycoconj. J*., 21, 35, 2004.
15. Hirabayashi, J., Arata, Y., and Kasai, K. Frontal affinity chromatography as a tool for elucidation of sugar recognition properties of lectins, *Methods Enzymol*., 362, 353, 2003.
16. Nakamura-Tsuruta, S., Uchiyama, N., and Hirabayashi, J. High-throughput analysis of lectin-oligosaccharide interactions by automated frontal affinity chromatography, *Methods Enzymol*., 415, 311, 2006.
17. Nakamura-Tsuruta, S. et al. Frontal affinity chromatography: Systematization for quantitative interaction analysis between lectins and glycans, in *Lectins: Analytical Technologies*, Nilsson, C.L., Ed., Elsevier, Amsterdam, 2007, Chapter 10.
18. Hase, S., Ikenaka, T., and Matsushima, Y. Structure analyses of oligosaccharides by tagging of the reducing end sugars with a fluorescent compound, *Biochem. Biophys. Res. Commun*., 85, 257, 1978.
19. Takahashi, Y. and Hirabayashi, J. A novel lectin-affinity database for structural glycomics, in *Glycoscience Lab Manual*, Springer Japan, Tokyo, 2007, in press.
20. Angeloni, S. et al. Glycoprofiling with micro-arrays of glycoconjugates and lectins, *Glycobiology*, 15, 31, 2005.
21. Pilobello, K.T. et al. Development of a lectin microarray for the rapid analysis of protein glycopatterns, *Chem. BioChem*., 6, 985, 2005.
22. Zheng, T. et al. Lectin arrays for profiling cell surface carbohydrate expression. *J. Am. Chem. Soc*., 127, 9982, 2005.
23. Kuno, A. et al. Evanescent-field fluorescence-assisted lectin microarray: A new strategy for glycan profiling, *Nat. Methods*, 2, 851, 2005.
24. Koshi, Y. et al. A fluorescent lectin array using supramolecular hydrogel for simple detection and pattern profiling for various glycoconjugates, *J. Am. Chem. Soc*., 128, 10413, 2006.
25. Rosenfeld, R. et al. A lectin array-based methodology for the analysis of protein glycosylation, *J. Biochem. Biophys. Methods*, 70, 415, 2007.
26. Ebe, Y. et al. Application of lectin microarray to crude samples: Differential glycan profiling of lec mutants, *J. Biochem. (Tokyo)*, 139, 323, 2006.
27. Hsu, K.L., Pilobello, K.T., and Mahal, L.K. Analyzing the dynamic bacterial glycome with a lectin microarray approach, *Nat. Chem. Biol*., 2, 153, 2006.
28. Kato, Y. et al. Inhibition of tumor cell-induced platelet aggregation using a novel anti-podoplanin antibody reacting with its platelet-aggregation-stimulating domain, *Biochem. Biophys. Res. Commun*., 349, 1301, 2006.

29. Kaneko, M.K. et al. Functional glycosylation of human podoplanin: Glycan structure of platelet aggregation-inducing factor, *FEBS Lett*., 23, 331, 2007.
30. Hirabayashi, J. Oligosaccharide microarrays for glycomics, *Trends Biotechnol*., 21, 141, 2003.
31. Uchiyama, N. et al. Development of a lectin microarray based on an evanescent-field fluorescence principle, *Methods Enzymol*., 415, 341, 2006.

6 Analysis of Whole-Genome and Other Data Resources to Characterize the Molecular, Structural, and Evolutionary Diversity of C-Type Lectins and Discover New Genes

Alex N. Zelensky and Jill E. Gready

CONTENTS

6.1 INTRODUCTION—GLYCOBIOLOGY IN THE "OMICS" ERA

The last decade saw an explosive growth in the volume of publicly accessible data due to the development of high-throughput research technologies such as large-scale DNA cloning and sequencing, semiautomated protein structure determination, systematic knockdown of individual genes by RNA interference and simultaneous analysis of variation in the expression of thousands of genes by microarray assays, and plate- and chip-based protein/ligand binding assays. Concurrently, bioinformaticians have devoted much effort to integrating this heterogeneous data, developing efficient algorithms for its analysis and making it accessible to the wider biological research community. In this chapter, we describe how various data resources can be explored productively in the field of glycobiology based on our own experience of *in silico* analysis of the evolution, structure, and function of animal calcium-binding (C-type) lectins.

C-type lectins constitute a superfamily of proteins containing C-type lectin-like domains (CTLDs, Figure 6.1), a large and heterogeneous group of extracellular Metazoan proteins with diverse functions. Evolutionary data suggest that sugar binding is the original function of the superfamily, as it is observed in its most distant members. However, many CTLDs have evolved to recognize ligands other than carbohydrates, including proteins, inorganic substances such as ice and calcium carbonate, and lipids. The unique loop-in-a-loop structure of the domain, in which a large flexible region is maintained on a stable core (Figure 6.1), can accommodate substantial variation in the shape of the ligand-binding region, and allows specific binding of large multivalent ligands such

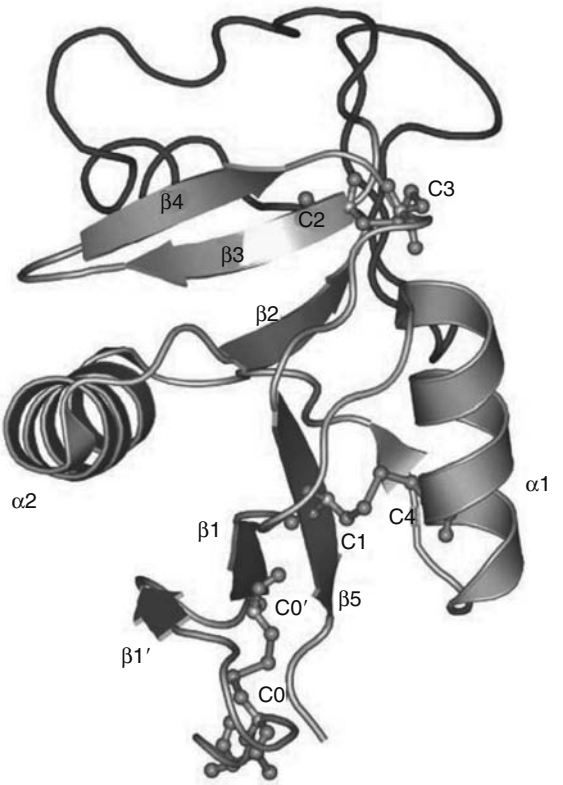


FIGURE 6.1 (See color insert following blank page 170. Also see CD for color figure.) CTLD structure. A cartoon representation of a typical CTLD structure (1k9i). Our suggested universal numbering scheme for secondary structure elements is used [1]. The long loop region is shown in blue. Disulfide bridges are shown as orange sticks, with cysteines numbered. The archtypical disulfide bridges of the fold are between C1 and C4, and C2 and C3. The disulfide bridge (C0–C0′) specific for long-form CTLDs, which have an N-terminal extension, is also shown.

as complex oligosaccharides. Formation of quaternary complexes further increases both specificity and affinity of carbohydrate recognition by C-type lectins, and allows them to function in diverse extracellular contexts, such as innate immune defence, turnover of glycosylated proteins, maintenance of the extracellular matrix structure, and cellular adhesion.

Comparative analysis and data mining have played an integral role in studies of the C-type lectin family. In a review of the family in 1993, Drickamer divided the CTLD-containing proteins (CTLDcps) known at that time into seven groups (I–VII) based on their domain architecture, and showed that such grouping correlated well with the results of phylogenetic analysis of the CTLD sequences and captured the functional similarities among the proteins [2]. The classification was revised in 2002 [3] with the addition of seven new groups (VIII–XIV), and later we suggested some re-definitions and addition of a further three new groups (XV–XVII) to make the classification consistent and able to cover all conserved vertebrate CTLDs [4]. Four studies using the whole-genome approach have been published analyzing the distribution of the superfamily in *Caenorhabditis elegans* [5], *Drosophila melanogaster* [6], human [3], and fish [7]. These studies and genome annotation projects demonstrated the relatively high genome-wide abundance of CTLDcps, compared with proteins containing other domains, and the strong conservation of the groups within the vertebrate lineage but with little or no similarity between vertebrate and invertebrate CTLDcps. These studies also led to discovery of a number of novel CTLDcp genes.

6.2 DISCOVERY OF NEW C-TYPE LECTINS

Large-scale cDNA sequencing, free public access to sequence databases such as GenBank or EMBL, and tools such as BLAST [8] to search them for homologous sequences, has changed the way new genes are discovered. These methods have been quickly adopted by the biological research community, and allow not only identification of a novel sequence, but also tissues in which it is present. Using iterative and profile-based search tools such as PSI-BLAST [8] or HMMER [9], very remote homologs can be detected with weak but significant sequence similarity to the query sequences (Figure 6.2). The sensitivity of the searches can be further improved by incorporating available structural information, which is often quite extensive. In the case of C-type lectins, comparison of the CTLD structures available in the Protein Data Bank [10] allowed us to identify determinants of the CTLD fold which are shared by the protein sequences but for which no significant similarity could be detected using sequence comparison alone [1]. Incorporation of these findings, i.e., the sequence "signature" of the CTLD fold, into searches for new family members can help to discriminate between true homologs and spurious sequence similarities, a critical step of the search process.

A comprehensive study of a family the size of the CTLD family, even within one organism [3,5–7], poses a formidable data-handling task. According to our estimate, the GenBank database contains more than 4000 CTLD-containing expressed sequence tags (ESTs) from various species. The figure is doubled if all the sequences predicted in various genome annotation projects are taken into account. The number of distinct domain sequences is even larger, as some CTLDcps contain more than one CTLD. The amount of bibliographic and other structural and functional information is also substantial. A PubMed search with CTLD-specific text queries returns more than 10,000 matches, and there are now more than 100 Protein Data Bank entries containing experimentally determined CTLD structures.

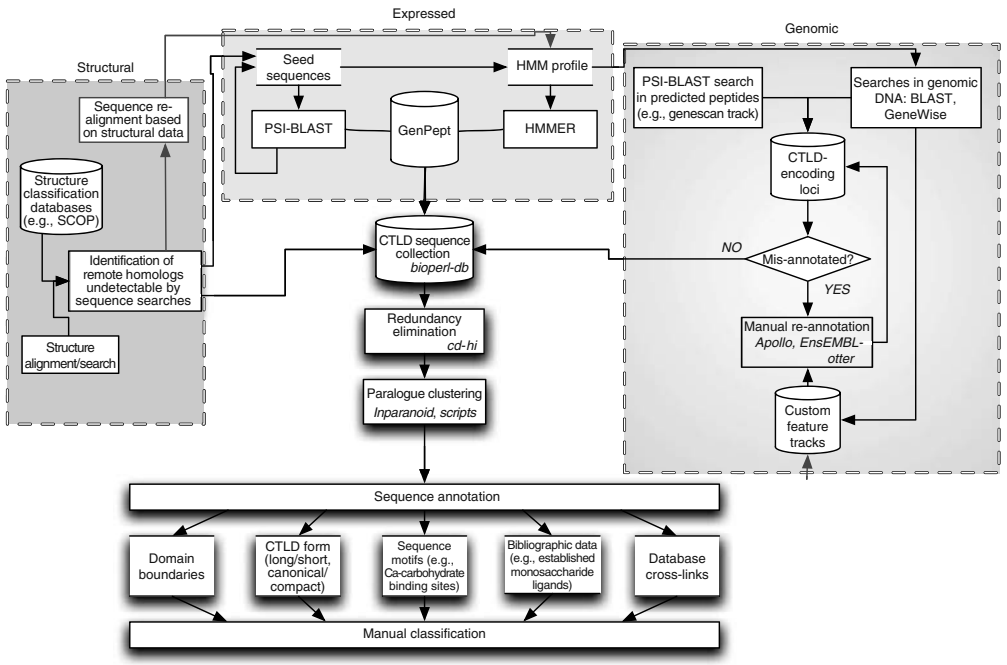


FIGURE 6.2 Building a CTLD database. Integration of sequence, structural, and genomic data on the CTLD family to create a comprehensive collection of CTLD sequences and to annotate it. Names of some of the software tools useful at various stages are shown in italics.

However, some of the sequence data is repetitive and it varies in quality. Minor variations in sequences deposited by different contributors result in the sequence corresponding to a single gene being present more than once, even in the nonredundant database subset. Therefore, a collection of sequences produced by a homology search needs to be clustered at several levels, first to determine duplicate occurrences of the same sequence (due to sequencing errors or polymorphisms), then to identify different protein forms expressed from the same gene (alternatively spliced or posttranslationally modified), and then further to define orthologs from closely related species, which share high levels of sequence similarity and often the same name. This process can be facilitated by specialized software such as cd-hi [11]. The collected and clustered sequences will require annotation before they can be analyzed conveniently to determine phylogenetic relationships and predict possible functions. As shown in Figure 6.2, the annotation (linking of additional information) can include known details on the protein structure, and sequence motif and bibliographic information on reported or predicted functions. The analysis can include, for example, selecting representative sequences from each cluster and replacing alphanumeric database accession numbers with meaningful identifiers containing information about protein name and species of origin. Automation of all steps in this process is possible and desirable, as it allows researchers to work with large numbers of sequences and to repeat the searches as the source databases are updated or to optimize the search parameters. Collections of software modules for popular computer programming languages are available such as Perl, Python, or Java, which simplify such automation. To build our own collection of CTLD sequences, and classify and annotate them with phylogenetic, functional, and bibliographic information, we used modules from the BioPerl toolkit [12].

The collection of sequences and annotation information can be stored as a set of flat-text files in Fasta or GenBank formats and accessed and maintained using common office software. Alternatively, if the number of sequences is large and annotations are complex, both sequence and annotated information can be stored in a relational database. Although requiring greater effort to set up, such an arrangement can be very powerful by allowing researchers to pose complex questions, such as "How many CTLDs from *C. elegans* are long-form and contain a predicted sugar-binding site with specificity for mannose," using simple queries. Our implementation of this approach was based on the MySQL database engine, BioPerl-db schema and a custom web interface and application programming interface (API) for accessing and modifying the information [13].

Compared with cDNA/EST sequence database searches, the whole-genome approach has the important advantages of being more comprehensive and allowing the identification of novel sequences that escaped large-scale and targeted cDNA cloning efforts due, for instance, to large transcript size or low expression levels. Analysis of the complete or nearly complete genome sequence data now available for all major model organisms will allow the evolution of individual lectins, and of the families as a whole, to be tracked. The genome of the puffer fish, *Fugu rubripes*, has proven to be a particularly useful tool to study the evolution of vertebrate protein families. It has been available since 2002 [14] and was the second vertebrate genome sequenced. It is eight times smaller than the human genome and, thus, is a convenient comparative model for analyzing the much larger mammalian genomes because of its compactness, low content of repetitive elements, and the relatively large evolutionary distance between fish and mammals, which is estimated to be about 430 Myr [15].

Genome sequence annotations can contain both systematic and sporadic mistakes for gene predictions. These need to be corrected before any conclusions can be made about the evolution of the protein. For example, in our study of the CTLDcps in the fish genome [7], we noticed that almost all *Fugu* proteins in the official Ensembl annotations were soluble proteins, whereas very few human CTLDcps are. Simple comparison with the GenScan [16] features overlapping the CTLD-encoding genes showed that absence of the transmembrane (TM) domains is a result of coding sequence (CDS) misprediction rather than a fundamental difference in *Fugu* CTLDcps. GenScan predictions, in turn, could not be used as a basis for our analysis because they sometimes contain regions that are absent from human or mouse orthologs, and often merge neighboring genes. Another general problem

with the Ensembl gene prediction was observed with proteins that had a previously unknown domain architecture. In such cases, individual domains had been separated into separate gene models. Yet another systematic problem we found was fragmentation of the predicted genes due to gaps in the assembly. In addition to these systematic problems, there were multiple sporadic ones. For example, a cluster of tandemly duplicated genes can be collapsed by the prediction algorithm into a single gene structure with several "alternatively spliced" transcripts, as was the case for the cluster of three selectin genes (P-, L-, and E-) in the *Fugu* genome. Manual revision of the predicted structures for misannotated genes using supporting evidence available in the original database, as well as additional evidence specific to the studied protein family, can solve such problems. Depending on how extensive the required reannotations are, various software can be useful; if working with individual loci, fragments of genomic DNA can be imported from the database and reannotated either manually or using an annotation editor such as Artemis [17]. To accommodate the annotations for more than 100 of the predicted CTLD-encoding loci in the *Fugu* genome, we used the Apollo genome browser [18] backed up by an Otter database [19] from the Ensembl project. Overall, as in our *Fugu* study, finding, checking, and annotating genome sequences can involve a considerable amount of work (cf. the steps in Figure 6.2) but can be very rewarding; in our case, we verified 32 gene structures and predicted 63 new ones [7].

Discovering novel superfamily members in existing database sequences is one of the most important and exciting outcomes of a systematic computer-based study. And despite all the limitations of the draft *Fugu* sequence, we found it to be a powerful instrument for gene discovery. We predicted putative *Fugu* orthologs for several uncharacterized mammalian CTLDcps (Bimlec, MGC3279, KIAA0534, CETM, SEEC, CBCP/Frem1/QBRICK, NLSLH) that are well conserved between *Fugu* and mammals [7]. Some of the predictions were supported by mammalian cDNA sequences from public databases, while for others (NLSLH and CBCP) no full-length cDNA from any organism was found in DBs. Three of the novel CTLDcps could not be attributed to any of the groups known at the time we performed our study, and we suggested extensions to the classification to accommodate them [4,7]. A large (~2100 aa) proteoglycan, which we called Calx-β CTLD binding protein (CBCP), which contains a set of chondroitin sulfate proteoglycan (CSPG) repeats [20] (homologous to the NG2 ectodomain [21]), a calcium-binding Calx-β domain [22], and a CTLD is a novel member of a well-studied protein family (NG2), which had not been reported previously to have members containing CTLDs. Our prediction has since been fully confirmed by two independent studies [23,24].

6.3 FUNCTIONAL ANNOTATION OF NOVEL C-TYPE LECTINS

The integrated coherent experimental and analysis approach of Drickamer and his colleagues has provided an in-depth understanding of many of aspects of the structural mechanisms of sugar binding by C-type lectin carbohydrate recognition domains (CRDs), and established a basis for developing bioinformatics techniques for predicting CTLD sugar-binding properties with substantial reliability by sequence analysis [25].

Residues with carbonyl side chains involved in Ca^{2+} coordination in site 2 form two characteristic motifs in the CTLD sequence and together with the calcium atom itself are directly involved in monosaccharide binding [2]. The molecular mechanism of Ca^{2+}-dependent carbohydrate binding is conserved in all family members studied; the amino acids that form the core of the binding sites have characteristic motifs ("EPN" and "WND") that can be identified by sequence similarity, and are indicative of the binding specificity (mannose vs. galactose). These observations provide a simple and very popular approach to predicting whether a CTLD of unknown function is likely to bind sugar ("EPN" and "WND" present), and whether it would preferentially bind mannose- or galactose-type ligands ("EPN" vs. "QPD"). This simple prediction technique [25] is widely used and has proven to be reliable in many cases, although, as discussed next, the mechanism of such selectivity is not completely explained.

The first crystallographic study of a complex between a CTLD and a carbohydrate was carried out on rat mannose-binding protein (MBP-A) and the *N*-glycan Man_6-$GalNAc_2$-Asn [26]. In the structure obtained, a ternary complex between the terminal mannose moiety of the oligosaccharide, the Ca^{2+} ion bound in site 2 and the protein was observed. The complex is stabilized by a network of coordination and hydrogen bonds: oxygen atoms from the 4- and 3-hydroxyls of the mannose form two coordination bonds with the Ca^{2+} ion and four hydrogen bonds with the carbonyl side chains that form the Ca^{2+}-binding site 2 (Figure 6.3). This bonding pattern is fundamental for CTLD/Ca^{2+}/monosaccharide complexes, and is observed in all known structures. It is also a major contributor to the binding affinity, especially in CTLDs specific to the mannose group of monosaccharides. For example in MBP-A, mannose atoms form very few interactions with the protein other than hydrogen/coordination bond formation by the two equatorial hydroxyls, and extensive mutagenesis screening has shown that the only other significant contributor to mannose binding is the Cβ group from His^{189} that forms a hydrophobic interaction with the sugar [27].

The positioning of hydrogen donors and acceptors in the binding sites has two important consequences. First, it determines the overall positioning and orientation of the ligand in the binding site. It may be seen from Figure 6.3A that the sugar-binding site of CTLDs has a twofold symmetry axis relating the sugar hydroxyls and that the hypothetical sugar shown could be rotated by 180° without introducing any changes to the bonding scheme. It is now known that this is indeed the case, although some early modeling and mutagenesis studies were based on the assumption that the orientation of the sugar was fixed. However, when the structure of a complex between rat MBP-C with mannose was determined, the orientation of the bound mannose was opposite to the orientation observed in

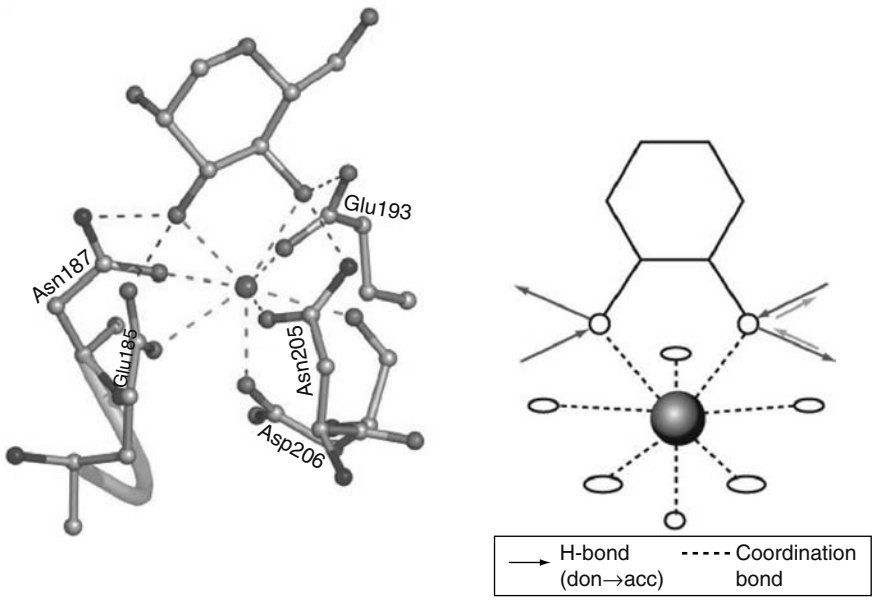


FIGURE 6.3 (See color insert following blank page 170. Also see CD for color figure.) Ca^{2+}-dependent monosaccharide binding by CTLDs. (A) Structure of the complex of mannose-binding protein A, Ca^{2+} ion and mannose (PDB 2msb). Coordination bonds are orange. Hydrogen bonds where the sugar hydroxyls act as acceptor and donor are red and blue, respectively. The Ca^{2+} atom is shown as a blue sphere. (B) A schematic representation of a Ca^{2+}-hexose-CTLD complex. Two hydroxyl oxygens and the ring of the hexose are shown. The Ca^{2+} atom is shown as a large gray sphere, and oxygens as circles and ovals. Protein groups that act as hydrogen donors and acceptors are not shown. Arrows show the direction of hydrogen bonds in mannose-specific CTLDs, while light-gray arrows indicate the changed directions in galactose-specific CTLDs.

MBP-A [28], and further studies revealed some of the factors that determine the preferred orientation [29]. Although the rat MBPs are the only established example of a CTLD that can bind carbohydrates in both orientations, it is known that different CTLDs bind the same monosaccharide in different orientations, for example binding of galactose to the galactose-binding MBP-A mutant and CEL-I versus that for TC-14 lectin.

The second constraint imposed by the Ca^{2+}-coordination site on the ligand defines the properties of the carbohydrate hydroxyls that the site can accept. This is best demonstrated by the mechanism of discrimination between the mannose and galactose groups of monosaccharides by CTLDs. Early in the history of CTLDs, an important correlation between the residues flanking the conserved *cis*-proline in the long loop region (Figure 6.1), which are involved in Ca^{2+}-binding site formation and the specificity for either galactose or mannose was made. In all mannose-binding proteins known at that time, the sequence of the motif was EPN (E185 and N187 in MBP-A), whereas in the galactose-specific CTLDs it was QPD (Gln-Pro-Asp). In a series of elegant mutagenesis experiments, Drickamer has shown that replacing the EPN sequence in MBP-A with a galactose-type QPD sequence is sufficient to switch the specificity to galactose [30], and that further modifications around the binding site (mainly introduction of a properly positioned aromatic ring to form a hydrophobic interaction with the apolar face of the sugar) can increase the affinity and specificity of the mutant MBP-A for galactose to the level observed in natural galactose-binding CTLDs [27].

Crystallographic analysis of the galactose-specific MBP-A mutant showed that the EPN to QPD change does not cause any serious restructuring of the Ca^{2+}-binding site 2 geometry [31]; this suggested that the key switch in the specificity was induced by swapping the hydrogen-bond donor and acceptor across the monosaccharide-binding plane and changing the hydrogen-bonding pattern from the mannose-type asymmetrical (Figure 6.3B, dark-gray arrows) to galactose-type symmetrical (Figure 6.3B, light-gray arrows). The same distribution of hydrogen-bonding partners was observed in the galactose-binding lectin TC-14 from the tunicate *Polyandrocarpa misakiensis* [32]. The TC-14 CTLD contains an unusual EPS motif in the long loop region, which is similar to the motifs of the mannose-binding proteins but contains a serine as a hydrogen-bond donor instead of the asparagine in MBP-A. The crystal structure revealed that due to a compensatory change on the opposite side of the ligand-binding site (the "WND" motif is changed to LDD), and a 180° rotation of the galactose residue compared with the orientation observed in the galactose-binding MBP-A mutant, the symmetrical pattern of the hydrogen bonding is maintained.

Although many of the determinants of the monosaccharide-binding specificity have been established experimentally, the mechanism underlying them is still unclear [4]. Mutual spatial disposition of bonded hydroxyls, which was initially suggested to be the main contributor to the specificity, is no longer considered so important; a growing number of crystal structures of CTLDs with the MBP-A-like ("asymmetrical") distribution of hydrogen-bond donors and acceptors have demonstrated that the core binding site is compatible not only with any two equatorial hydroxyls (3- and 4-OH of mannose and glucose, 2- and 3-OH of fucose), but also with a combination of an axial and an equatorial hydroxyl (3- and 4-OH of fucose, as in E- and P-selectin structures). A comparative study of different lectin–carbohydrate complexes published by Elgavish and Shaanan [33] suggests that additional stereochemical factors need to be taken into consideration. These authors noted the unique clustering of hydrogen-bond donors and acceptors around the 4-OH hydroxyl group in all structures they compared, which was not observed for other hydroxyls: in a Newman projection along the O4–C4 bond, hydrogen bond acceptors are never *gauche* to both vicinal ring carbons (C3 and C5), and thus the 4-OH proton is always pointing outside the ring. Poget et al. [32] confirmed this observation and also noted that in CTLDs the same rule is also true for the 3-OH proton. However, no explanation of the unique stereochemistry of the 4-OH binding orientation has been offered.

In summary, whole-genome studies of the CTLD family published by Drickamer and his colleagues focused on the evolution of the carbohydrate-binding properties and used the prediction methods discussed above to classify the CTLDcps they found according to their likelihood or not to bind carbohydrate at the main sugar-binding site, and to ascribe monosaccharide specificity for the

former [3,5,6]. Although our approach and scope for the *F. rubripes* genome study was somewhat different [7], for the carbohydrate-binding prediction part, we used the techniques developed by Drickamer and coworkers.

6.4 CONCLUSION AND FUTURE DIRECTIONS

Systematic comparative studies have provided important insights into the function of the C-type lectin family [4]. Sequence and tertiary structure comparison have led to definition of residue motifs associated with the Ca^{2+} and carbohydrate binding, as well as crucial interactions between different parts of the domain stabilizing the fold [1,2]. Classification of the family members into groups has facilitated prediction of the oligomerization and ligand-binding properties for newly found members [2,3,7,25]. Comprehensive genome-wide surveys and interspecies comparisons have led to detection of novel members of this otherwise well-studied family [3,5–7]. Integration of various bioinformatic tools and databases with custom software has made it possible to combine sequence, structural, and genomic information for thousands of proteins from a family the size of CTLDcps to revise and annotate it in a manual or semiautomated manner, and to query the resulting high-value-added data in a flexible way, finding quick answers to questions that would otherwise be intractable [7,13].

Application of these systematic and comprehensive approaches to CTLDcps has greatly expedited understanding of the evolution of structure and function of the whole superfamily, especially the carbohydrate-binding C-type lectin groups, and provided a molecular level view of how the functional versatility of the domain is related to its sequence and 3D structure. The results of these studies have posed many new questions and provided a strong framework to guide design of future experimental and bioinformatic investigations.

REFERENCES

1. Zelensky, A. N. and Gready, J. E. 2003. Comparative analysis of structural properties of the C-type-lectin-like domain (CTLD). *Proteins* **52**, 466–477.
2. Drickamer, K. 1993. Evolution of Ca^{2+}-dependent animal lectins. *Prog Nucleic Acid Res Mol Biol* **45**, 207–232.
3. Drickamer, K. and Fadden, A. J. 2002. Genomic analysis of C-type lectins. *Biochem Soc Symp* 59–72.
4. Zelensky, A. N. and Gready, J. E. 2005. The C-type lectin-like domain superfamily. *FEBS J* **272**, 6179–6217.
5. Drickamer, K. and Dodd, R. B. 1999. C-Type lectin-like domains in *Caenorhabditis elegans*: Predictions from the complete genome sequence. *Glycobiology* **9**, 1357–1369.
6. Dodd, R. B. and Drickamer, K. 2001. Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity. *Glycobiology* **11**, 71R–79R.
7. Zelensky, A. N. and Gready, J. E. 2004. C-type lectin-like domains in *Fugu rubripes*. *BMC Genomics* **5**: 51.
8. Altschul, S. F., et al. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
9. Eddy, S. R. 1998. Profile hidden Markov models. *Bioinformatics* **14**, 755–763.
10. Berman, H. M., et al. 2002. The protein data bank. *Acta Crystallogr D Biol Crystallogr* **58**, 899–907.
11. Li, W., Jaroszewski, L., and Godzik, A. 2001. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**, 282–283.
12. Stajich, J. E., et al. 2002. The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* **12**, 1611–1618.
13. Zelensky, A. N. 2004. *In silico analysis of C-type lectin domains' structure and properties*, PhD thesis, Australian National University.
14. Aparicio, S., et al. 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301–1310.
15. Ahlberg, P. E. and Milner, A. R. 1994. The origin and early diversification of tetrapods. *Nature* **368**, 507–514.

16. Burge, C. and Karlin, S. 1997. Prediction of complete gene structures in human genomic DNA. *J Mol Biol* **268**, 78–94.
17. Berriman, M. and Rutherford, K. 2003. Viewing and annotating sequence data with Artemis. *Brief Bioinform* **4**, 124–132.
18. Lewis, S. E., et al. 2002. Apollo: A sequence annotation editor. *Genome Biol* **3**, 1–14.
19. Searle, S. M., et al. 2004. The Otter annotation system. *Genome Res* **14**, 963–970.
20. Staub, E., Hinzmann, B., and Rosenthal, A. 2002. A novel repeat in the melanoma-associated chondroitin sulfate proteoglycan defines a new protein family. *FEBS Lett* **527**, 114–118.
21. Nishiyama, A., et al. 1991. The primary structure of NG2, a novel membrane-spanning proteoglycan. *J Cell Biol* **114**, 359–371.
22. Schwarz, E. M. and Benzer, S. 1997. Calx, a Na–Ca exchanger gene of *Drosophila melanogaster. Proc Natl Acad Sci USA* **94**, 10249–10254.
23. Smyth, I., et al. 2004. The extracellular matrix gene Frem1 is essential for the normal adhesion of the embryonic epidermis. *Proc Natl Acad Sci USA* **101**, 13560–13565.
24. Kiyozumi, D., et al. 2005. Identification of a novel cell-adhesive protein spatiotemporally expressed in the basement membrane of mouse developing hair follicle. *Exp Cell Res* **306**, 9–23.
25. Drickamer, K. and Taylor, M. E. 2003. Identification of lectins from genomic sequence data. *Methods Enzymol* **362**, 560–567.
26. Weis, W. I., Drickamer, K., and Hendrickson, W. A. 1992. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* **360**, 127–134.
27. Iobst, S. T. and Drickamer, K. 1994. Binding of sugar ligands to Ca^{2+}-dependent animal lectins. II. Generation of high-affinity galactose binding by site-directed mutagenesis. *J Biol Chem* **269**, 15512–15519.
28. Ng, K. K., Drickamer, K., and Weis, W. I. 1996. Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. *J Biol Chem* **271**, 663–674.
29. Ng, K. K., et al. 2002. Orientation of bound ligands in mannose-binding proteins. Implications for multivalent ligand recognition. *J Biol Chem* **277**, 16088–16095.
30. Drickamer, K. 1992. Engineering galactose-binding activity into a C-type mannose-binding protein. *Nature* **360**, 183–186.
31. Kolatkar, A. R. and Weis, W. I. 1996. Structural basis of galactose recognition by C-type animal lectins. *J Biol Chem* **271**, 6679–6685.
32. Poget, S. F., et al. 1999. The structure of a tunicate C-type lectin from *Polyandrocarpa misakiensis* complexed with D -galactose. *J Mol Biol* **290**, 867–879.
33. Elgavish, S. and Shaanan, B. 1997. Lectin-carbohydrate interactions: Different folds, common recognition principles. *Trends Biochem Sci* **22**, 462–467.

7 Animal Models for Assessing the Biological Roles of Lectins

Hafiz Ahmed, Gabriel A. Rabinovich, Shawn S. Jackson, Mariana Salatino, Keiko Saito, German Bianco, Satoshi Tasumi, Shao-J. Du, and Gerardo R. Vasta

CONTENTS

7.1 INTRODUCTION

Cell surface glycans, such as glycocoproteins and glycolipids, encode information that modulates interactions between cells, or between cells and the extracellular matrix (ECM), by specifically regulating the binding to cell surface-associated or soluble carbohydrate-binding receptors such as lectins. Since the time of their discovery in the late nineteenth century and beyond a strictly utilitarian role as useful reagents, lectins from numerous invertebrate and vertebrate species have been subject to a variety of studies aimed not only at the characterization of their biochemical, molecular and structural properties, but also to gain insight into their biological roles. In this regard, lectins have constituted an enigmatic group of proteins and it has long been deemed reasonable to postulate that their physiological roles in any particular cell or tissue relate in some way to their carbohydrate-binding properties. Thus, detailed and useful descriptions of biochemical properties such as sugar specificity and requirement of cations for binding, as well as subcellular and tissue distribution are available for many animal lectins. For some animal lectin families, comparative approaches have made substantial contributions to the understanding of their evolution in their gene organization, structural aspects, and biochemical properties, and the early reports focused primarily on their possible roles in fertilization and immune responses. Although these studies provide a solid foundation for gaining a better understanding of these proteins, conclusions drawn in regard to their functional aspects have been mostly speculative. The modern era of research on animal lectins has seen a vast expansion on these foundations and a considerable body of experimental evidence from biochemical and structural approaches has led to fairly sound conclusions about their diverse roles within intracellular compartments, at the cell surface, and between cells or cells and the ECM, thereby mediating roles in intracellular trafficking, protein folding, and signaling in various aspects of development and immunity (see the appropriate chapters in later sections of this volume). It is only in the past decade, however, that the use of genetically tractable animal model systems and the availability of their fully sequenced genomes has allowed us to fully appreciate the diversity of their lectin repertoires and has enabled the rigorous demonstration of function for any given member.

7.2 INVERTEBRATES AND PROTOCHORDATES

7.2.1 Studies on Lectins from Invertebrate and Protochordate Species

Lectin repertoires of invertebrates and protochordates are ample and complex, with representatives of most lectin families described so far in vertebrates, including mammals. The identification of lectin consensus carbohydrate-recognition domain (CRD) sequences in the genomes of invertebrate models, including *Drosophila melanogaster* [1], *Caenorhabditis elegans* [2], and *Anopheles gambiae* [3], has significantly contributed to the full characterization of their lectin repertoires, and provided insight into their biological and evolutionary roles [4,5].

In 1899, a lectin secreted from the albumin gland of the snail *Helix pomatia* was described as having agglutination properties for erythrocytes [6]. In later studies, the designation of lectins from snail eggs as "protectins" was proposed to suggest that the lectins provide internal protection for the eggs and embryos against microbial or fungal infection, possibly by immobilizing the microbes into cross-linked aggregates that would be readily phagocytosed and digested by phagocytic cells [7]. Further characterization of the *H. pomatia* hemagglutinin (HPA) [8] led to its use for bacterial classification for determining the antigenic structures of cell surfaces and for analyzing glycolipids and glycoproteins [9]. More recently, resolution of the structure of HPA has further advanced our understanding of animal lectin structure though a novel β-sandwich lectin fold as well as a novel hexameric quaternary state [10].

Studies on arthropod species, including chelicerates such as the American horseshoe crab *Limulus polyphemus*, and crustaceans such as the lobster *Homarus americanus*, the crab *Eupagurus*

prideauxii, and the spider crab *Maia squinado*, have provided evidence that lectin expression could be induced by immune challenge and exhibited opsonic properties [11–15]. Further studies on horseshoe crabs, including *L. polyphemus*, the Japanese *Tachypleus tridentatus*, and the Indian *Carcinoscorpius rotundicauda*, and other chelicerates have yielded useful information on the biochemical properties, molecular structures, and biological functions of various defense molecules, including sialic acid binding lectins [15–17]. The *T. tridentatus* innate immune recognition system includes humoral and cellular lectin repertoires with different carbohydrate specificities, with five types of tachylectins (TL-1 to -5) and several bacterial agglutinins, which act synergistically to defend the horseshoe crab against invading microbes and foreign substances [16–18]. Further, tachylectin TL-P, which has a similar sequence to that of TL-1, is also present in the perivitelline fluid of the horseshoe crab [19] and may thus represent a lectin important during embryogenesis [17]. Horseshoe crab hemocytes are filled with secretory granules that store these defense molecules, while its hemolymph plasma contains additional proteins involved in innate immunity such as C-reactive proteins (CRPs) [20], α2-macroglobulin [21,22], and hemagglutinins [16–18,20,23].

Among insects, because of their large size and hemolymph volumes, lepidopteran larvae, particularly those of saturnid sphingid moths, have been used for studies on the innate immune responses including biochemical analyses of lectins and antimicrobial peptides, and studies of hemocyte function. The tobacco hornworm *Manduca sexta*, for example, is easy to rear and is well suited for studies of hemocytes and hemolymph proteins as the last instar larva reaches 10–12 g, and ~10^6 hemocytes can be collected from each individual. Although genetic studies are impractical with *M. sexta*, it is an excellent subject for experiments requiring protein purification and for the study of protein interactions that occur in hemolymph [24]. *M. sexta* hemolymph harbors a group of proteins which bind to microbial surface molecules: hemolin, peptidoglycan recognition proteins (PGRPs), β-1,3-glucan recognition proteins (βGRPs), and C-type lectins (immulectins, IMLs) [24,25–31]. These hemolymph proteins function as pattern recognition receptors (PRRs) [30], forming complexes that stimulate immune responses including the prophenoloxidase (proPO) activation system. Hemolin synthesis in *M. sexta* is both developmentally regulated and strongly induced by microbial challenge [26,32]. *M. sexta* immulectin functions are discussed in detail in Chapter 26.

Similarly, studies on a variety of insect species such as the fleshfly *Sarcophaga peregrina*, the American cockroach *Periplaneta americana*, the silkworm *Bombyx mori*, and the West Indian leaf cockroach *Blaberus discoidalis* have contributed to our understanding of lectins involved in pathogen recognition and engulfment. These large insects are again useful for the biochemical purification of candidate molecules as their volume of hemolymph is substantial [33]. *Sarcophaga* lectin was the first C-type lectin-like domain (CTLD)-containing protein (CTLDcp) discovered in insects [34,35]. Although expressed in response to injury [35], its developmental expression pattern suggests it may also play a role in the development of the wing and leg imaginal disk [36]. Another C-type lectin, granulocytin, was purified from *Sarcophaga* hemocytes. Granulocytin is secreted into hemolymph in response to immune stimulation and shows significant similarity to *Drosophila* lectin CRD-like sequences [37]. Two lipopolysaccharide (LPS)-binding C-type lectins identified in *P. americana* [38,39] are suggested to function as opsonins upon infection [39,40]. One, regenectin, is also expressed in regenerating leg [41–43]. Three C-type lectins with different sugar specificities as well as a β-1,3-glucan-specific lectin have been identified in *B. discoidalis* [44,45]; these activate both proPO and the phagocytosis of microorganisms [46–48]. A mannose-binding homolog to the *B. discoidalis* β-1,3-glucan-specific lectin has been found in the mosquito *A. stephensi* [49], while in *A. gambiae*, candidate antimalarial immune genes include a mannose- and a galactose-binding lectin [50]. In addition, two lectins found in the mosquito *A. stephensi* represent the first developmental stage-specific and/or sex-related lectins identified in an insect [49].

Oysters are bivalve mollusks relevant to fisheries and aquaculture industries as well as to the ecosystem. As filter-feeding organisms, they play a critical role in maintaining water quality [51]. The eastern oyster (*Crassostrea virginica*) has been used as an important model to study the effects of environmental stressors [52]. However, the protozoan parasite *Perkinsus marinus* causes Dermo disease in the *C. virginica* and is responsible for catastrophic damage to shellfisheries and the estuarine environment in North America. The infection mechanisms remain unclear, but it is likely that, while filter feeding, the healthy oysters ingest *P. marinus* trophozoites released to the water column by the infected neighboring individuals. Recently, large-scale expressed sequence tag (EST) resources have revealed at least four C-type lectins and a tandem-repeat galectin [53]. Additionally, a novel galectin of unique domain organization (four carbohydrate-binding domains) from hemocytes of the *C. virginica* has recently been identified and cloned [54]. The evidence indicates that this new galectin facilitates recognition of selected microbes and algae, thereby promoting phagocytosis of both potential infectious challenges and phytoplankton components, and that *P. marinus* subverts the host's immune/feeding recognition mechanism to passively gain entry into the hemocytes.

Within the deuterostome lineage, echinoderms such as the purple sea urchin (*Strongylocentrotus purpuratus*) are important model organisms for developmental biology, immunobiology, and gene regulation studies. The Sea Urchin Genome Sequencing Consortium [55] has revealed several lectins including C-type lectins, and genes (homologs of C3, factor B, and mannose binding protein associated serine protease) that mediate the alternative and lectin complement pathways [56]. The unique intermediary position in evolution occupied by tunicates (Urochordata), between invertebrates and vertebrates, suggests these organisms may encompass extant examples of ancient defense mechanisms. Multiple lectins (C-type, galectins, ficolins, and others) are present in plasma of the colonial tunicate *Clavelina picta*, among which a fucose-binding protein is a homolog of the mammalian mannose-binding lectin (MBL) [57]. The discovery of MBL-associated serine proteases in another tunicate species [58], key molecular partners in the lectin-mediated complement pathway found in humans [59,60], is further proof that these model organisms harbor homologs of mammalian defense molecules. Study of the lectin-mediated defense pathways in these model organisms is very useful since factors that would otherwise compete for a lead role in defense in a higher model system, such as antibodies, are absent in tunicates.

7.2.2 Assessment of Lectin Function Using Genetically Tractable Invertebrate and Protochordate Model Systems

7.2.2.1 *Caenorhabditis elegans*

Relatively recently, insight into lectin functions has been gained from the model organism *C. elegans* [2], considered to be the most completely understood metazoan in terms of anatomy, genetics, development, and behavior [61]. The complete genome sequence of *C. elegans* provides an opportunity to gain a global picture of the role of proteins in a simple multicellular organism [2,62]. Extensive and powerful genetic methodologies have been developed for this organism [61]. *C. elegans* has further emerged as a model system for investigating innate immunity. Since *C. elegans* has no cellular immune system, the coating of pathogens by *C. elegans* lectins might play a simple yet essential neutralization role [5]. The worm has been found to mount pathogen-specific protective responses to a variety of fungal and bacterial pathogens. Proteins which might enable such recognition include numerous CTLDcps found within the worm genome [5]. Screening methods specifically designed to monitor increased sensitivity to infection have succeeded in finding components of the immune signaling apparatus, suggesting that further experiments in the same vein will, in time, define these molecules of recognition [5,63]. In addition to numerous C-type lectins (discussed in the following sections), the completed *C. elegans* genome [2] revealed 26 putative galectin sequences [64]. Further, homologs for two types of lectins involved in sorting events within lumenal compartments of cells, including a calnexin precursor [65], and two proteins homologous to the L-type lectins, ERGIC-53 and VIP-36

[66,67], were found. However, no mannose 6-phosphate receptors homologs have been identified in *C. elegans* [62]. A comprehensive analysis of the *C. elegans* genome identified numerous CTLDcps. Nineteen of the 183 CTLDs (within 135 proteins [5,68]) identified in *C. elegans* contain most of the five residues needed to form the primary Ca^{2+}-binding site found in vertebrate C-type CRDs [4,62]. Four of these 19 show perfect conservation of the five Ca^{2+} -binding site residues found in rat serum mannose-binding protein (MBP) [62]. The preferential binding of mannose and *N*-acetyl-glucosamine afforded by these residues in MBP enables this protein to specifically recognize carbohydrates on pathogenic microorganisms. Thus, these *C. elegans* C-type-lectin-like proteins may function similarly to MBP [5,69], while the remainder of the identified CTLDcps may possess novel ligand binding properties. The anatomical expression patterns and regulation of expression of a number of the identified CTLDcps are consistent with functions in pathogen surveillance [5].

7.2.2.2 *Drosophila melanogaster*

Although large insects have proven extremely useful in purifying hemolymph lectin components for characterization, *in vivo* studies of these molecules are hampered by the lack of genetic tools in these insect models. In contrast, the genetic tools available in the fruitfly *D. melanogaster* are extremely powerful, aided by the available genomic sequence of this organism [1,33]. The close phylogenetic relationship of, for example, the fleshfly to the fruitfly, has enabled relatively easy identification of *Drosophila* homologs of lectins of interest [70] and approaches have been developed to test the role of candidate molecules *in vivo*. *Drosophila* is thus extremely useful in genetically dissecting the mechanisms of lectin recognition and will continue to provide new insights on the potential roles of lectins in immunity and development [33]. A prototype galectin and five tandem-repeat type galectins have been identified in the genome of *D. melanogaster* [1,71]. Homologs of both calnexin and calreticulin, important chaperones ensuring proper glycoprotein folding in the endoplasmic reticulum, are found in *Drosophila*. Orthologs of members of the L-type lectin family, ERGIC-53 and VIP-36, which are involved in protein sorting and trafficking, are also found in *Drosophila* [4].

Comparisons of protein sequences of known *Drosophila* lectins with EST sequences have identified 19 novel C-type lectin open reading frames in *Drosophila* [72]. Comparative analysis of these *Drosophila* C-type lectin genes continues to shed light on the evolution of C-type lectins [73]. Further, a cluster of three C-type lectin genes on the genome of *D. melanogaster* revealed the presence of more than 30 C-type lectin-like genes [4]. Analysis of these 32 *Drosophila* CTLDs revealed that only six show conservation of the primary Ca^{2+}- and sugar-binding sites characterized in mammalian C-type CRDs [4]. Functions of such CTLDs (and their respective proteins) remain to be determined. The domain organization of these 32 CTLDcps is strikingly different from the organization of both the known mammalian C-type lectins and the *C. elegans* CTLDcps (discussed below) [62]. The lack of similarity suggests that most of these CTLDcps serve distinct functions in each of the organisms [4]. Further analysis of lectins in *Drosophila* will elucidate their biological role in immune systems since little is yet known about the participation of these C-type lectin genes in immunity [4,71,73].

7.2.2.3 *Ciona intestinalis*

A large EST project [74] and the draft genome sequence [75] of another model organism, the urochordate *Ciona intestinalis*, has also enabled *in silico* searches for lectin molecules. *Ciona* and other urochordates are attractive models in comparative and evolutionary immunology as studies of their immune systems may shed light on the evolution of the innate immune system in deuterostomes as well as the emergence of adaptive immunity [76]. The ascidian *C. lavelina picta* has a prototype galectin, while prototype and tandem-repeat galectins have been identified in *Ciona* [62,75].

7.3 VERTEBRATES

7.3.1 Studies on Lectins from Vertebrate Species

Among ectothermic vertebrates such as fish, the C-type lectin, galectin, and pentraxin repertoires exhibit significant diversification [62]. In addition, selectin and macrophage mannose receptor gene sequences have been found in the genomes of zebrafish and pufferfish (zebrafish: www.zfin.org; pufferfish: Ensembl accession code SINFRUT00000162546). Furthermore, putative natural killer (NK) cell C-type lectin receptors (CLRs) have been identified in cichlid fish, demonstrating that these receptors appeared early in vertebrate evolution [62,77]. CLRs containing immunoreceptor tyrosine-based inhibition motifs have been identified in the rainbow trout [78]. Pentraxins are also present in teleost fish, although the architecture of the native C-reactive proteins may be very different from the human prototype [79–81]. Calreticulin and calnexin have also been identified in teleost fish [82,83]. In amphibians, lectin repertoires are equally diversified and the *Xenopus* spp. genomic databases have greatly contributed to their comprehensive analysis [64,84,85].

In fish, C-type lectins, galectins, and pentraxins have been identified from the earliest jawed vertebrate to the more advanced teleost species, while ficolins and P-type lectins have yet to be identified [64]. Considerable heterogeneity in lectins is seen in fish. Such heterogeneity is best demonstrated by the fucolectins of the Japanese eel (*Anguilla japonica*) [86] which exists in at least seven expressed isoforms. Studies on the inducibility of C-reactive proteins from serum of the Indian major carp (*Labeo rohita*) have demonstrated a shift in expression from the normal form of C-reactive proteins to several structurally different isoforms [87].

Several fish have proved to be useful model organisms for gaining insight into structural, functional, and evolutionary aspects of lectin biology. For example, the novel structure of the European eel *(A. anguilla*) agglutinin (AAA) defined not only a novel CRD sequence motif but also a novel lectin fold [88]. F-type lectins have since been identified in a variety of taxa, including teleosts (zebrafish, steelhead trout, stickleback, and pufferfish) and *Xenopus* spp. [64]. Similarity searches in genomic databases revealed that the F-type sequence motif is phylogenetically broadly distributed, present in molluscs and planaria, horseshoe crabs [89] and insects [1,3], echinoderms [90,91], and skates, as well as in *X. laevis* and salamander [64].

Additional novel lectin families discovered in fish include the rhamnose-binding lectins from salmonids and the pufflectins from *Fugu*. The rhamnose-binding lectins found in eggs of steelhead trout, catfish (*Silurus asetus*), and the white-spotted char (*Salvelinus leucomaenis*) are homologous to members of the low-density lipoprotein receptor superfamily [92]. Pufflectins are MBLs that exhibit intriguing homology to lectins of monocotyledonous plants with similar sugar specificities [93,94].

Fish harbor a broad range of functionally active lectins with variable specificity for a number of oligosaccharide targets, and are thus able to recognize an expansive repertoire of patterns on pathogen surfaces [95]. Such lectins include the MBL of Atlantic salmon (*Salmo salar*) [96] and the structurally similar ladderlectin of rainbow trout (*Oncorhynchus mykiss*) [97,98], the *N*-acetylgalactosamine-binding lectin of the blue gourami (*Trichogaster trichopterus*) [99,100], and the skin lectins of the Japanese eel (*A. japonica*) [101,102]. Pentraxins have been isolated in channel catfish, carp, snapper (*Pagrus auratus*), Atlantic salmon, common wolffish (*Anarhichas lupus*), cod (*Gadus morhua*), halibut (*Hippoglossus hippoglossus*), and rainbow trout [79–81]. Homologs of both calnexin and calreticulin are found in teleost fish [82,83], and of selectin and macrophage mannose receptor in zebrafish and pufferfish (zebrafish: www.zfin.org; pufferfish: Ensembl accession code SINFRUT00000162546).

Among teleost fish, prototype galectins have been identified in flounder (*Paralichthys olivaceus*), medaka (*Oryzias latipes*), Atlantic salmon (*S. salar*), rainbow trout (*O. mykiss*), striped bass (*Morone saxatalis*), Japanese pufferfish (*F. rupribes*) and green-spotted pufferfish (*Tetraodon nigroviridis*), whereas tandem-repeat galectins are present in catfish (*Ictalurus punctatus*), salmon (*S. salar*), and both pufferfish species (reviewed in Ref. [64]). Several galectin sequences (three prototypes, one chimera type, and two tandem-repeat types) have been identified and characterized

in zebrafish (*D. rerio*) (see Section 7.3.2.2) [103]. The chimeric tandem-repeat type galectin, identified in zebrafish (Ahmed and Vasta, submitted), and the tandem-repeat 4 type, present in the eastern oyster [54], are yet to be identified in mammals.

The best-studied of the fish CTLDcp are the serum antifreeze proteins (AFPs) found in cold-water-living sea raven, Atlantic herring, and smelt. Herring AFP provides an interesting example of a CTLD in which bound Ca^{2+} interacts with a noncarbohydrate ligand [104–110] (reviewed in Refs. [111,112]). Several other soluble fish CTLDcps have been described (reviewed in Ref. [111]): five isoforms of *S. salar* serum lectin [113], three collectins from different *Cyprinidae* carp family species [114], two C-type lectins from gills of Japanese eel [115], two lectins from rainbow trout liver [116], a carp lectin [117], and a liver lectin from *Gillichthys mirabilis* (long-jawed mudsucker) [118]. Known membrane-bound CTLDcps from bony fishes include a polycystic kidney disease protein 1 ortholog from *Fugu* [119], a rainbow trout Kupffer cell receptor homolog [120], and a set of putative killer cell receptors from cichlids [77]. The only known CTLDcp sequence from cartilaginous fishes is a tetranectin homolog from reef shark cartilage [121].

The *F. rubripes* genome [122] was the second vertebrate genome to be sequenced. It is eight times smaller than the human genome and is an effective genomic comparator because of its low content of repetitive elements and the relatively large evolutionary distance between fish and mammals. Further, the compactness of the *Fugu* genome makes it an extremely convenient reference sequence for identification of new genes [111]. An analysis of CTLD family members present in *F. rubripes* was conducted using the draft genome sequence [111]. The study verified 32 CTLD-containing gene structures and predicted 63 new ones. In general, the results showed that all but two groups of CTLDcps identified in mammals are also found in fish, and that most of the groups have the same members as in mammals. The *Fugu* genome study suggested a tentative link between the vertebrate and invertebrate CTLD families: CTLDcps with domain architecture similar to the *Fugu* dual-CTLD group are present in invertebrates. The F1 group of predicted dual-CTLD-containing proteins in *Fugu* may therefore be the first vertebrate group that has detectable homologs in invertebrates. Evidence suggests that F1 members from *Fugu* and zebrafish, and scarf proteins from *Girardia tigrina*, may have evolved from the same predecessor [111,112]. Further invertebrate dual-CTLD proteins include the insect immulectins [27] and nine proteins from *C. elegans* [62]. Further, putative *Fugu* orthologs were predicted for several uncharacterized mammalian CTLDcps. Specifically, two loci with similarity to the salmon AFPs were found. All three selectin genes found in other vertebrates are present in *Fugu* and have the same genome arrangement, while the NK cell receptor and lithostathine/Reg groups are not represented. In contrast, gene family expansions were seen in the DC-SIGN and MmanR groups [111]. In mammals, members of these subgroups are implicated in the binding and phagocytosis of a wide range of microorganisms [123,124]. Expansion of these groups in *Fugu* may reflect a larger role for innate immunity in host defense in lower vertebrates [111]. Both of the *Fugu* collectins are well conserved compared with their human orthologs (~76% identity), suggesting their functional importance in both organisms. Surprisingly, however, *Fugu* has no orthologs for MBL (or pulmonary surfactant proteins) as have been found in other fish species (*D. rerio*, *Cyprinus carpio*, and *Carassius auratus* [114]). Given the role of MBLs in complement activation in mammals, and their presence and level of conservation in the carp family, it is possible that the *Fugu* MBL ortholog exists but is not covered by the draft genome sequence [111]. A novel group of fish dual-CTLD proteins found through analysis of the *Fugu* genome is not found in other vertebrates. These predicted proteins of unknown function may be homologous to invertebrate proteins of similar domain structure and likely appeared at the earliest stages of CTLD family evolution. Further, they may represent the first CTLDcp group reported to be shared by both invertebrates and vertebrates [111,112].

Among the amphibians, galectins have been identified in members of the subclass Anura (frogs and toads). For example, a prototype galectin has been identified in *Rana catesbeiana* [125,126], several members of all three types of galectins in *Xenopus laevis* [127–129], and a prototype galectin in *Bufo arenarum* [130–134]. Galectin-like lectins have been found also in subclass Urodela, but

their structural features are not known [135]. Biochemical (subunit structure) and molecular characterization indicates that each amphibian species appears to have multiple galectins—the diversity is generally observed in other organisms [129,136]. All amphibian protogalectins showed remarkable structural similarities with the mammalian galectin-1. Sequencing analyses resulted in at least two galectin-1 like sequences for *X. laevis* and *R. pipiens*. Three additional protogalectins (two galectin-1 like sequences and one isoform of the previously described 16 kDa skin galectin) have been reported in *X. laevis* [128,129].

The largest and the best-studied group of nonmammalian vertebrate CTLDcps is that of the snake venom CTLDs [112]. Homologous sequences have been found in *X. laevis*, but not in fish, implicating snake CTLDs as representative of an ancestral group of CTLD proteins, which appeared after the split between Actynopterigia and Sarcopterygia [111,112]. Based on the occurrence of CTLDs in the venoms of snakes from the Viperidae and Elapidae families, it has been suggested that two different groups of CTLDs were recruited independently as venom toxins [137]. A third group of CTLDcps that independently evolved to support a newly acquired clade-specific function may then be represented by the phospholipase A2 inhibitors (PLIs), thought to be important in self-protection from venom toxicity [112]. The abundance and functional diversity of the CTLDcps from these subgroups provides a good example of the suitability of the domain for rapid generation of new physiological activities, and a parallel to the independent expansion of CTLDcps in mammals [112].

Avian lectin representatives have been identified in chicken, duck, and quail. The sequencing of the chicken genome [138] has confirmed that counterparts of the mammalian immune system, including paradigms of cellular immunity, exist in avian species [139–141]. The complete sequencing of the chicken major histocompatibility complex (MHC) identified lectins of two types (B-NK and B-Lec) [142]. The quail MHC carries six lectin genes similar to B-Lec and four NK lectin genes [143]. The presence of the avian lectin genes within the MHC suggests an evolutionary connection between these two arms of the immune defense against pathogens in birds [144]. An EST project to identify immune-relevant genes in the White Peking duck (*Anas platyrhynchos*), an important animal model for hepatitis B research, identified over 200 immune-relevant genes, including lectin-like immunoreceptors [144,145]. Two distinct families of lectins were identified in the duck. One family clusters with C-type lectins and shows similarity with the chicken gene 17.5 [145,146]. Other closely related genes include the MHC-encoded B-lectin of the chicken and the NKC-encoded murine CLR family [144,147]. The other family is weakly homologous to several antigen-presenting cell lectins, including dendritic cell immunoreceptor (DCIR) [145,148]. A group of CTLDcps has been found as a major component of bird eggshell matrix and includes proteins from chicken, goose, and ostrich [112,149–151]. Interestingly, additional CTLDcps interacting with the Ca-rich mineral phase have been reported in mollusk [152], sea urchin [153], and mammals [154,155]. Of these, only mammalian lithostathine can be considered homologous to the bird sequences, indicating that these disparate examples of CTLD involvement in calcite-containing biocomposite formation are the result of convergent evolution [112].

7.3.2 Assessment of Galectin Function Using Genetically Tractable Vertebrate Model Systems

7.3.2.1 Studies on Murine Models

Rodents, particularly mice, are well-established animal models in experimental research by virtue of their remarkable similarities with humans, short lifespan, rapid reproductive cycle, and their relatively simple maintenance and handling. A number of knockout models have been generated to examine the biological functions of endogenous glycan-binding proteins. Targeted disruption of lectin genes *in vivo* has provided a first round picture of the physiopathological relevance of lectins in a wide spectrum of biological events. In addition, compelling evidence demonstrates the therapeutic benefit of exogenous lectins or the inhibition of protein–glycan interactions in a variety of

pathological disorders including inflammation, neurodegeneration, and cancer [156]. We will highlight here some examples illustrating the use of mouse models to study the role of glycan-binding proteins, particularly galectins, in a number of physiopathological settings.

7.3.2.1.1 Role of Endogenous Galectins in Vivo: Lessons from the Study of Knockout Mice

In 1993, Poirier and Robertson reported the generation of galectin-1-deficient (*Lgals1*$^{-/-}$) mice; these mice were viable, fertile, and did not show major spontaneous phenotypic abnormalities, suggesting that other proteins might potentially compensate for the absence of this lectin [157]. However, after a careful examination of particular systems, these mice showed a deficiency in the olfactory network [158] and a reduced thermal sensitivity [159]. In these mice, the neuronal subpopulation in the olfactory bulb, which normally expresses galectin-1, does not reach the appropriate targets in the olfactory glomeruli and altered topography of the olfactory network is observed [158].

Regarding the modulation of neuronal processes, Plachta et al. recently showed, using *Lgals1*$^{-/-}$ mice, that galectin-1 actively participates in the elimination of neuronal processes by induction of cell death [160]. On the other hand, Sakaguchi et al. showed, following phenotypic analysis of these mice, that galectin-1 promotes proliferation but not death of adult neural stem cells in the adult brain [161]. In addition, evidence from the study of knockout mice revealed a positive role for galectin-1 in fusion of myoblasts and muscle regeneration *in vivo* after recovery from induced injury [162]. Therefore, it seems possible that galectin-1 may contribute to proliferation, death, or differentiation of distinct cell types depending on different circumstances including the differentiation and activation stage of target cells and the dominant physiological or pathological microenvironment.

Given the immunomodulatory effects of galectins in different experimental settings (Ref. [156] and Chapter 27), it has been hypothesized that these proteins may play a critical, but specific role in the initiation, amplification or resolution of innate and adaptive immune responses. Recent findings from our laboratory demonstrated that *Lgals1*$^{-/-}$ mice have enhanced susceptibility to Th1- and Th17-mediated neuroinflammation when challenged with pathogenic stimuli [163]. These findings were consistent with the ability of galectin-1 to selectively eliminate Th1 and Th17 cells, while sparing Th2 cells [163]. Further investigation of the mechanisms involved in this selective effect revealed that Th1 and Th17 cells express the repertoire of cell surface glycans that are critical for galectin-1 binding and subsequent cell death, while Th2 cells are protected from galectin-1 through increased α2,6-sialylation of cell surface glycoproteins [163]. In addition, we found that macrophages recruited to the peritoneal cavity of *Lgals1*$^{-/-}$ mice in response to inflammatory stimuli showed significantly increased expression of MHC-II, which conferred to these macrophages an enhanced capacity to stimulate the allogeneic T cell response *in vitro* [164]. In this regard, Garin et al. observed a reduced regulatory activity of CD4^{+} CD25^{+} T cells obtained from galectin-1 null mutant mice, indicating that galectin-1 is a key mediator of the immunosuppressive activity of regulatory T cells [165]. In addition, very recent work using *Lgals1*$^{-/-}$ mice revealed that galectin-1 also influences vascular responses by modulating chronic hypoxia-induced pulmonary hypertension [166].

Regarding other members of the galectin family, Colnot et al. [167] and Hsu et al. [168] independently reported a similar phenotype for galectin-3 null mutant (*Lgals3*$^{-/-}$) mice. Similarly to galectin-1-deficient mice, *Lgals3*$^{-/-}$ and double mutant mice [*Lgals1*$^{-/-}$; *Lgals3*$^{-/-}$] were viable, fertile, and had no developmental abnormalities [169]. However, further characterization of *Lgals3*$^{-/-}$ mice revealed an altered inflammatory response. A reduced inflammatory cell infiltration was observed in the peritoneal cavity of *Lgal3*$^{-/-}$ mice in response to challenge with thioglycollate [168]. Hence, *Lgals3*$^{-/-}$ mice were useful in elucidating the critical role of galectin-3 as a novel proinflammatory mediator. Most recently, Liu's group has investigated the relevance of galectin-3 as a modulator of mast cell functions, showing that *Lgals3*$^{-/-}$ mice exhibit reduced mast cell degranulation and diminished passive anaphylactic reactions compared with wild-type littermates [170]. Moreover, *Lgals3*$^{-/-}$ mice exhibited lower airway allergic inflammation in an experimental model of systemic sensitization with ovoalbumin followed by an airway antigenic challenge compared with wild-type

mice [171]. In addition, challenge of these mice with the intracellular parasite *Toxoplasma gondii* recently revealed a critical proinflammatory role of galectin-3 in the regulation of innate and adaptive immune responses [172]. *Lgals3*$^{-/-}$ mice infected with this parasite exhibited lower inflammatory infiltrates in the gut, liver and brain compared with wild-type mice [172]. It is noteworthy, however, that inflammatory foci were more pronounced in the lungs of *Lgals3*$^{-/-}$ mice in the late stages of infection when compared with their wild-type counterpart. Interestingly, studies using these mutant mice have also revealed the ability of galectin-3 to regulate the Th1/Th2 cytokine balance. In the model of allergic airway inflammation, *Lgals3*$^{-/-}$ mice showed decreased IL-4 and IgE levels and elevated amounts of IFN-γ and IgG2a antibodies in alveolar fluid and serum compared with wild-type mice [171]. Accordingly, *Lgals3*$^{-/-}$ mice exhibited a significantly higher Th1 response upon infection with *T. gondii* compared with their wild-type counterparts [172].

Furthermore, using a model of corneal wound healing *Lgals3*$^{-/-}$ mice evidenced a slower regeneration of injured corneal tissue, indicating that galectin-3 may positively influence the healing process [173].

Although the abovementioned studies support an essential role for galectin-1 and -3 in the regulation of neuronal, inflammatory, and vascular processes, a more careful examination of the role of galectins *in vivo* is still required, probably through the generation of single or double null mutant mice for other galectin family members. These results are critical to clearly demonstrate the redundant or critical roles of individual galectins in physiological and pathological processes.

7.3.2.1.2 Protective Role of Galectins during Inflammatory Processes

A variety of experimental mouse models have been extremely useful to address the role of galectins during inflammatory and autoimmune responses [156]. Early in the 1980s, Levi et al. [174] reported the preventive and therapeutic effects of electrolectin, a galectin-1 homolog purified from the fish *Electrophorus electricus*, in an experimental model of autoimmune myasthenia gravis in rabbits. Since then, the anti-inflammatory properties of galectin-1 have been evaluated in several rodent models of chronic inflammation and autoimmunity including experimental autoimmune encephalomyelitis [163,175], collagen-induced arthritis [176], concanavalin A-induced hepatitis [177], hapten-induced colitis [178], interphotoreceptor-binding protein-induced uveitis [179], autoimmune diabetes [180], and graft versus host disease [181]. In 1999, our group demonstrated that a single cell injection of syngeneic fibroblasts engineered to secrete galectin-1 on the day of the disease onset abrogated clinical and histopathological manifestations of collagen-induced arthritis in an experimental model of rheumatoid arthritis in DBA/ 1 mice [176]. This effect was also observed in response to daily injection of recombinant galectin-1. Similarly, Santucci et al. [177,178] found that galectin-1 treatment prevented tissue injury and T-cell-mediated inflammation in animal models of concanavalin A-induced hepatitis and inflammatory bowel disease [177,178]. Furthermore, the ability of galectin-1 to restore immune cell tolerance and homeostasis *in vivo* has also been investigated in experimental autoimmune uveitis, a T-cell-mediated mouse model of retinal disease [179]. Treatment with galectin-1 either early or late during the course of experimental autoimmune disease was sufficient to suppress clinical pathology and counteract pathogenic Th1 cells [179]. These results highlight the ability of this endogenous lectin to counteract Th1- and Th17-mediated responses through different, but potentially overlapping anti-inflammatory mechanisms (i.e., specific elimination of pathogenic Th1/Th17 responses, expansion of regulatory T cells, and modulation of the Th1/Th2 cytokine balance). In addition, recent evidence indicates that dendritic cells engineered to overexpress galectin-1 can delay the onset of autoimmune diabetes and insulitis when targeted to inflammatory sites [180]. Finally, Baum et al. investigated the protective effects of galectin-1 in a murine model of graft versus host disease [181]. Thus, galectin-1 can restore immune cell tolerance in several autoimmune, transplantation, and inflammation settings by acting as an anti-inflammatory and immunoregulatory cytokine. In addition, other members of the galectin family have also been evaluated *in vivo* including galectin-9, which has been shown to ameliorate neuroinflammation in a mouse model

of multiple sclerosis by targeting Tim-3-positive Th1 cells [182]. From a therapeutic standpoint, these findings suggest the potential use of galectin-1 for the selective treatment of Th1- and Th-17-mediated inflammatory disorders.

7.3.2.1.3 Dissecting the Role of Galectins in Tumor Progression

In addition to the role of galectins in inflammatory processes, galectins can also contribute to tumor progression through many different mechanisms including modulation of cell migration, adhesion, angiogenesis, and tumor-immune escape (extensively reviewed in Refs. [183,184]). Even though the targeted disruption of galectin genes in mice does not result in overt phenotype abnormalities regarding tumor progression, these mice do exhibit subtle but complex alterations during development and inflammation, which might also compromise tumor development and tumor-associated inflammatory responses.

The overall effects of galectins *in vivo* could be the combination of some or all of the effects described above [183,184]. Studies using mouse models have provided evidence for the role of galectins in tumor growth and metastasis [185,186]. Bresalier et al. demonstrated, using antisense strategies, a positive correlation between galectin-3 expression and the metastatic potential of colon cancer cells [185]. In addition, liver metastases of human adenocarcinoma xenotransplants in severe combined immunodeficient (SCID) mice were found to be inhibited by administration of an antigalectin-3 antibody [187]. Furthermore, in a metastatic model, breast carcinoma cells that overexpress transgenic galectin-3 are more resistant to apoptosis induced by reactive nitrogen and oxygen species than control breast carcinoma cells, and therefore have a higher metastatic potential [188]. In this regard, using knockdown tumor transfectants, we found a link between galectin-1-mediated immunosuppression and tumor-immune escape [189]. Blockade of the inhibitory effects of galectin-1 within tumor tissue resulted in reduced tumor mass and enhanced tumor rejection in a syngeneic model of murine B16 melanoma [189]. These effects were accompanied by the generation of a potent tumor-specific T cell-mediated response [189]. Thus, galectins have emerged as promising molecular targets for cancer therapy and galectin inhibitors have the potential to be used as antitumor and antimetastatic agents [184]. Challenges for the future will be to select the most appropriate tumor models to elucidate the therapeutic efficacy of different galectin inhibitors *in vivo*.

7.3.2.1.4 Other Lectins

Targeted gene disruption has been also of great value to establish the function of other glycan-binding proteins or lectins including the role of selectins in leukocyte migration and homing [190]. Disruption of P-selectin gene revealed the essential role of this lectin in mediating leukocyte rolling in the absence of inflammation. In addition, the use of knockout mice provided clear-cut evidence of the role of L-, E-, and P-selectins in leukocyte rolling during an inflammatory response. Similarly, single gene disruption clearly demonstrated that P-selectin is responsible for early neutrophil recruitment while the other selectins contribute to inflammation at later stages [190]. Gene-deficient mice also confirmed the important role of L-selectin in the homing of naïve B and T cells through high endothelial venules of lymph nodes. Moreover, removal of endothelial selectins uncovered the hidden importance of E-selectin in leukocyte homeostasis and showed that the endothelial selectins were as critical for leukocyte extravasation as the leukocyte β_2 integrins [190].

Surfactant proteins, SP-A and SP-D, are collagen-containing C-type lectins (collectins), which contribute to surfactant homeostasis and innate immunity [191]. These highly versatile carbohydrate-binding proteins are involved in a wide range of immune functions including viral neutralization and clearance of bacteria, fungi and apoptotic cells, downregulation of allergic reactions and resolution of inflammation [191]. Studies involving gene knockout mice and models of lung hypersensitivity and infection have revealed the diverse roles of SP-A and SP-D in the control of lung inflammation (reviewed in Refs. [192,193]). A careful examination gene-deficient mice challenged with a variety of inflammatory, neoplastic, and pathogenic insults will be critical for the elucidation of the pathophysiological relevance of lectin–carbohydrate lattices *in vivo*.

7.3.2.2 Zebrafish (*Danio rerio*) as an Alternative Vertebrate Model System

Recent advances in the cell biology and technical manipulations of zebrafish are making this model increasingly more attractive for studies on growth development, immune function, and oncology [194]. It offers many advantages over mammalian systems [195,196]. For example, its external fertilization, transparent embryos, and rapid development allow visualization the effects of genes involved in developmental processes, such as cell migration, organogenesis, etc. Moreover, the gene expression can be manipulated in zebrafish embryos and the effect of its expression can be analyzed easily. However, identification of orthologs of the mammalian counterparts may require caution. Because it is believed that a genome duplication event took place after the divergence of teleost fish from the mammalian lineage, it remains unclear how much of the hypothetical duplicated genome is present in zebrafish [194]. If both orthologs are present, structural and functional divergence should be considered because the duplicated genes would tolerate a higher rate of mutation in sequence and function [194].

7.3.2.2.1 Embryogenesis

In recent years, zebrafish has been demonstrated to be a useful model for addressing developmental questions in higher vertebrates, including mammals [195–197]. In addition to the transparent and rapid developing embryos, a large number of mutations that affect early embryonic development have been characterized and mapped, making this model a powerful resource for genetic studies on the function and mechanisms of action of developmentally regulated genes.

7.3.2.2.2 Immune Function during Embryogenesis

In addition to its usefulness as a model for genetics and development, recent studies support the notion that zebrafish can be an equally attractive model for addressing fundamental questions in immunobiology [194,198–202]. Because of its external fertilization of transparent embryos, morphological and cellular aspects, such as development of lymphoid tissues and inflammatory cellular responses, can be directly examined and potentially manipulated at far earlier points in development. Transgenic approaches can be employed more efficiently and may be more informative for developmental studies of the immune system.

7.3.2.2.2.1 Innate Immunity

Zebrafish has emerged as a valuable model for studying innate immunity. In zebrafish, there is no expression of adaptive immunity components (such as T cells, B cells, *rag* 1 and 2, or *Ig*) before 4 days postfertilization (dpf) [198], and therefore it can be assumed that until that time, the embryo relies heavily in innate immune mechanisms to fight infection. Among humoral components of innate immunity, the zebrafish exhibits cytokines and interferon [203,204], and pathogen-recognition molecules such as F-type lectins [88,205]. Cells that constitute hallmarks of innate immunity, such as granulocytes and macrophages, have been identified and functionally characterized in zebrafish. These two distinct granulocytes and macrophages start circulating in blood as early as 48 h postfertilization (hpf) [200]. Embryonic macrophages are actively phagocytic and can remove carbon particles from the circulation [200], phagocytose apoptotic corpses, and engulf and destroy large amounts of bacteria injected intravenously [199]. Both macrophages and granulocytes accumulate at inflammation sites in experimental wounds [200], and in experimental mycobacterial infections can form granulomas typical of mammalian models [201]. Molecular markers such as myeloperoxidase for granulocytes [200], and draculin and leukocyte-specific plastin for macrophages [199] have been cloned and used to study myelopoiesis and inflammatory responses in zebrafish early embryogenesis [199,200]. Furthermore it has been shown that although only a fraction of the macrophage population goes to the site of infection, the entire population of macrophages get activated, a behavior similar to macrophages in mammals [199]. With regard to cellular responses, the transparent embryos are ideal for investigating early stages of lymphopoiesis and inflammation

during early embryogenesis. Like mammals, zebrafish possess neutrophils and eosinophils [200]. Recently, several novel immune-type receptor (NITR) genes have been described in zebrafish [206]. Resolution of the complete NITR gene cluster has revealed the presence of C-type lectins within the same locus.

7.3.2.2.2.2 Adaptive Immunity

Similar to inflammatory responses, the adaptive immune system of the zebrafish resembles that of higher vertebrates. This includes thymus organization and ultrastructure, germ line, and somatic diversity of Igs, general aspects of MHC, and patterns of expression of *rag*1 and *rag*2 genes [194]. Moreover, although only TCRα has been described in zebrafish [207], all four TCR isotypes are presumed to be present in zebrafish, as evidenced in related taxa (elasmobranchs). Furthermore, zebrafish is known to express sid1, which encodes a secreted immunoglobulin domain and shares structural properties with VpreB, a surrogate light chain that functions in early stages of B-cell receptor expression [194]. In mammals, maturation of B cells in bone marrow is modulated by interactions of the surrogate chain in preB cells with galectin [208].

7.3.2.3 Experimental Infection and Disease

The well-developed adaptive and innate cellular immune systems make zebrafish an attractive model for the study of infectious diseases [202]. Zebrafish have been associated with disease developed from infections by Gram-positive and Gram-negative bacteria, fungus (*L. mutabilis*), and parasites such as nematodes, microsporidians, and dinoflagellates [194]. Experimental models for infectious disease by *Streptococcus* sp. and *Mycobacterium* sp. have been recently developed and characterized [201,202]. Following *Streptococcus* injection, zebrafish develop infections resembling not only those observed in farmed fish populations, but also human streptococcal diseases [202]. Like in human tuberculosis, infection of zebrafish with pathogenic mycobacteria produces granulomas, highly organized structures containing differentiated macrophages, and lymphocytes that sequester the pathogen. The transparency of zebrafish embryos, which in early stages have macrophages and granulocytes but lack lymphocytes, allows to directly visualize the events of mycobacterial infection and formation of granulomas *in vivo*, solely in the context of innate immunity [201]. Galectins have been proposed to mediate multiple roles in innate and adaptive immunity. Therefore, the zebrafish model should constitute an ideal model to dissect the functions of galectins before and after adaptive immunity becomes active in resistance to disease. As the study of galectins in experimental models of infection is developed, mutagenesis screens can be created to examine the genetics of galectin-mediated disease resistance and susceptibility.

7.3.3 Zebrafish as a Model for Studies on Cancer

The zebrafish has become an important model system not only for developmental studies, but also for cancer research [209–213]. The molecular conservation between human and zebrafish liver tumors was recently evaluated using comparative functional genomics [214]. Significant similarities between fish and human in the expression profiles of orthologous genes according to histopathologic tumor grade were identified, thus strengthening the rationale for using zebrafish as a cancer model system. Besides transparent embryos, fast growing and other advantages, zebrafish can easily be maintained in large numbers and all major organs can be conveniently examined in a few histological sections, thus a zebrafish carcinogenesis study is more cost-effective, with greater statistical power than a mammalian study [215]. In the zebrafish system, both reverse and forward genetics approaches that allow genetic manipulation of cancer genes are amenable. For example, transgenic technology to drive the expression of specific oncogenes under tissue specific promoters to model certain cancer types [211,216] and target-selected mutagenesis strategy to generate and identify tumor suppressor mutants [217] can be performed in zebrafish. Moreover, the zebrafish model is

unique among other vertebrates because of its capacity for large-scale forward genetics and the use of phenotype-driven mutational screens, thus making it an ideal system for gene function discovery. Particularly, zebrafish can be manipulated genetically to produce tumor, thus making them valuable in dissecting pathways and identifying novel genes involved in neoplastic transformation [211,213,216,218,219]. Furthermore, these genetically manipulated cancer-prone fish can be used for high-throughput chemical screening for anticancer drugs and therapies [209,212,213,219–221]. Since large number of embryos can be produced per week, high throughput chemical screening for anticancer drugs and therapies can easily be performed. As galectins are known to be involved in apoptosis and tumor progression, the zebrafish cancer models can be examined for galectin-mediated tumor progression pathways.

7.3.4 Zebrafish Cell Lines

Various somatic cell lines (ZEM2S, ZF4, ZFL, SJD.1, AB.9) from zebrafish have been established (American Type Culture Collection [ATCC], Manassas, VA) and can be used for studying the expression of gene of interest and protein export at a cellular level. The cell lines ZEM2S [222] and ZF4 [223] were established from zebrafish embryos (blastula and 1-day-old, respectively). The ZF4 cell line was derived from adult liver [224]. The SJD.1 and AB.9 are fibroblast cell lines derived from caudal fins of an adult zebrafish from strain SJD and AB, respectively [225]. In the presence of cells from a rainbow trout spleen cell line (RTS34st), zebrafish embryo cells were able to produce germ-line chimeras when introduced into a host embryo [226].

7.3.5 Galectin Repertoires in Zebrafish

By using various approaches (protein purification and characterization, cloning, and *in silico* data mining), we have identified and characterized the zebrafish galectin repertoire: six prototype galectin-1 like lectin 1, 2, 3, 4, 5 from *Danio rerio* (Drgal1-L1, Drgal1-L2, Drgal1-L3, Drgal1-L4, Drgal5, DrGRIFIN), two chimera-type galectin-3 like lectin 1, 2 from *D. rerio* (Drgal3-L1, Drgal3-L2), and at least five tandem-repeat type galectins (Drgal4, Drgal8, Drgal9L1, Drgal9-L2, Drgal9-L3 [103,227,227a]) (see Table 7.1). All zebrafish galectins characterized showed remarkable structural similarities with mammalian galectins and this enabled their unambiguous classification within the three well-established galectin groups. Galectin repertoires of protostome invertebrates and lower vertebrates appear to be limited to one or two galectin types [136]. For example, *C. elegans* and *Drosophila* have proto and tandem-repeat type galectins [136,228,229]. In this aspect, it is noteworthy that at the evolutionary level of teleosts, all three galectin groups (proto, chimera, and tandem-repeat) are already represented. Interestingly, a novel chimeric tandem-repeat galectin (Drgal Ch-TR) fused with a kinase suppressor gene has recently been identified in zebrafish and shown to be upregulated upon immune challenge (Ahmed and Vasta, submitted).

7.3.5.1 Carbohydrate-Binding Specificities and Structures of Zebrafish Galectins Are Similar to Mammalian Galectins

Homology modeling of Drgal1-L2 [SWISS-MODEL, Version 36.0003 [230] at the SWISS-MODEL Protein Modeling Server (http://swissmodel.expasy.org)] based on the bovine (*Bos taurus*) spleen galectin/*N*-acetyllactosamine complex structure [231], revealed that all nine conserved residues forming the carbohydrate-binding site in most mammalian galectins are present in the putative binding site of Drgal1-L2, and that all side chains of these residues were within 0.5 Å of the equivalent side chains of the bovine spleen galectin (Figure 7.1A). Drgal1-L2 also has a histidine at position 52, like mammalian and toad galectin-1, and chicken-16 galectin, whereas the congerins, skin mucus galectins from the conger eel (*C. myriaster*) [232,233] have a glycine in the same position. The crystal structure of congerin II suggests that the tyrosine at position 51 in congerin I and II takes the

TABLE 7.1
Current Status of the Characterization of the Zebrafish Galectin Repertoire

Subunit Structure	Designation[a]	ORF cDNA	Full mRNA	Recomb. Express.[b]	Carbo. Spec.[c]	Ab[d]	Crystal.[e]	Ontogenic/Tissue Expression	gal-MO Probe[f]	CRD Mutant[g]	In Situ Hybridization/ Ab Stain[h]	Tg Cons.[i]
Proto	Drgal1-L1	√	√	√	Ip	—	—	√/√	√	—	–/–	—
	Drgal1-L2	√	√	√	√	√	√	√/√	√	√	√/√	√[j]
	Drgal1-L2d	√	√	√	—	—	—	√/–		—	–/–	—
	Drgal1-L3	√	—	√	Ip	—	—	√/√	√	—	√/√/–	—
	Drgal1-L4	√	gb	ip	—	—	—	√/√	√	—	√/–	ip
	DrGRIFIN	√	gb	√	—	—	—	√/√	√	—	√/–	—
	Drgal5	gb	—	ip	—	—	—	–/–	—	—	–/–	—
Chimera	Drgal3-L1	√	√	√	√	√	—	√/√	√	—	–/√	—
	Drgal3-L2	√	√	—	—	—	—	–/–	—	—	–/–	—
Tandem-repeat	Drgal4	gb	—	—	—	—	—	–/–	—	—	–/–	—
	Drgal8	gb	—	—	—	—	—	–/–	—	—	–/–	—
	Drgal9-L1	√	gb	ip	—	—	—	√/√	√	—	–/–	—
	Drgal9-L2	√	√	√	√	√	—	–/–	—	√	–/–	—
	Drgal9-L3	√	—	ip	—	—	—	–/–	—	—	–/–	—
	Drgal9-L4	√	—	—	—	—	—	–/–	—	—	–/–	—
Chimeric tandem-repeat	Drgal-Ch-TR	ip	—	—	—	—	—	–/–	—	—	–/–	—

[a] Designation based on Ref. [136]; Drgal1-L1 represents galectin-1-like lectin 1 from *D. rerio*.
[b] Recomb. expr., Recombinant protein obtained in *E. coli* expression system.
[c] Carbo. spec., Carbohydrate-binding specificity.
[d] Ab., Anti-Drgal-antibodies (polyclonal).
[e] Crystal., Crystallization for structural analysis.
[f] gal-MO., Morpholino-modified antisense oligonucleotides.
[g] CRD mutant., Clone with mutated CRD.
[h] In situ hyb./Ab stain., Whole-mount in situ hybridization using RNA probe; whole mount immunostaining using Drgal or Flag-specific antibodies.
[i] Tg cons., Transgenic construct.
[j] Tg constructs for wild-type, CRD single mutant, and CRD double mutant Drgal1-L2; ip, in progress, gb, GenBank information.

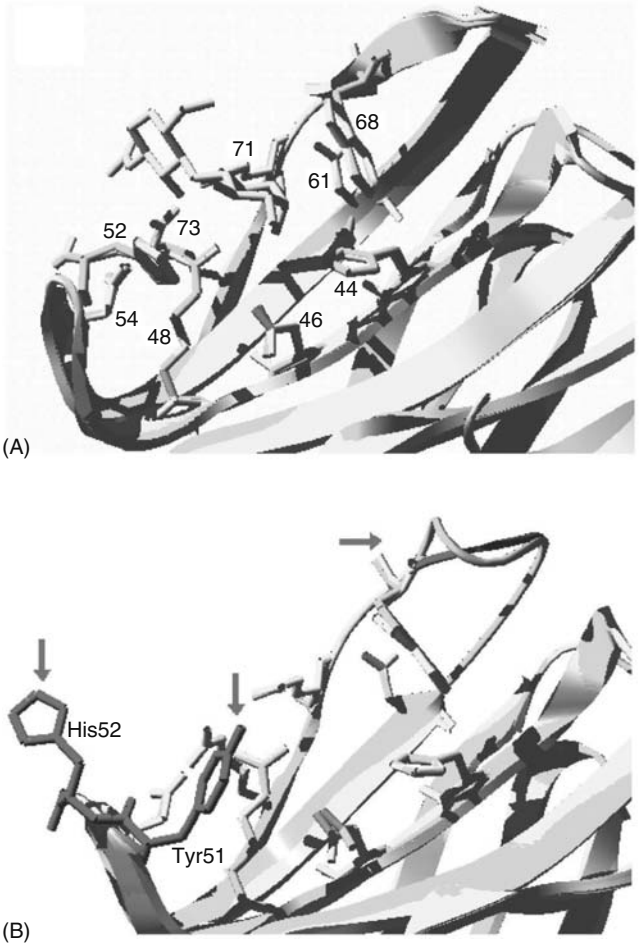


FIGURE 7.1 (See color insert following blank page 170. Also see CD for color figure.) Homology modeling of Drgal1-L2. (A) Drgal-L2 (yellow) was modeled to bovine spleen galectin-1 (blue), showing the side chains of nine conserved binding site residues of mammalian galectin-1. (B) Drgal-L2 (yellow) was modeled to congerin II (blue). The same nine side chains are shown as previously. Major differences in the binding site architecture are the His52 of Drgal1-L2 versus Tyr51 of congerin II, and the extended loop, highlighted by the red arrow, where the model is not supported due to additional residue (Ser62) in congerin II. (Reproduced from Vasta, G.R., Ahmed, H., Du, S., and Henrikson, D., *Glycoconj. J.*, 21, 503, 2005. With permission from Springer.)

place of His52 in protein–carbohydrate interactions. The Drgal1-L2 model also reveals that in the loop between F4 and F5 from bovine spleen galectin, a tyrosine has been replaced by a serine in Drgal1-L2. Primary structure analysis reveals the presence of two additional residues at positions 66 and 67 in congerin I, and an additional residue at position 67 in congerin II, as compared to the zebrafish and bovine spleen prototype galectins. When Drgal1-L2 was modeled on the structures of congerins I and II [232,233], these differences resulted in a larger loop between S5 and S6, thus generating a backbone discrepancy between the congerins structure and the Drgal1-L2 model (Figure 7.1B). The only other backbone difference is found where congerin I has the conformationally extended N- and C-termini, where it is believed that congerin I undergoes intersubunit strand swapping when dimerized. With respect to energy minimization, Drgal1-L2 appears to favor the bovine spleen galectin and congerin II structure at the N- and C-termini. Congerin II may have an extended binding site beyond the nonreducing end of bound galactosides. Of the six residues suggested

interacting with the extended sugar, Drgal1-L2 possesses only three identical and two similar with congerin II. Drgal1-L2 has even fewer of these six residues in common with bovine spleen galectin. This analysis suggests that binding sites of the prototype zebrafish galectins are structurally similar to the mammalian homologs. However, they appear as less similar to the galectins from the conger eel, which are clearly nonorthologous gene products.

7.3.6 Current Approaches to Address Biological Roles of Galectins in Zebrafish

Several unique approaches have been established in the zebrafish model to address biological roles of genes of interest. These approaches are currently implemented in our laboratory to examine the biological roles of galectins in zebrafish. In the first approach, the expression of the selected gene is blocked by injecting a morpholino-modified antisense oligo against the corresponding mRNA that blocks protein translation [234,235] or a splicing blocker that interferes with intron/exon splicing [236]. In the second approach, ectopic expression is used to manipulate the activity of a gene product in zebrafish embryos by either mRNA injection or DNA injection [237]. The expression and activity of the gene product can be inhibited or increased by specifically expressing a dominant negative or a full-length active form of that protein, respectively. The expression can be targeted to the specific tissue/organ of interest by using tissue-specific promoters [237]. In the third approach, mutants that affect the expression of the gene of interest are analyzed to determine possible interactions between the mutant and the gene of interest and the potential functions of gene of interest. The last category includes a series of approaches such as bacterial challenges through incubation or injection into zebrafish embryos that pertain to elucidation of the potential roles of the proteins of interest in innate immunity during embryogenesis [238].

7.3.6.1 Antisense Morpholino Knockdowns

To examine the possible functions of galectins in zebrafish embryos, blocking of galectins expression is performed by injection of morpholino-modified antisense oligo targeted to the galectin 5′UTR sequence near the ATG start site. To determine if the injection of antisense oligos effectively block galectin expression, whole mount antibody staining is performed on injected and uninjected embryos and compared the results. Histological analysis is performed with injected embryos to determine any developmental defect. Coinjections are also performed with the galectin mRNAs to rescue the developmental defect. Coinjection should rescue the development defect if the phenotype is galectin-specific [227].

7.3.6.2 Expression of Ectopic Galectin mRNA

7.3.6.2.1 Overexpression of the Gene of Interest

The phenotypic consequences of overexpression of the gene of interest can provide substantial information about its biological role [237,239]. For this purpose, the cDNA is cloned into a suitable vector and capped mRNAs are transcribed *in vitro* from the linearized plasmid DNA. The capped mRNAs are injected into 1–4 cell stage zebrafish embryos through microinjection. The injected mRNAs are distributed globally in the embryos. Consequently, proteins are made in ectopic sites that normally do not express these proteins. For example, injection of mRNAs encoding Drgal1-L2 into zebrafish embryos results in ectopic expression in many cell types in addition to notochord cells that express the endogenous Drgal1-L2. Thus, this approach provides a useful tool to analyze the biological role of Drgal1-L2.

7.3.6.2.2 Repression of Galectin Function by a Galectin Dominant Negative Construct

Another powerful approach to analyze gene function in zebrafish embryos is the use of “dominant negative” approach, which is used to specifically block the function of a gene product in a specific

location. Protogalectin is a dimer of two identical noncovalently linked subunits, each of which houses carbohydrate-binding site on the opposite face [231]. Mutating the carbohydrate-binding site but keeping the dimer interface unaltered will create a mutant protein that can dimerize with the endogenous galectin but is unable to bind to carbohydrate on the mutant subunit side. Consequently, the wild-type mutant heterodimer will fail to establish an interaction between cells or between cell and ECM. Thus, the mutant protein will act as a dominant negative that can interfere with activity of the endogenous galectin.

7.3.6.2.3 Identification of Notochord- and Muscle-Specific Promoters

Although ectopic expression through mRNA injection has been successfully used to study the function of many genes, the drawback of this approach is that the expression is not tissue-specific. As indicated above, Drgal1-L2 is expressed in the notochord of zebrafish embryos and the ideal approach to analyze its function is to disrupt its activity specifically in the notochord. For this purpose, a promoter that targets protein expression to the notochord is identified and characterized from the *tiggywinkle hedgehog* (*twhh)* gene [240]. As notochord influences the specification and differentiation of skeletal muscles, it is of interest to specifically express wild-type or mutant galectins in muscle cells. To this goal, muscle-specific promoters derived from zebrafish myoD and myogenin genes have also been identified and their activity analyzed in zebrafish embryos [241]. These promoters will be useful for expressing galectin genes in a tissue-restricted manner.

7.3.6.3 Analyses of Notochord Mutants for Expression of Drgal1-L2

The notochord-specific expression of Drgal1-L2 suggests that it may play a role in notochord formation. The *in vivo* requirement of laminin β1 and laminin γ1, which are galectin ligands for the differentiation of chordamesoderm to notochord [242], suggest that the mutants *grumpy* and *sleepy*, two zebrafish loci known to control notochord formation and encode laminin β1 and laminin γ1, respectively, should be examined in detail with regard to galectin function. Several zebrafish mutants with defects in notochord development have been generated [243,244]. These mutants have been classified into early or late defects. The early defect mutants include *ntl*, *flh*, *doc*, and *mom,* and are involved in initial specification and differentiation of notochord cells. Later defect mutants, such as *sly*, *gup*, *qam*, *sno*, *drb*, *git*, *bal*, *blo*, *pun*, and *kon* are primarily involved in maturation of the notochord. Although initial notochord formation appears normal in these mutants, the notochord cells failed to vacuolate. Except for *sly* and *qam*, which have been cloned and encode laminin β1 and laminin γ1, respectively, most of the mutants have not been mapped. It is interesting to note that some of the mutants, such as *kon*, showed both notochord and somite defect, and reduced mobility [243].

7.3.6.4 Roles of Galectins in Innate Immunity

Several approaches are currently taken to study the functions of galectins in zebrafish upon immune challenge. These include (1) challenge wild-type embryos, hatchlings (before and after 4 days) and adults with lipopolysaccharide (LPS), heat-killed and live bacteria (with phosphate-buffered saline injections and noninjected specimens as controls) in a dose–response format; (2) challenge knockdown embryos (for each galectin) and compare with the wild type; and (3) challenge adult transgenic zebrafish expressing dominant negative for each galectin and compare with the wild type. For each group, experiments are performed to assess (1) qualitative and quantitative changes on the patterns of temporal expression of galectins by standard and real-time reverse transcriptase polymerase chain reaction (RT-PCR); (2) qualitative and quantitative changes on the patterns of spatial expression (organ/tissue distribution) of galectins by in situ hybridization and immunostaining; (3) qualitative and quantitative aspects of the local acute inflammatory response at the injection site by histological analysis (magnitude of the cellular response), composition of the cellular

response (cell types) by using cell-specific markers for in situ hybridization and immunostaining; (4) qualitative and quantitative aspects of the systemic humoral acute inflammatory response, focusing on interferon expression by RT-PCR; and (5) developmental course of the embryos, hatching rates, hatchling and adult survival rates.

7.4 CONCLUSIONS AND FUTURE DIRECTIONS

A variety of *in vivo* and *in vitro* experiments and, more recently, *in silico* approaches have enabled greater insight into the diversity and complexity of lectin repertoires in invertebrates, protochordates, and vertebrates. Murine models are well-established by virtue of their remarkable similarities with humans, short lifespan, rapid reproductive cycle, and their relatively simple maintenance and handling. The use of knockouts has enabled the rigorous examination of biological roles of galectins and other lectins, but these models also have limitations. Relatively simple invertebrate organisms may serve as useful models for some of the functions of lectins in mammals. Ongoing genome, transcriptome, and proteome projects on model organisms representative of nonmammalian taxa will reveal the extent of their full lectin repertoires. The identification of members of lectin families in these taxa has revealed novel structural features, most probably reflecting functional adaptations along the lineages leading to higher vertebrate taxa. Furthermore, the identification of novel lectin families, such as the F-type lectins, underscores the fact that more research on nonmammalian model organisms will provide new information on structural, functional, and evolutionary aspects of lectin repertoires. For example, the early intracellular sorting events involving calnexin and L-type lectins as well as the role of R-type CRDs in glycosyltransferases are likely to be quite similar, whereas later sorting events involving the mannose 6-phosphate receptors will probably be different. At the cell surface, the role of some of the galectins may be similar in all animals, so that genetic and developmental analysis of the model invertebrates is likely to illuminate studies of the vertebrate proteins as well. In contrast, the greater diversity of invertebrate and vertebrate proteins containing CTLDs suggests that these proteins probably participate in more specialized functions of glycans that are unique to different groups of animals [4]. Studies that have highlighted the functional versatility of the CTLD superfamily suggest that many further interesting discoveries have yet to be made [112].

Because of its external fertilization, rapidly developing transparent embryos, variety of established techniques for manipulation of gene expression and a growing collection of mutations affecting early embryonic development that have been characterized and mapped, zebrafish (*D. rerio*) exhibits substantial experimental advantages over murine models for developmental studies. Zebrafish is endowed of a galectin repertoire that includes members of the three galectin subtypes, proto, chimera, and tandem-repeat, although less diversified relative to that of mammals. Further, structural analysis of selected zebrafish galectins suggests that their binding properties are very similar to the mammalian equivalents. Members of the zebrafish galectin repertoire exhibit unique temporal and spatial gene expression during early development. In this context, we propose the use of this species as a model organism for the elucidation and characterization of the biological roles of galectins in vertebrates, including mammals.

ACKNOWLEDGMENTS

Supported by grants R01 GM070589-01 from the National Institutes of Health, IOB 0618409 and MCB-0077928 from the National Science Foundation, and NA05NMF4571243 from the National Oceanic and Atmospheric Administration to G.R.V.; grant F32GM083352 from the National Institute of General Medical Sciences to S.S.J.; grant BC063627 from the Department of Defense to H.A.; and The Cancer Research Institute "Elaine R. Shepard Memorial Investigator Award," National Agency for Promotion of Science and Technology (PICT 2003-05-13787), University of Buenos Aires (M091), and a Program of Fundación Sales/CONICET to G.A.R.

REFERENCES

1. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al.: The genome sequence of *Drosophila melanogaster. Science* 2000, 287:2185–2195.
2. Consortium TCes: Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* 1998, 282:2012–2018.
3. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, et al.: The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 2002, 298:129–149.
4. Dodd RB and Drickamer K: Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity. *Glycobiology* 2001, 11:R71–R79.
5. Nicholas HR and Hodgkin J: Responses to infection and possible recognition strategies in the innate immune system of *Caenorhabditis elegans. Mol Immunol* 2004, 41:479–493.
6. Camus ML: Recherches experimentales sur une agglutinine produite par la glande de l'albumen chez l'. *Helix pomatia. CR Acad Sci* 1899, 129:233.
7. Prokop O, Uhlenbruck G, and Kohler W: A "new" human blood group receptor A_{hel} tested with saline extracts from *Helix hortensis* (garden snail). *J Forensic Med* 1965, 12:108.
8. Uhlenbruck G and Prokop O: An agglutinin from *Helix pomatia*, which reacts with terminal *N*-acetyl-D-galactosamine. *Vox Sang* 1966, 11:519–520.
9. Prokop O, Uhlenbruck G, and Kohler W: A new source of antibody-like substances having anti-blood group specificity. A discussion on the specificity of Helix agglutinins. *Vox Sang* 1968, 14:321–333.
10. Sanchez JF, Lescar J, Chazalet V, Audfray A, Gagnon J, Alvarez R, Breton C, Imberty A, and Mitchell EP: Biochemical and structural analysis of *Helix pomatia* agglutinin. A hexameric lectin with a novel fold. *J Biol Chem* 2006, 281:20171–20180.
11. Noguchi H: The interaction of the blood of cold-blooded animals, with reference to haemolysis, agglutination and precipitation. *Univ Pa Med Bull* 1902, 15.
12. Noguchi H: On the multiplicity of the serum haemagglutinins of cold-blooded animals. *Zentralb. Bakteriol Abt l Orig* 1903, 34:286.
13. Cantacuzene J: Sur certains anticorps naturels observes chez *Eupagurus prideauxii. CR Soc Biol* 1912, 73:663.
14. Canacuzene J: Anticorps normaux et experimentaux chez quelques invertebres marins. *CR Soc Biol* 1919, 82:1087.
15. Vasta GR and Marchalonis JJ: Humoral recognition factors in the Arthropoda. The specificity of Chelicerata serum lectins. *Am Zool* 1983, 23:157–171.
16. Iwanaga S, Kawabata S, and Muta T: New types of clotting factors and defense molecules found in horseshoe crab hemolymph: Their structures and functions. *J Biochem (Tokyo)* 1998, 123:1–15.
17. Iwanaga S: The molecular basis of innate immunity in the horseshoe crab. *Curr Opin Immunol* 2002, 14:87–95.
18. Kawabata S and Iwanaga S: Role of lectins in the innate immunity of horseshoe crab. *Dev Comp Immunol* 1999, 23:391–400.
19. Nagai T, Kawabata S, Shishikura F, and Sugita H: Purification, characterization, and amino acid sequence of an embryonic lectin in perivitelline fluid of the horseshoe crab. *J Biol Chem* 1999, 274:37673–37678.
20. Iwaki D, Osaki T, Mizunoe Y, Wai SN, Iwanaga S, and Kawabata S: Functional and structural diversities of C-reactive proteins present in horseshoe crab hemolymph plasma. *Eur J Biochem* 1999, 264:314–326.
21. Quigley JP and Armstrong PB: An endopeptidase inhibitor, similar to mammalian alpha 2-macroglobulin, detected in the hemolymph of an invertebrate, *Limulus polyphemus. J Biol Chem* 1983, 258: 7903–7906.
22. Iwaki D, Kawabata S, Miura Y, Kato A, Armstrong PB, Quigley JP, Nielsen KL, Dolmer K, Sottrup-Jensen L, and Iwanaga S: Molecular cloning of Limulus alpha 2-macroglobulin. *Eur J Biochem* 1996, 242:822–831.
23. Robey FA and Liu TY: Limulin: A C-reactive protein from *Limulus polyphemus*. *J Biol Chem* 1981, 256:969–975.
24. Kanost MR, Jiang H, and Yu XQ: Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunol Rev* 2004, 198:97–105.

25. Zhu Y, Johnson TJ, Myers AA, and Kanost MR: Identification by subtractive suppression hybridization of bacteria-induced genes expressed in *Manduca sexta* fat body. *Insect Biochem Mol Biol* 2003, 33:541–559.
26. Ladendorff NE and Kanost MR: Bacteria-induced protein P4 (hemolin) from *Manduca sexta*: A member of the immunoglobulin superfamily which can inhibit hemocyte aggregation. *Arch Insect Biochem Physiol* 1991, 18:285–300.
27. Yu XQ, Gan H, and Kanost MR: Immulectin, an inducible C-type lectin from an insect, *Manduca sexta*, stimulates activation of plasma prophenol oxidase. *Insect Biochem Mol Biol* 1999, 29:585–597.
28. Yu XQ and Kanost MR: Immulectin-2, a lipopolysaccharide-specific lectin from an insect, *Manduca sexta*, is induced in response to gram-negative bacteria. *J Biol Chem* 2000, 275:37373–37381.
29. Yu XQ and Kanost MR: Binding of hemolin to bacterial lipopolysaccharide and lipoteichoic acid. An immunoglobulin superfamily member from insects as a pattern-recognition receptor. *Eur J Biochem* 2002, 269:1827–1834.
30. Yu XQ, Zhu YF, Ma C, Fabrick JA, and Kanost MR: Pattern recognition proteins in *Manduca sexta* plasma. *Insect Biochem Mol Biol* 2002, 32:1287–1293.
31. Ma C and Kanost MR: A beta1,3-glucan recognition protein from an insect, *Manduca sexta*, agglutinates microorganisms and activates the phenoloxidase cascade. *J Biol Chem* 2000, 275:7505–7514.
32. Wang Y, Willott E, and Kanost MR: Organization and expression of the hemolin gene, a member of the immunoglobulin superfamily in an insect, *Manduca sexta*. *Insect Mol Biol* 1995, 4:113–123.
33. Franc NC and White K: Innate recognition systems in insect immunity and development: New approaches in Drosophila. *Microbes Infect* 2000, 2:243–250.
34. Takahashi H, Komano H, Kawaguchi N, Kitamura N, Nakanishi S, and Natori S: Cloning and sequencing of cDNA of *Sarcophaga peregrina* humoral lectin induced on injury of the body wall. *J Biol Chem* 1985, 260:12228–12233.
35. Komano H, Mizuno D, and Natori S: Purification of lectin induced in the hemolymph of *Sarcophaga peregrina* larvae on injury. *J Biol Chem* 1980, 255:2919–2924.
36. Kawaguchi N, Komano H, and Natori S: Involvement of Sarcophaga lectin in the development of imaginal discs of *Sarcophaga peregrina* in an autocrine manner. *Dev Biol* 1991, 144:86–93.
37. Fujita Y, Kurata S, Homma K, and Natori S: A novel lectin from Sarcophaga. Its purification, characterization, and cDNA cloning. *J Biol Chem* 1998, 273:9667–9672.
38. Jomori T, Kubo T, and Natori S: Purification and characterization of lipopolysaccharide-binding protein from hemolymph of American cockroach *Periplaneta americana*. *Eur J Biochem* 1990, 190:201–206.
39. Kawasaki K, Kubo T, and Natori S: A novel role of Periplaneta lectin as an opsonin to recognize 2-keto-3-deoxy octonate residues of bacterial lipopolysaccharides. *Comp Biochem Physiol B* 1993, 106:675–680.
40. Jomori T and Natori S: Function of the lipopolysaccharide-binding protein of *Periplaneta americana* as an opsonin. *FEBS Lett* 1992, 296:283–286.
41. Kubo T, Kawasaki K, Nonomura Y, and Natori S: Localization of regenectin in regenerates of American cockroach (*Periplaneta americana*) legs. *Int J Dev Biol* 1991, 35:83–90.
42. Kubo T, Kawasaki K, and Natori S: Transient appearance and localization of a 26-kDa lectin, a novel member of the Periplaneta lectin family, in regenerating cockroach leg. *Dev Biol* 1993, 156:381–390.
43. Arai T, Kawasaki K, Kubo T, and Natori S: Cloning of cDNA for regenectin, a humoral C-type lectin of *Periplaneta americana*, and expression of the regenectin gene during leg regeneration. *Insect Biochem Mol Biol* 1998, 28:987–994.
44. Chen C, Ratcliffe NA, and Rowley AF: Detection, isolation and characterization of multiple lectins from the haemolymph of the cockroach *Blaberus discoidalis*. *Biochem J* 1993, 294 (Pt 1):181–190.
45. Chen C, Rowley AF, Newton RP, and Ratcliffe NA: Identification, purification and properties of a beta-1,3-glucan-specific lectin from the serum of the cockroach, *Blaberus discoidalis* which is implicated in immune defence reactions. *Comp Biochem Physiol B Biochem Mol Biol* 1999, 122:309–319.
46. Chen C, Durrant HJ, Newton RP, and Ratcliffe NA: A study of novel lectins and their involvement in the activation of the prophenoloxidase system in *Blaberus discoidalis*. *Biochem J* 1995, 310 (Pt 1):23–31.
47. Wilson R and Ratcliffe NA: Effect of lysozyme on the lectin-mediated phagocytosis of *Bacillus cereus* by haemocytes of the cockroach, *Blaberus discoidalis*. *J Insect Physiol* 2000, 46:663–670.
48. Wilson R, Chen C, and Ratcliffe NA: Innate immunity in insects: The role of multiple, endogenous serum lectins in the recognition of foreign invaders in the cockroach, *Blaberus discoidalis*. *J Immunol* 1999, 162:1590–1596.
49. Chen C and Billingsley PF: Detection and characterization of a mannan-binding lectin from the mosquito, *Anopheles stephensi* (Liston). *Eur J Biochem* 1999, 263:360–366.

50. Dimopoulos G, Casavant TL, Chang S, Scheetz T, Roberts C, Donohue M, Schultz J, Benes V, Bork P, Ansorge W, et al.: Anopheles gambiae pilot gene discovery project: Identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines. *Proc Natl Acad Sci USA* 2000, 97:6619–6624.
51. Nelson KA, Leonard LA, Posey MH, Alphin TD, and Mallin MA: Using transplanted oyster (*Crassostrea virginica*) beds to improve water quality in small tidal creeks: A pilot study. *J Exp Mar Biol Ecol* 2004, 298:347–368.
52. Robinson WA, Maher WA, Krikowa F, Nell JA, and Hand R: The use of the oyster *Saccostrea glomerata* as a biomonitor of trace metal contamination: intra-sample, local scale and temporal variability and its implications for biomonitoring. *J Environ Monit* 2005, 7:208–223.
53. Quilang J, Wang S, Li P, Abernathy J, Peatman E, Wang Y, Wang L, Shi Y, Wallace R, Guo X, and Liu Z: Generation and analysis of ESTs from the eastern oyster, *Crassostrea virginica* Gmelin and identification of microsatellite and SNP markers. *BMC Genomics* 2007, 8:157.
54. Tasumi S and Vasta GR: A galectin of unique domain organization from hemocytes of the eastern oyster (*Crassostrea virginica*) is a receptor for the protistan parasite *Perkinsus marinus. J Immunol* 2007, 179:3086–3098.
55. Sea Urchin Genome Sequencing Consortium. The genome of the sea urchin *Strongylocentrotus purpuratus. Science* 2006, 314:941–952.
56. Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH, Terwilliger DP, Buckley KM, Brockton V, Nair SV, Berney K, Fugmann SD, Anderson MK, Pancer Z, Cameron RA, Smith LC, and Rast JP: The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol* 2006, 300:349–365.
57. Vasta, GR, Quesenberry MS, and Ahmed H: A tunicate fucose-binding lectin is a homologue of the mammalian binding proteins, In *New Directions in Invertebrate Immunology*, eds. Söderhäll, K., Iwanaga, S., Vasta, G.R. SOS Publications, Fair Haven, NJ, 1996, pp. 189–227.
58. Ji X, Azumi K, Sasaki M, and Nonaka M: Ancient origin of the complement lectin pathway revealed by molecular cloning of mannan binding protein-associated serine protease from a urochordate, the Japanese ascidian, *Halocynthia roretzi. Proc Natl Acad Sci USA*, 1997, 94:6340–6345.
59. Thiel, S., Vorup-Jensen T, Stover CM, Schwaeble W, Laursen SB, Poulsen K, Willis AC, Eggleton P, Hansen S, Holmskov U, Reid KBM, and Jensenius JC: A second serine protease associated with mannan-binding lectin that activates complement. *Nature* 1997, 386:506–510.
60. Matsushita M and Fujita T: Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. *J Exp Med* 1992, 176:1497–1502.
61. Schachter H: Protein glycosylation lessons from *Caenorhabditis elegans. Curr Opin Struct Biol* 2004, 14:607–616.
62. Drickamer K and Dodd RB: C-Type lectin-like domains in *Caenorhabditis elegans*: Predictions from the complete genome sequence. *Glycobiology* 1999, 9:1357–1369.
63. Kim DH, Feinbaum R, Alloing G, Emerson FE, Garsin DA, Inoue H, Tanaka-Hino M, Hisamoto N, Matsumoto K, Tan MW, et al.: A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* 2002, 297:623–626.
64. Vasta GR, Ahmed H, and Odom EW: Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. *Curr Opin Struct Biol* 2004, 14:617–630.
65. Trombetta ES and Helenius A: Lectins as chaperones in glycoprotein folding. *Curr Opin Struct Biol* 1998, 8:587–592.
66. Fiedler K and Simons K: A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. *Cell* 1994, 77:625–626.
67. Drickamer K: Increasing diversity of animal lectin structures. *Curr Opin Struct Biol* 1995, 5:612–616.
68. Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N, Chinwalla A, Clarke L, Clee C, Coghlan A, et al.: The genome sequence of *Caenorhabditis briggsae*: A platform for comparative genomics. *PLoS Biol* 2003, 1:E45.
69. Fujita T: Evolution of the lectin-complement pathway and its role in innate immunity. *Nat Rev Immunol* 2002, 2:346–353.
70. Natori S, Shiraishi H, Hori S, and Kobayashi A: The roles of Sarcophaga defense molecules in immunity and metamorphosis. *Dev Comp Immunol* 1999, 23:317–328.
71. Pace KE, Lebestky T, Hummel T, Arnoux P, Kwan K, and Baum LG: Characterization of a novel *Drosophila melanogaster* galectin. Expression in developing immune, neural, and muscle tissues. *J Biol Chem* 2002, 277:13091–13098.
72. Theopold U, Rissler M, Fabbri M, Schmidt O, and Natori S: Insect glycobiology: A lectin multigene family in *Drosophila melanogaster. Biochem Biophys Res Commun* 1999, 261:923–927.

73. Tanji T, Ohashi-Kobayashi A, and Natori S: Participation of a galactose-specific C-type lectin in Drosophila immunity. *Biochem J* 2006, 396:127–138.
74. Satou Y, Yamada L, Mochizuki Y, Takatori N, Kawashima T, Sasaki A, Hamaguchi M, Awazu S, Yagi K, Sasakura Y, et al.: A cDNA resource from the basal chordate *Ciona intestinalis*. *Genesis* 2002, 33:153–154.
75. Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, et al.: The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. *Science* 2002, 298:2157–2167.
76. Khalturin K, Panzer Z, Cooper MD, and Bosch TC: Recognition strategies in the innate immune system of ancestral chordates. *Mol Immunol* 2004, 41:1077–1087.
77. Sato A, Mayer WE, Overath P, and Klein J: Genes encoding putative natural killer cell C-type lectin receptors in teleostean fishes. *Proc Natl Acad Sci USA* 2003, 100:7779–7784.
78. Yoder JA, Mueller MG, Nichols KM, Ristow SS, Thorgaard GH, Ota T, and Litman GW: Cloning novel immune-type inhibitory receptors from the rainbow trout, *Oncorhynchus mykiss*. *Immunogenetics* 2002, 54:662–670.
79. Lund V and Olafsen JA: Changes in serum concentration of a serum amyloid P-like pentraxin in Atlantic salmon, *Salmo salar* L., during infection and inflammation. *Dev Comp Immunol* 1999, 23:61–70.
80. Cartwright JR, Tharia HA, Burns I, Shrive AK, Hoole D, and Greenhough TJ: Isolation and characterisation of pentraxin-like serum proteins from the common carp *Cyprinus carpio*. *Dev Comp Immunol* 2004, 28:113–125.
81. Cook MT, Hayball PJ, Nowak BF, and Hayball JD: The opsonising activity of a pentraxin-like protein isolated from snapper (*Pagrus auratus*, Sparidae) serum. *Dev Comp Immunol* 2005, 29:703–712.
82. Fuller JR, Pitzer JE, Godwin U, Albertino M, Machon BD, Kearse KP, and McConnell TJ: Characterization of the molecular chaperone calnexin in the channel catfish, *Ictalurus punctatus*, and its association with MHC class II molecules. *Dev Comp Immunol* 2004, 28:603–617.
83. Kales S, Fujiki K, and Dixon B: Molecular cloning and characterization of calreticulin from rainbow trout (*Oncorhynchus mykiss*). *Immunogenetics* 2004, 55:717–723.
84. Peavy TR, Hernandez C, and Carroll EJ, Jr.: Jeltraxin, a frog egg jelly glycoprotein, has calcium-dependent lectin properties and is related to human serum pentraxins CRP and SAP. *Biochemistry* 2003, 42:12761–12769.
85. Shoji H, Nishi N, Hirashima M, and Nakamura T: Characterization of the *Xenopus galectin* family. Three structurally different types as in mammals and regulated expression during embryogenesis. *J Biol Chem* 2003, 278:12285–12293.
86. Honda S, Kashiwagi M, Miyamoto K, Takei Y, and Hirose S: Multiplicity, structures, and endocrine and exocrine natures of eel fucose-binding lectins. *J Biol Chem* 2000, 275:33151–33157.
87. Sinha S, Mandal C, and Allen AK: Acute phase response of C-reactive protein of *Labeo rohita* to aquatic pollutants is accompanied by the appearance of distinct molecular forms. *Arch Biochem Biophys* 2001, 396:139–150.
88. Bianchet MA, Odom EW, Vasta GR, and Amzel LM: A novel fucose recognition fold involved in innate immunity. *Nat Struct Biol* 2002, 9:628–634.
89. Saito T, Hatada M, Iwanaga S, and Kawabata S: A newly identified horseshoe crab lectin with binding specificity to O-antigen of bacterial lipopolysaccharides. *J Biol Chem* 1997, 272:30703–30708.
90. Zhu X, Mahairas G, Illies M, Cameron RA, Davidson EH, and Ettensohn CA: A large-scale analysis of mRNAs expressed by primary mesenchyme cells of the sea urchin embryo. *Development* 2001, 128:2615–2627.
91. Multerer KA and Smith LC: Two cDNAs from the purple sea urchin, *Strongylocentrotus purpuratus*, encoding mosaic proteins with domains found in factor H, factor I, and complement components C6 and C7. *Immunogenetics* 2004, 56:89–106.
92. Tateno H, Ogawa T, Muramoto K, Kamiya H, and Saneyoshi M: Rhamnose-binding lectins from steelhead trout (*Oncorhynchus mykiss*) eggs recognize bacterial lipopolysaccharides and lipoteichoic acid. *Biosci Biotechnol Biochem* 2002, 66:604–612.
93. Tsutsui S, Tasumi S, Suetake H, Kikuchi K, and Suzuki Y: Carbohydrate-binding site of a novel mannose-specific lectin from fugu (*Takifugu rubripes*) skin mucus. *Comp Biochem Physiol B Biochem Mol Biol* 2006, 143:514–519.
94. Tsutsui S, Tasumi S, Suetake H, and Suzuki Y: Lectins homologous to those of monocotyledonous plants in the skin mucus and intestine of pufferfish, *Fugu rubripes*. *J Biol Chem* 2003, 278:20882–20889.
95. Russell S and Lumsden JS: Function and heterogeneity of fish lectins. *Vet Immunol Immunopathol* 2005, 108:111–120.

96. Ewart KV, Johnson SC, and Ross NW: Identification of a pathogen-binding lectin in salmon serum. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 1999, 123:9–15.
97. Jensen LE, Thiel S, Petersen TE, and Jensenius JC: A rainbow trout lectin with multimeric structure. *Comp Biochem Physiol B Biochem Mol Biol* 1997, 116:385–390.
98. Hoover GJ, el-Mowafi A, Simko E, Kocal TE, Ferguson HW, and Hayes MA: Plasma proteins of rainbow trout (*Oncorhynchus mykiss*) isolated by binding to lipopolysaccharide from *Aeromonas salmonicida*. *Comp Biochem Physiol B Biochem Mol Biol* 1998, 120:559–569.
99. Fock WL, Chen CL, Lam TJ, and Sin YM: Isolation and characterisation of a serum lectin from blue gourami, *Trichogaster trichopterus* (Pallus). *Fish Shellfish Immunol* 2000, 10:489–504.
100. Fock WL, Chen CL, Lam TJ, and Sin YM: Roles of an endogenous serum lectin in the immune protection of blue gourami, *Trichogaster trichopterus* (Pallus) against Aeromonas hydrophila. *Fish Shellfish Immunol* 2001, 11:101–113.
101. Tasumi S, Ohira T, Kawazoe I, Suetake H, Suzuki Y, and Aida K: Primary structure and characteristics of a lectin from skin mucus of the Japanese eel *Anguilla japonica*. *J Biol Chem* 2002, 277:27305–27311.
102. Tasumi S, Yang WJ, Usami T, Tsutsui S, Ohira T, Kawazoe I, Wilder MN, Aida K, and Suzuki Y: Characteristics and primary structure of a galectin in the skin mucus of the Japanese eel, *Anguilla japonica*. *Dev Comp Immunol* 2004, 28:325–335.
103. Ahmed H, Du SJ, O'Leary N, and Vasta GR: Biochemical and molecular characterization of galectins from zebrafish (*Danio rerio*): Notochord-specific expression of a prototype galectin during early embryogenesis. *Glycobiology* 2004, 14:219–232.
104. Ng NF, Trinh KY, and Hew CL: Structure of an antifreeze polypeptide precursor from the sea raven, *Hemitripterus americanus*. *J Biol Chem* 1986, 261:15690–15695.
105. Ng NF and Hew CL: Structure of an antifreeze polypeptide from the sea raven. Disulfide bonds and similarity to lectin-binding proteins. *J Biol Chem* 1992, 267:16069–16075.
106. Slaughter D, Fletcher GL, Ananthanarayanan VS, and Hew CL: Antifreeze proteins from the sea raven, *Hemitripterus americanus*. Further evidence for diversity among fish polypeptide antifreezes. *J Biol Chem* 1981, 256:2022–2026.
107. Ewart KV, Rubinsky B, and Fletcher GL: Structural and functional similarity between fish antifreeze proteins and calcium-dependent lectins. *Biochem Biophys Res Commun* 1992, 185:335–340.
108. Ewart KV and Fletcher GL: Herring antifreeze protein: Primary structure and evidence for a C-type lectin evolutionary origin. *Mol Mar Biol Biotechnol* 1993, 2:20–27.
109. Ewart KV, Li Z, Yang DS, Fletcher GL, and Hew CL: The ice-binding site of Atlantic herring antifreeze protein corresponds to the carbohydrate-binding site of C-type lectins. *Biochemistry* 1998, 37:4080–4085.
110. Achenbach JC and Ewart KV: Structural and functional characterization of a C-type lectin-like antifreeze protein from rainbow smelt (*Osmerus mordax*). *Eur J Biochem* 2002, 269:1219–1226.
111. Zelensky AN and Gready JE: C-type lectin-like domains in *Fugu rubripes*. *BMC Genomics* 2004, 5:51.
112. Zelensky AN and Gready JE: The C-type lectin-like domain superfamily. *FEBS J* 2005, 272:6179–6217.
113. Richards RC, Hudson DM, Thibault P, and Ewart KV: Cloning and characterization of the Atlantic salmon serum lectin, a long-form C-type lectin expressed in kidney. *Biochim Biophys Acta* 2003, 1621:110–115.
114. Vitved L, Holmskov U, Koch C, Teisner B, Hansen S, Salomonsen J, and Skjodt K: The homologue of mannose-binding lectin in the carp family Cyprinidae is expressed at high level in spleen, and the deduced primary structure predicts affinity for galactose. *Immunogenetics* 2000, 51:955–964.
115. Mistry AC, Honda S, and Hirose S: Structure, properties and enhanced expression of galactose-binding C-type lectins in mucous cells of gills from freshwater Japanese eels (*Anguilla japonica*). *Biochem J* 2001, 360:107–115.
116. Bayne CJ, Gerwick L, Fujiki K, Nakao M, and Yano T: Immune-relevant (including acute phase) genes identified in the livers of rainbow trout, *Oncorhynchus mykiss*, by means of suppression subtractive hybridization. *Dev Comp Immunol* 2001, 25:205–217.
117. Fujiki K, Bayne CJ, Shin DH, Nakao M, and Yano T: Molecular cloning of carp (*Cyprinus carpio*) C-type lectin and pentraxin by use of suppression subtractive hybridisation. *Fish Shellfish Immunol* 2001, 11:275–279.
118. Gracey AY, Troll JV, and Somero GN: Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*. *Proc Natl Acad Sci USA* 2001, 98:1993–1998.

119. Sandford R, Sgotto B, Aparicio S, Brenner S, Vaudin M, Wilson RK, Chissoe S, Pepin K, Bateman A, Chothia C, et al.: Comparative analysis of the polycystic kidney disease 1 (PKD1) gene reveals an integral membrane glycoprotein with multiple evolutionary conserved domains. *Hum Mol Genet* 1997, 6:1483–1489.
120. Zhang H, Robison B, Thorgaard GH, and Ristow SS: Cloning, mapping and genomic organization of a fish C-type lectin gene from homozygous clones of rainbow trout (*Oncorhynchus mykiss*). *Biochim Biophys Acta* 2000, 1494:14–22.
121. Neame PJ, Young CN, and Treep JT: Primary structure of a protein isolated from reef shark (*Carcharhinus springeri*) cartilage that is similar to the mammalian C-type lectin homolog, tetranectin. *Protein Sci* 1992, 1:161–168.
122. Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S, Smit A, et al.: Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. *Science* 2002, 297:1301–1310.
123. van Kooyk Y and Geijtenbeek TB: DC-SIGN: Escape mechanism for pathogens. *Nat Rev Immunol* 2003, 3:697–709.
124. Linehan SA, Martinez-Pomares L, and Gordon S: Macrophage lectins in host defence. *Microbes Infect* 2000, 2:279–288.
125. Ozeki Y, Matsui T, Nitta K, Kawauchi H, Takayanagi Y, and Titani K: Purification and characterization of beta-galactoside binding lectin from frog (*Rana catesbeiana*) eggs. *Biochem Biophys Res Commun* 1991, 178:407–413.
126. Uchiyama H, Komazaki S, Oyama M, Matsui T, and Ozeki Y: Distribution and localization of galectin purified from *Rana catesbeiana* oocytes. *Glycobiology* 1997, 7:1159–1165.
127. Marschal P, Herrmann J, Leffler H, Barondes SH, and Cooper DNW: Sequence and specificity of a soluble lactose-binding lectin from *Xenopus laevis* skin. *J Biol Chem* 1992, 267:12942–12949.
128. Shoji H, Nishi N, Hirashima M, and Nakamura T: Purification and cDNA cloning of Xenopus liver galectins and their expression. *Glycobiology* 2002, 12:163–172.
129. Shoji H, Nishi N, Hirashima M, and Nakamura T: Characterization of the Xenopus galectin family. Three structurally different types as in mammals and regulated expression during embryogenesis. *J Biol Chem* 2003, 278:12285–12293.
130. Fink NE, Caron M, Joubert R, Elola M T, Bladier D, and Herkovits J: Purification and some characteristics of a beta-galactoside binding soluble lectin from amphibian ovary. *FEBS Lett* 1987, 223:330–334.
131. Ahmed H, Pohl J, Fink NE, Strobel F, and Vasta GR. The primary structure and carbohydrate specificity of a beta-galactosyl-binding lectin from toad (*Bufo arenarum* Hensel) ovary reveal closer similarities to the mammalian galectin-1 than to the galectin from the clawed frog *Xenopus laevis*. *J Biol Chem* 1996, 271:33083–33094.
132. Vasta GR, Ahmed H, Amzel LM, and Bianchet MA: Galectins from amphibian species: Carbohydrate specificity, molecular structure, and evolution. *Trends Glycosci Glycotechnol* 1997, 9:131–144.
133. Elola MT, Cabada MO, Barisone GA, and Fink NE: Immunohistochemical localisation of a galectin from Bufo arenarum ovary. *Zygote* 1998, 6:1–9.
134. Bianchet MA, Ahmed H, Vasta GR, and Amzel LM: A soluble β-galactosyl-binding lectin (galectin) from toad (*Bufo arenarum* Hensel) ovary: Crystallographic studies of two protein–sugar complexes *Proteins* 2000, 40:378–388.
135. Allen HJ, Ahmed H, and Sharma A: Isolation of lactose-binding lectins from axolotl (*Ambystoma mexicanum*). *Comp Biochem Physiol* 1992, 103B:313–315.
136. Cooper DN: Galectinomics: A lesson in complexity. *Biochim Biophys Acta* 2002, 1572:209–231.
137. Fry BG and Wuster W: Assembling an arsenal: Origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences. *Mol Biol Evol* 2004, 21:870–883.
138. Corsortium ICGS: Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 2004, 432:695–716.
139. Smith J, Speed D, Law AS, Glass EJ, and Burt DW: In-silico identification of chicken immune-related genes. *Immunogenetics* 2004, 56:122–133.
140. Avery S, Rothwell L, Degen WD, Schijns VE, Young J, Kaufman J, and Kaiser P: Characterization of the first nonmammalian T2 cytokine gene cluster: The cluster contains functional single-copy genes for IL-3, IL-4, IL-13, and GM-CSF, a gene for IL-5 that appears to be a pseudogene, and a gene encoding another cytokinelike transcript, KK34. *J Interferon Cytokine Res* 2004, 24:600–610.
141. Degen WG, Daal N, Rothwell L, Kaiser P, and Schijns VE: Th1/Th2 polarization by viral and helminth infection in birds. *Vet Microbiol* 2005, 105:163–167.

142. Kaufman J, Milne S, Gobel TW, Walker BA, Jacob JP, Auffray C, Zoorob R, and Beck S: The chicken B locus is a minimal essential major histocompatibility complex. *Nature* 1999, 401:923–925.
143. Shiina T, Shimizu S, Hosomichi K, Kohara S, Watanabe S, Hanzawa K, Beck S, Kulski JK, and Inoko H: Comparative genomic analysis of two avian (quail and chicken) MHC regions. *J Immunol* 2004, 172:6751–6763.
144. MacDonald MR, Veniamin SM, Guo X, Xia J, Moon DA, and Magor KE: Genomics of antiviral defenses in the duck, a natural host of influenza and hepatitis B viruses. *Cytogenet Genome Res* 2007, 117:195–206.
145. Xia J, Radford C, Guo X, and Magor KE: Immune gene discovery by expressed sequence tag analysis of spleen in the duck (*Anas platyrhynchos*). *Dev Comp Immunol* 2007, 31:272–285.
146. Bernot A, Zoorob R, and Auffray C: Linkage of a new member of the lectin supergene family to chicken Mhc genes. *Immunogenetics* 1994, 39:221–229.
147. Plougastel B, Dubbelde C, and Yokoyama WM: Cloning of Clr, a new family of lectin-like genes localized between mouse Nkrp1a and Cd69. *Immunogenetics* 2001, 53:209–214.
148. Bates EE, Fournier N, Garcia E, Valladeau J, Durand I, Pin JJ, Zurawski SM, Patel S, Abrams JS, Lebecque S, et al.: APCs express DCIR, a novel C-type lectin surface receptor containing an immunoreceptor tyrosine-based inhibitory motif. *J Immunol* 1999, 163:1973–1983.
149. Mann K and Siedler F: The amino acid sequence of ovocleidin 17, a major protein of the avian eggshell calcified layer. *Biochem Mol Biol Int* 1999, 47:997–1007.
150. Mann K and Siedler F: Ostrich (*Struthio camelus*) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications. *Biochim Biophys Acta* 2004, 1696:41–50.
151. Lakshminarayanan R, Kini RM, and Valiyaveettil S: Investigation of the role of ansocalcin in the biomineralization in goose eggshell matrix. *Proc Natl Acad Sci USA* 2002, 99:5155–5159.
152. Mann K, Weiss IM, Andre S, Gabius HJ, and Fritz M: The amino-acid sequence of the abalone (*Haliotis laevigata*) nacre protein perlucin. Detection of a functional C-type lectin domain with galactose/mannose specificity. *Eur J Biochem* 2000, 267:5257–5264.
153. Killian CE and Wilt FH: Characterization of the proteins comprising the integral matrix of *Strongylocentrotus purpuratus* embryonic spicules. *J Biol Chem* 1996, 271:9150–9159.
154. Geider S, Baronnet A, Cerini C, Nitsche S, Astier JP, Michel R, Boistelle R, Berland Y, Dagorn JC, and Verdier JM: Pancreatic lithostathine as a calcite habit modifier. *J Biol Chem* 1996, 271:26302–26306.
155. Gerbaud V, Pignol D, Loret E, Bertrand JA, Berland Y, Fontecilla-Camps JC, Canselier JP, Gabas N, and Verdier JM: Mechanism of calcite crystal growth inhibition by the N-terminal undecapeptide of lithostathine. *J Biol Chem* 2000, 275:1057–1064.
156. Rabinovich GA, Liu FT, Hirashima M, and Anderson A: An emerging role for galectins in tuning the immune response: Lessons from experimental models of inflammatory disease, autoimmunity and cancer. *Scand J Immunol* 2007, 66:143–148.
157. Poirier F and Robertson EJ: Normal development of mice carrying a null mutation in the gene encoding the L14 S-type lectin. *Development* 1993, 119:1229–1236.
158. Puche AC, Poirier F, Hair M, Bartlett PF, and Key B: Role of galectin-1 in developing mouse olfactory system. *Dev Biol* 1996, 179:274–287.
159. McGraw J, Gaudet AD, Oschipok LW, Steeves JD, Poirier F, Tetzlaff W, and Ramer MS: Altered primary afferent anatomy and reduced thermal sensitivity in mice lacking galectin-1. *Pain* 2005, 114:7–18.
160. Plachta N, Annalheim C, Bissiere S, Lin S, Rüegg M, Hoving S, Müller D, Poirier F, Bibel M, and Barde YA: Identification of a lectin causing the degeneration of neuronal processes using engineered embryonic stem cells. *Nat Neurosci* 2007, 10:712–719.
161. Sakaguchi M, Shingo T, Shimazaki T, Okano HJ, Shiwa M, Ishibashi S et al.: A carbohydrate-binding protein, galectin-1, promotes proliferation of adult neural stem cells. *Proc Natl Acad Sci USA* 2006, 103:7112–7117.
162. Georgiadis V, Stewart HJ, Pollard HJ, Tavsanoglu Y, Prasad R, Horwood J, Deltour L, Gordring K, Poirier F, and Lawrence-Watt DJ: Lack of galectin-1 results in defects in myoblast fusion and muscle regeneration. *Dev Dyn* 2007, 236:1014–1024.
163. Toscano MA, Bianco GA, Ilarregui JM, Croci DO, Correale J, Hernandez JD, Zwirner NW, Poirier F, Riley EM, Baum LG, and Rabinovich GA: Differential glycosylation of T(H)1, T(H)2 and T(H)-17 effector cells selectively regulates susceptibility to cell death. *Nat Immunol* 2007, 8:825–834.
164. Barrionuevo P, Beigier-Bompadre M, Ilarregui JM, et al: A novel function for galectin-1 at the crossroad of innate and adaptive immunity: Galectin-1 regulates monocyte/macrophage physiology through a nonapoptotic ERK-dependent pathway. *J Immunol* 2007, 178:436–445.

165. Garin MI, Chu CC, Golshayan D, Cernuda-Morollon E, Wait R, and Lechler RI: Galectin-1: A key effector of regulation mediated by CD4 + CD25 + T cells. *Blood* 2007, 109:2058–2065.
166. Case D, Irwin D, Ivester C, Harral J, Morris K, Imamura M et al.: Mice deficient in galectin-1 exhibit attenuated physiological responses to chronic hypoxia-induced pulmonary hypertension. *Am J Physiol Lung Cell Mol Physiol* 2007, 292:L154–L164.
167. Colnot C, Ripoche MA, Milon G, et al.: Maintenance of granulocyte numbers during acute peritonitis is defective in galectin-3-null mutant mice. *Immunology* 1998, 94:290–296.
168. Hsu DK, Yang RY, Yu L, et al.: Targeted disruption of the galectin-3 gene results in attenuated peritoneal inflammatory responses. *Am J Pathol* 2000, 156:1073–1083.
169. Colnot C, Fowlis D, Ripoche MA, Bouchaert I, and Poirier F: Embryonic implantation in galectin-1/galectin-3 double mutant mice. *Dev Dyn* 1998, 211:306–313.
170. Chen HY, Sharma BB, Yu L et al.: Role of galectin-3 in mast cell functions: galectin-3-deficient mast cells exhibit impaired mediator release and defective JNK expression. *J Immunol* 2006, 177:4991–4997.
171. Zuberi RI, Hsu DK, Kalayci O et al.: Critical role for galectin-3 in airway inflammation and bronchial hyperresponsiveness in a murine model of asthma. *Am J Pathol* 2004, 165:2045–2053.
172. Bernardes ES, Silva NM, Ruas LP et al.: *Toxoplasma gondii* infection reveals a novel regulatory role for galectin-3 in the interface of innate and adaptive immunity. *Am J Pathol* 2006, 168:1910–1920.
173. Cao Z, Said N, Amin S, Wu HK, Bruce A, Garate M, Hsu DK, Kuwabara I, Liu FT, and Panjwani N: Galectins-3 and -7, but not galectin-1, play a role in re-epithelialization of wounds. *J Biol Chem* 2002, 277:42299–42305.
174. Levi G, Tarrab-Hazdai R, and Teichberg VI: Prevention and therapy with electrolectin of experimental autoimmune *Myasthenia gravis* in rabbits. *Eur J Immunol* 1983,13:500–507.
175. Offner H, Celnik B, Bringman TS, Casentini-Borocz D, Nedwin GE, and Vandenbark AA: Recombinant human b-galactoside binding lectin suppresses clinical and histological signs of experimental autoimmune encephalomyelitis. *J Neuroimmunol* 1990, 28:177–184.
176. Rabinovich GA, Daly G, Dreja H et al.: Recombinant galectin-1 and its genetic delivery suppress collagen-induced arthritis via T cell apoptosis. *J Exp Med* 1999, 190:385–398.
177. Santucci L, Fiorucci S, Cammilleri F, Servillo G, Federici B, and Morelli A: Galectin-1 exerts immunomodulatory and protective effects on concanavalin A-induced hepatitis in mice. *Hepatology* 2000, 31:399–406.
178. Santucci L, Fiorucci S, Rubinstein N et al.: Galectin-1 suppresses experimental colitis in mice. *Gastroenterology* 2003, 124:1381–1394.
179. Toscano MA, Commodaro AG, Ilarregui JM et al.: Galectin-1 suppresses autoimmune retinal disease by promoting concomitant Th2- and T regulatory-mediated anti-inflammatory responses. *J Immunol* 2006, 176:6323–6332.
180. Perone MJ, Bertera S, Tawadrous ZS et al.: Dendritic cells expressing transgenic galectin-1 delay onset of autoimmune diabetes in mice. *J Immunol* 2006a, 177:5278–5289.
181. Baum LG, Blackall DP, Arias-Magallano S et al.: Amelioration of graft versus host disease by galectin-1. *Clin Immunol* 2003, 109:295–307.
182. Zhu C, Anderson AC, Schubart A, et al.: The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nat Immunol* 2005, 6:1245–1252.
183. Liu FT and Rabinovich GA: Galectins as modulators of tumour progression. *Nat Rev Cancer* 2005, 5:29–41.
184. Ingrassia L, Camby I, Lefranc F, et al.: Anti-galectin compounds as potential anti-cancer drugs. *Curr Med Chem* 2006, 13:3513–3527.
185. Bresalier RS et al.: Metastasis of human colon cancer is altered by modifying expression of the β-galactoside-binding protein galectin-3. *Gastroenterology* 1998, 115:287–296.
186. Takenaka Y, Fukumori T, and Raz A: Galectin-3 and metastasis. *Glycoconj J* 2004, 19:543–549.
187. Inufusa H et al.: Role of galectin-3 in adenocarcinoma liver metastasis. *Int J Oncol* 2001, 19:913–919.
188. Song YK, Billiar TR, and Lee YJ: Role of galectin-3 in breast cancer metastasis: Involvement of nitric oxide. *Am J Pathol* 2002, 160:1069–1075.
189. Rubinstein N, Alvarez M, Zwirner NW, Toscano MA, Ilarregui JM, Bravo A, Mordoh J, Fainboim L, Podhajcer OL, and Rabinovich GA: Targeted inhibition of galectin-1 gene expression in tumor cells results in heightened T cell-mediated rejection: A potential mechanism of tumor-immune privilege. *Cancer Cell* 2004, 5:241–251.
190. Ley K: The role of selectins in inflammation and disease. *Trends Mol Med*. 2003, 9:263–268.
191. Kishore U, Greenhough TJ, Waters P, Shrive AK, Ghai R, Kamran MF, Bernal AL, Reid KB, Madan T, and Chakraborty T: Surfactant proteins SP-A and SP-D: structure, function and receptors. *Mol Immunol* 2006, 43:1293–1315.

192. Crocker PR, Paulson JC, and Varki A: Siglecs and their roles in the immune system. *Nat Rev Immunol* 2007, 7:255–266.
193. Toscano MA, Ilarregui JM, Bianco GA, et al.: Dissecting the pathophysiologic role of endogenous lectins: Glycan-binding proteins with cytokine-like activity? *Cytokine Growth Factor Rev* 2007, 18:57–71.
194. Yoder JA, Nielsen ME, Amemiya CT, and Litman GW: Zebrafish as an immunological model system. *Microbes Infect* 2002, 4:1469–1478.
195. Patton EE and Zon LI: The art and design of genetic screens: Zebrafish. *Nature Rev* 2001, 2:95666.
196. Jesuthasan S: Genetics and development: Zebrafish in the spotlight. *Science* 2002, 297:1484–1485.
197. Stickney HL, Barresi MJF, and Devoto SH: Somite development in zebrafish. *Develop Dynamics* 2000, 219, 287–303.
198. Danilova N and Steiner LA: B cells develop in the zebrafish pancreas. *Proc Natl Acad Sci USA* 2002, 99:13711–13716.
199. Herbomel P, Thisse B, and Thisse C, Ontogeny and behavior of early macrophages in the zebrafish embryo. *Development* 1999, 126:3735–3745.
200. Lieschke GJ, Oates AC, Crowhurst MO, Ward AC, and Layton JE: Morphologic and functional characterization of granulocytes and macrophages in embryonic and adult zebrafish. *Blood* 2001, 98:3087–3096.
201. Davis JM, Clay H, Lewis JL, Ghori N, Herbomel P, and Ramakrishnan L, Real-time visualization of mycobacterium-macrophage interactions leading to initiation of granuloma formation in zebrafish embryos. *Immunity* 2002, 17:693–702.
202. Neely MN, Pfeifer JD, and Caparon M: Streptococcus–Zebrafish model of bacterial pathogenesis. *Infect Immun* 2002, 70:3904–3914.
203. Altmann SM, Mellon MT, Distel DL, and Kim CH: Molecular and functional analysis of an interferon gene from the zebrafish. *Danio rerio, J Virol* 2003, 77:1992–2002.
204. Knaut H, Werz C, Geisler R, and Nusslein-Volhard C: Tubingen Screen Consortium 2003, A zebrafish homologue of the chemokine receptor Cxcr4 is a germ-cell guidance receptor. *Nature* 2000, 421:279–282.
205. Odom EW and Vasta GR: Characterization of a binary tandem domain F-type lectin from striped bass (*Morone saxatilis*). *J Biol Chem* 2006, 281:1698–1713.
206. Yoder JA, Litman RT, Mueller MG, Desai S, Dobrinski KP, Montgomery JS, Buzzeo MP, Ota T, Amemiya CT, Trede NS, et al.: Resolution of the novel immune-type receptor gene cluster in zebrafish. *Proc Natl Acad Sci USA* 2004, 101:15706–15711.
207. Haire RN, Rast JP, Litman RT, and Litman GW: Characterization of three isotypes of immunoglobulin light chains and T-cell antigen receptor alpha in zebrafish. *Immunogenetics* 2000, 51:915–923.
208. Gauthier L, Rossi B, Roux F, Termine E, and Schiff C: Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. *Proc Natl Acad Sci USA* 2002, 99:13014–13019.
209. Amatruda JF, Shepard JL, Stern HM, and Zon LI: Zebrafish as a cancer model system. *Cancer Cell* 2002, 1:229–231.
210. Grunwald DJ and Eisen JS: Headwaters of the zebrafish—Emergence of a new model vertebrate. *Nat Rev Genet* 2002, 3:717–724.
211. Langenau DM, Traver D, Ferrando AA, Kutok JL, Aster JC, Kanki JP, Shuo L, Prochownik E, Trede NS, Zon LI, and Look AT: Myc-induced T cell leukemia in transgenic zebrafish. *Science* 2003, 299:887–890.
212. Stern HM and Zon LI: Cancer genetics and drug discovery in the zebrafish. *Nat Rev Cancer* 2003, 3:1–7.
213. Berghmans S, Jette C, Langenau D, Hsu K, Stewart R, Look T, and Kanki JP: Making waves in cancer research: New models in the zebrafish. *Biotechniques* 2005, 39:227–237.
214. Lam SH, Wu YL, Vega VB, Miller LD, Spitsbergen J, Tong Y, Zhan H, Govindarajan KR, Lee S, Mathavan S, Krishna Murthy KR, Buhler DR, Liu ET, and Gong Z: Conservation of gene expression signatures between zebrafish and human liver tumors and tumor progression. *Nat Biotechnol* 2006, 24:73–75.
215. Lam SH and Gong Z: Modeling liver cancer using zebrafish: A comparative oncogenomics approach. *Cell Cycle* 2006, 5:573–577.
216. Yang HW, Kutok JL, Lee NH, Piao HY, Fletcher CD, Kanki JP, and Look AT: Targeted expression of human MYCN selectively causes pancreatic neuroendocrine tumors in transgenic zebrafish. *Cancer Res* 2004, 64:7256–7262.

217. Berghmans S, Murphey RD, Wienholds E, Neuberg D, Kutok JL, Fletcher CD, Morris JP, Liu TX, Schulte-Merker S, Kanki JP, Plasterk R, and Zon LI, Look AT: tp53 mutant zebrafish develop malignant peripheral nerve sheath tumors. *Proc Natl Acad Sci USA* 2005, 102:407–412.
218. Amsterdam A, Sadler KC, Lai K, Farrington S, Bronson RT, Lees JA, and Hopkins N: Many ribosomal protein genes are cancer genes in zebrafish. *PLOS Biol* 2004, 2:690–698.
219. Shepard JL, Amatruda JF, Stern HM, Subramanian A, Finkelstein D, Ziai J, Finley KR, Pfaff KL, Hersey C, Zhou Y, Barut B, Freedman M, Lee C, Spitsbergen J, Neuberg D, Weber G, Golub TR, Glickman JN, Kutok JL, Aster JC, and Zon LI: A zebrafish bmyb mutation causes genome instability and increased cancer susceptibility. *Proc Natl Acad Sci USA* 2005, 102:13194–13199.
220. Langheinrich U, Hennen E, Stott G, and Vacun G: Zebrafish as a model organism for the identification and characterization of drugs and genes affecting p53 signaling. *Curr Biol* 2002, 12:2023–2028.
221. Stern HM, Murphey RD, Shepard JL, Amatruda JF, Straub CT, Pfaff KL, Weber G, Tallarico JA, King RW, and Zon LI: Small molecules that delay S phase suppress a zebrafish bmyb mutant. *Nat Chem Biol* 2005, 1:366–370.
222. Ghosh C and Collodi P: Culture of cells from zebrafish (*Brachydanio rerio*) blastula-stage embryos. *Cytotechnology* 1994, 14:21–26.
223. Driever W and Rangini Z: Characterization of a cell line derived from zebrafish (*Brachydanio rerio*) embryos, *in vitro*. *Cell Dev Biol Anim* 1993, 29A:749–754.
224. Ghosh C, Zhou YL, and Collodi P: Derivation and characterization of a zebrafish liver cell line. *Cell Biol Toxicol* 1994, 10:167–176.
225. Paw BH and Zon LI: Primary fibroblast cell culture. *Methods Cell Biol* 1999, 59:39–43.
226. Ma C, Fan L, Ganassin R, Bols N, and Collodi P: Production of zebrafish germ-line chimeras from embryo cell cultures. *Proc Natl Acad Sci USA* 2001, 98:2461–2466.
227. Vasta GR, Ahmed H, Du S, and Henrikson D: Galectins in teleost fish: Zebrafish (*Danio rerio*) as a model species to address their biological roles in development and innate immunity. *Glycoconj J* 2004, 21:503–521.
227a. Ahmed H and Vasta GR: Unlike mammalian GRIFIN, the Zebrafish homologue (DrGRIFIN) represents a functional carbohydrate-binding galectin. *Biochem Biophys Res Commun* 2008, 371:350–355.
228. Arata Y, Hirabayashi J, and Kasai K: Structure of the 32-kDa galectin gene of the nematode *Caenorhabditis elegans*. *J Biol Chem* 1997, 272:26669–26677.
229. Pace KE, Lebestky T, Hummel T, Arnoux P, Kwan K, and Baum LG: Characterization of a novel *Drosophila melanogaster* galectin. Expression in developing immune, neural, and muscle tissues. *J Biol Chem* 2002, 277:13091–13098.
230. Shwede T, Kopp J, Guex N, and Peitsch MC: SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Research* 2003, 31:3381–3385.
231. Liao DI, Kapadia G, Ahmed H, Vasta GR, and Herzberg O: Structure of S-lectin, a developmentally regulated vertebrate β-galactoside-binding protein. *Proc Natl Acad Sci USA* 1994, 91:1428–1432.
232. Shirai T, Mitsuyama C, Niwa Y, Matsui Y, Hotta H, Yamane T, Kamiya H, Ishii C, Ogawa T, and Muramoto K: High resolution structure of Conger eel galectin, congerin I, in lactose-liganded and ligand-free forms: Emergence of a new structure class by accelerated evolution. *Structure* 1999, 7:1223–1233.
233. Shirai T, Matsui Y, Shionyu-Mitsuyama C, Yamane T, Kamiya H, Ishii C, Ogawa T, and Muramoto K: Crystal structure of a conger eel galectin (congerin II) at 1.45 Å resolution: Implication for the accelerated evolution of a new ligand-binding site following gene duplication. *J Mol Biol* 2002, 321:879–889.
234. Nasevicius A and Ekker SC: Effective targeted gene "knockdown" in zebrafish. *Nature Genet* 2000, 26:216–220.
235. Ekker SC and Larson JD: Morphant technology in model developmental systems. *Genesis* 2001, 30:89–93.
236. Draper BW, Morcos PA, and Kimmel, CB: Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholino oligos: A quantifiable method for gene knockdown. *Genesis* 2001, 30:154–156.
237. Du SJ, Devoto S, Westerfield M, and Moon RT: Positive and negative regulation of muscle cell identity by members of the *hedgehog* and *TGF* gene families. *J Cell Biol* 1997, 139:145–156.
238. van der Sar AM, Musters RJ, van Eeden FJ, Appelmelk BJ, Vandenbroucke-Grauls CM, and Bitter W: Zebrafish embryos as a model host for the real time analysis of *Salmonella typhimurium* infections. *Cell Microbiol* 2003, 5:601–611.
239. Lewis KE, Currie PD, Roy S, Schauerte H, Haffter P, and Ingham PW: Control of muscle cell-type specification in the zebrafish embryo by Hedgehog signaling. *Dev Biol* 1999, 216:469–480.
240. Du SJ and Dienhart M: Zebrafish *Tiggy-Winkle Hedgehog* promoter directs notochord and floor plate green fluorescent protein expression in transgenic zebrafish embryos. *Dev Dynamics* 2001, 222:655–666.

241. Du SJ and Dienhart M: Gli2 mediation of hedgehog signals in slow muscle induction in zebrafish. *Differentiation* 2001, 67:84–91.
243. Odenthal J, Haffter P, Vogelsang E, Brand M, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Kelsh RN, Mullins MC, Warga RM, Allende ML, Weinberg ES, and Nüsslein-Volhard C: Mutations affecting the formation of the notochord in the zebrafish. *Danio rerio*, *Development* 1996, 123:103–115.
244. Stemple DL, Solnica-Krezel L, Zwartkruis F, Neuhauss SC, Schier AF, Malicki J, Stainier DY, Abdelilah S, Rangini Z, Mountcastle-Shah E, and Driever W: Mutations affecting development of the notochord in zebrafish. *Development* 1996, 123:117–128.

Part III

Glycoprotein Folding, Sorting and Secretion, Targeting, Degradation, and Clearance

8 Calreticulin and Calnexin as Chaperones in Glycoprotein Folding

Julio J. Caramelo and Armando J. Parodi

CONTENTS

8.1 INTRODUCTION

Protein folding efficiency is enhanced *in vivo* by an array of chaperones and folding-assisting enzymes that cover the entire protein folding pathway, from the exit of the polypeptide from the ribosome, its translocation through membranes, disulfide bond formation, tertiary structure acquisition, and oligomer assembly. A protein acquires diverse conformational states during these processes, from extended and disordered structures when exiting the ribosome to more structured intermediates such as molten globule-like forms. Accordingly, the high diversity of chaperones reflects the conformational variety of their substrates. From a statistical point of view, the native state represents an extremely small fraction of the conformational space available to a protein, which may easily fall into irreversible kinetic traps during the folding process. In this sense, the crowded environment inside a cell represents a hostile scenario for protein folding, where protein aggregation could be the rule rather than the exception. One of the main functions of chaperones, perhaps the most important, is to inhibit nonspecific protein interactions, directing the protein to a productive folding pathway.

As will be described below, N-glycosylation plays an important role during glycoprotein maturation *in vivo*. It represents the most abundant posttranslational modification as about 80% of the proteins that enter the secretory pathway (that is about 40% of total proteins synthesized in a mammalian cell) are N-glycosylated in the consensus sequence Asn-X-Ser/Thr (where X may be any amino acid except for proline) by the oligosaccharyltransferase, a multimeric protein complex associated with the translocon. The structural variety of N-glycans allows them to fulfill very diverse roles, from modulating the structural properties of the protein moiety, to be key mediators of cellular

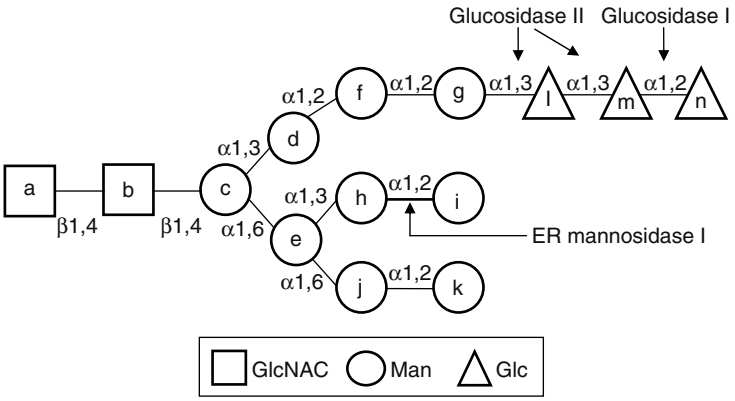


FIGURE 8.1 Structure of glycans. The figure represents the glycan transferred by the oligosaccharyltransferase. Lettering (a, b, c) corresponds to the order of addition of monosaccharides in the synthesis of the dolichol-P-P derivative. Bonds cleaved by glucosidases I and II (GI and GII) as well as by ER mannosidase I are indicated.

recognition processes. The glycan structural diversity present in mature, fully processed glycoproteins arises mainly as a consequence of the activity of Golgi-resident glycosidases and glycosyltransferases, and is in sharp contrast with the conserved structure of the glycan that is originally transferred from dolichol-P-P derivatives to proteins in the endoplasmic reticulum (ER) of most eukaryotic cells ($Glc_3Man_9GlcNAc_2$) (Figure 8.1). Prior to its processing in the Golgi, the transferred glycan is trimmed in the ER, where the three glucoses are removed by the successive action of glucosidases I and II (GI and GII). Moreover, several mannoses may be excised also. A puzzling issue is why a cell would transfer a glycan that is first trimmed in the ER to be next rebuilt with a modified structure? Why not simply transfer the core glycan and elongate it in the Golgi? The answer to this question took more than 15 years of research, and many of its aspects still remain obscure. It is now clear that N-glycans are deeply involved in protein folding in the ER, as they behave as an information platform that encodes the conformational status of glycoproteins. N-glycans are also involved in the pathways that divert irreparably misfolded proteins to degradation or, alternatively, correctly folded species to the Golgi apparatus.

8.2 CALRETICULIN AND CALNEXIN: UNCONVENTIONAL CHAPERONES

Proteins entering the ER face common challenges to acquire their native conformation. To this end, they take advantage of a limited array of chaperones and folding-catalyzing enzymes, and use a common N-glycan-based system to assist their folding. Why N-glycosylation has arisen during evolution as a fundamental process only in the secretory pathway and not in other organelles is not clear, but the dominant conditions in the ER, with an elevated calcium concentration, an oxidizing redox potential, and a high protein concentration, make it the perfect environment for protein aggregation, probably worse than any other cellular compartment. Central to the role of N-glycans in the ER are two proteins of the legume lectin family, calreticulin (CRT), a 46 kDa (400 residues) soluble ER-resident protein, and its paralog calnexin (CNX), a 65 kDa (572 residues) type I membrane protein. These lectins are part of a system known as glycoprotein-folding quality control (QC) that supervises and assists the folding maturation of glycoproteins in the ER [1]. At variance with other chaperones that rely on polypeptide-based interactions for their activity, the biological function of CRT and CNX is centered on glycan-mediated interactions. CRT has been localized to many subcellular compartments and has been implicated in diverse functions, such as protein folding, calcium homeostasis, gene transcription regulation, complement activation, and angiogenesis. This chapter will be mainly focused on the role of CRT and CNX in glycoprotein folding maturation in the ER.

CRT was first isolated from bovine liver and characterized as a calcium-binding protein [2]. Its cloning from rabbit [3] and mouse [4] showed that it is a typical ER-resident protein, endowed with a NH_2-terminal signal sequence and a KDEL-like C-terminal ER retention/retrieval signal. CNX was isolated and cloned from dog pancreas and firstly characterized as a calcium-binding protein displaying high homology to CRT [5]. The fact that CNX interacts preferentially with glycoproteins, and that such interaction can be disrupted by N-glycosylation inhibitors, prompted further research on the role of early N-glycan intermediates on CNX activity. Incubation of CNX with a mixture of $Glc_{0-3}Man_9GlcNAc_2$ glycans showed its exquisite specificity for $Glc_1Man_9GlcNAc_2$ [6]. This report was followed shortly by a similar observation for CRT [7]. Similarly to other ER chaperones, it was found that the expression of these lectins is induced by heat shock [8,9], amino acid deprivation [10], heavy metals, and calcium-mobilizing agents [11,12], thus indicating their role in protein folding. In agreement with these results, the 5′ promoter region of human CRT gene displays regulatory regions also present in the promoter region of other ER chaperones such as BiP, GRp94, and protein disulfide isomerase [13].

8.2.1 Structure

CNX is a type-I integral membrane protein with a transmembrane segment near its C-terminal end and a RKPRRE ER-localization signal located in a short cytosolic C-terminal tail. The crystal structure of CNX's ectodomain reveals a 140 Å long protruding arm inserted into a globular domain, which shows a fold similar to that of the legume lectin family [14] (Figure 8.2A). The

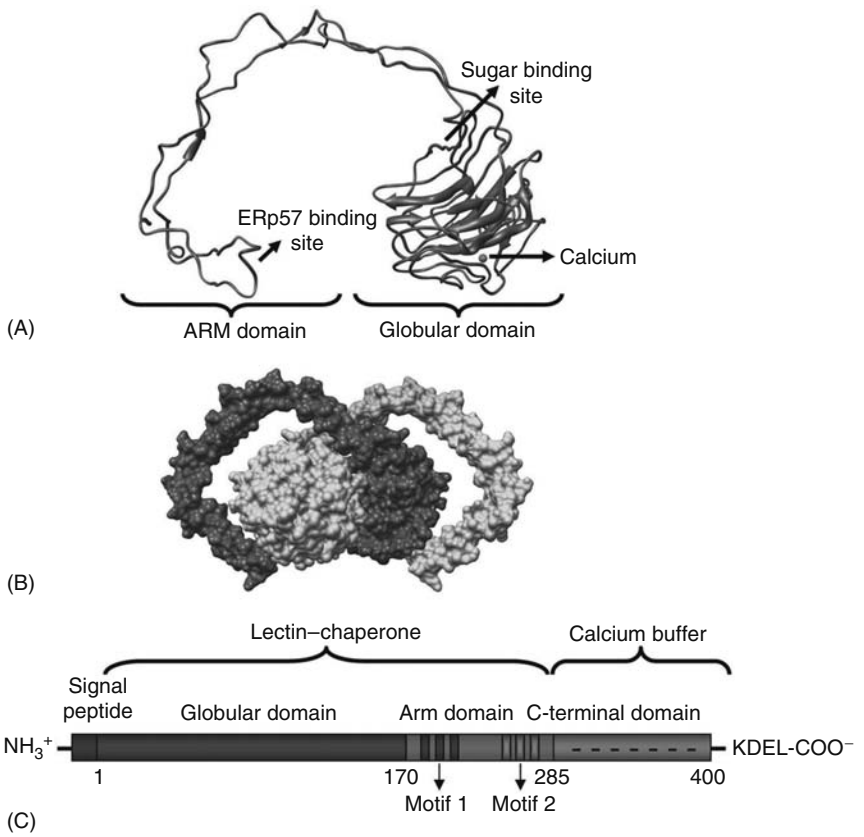


FIGURE 8.2 (See CD for color figure.) Structure of lectin–chaperones. (A) Crystal structure of the lumenal portion of CNX (pdb 1JHN). (B) The arm domain of CNX embraces the globular domain of a symmetry-related molecule. (C) Domain organization of CRT.

globular domain covers two regions of the primary sequence (residues 1–262 and 415–458), and is organized in two antiparallel β sheets, one concave (six strands) and the other convex (seven strands). Glucose unit binds to a shallow depression located on the surface of the concave β sheet, with its ring resting on Met189 and its hydroxyl groups hydrogen bonded to Tyr165, Lys186, and Glu217. Adjacent to this site, there are several hydrogen bond donors and acceptors located on the surface of the concave β sheet, which may contact the mannose residues of the glycan. Also in the globular domain, but opposite to the sugar binding face, a calcium ion is present, which accounts for the high-affinity, low-capacity calcium-binding activity of the protein (K_d ~ 2 μM) [15]. This calcium ion fulfills a structural role by stabilizing the lectin and, in principle, does not play a direct role in sugar binding. The arm domain, also known as P-domain, is a unique feature of these proteins not shared with the legume lectins. It is formed by a proline-rich tandem sequence repeats in which four copies of motif 1, IxDP(D/E)(A/D)xKP(E/D)DWD(D/E), and four copies of motif 2, GxWxxPxIxNPxY, are arranged following a 11112222 pattern. This domain can be divided into four modules, each formed by a copy of motif 1 that interacts head-to-tail with a copy of motif 2. A similar arrangement is observed in CRT, where three tandem repeats of motifs 1 and 2 are arranged following a 111222 pattern. Indeed, the structure of CRT arm domain solved by nuclear magnetic resonance (NMR) shows a highly flexible hairpin structure similar (but shorter, 110 Å) to that of CNX [16]. The flexibility of the CRT arm domain was also observed by electron microscopy, where it adopts various curved shapes [17]. It has been proposed that the arm domain stabilizes the lectin–glycoprotein complex by providing additional points of contact, thus increasing the chaperone ability [18], and is also necessary for optimal lectin activity [19,20]. Interestingly, the intermolecular contacts in the CNX crystal are mediated by the arm, which embraces the globular domain of a symmetry-related molecule (Figure 8.2B), suggesting a similar disposition for the interaction of the lectins with folding glycoproteins. In such an arrangement, the arm domain would protect the bound glycoprotein from interacting with other incompletely folded proteins. Noticeably, the tip of the P-domain constitutes an autonomous folding unit, being one of the smallest natural sequences that form a stable nonhelical fold in the absence of disulfide bonds or bound metals [21].

No structure of the globular domain of CRT is available, but since both lectins share ~40% sequence identity and present indistinguishable lectin specificity, a similar fold is predicted. In addition, CRT has a highly acidic C-terminal domain (residues 285–400) that bind approximately 20 calcium ions with low affinity (K_d ~ 2 mM), which is involved in calcium storage in the lumen of the ER (Figure 8.2C) [15]. This domain is dispensable for the lectin activity of CRT [22]. Given its very low hydrophobicity, the C-terminal domain is predicted to adopt a natively unfolded structure, in agreement with its high susceptibility toward proteases [23]. This domain is probably loosely connected to the rest of the protein, as observed by electron microscopy [17]. The C-terminal domain has been implicated also in the retention of CRT in the ER in a novel mechanism that would operate in conjunction with the KDEL-based retention/retrieval system [24].

8.2.2 ERp57

CRT and CNX associate *in vitro* and *in vivo* with ERp57, a member of the protein disulfide isomerase (PDI) family [25,26]. This association is independent of ligand binding to the lectins, since it is observed also in cells lacking GI, that is, cells in which no monoglucosylated glycans and thus no glycoprotein–CNX/CRT complexes are formed [25]. ERp57 is composed of four thioredoxin-like domains (a, b, b′ and a′), in which the a and a′ domains display the active -CXXC- motifs [27,28]. The amino-terminal cysteine residue of these motifs form intermolecular disulfide bridges with glycoproteins bound to CRT or CNX, and usually their association with ERp57 is also abrogated when glycoprotein binding to the lectins is inhibited [29–32]. ERp57 activity improves when its substrates are bound to the lectins. For instance, the ability of ERp57 to catalyze disulfide bridge acquisition of denatured monoglucosylated RNAseB *in vitro* is enhanced in the presence of CRT

and CNX, while the opposite effect is observed during refolding catalyzed by PDI [33]. Conversely, BiP enhances PDI activity but no effect is observed with respect to ERp57, suggesting that both enzymes have evolved to operate in conjunction with only one of the main ER chaperone systems, i.e., PDI working preferentially with the chaperone system centered on BiP and ERp57 with that based on the lectins [34]. Nevertheless, in some cases, ERp57 has been detected in complexes with ER proteins that are not associated with the lectins [35,36]. The ERp57–CRT association is mediated by residues 225–251 located on the tip of CRT's arm domain and the b′ domain of ERp57 [37–39], and a similar interaction has been described with respect to CNX [20,40]. Given the high flexibility of the arm domain, this association would allow the variation of the distance between ERp57 and the folding glycoprotein bound to the lectins, enabling the enzyme to scan for disulfide bridges located at distant positions. The moderate stability of the binary ERp57–CRT complex ($K_d \sim 9\,\mu M$) and its fast off-rate ($k_{off} > 1000\,s^{-1}$) probably allows ERp57 to rapidly sense for preexisting lectin–glycoprotein complexes [37]. PDI and ERp57 share a similar domain architecture, although both enzymes have developed different strategies to bind their substrates, since the PDI b and b′ domains are able to directly contact their substrates whereas the homologous ERp57 domains have evolved to interact with the arm domains of CNX and CRT, which in turn present most known substrates to ERp57. Therefore, CRT and CNX constitute a system specially fitted to present glycosylated substrates to ERp57.

8.2.3 Sugar Binding

CRT and CNX specifically bind monoglucosylated high mannose glycans [41]. The terminal glucose is essential for binding, although CRT affinity for this monosaccharide is very low. In addition, the following three mannoses present in the α1–3 branch make significant contributions to complex formation ($K_b = 2.2 \times 10^4\ M^{-1}$, $56.3 \times 10^4\ M^{-1}$, and $102 \times 10^4\ M^{-1}$ for Glcα1–3Man, Glcα1–3Manα1–2Man, and Glcα1–3Manα1–2Manα1–2Man, respectively) (Figure 8.1) [42]. Accordingly, site-directed mutations in positions that bind directly to the Glc residue (Y109 and D135 in rat CRT) are less tolerated than changing residues that contact the following mannose residues [43]. In addition, the glucose equatorially oriented 2-hydroxyl group is fundamental to establish the complex, since CRT is unable to bind 2-deoxygucose α1–3 mannose [43]. The lectins display substantial interaction with $Glc_1Man_5GlcNAc_2$, albeit with a lower affinity than with $Glc_1Man_9GlcNAc_2$, but they are unable to bind $Glc_1Man_4GlcNAc_2$, showing that the α1–6-mannose branch point of the glycan core also contributes to sugar binding [7,19]. The moderate affinity of CRT for $Glc_1Man_9GlcNAc_2$ ($K_b = 4.11 \times 10^5\ M^{-1}$) may have functional implications, as it probably allows the action of sugar-modifying enzymes such as GII during the dissociation phase of the lectin–glycan equilibrium [44]. As expected, many amino acids involved in sugar binding are shared by CRT and CNX, but this conservation is not absolute. For instance, residues Y109, K111, Y128, and D317 (numbered following the sequence of rabbit CRT) are vital for the lectin activity of both proteins, but M131 and D160, also present in both lectins, are important only for CNX [45].

8.2.4 Protein Binding

During their transit through the ER, most glycoproteins studied so far interact with CNX, CRT, or with both, either sequentially or simultaneously [46]. The underlying reasons for an N-glycan to interact with a particular lectin can be partially explained by taking into account their topological differences. CNX, being attached to the membrane, interacts preferentially with glycans placed near the ER membrane. In contrast, CRT binds preferentially glycoproteins soluble in the ER lumen or N-glycans placed far from the membrane [47,48]. For example, growth stages of tyrosinase, coagulation factor V, and influenza hemagglutinin exhibiting N-glycans close to the membrane are bound preferentially to CNX, but as proteins elongate and their glycans move away from the membrane they interact with

CRT [49,50]. Furthermore, a CRT mutant endowed with a membrane anchor exhibits a binding specificity similar to CNX and, reciprocally, when a soluble version of CNX is expressed, its specificity shifts to that observed for CRT [51]. Nevertheless, this interchangeability is partial since a soluble version of CNX is unable to replace CRT in assisting the peptide loading of the major histocompatibility complex (MHC) class I complex [52]. Besides, many proteins normally associated with CNX in wild-type cells, associate instead with BiP and not with CRT in $CNX^{-/-}$ cells [53], although under similar conditions, proteins of viral origin appeared in complexes with CRT. This observation suggests a more stringent requirement of certain viral proteins for the lectin-based folding enhancement compared to cellular proteins and may explain why viral replication is more sensitive to the presence of GI/GII inhibitors than cellular viability [54–56]. In addition, CRT- or CNX-encoding genes knockout is lethal in mice, showing that the remaining lectin cannot fully replace the deleted one [57,58]. On the other hand, certain microorganisms only code for one of the lectins, as for instance *Saccharomyces cerevisiae* or *Schizosaccharomyces pombe*, which only code for CNX or trypanosomatid protozoa, in which only CRT coding information is found.

Glycoprotein binding to these lectins is mediated by their glycan moieties, although polypeptide-based interactions may be important in some cases. Two alternative binding models have been proposed. In the "lectin-only" model, glycoprotein binding is mediated solely through their glycan moiety, while in the "dual-binding" model, there is a contribution of both polypeptide and glycan moieties. Substantial experimental evidence supports both models, and probably the relative contribution of polypeptide-based contacts varies with the inherent characteristics of a particular protein [59]. The "lectin-only" model is supported mainly by experiments using glycosylation or GI/GII inhibitors or by expressing proteins bearing site-specific mutations of consensus glycosylation sites, as under those conditions, protein association with the lectins is severely affected [1,32,60–65]. Furthermore, protein binding to CNX is almost totally impaired in cells lacking GI or GII [66]. Dissociation of the already formed lectin–glycoprotein complexes is prevented upon addition of GI/GII inhibitors, thus showing the importance of GII during the dissociating phase of the cycle [67]. Nevertheless, it has been shown that some nonglycosylated proteins remain associated with the lectins if care is taken during cell lysis in order to preserve weak complexes [68,69]. This result could be due to an unspecific association between the lectins and protein aggregates, but this possibility has been ruled out in some cases [68,70]. CNX and CRT can inhibit *in vitro* the aggregation of both glycosylated and nonglycosylated proteins, and they are able to form stable complexes with unfolded, nonglycosylated substrates but not with their native forms [71,72]. Interestingly, the addition of monoglucosylated glycans to these assays impaired the ability of CRT to inhibit the aggregation of glycosylated and nonglycosylated proteins [71], although the presence of a monoglucosylated glycan moiety on the substrates improved the chaperone capacity of CNX but not that of a classic chaperone such as BiP [73]. In addition, CNX and CRT mutants devoid of lectin activity inhibited the thermal aggregation of nonglycosylated substrates, but not that of the monoglucosylated substrate jack bean α-mannosidase [45,74]. Furthermore, the CNX lectin negative mutant was able to bind MHC I heavy chain *in vivo* [74]. Interestingly, CNX interacts with integral membrane proteins through their transmembrane domains in a glycan-independent mode [75,76]. For instance, CNX associates *in vivo* with EDEM, a lectin that targets high-mannose glycoproteins for degradation, through their transmembrane domains, physically linking the retention and degradation of irreparable misfolded species [77].

Although CRT and CNX are proteins dedicated to deal with partially folded species, their structure is marginally stable ($T_m \sim 42°C–45°C$) [23]. This property precludes a neat explanation for their precise role in many thermal-induced aggregation assays carried out near the lectin melting temperatures. Nevertheless, CRT inhibits protein aggregation of some substrates in experiments developed at temperatures far from any detectable conformational change (MHC class I and deglycosylated IgY at 37°C and 31°C, respectively) [71]. On the other hand, the crystal structure of CNX does not reveal any obvious site that could bind hydrophobic surfaces or segments exposed on misfolded

glycoproteins, at variance with classic chaperones that display binding cages or hydrophobic clefts. This constitutes a major drawback for the "dual-binding" model, but the static picture captured in the crystal may hinder an alternative conformation unfavorable under the crystallization conditions. For instance, upon heat shock or calcium depletion, both CRT and CNX undergo a conformational change that induces their oligomerization and increases their ability to bind nonglycosylated substrates [78,79]. This result suggests that changes in the ER environment may modulate their chaperone activity.

The scarcity of suitable model substrates has precluded a definitive answer regarding the relative importance of polypeptide-based interactions. Binding of ribonuclease B to CNX or CRT depends exclusively on the presence of monoglucosylated glycans, regardless of the conformational status of the substrate [80,81], although this highly hydrophilic protein is a poor candidate to display hydrophobic interactions. One of the few natural monoglucosylated glycoproteins readily available is IgG from egg yolk (IgY). The affinity of CRT for $Glc_1Man_9GlcNAc_2$ and IgY, measured by surface plasmon resonance, is very similar [44], showing that the main determinant for binding is the N-glycan (the presence of additional polypeptide-based contact sites would otherwise have induced a higher affinity for the glycoprotein). An alternative approach is to use a genetically tailored *S. cerevisiae* strain with disrupted *gls2* (GII), *mns1* (ER mannosidase), and *alg8 genes* (the last gene is responsible for the addition of the second glucose residue to the dolichol pyrophosphate-linked glycan precursor). Glycoproteins synthesized in this triple mutant display $Glc_1Man_9GlcNAc_2$ N-glycans. CRT binding *in vitro* to monoglucosylated MHC I expressed in this strain is absolutely dependent on the glycan and independent of the glycoprotein folding status, since similar binding constants were measured for free heavy chain and the empty heterodimer of the heavy chain, β2 microglobulin, and the heterodimer loaded with a peptide [82]. These examples support the "lectin-only" model, but clearly more experimental systems are needed to finally settle the issue. There is a consensus that for glycoproteins glycan-based interactions are determinant for the initial establishment of the lectin–glycoprotein complex [6,83], and that the significance of subsequent protein–protein interactions depends on the particular glycoprotein involved.

8.3 GLYCOPROTEIN FOLDING QUALITY CONTROL

Monoglucosylated glycoproteins may form in the ER either by partial deglucosylation of the transferred glycan by the sequential action of GI and GII, or by reglucosylation of the trimmed glycan by the UDP-Glc:glycoprotein glucosyltransferase (UGGT) (Figure 8.3). UGGT activity was first detected in protozoan parasites, since at variance with other wild-type eukaryotic cells, these microorganisms transfer truncated versions of the canonical structure, which are devoid of glucose residues ($Man_{6,7,9}GlcNAc_2$, depending on the species) [84]. Nevertheless, as monoglucosylated protein-linked glycans were detected, it was inferred that glucose units had necessarily been added to glycans after their transfer from the dolichol derivative [85]. Glucose removal from $Glc_3Man_9GlcNAc_2$ starts while glycoproteins are still in the process of translocation into the ER, immediately after glycan transfer by the oligosaccharyltransferase. The terminal glucose is removed by GI, an ER type II membrane-bound enzyme [86–88], followed by the removal of the glucoses by GII. At variance with other glycosidases of the secretory pathway, which are type II membrane proteins, GII is a soluble enzyme in the lumen of the ER. GII is a heterodimeric protein, in which the catalytic subunit (α) is retained in the ER due to the presence of an associated small subunit (β), which in turn displays an ER retention/retrieval sequence. In principle, this association is only needed for the intracellular localization of the α subunit since its activity *in vitro* is not affected by the absence of the β subunit although both subunits are required for activity *in vivo* [89]. It has been described that GII-catalyzed removal of the second glucose residue is a regulated process, as it is more active on glycoproteins bearing more than one N-glycan [90].

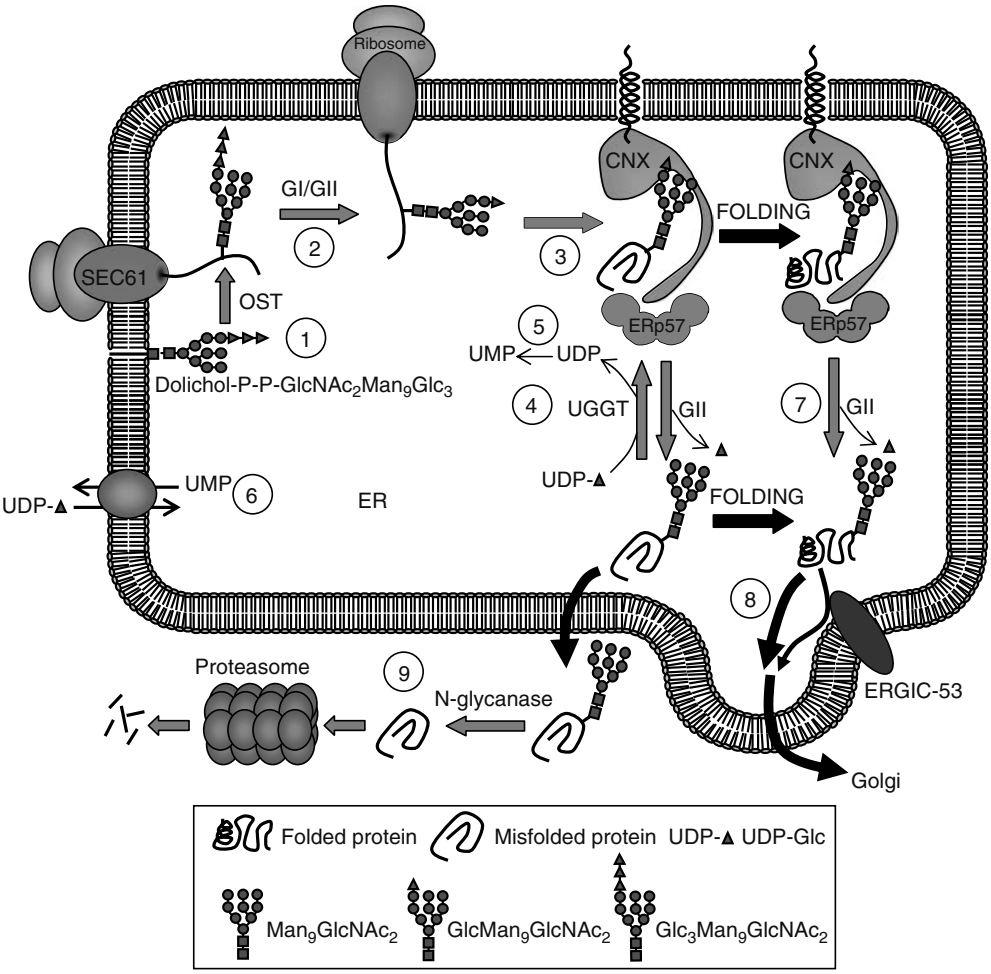


FIGURE 8.3 (See color insert following blank page 170. Also see CD for color figure.) Glycoprotein folding quality control. Proteins entering the ER are N-glycosylated by the oligosaccharyltranferase (OST) as they emerge from the translocon (Sec61) (1). Both outer glucoses are removed by the sequential action of GI and GII to generate monoglucosylated species (2) that are recognized by CNX and/or CRT (only CNX is shown), which are associated with ERp57 (3). The complex between the lectins and folding intermediates/misfolded glycoproteins dissociates upon removal of the last glucose by GII, and is reformed by UGGT activity (4). UDP generated by UGGT is hydrolysed to UMP (5), which is antiported to the cytosol by UDP-Glc entrance (6). Once glycoproteins have acquired their native conformations, either free or complexed with the lectins, GII hydrolyses the remaining glucose residue and releases the glycoproteins from the lectin anchors (7). These species are not recognized by UGGT and are transport to the Golgi (some of them by ERGIC53, a lectin that binds high mannose glycans) (8). Glycoproteins remaining in misfolded conformations are retrotranslocated to the cytosol where they are deglycosylated and degraded by the proteasome (9). One or more mannose residues may be removed during the whole folding process.

The presence of monoglucosylated glycans triggers recognition by CRT and CNX, an interaction that retains glycoprotein species bearing those markers in the ER. Association with CRT and/or CNX plays a vital role during the conformational maturation of many glycoproteins. Preventing the association usually diminishes the folding efficiency although under such situations the folding rate of most glycoproteins is increased. CRT and CNX are thought to improve the folding efficiency by preventing nonspecific interactions between folding species and, in addition, they enhance the correct pairing of disulfide bridges by promoting ERp57 intervention.

8.4 GLUCOSYLTRANSFERASE

At variance with other ER glucosyltransferases that employ dolichol-P-Glc as glucosyl donor, UGGT uses UDP-Glc, which is presumably antitransported by UMP into the ER lumen [91]. UGGT activity on high-mannose glycoproteins generates a glycan structure identical to that obtained from partial deglucosylation of $Glc_3Man_9GlcNAc_2$ by GI and GII. The activity of the enzyme is maximal with $Man_9GlcNAc_2$-bearing glycoproteins and diminishes with mannose trimming (the relative activity toward $Man_9GlcNAc_2$, $Man_8GlcNAc_2$ and $Man_7GlcNAc_2$, displaying residue *g*, that is, the mannose unit to which the glucose is attached is 1.00, 0.50, and 0.15, respectively; see Figure 8.1) [92], thus showing a trend similar to that observed for GII activity [93]. UGGT activity was first described in *Trypanosoma cruzi* cells and it was later detected in most eukaryotic cells, with the notable exception of *S. cerevisiae* [94,95]. UGGT is a rather large protein with a molecular mass of 170 kDa, endowed with an N-terminal signal peptide and a C-terminal ER retention/retrieval signal whose sequence varies among species (HEEL and HGEL in rat and drosophila, respectively; interestingly, the enzyme from *T. cruzi* lacks such signal) [96–98]. There are two genes coding for UGGTs in humans, HUGT1 and HUGT2, which are 55% identical. Although both genes are expressed, only the first one is active, the biological function of HUGT2 being unknown [99].

UGGT presents a highly conserved region comprising approximately 20% of the C-terminal region and poor sequence conservation in the rest of the sequence. Mild protease treatment of UGGT yields two fragments that remain tightly bound by hydrophobic interactions and display enzymatic activity. The fragments coincide with the conserved and variable regions of the sequence, suggesting the presence of at least two domains [100]. The conserved C-terminal region displays a similar size and significant similarity to members of glycosyltransferase family 8, it binds UDP-Glc [96], and it has been proposed to bear the catalytic center. Hybrid enzymes made by combining the C-terminal and N-terminal regions from different species are functional, even though in the case of *S. pombe* and *Drosophila*, the N-terminal domains (80% of the molecules) only share a 16% similarity. Moreover, both fragments can complement each other when expressed in *trans* [100]. The most prominent feature of UGGT is that in *in vitro* assays it glucosylates high-mannose glycoproteins preferentially when displaying perturbed conformations [101]. This constitutes a unique example of an enzyme that presents features both of a chaperone and a glycosyltransferase. The variable N-terminal domain is probably responsible for recognition of the misfolded polypeptide moiety [100,102]. For optimal activity, the glycan must be covalently attached to a misfolded protein moiety, since short glycopeptides derived from good glycoprotein acceptors such as bovine thyroglobulin are poor substrates [92]. Nevertheless, the enzyme is active against glycopeptides longer than 12 amino acid residues provided they display hydrophobic residues at either side of the N-glycan [103]. In addition, proteins chemically modified with a high-mannose glycan are recognized by UGGT [104]. Since UGGT recognizes a heterogeneous protein population, the molecular determinant for this process should be common to folding intermediates of all glycoproteins. It is for this reason that the misfolded substrate-binding domain in UGGTs from different species might not necessarily be expected to display significant similarity in their primary sequences. The enzyme recognizes exposed hydrophobic residues displayed in glycoprotein substrates preferentially when they are forming patches in molten globule-like conformers [105,106]. This feature allows recognition of subtle defects in acceptor glycoproteins [107]. Due to this selectivity, UGGT preferentially reglucosylates advanced folding intermediates rather than highly misfolded, random coiled species [108]. The enzyme thus focuses its biological function in the last stages of folding maturation in the ER [109]. Chaperones more fitted to deal with extended folding intermediates such as BiP would be the main partners during the initial stages of protein folding in the ER. The selectivity of UGGT enables it to participate also in the quality control of quaternary structures, since it can recognize folded complex subunits lacking the full complement of oligomer components provided that the orphan subunits expose a hydrophobic patch hindered upon oligomer assembly [110]. The distance between a local folding defect and the targeted glycan may vary depending on the particular characteristic of the

substrate [107,111]. For instance, only the N-glycan attached to the misfolded subunit of an artificial dimer formed by properly folded and misfolded RNase B monomers was glucosylated *in vitro* by UGGT [112].

Folding glycoproteins in the ER cycle between monoglucosylated and nonglucosylated states catalyzed by the opposite activities of GII and UGGT. The rate of conversion between the states is a matter of debate and may depend on the particular glycoprotein involved. Regarding both models proposed for glycoprotein binding to the lectins, in the "lectin-only" model UGGT would be the only conformational sensor, while in the "dual binding" model, both UGGT and the lectins would be involved in the recognition of misfolded species.

8.5 CALCIUM BINDING

In noncontractile cells, the ER is also the main intracellular calcium store. Calcium concentration in the ER fluctuates from 1 to 5 mM when it is full to 1–50 μM upon opening of the membrane calcium channels [113]. Most of the calcium is not free in the ER lumen, but it is bound instead to proteins that operate as calcium buffers by binding high amounts of calcium ions with low affinity [114]. The effect of ER calcium fluctuations on protein folding is variable. While in some cases, a particular protein may intrinsically need calcium to attain its native conformation, a more general effect arises from the calcium dependence of several ER chaperones and folding assisting enzymes as UGGT. Besides its role in QC, CRT is one of the main calcium buffers in the ER, accounting for about 45% of the total ion present in that subcellular location. Accordingly, CRT overexpression increases the ER calcium storage capacity [115–117] and as expected, the C-terminal domain is responsible for this effect [118,119]. Conversely, cell lines lacking CRT present lower calcium content in the ER, a condition that may be reversed by expressing the P and C domains [120]. Transgenic mice lacking CRT die due to compromised cardiac development [57] although cell lines isolated from them are viable. These cells show a diminished ER calcium storage capacity and $InsP_3$ receptor–mediated calcium release. Remarkably, overexpression of constitutively active calcineurin in the hearts of CRT-deficient mice reverses the embryonic lethality and produces viable CRT-deficient animals, demonstrating that CRT is an upstream regulator of calcineurin in the calcium signaling cascade [121,122].

The effect of the ER calcium level in the lectin–chaperone properties of CRT and CNX is unclear. CRT and CNX need calcium for their lectin activity when they are immobilized on a chelating matrix [19], and calcium increases the stability of CRT, making it more rigid and less susceptible to proteases [123,124]. Conversely, as mentioned above, calcium depletion induces the oligomerization of CRT and CNX and increases their polypeptide-binding capacity [79,125]. On the other hand, the interaction of CRT with other folding facilitating enzymes can be modulated by calcium, as has been shown *in vitro* for PDI and ERp57 [126]. Clearly, the ER environment may play an important role in regulating the biological activities of the lectins, although we still lack a complete picture of the situation.

8.6 CONCLUSION AND FUTURE DIRECTIONS

Conventional chaperone cycles of substrate binding and release is driven from prokaryotes to eukaryotes by the hydrolysis of high-energy molecules, usually ATP. A novel, unconventional chaperone system evolved in eukaryotes in which the energy in provided by a nucleotide sugar (UDP-Glc). Whereas in the first system a single macromolecular component may be responsible for the chaperone activity, the glycoprotein folding quality control system involves several proteins, UGGT, CNX/CRT, and GII. In this scenario, substrate recognition is a function of a glycosyltransferase, UGGT. Lectins (CNX/CRT) are responsible for substrate binding and finally substrate release is mediated by a glycosidase, GII. Although protein N-glycosylation may be found in eubacteria, it is

only in eukaryotes that the unconventional, N-glycan-dependent chaperone system occurs. One reason for this late evolutionary appearance is probably the fact that whereas glycans are added to unfolded proteins in eukaryotes, in eubacteria, the oligosaccharyltransferase glycosylates already properly folded species [127].

REFERENCES

1. Hammond, C., Braakman, I., and Helenius, A., Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control, *Proc. Natl Acad. Sci. USA*, 91, 913, 1994.
2. Waisman, D.M., Salimath, B.P., and Anderson, M.J., Isolation and characterization of CAB-63, a novel calcium-binding protein, *J. Biol. Chem*., 260, 1652, 1985.
3. Fliegel, L. et al., Molecular cloning of the high affinity calcium-binding protein (calreticulin) of skeletal muscle sarcoplasmic reticulum, *J. Biol. Chem*., 264, 21522, 1989.
4. Smith, M.J. and Koch, G.L., Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein, *EMBO J*., 8, 3581, 1989.
5. Wada, I. et al., SSR alpha and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane, *J. Biol. Chem*., 266, 19599, 1991.
6. Ware, F.E. et al, The molecular chaperone calnexin binds $Glc_1Man_9GlcNAc_2$ oligosaccharide as an initial step in recognizing unfolded glycoproteins, *J. Biol. Chem*., 270, 4697, 1995.
7. Spiro, R.G., Zhu, Q., Bhoyroo, V., and Soling, H.D., Definition of the lectin-like properties of the molecular chaperone, calreticulin, and demonstration of its copurification with endomannosidase from rat liver Golgi. *J. Biol. Chem*., 271, 11588, 1996.
8. Conway, E.M. et al., Heat shock-sensitive expression of calreticulin. In vitro and in vivo up-regulation, *J. Biol. Chem*., 270, 17011, 1995.
9. Szewczenko-Pawlikowski, M. et al., Heat shock-regulated expression of calreticulin in retinal pigment epithelium, *Mol. Cell. Biochem*., 177, 145, 1997.
10. Heal, R. and McGivan, J., Induction of calreticulin expression in response to amino acid deprivation in Chinese hamster ovary cells, *Biochem. J*., 329, 389, 1998.
11. Nguyen, T.O., Capra, J.D., and Sontheimer, R.D., Calreticulin is transcriptionally upregulated by heat shock, calcium and heavy metals, *Mol. Immunol*., 33, 379, 1996.
12. Llewellyn, D.H. et al., Induction of calreticulin expression in HeLa cells by depletion of the endoplasmic reticulum Ca^{2+} store and inhibition of N-linked glycosylation, *Biochem. J*., 318, 555, 1996.
13. McCauliffe, D.P. et al., The 5′-flanking region of the human calreticulin gene shares homology with the human GRP78, GRP94, and protein disulfide isomerase promoters, *J. Biol. Chem*., 267, 2557, 1992.
14. Schrag, J.D. et al., The structure of calnexin, an ER chaperone involved in quality control of protein folding, *Mol. Cell*, 8, 633, 2001.
15. Baksh, S. and Michalak, M., Expression of calreticulin in *Escherichia coli* and identification of its Ca^{2+} binding domains, *J. Biol. Chem*., 266, 21458, 1991.
16. Ellgaard, L. et al., NMR structure of the calreticulin P-domain, *Proc. Natl Acad. Sci. USA* 98, 3133, 2001.
17. Tan, Y. et al., The calcium- and zinc-responsive regions of calreticulin reside strictly in the N-/C-domain, *Biochim. Biophys. Acta*, 1760, 745, 2006.
18. Xu, X., Azakami, H., and Kato, A., P-domain and lectin site are involved in the chaperone function of *Saccharomyces cerevisiae* calnexin homologue, *FEBS Lett*., 570, 155, 2004.
19. Vassilakos, A. et al., Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin, *Biochemistry*, 37, 3480, 1998.
20. Leach, M.R. et al., Localization of the lectin, ERp57 binding, and polypeptide binding sites of calnexin and calreticulin, *J. Biol. Chem*., 277, 29686, 2002.
21. Ellgaard, L. et al., NMR structures of 36 and 73-residue fragments of the calreticulin P-domain, *J. Mol. Biol*., 322, 773, 2002.
22. Peterson, J.R. and Helenius, A., In vitro reconstitution of calreticulin–substrate interactions, *J. Cell Sci*., 112, 2775, 1999.
23. Bouvier, M. and Stafford, W.F., Probing the three-dimensional structure of human calreticulin, *Biochemistry*, 39, 14950, 2000.
24. Sonnichsen, B. et al., Retention and retrieval: Both mechanisms cooperate to maintain calreticulin in the endoplasmic reticulum, *J. Cell Sci*., 107, 2705, 1994.

25. Oliver, J.D., et al., ERp57 functions as a subunit of specific complexes formed with the ER lectins calreticulin and calnexin, *Mol. Biol. Cell*., 10, 2573, 1999.
26. Oliver, J.D. et al., Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins, *Science*, 275, 86, 1997.
27. Silvennoinen, L. et al., Identification and characterization of structural domains of human ERp57: Association with calreticulin requires several domains, *J. Biol. Chem*., 279, 13607, 2004.
28. Frickel, E.M. et al., ERp57 is a multifunctional thiol–disulfide oxidoreductase, *J. Biol. Chem*., 279, 18277, 2004.
29. Di Jeso, B. et al., Mixed-disulfide folding intermediates between thyroglobulin and endoplasmic reticulum resident oxidoreductases ERp57 and protein disulfide isomerase, *Mol. Cell. Biol*., 25, 9793, 2005.
30. Van der Wal, F.J., Oliver, J.D., and High, S., The transient association of ERp57 with N-glycosylated proteins is regulated by glucose trimming, *Eur. J. Biochem*., 256, 51, 1998.
31. Morrice, N.A. and Powis, S.J., A role for the thiol-dependent reductase ERp57 in the assembly of MHC class I molecules, *Curr. Biol*., 8, 713, 1998.
32. Kang, S.J. and Cresswell, P., Calnexin, calreticulin, and ERp57 cooperate in disulfide bond formation in human CD1d heavy chain, *J. Biol. Chem*., 277, 44838, 2002.
33. Zapun, A. et al., Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57, *J. Biol. Chem*., 273, 6009, 1998.
34. Mayer, M. et al, Influence of the oxidoreductase ER57 on the folding of an antibody fab fragment, *J. Mol. Biol*., 341, 1077, 2004.
35. Peaper, D.R., Wearsch, P.A., and Cresswell, P., Tapasin and ERp57 form a stable disulfide-linked dimer within the MHC class I peptide-loading complex, *EMBO J*., 24, 3613, 2005.
36. McCormick, L.M. et al., Independent and cooperative roles of N-glycans and molecular chaperones in the folding and disulfide bond formation of the low-density lipoprotein (LDL) receptor-related protein, *Biochemistry*, 44, 5794, 2005.
37. Frickel, E.M. et al., TROSY-NMR reveals interaction between ERp57 and the tip of the calreticulin P-domain, *Proc. Natl Acad. Sci. USA*, 99, 1954, 2002.
38. Russell, S.J. et al., The primary substrate binding site in the b′ domain of ERp57 is adapted for endoplasmic reticulum lectin association, *J. Biol. Chem*., 279, 18861, 2004.
39. Kozlov, G. et al., Crystal structure of the bb′ domains of the protein disulfide isomerase ERp57, *Structure*, 14, 1331, 2006.
40. Pollock, S. et al., Specific interaction of ERp57 and calnexin determined by NMR spectroscopy and an ER two-hybrid system, *EMBO J*., 23, 1020, 2004.
41. Ware, F.E. et al., The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins, *J. Biol. Chem*., 270, 4697, 1995.
42. Kapoor, M. et al., Interactions of substrate with calreticulin, an endoplasmic reticulum chaperone, *J. Biol. Chem*., 278, 6194, 2003.
43. Kapoor, M. et al., Mutational analysis provides molecular insight into the carbohydrate-binding region of calreticulin: Pivotal roles of tyrosine-109 and aspartate-135 in carbohydrate recognition, *Biochemistry*, 43, 97, 2004.
44. Patil, A.R., Thomas, C.J., and Surolia, A., Kinetics and the mechanism of interaction of the endoplasmic reticulum chaperone, calreticulin, with monoglucosylated (Glc1Man9GlcNAc2) substrate, *J. Biol. Chem*., 275, 24348, 2000.
45. Thomson, S.P. and Williams, D.B., Delineation of the lectin site of the molecular chaperone calreticulin, *Cell Stress Chaperones*, 10, 242, 2005.
46. Helenius, A. et al., Calnexin, calreticulin and the folding of glycoproteins, *Trends Cell Biol*., 7, 193, 1997.
47. Andersson, H., Nilsson, I., and von Heijne, G., Calnexin can interact with N-linked glycans located close to the endoplasmic reticulum membrane, *FEBS Lett*., 397, 321, 1996.
48. Hebert, D.N. et al., The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin, *J. Cell Biol*., 139, 613, 1997.
49. Wang, N., Daniels, R., and Hebert, D.N., The cotranslational maturation of the type I membrane glycoprotein tyrosinase: The heat shock protein 70 system hands off to the lectin-based chaperone system, *Mol. Biol. Cell*, 16, 3740, 2005.
50. Daniels, R. et al., N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin, *Mol. Cell*, 11, 79, 2003.
51. Danilczyk, U.G., Cohen-Doyle, M.F., and Williams, D.B., Functional relationship between calreticulin, calnexin, and the endoplasmic reticulum luminal domain of calnexin, *J. Biol. Chem*., 275, 13089, 2000.

52. Gao, B. et al., Assembly and antigen-presenting function of MHC class I molecules in cells lacking the ER chaperone calreticulin, *Immunity*, 16, 99, 2002.
53. Pieren, M. et al., The use of calnexin and calreticulin by cellular and viral glycoproteins, *J. Biol. Chem*., 280, 28265, 2005.
54. Gruters, R.A. et al., Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase, *Nature*, 330, 74, 1987.
55. Fischer, P.B. et al., *N*-butyldeoxynojirimycin-mediated inhibition of human immunodeficiency virus entry correlates with changes in antibody recognition of the V1/V2 region of gp120, *J. Virol*., 70, 7143, 1996.
56. Mehta, A. et al., Hepatitis B virus (HBV) envelope glycoproteins vary drastically in their sensitivity to glycan processing: Evidence that alteration of a single N-linked glycosylation site can regulate HBV secretion, *Proc. Natl Acad. Sci. USA*, 94, 1822, 1997.
57. Mesaeli, N. et al., Calreticulin is essential for cardiac development, *J. Cell Biol*., 144, 857, 1999.
58. Denzel, A. et al., Early postnatal death and motor disorders in mice congenitally deficient in calnexin expression, *Mol. Cell. Biol*., 22, 7398, 2002.
59. Williams, D.B., Beyond lectins: The calnexin/calreticulin chaperone system of the endoplasmic reticulum, *J. Cell Sci*., 119, 615, 2006.
60. Nakhasi, H.L. et al., Rubella virus glycoprotein interaction with the endoplasmic reticulum calreticulin and calnexin, *Arch. Virol*., 146, 1, 2001.
61. Branza-Nichita, N. et al., Antiviral effect of *N*-butyldeoxynojirimycin against bovine viral diarrhea virus correlates with misfolding of E2 envelope proteins and impairment of their association into E1-E2 heterodimers, *J. Virol*., 75, 3527, 2001.
62. Tatu, U. and Helenius, A., Interaction of newly synthesized apolipoprotein B with calnexin and calreticulin requires glucose trimming in the endoplasmic reticulum, *Biosci. Rep*., 19, 189, 1999.
63. Keller, S.H., Lindstrom, J., and Taylor, P., Inhibition of glucose trimming with castanospermine reduces calnexin association and promotes proteasome degradation of the alpha-subunit of the nicotinic acetylcholine receptor, *J. Biol. Chem*., 273, 17064, 1998.
64. Zhang, L. et al., Calreticulin promotes folding/dimerization of human lipoprotein lipase expressed in insect cells (sf21), *J. Biol. Chem*., 278, 29344, 2003.
65. Di Jeso, B. et al., Folding of thyroglobulin in the calnexin/calreticulin pathway and its alteration by loss of Ca^{2+} from the endoplasmic reticulum, *Biochem. J*., 370, 449, 2003.
66. Ora, A. and Helenius, A., Calnexin fails to associate with substrate proteins in glucosidase-deficient cell lines, *J. Biol. Chem*., 270, 26060, 1995.
67. Hebert, D.N., Foellmer, B., and Helenius, A., Glucose trimming and reglucosylation determine glycoprotein association with calnexin in the endoplasmic reticulum, *Cell*, 81, 425, 1995.
68. Danilczyk, U.G. and Williams, D.B., The lectin chaperone calnexin utilizes polypeptide-based interactions to associate with many of its substrates in vivo, *J. Biol. Chem*., 276, 25532, 2001.
69. Mizrachi, D. and Segaloff, D.L., Intracellularly located misfolded glycoprotein hormone receptors associate with different chaperone proteins than their cognate wild-type receptors, *Mol. Endocrinol*., 18, 1768, 2004.
70. Wanamaker, C.P. and Green, W.N., N-linked glycosylation is required for nicotinic receptor assembly but not for subunit associations with calnexin, *J. Biol. Chem*., 280, 33800, 2005.
71. Saito, Y. et al., Calreticulin functions *in vitro* as a molecular chaperone for both glycosylated and non-glycosylated proteins, *EMBO J*., 18, 6718, 1999.
72. Ihara, Y. et al., Calnexin discriminates between protein conformational states and functions as a molecular chaperone in vitro, *Mol. Cell*, 4, 331, 1999.
73. Stronge, V.S. et al., Relationship between calnexin and BiP in suppressing aggregation and promoting refolding of protein and glycoprotein substrates, *J. Biol. Chem*., 276, 39779, 2001.
74. Leach, M.R. and Williams, D.B., Lectin-deficient calnexin is capable of binding class I histocompatibility molecules in vivo and preventing their degradation, *J. Biol. Chem*., 279, 9072, 2004.
75. Fontanini, A. et al., Glycan-independent role of calnexin in the intracellular retention of Charcot-Marie-tooth 1A Gas3/PMP22 mutants, *J. Biol. Chem*., 280, 2378, 2005.
76. Swanton, E., High, S., and Woodman, P., Role of calnexin in the glycan-independent quality control of proteolipid protein, *EMBO J*., 22, 2948, 2003.
77. Oda, Y. et al., EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin, *Science*, 299, 1394, 2003.
78. Thammavongsa, V., Mancino, L., and Raghavan, M., Polypeptide substrate recognition by calnexin requires specific conformations of the calnexin protein, *J. Biol. Chem*., 280, 33497, 2005.

79. Rizvi, S.M. et al., A polypeptide binding conformation of calreticulin is induced by heat shock, calcium depletion, or by deletion of the C-terminal acidic region, *Mol. Cell*, 15, 913, 2004.
80. Zapun, A. et al., Conformation-independent binding of monoglucosylated ribonuclease B to calnexin, *Cell*, 88, 29, 1997.
81. Rodan, A.R. et al., N-linked oligosaccharides are necessary and sufficient for association of glycosylated forms of bovine RNase with calnexin and calreticulin, *EMBO J.*, 15, 6921, 1996.
82. Wearsch, P.A. et al., Major histocompatibility complex class I molecules expressed with monoglucosylated N-linked glycans bind calreticulin independently of their assembly status, *J. Biol. Chem.*, 279, 25112, 2004.
83. Zhang, Q., Tector, M., and Salter, R.D., Calnexin recognizes carbohydrate and protein determinants of class I major histocompatibility complex molecules, *J. Biol. Chem.*, 270, 3944, 1995.
84. Parodi, A.J., N-glycosylation in trypanosomatid protozoa, *Glycobiology*, 3, 193, 1993.
85. Parodi, A.J. and Cazzulo, J.J., Protein glycosylation in *Trypanosoma cruzi*. II. Partial characterization of protein-bound oligosaccharides labeled "in vivo," *J. Biol. Chem.*, 257, 7641, 1982.
86. Kalz-Fuller, B., Bieberich, E., and Bause, E., Cloning and expression of glucosidase I from human hippocampus, *Eur. J. Biochem.*, 231, 344, 1995.
87. Romero, P.A. et al., The yeast CWH41 gene encodes glucosidase I, *Glycobiology*, 7, 997, 1997.
88. Simons, J.F., Ebersold, M., and Helenius, A., Cell wall 1,6-beta-glucan synthesis in *Saccharomyces cerevisiae* depends on ER glucosidases I and II, and the molecular chaperone BiP/Kar2p, *EMBO J.*, 17, 396, 1998.
89. D'Alessio, C. et al., Genetic evidence for the heterodimeric structure of glucosidase II. The effect of disrupting the subunit-encoding genes on glycoprotein folding, *J. Biol. Chem.*, 274, 25899, 1999.
90. Deprez, P., Gautschi, M., and Helenius, A., More than one glycan is needed for ER glucosidase II to allow entry of glycoproteins into the calnexin/calreticulin cycle, *Mol. Cell*, 19, 183, 2005.
91. Castro, O. et al., Uridine diphosphate-glucose transport into the endoplasmic reticulum of *Saccharomyces cerevisiae*: In vivo and in vitro evidence, *Mol. Biol. Cell*, 10, 1019, 1999.
92. Sousa, M.C., Ferrero-Garcia, M.A., and Parodi, A.J., Recognition of the oligosaccharide and protein moieties of glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase, *Biochemistry*, 31, 97, 1992.
93. Grinna, L.S. and Robbins, P.W., Substrate specificities of rat liver microsomal glucosidases which process glycoproteins, *J. Biol. Chem.*, 255, 2255, 1980.
94. Trombetta, S.E., Bosch, M., and Parodi, A.J., Glucosylation of glycoproteins by mammalian, plant, fungal, and trypanosomatid protozoa microsomal membranes, *Biochemistry*, 28, 8108, 1989.
95. Fernandez, F.S. et al., Purification to homogeneity of UDP-glucose:glycoprotein glucosyltransferase from *Schizosaccharomyces pombe* and apparent absence of the enzyme fro *Saccharomyces cerevisiae*, *J. Biol. Chem.*, 269, 30701, 1994.
96. Tessier, D.C. et al., Cloning and characterization of mammalian UDP-glucose glycoprotein: Glucosyltransferase and the development of a specific substrate for this enzyme, *Glycobiology*, 10, 403, 2000.
97. Parker, C.G. et al., Drosophila UDP-glucose:glycoprotein glucosyltransferase: Sequence and characterization of an enzyme that distinguishes between denatured and native proteins, *EMBO J.*, 14, 1294, 1995.
98. Conte, I. et al., The interplay between folding-facilitating mechanisms in *Trypanosoma cruzi* endoplasmic reticulum, *Mol. Biol. Cell*, 14, 3529, 2003.
99. Arnold, S.M. et al., Two homologues encoding human UDP-glucose:glycoprotein glucosyltransferase differ in mRNA expression and enzymatic activity, *Biochemistry*, 39, 2149, 2000.
100. Guerin, M. and Parodi, A.J., The UDP-glucose:glycoprotein glucosyltransferase is organized in at least two tightly bound domains from yeast to mammals, *J. Biol. Chem.*, 278, 20540, 2003.
101. Trombetta, S.E. and Parodi, A.J., Purification to apparent homogeneity and partial characterization of rat liver UDP-glucose:glycoprotein glucosyltransferase, *J. Biol. Chem.*, 267, 9236, 1992.
102. Arnold, S.M. and Kaufman, R.J., The noncatalytic portion of human UDP-glucose: Glycoprotein glucosyltransferase I confers UDP-glucose binding and transferase function to the catalytic domain, *J. Biol. Chem.*, 278, 43320, 2003.
103. Taylor, S.C. et al., Glycopeptide specificity of the secretory protein folding sensor UDP-glucose glycop rotein:glucosyltransferase, *EMBO Rep.*, 4, 405, 2003.
104. Sousa, M. and Parodi, A.J., The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase, *EMBO J.*, 14, 4196, 1995.
105. Caramelo, J.J. et al., The endoplasmic reticulum glucosyltransferase recognizes nearly native glycoprotein folding intermediates, *J. Biol. Chem.*, 279, 46280, 2004.

106. Caramelo, J.J. et al., UDP-Glc:glycoprotein glucosyltransferase recognizes structured and solvent accessible hydrophobic patches in molten globule-like folding intermediates, *Proc. Natl Acad. Sci. USA*, 100, 86, 2003.
107. Ritter, C. et al., Minor folding defects trigger local modification of glycoproteins by the ER folding sensor GT, *EMBO J.*, 24, 1730, 2005.
108. Trombetta, E.S. and Helenius, A., Conformational requirements for glycoprotein reglucosylation in the endoplasmic reticulum, *J. Cell Biol.*, 148, 1123, 2000.
109. Labriola, C., Cazzulo, J.J., and Parodi, A.J., *Trypanosoma cruzi* calreticulin is a lectin that binds monoglucosylated oligosaccharides but not protein moieties of glycoproteins, *Mol. Biol. Cell*, 10, 1381, 1999.
110. Keith, N., Parodi, A.J., and Caramelo, J.J., Glycoprotein tertiary and quaternary structures are monitored by the same quality control mechanism, *J. Biol. Chem.*, 280, 18138, 2005.
111. Taylor, S.C. et al., The ER protein folding sensor UDP-glucose glycoprotein–glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation, *Nat. Struct. Mol. Biol.*, 11, 128, 2004.
112. Ritter, C. and Helenius, A., Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase, *Nat. Struct. Biol.*, 7, 278, 2000.
113. Llopis, J. et al., Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin, *Nature*, 388, 882, 1997.
114. Brostrom, M.A. and Brostrom, C.O., Calcium dynamics and endoplasmic reticular function in the regulation of protein synthesis: Implications for cell growth and adaptability, *Cell Calcium*, 34, 345, 2003.
115. Mery, L. et al., Overexpression of calreticulin increases intracellular Ca^{2+} storage and decreases store-operated Ca^{2+} influx, *J. Biol. Chem.*, 271, 9332, 1996.
116. Bastianutto, C. et al., Overexpression of calreticulin increases the Ca^{2+} capacity of rapidly exchanging Ca^{2+} stores and reveals aspects of their lumenal microenvironment and function, *J. Cell Biol.*, 130, 847, 1995.
117. Persson, S. et al., The Ca^{2+} status of the endoplasmic reticulum is altered by induction of calreticulin expression in transgenic plants, *Plant Physiol.*, 126, 1092, 2001.
118. Xu, W. et al., Calreticulin modulates capacitative Ca^{2+} influx by controlling the extent of inositol 1,4,5-trisphosphate-induced Ca^{2+} store depletion, *J. Biol. Chem.*, 275, 36676, 2000.
119. Wyatt, S.E., Tsou, P.L., and Robertson, D., Expression of the high capacity calcium-binding domain of calreticulin increases bioavailable calcium stores in plants, *Transgenic Res.*, 11, 1, 2002.
120. Nakamura, K. et al., Functional specialization of calreticulin domains, *J. Cell Biol.*, 154, 961, 2001.
121. Lynch, J., and Michalak, M., Calreticulin is an upstream regulator of calcineurin, *Biochem. Biophys. Res. Commun.*, 311, 1173, 2003.
122. Guo, L. et al., Cardiac-specific expression of calcineurin reverses embryonic lethality in calreticulin-deficient mouse, *J. Biol. Chem.*, 277, 50776, 2002.
123. Li, Z., Stafford, W.F., and Bouvier, M., The metal ion binding properties of calreticulin modulate its conformational flexibility and thermal stability, *Biochemistry*, 40, 11193, 2001.
124. Corbett, E.F. et al., The conformation of calreticulin is influenced by the endoplasmic reticulum luminal environment, *J. Biol. Chem.*, 275, 27177, 2000.
125. Thammavongsa, V., Mancino, L., and Raghavan, M., Polypeptide substrate recognition by calnexin requires specific conformations of the calnexin protein, *J. Biol. Chem.*, 280, 33497, 2005.
126. Corbett, E.F. et al., Ca^{2+} regulation of interactions between endoplasmic reticulum chaperones, *J. Biol. Chem.*, 274, 6203, 1999.
127. Kowarik, M. et al., N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase, *Science*, 314, 1148, 2006.

9 Role of L-Type Lectins in Glycoprotein Sorting and Trafficking

Beat Nyfeler, Markus W. Wendeler, and Hans-Peter Hauri

CONTENTS

9.1 INTRODUCTION

One-third of all newly synthesized proteins are cotranslationally inserted into the endoplasmic reticulum (ER) where they are folded, modified, and subjected to quality control prior to secretion or transport to various intracellular organelles. During translocation, a majority of these proteins are N-glycosylated by the addition of a 14-saccharide core glycan at conserved Asn-X-Ser/Thr consensus sequences [1]. N-linked glycans are important determinants of protein recognition and modification involving intracellular lectins at various steps along the secretory pathway. The first lectins encountered by a newly synthesized glycoprotein are calnexin and calreticulin, two ER-localized proteins that support protein folding and quality control [2]. Correctly folded glycoproteins can then be captured by members of an additional class of animal lectins termed L-type lectins [3]. These lectins are characterized by a luminal carbohydrate recognition domain (CRD) that corresponds to the single-folded domain of soluble lectins found in the seeds of leguminous plants [4,5].

Unlike leguminous plant lectins, which are soluble secretory proteins, animal L-type lectins are membrane proteins and localize to the early secretory pathway. The early secretory pathway of animal cells is composed of the ER, the ER Golgi intermediate compartment (ERGIC) and the *cis*-Golgi [6], which are interconnected by vesicular transport pathways [7]. In the anterograde pathway, coat protein II (COPII)-coated vesicles bud from the ER and fuse with ERGIC membranes. In the retrograde pathway, coat protein I (COPI)-coated vesicles bud from *cis*-Golgi and ERGIC and retrieve proteins back to the ER.

How can proteins be sorted into the different transport vesicles? Membrane proteins are packaged into budding vesicles by selective interaction of their cytosolic tails with subunits of the COPI and/or

COPII coats. The interaction with COPII is mediated by diacidic [8] or dihydrophobic [9,10] ER-export motifs that bind to the Sec24 coat subunit [11]. The interaction with COPI is mediated by a dilysine signal that binds the cytosolic coatomer complex [12,13]. Soluble luminal proteins cannot directly interact with cytosolic coat subunits due to the membrane barrier. Although some soluble proteins might be incorporated passively into budding vesicles by bulk flow [14], efficient transport clearly requires active, receptor-mediated recruitment into coated vesicles [15]. One example for a nonlectin membrane cargo receptor is the KDEL (Lys-Asp-Glu-Leu)-receptor that retrieves soluble proteins back to the ER by recruiting them into COPI-coated vesicles in the ERGIC and Golgi [16,17]. With its luminal domain, the KDEL-receptor captures proteins carrying a C-terminal tetrapeptide transport signal while the cytosolic tail of the receptor interacts with COPI via a dilysine-related signal. Similarly, animal L-type lectins bind with their luminal CRD to glycoproteins while their cytosolic tail interacts with cytosolic coat subunits. With one exception, all animal L-type lectins bear ER export and ER retention/retrieval signals that mediate incorporation into COPII- and/or COPI-coated vesicles. Many properties of the animal L-type lectin family suggest a role in the transport of glycoproteins. In this chapter, we summarize the current knowledge of L-type lectins and discuss their known and presumed functions in the secretory pathway of eukaryotic cells.

9.2 FEATURES OF ANIMAL L-TYPE LECTINS

Animal cells express four different L-type lectins known as ERGIC-53 [18,19] (ER Golgi intermediate compartment protein of 53 kDa), VIP36 [5] (vesicular integral membrane protein of 36 kDa), ERGL [20] (ERGIC-53-like protein), and VIPL [21,22] (VIP36-like protein). These family members share a common luminal L-type lectin domain but differ with regard to the presence of a coiled-coil domain, an N-glycosylation site, and cytosolic sorting signals, as depicted in Figure 9.1 and summarized in Table 9.1.

9.2.1 ERGIC-53

ERGIC-53 was the first L-type lectin discovered and is currently the best characterized family member. It was identified as a marker protein for the ERGIC in a monoclonal antibody screen for organelle markers [18]. Independently, p58 was identified by a polyclonal antibody approach and subsequently found to be the rat homolog of ERGIC-53 [23,24]. ERGIC-53 is a ubiquitously expressed type-I transmembrane protein with a large luminal domain (447 amino acids) and a short cytosolic domain (12 amino acids). The protein continuously cycles between ERGIC and ER, but its concentration is highest in the ERGIC at steady state [25,26]. Upon treatment of cells with the fungal metabolite brefeldin A (BFA), COPI dissociates from Golgi membranes and ERGIC-53 accumulates in ERGIC structures [27]. This BFA-dependent redistribution turned out to be typical of rapidly cycling proteins of the early secretory pathway, very much in contrast to Golgi proteins, most of which redistribute to the ER in response to BFA. The cycling properties of ERGIC-53 are determined by different sorting motifs in the cytosolic tail of the protein. The C-terminal diphenylalanine motif operates as an ER export signal by interaction with the Sec24 subunit of the COPII [10,28,29]. The dilysine signal close to the diphenylalanine motif binds COPI and thereby mediates retrieval of ERGIC-53 back to the ER [30]. ERGIC-53 forms disulfide-linked homodimers and homohexamers [28,31]. Oligomerization is supported by a luminal coiled-coil domain (Figure 9.1) and is required for the optimal presentation of the diphenylalanine motif to Sec24 [28].

The N-terminal part of ERGIC-53 contains a ~200-residue luminal domain that shares sequence homology with the CRD of leguminous plant lectins [4]. The crystal structure of ERGIC-53 reveals that the protein fold including the ligand binding site is indeed highly similar to that of leguminous plant lectins [32]. The structure shows an overall β-sandwich composed of one concave and one convex β-sheet with the ligand-binding site in a negatively charged cleft. The CRD of ERGIC-53

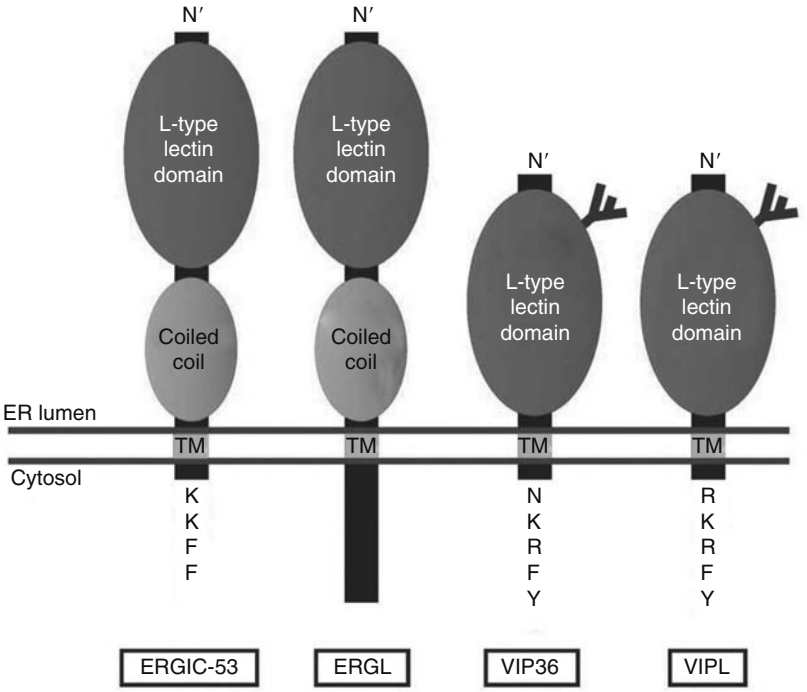


FIGURE 9.1 (See color insert following blank page 170. Also see CD for color figure.) Schematic representation of animal L-type lectins. The domain organization and the topology of the four animal L-type lectins are depicted: luminal N-terminus (N′), luminal L-type lectin domain (red), luminal coiled-coil domain (green), and N-linked glycans (gray). Potential ER export motifs in the cytoplasmic tails that interact with COPII are indicated in orange (single letter code of amino acids). ER retention/retrieval motifs that interact with COPI are indicated in blue. Note that the KR motif does not entirely conform to the consensus of an ER retention/retrieval motif and has not been proven to be a functional transport signal. (From Teasdale, R.D. and Jackson, M.R., *Annu. Rev. Cell Dev. Biol.*, 12, 27, 1996.)

binds mannose residues in a calcium-dependent way and requires the presence of a highly conserved asparagine residue in the lectin domain of ERGIC-53 [4]. The mannose specificity together with the ER-ERGIC recycling properties led to the notion that ERGIC-53 may act as a sorting receptor for glycoproteins [4]. This notion was substantiated when ERGIC-53 was found to be required for efficient transport of cathepsin C [33] and genetically linked to the inherited disease combined blood coagulation factor V and VIII deficiency (F5F8D; OMIM #227′300) [34]. Patients suffering from F5F8D have reduced plasma levels of coagulation factors V and VIII. This is caused by inefficient secretion of the two coagulation factors due to the lack of functional ERGIC-53. Subsequently, additional glycoproteins were identified as ERGIC-53 cargo proteins including cathepsin Z [35,36], nicastrin [37], and mutant Ig μ [38]. Furthermore, ERGIC-53 interacts in a lectin-independent manner with multiple coagulation factor deficiency protein 2 (MCFD2). The MCFD2 gene is a second locus associated with F5F8D. It encodes a soluble luminal protein of 16 kDa [39]. The MCFD2 protein contains two EF-hands and interacts with ERGIC-53 in a calcium-dependent manner. MCFD2 and ERGIC-53 form a stable protein complex and cycle together in the early secretory pathway [40]. The intracellular localization of MCFD2 strictly depends on ERGIC-53 since a knockdown of ERGIC-53 leads to the secretion of MCFD2. Interestingly, the expression of both MCFD2 and ERGIC-53 is inducible in response to cell stress [41,42]. The potential function of MCFD2 in modulating the cargo-binding properties of ERGIC-53 will be discussed below.

TABLE 9.1
Features of Animal L-Type Lectin Family Members

Protein Name	Gene Name	Subcellular Localization	Molecular Properties	Sorting Motifs	Carbohydrate Specificity	Cargo Proteins	Tissue Distribution	References
ERGIC-53 (ER Golgi intermediate compartment protein of 53 kDa)	LMAN1 (lectin, mannose-binding 1)	ERGIC	• 53 kDa • Coiled-coil domain	• C-terminal KKFF ER export signal • Dilysine ER retrieval signal	• Mannose • Ca^{2+} required • pH optimum ~7.4	• Coagulation factors V + VIII • Cathepsin Z • Cathepsin C • Nicastrin • Mutant Ig μ	Ubiquitous	[19,26,28, 30,32–40, 55,61]
VIP36 (vesicular integral membrane protein of 36 kDa)	LMAN2 (lectin, mannose-binding 2)	ERGIC Golgi	• 36 kDa • N-glycosylated	• C-terminal KRFY dihydrophobic ER export signal • potential lysine-arginine ER retrieval signal	• High-mannose (D1 branch) • Ca^{2+} required • pH optimum ~ 6.4	• α-Amylase?	• Ubiquitous	[5,43,44, 46–48, 64–67]
ERGL (ERGIC-53-like protein)	LMAN1L (lectin, mannose-binding 1-like)	n.d.	• 55 kDa • Coiled-coil domain	• No characterized sorting motifs	• n.d.	• n.d.	• Prostate • Cardiac atrium • Salivary gland • Spleen • Central nervous system	[20]
VIPL (VIP36-like protein)	LMAN2L (lectin, mannose-binding 2-like)	ER	• 36 kDa • N-glycosylated	• C-terminal RKRFY di-hydrophobic ER export signal • Diarginine ER retention signal	n.d.	• Uncharacterized glycoproteins	• Ubiquitous	[21,22]

Note: n.d., not determined.

9.2.2 VIP36

VIP36 is an N-glycosylated calcium-binding membrane protein whose luminal domain is homologous to that of ERGIC-53 and leguminous plant lectins. In contrast to ERGIC-53, VIP36 lacks a coiled-coil domain and might function as a monomer [43]. VIP36 was originally purified from a detergent-insoluble fraction of madin-darbin canine kidney (MDCK) cells in an attempt to isolate novel vesicular components of glycolipid rafts [5]. In a subsequent study, however, VIP36 was found not to be associated with the detergent-insoluble fraction and hence is unlikely to be a raft component [44]. VIP36 is ubiquitously expressed in kidney, liver, intestine, lung, and spleen with very low levels in heart and brain as shown by Northern blotting of adult mouse RNA [43]. Endogenous VIP36 partially colocalizes with ERGIC-53 but also with Golgi markers [44]. Upon overexpression, VIP36 can also be detected at the plasma membrane and in endosomes [5]. VIP36 cycles between Golgi and ER and colocalizes with ERGIC-53 to the ERGIC in BFA-treated cells. Passage through the Golgi is evident from complex glycosylation of VIP36 [44], and cycling of green fluorescent protein (GFP)-tagged VIP36 was directly visualized in living COS cells at moderate expression levels [45]. The presence of a phenylalanine–tyrosine putative ER export motif and a dibasic motif resembling (but not strictly conforming to) an ER retrieval signal in the C-terminal cytosolic domain (Figure 9.1) may mediate the cycling of VIP36 between ER, ERGIC, and Golgi. It should be emphasized, however, that there is no direct proof for the functionality of these putative transport motifs in VIP36.

Data regarding carbohydrate specificity and calcium requirement of VIP36 are controversial. While the recombinant luminal domain of VIP36 was initially shown to bind *N*-acetyl-D-galactosamine in a calcium-dependent manner [43], a glutathion-*S*-transferase (GST) fusion protein of the luminal domain of VIP36 was later reported to bind preferentially high-mannose-type glycans in a calcium-independent way [46]. Recently, the carbohydrate specificity of VIP36 was analyzed in detail by frontal affinity chromatography using a carbohydrate library of 21 different oligosaccharides [47]. This study revealed that VIP36 binds high-mannose-type glycans containing the D1 branch ($Man\alpha_{1\rightarrow2}$ $Man\alpha_{1\rightarrow2}$ Man branch) and that this binding is calcium-dependent. The conclusion that VIP36 is a calcium-dependent lectin with specificity for high-mannose-type glycans is supported by the localization of VIP36 to the early secretory pathway. Cargo proteins of VIP36 have not been identified with certainty, although α-amylase was proposed as candidate cargo [48]. The evidence for a dependence of α-amylase secretion on VIP36 requires further validation.

9.2.3 ERGL

In 2001, a novel L-type animal lectin was identified from a human prostate-specific expressed sequence tag (EST) cluster [20]. The deduced protein was termed ERGIC-53-like protein (ERGL) due to its high similarity with the amino terminal region of ERGIC-53. ERGL is a type-I membrane protein with a luminal CRD and a luminal coiled-coil domain. In contrast to the other L-type animal lectins, ERGL lacks typical transport motifs in its cytoplasmic domain and, therefore, might not be able to traffic in the secretory pathway independently, or may even operate as an ER protein. The presence of a coiled-coil domain in ERGIC-53 and ERGL may indicate that both proteins form heterooligomers and that ERGL is transported indirectly via an interaction with ERGIC-53. In contrast to the ubiquitous tissue distribution of the other L-type animal lectins, expression of ERGL mRNA is restricted to prostate, spleen, salivary gland, cardiac atrium, and the central nervous system. So far, endogenous human ERGL protein has not been studied. In human prostate, two ERGL splice variants exist, one of which lacks the transmembrane region, resulting in a potentially secreted protein [20]. Additionally, a putative rat homolog of ERGL was cloned and termed sublingual acinar membrane protein (SLAMP) [49]. SLAMP shows an even more restricted expression pattern than ERGL and appears to be only expressed in the sublingual gland as revealed by Northern and Western blotting [49]. Currently, no clear data exist on the subcellular localization and carbohydrate specificity of ERGL or SLAMP, nor have potential cargo proteins been identified.

9.2.4 VIPL

VIPL was identified in 2003 as the fourth member of the L-type lectin family in a databank search for novel lectins using the conserved CRD of ERGIC-53 and VIP36 as a search string [21,22]. Due to its high similarity to VIP36, it was termed VIP36-like protein (VIPL). VIPL is an N-glycosylated type-I membrane protein with a short half-life of approximately 30 min in HeLa cells [21,22]. In contrast to ERGIC-53 and VIP36, VIPL is a noncycling ER-resident protein. ER localization is in agreement with the high-mannose N-glycan residue of VIPL, which is not processed by Golgi enzymes. Furthermore, VIPL does not redistribute to ERGIC clusters in response to BFA [22]. ER retention of VIPL is mediated by a RKR retention motif in the cytosolic domain. This diarginine motif is dominant over the phenylalanine/tyrosine ER export signal at the C-terminus of VIPL since mutation of the RKR motif to RAA abolishes retention and leads to transport of VIPL to the plasma membrane. VIPL is ubiquitously expressed with high levels found in skeletal muscle, kidney, heart, liver, and placenta and minor levels found in brain, thymus, spleen, small intestine, and lung [21]. Currently, there exist no data concerning the carbohydrate specificity of VIPL. A hint that VIPL may play a role in ER export of glycoproteins is based on knockdown experiments using VIPL-specific short interference RNA (siRNA) [21]. Downregulation of VIPL leads to inefficient secretion of two unknown glycoproteins with a molecular mass of 35 and 250 kDa. A direct dependence of glycoprotein secretion on VIPL remains to be shown.

9.3 L-TYPE LECTINS IN SIMPLER EUKARYOTES

The expression of L-type lectins is not restricted to animal cells since they were also identified in simpler eukaryotes such as the slime mold *Dictyostelium dyscoideum* [50] and yeast. The baker's yeast *Saccharomyces cerevisiae* expresses two L-type lectins, Emp46p and Emp47p, which show sequence homology with ERGIC-53 [51]. Emp46p and Emp47p are type I membrane proteins, which heterooligomerize and cycle between ER and Golgi. Like ERGIC-53, the two proteins contain a luminal lectin domain, a luminal coiled-coil domain, and a short cytoplasmic tail that interacts with COPI and COPII [52]. The crystal structures of the CRD of Emp46p and Emp47p reveal a β-sandwich fold that resembles the structure of human ERGIC-53 [53]. Emp46p does not bind calcium but potassium ions while Emp47p does not bind metal ions at all. No specific cargo proteins have been identified for Emp46p and Emp47p but gene disruption of both EMP46 and EMP47 results in a secretion defect of a subset of glycoproteins [51]. Thus, Emp46p and Emp47p are thought to operate as calcium-independent transport receptors that facilitate the recruitment of some glycoproteins into vesicles budding from the ER.

9.4 FUNCTION OF L-TYPE LECTINS

ERGIC-53 is the most extensively studied L-type lectin, whereas functional data is scarce for VIP36, ERGL, and VIPL. Nevertheless, the current information suggests a function of the four animal L-type lectins in glycoprotein sorting and trafficking as illustrated in Figure 9.2.

Genetic and biochemical evidence suggests a role of ERGIC-53 in selective ER to ERGIC transport of a subset of glycoproteins. The genetic data derives from the analysis of patients suffering from F5F8D and link ERGIC-53 to the efficient secretion of blood coagulation factors V and VIII [34]. This finding was confirmed in mammalian cell culture studies demonstrating reduced secretion of blood coagulation factors V and VIII when ERGIC-53 was mislocalized to the ER by the expression of a dominant-negative mutant of ERGIC-53 [54]. Furthermore, cross-linking experiments showed a direct physical interaction between blood coagulation factor VIII and ERGIC-53 [55].

Most information on the mode of action of ERGIC-53 derives from studies with ERGIC-53's cargo protein cathepsin Z. Cathepsin Z is a lysosomal glycoprotein that coprecipitates with ERGIC-53 after chemical cross-linking [35]. The requirement for chemical cross-linking reflects the low affinity

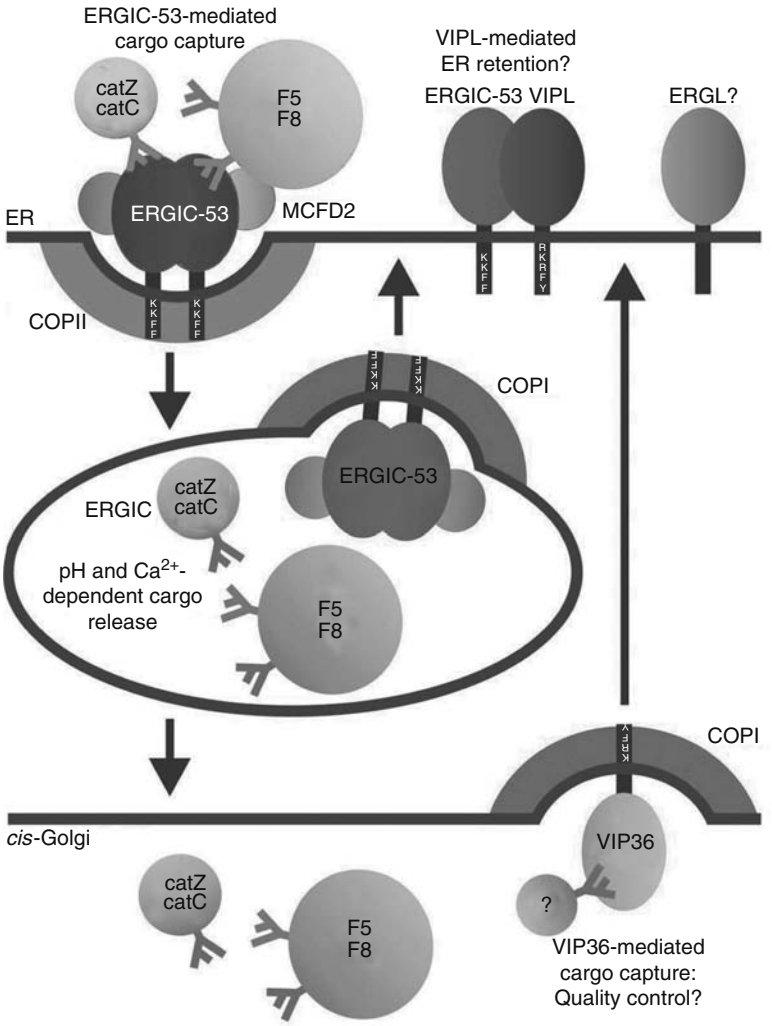


FIGURE 9.2 (See color insert following blank page 170. Also see CD for color figure.) Role of L-type lectins in glycoprotein sorting and trafficking. ERGIC-53, VIP36, ERGL, and VIPL localize to the early secretory pathway. A transport receptor function is well documented for ERGIC-53, which, for simplicity, is depicted as a homodimeric rather than a homohexameric protein. In the ER, ERGIC-53 captures glycoproteins, such as cathepsin Z (catZ) and cathepsin C (catC). The cytosolic dihydrophobic motif of ERGIC-53 (orange) binds to the Sec24 subunit of the COPII coat. Thereby the ERGIC-53/cargo complex is packaged into COPII-coated vesicles and transported to the ERGIC. In the ERGIC, a drop in pH is thought to protonate a conserved histidine residue in the CRD of ERGIC-53, which together with a drop of free calcium triggers cargo release. After dissociation, ERGIC-53 is packaged into COPI-coated vesicles via its dilysine motif (blue) and is retrieved back to the ER for another round of transport. The transport signal in cathepsin Z and cathepsin C that is recognized by ERGIC-53 is composed of a high-mannose N-glycan and a β-hairpin in the protein backbone. This structural motif is present in both fully folded cathepsin Z and cathepsin C. The ERGIC-53 interaction partner MCFD2 is a luminal factor, which recruits blood coagulation factors V and VIII (F5 and F8) to ERGIC-53. VIPL, a short-lived ER resident protein, is a potential regulator of ERGIC-53. It may control passage of ERGIC-53 through the ER and perhaps regulate cargo loading. Overexpressed ERGL localizes to the ER and also binds to ERGIC-53 (L. Liang and H.P. Hauri, unpublished), but endogenous ERGL has not been studied yet. VIP36 is likely to act as a transport receptor. In contrast to ERGIC-53, VIP36 binds high-mannose glycans with a pH optimum of about 6.4, which favors the notion that this lectin captures cargo in the Golgi. VIP36 might retrieve incompletely trimmed glycoproteins back to the ER in a secondary quality control process.

of the underlying protein interaction. Unlike stable protein–protein interactions, receptor–cargo interactions are transient and need to be reversible. Binding of cathepsin Z occurs in the ER and is carbohydrate- and calcium-dependent. Dissociation of cathepsin Z takes place in the ERGIC and is delayed if ERGIC-53 is mislocalized to the ER [35]. These findings strongly support a transport receptor function of ERGIC-53. Thus, ERGIC-53 captures cargo glycoproteins in the ER, recruits them to COPII-coated vesicles, and thereby transports them to the ERGIC where they are released (Figure 9.2).

How are cargo proteins released in the ERGIC? A recent study suggests that the release is induced by a combined change of pH and calcium in the ERGIC. Purified ERGIC-53 binds to immobilized mannose at pH 7.4, the pH of the ER, but not at lower pH. This binding is calcium-sensitive [36]. Although pH and calcium levels in the ERGIC are unknown, a pH drop is in line with the known progressive acidification of the secretory pathway from the ER (pH 7.1–7.4) to the *trans*-Golgi network (pH 5.9–6.3) [56,57]. On a molecular level, pH and calcium-dependent carbohydrate binding of ERGIC-53 can be assigned to a conserved histidine residue in its CRD that acts as a pH sensor. Determination of the crystal structure of ERGIC-53 in its calcium-bound form confirmed that this conserved histidine residue complexes one of the two calcium ions present in ERGIC-53 [58]. Importantly, these biochemical results are supported by *in vivo* data. Cell acidification results in less efficient binding of cathepsin Z to ERGIC-53, as opposed to organelle neutralization that delays cargo dissociation [36].

The identification of MCFD2 as an ERGIC-53 interacting protein required for the efficient secretion of blood coagulation factors V and VIII raised the possibility that MCFD2 may regulate the lectin activity of ERGIC-53. This possibility was tested by a siRNA approach in HeLa cells [40]. Depletion of MCFD2, however, did not affect binding of cathepsin Z and cathepsin C to ERGIC-53, indicating that the lectin activity of ERGIC-53 does not require the presence of MCFD2 *in vivo*. While MCFD2 was found in a complex with blood coagulation factor VIII [55], no direct interaction was evident between MCFD2 and cathepsin Z or cathepsin C [40]. This suggests that MCFD2 is a specific recruitment factor for blood coagulation factors V and VIII and that the MCFD2/ERGIC-53 receptor complex possesses dual-binding properties. According to this model, ERGIC-53 binds cargo proteins such as cathepsin Z and cathepsin C in a lectin-dependent but MCFD2-independent manner. In contrast, blood coagulation factors V and VIII are specifically recruited by MCFD2. This notion is in line with the finding that unglycosylated blood coagulation factor VIII but not unglycosylated cathepsin Z can be cross-linked to ERGIC-53 [35,55].

The current list of ERGIC-53 cargo proteins includes blood coagulation factors V and VIII, cathepsin Z, cathepsin C, nicastrin, and mutant Ig μ. What defines an ERGIC-53 cargo protein? Obviously, the mere presence of N-glycans is insufficient to promote efficient binding of glycoproteins to ERGIC-53. Recent studies on cathepsin Z identified a combined oligosaccharide/β-hairpin motif to be required for the ERGIC-53-assisted ER export [59]. This ER export motif is composed of a high-mannose-type glycan intimately associated with a surface-exposed peptide β-hairpin loop and is only present in the fully folded cathepsin Z. Interestingly, this oligosaccharide/β-hairpin motif is also present in cathepsin C, a close relative of cathepsin Z, whose transport is also dependent of functional ERGIC-53 [33]. The oligosaccharide/β-hairpin motif of cathepsin C can be superimposed with that of cathepsin Z [59]. The finding that ERGIC-53 recognizes both glycan and peptide structures in a conformation-dependent manner suggests that ERGIC-53 acts in a secondary quality control process following the primary folding and quality control processes in the ER. Cocrystallization of ERGIC-53 with one of its cargo proteins will be required to understand the molecular mechanisms underlying the cargo selection process in greater detail.

A recent study reporting an interaction of mutant Ig μ and ERGIC-53 points to an additional function of ERGIC-53 as a modulator of the oligomerization process of glycoproteins [38]. In this study, ERGIC-53 was found to be involved in the formation of dilated ER membranes containing aggregated immunoglobin heavy chains, also known as Russell bodies. In the absence of light chains, mutant Igμ heavy chains interact with ERGIC-53 and condense into aggregates. Hexamerization of

ERGIC-53 may seed this condensation and aggregation process. In more general terms, ERGIC-53 monomers might capture cargo molecules and by hexamerization promote the oligomerization of bound cargo proteins.

Does VIP36 operate as a cargo receptor for glycoproteins similar to ERGIC-53? So far, no glycoproteins have been identified whose secretion depends on VIP36. However, the localization and dynamics in the early secretory pathway [44,45] as well as the carbohydrate specificity [47] show a high degree of similarity of VIP36 and ERGIC-53 and support the notion of a transport receptor function. Although sharing a carbohydrate specificity for mannose, the glycan-binding parameters of VIP36 differ from those of ERGIC-53, particularly in terms of pH optimum for the carbohydrate interaction. ERGIC-53 binds high-mannose residues with highest affinity at pH 7.4 and at high calcium levels [36], whereas VIP36 shows optimal binding of the mannose D1 branch at a pH of 6.4. Based on the difference in pH optima, a plausible model is that VIP36 acts complementary to ERGIC-53 in secondary quality control by retrieving glycoproteins back to the ER [47,60]. According to this model, ERGIC-53 binds glycoproteins with high-mannose residues in the ER and releases the cargo due to the decrease of the pH in the ERGIC, whereas VIP36 binds high-mannose glycoproteins in the Golgi that have escaped correct carbohydrate processing and retrieves them back to the ER (Figure 9.2). However, this model is largely based on *in vitro* binding studies and requires further validation.

Due to limited knowledge, one can only speculate about possible functions of ERGL and VIPL. So far, no lectin activity has been reported for the two proteins. Nevertheless, the conservation of key residues, known to be essential for metal and sugar binding in the CRD of ERGIC-53 and VIP36, suggests that at least VIPL can bind glycans. A role of VIPL in glycoprotein trafficking is endorsed by the finding that siRNA-mediated depletion of VIPL impairs the secretion of two glycoproteins [21]. The ER localization of VIPL, however, argues against a role as a conventional transport receptor that captures cargo proteins in the ER and escorts them to the ERGIC. An interesting observation comes from the finding that overexpression of VIPL redistributes ERGIC-53 to the ER without affecting the morphology of the early secretory pathway or the localization of other cycling proteins, such as the KDEL-receptor [22]. It seems possible, that VIPL regulates the localization and function of ERGIC-53 and perhaps other lectins in the ER. Consistent with such a regulatory function of VIPL is its short half-life [21].

9.5 CONCLUSION AND FUTURE DIRECTIONS

In conclusion, the following findings define a role of animal L-type lectins in glycoprotein trafficking:

1. A knockout or knockdown of L-type lectins results in reduced secretion of a subset of glycoproteins. A knockout of ERGIC-53 is associated with F5F8D and leads to a secretion defect of blood coagulation factors V and VIII [34]. A knockdown of VIPL leads to the impaired secretion of two uncharacterized glycoproteins [21]. Furthermore, a secretion defect was observed for of a subset of glycoprotein upon disruption of EMP46 and EMP47, the two yeast gene orthologs of animal L-type lectins [51].
2. L-type lectins directly interact with cargo glycoproteins. Direct protein interactions have been established for ERGIC-53 and blood coagulation factors V and VIII, cathepsin Z, cathepsin C, nicastrin, and mutant Ig μ [35,37,38,55,61]. α-Amylase is discussed as a putative cargo protein for VIP36 [48].
3. The interaction of ERGIC-53 with its cargo protein cathepsin Z is dependent on the lectin activity of ERGIC-53 and the glycosylation state of cathepsin Z. *In vivo*, binding of cargo to ERGIC-53 occurs in the ER and dissociation in the ERGIC [35].
4. The carbohydrate-binding properties of L-type lectins can be modulated by pH and calcium levels. Glycan binding of ERGIC-53 requires a pH optimum of 7.4 and high calcium

levels [36], whereas VIP36 has a pH optimum of 6.4 [47]. Due to progressive acidification along the secretory pathway, pH differences potentially regulate cargo binding and cargo release.

5. L-type lectins possess ER export and ER retention/retrieval signals in their cytosolic tails. These signals interact with coat subunits and mediate recruitment into transport vesicles. The continuous cycling of ERGIC-53 and VIP36 is in line with the proposed transport receptor function.

It will be important to functionally characterize the less well-defined members of the animal L-type lectin family. ERGL and VIPL have been defined as L-type lectins only due to their homology to ERGIC-53. The conservation of key residues essential for metal and sugar binding suggests that at least VIPL is a functional lectin [21,22] but binding studies with glycans will be required to clarify this point. An attractive method to determine the carbohydrate-binding properties of animal L-type lectins is frontal affinity chromatography, a method which was successfully applied to VIP36 [47]. Using frontal affinity chromatography, a sugar library is screened for binding to the recombinant CRD of the lectin. The method provides information on carbohydrate specificity and affinity. Moreover, one can easily change the binding environment and thereby investigate pH and calcium dependence of the carbohydrate binding. Nevertheless, the findings obtained by frontal affinity chromatography will require validation *in vivo*, for instance, by studying protein secretion in cells in which a given lectin is silenced by siRNA. Since siRNA-mediated knockdowns often suffer from incomplete protein depletion, functional studies should also be conducted in L-type lectin knockout mice.

Our understanding of L-type lectins greatly profited from the characterization of ERGIC-53 and the identification of cargo proteins, which directly interact with ERGIC-53. Cargo proteins for

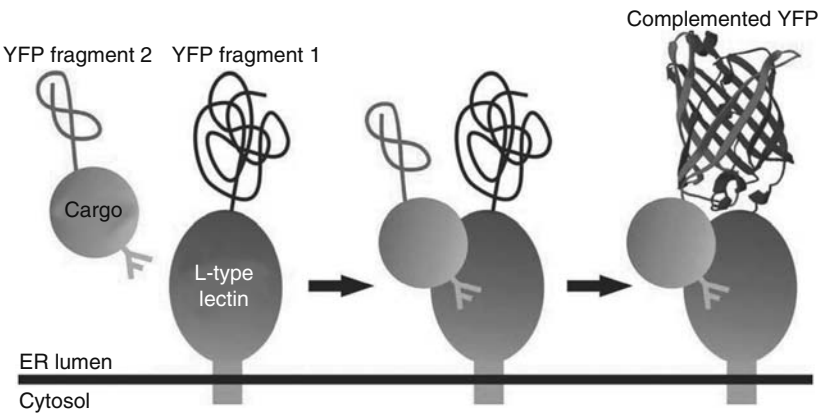


FIGURE 9.3 (See color insert following blank page 170. Also see CD for color figure.) Luminal YFP complementation as a tool to identify cargo proteins of L-type lectins. L-type lectins bind their cargo proteins in the lumen of the ER, which is an oxidizing environment with a high calcium concentration. This interaction is carbohydrate-dependent, transient, and of low affinity, and hence difficult to capture. The luminal YFP complementation assay was developed as a powerful tool to trap and visualize protein–protein interactions in the lumen of the secretory pathway of living cells. (From Nyfeler, B., Michnick, S.W., and Hauri, H.P., *Proc. Natl Acad. Sci. USA*, 102, 6350, 2005.) The basic concept of the luminal YFP complementation assay relies on splitting YFP into two fragments, YFP fragment 1 and YFP fragment 2, which show no fluorescence by themselves and do not spontaneously fold. YFP fragments 1 and 2 are fused to an L-type lectin and a potential cargo protein, respectively. The fusion proteins are expressed in mammalian cells. If they interact, the YFP fragments are brought into close proximity and are able to reconstitute fluorescent YFP. The luminal YFP complementation approach was shown to be specific and capable of capturing the interactions of ERGIC-53 with its cargo proteins cathepsin Z and cathepsin C. (From Nyfeler, B., Zhang, B., Ginsburg, D., Kaufman, R.J., and Hauri, H.P., *Traffic*, 7, 1473, 2006; Nyfeler, B., Michnick, S.W., and Hauri, H.P., *Proc. Natl Acad. Sci. USA*, 102, 6350, 2005.)

VIP36, ERGL, and VIPL remain to be found. The identification of cargo proteins of L-type lectins is far from trivial due to the weak and transient nature of the underlying protein interactions. The yeast two-hybrid system and analysis of pull-downs by mass spectrometry are currently the most popular methods to identify interacting proteins. Although highly successful in the dissection of cytosolic protein complexes, these methods are of limited value to detect protein interactions of L-type lectins and cargo proteins in the lumen of the ER, ERGIC, and Golgi. These limitations can be overcome with the adaptation of the recently developed yellow fluorescence protein (YFP) complementation assay to the lumen of the secretory pathway [61,62]. The basic concept of the luminal YFP complementation is depicted in Figure 9.3 and relies on splitting the YFP reporter protein into two fragments that exhibit no fluorescence by themselves. These YFP fragments are fused to interacting proteins whose interaction brings the two YFP fragments into close proximity and induces the correct folding and complementation of fluorescent YFP. The proof-of-concept of luminal YFP complementation was provided using ERGIC-53 and its interaction partners as a model [61]. YFP fluorescence complementation revealed the oligomerization of ERGIC-53, its interaction with MCFD2, and its lectin-dependent interaction with cathepsin Z and cathepsin C. The high specificity of the assay was demonstrated by the finding that inactivation of ERGIC-53's lectin domain by a point mutation selectively impaired YFP complementation of the lectin-dependent interaction with cathepsin Z [61]. Due to the irreversible nature of YFP complementation, transient protein interactions can be trapped, which is highly desirable for the identification of novel L-type lectin cargo proteins. As demonstrated by Remy and Michnick, YFP complementation can be used to screen for protein interaction partners in a genomewide setting [63]. The potential setup for a genomewide L-type lectin cargo hunt would be as follows: One fragment of YFP is fused to a cDNA library while the second YFP fragment is fused to an L-type lectin. The L-type lectin bait protein is coexpressed with the cDNA fusion library, cells expressing complemented YFP are isolated by fluorescence activated cell sorting (FACS), and library plasmids are recovered and analyzed. Such a genomewide cargo hunt by YFP complementation holds the promise to identify novel glycoprotein cargo of L-type lectins, which may increase our knowledge of the role of animal L-type lectins in glycoprotein sorting and trafficking.

REFERENCES

1. Helenius, A. and Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73**, 1019–1049, 2004.
2. Ruddock, L.W. and Molinari, M. N-glycan processing in ER quality control. *J Cell Sci* **119**, 4373–4380, 2006.
3. Hauri, H.P., Nufer, O., Breuza, L., Tekaya, H.B., and Liang, L. Lectins and protein traffic early in the secretory pathway. *Biochem Soc Symp* 73–82, 2002.
4. Itin, C., Roche, A.C., Monsigny, M., and Hauri, H.P. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. *Mol Biol Cell* **7**, 483–493, 1996.
5. Fiedler, K., Parton, R.G., Kellner, R., Etzold, T., and Simons, K. VIP36, a novel component of glycolipid rafts and exocytic carrier vesicles in epithelial cells. *EMBO J* **13**, 1729–1740, 1994.
6. Appenzeller-Herzog, C., and Hauri, H.P. The ER–Golgi intermediate compartment (ERGIC): In search of its identity and function. *J Cell Sci* **119**, 2173–2183, 2006.
7. Lee, M.C., Miller, E.A., Goldberg, J., Orci, L., and Schekman, R. Bi-directional protein transport between the ER and Golgi. *Annu Rev Cell Dev Biol* **20**, 87–123, 2004.
8. Nishimura, N. and Balch, W.E. A di-acidic signal required for selective export from the endoplasmic reticulum. *Science* **277**, 556–558, 1997.
9. Nufer, O. et al. Role of cytoplasmic C-terminal amino acids of membrane proteins in ER export. *J Cell Sci* **115**, 619–628, 2002.
10. Kappeler, F., Klopfenstein, D.R., Foguet, M., Paccaud, J.P., and Hauri, H.P. The recycling of ERGIC-53 in the early secretory pathway. ERGIC-53 carries a cytosolic endoplasmic reticulum-exit determinant interacting with COPII. *J Biol Chem* **272**, 31801–31808, 1997.
11. Miller, E., Antonny, B., Hamamoto, S., and Schekman, R. Cargo selection into COPII vesicles is driven by the Sec24p subunit. *EMBO J* **21**, 6105–6113, 2002.

12. Jackson, M.R., Nilsson, T., and Peterson, P.A. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *Embo J* **9**, 3153–3162, 1990.
13. Cosson, P. and Letourneur, F. Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. *Science* **263**, 1629–1631, 1994.
14. Karrenbauer, A. et al. The rate of bulk flow from the Golgi to the plasma membrane. *Cell* **63**, 259–267, 1990.
15. Malkus, P., Jiang, F., and Schekman, R. Concentrative sorting of secretory cargo proteins into COPII-coated vesicles. *J Cell Biol* **159**, 915–921, 2002.
16. Lewis, M.J. and Pelham, H.R. Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. *Cell* **68**, 353–364, 1992.
17. Lewis, M.J. and Pelham, H.R. A human homologue of the yeast HDEL receptor. *Nature* **348**, 162–163, 1990.
18. Schweizer, A., Fransen, J.A., Bachi, T., Ginsel, L., and Hauri, H.P. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the cis-side of the Golgi apparatus. *J Cell Biol* **107**, 1643–1653, 1988.
19. Hauri, H.P., Kappeler, F., Andersson, H., and Appenzeller, C. ERGIC-53 and traffic in the secretory pathway. *J Cell Sci* **113** (Pt 4), 587–596, 2000.
20. Yerushalmi, N. et al. ERGL, a novel gene related to ERGIC-53 that is highly expressed in normal and neoplastic prostate and several other tissues. *Gene* **265**, 55–60, 2001.
21. Neve, E.P., Svensson, K., Fuxe, J., and Pettersson, R.F. VIPL, a VIP36-like membrane protein with a putative function in the export of glycoproteins from the endoplasmic reticulum. *Exp Cell Res* **288**, 70–83, 2003.
22. Nufer, O., Mitrovic, S., and Hauri, H.P. Profile-based data base scanning for animal L-type lectins and characterization of VIPL, a novel VIP36-like endoplasmic reticulum protein. *J Biol Chem* **278**, 15886–15896, 2003.
23. Lahtinen, U., Hellman, U., Wernstedt, C., Saraste, J., and Pettersson, R.F. Molecular cloning and expression of a 58-kDa *cis*-Golgi and intermediate compartment protein. *J Biol Chem* **271**, 4031–4037, 1996.
24. Saraste, J., Palade, G.E., and Farquhar, M.G. Antibodies to rat pancreas Golgi subfractions: Identification of a 58-kD *cis*-Golgi protein. *J Cell Biol* **105**, 2021–2029, 1987.
25. Ben-Tekaya, H., Miura, K., Pepperkok, R., and Hauri, H.P. Live imaging of bidirectional traffic from the ERGIC. *J Cell Sci* **118**, 357–367, 2005.
26. Klumperman, J. et al. The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. *J Cell Sci* **111** (Pt 22), 3411–3425, 1998.
27. Lippincott-Schwartz, J. et al. Microtubule-dependent retrograde transport of proteins into the ER in the presence of brefeldin A suggests an ER recycling pathway. *Cell* **60**, 821–836, 1990.
28. Nufer, O., Kappeler, F., Guldbrandsen, S., and Hauri, H.P. ER export of ERGIC-53 is controlled by cooperation of targeting determinants in all three of its domains. *J Cell Sci* **116**, 4429–4440, 2003.
29. Wendeler, M.W., Paccaud, J.P., and Hauri, H.P. Role of Sec24 isoforms in selective export of membrane proteins from the endoplasmic reticulum. *EMBO Rep* **8**, 258–264, 2007.
30. Itin, C., Schindler, R., and Hauri, H.P. Targeting of protein ERGIC-53 to the ER/ERGIC/cis-Golgi recycling pathway. *J Cell Biol* **131**, 57–67, 1995.
31. Neve, E.P., Lahtinen, U., and Pettersson, R.F. Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. *J Mol Biol* **354**, 556–568, 2005.
32. Velloso, L.M., Svensson, K., Schneider, G., Pettersson, R.F., and Lindqvist, Y. Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum. *J Biol Chem* **277**, 15979–15984, 2002.
33. Vollenweider, F., Kappeler, F., Itin, C., and Hauri, H.P. Mistargeting of the lectin ERGIC-53 to the endoplasmic reticulum of HeLa cells impairs the secretion of a lysosomal enzyme. *J Cell Biol* **142**, 377–389, 1998.
34. Nichols, W.C. et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell* **93**, 61–70, 1998.
35. Appenzeller, C., Andersson, H., Kappeler, F., and Hauri, H.P. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nat Cell Biol* **1**, 330–334, 1999.
36. Appenzeller-Herzog, C., Roche, A.C., Nufer, O., and Hauri, H.P. pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. *J Biol Chem* **279**, 12943–12950, 2004.
37. Morais, V.A. et al. N-glycosylation of human nicastrin is required for interaction with the lectins from the secretory pathway calnexin and ERGIC-53. *Biochim Biophys Acta* **1762**, 802–810, 2006.

38. Mattioli, L. et al. ER storage diseases: A role for ERGIC-53 in controlling the formation and shape of Russell bodies. *J Cell Sci* **119**, 2532–2541, 2006.
39. Zhang, B. et al. Bleeding due to disruption of a cargo-specific ER-to-Golgi transport complex. *Nat Genet* **34**, 220–225, 2003.
40. Nyfeler, B., Zhang, B., Ginsburg, D., Kaufman, R.J., and Hauri, H.P. Cargo selectivity of the ERGIC-53/MCFD2 transport receptor complex. *Traffic* **7**, 1473–1481, 2006.
41. Nyfeler, B., Nufer, O., Matsui, T., Mori, K., and Hauri, H.P. The cargo receptor ERGIC-53 is a target of the unfolded protein response. *Biochem Biophys Res Commun* **304**, 599–604, 2003.
42. Spatuzza, C. et al. Heat shock induces preferential translation of ERGIC-53 and affects its recycling pathway. *J Biol Chem* **279**, 42535–42544, 2004.
43. Fiedler, K. and Simons, K. Characterization of VIP36, an animal lectin homologous to leguminous lectins. *J Cell Sci* **109** (Pt 1), 271–276, 1996.
44. Fullekrug, J., Scheiffele, P., and Simons, K. VIP36 localisation to the early secretory pathway. *J Cell Sci* **112** (Pt 17), 2813–2821, 1999.
45. Dahm, T., White, J., Grill, S., Fullekrug, J., and Stelzer, E.H. Quantitative ER ↔ Golgi transport kinetics and protein separation upon Golgi exit revealed by vesicular integral membrane protein 36 dynamics in live cells. *Mol Biol Cell* **12**, 1481–1498 2001.
46. Hara-Kuge, S., Ohkura, T., Seko, A., and Yamashita, K. Vesicular-integral membrane protein, VIP36, recognizes high-mannose type glycans containing alpha1 → 2 mannosyl residues in MDCK cells. *Glycobiology* **9**, 833–839, 1999.
47. Kamiya, Y. et al. Sugar-binding properties of VIP36, an intracellular animal lectin operating as a cargo receptor. *J Biol Chem* **280**, 37178–37182, 2005.
48. Hara-Kuge, S., Seko, A., Shimada, O., Tosaka-Shimada, H., and Yamashita, K. The binding of VIP36 and alpha-amylase in the secretory vesicles via high-mannose type glycans. *Glycobiology* **14**, 739–744, 2004.
49. Sakulsak, N., Wakayama, T., Hipkaeo, W., Yamamoto, M., and Iseki, S. Cloning and characterization of a novel animal lectin expressed in the rat sublingual gland. *J Histochem Cytochem* **53**, 1335–1343, 2005.
50. Glockner, G. et al. Sequence and analysis of chromosome 2 of *Dictyostelium discoideum*. *Nature* **418**, 79–85, 2002.
51. Sato, K. and Nakano, A. Emp47p and its close homolog Emp46p have a tyrosine-containing endoplasmic reticulum exit signal and function in glycoprotein secretion in *Saccharomyces cerevisiae*. *Mol Biol Cell* **13**, 2518–2532, 2002.
52. Sato, K. and Nakano, A. Oligomerization of a cargo receptor directs protein sorting into COPII-coated transport vesicles. *Mol Biol Cell* **14**, 3055–3063, 2003.
53. Satoh, T. et al. Structures of the carbohydrate recognition domain of Ca^{2+}-independent cargo receptors Emp46p and Emp47p. *J Biol Chem* **281**, 10410–10419, 2006.
54. Moussalli, M. et al. Mannose-dependent endoplasmic reticulum (ER)-Golgi intermediate compartment-53-mediated ER to Golgi trafficking of coagulation factors V and VIII. *J Biol Chem* **274**, 32539–32542, 1999.
55. Zhang, B., Kaufman, R.J., and Ginsburg, D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. *J Biol Chem* **280**, 25881–25886, 2005.
56. Wu, M.M. et al. Mechanisms of pH regulation in the regulated secretory pathway. *J Biol Chem* **276**, 33027–33035, 2001.
57. Weisz, O.A. Acidification and protein traffic. *Int Rev Cytol* **226**, 259–319, 2003.
58. Velloso, L.M., Svensson, K., Pettersson, R.F., and Lindqvist, Y. The crystal structure of the carbo hydrate-recognition domain of the glycoprotein sorting receptor p58/ERGIC-53 reveals an unpredicted metal-binding site and conformational changes associated with calcium ion binding. *J Mol Biol* **334**, 845–851, 2003.
59. Appenzeller-Herzog, C. et al. Carbohydrate- and conformation-dependent cargo capture for ER-exit. *Mol Biol Cell* **16**, 1258–1267, 2005.
60. Hauri, H., Appenzeller, C., Kuhn, F., and Nufer, O. Lectins and traffic in the secretory pathway. *FEBS Lett* **476**, 32–37, 2000.
61. Nyfeler, B., Michnick, S.W., and Hauri, H.P. Capturing protein interactions in the secretory pathway of living cells. *Proc Natl Acad Sci USA* **102**, 6350–6355, 2005.
62. Kerppola, T.K. Visualization of molecular interactions by fluorescence complementation. *Nat Rev Mol Cell Biol* **7**, 449–456, 2006.

63. Remy, I. and Michnick, S.W. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. *Methods* **32**, 381–388, 2004.
64. Hara-Kuge, S. et al. Involvement of VIP36 in intracellular transport and secretion of glycoproteins in polarized Madin–Darby canine kidney (MDCK) cells. *J Biol Chem* **277**, 16332–16339, 2002.
65. Hara-Kuge, S., Seko, A., and Yamashita, K. Carbohydrate recognition of vesicular integral protein of 36 kDa (ViP36) in intracellular transport of newly synthesized glycoproteins. *Methods Enzymol* **363**, 525–532, 2003.
66. Shimada, O. et al. Localization of VIP36 in the post-Golgi secretory pathway also of rat parotid acinar cells. *J Histochem Cytochem* **51**, 1057–1063, 2003.
67. Shimada, O. et al. Clusters of VIP-36-positive vesicles between endoplasmic reticulum and Golgi apparatus in GH3 cells. *Cell Struct Funct* **28**, 155–163, 2003.
68. Teasdale, R.D. and Jackson, M.R. Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the Golgi apparatus. *Annu Rev Cell Dev Biol* **12**, 27–54, 1996.

10 P-Type Lectins and Lysosomal Enzyme Targeting

Nancy M. Dahms, Linda J. Olson, and Jung-Ja P. Kim

CONTENTS

10.1 INTRODUCTION

The mannose-6-phosphate receptors (MPRs) play an essential role in the formation of functional lysosomes by directing newly synthesized soluble acid hydrolases containing a mannose 6-phosphate (Man-6-P) tag to lysosomes. The importance of this targeting process in the generation of lysosomes containing a full-complement of hydrolytic enzymes is evidenced by the existence of over 40 different lysosomal storage diseases (LSDs) [1–3]. LSDs constitute a significant portion of inborn metabolic disorders: Although individually rare, the collective frequency of LSDs is estimated to be approximately 1 in 8000 live births [4]. The majority of LSDs are caused by a deficiency of a single lysosomal enzyme that results in the accumulation or storage of undigested endogenous macromolecules within lysosomes, with the defective lysosomes appearing as phase-dense inclusions in the cytoplasm. How the accumulation of macromolecules in lysosomes leads to progressive cellular and organ dysfunction and to the enormous heterogeneity of clinical phenotypes observed in LSD patients remains an unanswered question [3]. *In vitro* cell culture studies demonstrating that the addition of exogenous enzyme to fibroblasts from LSD patients could decrease the intracellular accumulation of stored products in lysosomes [5–7] paved the way for the treatment of patients by enzyme replacement therapy. In these patients, lysosomal storage can be partially or completely reversed in many target tissues by intravenous injection, typically on a weekly basis, of the missing lysosomal enzyme, and the injected enzyme is internalized by receptor-mediated endocytosis via endogenous receptors. To date, three (Fabry disease, Mucopolysaccharidosis I, and Pompe disease) out of the four Food and Drug Administration

(FDA)-approved enzyme replacement therapies target the MPRs for uptake of the infused Man-6-P-containing enzyme [8]. Understanding the molecular basis of how MPRs function in the biogenesis of lysosomes is of fundamental importance in the development of new and improved therapies for the treatment of LSDs.

10.2 SYNTHESIS OF LYSOSOMAL ENZYMES AND GENERATION OF THE MAN-6-P TAG

Generation of the Man-6-P signal occurs by a two-step process during transit of lysosomal enzymes through the endoplasmic reticulum (ER)–Golgi biosynthetic pathway (Figure 10.1A). The first enzyme, UDP-*N*-acetylglucosamine:lysosomal enzyme *N*-acetylglucosamine-1-phosphotransferase (phosphotransferase; IUBMB accession number EC 2.7.8.17), attaches *N*-acetylglucosamine (GlcNAc) 1-phosphate to the C-6 hydroxyl group of one or more mannose residues to form the Man-6-P-OGlcNAc phosphodiester intermediate [9–12] (Figure 10.1B). This selective phosphorylation of N-linked high-mannose-type oligosaccharides is achieved by the ability of the phosphotransferase to recognize a surface patch on lysosomal enzymes, with two or more lysine residues that are correctly spaced relative to each other and to the oligosaccharide chain serving as critical elements of a more extensive three-dimensional recognition marker [13–15]. The second enzyme, *N*-acetylglucosamine-1-phosphodiesterα-*N*-acetylglucosaminidase, which is often referred to as the "uncovering enzyme" (IUBMB accession number EC 3.1.4.45), removes the GlcNAc moiety in the *trans*-Golgi network (TGN) to generate the phosphomonoester [16–18] (Figure 10.1C). The uncovering enzyme is synthesized as a proenzyme with little or no enzymatic activity and the removal of the propeptide by furin in the TGN has been shown to be essential for the generation of an active enzyme [19,20]. The MPRs encounter a diverse array of Man-6-P-tagged lysosomal

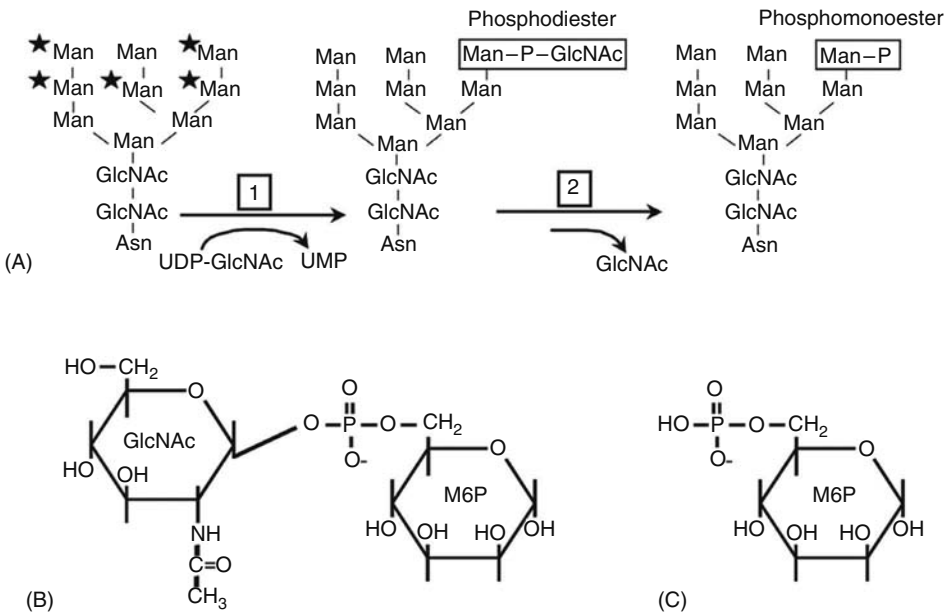


FIGURE 10.1 Generation of the Man-6-P tag on N-linked oligosaccharides. (A) Phosphorylation of mannose residues on N-linked oligosaccharides occurs in two steps. First, the phosphotransferase transfers GlcNAc-1-phosphate from UDP-GlcNAc to the C-6 hydroxyl group of mannose to form the Man-6-P-OGlcNAc phosphodiester intermediate. Second, the uncovering enzyme removes the GlcNAc moiety in the TGN, revealing the Man-6-P phosphomonoester. The stars indicate the five potential sites of phosphorylation. (B) Structure of Man-6-P-OGlcNAc phosphodiester. (C) Structure of Man-6-P phosphomonoester.

enzymes in the TGN since phosphorylated oligosaccharides isolated from lysosomal enzymes have been shown to be quite heterogeneous, containing one or two phosphomannosyl residues (phosphomonoester or GlcNAc phosphodiester) that can be located at five different positions in the oligosaccharide chain (Figure 10.1A) [9–11].

Genetic defects in the phosphotransferase cause the autosomal recessive LSD, mucolipidosis II (ML II; also referred to as "I-cell disease"), and mucolipidosis III (ML III; also referred to as "pseudo-Hurler polydystrophy") in which the activity of the enzyme is absent or reduced, respectively [21]. Unlike LSDs affecting single enzymes in a catabolic pathway, these disorders are characterized by impaired sorting of multiple newly synthesized enzymes to lysosomes and their subsequent secretion from cells due to the inability of the enzymes to acquire the Man-6-P tag and be recognized by the MPRs. ML II and ML III share similar clinical features. However, ML II is more severe, often noted shortly after birth whereas ML III has a later onset and a slower course. Consequently, ML II patients typically die during the first decade of life due to cardiovascular system dysfunction resulting from progressive storage of undigested material in the lysosomes of fibroblasts in the connective tissue of the heart valves, endocardium, myocardium, and perivascular areas [22].

The discovery of the receptors that mediate lysosomal enzyme targeting arose from investigations into the molecular basis of ML II. Hickman and Neufeld made key observations that ML II fibroblasts were capable of endocytosing lysosomal enzymes secreted by normal cells while, in contrast, normal cells were incapable of internalizing the enzymes secreted by ML II fibroblasts [5]. Their suggestion that lysosomal enzymes contained a recognition marker required for uptake and transport to lysosomes was later confirmed with the identification of the marker as Man-6-P [23–25]. Compared to ML II patients, transgenic mice lacking both MPRs demonstrate a more severe phenotype of embryonic lethality that has been attributed to the loss of insulin-like growth factor II (IGF-II) regulation resulting from the absence of the CI-MPR [26–29]. Analyses of fibroblasts derived from these transgenic animals show that these cells secrete the majority of their lysosomal enzymes and accumulate undigested material in their lysosomes in a manner similar to that observed in fibroblasts from ML II patients [30,31], demonstrating that a similar phenotype can result from either a deficiency of the Man-6-P tag or the MPRs.

10.3 PRIMARY STRUCTURE OF THE MPRs

The P-type lectins are type I transmembrane glycoproteins that exist as oligomers: the cation-dependent mannose-6-phosphate receptor (CD-MPR) is a stable homodimer and the cation-independent mannose-6-phosphate receptor (CI-MPR) most likely exists in the form of a dimer (see Ref. [32] for a review) (Figure 10.2). The bovine CD-MPR is composed of a 28-residue amino-terminal signal sequence, a 159-residue extracytoplasmic region, a 25-residue transmembrane region, and a 67-residue carboxyl-terminal cytoplasmic domain. The extracytoplasmic region of the CD-MPR contains six cysteine residues that are involved in the formation of three intramolecular disulfide bonds that play an essential role in the folding of the receptor [33]. The CI-MPR contains a 44-residue amino-terminal signal sequence, a large 2269-residue extracytoplasmic region, a 23-residue transmembrane region, and a 163-residue carboxyl-terminal cytoplasmic domain. The large extracytoplasmic region is composed of 15 contiguous domains that display a similar size (~150 residues) and cysteine distribution, and exhibit significant amino acid identity (14%–38%) when compared to each other and to the CD-MPR, giving rise to the prediction that they have a similar tertiary structure. This hypothesis has been confirmed, in part, by crystal structure determinations, which show that the extracytoplasmic region of the CD-MPR and domains 1, 2, 3, and 11 of the CI-MPR all exhibit the same fold (see below). The MPRs undergo several types of co- and posttranslational modifications. The CD-MPR contains five potential N-glycosylation sites, four of which are utilized. The bovine CI-MPR contains 19 potential N-glycosylation sites,

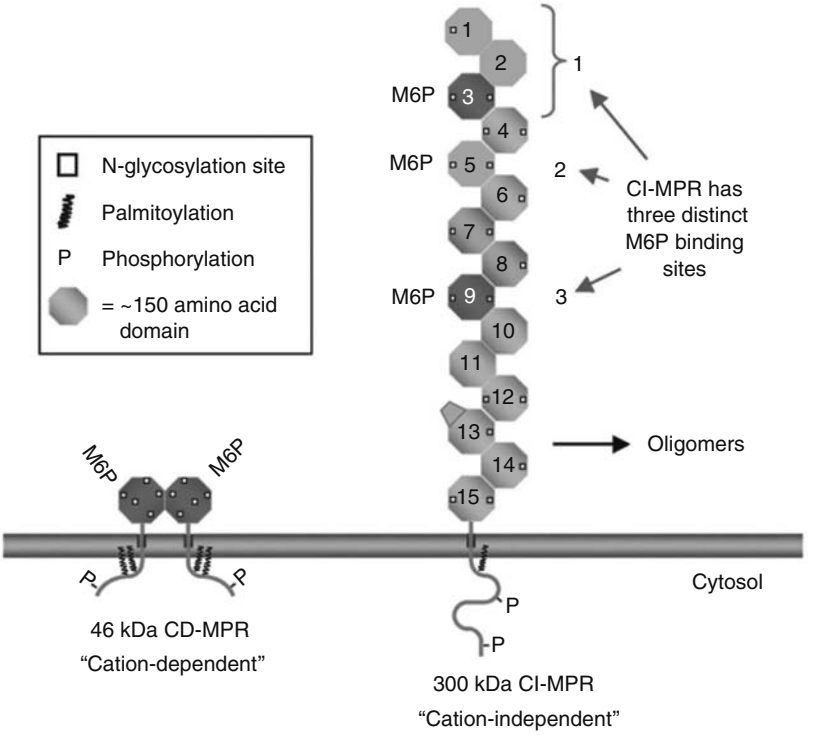


FIGURE 10.2 (See CD for color figure.) Schematic diagram of the CD-MPR and CI-MPR. The MPRs are type I transmembrane glycoproteins. The CD-MPR exists as a homodimer. The CI-MPR also undergoes oligomerization and most likely exists as a dimer. The extracytoplasmic region of the CI-MPR is comprised of 15 contiguous domains, each approximately 150 residues in length, which exhibit 14%–38% amino acid identity when compared to each other or to the extracytoplasmic domain of the CD-MPR. The Man-6-P-binding sites are highlighted in pink. The CD-MPR contains a single high-affinity Man-6-P binding site per polypeptide. In contrast, the CI-MPR contains three carbohydrate recognition sites: two high affinity sites are localized to domains 1–3 and domain 9, and one low affinity site is contained within domain 5.

and it is likely that the majority of these sites are modified based on a comparison of glycosylated versus nonglycosylated species by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis [34,35]. Like the CD-MPR, the cytoplasmic region of the CI-MPR undergoes phosphorylation and palmitoylation. Palmitoylation of the CD-MPR has been shown to prevent its degradation in lysosomes [36] while serine phosphorylation of both MPRs has been shown to influence their intracellular transport [37–39].

10.4 TRAFFICKING OF THE MPRs AND THEIR CARGO

The MPRs cycle between the TGN, endosomes, and the plasma membrane (Figure 10.3) and recent live video microscopy using fluorescently tagged proteins, such as green fluorescent protein fused to the CI-MPR [40,41], has highlighted the dynamic nature of this sorting process. Although the MPRs direct newly synthesized soluble acid hydrolases to the lysosome, the receptors are not detected in lysosomal compartments. In contrast to many endocytic receptors, only ~10% of the MPRs are present at the cell surface while the remainder of the receptors are found predominantly in endosomal compartments and the TGN. Numerous studies have provided insight into the complex intracellular pathways traversed by the P-type lectins (for reviews, see Refs. [42–44]). A brief overview of these pathways and their associated signals and cytosolic transport factors is provided in the following sections.

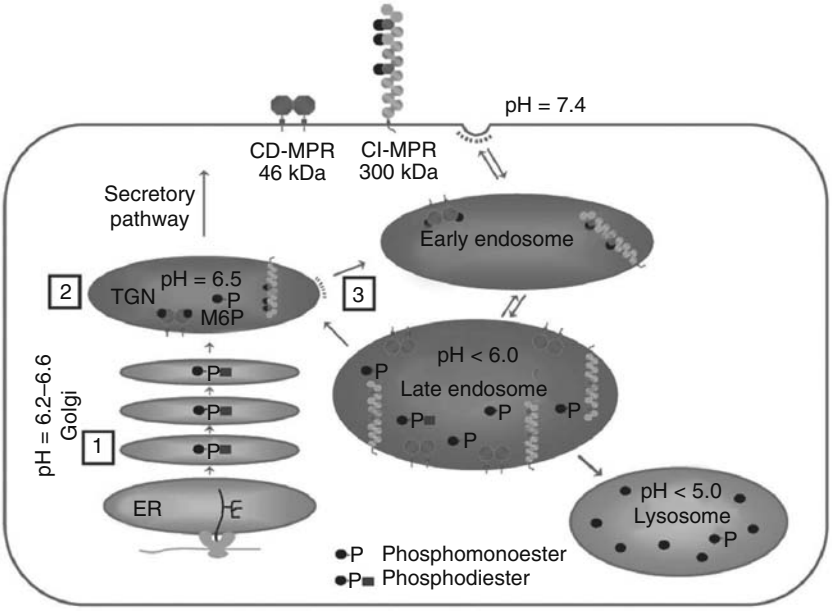


FIGURE 10.3 (See color insert following blank page 170. Also see CD for color figure.) Subcellular localization and trafficking of the MPRs. Delivery of newly synthesized lysosomal enzymes to lysosomes requires three major events. First, the phosphotransferase distinguishes lysosomal enzymes from other proteins in the Golgi and attaches a phosphorylated GlcNAc to mannose residues on their N-linked oligosaccharides. Second, the uncovering enzyme acts in the TGN to remove the GlcNAc residue, generating the Man-6-P phosphomonoester. Third, the MPRs bind the Man-6-P-tagged lysosomal enzymes in the TGN and divert them from the secretory pathway. The enzyme-bound receptors are concentrated into clathrin-coated vesicles and transported from the TGN to endosomal compartments. Within the acidic pH of the late endosomal compartment, the lysosomal enzymes dissociate from the MPRs and are subsequently packaged into lysosomes where they can be dephoshorylated by the action of lysosomal phosphatases. The MPRs recycle back to the TGN to repeat the process or move to the cell surface where the CI-MPR, but not the CD-MPR functions to internalize exogenous ligands. The MPRs cycle continuously between intracellular compartments (TGN, endosomes) and the plasma membrane (blue shaded organelles). At steady state, little or no receptors are detected in the ER, Golgi stacks, or lysosome (gray shaded organelles). Clathrin (red dots).

10.4.1 Intracellular Ligands and the "Synthetic Pathway"

Recognition of Man-6-P-tagged acid hydrolases by the P-type lectins occurs in the TGN. The resulting lysosomal enzyme–receptor complexes are subsequently transported from the TGN to endosomal compartments via clathrin-coated vesicles (Figure 10.3). This process selectively sorts lysosomal enzymes from literally thousands of other proteins within the secretory pathway. The localization and trafficking of the MPRs relies upon canonical sorting motifs within their cytosolic domains, which are recognized by components of the vesicular trafficking machinery. However, there is some evidence to suggest that the lumenal domain of the CI-MPR also influences its intracellular sorting [41,45,46]. Although a role for the clathrin adaptor protein 1 (AP-1) in the transport of MPRs from the TGN to endosomes has not been ruled out [47], significant evidence exists that members of the Golgi-localized, Gamma-ear-containing, ADP-ribosylation (GGA) factor-binding protein family, which comprise three members in mammals (GGA1, GGA2, and GGA3), sort the MPRs into clathrin-coated vesicles at the TGN and may also mediate MPR trafficking between early and late endosomal compartments [48]. In addition, direct interaction between an acidic cluster–dileucine motif located near the C-terminus of the MPRs and the VHS (VPS27/Hrs/STAM) domain of the GGA family members has been shown [49–52].

Within the acidic pH (<6.0) environment of late endosomal compartments, the lysosomal enzymes are released from the MPRs and are packaged into lysosomes by an unknown mechanism [42,53]. The MPRs do not travel to lysosomes but recycle from endosomal compartments back to the TGN to retrieve additional enzymes from the secretory pathway (Figure 10.3). This endosome–TGN retrieval pathway, which does not appear to require clathrin, is mediated by a multiprotein complex termed "retromer" [54–56]. Other sorting molecules, phosphofurin acidic cluster sorting protein-1 (PACS-1), which binds to the C-terminal cluster of acidic residues in the cytosolic tail of the CI-MPR in addition to interacting with the AP-1 [57–59], tail-interacting protein of 47 kDa (TIP47) [60,61], and EpsinR [62] have also been implicated in mediating the endosome-to-TGN retrieval of the CI-MPR. PACS-1/AP-1 likely participates in receptor recycling from early endosomes whereas TIP47/Rab9 facilitates receptor recycling from late endosomes. Recent studies have provided additional support for a regulatory role of phosphorylation in the complex itinerary of the MPRs [63–65].

10.4.2 Extracellular Ligands and the "Recapture Pathway"

Cell surface CI-MPRs, but not CD-MPRs, carry out the internalization of a variety of Man-6-P-containing ligands for their subsequent clearance or activation (see below). With respect to lysosomal enzymes, cell surface CI-MPRs function in the endocytic recapture of Man-6-P-modified acid hydrolases that are not properly sorted from the secretory pathway at the TGN [30,66–68]. This property of the CI-MPR forms the basis of enzyme replacement therapy for several LSDs (see above) and explains, in part, the observation that the CI-MPR is more efficient than the CD-MPR in targeting lysosomal enzymes to the lysosome [68,69]. Although the CD-MPR cycles between the cell surface and intracellular compartments, it does not play a significant role in the recapture of secreted lysosomal enzymes due to its decreased ability to bind lysosomal enzymes at pH 7.4 [70]. Endocytosis of both MPRs from the plasma membrane involves formation of clathrin-coated vesicles mediated by AP-2 [71,72]. A single tyrosine-based internalization sequence, YKYSKV, has been identified in the cytoplasmic domain of the CI-MPR that is recognized by AP-2 [73–75]. In contrast, the cytosolic tail of the human CD-MPR contains three separate internalization sequences: A phenylalanine-containing sequence (AKGMEQF), a tyrosine-based motif (YRGV), and the C-terminal dileucine motif that also influences sorting of the receptor at the TGN and endosomes [76].

10.5 EXPANDING LIST OF MAN-6-P-CONTAINING PROTEINS

In addition to lysosomal enzymes, the repertoire of identified extracellular ligands of the CI-MPR has expanded to include a diverse spectrum of Man-6-P-containing proteins [32]. For example, two viruses, herpes simplex virus [77] and varicella-zoster virus [78], have been shown to express viral Man-6-P-containing glycoproteins, which function to facilitate the entry of the virus into mammalian cells via the CI-MPRs. Elegant studies by Chen et al. [79] have clarified the pathology of varicella-zoster virus infection. Intracellular CI-MPR diverts newly enveloped varicella-zoster virus to late endosomes, thereby preventing the spread of the virus. However, CI-MPR expression is lost in maturing superficial epidermal cells and thus these cells do not divert the virus to endosomes but rather constitutively secrete infectious virus particles. With respect to endogenous mammalian proteins, a number of secreted proteins have been shown to contain Man-6-P, such as transforming growth factor (TGF) precursor [80], the cytokine leukemia inhibitory factor (LIF) [81], the placental angiogenic hormone proliferin [82], the aspartic protease renin precursor [83], the serine proteases granzymes A and B [84], the T cell activation antigen CD26 [85], and the cysteine proteinase inhibitor cystatin F [86]. Roles for the interaction of these nonlysosomal proteins with cell surface CI-MPRs include activation (e.g., TGF- precursor [87], renin precursor [88]) and clearance from the plasma (e.g., LIF [89]). Recent proteomic approaches using affinity columns containing immobilized CI-MPR have identified new Man-6-P-containing lysosomal proteins [90,91]. In addition, this

methodology has demonstrated that a small proportion of abundant classical plasma proteins (e.g., α1-acid glycoprotein, ceruloplasmin, haptoglobin) exist as Man-6-P-containing glycoforms [91]. Although not directly demonstrated, it is assumed that these nonlysosomal proteins acquire Man-6-P by the same phosphotransferase that acts on acid hydrolases. In addition, it is not clear how these Man-6-P-containing proteins can escape interaction with the MPRs in the TGN and are secreted rather than being diverted to the endosome/lysosome system. Clearly, additional studies are needed to address these questions.

10.6 CARBOHYDRATE-BINDING PROPERTIES OF THE MPRs

Soluble acid hydrolases constitute a heterogeneous population of over 50 enzymes that differ in size, oligomeric state, number of N-linked oligosaccharides, extent of phosphorylation, and the position of the Man-6-P moiety and its linkage to the penultimate mannose residue in the oligosaccharide chain (Figure 10.1A). Due to the presence of both MPRs in most cell types, the relative contribution of each MPR to the targeting of this diverse population of enzymes to the lysosome has been examined. The two MPRs display different affinities and capacities for transport of the various acid hydrolases, and studies utilizing receptor-deficient fibroblasts demonstrate that both receptors are necessary for the efficient sorting of all lysosomal enzymes to the lysosome as neither MPR can fully compensate for the other [66–68,92].

The CD-MPR and CI-MPR share a number of similarities with respect to carbohydrate recognition. For example, both MPRs bind the monosaccharide Man-6-P with essentially the same affinity ($7–8 \times 10^{-6}$ M) [70,93]. The axial 2-hydroxyl group and the 6-phosphate monoester group are major determinants of binding specificity based on the observation that mannose or glucose 6-phosphate interact poorly with the MPRs ($K_i = 1–5 \times 10^{-2}$ M) [70]. Inhibition studies using chemically synthesized oligomannosides or neoglycoproteins demonstrated that the presence of the phosphomonoester Man-6-P at a terminal position is the major determinant of receptor binding. In addition, linear mannose sequences, which contain a terminal Man-6-P linked α1,2 to the penultimate mannose, were shown to be the most potent inhibitors [94,95], suggesting that the MPRs bind an extended oligosaccharide structure, which includes the Man-6-P α1,2 Man sequence. Furthermore, multivalent interactions between the receptor and a lysosomal enzyme result in high-affinity binding, typically on the order of 1–10 nM for both MPRs [69,70,92].

In contrast to these similarities, the two MPRs exhibit a number of differences in their binding properties, which include pH dependence, cation dependence, and recognition of phosphodiesters. The two MPRs display optimal ligand binding at ~pH 6.4 and no detectable binding below pH 5, which is consistent with their function of releasing ligands in the acidic environment of the endosome. The CI-MPR retains phosphomannosyl binding capabilities at neutral pH, which corresponds well with the ability of this receptor to bind and internalize lysosomal enzymes at the cell surface. In contrast, the ligand-binding ability of the CD-MPR is dramatically reduced at a pH > 6.4 [70,93], which is consistent with its decreased ability to bind and internalize lysosomal enzymes at the cell surface [96]. The inability to purify the CD-MPR by phosphomannosyl affinity chromatography performed in the absence of cations led to its designation as a "cation-dependent" receptor [97]. However, the presence of cations increases the binding affinity of the CD-MPR toward Man-6-P [70] and lysosomal enzymes only four-fold [98] but has no effect on the binding affinity of the CI-MPR. This finding differentiates the CD-MPR from C-type lectins, which have an absolute requirement for calcium to carry out their sugar-binding activities [99]. Mutagenesis studies [98] evaluated in the context of the crystal structure [100,101] indicate that a conserved aspartic acid residue at position 103 of the CD-MPR, which is not present in the CI-MPR, necessitates the presence of a divalent cation in the binding pocket to obtain high-affinity ligand binding by functioning to neutralize the negative charge of Asp-103 juxtaposed to the phosphate oxygen of Man-6-P. The CI-MPR, unlike the CD-MPR, is able to recognize Man-6-P-OGlcNAc phosphodiesters [70,93,94] as well as lysosomal enzymes derived from *Dictyostelium discoideum*, which contain

mannose-6-sulfate residues and small methyl phosphodiester, Man-6-P-OCH_3, but not phosphomonoesters [102,103]. Furthermore, the MPRs differ in the number of Man-6-P binding sites contained within their polypeptide chain, with the CD-MPR containing one [70] and the CI-MPR containing three Man-6-P binding sites [104].

Expression of recombinant truncated forms of the CI-MPR has mapped its three carbohydrate-binding sites to discrete regions within its large extracellular region: Two high-affinity sites (K_i = ~10 μ M for Man-6-P) map to domains 1–3 and domain 9 [105] while domain 5 houses a low-affinity (K_i = ~5 mM for Man-6-P) binding site [104] (Figure 10.2). A comparison of the binding properties of the individual carbohydrate recognition sites demonstrated that domain 9 of the CI-MPR exhibits optimal binding at pH 6.4–6.5, similar to that of the CD-MPR. In contrast, the N-terminal Man-6-P binding site (i.e., domains 1–3) has a significantly higher optimal binding pH of 6.9–7.0 [106]. This observation may not only explain the relatively broad pH range of ligand binding by the CI-MPR but also likely is a main contributor to the ability of the CI-MPR, as opposed to the CD-MPR, to internalize exogenous ligands at the slightly alkaline pH 7.4 present at the cell surface. Domain 9 of the CI-MPR, like the CD-MPR, is highly specific for phosphomonoesters. In contrast, the N-terminal carbohydrate recognition site of the CI-MPR is promiscuous in that, in addition to Man-6-P, it is able to efficiently bind the Man-6-P-OCH_3 phosphodiester and mannose 6-sulfate [106]. Taken together, the presence of three distinct carbohydrate recognition sites in the CI-MPR likely accounts for the ability of the CI-MPR to recognize a greater diversity of ligands than the CD-MPR both *in vitro* [92,107] and *in vivo* [30,66,67,108].

10.7 CRYSTAL STRUCTURE OF TRUNCATED FORMS OF THE MPRs

To date, the crystal structure of the extracytoplasmic region of the CD-MPR bound to either a single sugar, Man-6-P [100], or to an oligosaccharide, pentamannosyl phosphate [101], or in the unbound state [109] has been determined. Four out of the 15 domains of the CI-MPR have also been crystallized: the N-terminal 432 residues (domains 1, 2, and 3), which houses a high-affinity Man-6-P binding site, bound to either a mannose residue from a crystallographic neighbor [110] or Man-6-P [111]; the IGF-II binding site (domain 11) in the unbound state [112,113]. These structures have provided a framework from which the mechanism of ligand binding by the MPRs can be inferred.

The extracytoplasmic domain of the CD-MPR is crystallized as a dimer (Figure 10.4A), which is consistent with previous biochemical data. Each polypeptide chain folds into an N-terminal helix followed by four antiparallel β-strands, which together comprise the solvent-exposed front face. The dimer interface sheet (β5–β9), which accounts for approximately 20% of the surface area of the monomer, is composed of five antiparallel β-strands, with strand 9 interjecting between strands 7 and 8 (Figure 10.4C). Subsequent determination of the structure of domains 1–3 (Figure 10.4B) and domain 11 of the CI-MPR showed this overall topology is conserved with the exception of the N-terminal region: Neither domains 1–3 nor domain 11 contains the α-helix; rather, this secondary structural element is replaced by two β-strands (Figure 10.4C). The quaternary domain arrangement of CD-MPR is not conserved in the structure of domains 1–3 of the CI-MPR. The structure of the N-terminal region of the CI-MPR shows the three domains form a wedge with domains 1 and 2 oriented such that the four-stranded N-terminal β–sheet (β1–β4) of domain 1 and the five-stranded C-terminal β–sheet (β5–β9) of domain 2 form a continuous surface (Figure 10.4B). In comparison to the CD-MPR in which extensive contacts exist between the two dimer interface β-sheets (β5–β9) (Figure 10.4A), the interaction between the three N-terminal domains of the CI-MPR is quite different and is much less extensive: The contacts between the three domains are mediated mainly by residues within the linker regions and loops (Figure 10.4B). However, the contacts between domains 1, 2, and 3 are important for maintaining the integrity of the binding pocket housed within domain 3. The multiple interactions between residues of domains 1 and 2 with residues of loops C and D of domain 3 are likely to aid in the stabilization of the binding pocket and provide an explanation for the inability of a construct-encoding domain 3 alone to generate a high-affinity carbohydrate binding site [105].

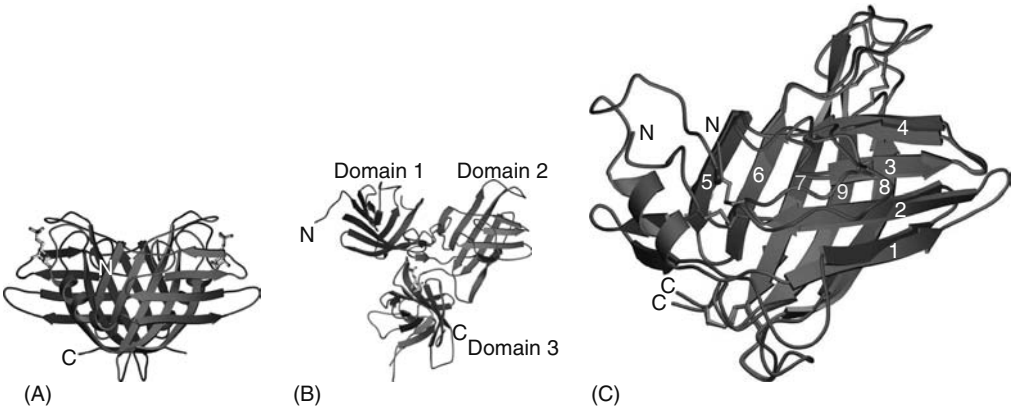
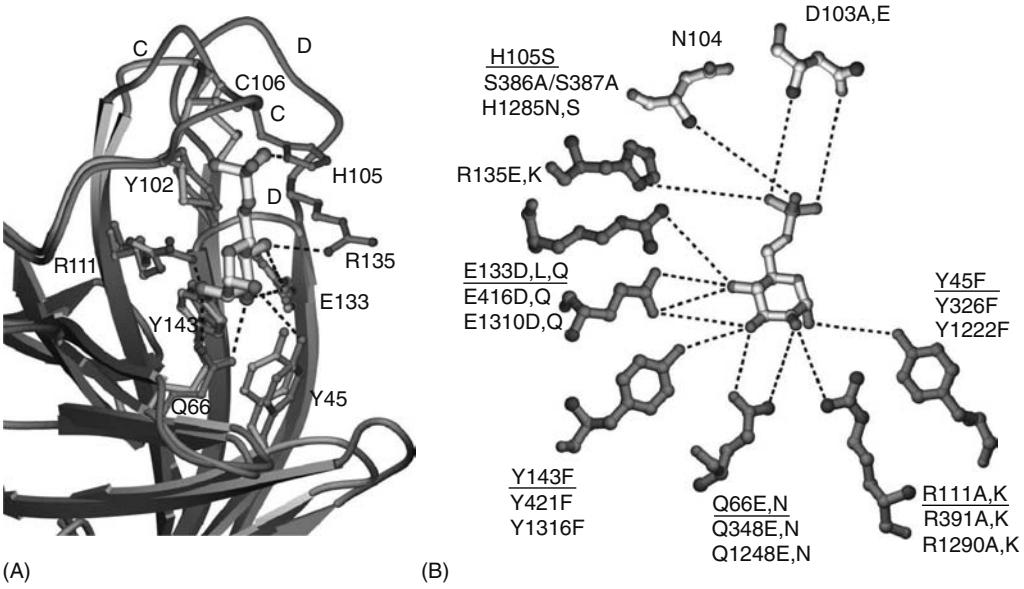


FIGURE 10.4 (See color insert following blank page 170. Also see CD for color figure.) Three-dimensional structure of the CD-MPR and domains 1–3 of the CI-MPR. (A) Crystal structure of the extracytoplasmic region (residues 3–154) of the bovine CD-MPR in the presence of an oligosaccharide, pentamannosyl phosphate (PDB 1C39). Note that only the terminal Man-6-P (gold ball-and-stick model) is shown for clarity. Both monomers (light purple and dark purple) of the CD-MPR dimer are shown in this ribbon diagram. The location of the N-terminus (N) and C-terminus (C) of one monomer are indicated. (B) Crystal structure of the N-terminal three domains (residues 7–432) of the bovine CI-MPR (PDB 1SZO). The N- and C-terminus of the protein-encoding domain 1 (blue), domain 2 (pink), and domain 3 (green) are indicated. The location of Man-6-P (gold ball-and-stick model) is shown. (C) Overlay of the structures of the CD-MPR (purple) and domain 3 (green) of the CI-MPR. The strands are sequentially numbered. The disulfide bridges are shown in gold, and the N- and C-terminus are indicated.

A comparison of the sugar-binding pocket of the CD-MPR and domain 3 of the CI-MPR reveals that residues, which interact with the mannose ring (Gln, Arg, Glu, and Tyr), are located in a strikingly similar position in the "base" of the pocket (Figure 10.5A) and form the same contacts with the ligand (Figure 10.5A and B). These four residues have been shown to be essential for Man-6-P recognition by mutagenesis studies (Figure 10.5B), and are conserved in all species and in the other two Man-6-P binding sites of the CI-MPR (i.e., domains 5 and 9) (Figure 10.5C). The presence of this "signature motif" for phosphomannosyl binding (Gln, Arg, Glu, and Tyr) along with conserved cysteine residues (Figure 10.5C) allowed for the prediction that domain 5 of the CI-MPR contains a Man-6-P binding site, a hypothesis which was subsequently confirmed [104]. Furthermore, mutation of Gln, Arg, Glu, and Tyr in domain 9 demonstrates their essential role in carbohydrate recognition by this binding site [114] (Figure 10.5B). Thus, the strict requirement for a terminal mannose residue by both receptors is reflected in the similarities in that region of the binding pocket responsible for sugar recognition.

In contrast, the phosphate recognition region (lid) of the binding site appears to have the most variability both in amino acid composition and structure. In both receptors, the lid is formed by residues joining β-strands 6 and 7 (loop C). This lid region is larger in the CD-MPR and the positioning of the disulfide anchors the loop in a more closed position, which translates into a more sterically confined binding region (Figure 10.5A). The conformationally constrained lid may account for the inability of CD-MPR to bind phosphodiesters. Shortening of both loops C and D effectively makes the binding pocket of domains 1–3 more open than that of the CD-MPR (Figure 10.5A), allowing for this region of the CI-MPR to bind a larger repertoire of ligands, including phosphomonoesters, mannose 6-sulfate, and phosphodiesters. Thus, the diversity in ligand recognition by the two receptors appears to be accomplished by alterations in the receptor binding site architecture surrounding the phosphate moiety.



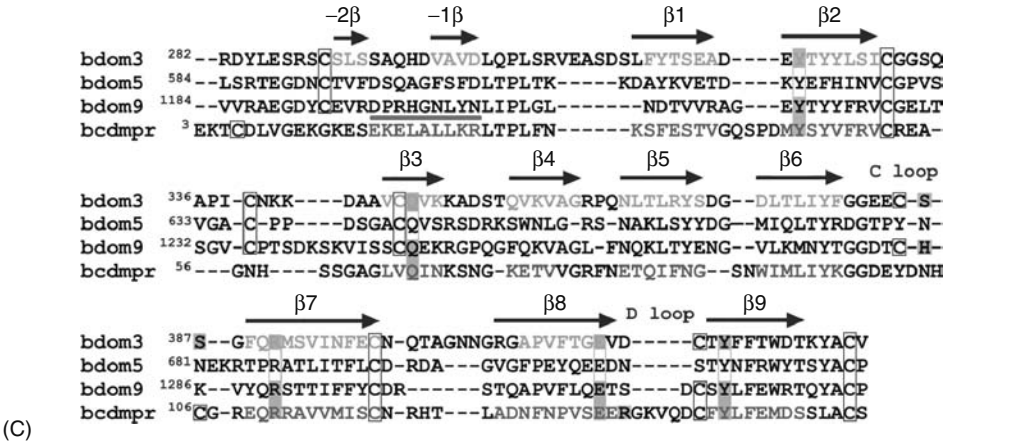


FIGURE 10.5 (See CD for color figure.) Comparison of the carbohydrate-binding pocket of the CD-MPR and domain 3 of the CI-MPR. (A) Ribbon diagram showing the binding site of the CD-MPR (purple) superimposed onto domain 3 (green) of the CI-MPR. Loops C and D are labeled. The disulfides are shown in gold and Man-6-P (gold ball-and-stick model) is also indicated. (B) Schematic view of the potential hydrogen bond and ionic interactions between the binding pocket residues of the CD-MPR and Man-6-P (gold ball-and-stick). Directly beneath the CD-MPR designations (underlined) are listed the corresponding carbohydrate-binding site residues of domains 3 and 9 of the CI-MPR. Shaded in light gray are the residues that have not been tested (black amino acid label) or when mutated (D103) retained wild-type Man-6-P binding ability. Shaded in dark gray are the residues which when mutated resulted in receptors with diminished (~50–150-fold) Man-6-P-binding ability as compared to wild-type receptors. Shaded in blue are the residues identified as essential for carbohydrate recognition by the MPRs (i.e., single amino acid substitution of these residues abolished (~1000-fold) the Man-6-P-binding ability of the receptors). (From Sun, G., Zhao, H., Kalyanaraman, B., and Dahms, N.M., *Glycobiology*, 15, 1136, 2005; Hancock, M.K., Haskins, D.J., Sun, G., and Dahms, N.M., *J. Biol. Chem.*, 277, 11255, 2002; Olson, L.J., Hancock, M.K., Dix, D., Kim, J.-J.P., and Dahms, N.M., *J. Biol. Chem.*, 274, 36905, 1999; Wendland, M., Waheed, A., von Figura, K., and Pohlmann, R., *J. Biol. Chem.*, 266, 2917, 1991.) (C) Structure-based amino acid sequence alignment of the extracytoplasmic region of the CD-MPR and domains 3, 5, and 9 of the CI-MPR. The secondary structure of domain 3 of the CI-MPR and the CD-MPR are shown, with arrows representing the β-strands and the cylinder representing the single α-helix. The cysteine residues are boxed. Residues that have been subjected to site-directed mutagenesis are shaded and color-coded as in panel B. Domain 5 of the CI-MPR has the conserved residues (open color-coded boxes), but they have not yet been mutated.

Unlike what has been observed in other lectins, the crystal structure of the ligand-free CD-MPR differs considerably from the ligand-bound form in that changes in both quaternary structure and positioning of loops involved in sugar binding are seen, along with changes in the spacing of the two carbohydrate binding sites in the dimeric receptor [109]. These results indicate that the CD-MPR protein is dynamic. It is intriguing to speculate that the receptor can alter the spacing of its two sugar binding sites in order to accommodate the diverse spatial array of phosphomannosyl-containing oligosaccharides found on the approximately 50 different lysosomal enzymes it encounters in the TGN.

10.8 EVOLUTION OF THE MPRs AND THE MANNOSE-6-PHOSPHATE RECEPTOR HOMOLOGY PROTEIN FAMILY

The Man-6-P-based system for targeting lysosomal hydrolases to lysosomes is conserved in mammals, birds, amphibians, and crustaceans but is absent in the unicellular protozoa *Trypanosoma* [115] and Leishmania [116]. *D. discoideum*, and *Acanthamoeba castellani* both exhibit phosphotransferase activity and can transfer GlcNAc-1-PO_4 to mannose residues [117]. However, MPRs have not been identified in these species. The recently reported sequence of the zebrafish (*Danio rerio*) CD-MPR and CI-MPR [118] indicates that targeting of lysosomal enzymes by MPRs represents an ancient pathway in vertebrate cell biology. The CD-MPR [119] and CI-MPR [120,121] have also been reported in the invertebrate mollusc *Unio*. Taken together, these studies and others demonstrate that numerous species throughout the animal kingdom express bona fide MPRs that are capable of binding phosphomannosyl residues.

The existence of a Man-6-P-dependent transport pathway for lysosomal enzymes in insects has been unclear since Man-6-P has been reported on a lysosomal enzyme (DNase I) derived from *Drosophila melanogaster* [122] whereas Man-6-P-containing oligosaccharides or phosphotransferase activity were not detected in the Sf9 insect cell line [123]. Recently, a *Drosophila* protein (lysosomal enzyme receptor protein [LERP]) that is related to the mammalian CI-MPR has been identified [124]. The amino acid sequence predicts a type I transmembrane glycoprotein containing five contiguous domains, each about 155 residues in length, in its lumenal region that correspond to domains 9–13 of the CI-MPR. The authors showed that LERP, which is able to interact with *Drosophila* and mammalian GGA adaptors, mediates lysosomal enzyme targeting and rescues the missorting of lysosomal enzymes that occurs in MPR-deficient mammalian cells. Although the nature of the interaction between *Drosophila* LERP and mammalian lysosomal enzymes has not yet been elucidated, it is unlikely to involve Man-6-P since the residues that are essential for Man-6-P recognition are not conserved in the *Drosophila* protein and no detectable binding was observed between LERP and a Man-6-P-containing affinity resin (phosphomannan). It is intriguing to speculate that LERP may represent the evolutionary intermediate between yeast and the animal kingdom: Targeting of hydrolytic enzymes to the yeast vacuole (functional equivalent of the lysosome) occurs in a Man-6-P-independent fashion, with a receptor (vps10) that recognizes a protein determinant, rather than a carbohydrate moiety, on hydrolytic enzymes [125].

In addition to the CD-MPR and CI-MPR, there are four other genes in the human genome that contain what has been termed "mannose 6-phosphate receptor homology" (MRH) domains: (1) the β-subunit of glucosidase II is the noncatalytic subunit of a dimeric ER-resident enzyme involved in the processing of N-glycans on nascent glycoproteins [126]; (2) the γ-subunit of the phosphotransferase is the noncatalytic subunit of the $\alpha_2\beta_2\gamma_2$ hexameric complex involved in generating the Man-6-P tag on lysosomal enzymes [126]; (3) erlectin is a lumenal ER-resident protein that appears to recognize N-glycans and may function as an ER chaperone [126,127]; (4) OS-9, originally identified as a protein upregulated in human osteosarcomas [128], is an ER-associated cytosolic protein implicated in the transport of proteins from the ER to the Golgi [126,129]. Recent studies on the Yos9 protein, the *Saccharomyces cerevisiae* homolog of OS-9, indicate that this ER-resident protein plays an essential role in the recognition of misfolded glycoproteins during ER-associated

degradation [130–132]. All of the MRH-containing proteins contain a single MRH domain, except erlectin, which contains two MRH domains. The carbohydrate-binding properties of these MRH-containing proteins have not been characterized. However, since their MRH domains contain the four conserved residues (Glu, Arg, Gln, Tyr) shown to interact with the 2-, 3-, and 4-hydroxyl groups of the mannose ring of Man-6-P by the MPRs, it is likely that these proteins bind specifically to high mannose-type oligosaccharides, which would be consistent with their proposed functions in the ER and early Golgi compartments.

Using an approach in which sequence databases are screened with various carbohydrate recognition domain profiles to identify proteins containing potential lectin activity, Dodd and Drickamer reported five proteins, three from *D. melanogaster*, one from *Caenorhabditis elegans*, and one from *S. cerevisiae*, which contain P-type lectin-like domains [133]. However, since these proteins lack most of the residues found in the Man-6-P binding pocket, it is unlikely that these proteins bind carbohydrate. Rather, these proteins may utilize the MPR fold for protein–protein interactions in a manner similar to that of domain 11 of the CI-MPR, which binds the mitogenic polypeptide, IGF-II, rather than Man-6-P.

10.9 CONCLUSION AND FUTURE DIRECTIONS

Although the first description of a LSD (Tay-Sachs disease) was reported in 1881 [134], the discovery of the lysosome by Christian de Duve did not occur until 1955 [135]. Soon after, Hers was able to demonstrate the first link between an enzyme deficiency and a LSD (Pompe disease) [136]. This key observation led the way for a series of seminal findings about the biology of these enzymes and the receptors that mediate their intracellular trafficking. Cell biological, biochemical, and biophysical studies have made significant contributions to our understanding of the molecular basis governing the intracellular transport of the MPRs and their mode of carbohydrate recognition. However, many important questions remain unanswered. It is essential for the MPRs to release their ligands in the acidic environment of endosomes in order to be able to recycle back to the TGN to retrieve additional lysosomal enzymes; yet the mechanism by which changes in pH affect the carbohydrate-binding pockets of the MPRs is unknown. All of the functional domains of the CI-MPR have been mapped to odd-numbered domains (e.g., domain 1, plasminogen; domains 3, 5, and 9, Man-6-P; domain 11, IGF-II). However, the role of the remaining 10 unassigned domains is unclear, as is the manner in which the MPRs acquired Man-6-P binding capabilities during evolution. The presence of a MRH domain in the phosphotransferase raises the interesting possibility that the proteins involved in the synthesis and recognition of Man-6-P evolved together from a common ancestor. The observation that domain 9 can be expressed as an individual domain, retaining high-affinity binding capabilities [105], indicates its mechanism of maintaining and stabilizing its binding pocket must be fundamentally different from that used by domains 1–3 of the CI-MPR. Therefore, structural studies of the CI-MPR to determine the mechanism of carbohydrate binding by domain 5 and domain 9, the arrangement of all 15 domains within its extracellular region and their relative degree of flexibility, and its mode of oligomerization are essential in order to understand how this large receptor functions to target a diverse array of ligands to the lysosome.

ACKNOWLEDGMENTS

We apologize to the authors of publications not cited due to space limitations. This work was supported by NIH grant DK42667 to N.M.D and J.J. P.K.

REFERENCES

1. Neufeld EF. 1991. Lysosomal storage diseases. *Annu. Rev. Biochem.* 60: 257–280.
2. Vellodi A. 2004. Lysosomal storage disorders. *Br. J. Haematol.* 128: 413–431.
3. Futerman AH, van Meer G. 2004. The cell biology of lysosomal storage disorders. *Nat. Rev. Mol. Cell Biol.* 5: 554–565.

4. Meikle PJ, Hopwood JJ, Clague AE, Carey WF. 1999. Prevalence of lysosomal storage disorders. *JAMA* 281: 249–254.
5. Hickman S, Neufeld EF. 1972. A hypothesis for I-cell disease: Defective hydrolases that do not enter lysosomes. *Biochem. Biophys. Res. Commun.* 49: 992–999.
6. Hickman S, Shapiro LJ, Neufeld EF. 1974. A recognition marker required for uptake of a lysosomal enzyme by cultured fibroblasts. *Biochem. Biophys. Res. Commun.* 57: 55–61.
7. O'Brien JS, Miller AL, Loverde AW, Veath ML. 1973. Sanfilippo disease type B: Enzyme replacement and metabolic correction in cultured fibroblasts. *Science* 181: 753–755.
8. Brady RO. 2006. Enzyme replacement for lysosomal diseases. *Annu. Rev. Med.* 57: 283–296.
9. Hasilik A, Klein U, Waheed A, Strecker G, von Figura K. 1980. Phosphorylated oligosaccharides in lysosomal enzymes: Identification of alpha-*N*-acetylglucosamine(1)phospho(6)mannose diester groups. *Proc. Natl Acad. Sci. USA* 77: 7074–7078.
10. Tabas I, Kornfeld S. 1980. Biosynthetic intermediates of beta-glucuronidase contain high mannose oligosaccharides with blocked phosphate residues. *J. Biol. Chem.* 255: 6633–6639.
11. Varki A, Kornfeld S. 1980. Structural studies of phosphorylated high mannose-type oligosaccharides. *J. Biol. Chem.* 255: 10847–10858.
12. Waheed A, Hasilik A, von Figura K. 1982. UDP-*N*-acetylglucosamine:lysosomal enzyme precursor *N*-acetylglucosamine-1-phosphotransferase. Partial purification and characterization of the rat liver Golgi enzyme. *J. Biol. Chem.* 257: 12322–12331.
13. Steet R, Lee WS, Kornfeld S. 2005. Identification of the minimal lysosomal enzyme recognition domain in cathepsin d. *J. Biol. Chem.* 280: 33318–33323.
14. Warner JB, Thalhauser C, Tao K, Sahagian GG. 2002. Role of N-linked oligosaccharide flexibility in mannose phosphorylation of lysosomal enzyme cathepsin L. *J. Biol. Chem.* 277: 41897–41905.
15. Reitman ML, Kornfeld S. 1981. Lysosomal enzyme targeting. N-Acetylglucosaminylphosphotransferase selectively phosphorylates native lysosomal enzymes. *J. Biol. Chem.* 256: 11977–11980.
16. Varki A, Kornfeld S. 1981. Purification and characterization of rat liver alpha-*N*-acetylglucosaminyl phosphodiesterase. *J. Biol. Chem.* 256: 9937–9943.
17. Waheed A, Hasilik A, von Figura K. 1981. Processing of the phosphorylated recognition marker in lysosomal enzymes. Characterization and partial purification of a microsomal alpha-*N*-acetylglucosaminyl phosphodiesterase. *J. Biol. Chem.* 256: 5717–5721.
18. Varki A, Sherman W, Kornfeld S. 1983. Demonstration of the enzymatic mechanisms of alpha-*N*-acetyl-D-glucosamine-1-phosphodiester *N*-acetylglucosaminidase (formerly called alpha-*N*-acetylglucosaminylphosphodiesterase) and lysosomal alpha-*N*-acetylglucosaminidase. *Arch. Biochem. Biophys.* 222: 145–149.
19. Rohrer J, Kornfeld R. 2001. Lysosomal hydrolase mannose 6-phosphate uncovering enzyme resides in the *trans*-Golgi network. *Mol. Biol. Cell.* 12: 1623–1631.
20. Do H, Lee WS, Ghosh P, Hollowell T, Canfield W, Kornfeld S. 2002. Human mannose 6-phosphate-uncovering enzyme is synthesized as a proenzyme that is activated by the endoprotease furin. *J. Biol. Chem.* 277: 29737–29744.
21. Kudo M, Brem MS, Canfield WM. 2006. Mucolipidosis II (I-cell disease) and mucolipidosis IIIA (classical pseudo-hurler polydystrophy) are caused by mutations in the GlcNAc-phosphotransferase alpha/beta-subunits precursor gene. *Am. J. Hum. Genet.* 78: 451–463.
22. Kornfeld S, Sly WS. 2001. In *Metabolic and Molecular Bases of Inherited Diseases*, CR Scriver, AL Beaudet, WS Sly, D Valle, eds. McGraw Hill, New York, pp. 3469–3482.
23. Kaplan A, Achord DT, Sly WS. 1977. Phosphohexosyl components of a lysosomal enzyme are recognized by pinocytosis receptors on human fibroblasts. *Proc. Natl Acad. Sci. USA* 74: 2026–2030.
24. Natowicz MR, Chi MM, Lowry OH, Sly WS. 1979. Enzymatic identification of mannose 6-phosphate on the recognition marker for receptor-mediated pinocytosis of beta-glucuronidase by human fibroblasts. *Proc. Natl Acad. Sci. USA* 76: 4322–4326.
25. Distler J, Hieber V, Sahagian G, Schmickel R, Jourdian GW. 1979. Identification of mannose 6-phosphate in glycoproteins that inhibit the assimilation of beta-galactosidase by fibroblasts. *Proc. Natl Acad. Sci. USA* 76: 4235–4239.
26. Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P, Stewart CL. 1994. Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes Dev.* 8: 2953–2963.
27. Wang ZQ, Fung MR, Barlow DP, Wagner EF. 1994. Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature* 372: 464–467.
28. Ludwig T, Eggenschwiler J, Fisher P, D'Ercole AJ, Davenport ML, Efstratiadis A. 1996. Mouse mutants lacking the type 2 IGF receptor (IGF2R) are rescued from perinatal lethality in Igf2 and Igf1r null backgrounds. *Dev. Biol.* 177: 517–535.

29. Dittmer F, Hafner A, Ulbrich EJ, Moritz JD, Schmidt P, et al. 1998. I-cell disease-like phenotype in mice deficient in mannose 6-phosphate receptors. *Transgenic Res*. 7: 473–483.
30. Ludwig T, Munier-Lehmann H, Bauer U, Hollinshead M, Ovitt C, et al. 1994. Differential sorting of lysosomal enzymes in mannose 6-phosphate receptor-deficient fibroblasts. *EMBO J*. 13: 3430–3437.
31. Schellens JP, Saftig P, von Figura K, Everts V. 2003. Deficiency of mannose 6-phosphate receptors and lysosomal storage: A morphometric analysis of hepatocytes of neonatal mice. *Cell Biol. Int*. 27: 897–902.
32. Dahms NM, Hancock MK. 2002. P-type lectins. *Biochim. Biophys. Acta* 1572: 317–340.
33. Wendland M, von Figura K, Pohlmann R. 1991. Mutational analysis of disulfide bridges in the Mr 46,000 mannose 6-phosphate receptor. Localization and role for ligand binding. *J. Biol. Chem*. 266: 7132–7136.
34. Goldberg DE, Gabel CA, Kornfeld S. 1983. Studies of the biosynthesis of the mannose 6-phosphate receptor in receptor-positive and -deficient cell lines. *J. Cell Biol*. 97: 1700–1706.
35. Sahagian GG, Neufeld EF. 1983. Biosynthesis and turnover of the mannose 6-phosphate receptor in cultured Chinese hamster ovary cells. *J. Biol. Chem*. 258: 7121–7128.
36. Rohrer J, Schweizer A, Johnson KF, Kornfeld S. 1995. A determinant in the cytoplasmic tail of the cation-dependent mannose 6-phosphate receptor prevents trafficking to lysosomes. *J. Cell Biol*. 130: 1297–1306.
37. Breuer P, Korner C, Boker C, Herzog A, Pohlmann R, Braulke T. 1997. Serine phosphorylation site of the 46-kDa mannose 6-phosphate receptor is required for transport to the plasma membrane in Madin–Darby canine kidney and mouse fibroblast cells. *Mol. Biol. Cell* 8: 567–576.
38. Meresse S, Hoflack B. 1993. Phosphorylation of the cation-independent mannose 6-phosphate receptor is closely associated with its exit from the *trans*-Golgi network. *J. Cell Biol*. 120: 67–75.
39. Braulke T, Mieskes G. 1992. Role of protein phosphatases in insulin-like growth factor II (IGF II)-stimulated mannose 6-phosphate/IGF II receptor redistribution. *J. Biol. Chem*. 267: 17347–17353.
40. Waguri S, Dewitte F, Le Borgne R, Rouille Y, Uchiyama Y, et al. 2003. Visualization of TGN to endosome trafficking through fluorescently labeled MPR and AP-1 in living cells. *Mol. Biol. Cell* 14: 142–155.
41. Waguri S, Tomiyama Y, Ikeda H, Hida T, Sakai N, et al. 2006. The luminal domain participates in the endosomal trafficking of the cation-independent mannose 6-phosphate receptor. *Exp. Cell Res*. 312: 4090–4107.
42. Mullins C, Bonifacino JS. 2001. The molecular machinery for lysosome biogenesis. *Bioessays* 23: 333–343.
43. Dell'Angelica EC, Payne GS. 2001. Intracellular cycling of lysosomal enzyme receptors: Cytoplasmic tails' tales. *Cell* 106: 395–398.
44. Ghosh P, Dahms NM, Kornfeld S. 2003. Mannose 6-phosphate receptors: new twists in the tale. *Nat. Rev. Mol. Cell Biol*. 4: 202–213.
45. Conibear E, Pearse B. 1994. A chimera of the cytoplasmic tail of the mannose 6-phosphate/IGF-II receptor and lysozyme localizes to the TGN rather than prelysosomes where the bulk of the endogenous receptor is found. *J. Cell Sci*. 107: 923–932.
46. Dintzis SM, Velculescu VE, Pfeffer SR. 1994. Receptor extracellular domains may contain trafficking information. Studies of the 300-kDa mannose 6-phosphate receptor. *J. Biol. Chem*. 269: 12159–12166.
47. Doray B, Ghosh P, Griffith J, Geuze HJ, Kornfeld S. 2002. Cooperation of GGAs and AP-1 in packaging MPRs at the *trans*-Golgi network. *Science* 297: 1700–1703.
48. Bonifacino JS. 2004. The GGA proteins: Adaptors on the move. *Nat. Rev. Mol. Cell Biol*. 5: 23–32.
49. Misra S, Puertollano R, Kato Y, Bonifacino JS, Hurley JH. 2002. Structural basis for acidic-cluster-dileucine sorting-signal recognition by VHS domains. *Nature* 415: 933–937.
50. Shiba T, Takatsu H, Nogi T, Matsugaki N, Kawasaki M, et al. 2002. Structural basis for recognition of acidic-cluster dileucine sequence by GGA1. *Nature* 415: 937–941.
51. Takatsu H, Katoh Y, Shiba Y, Nakayama K. 2001. Golgi-localizing, gamma-adaptin ear homology domain, ADP-ribosylation factor-binding (GGA) proteins interact with acidic dileucine sequences within the cytoplasmic domains of sorting receptors through their Vps27p/Hrs/STAM (VHS) domains. *J. Biol. Chem*. 276: 28541–28545.
52. Zhu Y, Doray B, Poussu A, Lehto VP, Kornfeld S. 2001. Binding of GGA2 to the lysosomal enzyme sorting motif of the mannose 6-phosphate receptor. *Science* 292: 1716–1718.
53. Le Borgne R, Hoflack B. 1998. Protein transport from the secretory to the endocytic pathway in mammalian cells. *Biochim. Biophys. Acta* 1404: 195–209.

54. Seaman MN. 2004. Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *J. Cell Biol*. 165: 111–122.
55. Arighi CN, Hartnell LM, Aguilar RC, Haft CR, Bonifacino JS. 2004. Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor. *J. Cell Biol*. 165: 123–133.
56. Seaman MN. 2005. Recycle your receptors with retromer. *Trends Cell Biol*. 15: 68–75.
57. Wan L, Molloy SS, Thomas L, Liu G, Xiang Y, et al. 1998. PACS-1 defines a novel gene family of cytosolic sorting proteins required for *trans*-Golgi network localization. *Cell* 94: 205–216.
58. Crump CM, Xiang Y, Thomas L, Gu F, Austin C, et al. 2001. PACS-1 binding to adaptors is required for acidic cluster motif-mediated protein traffic. *EMBO J*. 20: 2191–2201.
59. Ghosh P, Kornfeld S. 2004. The cytoplasmic tail of the cation-independent mannose 6-phosphate receptor contains four binding sites for AP-1. *Arch. Biochem. Biophys*. 426: 225–230.
60. Diaz E, Pfeffer SR. 1998. TIP47: A cargo selection device for mannose 6-phosphate receptor trafficking. *Cell* 93: 433–443.
61. Orsel JG, Sincock PM, Krise JP, Pfeffer SR. 2000. Recognition of the 300-kDa mannose 6-phosphate receptor cytoplasmic domain by 47-kDa tail-interacting protein. *Proc. Natl Acad. Sci. USA* 97: 9047–9051.
62. Saint-Pol A, Yelamos B, Amessou M, Mills IG, Dugast M, et al. 2004. Clathrin adaptor epsinR is required for retrograde sorting on early endosomal membranes. *Dev. Cell* 6: 525–538.
63. Doray B, Bruns K, Ghosh P, Kornfeld SA. 2002. Autoinhibition of the ligand-binding site of GGA1/3 VHS domains by an internal acidic cluster-dileucine motif. *Proc. Natl Acad. Sci. USA* 99: 8072–8077.
64. Scott GK, Fei H, Thomas L, Medigeshi GR, Thomas G. 2006. A PACS-1, GGA3 and CK2 complex regulates CI-MPR trafficking. *EMBO J*. 25: 4423–4435.
65. Damen E, Krieger E, Nielsen JE, Eygensteyn J, van Leeuwen JE. 2006. The human Vps29 retromer component is a metallo-phosphoesterase for a cation-independent mannose 6-phosphate receptor substrate peptide. *Biochem. J*. 398: 399–409.
66. Kasper D, Dittmer F, von Figura K, Pohlmann R. 1996. Neither type of mannose 6-phosphate receptor is sufficient for targeting of lysosomal enzymes along intracellular routes. *J. Cell Biol*. 134: 615–623.
67. Pohlmann R, Boeker MW, von Figura K. 1995. The two mannose 6-phosphate receptors transport distinct complements of lysosomal proteins. *J. Biol. Chem*. 270: 27311–27318.
68. Munier-Lehmann H, Mauxion F, Bauer U, Lobel P, Hoflack B. 1996. Re-expression of the mannose 6-phosphate receptors in receptor-deficient fibroblasts. Complementary function of the two mannose 6-phosphate receptors in lysosomal enzyme targeting. *J. Biol. Chem*. 271: 15166–15174.
69. Watanabe H, Grubb JH, Sly WS. 1990. The overexpressed human 46-kDa mannose 6-phosphate receptor mediates endocytosis and sorting of beta-glucuronidase. *Proc. Natl Acad. Sci. USA* 87: 8036–8040.
70. Tong PY, Kornfeld S. 1989. Ligand interactions of the cation-dependent mannose 6-phosphate receptor. Comparison with the cation-independent mannose 6-phosphate receptor. *J. Biol. Chem*. 264: 7970–7975.
71. Pearse BM. 1988. Receptors compete for adaptors found in plasma membrane coated pits. *EMBO J*. 7: 3331–3336.
72. Glickman JN, Conibear E, Pearse BM. 1989. Specificity of binding of clathrin adaptors to signals on the mannose-6-phosphate/insulin-like growth factor II receptor. *EMBO J*. 8: 1041–1047.
73. Lobel P, Fujimoto K, Ye RD, Griffiths G, Kornfeld S. 1989. Mutations in the cytoplasmic domain of the 275 kD mannose 6-phosphate receptor differentially alter lysosomal enzyme sorting and endocytosis. *Cell* 57: 787–796.
74. Canfield WM, Johnson KF, Ye RD, Gregory W, Kornfeld S. 1991. Localization of the signal for rapid internalization of the bovine cation-independent mannose 6-phosphate/insulin-like growth factor-II receptor to amino acids 24–29 of the cytoplasmic tail. *J. Biol. Chem*. 266: 5682–5688.
75. Jadot M, Canfield WM, Gregory W, Kornfeld S. 1992. Characterization of the signal for rapid internalization of the bovine mannose 6-phosphate/insulin-like growth factor-II receptor. *J. Biol. Chem*. 267: 11069–11077.
76. Denzer K, Weber B, Hille-Rehfeld A, Figura KV, Pohlmann R. 1997. Identification of three internalization sequences in the cytoplasmic tail of the 46 kDa mannose 6-phosphate receptor. *Biochem. J*. 326: 497–505.
77. Brunetti CR, Burke RL, Kornfeld S, Gregory W, Masiarz FR, et al. 1994. Herpes simplex virus glycoprotein D acquires mannose 6-phosphate residues and binds to mannose 6-phosphate receptors. *J. Biol. Chem*. 269: 17067–17074.
78. Gabel CA, Dubey L, Steinberg SP, Sherman D, Gershon MD, Gershon AA. 1989. Varicella-zoster virus glycoprotein oligosaccharides are phosphorylated during posttranslational maturation. *J. Virol*. 63: 4264–4276.

79. Chen JJ, Zhu Z, Gershon AA, Gershon MD. 2004. Mannose 6-phosphate receptor dependence of varicella zoster virus infection in vitro and in the epidermis during varicella and zoster. *Cell* 119: 915–926.
80. Purchio AF, Cooper JA, Brunner AM, Lioubin MN, Gentry LE, et al. 1988. Identification of mannose 6-phosphate in two asparagine-linked sugar chains of recombinant transforming growth factor-beta 1 precursor. *J. Biol. Chem*. 263: 14211–14215.
81. Blanchard F, Raher S, Duplomb L, Vusio P, Pitard V, et al. 1998. The mannose 6-phosphate/insulin-like growth factor II receptor is a nanomolar affinity receptor for glycosylated human leukemia inhibitory factor. *J. Biol. Chem*. 273: 20886–20893.
82. Lee SJ, Nathans D. 1988. Proliferin secreted by cultured cells binds to mannose 6-phosphate receptors. *J. Biol. Chem*. 263: 3521–3527.
83. Faust PL, Chirgwin JM, Kornfeld S. 1987. Renin, a secretory glycoprotein, acquires phosphomannosyl residues. *J. Cell Biol*. 105: 1947–1955.
84. Griffiths GM, Isaaz S. 1993. Granzymes A and B are targeted to the lytic granules of lymphocytes by the mannose-6-phosphate receptor. *J. Cell Biol*. 120: 885–896.
85. Ikushima H, Munakata Y, Ishii T, Iwata S, Terashima M, et al. 2000. Internalization of CD26 by mannose 6-phosphate/insulin-like growth factor II receptor contributes to T cell activation. *Proc. Natl Acad. Sci. USA* 97: 8439–8444.
86. Journet A, Chapel A, Kieffer S, Roux F, Garin J. 2002. Proteomic analysis of human lysosomes: Application to monocytic and breast cancer cells. *Proteomics* 2: 1026–1040.
87. Dennis PA, Rifkin DB. 1991. Cellular activation of latent transforming growth factor beta requires binding to the cation-independent mannose 6-phosphate/insulin-like growth factor type II receptor. *Proc. Natl Acad. Sci. USA* 88: 580–584.
88. van Kesteren CA, Danser AH, Derkx FH, Dekkers DH, Lamers JM, et al. 1997. Mannose 6-phosphate receptor-mediated internalization and activation of prorenin by cardiac cells. *Hypertension* 30: 1389–1396.
89. Blanchard F, Duplomb L, Raher S, Vusio P, Hoflack B, et al. 1999. Mannose 6-phosphate/insulin-like growth factor II receptor mediates internalization and degradation of leukemia inhibitory factor but not signal transduction. *J. Biol. Chem*. 274: 24685–24693.
90. Czupalla C, Mansukoski H, Riedl T, Thiel D, Krause E, Hoflack B. 2006. Proteomic analysis of lysosomal acid hydrolases secreted by osteoclasts: Implications for lytic enzyme transport and bone metabolism. *Mol. Cell. Proteomics* 5: 134–143.
91. Sleat DE, Wang Y, Sohar I, Lackland H, Li Y, et al. 2006. Identification and validation of mannose 6-phosphate glycoproteins in human plasma reveal a wide range of lysosomal and non-lysosomal proteins. *Mol. Cell. Proteomics* 5: 1942–1956.
92. Sleat DE, Lobel P. 1997. Ligand binding specificities of the two mannose 6-phosphate receptors. *J. Biol. Chem*. 272: 731–738.
93. Tong PY, Gregory W, Kornfeld S. 1989. Ligand interactions of the cation-independent mannose 6-phosphate receptor. The stoichiometry of mannose 6-phosphate binding. *J. Biol. Chem*. 264: 7962–7969.
94. Distler JJ, Guo JF, Jourdian GW, Srivastava OP, Hindsgaul O. 1991. The binding specificity of high and low molecular weight phosphomannosyl receptors from bovine testes. Inhibition studies with chemically synthesized 6-*O*-phosphorylated oligomannosides. *J. Biol. Chem*. 266: 21687–21692.
95. Tomoda H, Ohsumi Y, Ichikawa Y, Srivastava OP, Kishimoto Y, Lee YC. 1991. Binding specificity of D-mannose 6-phosphate receptor of rabbit alveolar macrophages. *Carbohydr. Res*. 213: 37–46.
96. Stein M, Zijderhand-Bleekemolen JE, Geuze H, Hasilik A, von Figura K. 1987. Mr 46,000 mannose 6-phosphate specific receptor: Its role in targeting of lysosomal enzymes. *EMBO J*. 6: 2677–2681.
97. Hoflack B, Kornfeld S. 1985. Lysosomal enzyme binding to mouse P388D1 macrophage membranes lacking the 215-kDa mannose 6-phosphate receptor: evidence for the existence of a second mannose 6-phosphate receptor. *Proc. Natl Acad. Sci. USA* 82: 4428–4432.
98. Sun G, Zhao H, Kalyanaraman B, Dahms NM. 2005. Identification of residues essential for carbohydrate recognition and cation dependence of the 46-kDa mannose 6-phosphate receptor. *Glycobiology* 15: 1136–1149.
99. Drickamer K. 1999. C-type lectin-like domains. *Curr. Opin. Struct. Biol*. 9: 585–590.
100. Roberts DL, Weix DJ, Dahms NM, Kim J-JP. 1998. Molecular basis of lysosomal enzyme recognition: Three-dimensional structure of the cation-dependent mannose 6-phosphate receptor. *Cell* 93: 639–648.
101. Olson LJ, Zhang J, Lee YC, Dahms NM, Kim J-JP. 1999. Structural basis for recognition of phosphorylated high mannose oligosaccharides by the cation-dependent mannose 6-phosphate receptor. *J. Biol. Chem*. 274: 29889–29896.
102. Gabel CA, Costello CE, Reinhold VN, Kurz L, Kornfeld S. 1984. Identification of methylphosphomannosyl residues as components of the high mannose oligosaccharides of *Dictyostelium discoideum* glycoproteins. *J. Biol. Chem*. 259: 13762–13769.

103. Freeze HH. 1986. Modifications of lysosomal enzymes in *Dictyostelium discoideum*. *Mol. Cell. Biochem*. 72: 47–65.
104. Reddy ST, Chai W, Childs RA, Page JD, Feizi T, Dahms NM. 2004. Identification of a low affinity mannose 6-phosphate-binding site in domain 5 of the cation-independent mannose 6-phosphate receptor. *J. Biol. Chem*. 279: 38658–38667.
105. Hancock MK, Yammani RD, Dahms NM. 2002. Localization of the carbohydrate recognition sites of the insulin-like growth factor II/mannose 6-phosphate receptor to domains 3 and 9 of the extracytoplasmic region. *J. Biol. Chem*. 277: 47205–47212.
106. Marron-Terada PG, Hancock MK, Haskins DJ, Dahms NM. 2000. Recognition of *Dictyostelium discoideum* lysosomal enzymes is conferred by the amino-terminal carbohydrate binding site of the insulin-like growth factor II/mannose 6-phosphate receptor. *Biochemistry* 39: 2243–2253.
107. Hoflack B, Fujimoto K, Kornfeld S. 1987. The interaction of phosphorylated oligosaccharides and lysosomal enzymes with bovine liver cation-dependent mannose 6-phosphate receptor. *J. Biol. Chem*. 262: 123–129.
108. Sohar I, Sleat D, Gong Liu C, Ludwig T, Lobel P. 1998. Mouse mutants lacking the cation-independent mannose 6-phosphate/insulin-like growth factor II receptor are impaired in lysosomal enzyme transport: Comparison of cation-independent and cation-dependent mannose 6-phosphate receptor-deficient mice. *Biochem. J*. 330: 903–908.
109. Olson LJ, Zhang J, Dahms NM, Kim J-JP. 2002. Twists and turns of the CD-MPR: Ligand-bound versus ligand-free receptor. *J. Biol. Chem*. 277: 10156–10161.
110. Olson LJ, Yammani RD, Dahms NM, Kim JJ. 2004. Structure of uPAR, plasminogen, and sugar-binding sites of the 300 kDa mannose 6-phosphate receptor. *EMBO J*. 23: 2019–2028.
111. Olson LJ, Dahms NM, Kim JJ. 2004. The N-terminal carbohydrate recognition site of the cation-independent mannose 6-phosphate receptor. *J. Biol. Chem*. 279: 34000–34009.
112. Brown J, Esnouf RM, Jones MA, Linnell J, Harlos K, et al. 2002. Structure of a functional IGF2R fragment determined from the anomalous scattering of sulfur. *EMBO J*. 21: 1054–1062.
113. Uson I, Schmidt B, von Bulow R, Grimme S, von Figura K, et al. 2003. Locating the anomalous scatterer substructures in halide and sulfur phasing. *Acta Crystallogr. D Biol. Crystallogr*. 59: 57–66.
114. Hancock MK, Haskins DJ, Sun G, Dahms NM. 2002. Identification of residues essential for carbohydrate recognition by the insulin-like growth factor II/mannose 6-phosphate receptor. *J. Biol. Chem*. 277: 11255–11264.
115. Huete-Perez JA, Engel JC, Brinen LS, Mottram JC, McKerrow JH. 1999. Protease trafficking in two primitive eukaryotes is mediated by a prodomain protein motif. *J. Biol. Chem*. 274: 16249–16256.
116. Clayton C, Hausler T, Blattner J. 1995. Protein trafficking in kinetoplastid protozoa. *Microbiol. Rev*. 59: 325–344.
117. Lang L, Couso R, Kornfeld S. 1986. Glycoprotein phosphorylation in simple eucaryotic organisms. Identification of UDP-GlcNAc:glycoprotein *N*-acetylglucosamine-1-phosphotransferase activity and analysis of substrate specificity. *J. Biol. Chem*. 261: 6320–6325.
118. Nolan CM, McCarthy K, Eivers E, Jirtle RL, Byrnes L. 2006. Mannose 6-phosphate receptors in an ancient vertebrate, zebrafish. *Dev. Genes Evol*. 216: 144–151.
119. Nadimpalli SK, von Figura K. 2002. Identification of the putative mannose 6-phosphate receptor (MPR 46) protein in the invertebrate mollusc. *Biosci. Rep*. 22: 513–521.
120. Lakshmi YU, Radha Y, Hille-Rehfeld A, von Figura K, Kumar NS. 1999. Identification of the putative mannose 6-phosphate receptor protein (MPR 300) in the invertebrate unio. *Biosci. Rep*. 19: 403–409.
121. Nadimpalli SK, Padmanabhan N, Koduru S. 2004. Biochemical and immunological characterization of a glycosylated alpha-fucosidase from the invertebrate Unio: Interaction of the enzyme with its in vivo binding partners. *Protein Expr. Purif*. 37: 279–287.
122. Gaszner M, Udvardy A. 1991. Purification of a lysosomal DNase from *Drosophila melanogaster*. *Biochem. Biophys. Res. Commun*. 181: 44–50.
123. Aeed PA, Elhammer AP. 1994. Glycosylation of recombinant prorenin in insect cells: The insect cell line Sf9 does not express the mannose 6-phosphate recognition signal. *Biochemistry* 33: 8793–8797.
124. Dennes A, Cromme C, Suresh K, Kumar NS, Eble JA, et al. 2005. The novel Drosophila lysosomal enzyme receptor protein mediates lysosomal sorting in mammalian cells and binds mammalian and Drosophila GGA adaptors. *J. Biol. Chem*. 280: 12849–12857.
125. Ni X, Canuel M, Morales CR. 2006. The sorting and trafficking of lysosomal proteins. *Histol. Histopathol*. 21: 899–913.
126. Munro S. 2001. The MRH domain suggests a shared ancestry for the mannose 6-phosphate receptors and other N-glycan-recognising proteins. *Curr. Biol*. 11: R499–501.
127. Cruciat CM, Hassler C, Niehrs C. 2006. The MRH protein erlectin is a member of the endoplasmic reticulum synexpression group and functions in N-glycan recognition. *J. Biol. Chem*. 281: 12986–12993.

128. Kimura Y, Nakazawa M, Yamada M. 1998. Cloning and characterization of three isoforms of OS-9 cDNA and expression of the OS-9 gene in various human tumor cell lines. *J. Biochem*. 123: 876–882.
129. Litovchick L, Friedmann E, Shaltiel S. 2002. A selective interaction between OS-9 and the carboxyl-terminal tail of meprin beta. *J. Biol. Chem*. 277: 34413–34423.
130. Bhamidipati A, Denic V, Quan EM, Weissman JS. 2005. Exploration of the topological requirements of ERAD identifies Yos9p as a lectin sensor of misfolded glycoproteins in the ER lumen. *Mol. Cell* 19: 741–751.
131. Kim W, Spear ED, Ng DT. 2005. Yos9p detects and targets misfolded glycoproteins for ER-associated degradation. *Mol. Cell* 19: 753–764.
132. Szathmary R, Bielmann R, Nita-Lazar M, Burda P, Jakob CA. 2005. Yos9 protein is essential for degradation of misfolded glycoproteins and may function as lectin in ERAD Mol. *Cell* 19: 765–775.
133. Dodd RB, Drickamer K. 2001. Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity. *Glycobiology* 11: 71R-79R.
134. Tay W. 1881. Symmetrical changes in the region of the yellow spot in each eye of an infant. *Trans. Ophthalmol. Soc. UK* 1: 55–57.
135. de Duve C, Pressman BC, Gianetto R, Wattiaux R, Appelmans F. 1955. Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochem. J*. 60: 604–617.
136. Hers HG. 1963. Alpha-glucosidase deficiency in generalized glycogen storage disease (Pompe's disease). *Biochem. J*. 86: 11–16.
137. Olson LJ, Hancock MK, Dix D, Kim J-JP, Dahms NM. 1999. Mutational analysis of the binding site residues of the bovine cation-dependent mannose 6-phosphate receptor. *J. Biol. Chem*. 274: 36905–36911.
138. Wendland M, Waheed A, von Figura K, Pohlmann R. 1991. Mr 46,000 mannose 6-phosphate receptor. The role of histidine and arginine residues for binding of ligand. *J. Biol. Chem*. 266: 2917–2923.

11 M-Type Lectins as Novel Components of Secretory Pathways

Nobuko Hosokawa and Kazuhiro Nagata

CONTENTS

11.1 INTRODUCTION

The role played by N-linked oligosaccharides in the endoplasmic reticulum (ER) quality control (ERQC) of newly synthesized glycoproteins is now well characterized and essentially conserved in all eukaryotes ranging from yeast to mammals [1,2]. In the mammalian ER, the addition and removal of glucose at the A branch terminus of N-linked oligosaccharides is involved in the productive folding of glycoproteins, which is mediated through interactions with the lectin-chaperones calnexin and calreticulin that recognize and bind to monoglucosylated N-glycans (Figure 11.1) (refer Chapter 8). Polypeptides that fail to adopt their native conformation are retained in the ER where they undergo multiple folding attempts. Failure to adopt the correct conformation results in their elimination from the ER through a mechanism known as ER-associated degradation (ERAD) [1–5] (Figure 11.1). ERAD is an elaborate mechanism for the elimination of misfolded proteins or misassembled protein complexes synthesized and accumulated in the ER. Polypeptides recognized as terminally misfolded are retrotranslocated out of the ER through a protein-conducting channel called the dislocon into the cytosol where they are degraded by the cytoplasmic proteasome. ATP hydrolysis is required for the retrotranslocation of ERAD substrates, which are polyubiquitinated by E3 ubiquitin ligases in the ER membrane. The N-linked sugars are removed by cytosolic peptide-N-glycanase (PNGase) before entry into the cavity of the proteasome.

In glycoprotein ERAD, mannose trimming from the middle branch (Branch B) of the N-linked oligosaccharide by ER α-mannosidase I is an important trigger for misfolded glycoproteins entering the disposal pathway both in yeast and in mammals [6]. Hence, a lectin capable of recognizing misfolded glycoproteins in the form of Man8B was predicted [7]. Mammalian ER-degradation enhancing α-mannosidase-like protein (EDEM1) and its yeast homolog Htm1p/Mnl1p were cloned as accelerators of glycoprotein ERAD [8–10]. EDEM1 and Htm1p/Mnl1p have a domain similar to class I α1,2-mannosidases (glycosylhydrolase family 47). Since they are enzymatically inactive, they were predicted to be lectins, and classified as M-type lectins. EDEM1 was proposed to extract

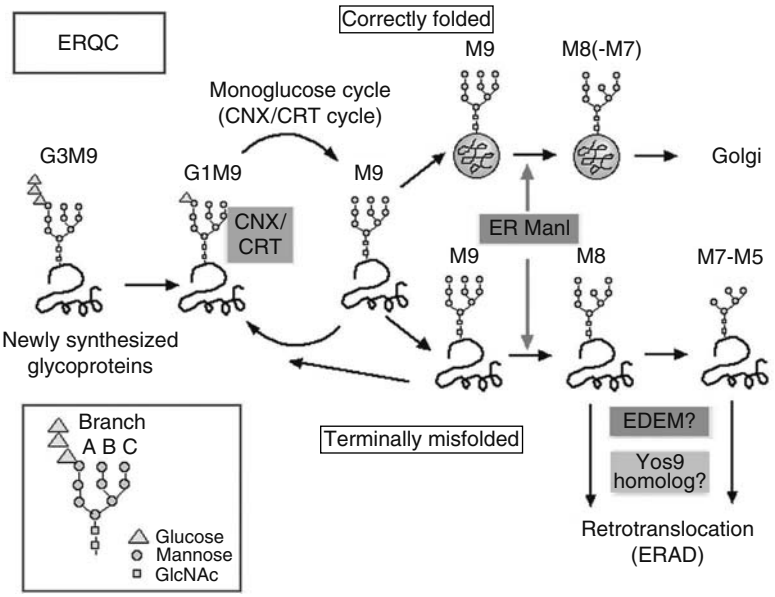


FIGURE 11.1 (See CD for color figure.) Scheme of ERQC of glycoproteins in mammals. The structure of the N-linked oligosaccharide is shown in the inset. CNX, calnexin; CRT, calreticulin.

misfolded glycoproteins from the calnexin/calreticulin cycle for entry into the degradation pathway [11,12]. Recently, two novel mammalian EDEM homologs, EDEM2 and 3, were cloned and found to accelerate ERAD of glycoproteins, similar to EDEM1 [13–15]. However, an analysis of these EDEM homologs shows that the molecular mechanism whereby EDEM proteins work in the ERAD pathway still remains an intriguing subject of investigation.

11.2 MAMMALIAN EDEM PROTEINS AND ER α-MANNOSIDASE I

The domain organization of three EDEM homologs is schematically shown in Figure 11.2 and compared with that of human ER α1,2-mannosidase I (ER ManI). The class I α1,2-mannosidase domain is well conserved among EDEM proteins with an amino acid identity of approximately 45%. The acidic amino acids conserved in this class of mannosidases are present, and important residues involved in hydrolysis by ER ManI have been shown to be structurally conserved in EDEM proteins by modeling on ER ManI [16]. ER ManI is a type II transmembrane protein, whereas EDEM2 and 3 are luminal proteins. EDEM1 and yeast Htm1p/Mln1p were originally described as membrane proteins, but recently, mammalian EDEM1 was reported to be luminal [14]. Although the precise significance of the N-terminal hydrophobic stretch of EDEM1 needs to be clarified, our preliminary data suggest that the function of EDEM1 in glycoprotein ERAD is not altered by cleavage of its N-terminal sequence (unpublished observation).

The expression of EDEM1 mRNA is upregulated by ER stress induced by the accumulation of misfolded proteins in the ER [8,17]. ER stress evokes a cellular response called unfolded protein response (UPR). In this process, ER chaperone proteins and ERAD machineries are transcriptionally upregulated to cope with this stress [18,19]. EDEM2 and, in some cells and tissues, EDEM3 is also transcriptionally induced by ER stress [14,15], whereas ER ManI is constitutively expressed [8]. ER ManI trims the mannose from the middle branch (Branch B) of the N-linked oligosaccharides attached to substrates both in the productive folding pathway and in ERAD, indicating that ER ManI does not discriminate between folded and unfolded forms of glycoproteins [1,2,20]. Overexpression of ER ManI causes the accelerated degradation of misfolded glycoproteins [21,22], as well as

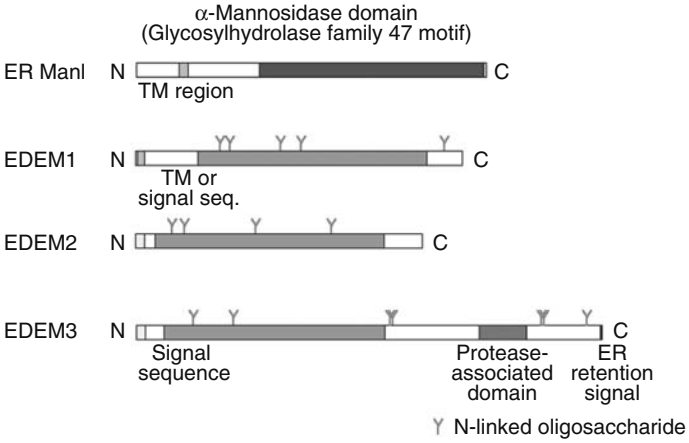


FIGURE 11.2 (See CD for color figure.) Domain organization of mammalian EDEM homologs (EDEM1, 2, and 3) and ER ManI. The α-mannosidase domains (class I α1,2-mannosidase domain or glycosylhydrolase family 47 motif) of EDEM proteins have approximately 45% of amino acid identity.

premature degradation of polypeptides on the way to acquiring their native conformations [22]. On the other hand, overexpression of EDEM1 does not affect the folding or secretion of correctly folding wild-type α1-antitrypsin [23], and cotransfection of ER ManI with EDEM1 further enhances glycoprotein ERAD in an additive manner [21]. It has also been reported that EDEM1 prevents dimerization or aggregation of ERAD substrates in the ER [23,24], suggesting that EDEM1 accommodates misfolded glycoproteins degradation-competent. Thus, ER ManI and EDEM1 may have diverse functions in glycoprotein quality control; EDEM acts on substrates destined for ERAD, whereas ER ManI operates in both productive folding and the ERAD pathway.

11.3 CLASS I α1,2-MANNOSIDASE FAMILY AND EDEM PROTEINS: LECTINS OR ENZYMES?

The class I α1,2-mannosidase family (glycosylhydrolase family 47) is composed of three groups, ER mannosidase I (ER ManI), Golgi α-mannosidase I (Golgi ManI), and EDEM [20]. Amino acid identity between the glycosylhydrolase family 47 domains of the three EDEM homologs is as high as 45%, whereas it is approximately 33% between ER ManI and each of the three EDEM proteins, 38% between Golgi ManI and ER ManI, and about 30% between Golgi ManI and EDEM (Figure 11.3).

As mentioned above, α-mannosidase enzyme activity was not detected in EDEM1, EDEM2, or in yeast Htm1p/Mln1p [8,9,13], but lectin activity has not yet been clearly demonstrated. We have shown recently that EDEM3 has a processing α-mannosidase activity in EDEM3-overexpressing cells, and that the degradation-enhancing effect of EDEM3 depends on its enzyme activity [15]. More recently, EDEM1 was also reported to have mannose-trimming activity in EDEM1-overexpressing cells, but importantly, the ERAD-enhancing capacity of EDEM1 was found not to depend on its enzyme activity [24]. The detection of the enzyme and lectin activities *in vitro* using recombinant EDEMs has not succeeded yet because it has been difficult to purify enough of these recombinant proteins.

In the fission yeast *Schizosaccharomyces pombe*, it is known that only a faint amount of glycoproteins with Man8B oligosaccharides are detected, and that, although ER α-mannosidase (ER Man) exists, its enzyme activity has not been observed. Interestingly, a recent study demonstrated that disruption of the gene encoding ER Man of *S. pombe* was found to prevent glycoprotein ERAD [25]. Therefore, they proposed that ER Man in *S. pombe* acts as a lectin rather than as an enzyme, similar to the case of Htm1p/Mnl1p/EDEM.

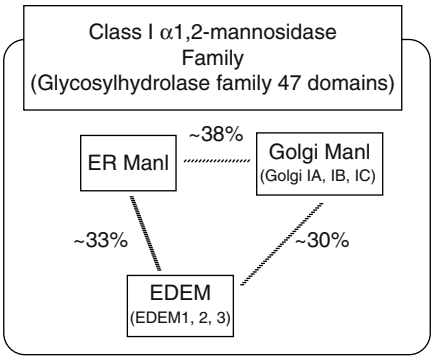


FIGURE 11.3 Class I α1,2-mannosidase family proteins. Class I α1,2-mannosidase family consists of three subgroups: ER ManI, Golgi ManI, and EDEM. The amino acid identity among the α-mannosidase domains or the glycosylhydrolase family 47 domains of the three groups is shown.

These observations indicate two possibilities for the function of EDEM proteins: EDEM proteins may act as M-type lectins to discriminate terminally misfolded glycoproteins destined for the ERAD pathway. Alternatively, EDEM proteins may be processing α-mannosidases that trim the mannose from the N-linked oligosaccharides of ERAD substrates. Further analysis will be required, including *in vitro* biochemical assays using recombinant proteins, to obtain a complete understanding of the function of these proteins.

11.4 CONCLUSION AND FUTURE DIRECTIONS

Collectively, it is still controversial at present whether mammalian EDEMs and yeast Htm1p/Mnl1p act as M-type lectins or processing α-mannosidases *in vivo* and *in vitro*. In mammalian cells, accumulating data shows that the N-glycan species attached to misfolded glycoproteins is not necessarily restricted to Man8B and that smaller oligosaccharides, such as Man5–7 ($Man_{5-7}GlcNAc_2$), are also preferentially recognized in the disposal pathway [26]. In this context, how can we incorporate the function and action mechanism of EDEMs in ERAD (Figure 11.1)? In yeast, does mannose-trimming from Man8B oligosaccharides proceeds further *in vivo*? Recently, Yos9p was reported in yeast to act as a lectin that discriminates the folding state of the glycoproteins, in the context of a large complex including a membrane-spanning E3 ubiquitin ligase in the ER [27,28] (refer Chapter11). Yos9p has a MRH domain (mannose 6-phosphate receptor homology domain) [29], and the binding to ERAD substrates bearing Man8B and Man5 oligosaccharides is reported [30]. There are several MRH domain-containing proteins in mammals, and analysis of these proteins in relation to the function of EDEM proteins would reveal the involvement of lectins in the ERAD pathway.

REFERENCES

1. Helenius, A. and Aebi, M. 2004. *Annu Rev Biochem* **73**, 1019–1049.
2. Trombetta, E. S. and Parodi, A. J. 2003. *Annu Rev Cell Dev Biol* **19**, 649–676.
3. Tsai, B., Ye, Y., and Rapoport, T. A. 2002. *Nat Rev Mol Cell Biol* **3**(4), 246–255.
4. Kostova, Z. and Wolf, D. H. 2003. *Embo J* **22**(10), 2309–2317.
5. McCracken, A. A. and Brodsky, J. L. 2003. *Bioessays* **25**(9), 868–877.
6. Cabral, C. M., Liu, Y., and Sifers, R. N. 2001. *Trends Biochem Sci* **26**(10), 619–624.
7. Jakob, C. A., Burda, P., Roth, J., and Aebi, M. 1998. *J Cell Biol* **142**(5), 1223–1233.
8. Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L. O., Herscovics, A., and Nagata, K. 2001. *EMBO Rep* **2**(5), 415–422.
9. Jakob, C. A., Bodmer, D., Spirig, U., Battig, P., Marcil, A., Dignard, D., Bergeron, J. J., Thomas, D. Y., and Aebi, M. 2001. *EMBO Rep* **2**(5), 423–430.

10. Nakatsukasa, K., Nishikawa, S., Hosokawa, N., Nagata, K., and Endo, T. 2001. *J Biol Chem* **276**(12), 8635–8638.
11. Molinari, M., Calanca, V., Galli, C., Lucca, P., and Paganetti, P. 2003. *Science* **299**(5611), 1397–1400.
12. Oda, Y., Hosokawa, N., Wada, I., and Nagata, K. 2003. *Science* **299**(5611), 1394–1397.
13. Mast, S. W., Diekman, K., Karaveg, K., Davis, A., Sifers, R. N., and Moremen, K. W. 2005. *Glycobiology* **15**(4), 421–436.
14. Olivari, S., Galli, C., Alanen, H., Ruddock, L., and Molinari, M. 2005. *J Biol Chem* **280**(4), 2424–2428.
15. Hirao, K., Natsuka, Y., Tamura, T., Wada, I., Morito, D., Natsuka, S., Romero, P., Sleno, B., Tremblay, L. O., Herscovics, A., Nagata, K., and Hosokawa, N. 2006. *J Biol Chem* **281**(14), 9650–9658.
16. Moremen, K. W. and Molinari, M. 2006. *Curr Opin Struct Biol* **16**(5), 592–599.
17. Yoshida, H., Matsui, T., Hosokawa, N., Kaufman, R. J., Nagata, K., and Mori, K. 2003. *Dev Cell* **4**(2), 265–271.
18. Bernales, S., Papa, F. R., and Walter, P. 2006. *Annu Rev Cell Dev Biol* **22**, 487–508.
19. Rutkowski, D. T. and Kaufman, R. J. 2004. *Trends Cell Biol* **14**(1), 20–28.
20. Herscovics, A. 2001. *Biochimie* **83**(8), 757–762.
21. Hosokawa, N., Tremblay, L. O., You, Z., Herscovics, A., Wada, I., and Nagata, K. 2003. *J Biol Chem* **278**(28), 26287–26294.
22. Wu, Y., Swulius, M. T., Moremen, K. W., and Sifers, R. N. 2003. *Proc Natl Acad Sci USA* **100**(14), 8229–8234.
23. Hosokawa, N., Wada, I., Natsuka, Y., and Nagata, K. 2006. *Genes Cells* **11**(5), 465–476.
24. Olivari, S., Cali, T., Salo, K. E., Paganetti, P., Ruddock, L. W., and Molinari, M. 2006. *Biochem Biophys Res Commun* **349**(4), 1278–1284.
25. Movsichoff, F., Castro, O. A., and Parodi, A. J. 2005. *Mol Biol Cell* **16**(10), 4714–4724.
26. Lederkremer, G. Z. and Glickman, M. H. 2005. *Trends Biochem Sci* **30**(6), 297–303.
27. Ismail, N. and Ng, D. T. 2006. *Cell* **126**(2), 237–239.
28. Cormier, J. H., Pearse, B. R., and Hebert, D. N. 2005. *Mol Cell* **19**(6), 717–719.
29. Munro, S. 2001. *Curr Biol* **11**(13), R499–501.
30. Szathmary, R., Bielmann, R., Nita-Lazar, M., Burda, P., and Jakob, C. A. 2005. *Mol Cell* **19**(6), 765–775.

12 Functional Aspects of the Hyaluronan and Chondroitin Sulfate Receptors

Edward N. Harris and Paul H. Weigel

CONTENTS

12.1 INTRODUCTION

Glycosaminoglycans (GAGs) are distinct from other sugars/oligosaccharides in that they are polymers of disaccharide units (Figure 12.1) composed of an amino sugar, *N*-acetylglucosamine (GlcNAc), or *N*-acetylgalactosamine (GalNAc), and uronic acid, glucuronic acid (GlcUA), or iduronic acid (IdUA). The exception is keratan sulfate (KS) in which the uronic acid is replaced by the neutral sugar galactose. Initially, GAGs were thought to be just part of an extracellular glue or ground substance that held tissues together and provided a liquid-like space between cells for the transport of nutrients. However, research using new technologies and techniques over the last two decades has revealed that the effects of GAGs on cells are more dynamic than previously thought. In addition to their structural roles, GAGs are fundamentally important as modifiers of cell behaviors that range from leukocyte recruitment during inflammation to the complex signaling events that help cancer cells survive during homing and metastasis. Cellular behaviors and responses

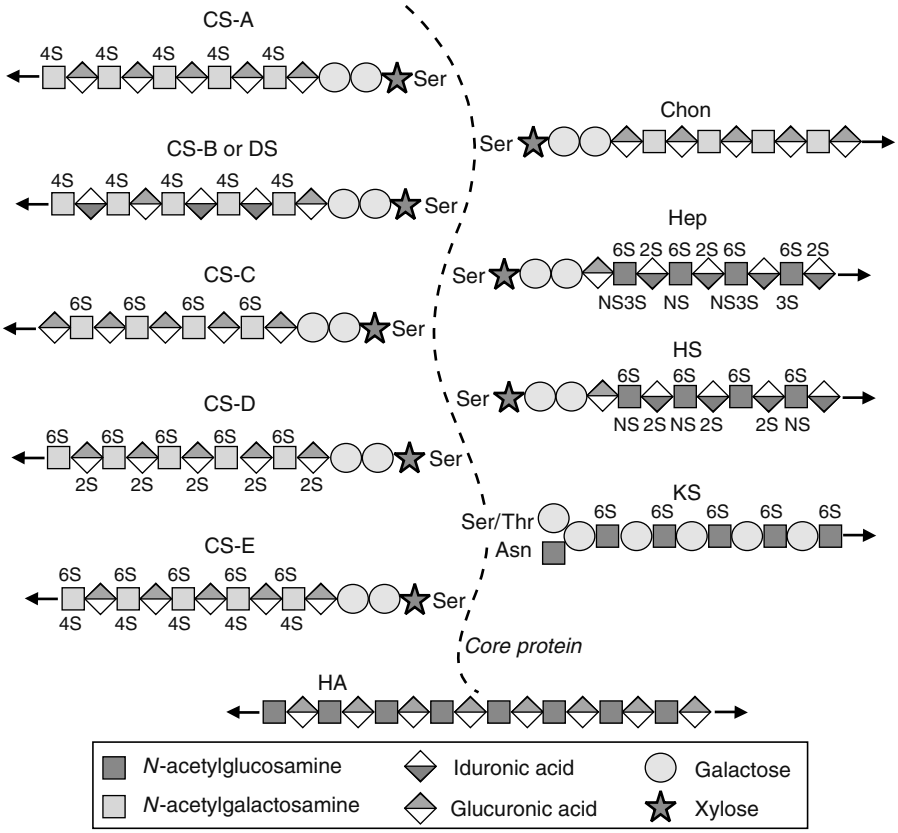


FIGURE 12.1 (See CD for color figure.) Schematic structures of the glycosaminoglycans. The GAGs are primarily polymers of disaccharide units containing the indicated amino and acidic sugars. Standard symbols for the sugar residues are used as defined by the Consortium for Functional Glycomics (http://grtc.ucsd.edu/essentials.html). All GAGs, with the exception of HA, are covalently linked to a variety of the ~40 known core proteins. The GAGs are attached via Ser residues, except for KS, which can also be linked via Thr or, less commonly, Asn residues. The arrows indicate that the repeating disaccharides continue to create larger polysaccharide chains.

depend in part on chain length of the GAG as well as GAG–sugar modifications such as sulfation and acetylation. Longer GAGs tend to be part of the extracellular matrix (ECM) in which long cable-like structures and networks may support tissue structure through cross-linking with other ECM components. Shorter GAGs or small oligosaccharides of less than 30–50 sugars may interact with cellular receptors to promote cell signaling involving antiapoptosis or immunostimulation [1].

There are four classes of GAGs (Figure 12.1): Chondroitin (Chon) sulfate (CS) A through E, heparin (HP) and heparan sulfate (HS), KS, and hyaluronan (HA). CS is a polymer of the disaccharide GlcUA-β1,3-GalNAc-β1,4 that can be sulfated at one or more positions: GalNAc-4 (type A), GalNAc-4 (type B), GalNAc-6 (type C), GalNAc-6, GlcUA-2 (type D), and GalNAc-4,6 (type E). CS-B, which is now called dermatan sulfate (DS), is an unusual CS family member in which iduronic acid substitutes for some, but not all, glucuronic acid (as in HP and HS) and GlcUA is in an alpha-1,3 linkage with GlcNAc. HP and HS are polymers of the disaccharide β1,4-GlcUA-β1,4-GlcNAc, in which some of the glucuronic acid is converted to iduronic acid. HS differs from HP in that the *N*-acetylglucosamine is de-*N*-acetylated and *N*-sulfated at position 2. HP is typically more highly sulfated than HS and is primarily found in mast cells of the immune system. In contrast, HS is ubiquitous throughout the body and its sugars are highly sulfated in patches along the chain, allowing differing ligands to bind to regions of higher or lower anionic charge density. Although KS

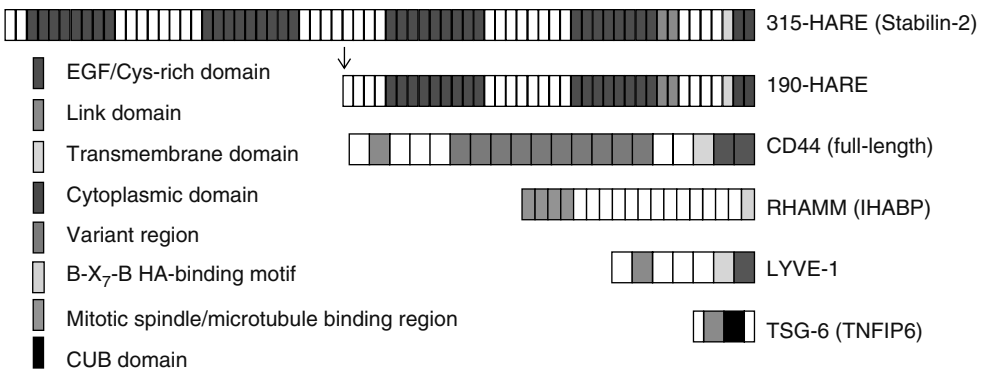


FIGURE 12.2 (See color insert following blank page 170. Also see CD for color figure.) HA-binding receptors/proteins. The domain organizations of the five HA-binding receptors or proteins discussed in this chapter are illustrated. The arrow indicates the proteolytic cleavage of the 315 kDa HARE to produce the smaller 190 kDa HARE isoform. Except for TSG-6, each colored segment represents exons encoding the indicated type of protein domain that is translated from the human gene. The colored segments for TSG-6 represent protein domains.

does not contain uronic acid, it is grouped with the family of GAGs because of its repeating disaccharide containing GlcNAc. It is assembled on a core protein with either an O-linkage to serine (Ser) or threonine (Thr) or an N-linkage to asparagine (Asn) [2]. KS is a major component of lumican and keratocan, proteoglycans that are important in the clarity of corneal matrix [3] as well as aggrecan, a structural molecule in cartilage and during brain development [4].

Unlike the other GAGs, HA is a "free" polysaccharide that is not assembled covalently onto a proteoglycan core protein in the normal biosynthetic organelles (Golgi and ER), but is often bound to proteins after its extrusion from the plasma membrane by HA synthase. After it is released by the synthase, HA is never sulfated, de-N-acetylated, or modified in any way (except by IαI as described below). All of the GAGs, except HA and some KS chains (as noted above), are attached to and assembled as O-linked glycans on proteoglycan core proteins and are sulfated to various degrees. CS, HP, and HS are attached to proteoglycans via Ser, or sometimes Thr residues, and their assembly starts with the common tetrasaccharide: GlcUA(β1,3)Gal(β1,3)Gal(β1,4)Xyl(β1)→Ser.

CS, HP, and HS often act as concomitant molecular signatures for their respective proteoglycans. Receptors or binding proteins for the GAGs can be grouped into three different classes. Receptors in the first class are soluble proteins that are ECM organizers, such as versican, aggrecan, and neurocan. The second class of receptors (e.g., CD44, RHAMM, and TSG6) is associated with cell-mediated processes stimulated by GAGs such as inflammation and cellular motility. The third class of receptors (e.g., CD44, HARE/Stabilin-2) is involved in GAG scavenging and turnover and is responsible for proper homeostasis of GAG levels in the plasma, tissues, and lymphatic fluids (Figure 12.2). CD44 is in two classes due its functional diversity.

12.2 TUMOR NECROSIS FACTOR-STIMULATED GENE 6

Tumor necrosis factor-stimulated gene 6 (TSG-6), which was formerly named tumor necrosis factor-induced protein-6 (TNFIP6), is a multi-GAG-binding protein that is secreted during inflammation in response to chemokines such as IL-1β, TNF, and PGE_2. TSG-6 is a serum protein with little or no expression under normal conditions, but is quickly secreted in response to inflammatory triggers, e.g., during ovulation and arthritis. This 35 kDa protein contains one CUB domain (complement protein subcomponents C1r/C1s, urchin embryonic growth factor, and bone morphogenic protein), and one Link domain and binds a variety of GAGs, including HA [5], CS-A [6], Hep/HS [7], and DS [8]. In addition to binding these GAGs, TSG-6 downregulates the inflammatory process by binding

to multiple proteins, such as the G1 domain of aggrecan [9], interalpha-(trypsin)-inhibitor (IαI) [10], pre-alpha-(trypsin)-inhibitor (PαI) [11], pentraxin-3 [12], and thrombospondin-1 (TSP-1) [13].

The role of HA in inflammatory processes such as wound healing and ovulation is well known. Ovulation begins when ovarian cells respond to a spike in luteinizing hormone (LH), which initiates a signaling cascade involving cAMP and PGE_2 mediators [14]. Granulosa and cumulus cells of the oocyte respond to these signal mediators by initiating transcription of a host of genes, one of which is TSG-6 [15–17]. TSG-6 also promotes cervical ripening by human cervical smooth muscle cells [18]. TSG-6 may be part of a network of factors that promote fertility by remodeling the ECM of the cumulus cell–oocyte complex, a matrix rich in HA [19]. TSG-6 knockout mice are subfertile or completely infertile [20]. Upregulation of TSG-6 is also most prominent in areas of chronic inflammation, such as the inflamed synovial tissues of rheumatoid arthritis patients [21] and the smooth muscles of the brochia in asthmatic patients [22].

12.2.1 TSG-6 AND HA

IαI is a unique proteoglycan in which the bikunin moiety, containing two Kunitz domains with antiprotease activity, is covalently linked to two protein heavy chains (HC1 and HC2) via its CS-A chain (Figure 12.3). CS-A is connected to bikunin by the standard GlcUAGalGalXyl-protein linkage at Ser-10. The C-terminal Asp648 alpha-carboxyl of each HC is esterified to the C-6 hydroxyl of an internal GalNAc of the CS-A chain [23]. During the inflammatory process, TSG-6 interacts with the HCs of IαI and mediates a reaction in which TSG-6 displaces CS-A

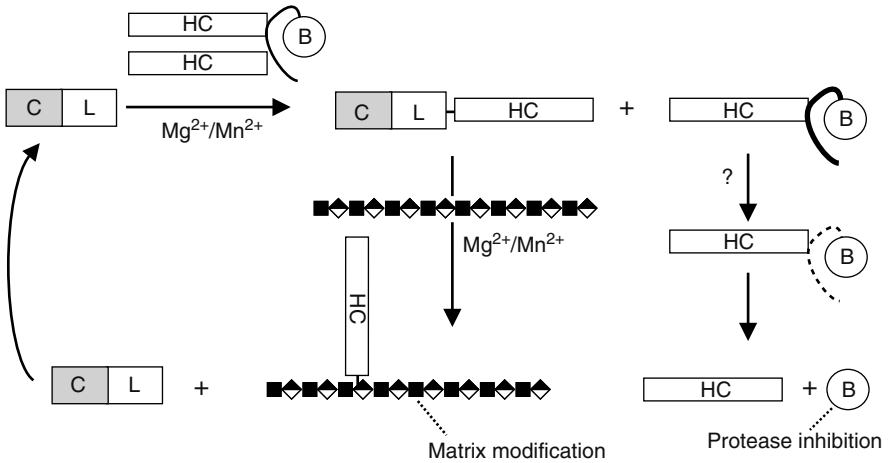


FIGURE 12.3 HA modification by the heavy chain of inter-α-trypsin inhibitor. IαI is a three-member proteoglycan that contains two protein heavy chains (HC in boxes), the metalloprotease inhibitor bikunin (B in the circles), and a single CS-A chain (solid line). The two HC chains are linked to CS-A by ester linkages at their C-terminal aspartate residues. The Cub (C in gray boxes) and Link (L in white box) domains are indicated for TSG-6. TSG-6 acts as a catalyst that transfers one of the heavy chains from IαI to HA (alternating squares and diamonds as in Figure 12.1) in two steps via transesterification reactions. The first reaction results in transfer of one HC chain from IαI to TSG-6, forming a covalent linkage between the HC C-terminal Asp and TSG-6. The second reaction then transfers the HC C-terminal Asp to a C-6 hydroxyl group of an internal GlcNAc residue in HA [25]. The HA–HC complex is utilized for matrix substrates while the TSG-6 can be reutilized for another HC transfer reaction. The CS-A chain of the modified IαI is degraded by an unknown mechanism (dashed line) that liberates bikunin, which can then inhibit metalloproteases.

and becomes attached to the C-terminal aspartic acid of HC1 or HC2 [24,25]. This reaction promotes transesterification of the HC from IαI, giving rise to two intermediate complexes, TSG-6-HC1 or TSG-6-HC2, in addition to the remaining substrate bikunin–HC1 or bikunin–HC2 [26]. During the inflammatory process, bikunin and the single HC linked to CS-A are cleaved by an unknown mechanism to produce free bikunin and HC. TSG-6 is recycled to its free state when it acts as a cofactor and catalyst to transfer HC to produce a covalent HC–HA complex, also called serum-derived HA-associated protein (SHAP) [27]. The HC undergoes a Mg^{2+}- or Mn^{2+}-dependent transesterification reaction between TSG-6 and a GlcUA of HA [25]. The end result is that as the HCs of IαI are transferred and sequestered by TSG-6 and HA. Bikunin, in its monomeric form, is the protease inhibitor neutralizing plasmin-induced metalloprotease activity [26].

It is the plasmin-induced metalloprotease activities involving leukocyte recruitment and the release of inflammatory cytokines (TNF-α, IL-1, IL-17, LPS, TGF-β, and PGE_2) that cause the red and swollen features of inflamed tissues. In the case of ovulation, a condition that mimics inflammation, cumulus cell ECM remodeling is dependent on TSG-6 and IαI or PαI activities to expand HA-rich matrix assembly [28]. PαI is an ortholog of IαI containing just one HC (HC3) connected to bikunin via CS-A [29]. The model for TSG-6 interaction with PαI is similar to the IαI-TSG-6 model, in which TSG-6 acts a catalyst to transfer HC3 to HA [20].

The HA-binding groove of TSG-6 is the best characterized of any of the HA-binding receptors. The amino acids essential for HA binding were first identified by site-directed mutagenesis [30] and then confirmed by nuclear magnetic resonance (NMR) spectroscopy [7]. Key contact residues in the HA-binding groove are Lys^{11}, Tyr^{12}, Tyr^{59}, Phe^{70}, Tyr^{78}, and Arg^{81}. Due to the high content of aromatic residues, it is likely that the hydrophobic face of pyranoside rings in HA stack onto the aromatic side chains of the amino acids that form two hydrophobic pockets. In addition, Cys^{47}, Val^{57}, Ile^{61}, Cys^{68}, and Trp^{88} are also likely candidates to interact closely with an HA chain. Computer modeling of the binding pocket reveals that five sugar rings of HA are required to fill the binding groove and that the hydrophobic pockets accommodate the acetamido side chain of GlcNAc [7]. Other cellular HA receptors and HA-binding proteins with Link domains also fit this model with some degrees of variation.

12.2.2 TSG-6 and Heparin

An increase in plasmin activity is associated with ECM reorganization during inflammation, such as occurs in rheumatoid arthritis [31], multiple sclerosis [32], bacterial infection [33], tumorigenesis/angiogenesis [34], and even follicle rupture during ovulation [35]. Plasmin activity induces the liberation of inflammatory cytokines, as well as activating matrix metalloproteases. TSG-6 is one of several molecules that regulate the antiplasmin activity or chondroprotective responses, through IαI and PαI, to prevent inflammation-mediated damage to the host. HP, in addition to the more heterogeneous HS, also binds TSG-6 and IαI and is a regulator of serine proteases through its inhibitory activity. Using mutagenesis and NMR spectroscopy, Mahoney et al. discovered that HP binds to TSG-6 at a distinct site from HA [36]. The amino acids involved are Lys^{20}, Lys^{34}, Lys^{41}, and Lys^{54}. Although HP does not affect HA-TSG-6 binding, HA and HP cannot both bind to TSG-6 at the same time, indicating that when bound to either GAG, TSG-6 is locked into a conformation that precludes binding to the other GAG. In addition, HP is a strong competitor for CS-A binding, suggesting that CS-A and HP bind at the same site. HP-TSG-6 binding was most optimal at pH 6 and binding affinities dropped when the pH shifted in either direction [36]. A lower pH reflects the conditions at sites of inflammation, indicating that HP-TSG-6 complexes could have increased antiplasmin activity. In contrast, pH changes had no effect on HA-TSG-6 binding. It is not clear how both HA and HP modulate the antiplasmin activity of TSG-6, but the current model requires dimerization of TSG-6 [36]. To further potentiate the antiplasmin activity of bikunin, two TSG-6 proteins may dimerize when bound to a single HP or HS chain (much like a sandwich) and

the TSG-6 dimers are able to bind to bikunin. This binding may change the orientation of the Kunitz domains of bikunin making them more accessible, thus increasing their antiplasmin activity [36].

To make this story a bit more complex, but perhaps at the same time more reasonable, TSG-6 also interacts with TSP1. TSP1 directly inhibits T-cell mediated T-cell activation [37] and dendritic cell activation [38], but promotes neutrophil oxidative burst responses [39]. TSP1 is one of many factors involved in early wound repair and mice lacking an effective TSP1 response are delayed in wound healing [40]. The N-terminal module of TSP1 is related to the TSG-6-binding pentraxin family of proteins and binds TSG-6 in a divalent cation-dependent manner. The N-terminal module of TSP1 also binds HP and is an effective inhibitor of the TSP1 interaction with TSG-6. TSP1 and HA do not bind with each other but both bind to TSG-6 at distinct sites. Furthermore, in the absence of HA, TSP1 enhances the interaction of TSG-6 with IαI, resulting in an increased degradation of IαI. In summary, a trimolecular complex consisting of TSP1-HA-TSG-6 acts upon IαI to give rise to HA–HC complexes for ECM remodeling. The presence of HP or HS may negatively regulate this activity [13]. Overall, both HA and HP or HS proteoglycans enhance the antiplasmin activities of TSG-6 and concomitant effectors during inflammation.

12.3 RECEPTOR FOR HYALURONIC ACID MEDIATED MOTILITY

Receptor for hyaluronic acid mediated motility (RHAMM) is present as both a plasma membrane GPI-anchored HA receptor and an intracellular HA-binding protein lacking both transmembrane and cytoplasmic signal sequences [41]. Consequently, it is also called intracellular hyaluronic acid-binding protein (IHABP) or CD168. RHAMM was first identified, with the use of anti-RHAMM antibodies, in murine 3T3 cells as a group of HA-binding proteins with molecular masses of 56, 66, and 70 kDa [42,43]. Following this initial discovery, the masses reported for RHAMM have been quite variable, ranging from 58 kDa on human B cells [44] to 120 kDa in fibroblasts from *src* null mice [45], based on immunoreactivity against anti-RHAMM antibodies. The wide variety of molecular masses could be due to proteolysis, different glycosylation states of the receptor [46], or the presence of splice variants [47]. The full-length human RHAMM, which is about 85% homologous to murine RHAMM, is identified as a 725-amino acid protein that migrates at ~84 kDa by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

RHAMM is the founding member of hyaladherins (HA-binding proteins) that contain the basic B-X_7-B motif, in which any seven amino acids are flanked by two basic amino acids, either lysine or arginine. RHAMM contains two of these motifs near the C-terminus that mediate high-affinity protein binding to HA ($K_d = 10^{-8}$ M) [43,48]. Mutagenesis of any basic residue flanking the B-X_7-B motifs led to a sharp decrease in HA binding [49]. Multidimensional NMR analysis of domains containing the B-X_7-B HA-binding motifs reveals that it is a helix-loop-helix with a coiled-coil tertiary structure, which is very different than the architecture of the Link domains found in most other HA-binding proteins [5,50]. RHAMM is a predominantly helical protein containing more basic residues than any of the other hyaladherins. Theoretically, the net positive charge of the protein could allow it to bind to any GAG due to the high negative charge densities of sulfates of the CS and HS families. However, none of the CS family members bind to RHAMM. Other than HA, only HP binds to RHAMM and competes with HA at the same binding sites. In *ras*-transformed cells, HP induces cell locomotion in a dose-dependent manner. Biotinylated HP specifically binds to GST-RHAMM and deletion of the two B-X_7-B motifs negates HP binding [49]. In contrast, other groups dispute the claim that HP binds to RHAMM, although all agree that HA binds to RHAMM with high affinity [51,52]. It is still unclear if HP specifically binds to RHAMM because different experimental methodologies and GAG sources were employed for determining HP-RHAMM binding, and few follow-up studies have appeared after the HP-RHAMM "controversy" between 1994 and 1998.

Expression of RHAMM is widespread in a variety of cell types including smooth muscles cells [53,54], and endothelial cells [55]. In addition, overexpression of full-length RHAMM and some of

its splice variants are found in cancers of the hematopoietic system [56,57], breast [58,59], pancreas [60], brain [61], and the lower GI tract [62]. RHAMM is primarily located in the cytoplasm within cells, hence the alternative name IHABP, but depending on cell type, it can also be expressed on the cell surface where it interacts with focal adhesion sites [63].

In the cytoplasm, RHAMM binds intracellular HA and localizes to both microtubules [54] and centrosomes, and participates in the maintenance of spindle pole integrity. The function of intracellular HA is unknown, but it is thought to be important during cytokinesis and for cellular movement; how HA functions in these capacities is not clear. Most hypotheses for functions of intracellular HA point to the hyaladherin (HA binding) activity of RHAMM and thus a role for intracellular RHAAM–HA interactions. There is still considerable debate on the source of intracellular HA since the three human (and other mammalian) isozymes of HA synthase (HAS) are plasma membrane enzymes whose only characterized activity is the extrusion of HA into the ECM [64]. Previous reports suggest that intracellular HA is a diffuse network of cable-connecting vesicles, which becomes more pronounced with serum stimulation or during mitosis at the G2/M phase [65]. In keratinocytes, extracellular HA is internalized in a novel endocytic pathway involving CD44 that is independent of clathrin-coated pits [66]. It is known that some, but not all internalized, HA is catabolized down to monosaccharides and simple organic molecules. The proposed HA-rich, nonacidic vesicles (hyaluronasomes), which would also contain hyaluronidases, are specialized structures that may, theoretically, be a source of intracellular HA that interact with microtubules and motor proteins to facilitate cytokinesis [54,67]. Despite several similar models in the literature, the hyaluronasome is still a hypothetical structure that has not been biochemically identified, isolated, or characterized.

RHAMM is a linker protein connecting microtubules via its N-terminus [68] and HA cables via its C-terminus [49]. More specifically, the C-terminus of RHAMM also contains a leucine zipper domain, essential for binding with dynein- and kinesin-like protein 2 (TPX2), which partially overlaps the HA-binding motifs [69]. RHAMM binds directly to TPX2 in a cell cycle-dependent manner, is essential for spindle pole assembly and maintenance throughout the cell cycle, and is required for cell progression through G_2 of mitosis [70]. The RHAAM interaction with microtubule motors may explain the role of HA in mitotic regulation and why increased RHAMM expression might enhance the oncogenic properties of cells, which is common in several cancer types including multiple myeloma [69], breast [59], stomach [71], and endometrial [72].

Multiple studies over the last several decades provide a connection between cell motility and HA [73–77]. One such study, comparing metastases and primary tumor staining, found that RHAMM expression was significantly higher in the metastatic cells. Presentation of RHAMM on the cell surface, where it can interact with external HA, can regulate signaling downstream of *ras* through ERK, the extracellular signal-regulated protein kinase [59] and through Src kinases [45]. Treatment of cells with nanogram amounts of HA promoted changes in downstream events such as phosphorylation and dephosphorylation of focal adhesion kinase, pp125FAK. This rapid and transient phosphorylation was regulated in part by RHAMM signal transduction via MAP kinases ERK-1/2 [63,78]. The interaction of ERK and RHAMM stimulates RHAMM binding to cytoskeletal elements that induce cellular locomotion [68,79].

The human gene encoding RHAMM is located on chromosome 5q33.2 and contains 18 exons, which therefore makes it a likely candidate for alternative splicing [80]. Likewise, the mouse gene contains 14 exons and Western analyses of fibroblast cell lysates revealed two bands, a major 70 kDa band and a minor larger 73 kDa band encoding exon 4 [41]. An analysis of both breast and cervical cancers versus normal tissue revealed the expression of RHAMM splice variants lacking sequences encoded by only exon 4, only exon 13, and both exons 4 and 13, suggesting that increased expression of RHAMM-like splice variants contribute to an oncogenic phenotype [51,68]. B cells and plasma cells taken from multiple myeloma patients showed two deletion splice variants of RHAMM lacking 48 ($RHAMM^{-48}$) or 147 DNA base pairs ($RHAMM^{-147}$), which were expressed in the cancerous cells but not in normal or resting B cells [47]. Increased expression of both the full-length and variant RHAMM proteins alters normal intracellular signaling pathways, inducing the oncogenic phenotype.

12.4 CD44

Human CD44, first known as a lymphocyte homing receptor, is a C-type lectin plasma membrane receptor that binds to HA and other ECM components, and plays major roles in cytokinesis, endocytosis, phagocytosis, cell signaling, and ligand internalization. The common or standard CD44 receptor is encoded by exons 1–5, 16–18, and 20, which together produce an ~80 kDa protein. Exons 6–15 are highly variable exons, commonly named v1–10, whose multiple combinations produce many CD44 variants that contribute to its wide array of receptor activities [81]. For example, CD44v6 expresses exon 6 of the variable group, which is really exon 11 if counting from exon 1.

12.4.1 CD44 Interactions with HA

CD44 binds HA via the Link domain encoded by exon 2. This Link domain, which is well conserved and similar to the Link domains found in other hyaladherins, is ~100 amino acids long, has two antiparallel β-sheets crossed over by two α-helices, and is held together by two disulfide bridges [82]. The receptor on resting leukocytes does not bind HA until it is "activated" by proinflammatory cytokines, which results in the modification of the N-glycans on CD44 [83]. It is thought that tumor necrosis factor-alpha (TNF-α) induces sialidase activity to remove sialic acid from the N-glycans, which then allows HA-binding to the receptor [84]. The critical residues involved with HA binding include Arg^{41}, Tyr^{42}, Arg^{78}, Tyr^{79} along with the other necessary residues, which are Lys^{38}, Lys^{68}, Asn^{100}, Asn^{101}, and Tyr^{105} [82].

12.4.2 CD44 Interactions with CS

The CD44 Link domain interacts with the CS family of GAGs. In a competition assay, versican binding to CD44 was inhibited by HA, Chon, CS-A, CS-B, CS-C, CS-D, and CS-E, but not by HS or KS. In a complementary biochemical assay, soluble CD44 also directly bound all these immobilized GAGs with the exception of KS and HS [85]. In a rolling cell experiment, murine leukocytes expressing CD44 initially bound and rolled on capillary tubes coated with CSs at physiological flow rates. The rolling action of these cells was abrogated when either the cells were treated with anti-CD44 antibodies or the tubes were treated with chrondroitinase ABC, but not hyaluronidase [86]. In a more relevant physiological context, one of the major proteoglycans, Serglycin, is a ligand for CD44. Serglycin (Mr ~17,600; previously known as gp600) is secreted by hematopoietic cells and contains CS-A and CS-C chains attached to a centrally located Ser-Gly repeat [87]. Serglycin plays a role in the packaging and inactivation of some of the proteases in granules of neutraphil granulocytes [88]. Granzyme B, a perforin component that can disrupt cell membranes, is packaged in part by Serglycin in the maturing granule of cytotoxic T-cells [89–91]. Serglycin is secreted primarily by lymphocytes, spleenoctyes, and thymocytes in small amounts, but its secretion is greatly increased by lipopolysaccharide or concanavalin A [87]. Murine T-cells expressing CD44 aggregate in the presence of Serglycin-containing serum. Furthermore, purified Serglycin binds to CD44 via its CS-A chain, which increases CD3-dependent release of granzyme B, suggesting that this proteoglycan may regulate lymphoid cell adherence and activation [92,93]. Serglycin-deficient T-cells fail to package granzyme B, but not granzyme A, Fas ligand, and perforin, although the mRNA levels for all three proteins were not affected and the late-stage apoptosis of target cells mediated by these cytotoxic cells was also unaffected. In summary, these studies point out that the CS chains of Serglycin bind to the HA-binding region of the Link domain of CD44 and is important in part for the cytotoxic T-cell response in cell-mediated apoptosis and cell rolling.

12.4.3 CD44 and HA Endocytosis

CD44 is the most common endocytic receptor for HA and possibly CS at the local tissue level. Initial studies of CD44-mediated endocytosis performed in alveolar macrophages and simian

virus transformed murine 3T3 fibroblasts suggested that CD44 is the primary endocytic receptor for HA in these cells and responsible for its delivery to lysosomes for degradation [94]. The endocytosis of HA also facilitates a higher cycling rate of CD44 on the cell surface of chondrocytes, which is blocked by anti-CD44 antibodies [95]. Chondrocytes are present in HA-rich environments where the balance of synthesis and catabolism are critical to maintain homeostasis of HA quantity and size (since chain length affects viscosity). Decreases in the mass of HA will increase its fluidity and could cause adverse cellular signaling events and, possibly, inflammation (e.g., in arthritis). Once CD44 binds HA, the complex is internalized followed by dissociation in an early endosome, most likely the recycling syntaxin13 and GTPase Rab11/25-rich endosome that promotes surface expression of ligand-free CD44 [96]. This mechanism is regarded as the "slow" recycling pathway with a half-life of about 30 min. Fast recycling regulated by Rab4 has not been shown to be associated with CD44 recycling.

The mechanisms of internalization and catabolism of HA are not the same in all tissues. Catabolic pathways for HA exist in smooth muscle cells [97], macrophages [94,98], chondrocytes [99], and some cancer cells [100]. Roughly half of all the HA in the human body is localized in the dermal layers of the skin. Keratinocytes possess a novel, noncoated pit-mediated endocytic pathway for internalization and catabolism of HA. Epidermal keratinocytes actively produce HA, which has a short half-life of ~1.5 days, indicating that the skin is one of the most active tissues in the body for the synthesis and degradation of HA [101,102]. Partially digested chains of high molecular mass HA [66] bound to CD44 are internalized, and trafficked to nonacidic, perinuclear vesicles from which the HA is eventually sorted to lysozymes and then degraded to monosaccharides by lysosomal hyaluronidase and the two exoglycosidases: β-*N*-acetylglucosaminidase and β-glucuronidase. The terminal HA-degrading vesicles are devoid of Cathepsin D, normally present in lysosomes, suggesting that the HA is delivered to a specialized lysosome [66].

The route of internalization in keratinocytes is independent of clathrin-coated pit and caveolin-mediated mechanisms. Consistent with this, the cytoplasmic domain of CD44 does not contain any AP-2 or clathrin consensus-binding motifs, suggesting that internalization is regulated by alternative mechanisms. Exon 20, the most 3′ exon, encodes the intracellular C-terminal 70 amino acids, which contains binding sites for ankyrin and the ezrin/radixin/moesin (ERM) linker proteins of the actin-based cytoskeletal network and the intracellular signaling cues for internalization, some of which have not been elucidated [103,104]. CD44 is widely thought to be pivotal for HA-induced chemotaxis and movements that require cytoskeletal structures. A study with rat endothelial cells in situ showed that CD44 and the ERM proteins colocalize with each other on actin filaments near the plasma membrane and this region was also enriched with phosphorylated protein kinase C α/β II. Induction of actin branching facilitated cellular movements needed for normal wound healing. Addition of protein kinase C (PKC) inhibitors caused F-actin structures to disassemble and the microdomains containing CD44 and associated proteins to disassociate. The end-result of this PKC inhibition was to decrease wound healing [105]. CD44 also contains two Ser phosphorylation sites. Inhibition of casein kinase II decreased CD44 association with the actin cytoskeleton, suggesting that CD44 is activated in a phosphylation- and dephosphorylation-dependent way during cytoskeletal movement [106]. Although increased CD44 expression is thought to promote tumorigenesis, at least one of the cytoskeletal proteins belonging to the ERM family of proteins may temper this activity. The NF2 tumor suppressor gene, which encodes Merlin, induces tumorigenesis in Schwann cells through a loss-of-function mutation [107]. In cell culture, overexpression of Merlin prevents CD44 from binding with HA. Additionally, knocking down the expression of endogenous Merlin in Tr6BC1 Schwannoma cells in Rag1 mice or overexpression of a dominant-negative Merlin (missing the first 50 amino acids from the N-terminus) results in the promotion of tumor cell growth [108].

12.4.4 CD44 and Phagocytosis

The HA-binding activity of CD44 is diverse enough to include phagocytosis of apoptotic bodies. Following tissue injury and the influx and resulting apoptosis of polymorphonuclear leukocytes

(PMN), clearance mechanisms must exist to remove cellular debris in order for normal healing to occur. The persistence of injury and the lack of healing in CD44-deficient mice demonstrate how CD44 contributes to the clearing of apoptotic cells [109]. The precise mechanism for how this occurs is still unknown. CD44 also contributes to the phagocytosis of bacteria by PMNs, further supporting the notion that ankyrin attached to CD44 assists in the maturation of phagosomes with the inclusion of lysosomal vesicles [110]. Confirming these earlier data, a more definitive study using erythrocytes coated with either anti-CD44 or control antibodies of the same isotype measured a 25-fold increase in CD44-selective phagocytosis by RAW 264.7 macrophages in comparison to controls. The induction of cell signaling molecules such as Syk, Rac1, and PI3-kinase during phagocytosis indicates that CD44 is a competent phagocytic receptor, in addition to its HA endocytosis and cell matrix-binding activities [111].

12.4.5 CD44 Splice Variants

The most common isoform of CD44 is CD44H (H = hematopoietic), also called CD44s (s = standard), which is encoded by exons 1–5, 16–18, and 20 with a mass of ~80 kDa. The extracellular domain contains a variable region encoded by exons 6–15 (also termed variants 1–10) that can increase the mass of CD44 to just above 200 kDa and could theoretically result in thousands of different CD44 splice variant combinations. The inclusion of exon 19 in place of exon 20 is a special case, in which the resulting cytoplasmic domain consists of only three amino acids, forming a "tail-less" receptor variant or CD44st (short-tail). CD44st does not bind or internalize HA efficiently and its expression varies from 5% to 33% of total CD44 receptors in normal articular chondrocytes. Interestingly, CD44st may be a one of several regulators for CD44H, since antisense inhibition of CD44st enhances HA binding and internalization by CD44H [112]. Other studies with CD44st indicate that the cytoplasmic domain of CD44 is also critical for HA binding, pericellular matrix assembly, and endocytosis of HA [113]. It is not known how CD44st interacts with CD44H, but one possibility is steric interference through common ECM-binding partners.

Most CD44 splice variants were discovered, with the use of reverse transcription polymerase chain reaction (RT-PCR) or immunohistochemistry in cancer specimens, as markers of metastasis [114,115]. The predominant hypothesis is that the increased expression of CD44 and its variants leads to increased cell motility, survival, and sometimes, proliferation of cancerous cells in many types of tissues. Hundreds of studies have explored the involvement of many variants in numerous types of cancers, so we will briefly note just a few common CD44 variants and their associated cancers. Culty et al. first discovered variants in breast cancer. Their findings indicated that not all of the higher mass species of CD44 bind HA and those that did bind HA had differing binding affinities [100]. CD44v6 and v9 splice variants are present in numerous malignant cancers. CD44v6 associates with phosphylated c-Met/hepatocyte growth factor/scatter factor and stimulates downstream signaling through the MAP kinase pathway, stimulating metastasis in colon cancer cells [116]. CD44v6 and v9 protein levels, but not mRNA levels, are also increased by osteopontin, a secreted phosphoprotein that regulates both chemotaxis and attachment, in breast cancer cells [117]. CD44v6 expression is increased in higher grades of malignancy in endometrial cancers, suggesting that this isoform is directly involved with stromal invasion of early squamous cervical carcinoma [118]. A large-scale study by Liu et al. [119] correlated metastasis of breast, gastric, lymph node, colorectal, and lung cancers with CD44v6.

Most studies have shown positive correlations between increased CD44 variant expression and disease. For example, CD44v3–10 is usually more highly expressed than normal in head and neck tumors, breast, ovarian, and general gynecological tumors [120–123]. Some studies indicate that CD44 variants may curtail the activities of CD44H. A study using SW620 cells, which do not express CD44, suggests that in this colon cancer cell model, overexpression of CD44 variants leads to a decrease in motility. When both CD44H and other CD44 variants are coexpressed, these variants interfere with the normal motility functions of CD44H [124]. Recently, the view is emerging that the

pattern of CD44 splice variants appears to be more relevant to disease outcomes than the amount of individual CD44 variants. For example, normal lung and nonsmall cell lung carcinoma express CD44 v10, v8–10, v6–10, v2–10, and v3–10. In lung adenocarcinoma, the expression of v6–10, v2–10, and v3–10 were lower, whereas in squamous cell lung carcinoma, the expression of these five variants was either the same as normal lung or higher [125]. Certainly, the issues of CD44 splice variant expression patterns and the different ligands available for the variant region of CD44 is just beginning to be explored.

12.4.6 CD44 Knockout Mice

Since CD44 has many functions including cell–matrix interactions, internalization of HA, and transduction of cellular signaling, one would assume that animals lacking CD44 would be developmentally impaired or severely deformed. It turned out that the CD44 knockout mice show no real change in phenotype [126]. Their skeleton, cartilage, and skin dermis are quite normal. In conditional knockouts in which a CD44 antisense transgene was controlled by the cytokeratin promoter, animals overaccumulated HA within the dermis of the skin and in the cornea of the eye [127]. The results suggest that during development in the CD44-KO background, compensatory mechanisms exist to take care of local HA turnover and ECM stabilization. These compensatory mechanisms either do not exist or are overwhelmed when there is a loss of CD44 in developed cells. It is not known what receptors might compensate for loss of CD44, but there are some good candidates. RHAMM could be one of them, since it binds HA, promotes cell signaling, and leukocyte migration. In the arthritic CD44−/− mouse, signaling through the RHAMM pathway upregulates proinflammatory genes such as IL-1B, TNFα, and MAP kinase. The paw swelling symptom of collagen-induced arthritis was partially alleviated with anti-RHAMM antibody treatment, suggesting that cell migration and ECM stabilization is mediated by RHAMM in the absence of CD44 [128]. Overall, this indicates that RHAMM is a compensatory receptor for CD44, but more research is needed to verify this.

12.5 LYMPHATIC VASCULAR ENDOTHELIAL HYALURONAN RECEPTOR 1

As noted above, the CD44−/− mouse has no apparent phenotype and the lymphatic endothelia are quite normal in these mice. Another candidate compensatory receptor in the absence of CD44 is lymphatic vascular endothelial hyaluronan receptor 1 (LYVE-1), a small HA-binding receptor expressed at high levels in the lymphatics. LYVE-1 is a type I, single-pass ~60 kDa plasma membrane glycoprotein containing 322 amino acids. A Link domain, which comprises the bulk of the receptor ectodomain, binds only HA, not CS or HS [129]. The cysteines (Cys) of this Link domain are highly conserved, but only three other residues (Lys^{46}, Try^{87}, and Asn^{109}) are conserved in the Link family and known to be involved with HA binding. Much of the LYVE-1 Link domain is nonconserved, which may explain why this particular Link domain only binds HA and none of the other GAGs. On the C-terminal side of the Link domain, there are several tracts of basic amino acids, e.g., RRKK,[198] which may also assist in HA binding. Like CD44, LYVE-1 contains one free Cys in the transmembrane region that may function as a dimerization stabilizer during receptor-mediated endocytosis [130]. Two common pathways for endocytosis are entered by targeting either to clathrin-coated pits utilizing tyrosine-based motifs, e.g., YXXB or NPXY, where X is any amino acid and B is an amino acid with a bulky hydrophobic side-group [131], or to caveolae utilizing a hydrophobic motif, e.g., YXXXFXF [132]. LYVE-1 does not contain consensus sequences for coated-pit-mediated endocytosis, and caveolin-mediated endocytosis by LYVE-1 has not been reported. It is quite possible that LYVE-1-mediated uptake is similar to CD44-mediated endocytosis in that the route of endocytosis is unique with a recycling receptor going to a specialized compartment where HA is dissociated from the receptor and then delivered to an acidic lysosome-like organelle [66]. However, this unusual endocytic mechanism is only proven in keratinocytes and is not known to exist in other cell types.

LYVE-1 expression is limited to lymph vessel endothelium, but not blood vascular endothelium, with the possible exception of the liver sinusoids [133]. Tissues expressing LYVE-1 include lymph vessel endothelia draining GI, skin, breast, lymph node, and salivary gland tissues [129]. Since LYVE-1 is restricted to lymphatic endothelia, it has been utilized as a distinct marker for lymphangiogenesis. Currently, the mechanisms stimulating and directing lymphangiogenesis are largely unknown. Lymphangiogenesis is a similar process to hemangiogenesis, which is the outgrowth of new blood vessels into tissues. Current evidence indicates that lymphangiogenesis is mediated by vascular endothelial growth factor (VEGF-C) and -D binding to VEGFR3 to expand lymphatic vessel outgrowth. Likewise, hemangiogenesis is mediated by VEGF-A binding with VEGFR1 and VEGFR2 to promote blood vessel growth [134]. Lymphangiogenesis occurs naturally in growing tissues and grafted tissues. Like hemangiogenesis, lymphangiogenesis facilitates the trafficking of monocytes and leukocytes throughout the tissue. Increased lymphangiogenesis is highly correlated with the rejection of graft tissue [135] and poor outcome due to premalignant breast cancer metastasis [136]. In contrast, decreased lymphangiogenesis and increased hemangiogenesis are implicated in the development of malignant cancer in which the trafficking of white blood cells is restricted [137]. A high probability of metastasis and poor outcome is observed if the cancer tissue retains the lymph vessel endothelium, allowing the trafficking of cancerous cells to spread to other areas of the body [138–140].

It is not clear how LYVE-1 fits into the lymphangiogenesis story besides being a good marker for lymph endothelium. The primary route of metastatic malignant cells and part of the circulatory route of white blood cells is through the lymphatic system. Like CD44, which is a homing receptor for leukocytes, LYVE-1 may facilitate the binding, rolling, and extravasation of leukocytes by its recognition of HA. Modeling the Link domain of LYVE-1 onto the CD44 Link crystal structure reveals that the LYVE-1 Link domain may have an on/off switching mechanism that works similarly to the N-glycan modification mechanism on CD44 that regulates HA binding [82]. Indirect evidence of this switch is evident since the receptor is functionally "off" when expressed in B-lymphoma and cervical carcinoma cells, but functionally "on" when expressed in 293 or COS cells. Treatment of 293T cells expressing LYVE-1 with phorbol 12-myristate 13-acetate increases HA binding to LYVE-1 [141]. LYVE-1 may also transport HA across the lymphatic endothelium from the interstitium to the lymph in a transcytosis process. This idea is supported by the findings that LYVE-1 is on both the luminal and abluminal sides of lymphatic capillaries and that the lymphatic endothelia contains intracytoplasmic vesicle clusters that may form stable channels that facilitate this type of transport [130,142]. The specific functional aspects of how LYVE-1, a distinct and interesting marker for the lymph endothelium, uses HA in cellular transport has yet to be discovered.

12.6 HYALURONIC ACID RECEPTOR FOR ENDOCYTOSIS

The hyaluronic acid receptor for endocytosis (HARE) is the primary scavenger receptor for systemic HA and CS turnover. The adult human body typically contains about 15 g of HA of which approximately one-third or 5 g is catabolized daily [143–145]. About 50% of the total HA is located in the epidermal layers with the remaining enriched in the bursa of the joints, the vitreous humor of the eye and, to varying extents, all other tissues. The catabolism of HA occurs at two levels: local, within the immediate areas of a tissue, and systemic, involving transport of HA to distal sites. As previously discussed, CD44, which is found on many cell types of the body, handles much of the local HA turnover. However, in the 1980s, Laurent and coworkers observed that most HA is not broken down at the local tissue level, but is rather transported through the lymphatic and vascular systems to distal sites to be catabolized in the sinusoidal endothelium of the lymph nodes and liver [145–147]. The general route of HA turnover involves local partial cleavage in tissues, possibly by hyaluronidases or mechanical/oxidative degradation in the interstitium. The large HA fragments (probably with bound proteoglycans) first perfuse through the lymphatic system, which catabolizes about 85% of it, with the remaining amount entering the peripheral circulation via the thoracic duct. This residual HA is then internalized and degraded in the liver sinusoids; maintaining steady-state HA blood levels at <50 ng/mL.

HARE was first identified, purified, and characterized from rat liver sinusoidal endothelial cells [148,149]. Using cross-reacting monoclonal antibodies, the HARE protein from human spleen was then characterized [150], and soon after, molecular cloning of the rat and human HARE were reported [151,152]. HARE is expressed as two isoforms; a large 300 kDa (rat) or 315 kDa (human) and a smaller proteolytically processed 175 kDa (rat) or 190 kDa (human) form. The smaller rat and human recombinant HARE isoforms were characterized in stable cell lines by Zhou et al. [151] and Harris et al. [152]. The full-length human isoform was cloned and expressed in 293 cell lines and named Stabilin-2 due to its close homology to Stabilin-1, which had no known function at that time [153,154]. Subsequently, HARE was also named FEEL-2 [155,156], by other groups, although in the field of glycobiology, it is commonly referred to as HARE as it will be in this chapter. Although earlier reports questioned the legitimacy of the smaller isoform [157], we confirmed that cells expressing the human full-length recombinant protein from cDNA always produce both HARE isoforms [154].

Human HARE is expressed primarily in the medullary sinuses of lymph nodes, sinusoids of the liver, and venous sinuses of the spleen. It has also been reported in mice in specialized tissues such as corneal and lens epithelium, mesenchymal cells of heart valves, ependymal cells lining the ventricles of the brain, and in prismatic epithelial cells covering the renal papillae [158,159]. The human gene, located on chromosome 12q23.3, is 180.2 kb, contains 69 exons and encodes a 2551 amino acid type I receptor (Figure 12.2). The glycoprotein mass by SDS-PAGE, including all glycosylation modifications, is 315/300 kDa and 190/175 kDa for both rat/human isoforms. The smaller isoform is identical to the C-terminal 1416 (human)/1431 (rat) amino acids of the full-length HARE protein. A subset of cellular HARE is proteolytically cleaved near or at the plasma membrane, but the protease involved remains unidentified, and the proteolytic site does not contain a consensus sequence for any known proteases [154]. Both isoforms contain numerous fascilin-like and EGF-like domains in complex Cys-rich areas and one Link domain near the plasma membrane domain. The function of either the fascilin- and EGF-like domains is unknown, although they may interact with extracellular/matrix components during adhesion, signaling, or ligand interactions. We recently showed that the HARE Link domain is the primary HA-binding domain [172,173]. Since the HARE and TSG-6 Link domains share the greatest degree of homology [7] and bind HA with high affinity, it is reasonable that the Link domain is the primary HA-binding domain of HARE, responsible for ~90% of the binding activity.

HA–HARE complex formation is high-affinity [154,160] and cation-independent [161] and internalization of complexes proceeds via the clathrin-coated pit pathway [151,162,163]. The cytoplasmic domain contains three motifs involved in clathrin/AP-2 targeting or intracellular trafficking after endocytosis (e.g., YSYFRI, FQHF and NPLY) [174]. Subcellular localization of recombinant rat HARE is primarily at the cell surface and in endocytic vesicles, but not lysosomes, suggesting that HARE is recycled back to the cell surface [151]. Additionally, cycloheximide treatment of isolated rat liver endothelial cells (LEC) revealed that endocytosis of ^{125}I-HA proceeds uninterrupted for several hours, and the receptors recycle from endosomes back to the cell surface every ~28 min to mediate uptake of additional HA [162]. Both rat and human HARE recycle in a similar fashion; recombinant 190 kDa human HARE in stable 293 Flp-In cell lines recycles every ~8.5 min, whereas the recombinant 175 kDa rat HARE in stable SK-Hep1 cell lines recycles every ~20 min [151,152]. The differences in receptor recycling times is more likely due to cellular machinery differences rather than intrinsic to the receptor itself.

12.6.1 HARE and GAG Binding

A primary role for HARE is to bind and endocytose HA, but it also internalizes CS-A, CS-C, CS-D, CS-E, DS, and Chon [152,172]. It has also been reported to bind advanced glycation end-products, collagen N-terminal propeptides, and Gram-negative and Gram-positive bacteria [155,156,163]. HA used in our standard endocytic experiments is radiolabeled with ^{125}I only at the reducing end, which

keeps virtually all of the HA polymer native and unmodified. To assess binding of free CS, HP, or HS chains, LECs were incubated with both the ^{125}I-HA and an unlabeled GAG in competition assays. Interference of ^{125}I-HA binding by CS and HP in rat LEC was discovered in the 1980s [162,164,165]. Oynebraten et al. [166] found that rat LECs efficiently endocytosed and degraded labeled Serglycin, the serum proteoglycans rich in CS and HS chains. Serglycin also interfered with HA binding; further supporting the conclusion that HARE is "promiscuous" and able to recognize many different GAGs [158,167,172]. As noted previously, CD44 also binds and internalizes Serglycin, but the efficiency of binding and endocytosis by HARE is much greater. Later, when the 190 kDa HARE was cloned from human spleen and expressed in 293 Flp-In cells, all of the binding and endocytic assays indicated that Chon and all of the CS types competed for the binding of labeled HA to varying degrees. Inhibition of HA endocytosis was most pronounced with CS-A, CS-C, and CS-D, but was not observed with HP, HS, or KS [152]. The discrepancy in results regarding HP competition in earlier [162] versus recent [152] reports leads us to believe the HP used in earlier experiments may have contained other GAG impurities. The broad GAG-binding specificity of HARE, its location in lymphatic and liver sinusoids, and its fast recycling activity as a coated-pit mediated endocytic receptor strongly argues that HARE is the primary scavenger receptor responsible for the clearance of multiple GAGs, including HA, rather than LYVE-1, CD44, RHAAM, or any other known HA/GAG receptor. For example, although LYVE-1 is located in lymphoid endothelium, it is not likely a systemic scavenger receptor for HA or CS because its binding affinity for HA is too low to be efficient and its endocytic uptake rate is less than half that for HARE. We also discovered that HARE is a clearance receptor for HP, and that the HA and HP binding sites are separate and independent [172,173].

12.6.2 HARE and Development

HA is a molecule that facilitates cell motility in cancer and development, for example, as seen in the HA/CD44/RHAMM studies. Since HARE is a recently discovered HA receptor, studies involving its role in development are few. HA and other proteoglycans are the foundation of endocardial development, chamber septation, and valve development. These ECM components participate in cell signaling cascades and creating space for cell infiltration that is necessary for cellular differentiation as the organ develops [168–170]. A study examining the role of fascilin-containing genes expressed in heart development shows that Stabilin-1, HARE, β-IgH3, and periostin are regulated both spatially and temporally as the heart develops [159,171]. HARE expression is restricted to the postnatal mature valve endothelial cells, an area rich in GAGs. Another preliminary study indicates that HARE expression oscillates as the fetal rat liver develops. Using histological methods, HARE expression is present as early as day 13, maximizes by day 15, disappears by day 17, and then reappears on day 18 (P.H. Weigel, unpublished data). It is not known what regulates this differential gene expression although the production and scavenging of HA by hyaladherins in organogenesis is critical for proper development.

12.7 CONCLUSION AND FUTURE DIRECTIONS

The basic ligand-binding functionalities of the HA/CS receptors are fairly well known. Tissue/ECM remodeling, organ development, and responses to diseases caused by either intrinsic pathologies (cancer, arthritis, etc.) or external invaders caused by bacterial or viral infections are more of the complex roles that involve the HA/CS receptors. There are considerable redundancies in place to handle the maintenance of basic life functions in case one of these receptors fails. This is most evident in the CD44 and LYVE-1 knockout mice, which contain no obvious phenotype, although multiple functions (especially for CD44) have been attributed to the respective receptor. However, each receptor is necessary for the optimal response in the organism to meet all of the different challenges associated with life. One of the future challenges for the HA/CS receptors is the understanding

of how the molecular machinery is tied in with physiological responses. Despite the numerous studies on the subject of endocytosis, the molecular pathways for receptor internalization, ligand uncoupling, gene regulation, and transport are only vaguely understood for these receptors. In addition, it is not known how these pathways and molecular mechanisms are altered in cases of cancer or infection. Current research is still on the observation stage in which we are dissecting the mechanical parts and we have much to learn on how these parts work together.

ACKNOWLEDGMENTS

We thank Janet A. Weigel for help assembling the manuscript and references. Some of the research described here was supported by National Institute of General Medical Sciences grants GM69961, GM35978 and F32 GM070262 from the National Institutes of Health.

REFERENCES

1. R. Stern, A.A. Asari, and K.N. Sugahara, Hyaluronan fragments: An information-rich system, *Eur. J. Cell Biol*. 85, 699–715, 2006.
2. F.P. Barry, L.C. Rosenberg, J.U. Gaw, T.J. Koob, and P.J. Neame, N- and O-linked keratan sulfate on the hyaluronan binding region of aggrecan from mature and immature bovine cartilage, *J. Biol. Chem*. 270, 20516–20524, 1995.
3. W.W. Kao and C.Y. Liu, Roles of lumican and keratocan on corneal transparency, *Glycoconj. J*. 19, 275–285, 2002.
4. J. Dudhia, Aggrecan, aging and assembly in articular cartilage, *Cell Mol. Life Sci*. 62, 2241–2256, 2005.
5. D. Kohda, C.J. Morton, A.A. Parkar, H. Hatanaka, F.M. Inagaki, I.D. Campbell, and A.J. Day, Solution structure of the link module: A hyaluronan-binding domain involved in extracellular matrix stability and cell migration, *Cell* 86, 767–775, 1996.
6. A.A. Parkar and A.J. Day, Overlapping sites on the Link module of human TSG-6 mediate binding to hyaluronan and chrondroitin-4-sulphate, *FEBS Lett*. 410, 413–417, 1997.
7. C.D. Blundell, A. Almond, D.J. Mahoney, P.L. DeAngelis, I.D. Campbell, and A.J. Day, Towards a structure for a TSG-6-hyaluronan complex by modeling and NMR spectroscopy: Insights into other members of the link module superfamily, *J. Biol. Chem*. 280, 18189–18201, 2005.
8. C.M. Milner, V.A. Higman, and A.J. Day, TSG-6: A pluripotent inflammatory mediator? *Biochem. Soc. Trans*. 34, 446–450, 2006.
9. A.A. Parkar, J.D. Kahmann, S.L.T. Howat, M.T. Bayliss, and A.J. Day, TSG-6 interacts with hyaluronan and aggrecan in a pH-dependent manner via a common functional element: Implications for its regulation in inflamed cartilage, *FEBS Lett*. 428, 171–176, 1998.
10. H.G. Wisniewski, J.C. Hua, D.M. Poppers, D. Naime, J. Vilcek, and B.N. Cronstein, TNF/IL-1-inducible protein TSG-6 potentiates plasmin inhibition by inter-alpha-inhibitor and exerts a strong anti-inflammatory effect in vivo, *J. Immunol*. 156, 1609–1615, 1996.
11. R. Forteza, S.M. Casalino-Matsuda, M.E. Monzon, E. Fries, M.S. Rugg, C.M. Milner, and A.J. Day, TSG-6 potentiates the anti-tissue kallikrein activity of inter-{alpha}-inhibitor through bikunin release, *Am. J. Respir. Cell Mol. Biol*. 36, 20–31, 2006.
12. A. Salustri, C. Garlanda, E. Hirsch, M. De Acetis, A. Maccagno, B. Bottazzi, A. Doni, A. Bastone, G. Mantovani, P.B. Peccoz, G. Salvatori, D.J. Mahoney, A.J. Day, G. Siracusa, L. Romani, and A. Mantovani, PTX3 plays a key role in the organization of the cumulus oophorus extracellular matrix and in in vivo fertilization, *Development* 131, 1577–1586, 2004.
13. S.A. Kuznetsova, A.J. Day, D.J. Mahoney, M.S. Rugg, D.F. Mosher, and D.D. Roberts, The N-terminal module of thrombospondin-1 interacts with the link domain of TSG-6 and enhances its covalent association with the heavy chains of inter-{alpha}-trypsin inhibitor, *J. Biol. Chem*. 280, 30899–30908, 2005.
14. J.Y. Park, Y.Q. Su, M. Ariga, E. Law, S.L.C. Jin, and M. Conti, EGF-like growth factors as mediators of LH action in the ovulatory follicle, *Science* 303, 682–684, 2004.
15. C. Fulop, A. Salustri, and V.C. Hascall, Coding sequence of a hyaluronan synthase homologue expressed during expansion of the mouse cumulus–oocyte complex, *Arch. Biochem. Biophys*. 337, 261–266, 1997.
16. O. Carrette, R.V. Nemade, A.J. Day, A. Brickner, and W.J. Larsen, TSG-6 is concentrated in the extracellular matrix of mouse cumulus oocyte complexes through hyaluronan and inter-alpha-inhibitor binding, *Biol. Reprod*. 65, 301–308, 2001.

17. D. Mukhopadhyay, V.C. Hascall, A.J. Day, A. Salustri, and C. Fulop, Two distinct populations of tumor necrosis factor-stimulated gene-6 protein in the extracellular matrix of expanded mouse cumulus cell–oocyte complexes, *Arch. Biochem. Biophys.* 394, 173–181, 2001.
18. T. Fujimoto, R.C. Savani, M. Watari, A.J. Day, and J.F. Strauss III, Induction of the hyaluronic acid-binding protein, tumor necrosis factor-stimulated gene-6, in cervical smooth muscle cells by tumor necrosis factor-{alpha} and prostaglandin E2, *Am. J. Pathol.* 160, 1495–1502, 2002.
19. A. Salustri, M. Yanagishita, and V.C. Hascall, Synthesis and accumulation of hyaluronic acid and proteoglycans in the mouse cumulus cell–oocyte complex during follicle-stimulating hormone-induced mucification, *J. Biol. Chem.* 264, 13840–13847, 1989.
20. C. Fulop, S. Szanto, D. Mukhopadhyay, T. Bardos, R.V. Kamath, M.S. Rugg, A.J. Day, A. Salustri, V.C. Hascall, T.T. Glant, and K. Mikecz, Impaired cumulus mucification and female sterility in tumor necrosis factor-induced protein-6 deficient mice, *Development* 130, 2253, 2003.
21. H.G. Wisniewski, W.H. Burgess, J.D. Oppenheim, and J. Vilcek, TSG-6, an arthritis-associated hyaluronan binding protein, forms a stable complex with the serum protein inter-alpha-inhibitor, *Biochemistry* 33, 7423–7429, 1994.
22. C.M. Lilly, H. Tateno, T. Oguma, E. Israel, and L.A. Sonna, Effects of allergen challenge on airway epithelial cell gene expression, *Am. J. Respir. Crit. Care Med.* 171, 579–586, 2005.
23. J.J. Enghild, G. Salvesen, I.B. Thogersen, Z. Valnickova, S.V. Pizzo, and S.A. Hefta, Presence of the protein—glycosaminoglycan—protein covalent cross-link in the inter-alpha-inhibitor-related proteinase inhibitor heavy chain 2/bikunin, *J. Biol. Chem.* 268, 8711–8716, 1993.
24. K.W. Sanggaard, H. Karring, Z. Valnickova, I.B. Thogersen, and J.J. Enghild, The TSG-6 and I{alpha} I interaction promotes a transesterification cleaving the protein–glycosaminoglycan–protein (PGP) cross-link, *J. Biol. Chem.* 280, 11936–11942, 2005.
25. M. Zhao, M. Yoneda, Y. Ohashi, S. Kurono, H. Iwata, Y. Ohnuki, and K. Kimata, Evidence for the covalent binding of SHAP, heavy chains of inter- alpha-trypsin inhibitor, to hyaluronan, *J. Biol. Chem.* 270, 26657–26663, 1995.
26. M.S. Rugg, A.C. Willis, D. Mukhopadhyay, V.C. Hascall, E. Fries, C. Fulop, C.M. Milner, and A.J. Day, Characterization of complexes formed between TSG-6 and inter-alpha-inhibitor that act as intermediates in the covalent transfer of heavy chains on to hyaluronan, *J. Biol. Chem.* 280, 25674–25686, 2005.
27. L. Huang, M. Yoneda, and K. Kimata, A serum-derived hyaluronan-associated protein (SHAP) is the heavy chain of the inter alpha-trypsin inhibitor, *J. Biol. Chem.* 268, 26725–26730, 1993.
28. L. Zhuo and K. Kimata, Cumulus oophorus extracellular matrix: Its construction and regulation, *Cell Struct. Funct.* 26, 189–196, 2001.
29. L. Chen, S.J.T. Mao, and W.J. Larsen, Identification of a factor in fetal bovine serum that stabilizes the cumulus extracellular matrix: A role for a member of the inter-a-trypsin inhibitor family, *J. Biol. Chem.* 267, 12380–12386, 1992.
30. D.J. Mahoney, C.D. Blundell, and A.J. Day, Mapping the hyaluronan-binding site on the link module from human tumor necrosis factor-stimulated gene-6 by site-directed mutagenesis, *J. Biol. Chem.* 276, 22764–22771, 2001.
31. N. Busso and J.A. Hamilton, Extravascular coagulation and the plasminogen activator/plasmin system in rheumatoid arthritis, *Arthritis Rheum.* 46, 2268–2279, 2002.
32. F.O. Akenami, V. Siren, M. Wessman, M. Koskiniemi, and A. Vaheri, Tissue plasminogen activator gene expression in multiple sclerosis brain tissue, *J. Neurol. Sci.* 165, 71–76, 1999.
33. M.W. Cunningham, Pathogenesis of group A streptococcal infections, *Clin. Microbiol. Rev.* 13, 470–511, 2000.
34. J.M. Rakic, C. Maillard, M. Jost, K. Bajou, V. Masson, L. Devy, V. Lambert, J.M. Foidart, and A. Noel, Role of plasminogen activator-plasmin system in tumor angiogenesis, *Cell Mol. Life Sci.* 60, 463–473, 2003.
35. W.J. Murdoch and M.L. Gottsch, Proteolytic mechanisms in the ovulatory folliculo-luteal transformation, *Connect Tissue Res.* 44, 50–57, 2003.
36. D.J. Mahoney, B. Mulloy, M.J. Forster, C.D. Blundell, E. Fries, C.M. Milner, and A.J. Day, Characterization of the interaction between tumor necrosis factor-stimulated Gene-6 and heparin: Implications for the inhibition of plasmin in extracellular matrix microenvironments, *J. Biol. Chem.* 280, 27044–27055, 2005.
37. A.I. Yudin, M.W. Li, K.R. Robertson, G.N. Cherr, and J.W. Overstreet, Characterization of the active site of monkey sperm hyaluronidase, *J. Reprod. Fertil.* 121, 735–743, 2001.
38. V. Doyen, M. Rubio, D. Braun, T. Nakajima, J. Abe, H. Saito, G. Delespesse, and M. Sarfati, Thrombospondin 1 is an autocrine negative regulator of human dendritic cell activation, *J. Exp. Med.* 198, 1277–1283, 2003.

39. S.J. Suchard, L.A. Boxer, and V.M. Dixit, Activation of human neutrophils increases thrombospondin receptor expression, *J. Immunol.* 147, 651–659, 1991.
40. A. Agah, T.R. Kyriakides, J. Lawler, and P. Bornstein, The lack of thrombospondin-1 (TSP1) dictates the course of wound healing in double-TSP1/TSP2-null mice, *Am. J. Pathol.* 161, 831–839, 2002.
41. J. Entwistle, S.W. Zhang, B.H. Yang, C. Wong, Q. Li, C.L. Hall, A. Jingbo, M. Mowat, A.H. Greenberg, and E.A. Turley, Characterization of the murine gene encoding the hyaluronan receptor RHAMM, *Gene* 163, 233–238, 1995.
42. E.A. Turley, D. Moore, and L.J. Hayden, Characterization of hyaluronate binding proteins isolated from 3T3 and murine sarcoma virus transformed 3T3 cells, *Biochemistry* 26, 2997–3005, 1987.
43. E.A. Turley, Purification of a hyaluronate-binding protein fraction that modifies cell social behavior, *Biochem. Biophys. Res. Commun.* 108, 1016–1024, 1982.
44. E.A. Turley, A.J. Belch, S. Poppema, and L.M. Pilarski, Expression and function of a receptor for hyaluronan-mediated motility on normal and malignant lymphocytes-B, *Blood* 81, 446–453, 1993.
45. C.L. Hall, L.A. Lange, D.A. Prober, S. Zhang, and E.A. Turley, pp60(c-src) is required for cell locomotion regulated by the hyaluronan receptor RHAMM, *Oncogene* 13, 2213–2224, 1996.
46. E.A. Turley, L. Austen, K. Vandeligt, and C. Clary, Hyaluronan and a cell-associated hyaluronan binding protein regulate the locomotion of ras-transformed cells, *J. Cell Biol.* 112, 1041–1047, 1991.
47. M. Crainie, A.R. Belch, M.J. Mant, and L.M. Pilarski, Overexpression of the receptor for hyaluronan-mediated motility (RHAMM) characterizes the malignant clone in multiple myeloma: Identification of three distinct RHAMM variants, *Blood* 93, 1684–1696, 1999.
48. B.H. Yang, L.Y. Zang, and E.A. Turley, Identification of 2 hyaluronan-binding domains in the hyaluronan receptor RHAMM, *J. Biol. Chem.* 268, 8617–8623, 1993.
49. B.H. Yang, B.L. Yang, R.C. Savani, and E.A. Turley, Identification of a common hyaluronan binding motif in the hyaluronan binding proteins rhamm, cd44 and link protein, *EMBO J.* 13, 286–296, 1994.
50. M.R. Ziebell, Z.G. Zhao, B. Luo, Y. Luo, E.A. Turley, and G.D. Prestwich, Peptides that mimic glycosaminoglycans: High-affinity ligands for a hyaluronan binding domain, *Chem. Biol.* 8, 1081–1094, 2001.
51. V. Assmann, J.F. Marshall, C. Fieber, M. Hofmann, and I.R. Hart, The human hyaluronan receptor RHAMM is expressed as an intracellular protein in breast cancer cells, *J. Cell Sci.* 111, 1685–1694, 1998.
52. C. Samuelsson and S. Gustafson, Studies on the interaction between hyaluronan and a rat colon cancer cell line, *Glycoconj. J.* 15, 169–175, 1998.
53. R.C. Savani, C. Wang, B.H. Yang, S.W. Zhang, M.G. Kinsella, T.N. Wight, R. Stern, D.M. Nance, and E.A. Turley, Migration of bovine aortic smooth muscle cells after wounding injury—the role of hyaluronan and RHAMM, *J. Clin. Invest.* 95, 1158–1168, 1995.
54. S.P. Evanko, W.T. Parks, and T.N. Wight, Intracellular hyaluronan in arterial smooth muscle cells: Association with microtubules, RHAMM, and the mitotic spindle, *J. Histochem. Cytochem.* 52, 1525–1536, 2004.
55. V.B. Lokeshwar and M.G. Selzer, Differences in hyaluronic acid-mediated functions and signaling in arterial, microvessel, and vein-derived human endothelial cells, *J. Biol. Chem.* 275, 27641–27649, 2000.
56. I. Hus, J. Rolinski, J. Tabarkiewicz, K. Wojas, A. Bojarska-Junak, J. Greiner, K. Giannopoulos, A. Dmoszynska, and M. Schmitt, Allogeneic dendritic cells pulsed with tumor lysates or apoptotic bodies as immunotherapy for patients with early-stage B-cell chronic lymphocytic leukemia, *Leukemia* 19, 1621–1627, 2005.
57. J. Greiner, M. Ringhoffer, M. Taniguchi, A. Schmitt, D. Kirchner, G. Krahn, V. Heilmann, J. Gschwend, L. Bergmann, H. Dohner, and M. Schmitt, Receptor for hyaluronan acid-mediated motility (RHAMM) is a new immunogenic leukemia-associated antigen in acute and chronic myeloid leukemia, *Exp. Hematol.* 30, 1029–1035, 2002.
58. T. Ahrens, V. Assmann, C. Fieber, C.C. Termeer, P. Herrlich, M. Hofmann, and J.C. Simon, CD44 is the principal mediator of hyaluronic-acid-induced melanoma cell proliferation, *J. Invest. Dermatol.* 116, 93–101, 2001.
59. C. Wang, A.D. Thor, D.H. Moore, Y. Zhao, R. Kerschmann, R. Stern, P.H. Watson, and E.A. Turley, The overexpression of RHAMM, a hyaluronan-binding protein that regulates ras signaling, correlates with overexpression of mitogen-activated protein kinase and is a significant parameter in breast cancer progression, *Clin. Cancer Res.* 4, 567–576, 1998.
60. V. Abetamann, H.F. Kern, and H.P. Elsasser, Differential expression of the hyaluronan receptors CD44 and RHAMM in human pancreatic cancer cells, *Clin. Cancer Res.* 2, 1607–1618, 1996.

61. Y. Akiyama, S. Jung, B. Salhia, S. Lee, S. Hubbard, M. Taylor, T. Mainprize, K. Akaishi, W. van Furth, and J.T. Rutka, Hyaluronate receptors mediating glioma cell migration and proliferation, *J. Neurooncol.* 53, 115–127, 2001.
62. A. Lugli, I. Zlobec, U. Gunthert, P. Minoo, K. Baker, L. Tornillo, L. Terracciano, and J.R. Jass, Overexpression of the receptor for hyaluronic acid mediated motility is an independent adverse prognostic factor in colorectal cancer, *Mod. Pathol.* 19, 1302–1309, 2006.
63. C.L. Hall, C. Wang, L.A. Lange, and E.A. Turley, Hyaluronan and the hyaluronan receptor RHAMM promote focal adhesion turnover and transient tyrosine kinase activity, *J. Cell Biol.* 126, 575–588, 1994.
64. P.H. Weigel, V.C. Hascall, and M. Tammi, Hyaluronan synthases, *J. Biol. Chem.* 272, 13997–14000, 1997.
65. S.P. Evanko and T.N. Wight, Intracellular localization of hyaluronan in proliferating cells, *J. Histochem. Cytochem.* 47, 1331–1341, 1999.
66. R. Tammi, K. Rilla, J.P. Pienimäki, D.K. Maccallum, M. Hogg, M. Luukkonen, V.C. Hascall, and M. Tammi, Hyaluronan enters keratinocytes by a novel endocytic route catabolism, *J. Biol. Chem.* 276, 35111–35122, 2001.
67. R. Stern, Devising a pathway for hyaluronan catabolism: Are we there yet?, *Glycobiology* 13, 105R–115R, 2003.
68. V. Assmann, D. Jenkinson, J.F. Marshall, and I.R. Hart, The intracellular hyaluronan receptor RHAMM/IHABP interacts with microtubules and actin filaments, *J. Cell Sci.* 112, 3943–3954, 1999.
69. C.A. Maxwell, J.J. Keats, M. Crainie, X. Sun, T. Yen, E. Shibuya, M. Hendzel, G. Chan, and L.M. Pilarski, RHAMM is a centrosomal protein that interacts with dynein and maintains spindle pole stability, *Mol. Biol. Cell* 14, 2262–2276, 2003.
70. S. Adamia, C.A. Maxwell, and L.M. Pilarski, Hyaluronan and hyaluronan synthases: Potential therapeutic targets in cancer, *Curr. Drug Targets Cardiovasc. Haematol. Disord.* 5, 3–14, 2005.
71. H. Li, L. Guo, J.W. Li, N. Liu, R. Qi, and J. Lui, Expression of hyaluronan receptors CD44 and RHAMM in stomach cancers: Relevance with tumor progression, *Int. J. Oncol.* 17, 927–932, 2000.
72. D.T. Rein, K. Roehrig, T. Schondorf, A. Lazar, M. Fleisch, D. Niederacher, H.G. Bender, and P. Dall, Expression of the hyaluronan receptor RHAMM in endometrial carcinomas suggests a role in tumour progression and metastasis, *J. Cancer Res. Clin. Oncol.* 129, 161–164, 2003.
73. S.L. Gares and L.M. Pilarski, Balancing thymocyte adhesion and motility: A functional linkage between beta1 integrins and the motility receptor RHAMM, *Dev. Immunol.* 7, 209–225, 2000.
74. V. Nehls and W. Hayen, Are hyaluronan receptors involved in three-dimensional cell migration?, *Histol. Histopathol.* 15, 629–636, 2000.
75. A. Masellissmith, A.R. Belch, M.J. Mant, E.A. Turley, and L.M. Pilarski, Hyaluronan-dependent moti lity of B cells and leukemic plasma cells in blood, but not of bone marrow plasma cells, in multiple myeloma: Alternate use of receptor for hyaluronan-mediated motility (RHAMM) and CD44, *Blood* 87, 1891–1899, 1996.
76. E.A. Turley, M.Z. Hossain, T. Sorokan, L.M. Jordan, and J.I. Nagy, Astrocyte and microglial motility in vitro is functionally dependent on the hyaluronan receptor RHAMM, *Glia* 12, 68–80, 1994.
77. P.H. Weigel, G.M. Fuller, and R.D. LeBoeuf, A model for the role of hyaluronic acid and fibrin in the early events during the inflammatory response and wound healing, *J. Theor. Biol.* 119, 219–234, 1986.
78. C.L. Hall and E.A. Turley, Hyaluronan: RHAMM mediated cell locomotion and signaling in tumorigenesis, *J. Neurooncol.* 26, 221–229, 1995.
79. E.A. Turley, P.W. Noble, and L.Y.W. Bourguignon, Signaling properties of hyaluronan receptors, *J. Biol. Chem.* 277, 4589–4592, 2002.
80. A.P. Spicer, M.L. Roller, S.A. Camper, J.D. McPherson, J.J. Wasmuth, S. Hakim, C. Wang, E.A. Turley, and J.A. McDonald, The human and mouse receptors for hyaluronan-mediated motility, RHAMM, genes (HMMR) map to human chromosome 5q33.2-qter and mouse chromosome 11, *Genomics* 30, 115–117, 1995.
81. H. Ponta, D. Wainwright, and P. Herrlich, The CD44 protein family, *Int. J. Biochem. Cell Biol.* 30, 299–305, 1998.
82. P. Teriete, S. Banerji, M. Noble, C.D. Blundell, A.J. Wright, A.R. Pickford, E. Lowe, D.J. Mahoney, M.I. Tammi, J.D. Kahmann, I.D. Campbell, A.J. Day, and D.G. Jackson, Structure of the regulatory hyaluronan binding domain in the inflammatory leukocyte homing receptor CD44, *Mol. Cell* 13, 483–496, 2004.
83. M.C. Levesque and B.F. Haynes, TNFα and IL-4 regulation of hyaluronan binding to monocyte CD44 involves posttranslational modification of CD44, *Cell. Immunol.* 193, 209–218, 1999.
84. K. Gee, M. Kozlowski, and A. Kumar, Tumor necrosis factor-{alpha} induces functionally active hyaluronan-adhesive CD44 by activating sialidase through p38 mitogen-activated protein kinase in lipopolysaccharide-stimulated human monocytic cells, *J. Biol. Chem.* 278, 37275–37287, 2003.

85. H. Kawashima, M. Hirose, J. Hirose, D. Nagakubo, A.H.K. Plaas, and M. Miyasaka, Binding of a large chondroitin sulfate/dermatan sulfate proteoglycan, versican, to L-selectin, P-selectin, and CD44, *J. Biol. Chem.* 275, 35448–35456, 2000.
86. T. Murai, N. Sougawa, H. Kawashima, K. Yamaguchi, and M. Miyasaka, CD44–chondroitin sulfate interactions mediate leukocyte rolling under physiological flow conditions, *Immunol. Lett.* 93, 163–170, 2004.
87. N. Toyama-Sorimachi, F. Kitamura, H. Habuchi, Y. Tobita, K. Kimata, and M. Miyasaka, Widespread expression of chondroitin sulfate-type serglycins with CD44 binding ability in hematopoietic cells, *J. Biol. Chem.* 272, 26714–26719, 1997.
88. S.O. Kolset and J.T. Gallagher, Proteoglycans in haemopoietic cells, *Biochim. Biophys. Acta* 1032, 191–211, 1990.
89. M. Grujic, T. Braga, A. Lukinius, M.L. Eloranta, S.D. Knight, G. Pejler, and M. Abrink, Serglycin-deficient cytotoxic T lymphocytes display defective secretory granule maturation and granzyme B storage, *J. Biol. Chem.* 280, 33411–33418, 2005.
90. S.M. Raja, B. Wang, M. Dantuluri, U.R. Desai, B. Demeler, K. Spiegel, S.S. Metkar, and C.J. Froelich, Cytotoxic cell granule-mediated apoptosis. Characterization of the macromolecular complex of granzyme B with serglycin, *J. Biol. Chem.* 277, 49523–49530, 2002.
91. S.S. Metkar, B. Wang, S. Aguilar, S.M. Raja, L. Uhlin-Hansen, E. Podack, J.A. Trapani, and C.J. Froelich, Cytotoxic cell granule-mediated apoptosis: Perforin delivers granzyme B-serglycin complexes into target cells without plasma membrane pore formation, *Immunity* 16, 417–428, 2002.
92. N. Toyamasorimachi and M. Miyasaka, A novel ligand for cd44 is sulfated proteoglycan, *Int. Immunol.* 6, 655–660, 1994.
93. N. Toyamasorimachi, H. Sorimachi, Y. Tobita, F. Kitamura, H. Yagita, K. Suzuki, and M. Miyasaka, A novel ligand for CD44 is serglycin, a hematopoietic cell lineage-specific proteoglycan—Possible involvement in lymphoid cell adherence and activation, *J. Biol. Chem.* 270, 7437–7444, 1995.
94. M. Culty, H.A. Nguyen, and C.B. Underhill, The hyaluronan receptor (CD44) participates in the uptake and degradation of hyaluronan, *J. Cell Biol.* 116, 1055–1062, 1992.
95. Q. Hua, C.B. Knudson, and W. Knudson, Internalization of hyaluronan by chondrocytes occurs via receptor-mediated endocytosis, *J. Cell Sci.* 106, 365–375, 1993.
96. A.C. Smith, J.T. Cirulis, J.E. Casanova, M.A. Scidmore, and J.H. Brumell, Interaction of the Salmonella-containing vacuole with the endocytic recycling system, *J. Biol. Chem.* 280, 24634–24641, 2005.
97. P.G. McGuire, J.J. Castellot, and R.W. Orkin, Size-dependent hyaluronate degradation by cultured cells, *J. Cell. Physiol.* 133, 267–276, 1987.
98. C.B. Underhill, H.A. Nguyen, M. Shizari, and M. Culty, CD44 positive macrophages take up hyaluronan during lung development, *Dev. Biol.* 155, 324–336, 1993.
99. D.J. Aguiar, W. Knudson, and C.B. Knudson, Internalization of the hyaluronan receptor CD44 by chondrocytes, *Exp. Cell Res.* 252, 292–302, 1999.
100. M. Culty, M. Shizari, E.W. Thompson, and C.B. Underhill, Binding and degradation of hyaluronan by human breast cancer cell lines expressing different forms of CD44: Correlation with invasive potential, *J. Cell Physiol.* 160, 275–286, 1994.
101. R. Tammi and M. Tammi, Influence of retinoic acid on the ultrastructure and hyaluronic acid synthesis of adult human epidermis in whole skin organ culture, *J. Cell. Physiol.* 126, 389–398, 1986.
102. R. Tammi and M. Tammi, Correlations between hyaluronan and epidermal proliferation as studied by glucosamine and thymidine incorporations and staining of hyaluronan on mitotic keratinocytes, *Exp. Cell Res.* 195, 524–527, 1991.
103. V.B. Lokeshwar, N. Fregien, and L.Y.W. Bourguignon, Ankyrin-binding domain of CD44(GP85) is required for the expression of hyaluronic acid-mediated adhesion function, *J. Cell Biol.* 126, 1099–1109, 1994.
104. S. Tsukita, K. Oishi, N. Sato, J. Sagara, A. Kawai, and S. Tsukita, ERM family members as molecular linkers between the cell surface glycoprotein CD44 and actin-based cytoskeletons, *J. Cell Biol.* 126, 391–401, 1994.
105. P.V. Jensen and L.I. Larsson, Actin microdomains on endothelial cells: Association with CD44, ERM proteins, and signaling molecules during quiescence and wound healing, *Histochem. Cell Biol.* 121, 361–369, 2004.
106. B. Formby and R. Stern, Phosphorylation stabilizes alternatively spliced CD44 mRNA transcripts in breast cancer cells: inhibition by antisense complementary to casein kinase II mRNA, *Mol. Cell Biochem.* 187, 23–31, 1998.
107. A.I. McClatchey and M. Giovannini, Membrane organization and tumorigenesis—the NF2 tumor suppressor, Merlin, *Genes Dev.* 19, 2265–2277, 2005.
108. Y. Bai, Y.J. Liu, H. Wang, Y. Xu, I. Stamenkovic, and Q. Yu, Inhibition of the hyaluronan-CD44 interaction by merlin contributes to the tumor-suppressor activity of merlin, *Oncogene* 26, 836–850, 2006.

109. P. Teder, R.W. Vandivier, D.H. Jiang, J.R. Liang, L. Cohn, E. Puré, P.M. Henson, and P.W. Noble, Resolution of lung inflammation by CD44, *Science* 296, 155–158, 2002.
110. F.L. Moffat, T. Han, Z.M. Li, M.D. Peck, R.E. Falk, P.B. Spalding, W. Jy, Y.S. Ahn, A.J. Chu, and L.Y. Bourguignon, Involvement of CD44 and the cytoskeletal linker protein ankyrin in human neutrophil bacterial phagocytosis, *J. Cell Physiol.* 168, 638–647, 1996.
111. E. Vachon, R. Martin, J. Plumb, V. Kwok, R.W. Vandivier, M. Gloguaer, A. Kapus, X. Wang, C.W. Chow, S. Grinstein, and G.P. Downey, CD44 is a phagocytic receptor, *Blood* 107, 4149–4158, 2006.
112. H. Jiang, C.B. Knudson, and W. Knudson, Antisense inhibition of CD44 tailless splice variant in human articular chondrocytes promotes hyaluronan internalization, *Arthritis Rheum.* 44, 2599–2610, 2001.
113. H. Jiang, R.S. Peterson, W.H. Wang, E. Bartnik, C.B. Knudson, and W. Knudson, A requirement for the CD44 cytoplasmic domain for hyaluronan binding, pericellular matrix assembly, and receptor-mediated endocytosis in COS-7 cells, *J. Biol. Chem.* 277, 10531–10538, 2002.
114. B. Dome, B. Somlai, A. Ladanyi, K. Fazekas, M. Zoller, and J. Timar, Expression of CD44v3 splice variant is associated with the visceral metastatic phenotype of human melanoma, *Virchows Arch.* 439, 628–635, 2001.
115. S. Maula, R.L. Huuhtanen, C.P. Blomqvist, T.A. Wiklund, P. Laurila, and R. Ristamaki, The adhesion molecule CD44v6 is associated with a high risk for local recurrence in adult soft tissue sarcomas, *Br. J. Cancer* 84, 244–252, 2001.
116. R. Singh, S. Subramanian, J.M. Rhodes, and B.J. Campbell, Peanut lectin stimulates proliferation of colon cancer cells by interaction with glycosylated CD44v6 isoforms and consequential activation of c-Met and MAPK: Functional implications for disease-associated glycosylation changes, *Glycobiology* 16, 594–601, 2006.
117. S.A. Khan, A.C. Cook, M. Kappil, U. Gunthert, A.F. Chambers, A.B. Tuck, and D.T. Denhardt, Enhanced cell surface CD44 variant (v6, v9) expression by osteopontin in breast cancer epithelial ells facilitates tumor migration: novel post-transcriptional, post-translational regulation, *Clin. Exp. Metastasis* 22, 663–673, 2005.
118. S.C. Hong, J.Y. Song, J.K. Lee, N.W. Lee, S.H. Kim, B.W. Yeom, and K.W. Lee, Significance of CD44v6 expression in gynecologic malignancies, *J. Obstet. Gynaecol. Res.* 32, 379–386, 2006.
119. Y.J. Liu, P.S. Yan, J. Li, and J.F. Jia, Expression and significance of CD44s, CD44v6, and nm23 mRNA in human cancer, *World J. Gastroenterol.* 11, 6601–6606, 2005.
120. N. Ohta, H. Saito, T. Kuzumaki, T. Takahashi, M.M. Ito, T. Saito, K. Nakahara, and M. Hiroi, Expression of CD44 in human cumulus and mural granulosa cells of individual patients in in-vitro fertilization programmes, *Mol. Hum. Reprod.* 5, 22–28, 1999.
121. H.S. Berner, B. Davidson, A. Berner, B. Risberg, and J.M. Nesland, Differential expression of CD44s and CD44v3–10 in adenocarcinoma cells and reactive mesothelial cells in effusions, *Virchows Arch.* 436, 330–335, 2000.
122. H.S. Berner, B. Davidson, A. Berner, B. Risberg, G.B. Kristensen, C.G. Trope, G. Van de Putte, and J.M. Nesland, Expression of CD44 in effusions of patients diagnosed with serous ovarian carcinoma—Diagnostic and prognostic implications, *Clin. Exp. Metastasis* 18, 197–202, 2000.
123. E.P. Reategui, A.A. de Mayolo, P.M. Das, F.C. Astor, R. Singal, K.L. Hamilton, W.J. Goodwin, K.L. Carraway, and E.J. Franzmann, Characterization of CD44v3-containing isoforms in head and neck cancer, *Cancer Biol. Ther.* 5, 1163–1168, 2006.
124. K. Wong, U. Rubenthiran, and S. Jothy, Motility of colon cancer cells: Modulation by CD44 isoform expression, *Exp. Mol. Pathol.* 75, 124–130, 2003.
125. L.N. Lee, S.H. Kuo, Y.C. Lee, Y.L. Chang, H.C. Chang, I.S. Jan, and P.C. Yang, CD44 splicing pattern is associated with disease progression in pulmonary adenocarcinoma, *J. Formos. Med. Assoc.* 104, 541–548, 2005.
126. U. Protin, T. Schweighoffer, W. Jochum, and F. Hilberg, CD44-deficient mice develop normally with changes in subpopulations and recirculation of lymphocytes subsets, *J. Immunol.* 163, 4917–4923, 1999.
127. G. Kaya, I. Rodriguez, J.L. Jorcano, P. Vassalli, and I. Stamenkovic, Selective suppression of CD44 in keratinocytes of mice bearing an antisense CD44 transgene driven by a tissue-specific promoter disrupts hyaluronate metabolism in the skin and impairs keratinocyte proliferation, *Genes Dev.* 11, 996–1007, 1997.
128. S. Nedvetzki, E. Gonen, N. Assayag, R. Reich, R.O. Williams, R.L. Thurmond, J.F. Huang, B.A. Neudecker, F.S. Wang, E.A. Turley, and D. Naor, RHAMM, a receptor for hyaluronan-mediated motility, compensates for CD44 in inflamed CD44-knockout mice: A different interpretation of redundancy, *Proc. Natl. Acad. Sci. USA* 101, 18081–18086, 2004.

129. S. Banerji, J. Ni, S.X. Wang, S. Clasper, J. Su, R. Tammi, M. Jones, and D.G. Jackson, LYVE-1, a new homologue of the CD44 glycoprotein, is a lymph- specific receptor for hyaluronan, *J. Cell Biol.* 144, 789–801, 1999.
130. R. Prevo, S. Banerji, D.J.P. Ferguson, S. Clasper, and D.G. Jackson, Mouse LYVE-1 is an endocytic receptor for hyaluronan in lymphatic endothelium, *J. Biol. Chem.* 276, 19420–19430, 2001.
131. J.S. Bonifacino and E.C. Dell'Angelica, Molecular bases for the recognition of tyrosine-based sorting signals, *J. Cell Biol.* 145, 923–926, 1999.
132. K.G. Rothberg, J.E. Heuser, W.C. Donzell, Y.S. Ying, J.R. Glenney, and R.G. Anderson, Caveolin, a protein component of caveolae membrane coats, *Cell* 68, 673–682, 1992.
133. C.M. Carreira, S.M. Nasser, E. di Tomaso, T.P. Padera, Y. Boucher, S.I. Tomarev, and R.K. Jain, LYVE-1 is not restricted to the lymph vessels: Expression in normal liver blood sinusoids and down-regulation in human liver cancer and cirrhosis, *Cancer Res.* 61, 8079–8084, 2001.
134. C. Cursiefen, L. Chen, L.P. Borges, D. Jackson, J. Cao, C. Radziejewski, P.A. D'Amore, M.R. Dana, S.J. Wiegand, and J.W. Streilein, VEGF-A stimulates lymphangiogenesis and hemangiogenesis in inflammatory neovascularization via macrophage recruitment, *J. Clin. Invest.* 113, 1040–1050, 2004.
135. C. Cursiefen, S. Ikeda, P.M. Nishina, R.S. Smith, A. Ikeda, D. Jackson, J.S. Mo, L. Chen, M.R. Dana, B. Pytowski, F.E. Kruse, and J.W. Streilein, Spontaneous corneal hem-and lymphangiogenesis in mice with destrin-mutation depend on VEGFR3 signaling, *Am. J. Pathol.* 166, 1367–1377, 2005.
136. P. Bono, V.M. Wasenius, P. Heikkila, J. Lundin, D.G. Jackson, and H. Joensuu, High LYVE-1-positive lymphatic vessel numbers are associated with poor outcome in breast cancer, *Clin. Cancer Res.* 10, 7144–7149, 2004.
137. L. Trojan, M.S. Michel, F. Rensch, D.G. Jackson, P. Alken, and R. Grobholz, Lymph and blood vessel architecture in benign and malignant prostatic tissue: Lack of lymphangiogenesis in prostate carcinoma assessed with novel lymphatic marker lymphatic vessel endothelial hyaluronan receptor (LYVE-1), *J. Urol.* 172, 103–107, 2004.
138. G.H. Cunnick, W.G. Jiang, K.F. Gomez, and R.E. Mansel, Lymphangiogenesis and breast cancer metastasis, *Histol. Histopathol.* 17, 863–870, 2002.
139. S.S. Sundar, H. Zhang, P. Brown, S. Manek, C. Han, K. Kaur, M.F. Charnock, D. Jackson, and T.S. Ganesan, Role of lymphangiogenesis in epithelial ovarian cancer, *Br. J. Cancer* 94, 1650–1657, 2006.
140. F. Gao, Y.M. Lu, M.L. Cao, Y.W. Liu, Y.Q. He, and Y. Wang, Expression and quantification of LYVE-1 in human colorectal cancer, *Clin. Exp. Med.* 6, 65–71, 2006.
141. D.G. Jackson, Hyaluronan and lymphedema, *Lymphology* 37, 1–5, 2004.
142. W.R. Jones, C.C. O'Morchoe, H.M. Jarosz, and P.J. O'Morchoe, Distribution of charged sites on lymphatic endothelium, *Lymphology* 19, 5–14, 1986.
143. J.R.E. Fraser, T.C. Laurent, A. Engstrom-Laurent, and U.B.G. Laurent, Elimination of hyaluronic acid from the blood stream in the human, *Clin. Exp. Pharmacol. Physiol.* 11, 17–25, 1984.
144. T.C. Laurent and J.R.E. Fraser, The properties and turnover of hyaluronan, *Ciba Found. Symp.* 124, 9–29, 1986.
145. J.R.E. Fraser, L.-E. Appelgren, and T.C. Laurent, Tissue uptake of circulating hyaluronic acid, *Cell Tissue Res.* 233, 285–293, 1983.
146. J.R.E. Fraser, T.C. Laurent, H. Pertoft, and E. Baxter, Plasma clearance, tissue distribution and metabolism of hyaluronic acid injected intravenously in the rabbit, *Biochem. J.* 200, 415–424, 1981.
147. T.C. Laurent, I.M.S. Dahl, L.B. Dahl, A. Engstrom-Laurent, S. Eriksson, J.R.E. Fraser, K.A. Granath, C. Laurent, U.B.G. Laurent, K. Lilja, H. Pertoft, B. Smedsrod, A. Tengblad, and O. Wik, The catabolic fate of hyaluronic acid, *Connect. Tissue Res.* 15, 33–41, 1986.
148. B. Zhou, J.A. Oka, A. Singh, and P.H. Weigel, Purification and subunit characterization of the rat liver endocytic hyaluronan receptor, *J. Biol. Chem.* 274, 33831–33834, 1999.
149. B. Zhou, J.A. Weigel, L.A. Fauss, and P.H. Weigel, Identification of the hyaluronan receptor for endocytosis (HARE), *J. Biol. Chem.* 275, 37733–37741, 2000.
150. B. Zhou, C.T. McGary, J.A. Weigel, A. Saxena, and P.H. Weigel, Purification and molecular identification of the human hyaluronan receptor for endocytosis, *Glycobiology* 13, 339–349, 2003.
151. B. Zhou, J.A. Weigel, A. Saxena, and P.H. Weigel, Molecular cloning and functional expression of the rat 175-kDa hyaluronan receptor for endocytosis, *Mol. Biol. Cell* 13, 2853–2868, 2002.
152. E.N. Harris, J.A. Weigel, and P.H. Weigel, Endocytic function, glycosaminoglycan specificity, and antibody sensitivity of the recombinant human 190 kDa HA receptor for endocytosis (HARE), *J. Biol. Chem.* 279, 36201–36209, 2004.

153. O. Politz, A. Gratchev, P.A.G. McCourt, K. Schledzewski, P. Guillot, S. Johansson, G. Svineng, P. Franke, C. Kannicht, J. Kzhyshkowska, P. Longati, F.W. Velten, and S. Goerdt, Stabilin-1 and-2 constitute a novel family of fasciclin-like hyaluronan receptor homologues, *Biochem. J.* 362, 155–164, 2002.
154. E.N. Harris, S. Kyosseva, J.A. Weigel, and P.H. Weigel, Expression, processing and glycosaminoglycan binding activity of the recombinant human 315-kDa HA receptor for endocytosis (HARE), *J. Biol. Chem.* 282, 2785–2797, 2006.
155. Y. Tamura, H. Adachi, J. Osuga, K. Ohashi, N. Yahagi, M. Sekiya, H. Okazaki, S. Tomita, Y. Iizuka, H. Shimano, R. Nagai, S. Kimura, M. Tsujimoto, and S. Ishibashi, FEEL-1 and FEEL-2 are endocytic receptors for advanced glycation end products, *J. Biol. Chem.* 278, 12613–12617, 2003.
156. H. Adachi and M. Tsujimoto, FEEL-1, a novel scavenger receptor with in vitro bacteria-binding and angiogenesis-modulating activities, *J. Biol. Chem.* 277, 34264–34270, 2002.
157. B. Smedsrod, S. Johansson, and S. Goerdt, Shooting HARE, *Glycobiology* 13, 11G–12G, 2003.
158. M. Falkowski, K. Schledzewski, B. Hansen, and S. Goerdt, Expression of stabilin-2, a novel fasciclin-like hyaluronan receptor protein, in murine sinusoidal endothelia, avascular, tissues, and at solid/liquid interfaces, *Histochem. Cell Biol.* 120, 361–369, 2003.
159. A. Lindsley, W. Li, J. Wang, N. Maeda, R. Rogers, and S.J. Conway, Comparison of the four mouse fasciclin-containing genes expression patterns during valvuloseptal morphogenesis, *Gene Exp. Patterns* 5, 593–600, 2005.
160. R.H. Raja, C.T. McGary, and P.H. Weigel, Affinity and distribution of surface and intracellular hyaluronic acid receptors in isolated rat liver endothelial cells, *J. Biol. Chem.* 263, 16661–16668, 1988.
161. J. Yannariello-Brown, S.J. Frost, and P.H. Weigel, Identification of the Ca^{2+}-independent endocytic hyaluronan receptor in rat liver sinusoidal endothelial cells using a photoaffinity cross-linking reagent, *J. Biol. Chem.* 267, 20451–20456, 1992.
162. C.T. McGary, R.H. Raja, and P.H. Weigel, Endocytosis of hyaluronic acid by rat liver endothelial cells: Evidence for receptor recycling, *Biochem. J.* 257, 875–884, 1989.
163. B. Hansen, P. Longati, K. Elvevold, G.-I. Nedredal, K. Schledzewski, R. Olsen, M. Falkowski, J. Kzhyshkowska, F. Carlsson, S. Johansson, B. Smedsrod, S. Goerdt, S. Johansson, and P. McCourt, Stabilin-1 and stabilin-2 are both directed into the early endocytic pathway in hepatic sinusoidal endothelium via interactions with clathrin/AP-2, independent of ligand binding, *Exp. Cell Res.* 303, 160–173, 2005.
164. R.H. Raja, M. Herzig, M. Grissom, and P.H. Weigel, Preparation and use of synthetic cell culture surfaces. A new reagent for the covalent immobilization of proteins and glycoproteins on a nonionic inert matrix, *J. Biol. Chem.* 261, 8505–8513, 1986.
165. B. Smedsrod, L. Kjellen, and H. Pertoft, Endocytosis and degradation of chondroitin sulphate by liver endothelial cells, *Biochem. J.* 229, 63–71, 1985.
166. I. Oynebraten, B. Hansen, B. Smedsrod, and L. Uhlin-Hansen, Serglycin secreted by leukocytes is efficiently eliminated from the circulation by sinusoidal scavenger endothelial cells in the liver, *J. Leukoc. Biol.* 67, 183–188, 2000.
167. B. Falkowska-Hansen, I. Oynebraten, L. Uhlin-Hansen, and B. Smedsrod, Endocytosis and degradation of serglycin in liver sinusoidal endothelial cells, *Mol. Cell Biochem.* 287, 43–52, 2006.
168. D.H. Bernanke and R.R. Markwald, Effects of two glycosaminoglycans on seeding of cardiac cushion tissue cells into a collagen–lattice culture system, *Anat. Rec.* 210, 25–31, 1984.
169. S.L. Shyng, S. Lehmann, K.L. Moulder, and D.A. Harris, Sulfated glycans stimulate endocytosis of the cellular isoform of the prion protein, PrPC in cultured cells, *J. Biol. Chem.* 270, 30221–30229, 1995.
170. J.A. Schroeder, L.F. Jackson, D.C. Lee, and T.D. Camenisch, Form and function of developing heart valves: Coordination by extracellular matrix and growth factor signaling, *J. Mol. Med.* 81, 392–403, 2003.
171. K.J. Grande-Allen, W.J. Mako, A. Calabro, Y. Shi, N.B. Ratliff, and I. Vesely, Loss of chondroitin 6-sulfate and hyaluronan from failed porcine bioprosthetic valves, *J. Biomed. Mater. Res.* 65A, 251–259, 2003.
172. E.N. Harris and P.H. Weigel, The ligand-binding profile of HARE/Stabilin-2: Hyaluronan and chondroitin sulfates A, C, and D bind to overlapping sites distinct from the sites for heparin, acetylated low-density lipoprotein, dermatan sulfate, and CS-E, *Glycobiology* Epub. 22 May, 2008.
173. E.N. Harris, J.A. Weigel, and P.H. Weigel, The human hyaluronan receptor for endocytosis (HARE) is a systemic clearance receptor for heparin, *J. Biol. Chem.* 283, Epub. 22 April, 2008.
174. M.S. Pandey, E.N. Harris, J.A. Weigel, and P.H. Weigel, The cytoplasmic domain of the hyaluronan receptor for endocytosis (HARE) contains multiple endocytic motifs targeting coated pit mediated internalization, *J. Biol. Chem.* 283, In press, 2008.

Part IV

Cell Adhesion and Cell Surface Lattice Formation

13 Galectin-3 and Cancer

Yi Wang, Vitaly Balan, and Avraham Raz

CONTENTS

13.1 INTRODUCTION

Galectins are members of animal lectins that bind β-galactosides through evolutionarily conserved sequence elements of the carbohydrate recognition domain (CRD) [1]. To date, 15 members of the galectin family have been identified, cloned, and classified into three subgroups based on their structure and numbers of the CRD: (1) prototype (galectin-1, -2, -5, -7, -10, -11, -13, -14, and -15); (2) chimera-type (galectin-3); and (3) tandem repeat type (galectin-4, -6, -8 -9, and -12), [2]. It has been suggested that these multivalent members of the galectin family might play different roles in various biological responses [3].

Galectin-3, previously described as IgE-binding protein, CBP35, CBP30, Mac-2, L-29, L-31, and L-34 plays an important role in tumorigenesis and progression through regulating cell proliferation, apoptosis, cell adhesion, invasion, angiogenesis, and metastasis by binding to the carbohydrate moiety of cell surface glycoproteins or glycolipids. In this chapter, multiple biological functions of galectin-3 in relation to tumorigenesis and its progression to malignant phenotype have been discussed.

13.2 STRUCTURE AND DISTRIBUTION OF GALECTIN-3

The coding sequence of galectin-3 gene produces a protein with 31 kDa molecular mass that comprises 250 amino acid residues. Galectin-3 has a unique chimera-type structure consisting of three different domains: an NH_2-terminal domain, a repeated collagen-like sequence rich in glycine, proline and tyrosine residues, and a COOH-terminal domain. The NH_2-terminal domain consists of 12 amino acids containing a serine phosphorylation site that regulates its cellular targeting. The collagen-like domain contains two matrix metalloproteinases (MMPs) recognition sites. The

COOH-terminal domain contains a single CRD consisting of 140 amino acid residues, which defines the molecular characteristics of galectin family [4–6].

Galectin-3 predominantly localizes in the cytoplasm, can import into the nucleus through at least two pathways: via passive diffusion and/or active transport [7,8] and secrete from the cytoplasm via the nonclassical secretion pathway since galectin-3 contains no consensuses signal sequence for either secretion or nuclear translocation [9,10].

13.3 GALECTIN-3 AND TUMOR CELL APOPTOSIS

Normal control of cell growth hinges on the delicate balance between cell proliferation and cell death. Dysregulation of either event can result in pathological consequences exemplified by diseases such as cancer, where unchecked proliferation results in continuous accumulation of cells. Apoptosis is a natural biological mechanism observed during cancer cell death after chemotherapy or radiotherapy. The resistance of tumor cells to apoptosis contributes to the increased malignancy of tumors and tumor progression.

Previous studies have shown that overexpression of galectin-3 correlated with increased resistance to apoptosis of cancer cells; however, more recent studies indicate its diverse effects on apoptosis depending on the intra- or extracellular localization of galectin-3.

13.3.1 Effect of Extracellular Galectin-3

It was reported that secreted extracellular galectin-3 could induce apoptosis of human T leukemia cell lines, human peripheral blood mononuclear cells, and activated mouse T cells. CD7 and CD29 (β1 integrin) were identified as the T cell surface glycoconjugate receptors for galectin-3, and the induction of apoptosis was triggered by extracellular galectin-3 binding to its receptors through carbohydrate-dependent interactions, which was inhibited by lactose, a specific sugar inhibitor. The proapoptotic signal generated by galectin-3 binding to its receptor is amplified within the cell into a cascade of actions that resulted in activation of mitochondrial apoptosis events including cytochrome *c* release and caspase-3 activation, but not caspase-8 activation.

However, the apoptosis sensitivity to galectin-3 varied among the different cell lines. Galectin-3 null Jurkat, CEM, and MOL T-4 cells were significantly more sensitive to exogenous galectin-3 than SKW6.4 and H9 cells, which express galectin-3, suggesting a cross-talk between the intracellular and extracellular galectin-3 in regulating apoptosis [11]. These results suggest that the induction of T-cell apoptosis by secreted extracellular galectin-3 may play a role in the immune escape mechanism, which contributes to tumor progression and metastasis.

13.3.2 Effect of Intracellular Galectin-3

It is well established that intracellular galectin-3 demonstrates antiapoptotic activity in response to various apoptotic stimuli in a wide variety of cells through diverse mechanisms in response to various apoptotic insults. It was reported that galectin-3 contains the NWGR amino acid sequence highly conserved in the BH1 domain of the bcl-2 gene family, which is implicated in its anti-apoptotic activity [12]. It was also found that phosphorylation of the Ser 6 of galectin-3 acts as a molecular switch for its cellular translocation from the nucleus to the cytoplasm and, as a result, regulates the antiapoptotic activity of galectin-3 [13]. However, the exact mechanisms on how it regulates the apoptotic activity are unclear. Three possible mechanisms have been suggested: (1) by affecting cell cycle processes; (2) by regulating mitochondrial integrity and cytochrome *c* release; and (3) by modulating Akt activity and the subsequent apoptosis-associated events (Figure 13.1).

The human breast carcinoma BT549 cell line has been widely used to investigate the function of galectin-3 since it does not express this protein. Cell–matrix interactions regulate many cellular

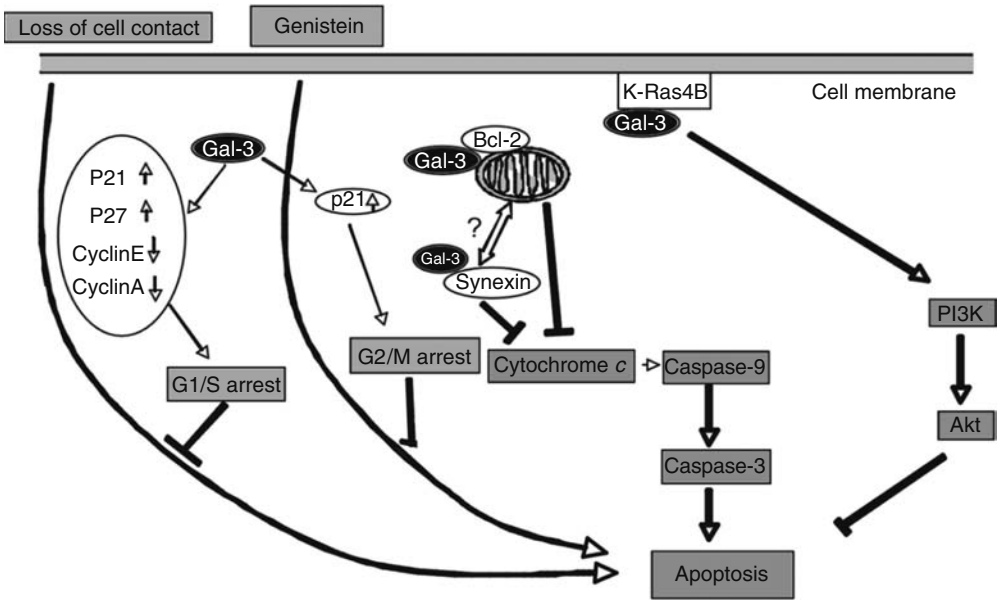


FIGURE 13.1 Antiapoptosis mechanisms of intracellular galectin-3. The three main mechanisms are (1) regulating the cell cycle process, (2) stabilizing mitochondrial integrity, and (3) inhibiting cytochrome *c* release and subsequent procaspases activation and adjusting the Akt signal transduction pathway.

responses including gene expression, differentiation, and cell survival [14–16]; it was hypothesized that the loss of cell–matrix interactions induce anoikis, a specific form of apoptosis, and cellular sensitivity to anoikis is associated with cell cycle regulation [17]. In contrast to the anoikis of BT549 cells induced by the loss of cell anchorage, galectin-3 overexpressing BT549 cells respond to the loss of cell adhesion by inducing G1 arrest without detectable cell death. Galectin-3-mediated G1 arrest involves downregulation of G1-S cyclin levels (cyclin E and cyclin A) and upregulation of their inhibitory protein levels ($p21^{WAF1/CIP1}$ and $p27^{KIP1}$). Rb protein becomes hypophosphorylated in galectin-3 overexpressing cells after the loss of cell anchorage, but galectin-3 induces cyclin D1 expression (an early G1 cyclin) and its associated kinase activity [18]. These results suggest that galectin-3 inhibition of anoikis may involve cell cycle arrest at an anoikis-insensitive point (late G1) through modulation of gene expression and activities of cell cycle regulators. Another report also demonstrated galectin-3-mediated inhibition of apoptosis through induction of cell cycle arrest. In another study, genistein, a natural isoflavonoid compound found in soy products, was used to induce apoptosis of BT549 cells. Genistein induced apoptosis and cell cycle arrest at G(2)/M in several cancer cell lines *in vitro*, which is associated with induction of $p21^{WAF1/CIP1}$. Genistein effectively induced apoptosis without detectable cell cycle arrest in BT549, but it induced cell cycle arrest at the G(2)/M phase without apoptosis induction in galectin-3-transfected BT549 cells. Interestingly, genistein induced $p21^{WAF1/CIP1}$ expression in galectin-3-expressing BT549 cells, but not in control BT549 cells undergoing apoptosis [19].

Another important mechanism of intracellular galectin-3 inhibition of apoptosis involves the influence of galectin-3 on mitochondrial integrity. By using the liver ischemia/reperfusion model, which induces nitrogen-free radical-mediated apoptosis and cell coculture system, galectin-3 overexpression was shown to protect BT549 cells against nitric oxide–induced apoptosis, which involved protection of mitochondrial integrity, inhibition of cytochrome *c* release and caspase activation [20,21]. Similar results were obtained in human prostate cancer LNCap cells, where the expression of galectin-3 in LNCap cells inhibited anticancer drug-induced apoptosis by inhibiting the mitochondria-mediated pathway [22]. To further understand how galectin-3 regulates mitochondrial stability, some research was done to investigate which molecular events contribute to

this process. It was reported that galectin-3 translocates to the perinuclear membrane following a variety of apoptotic stimuli. Confocal microscopy and biochemical analysis revealed that galectin-3 is enriched in the mitochondria and prevents mitochondrial damage and cytochrome *c* release. Using a yeast two-hybrid system, researchers screened for galectin-3-interacting proteins that regulate galectin-3 localization and antiapoptotic activity. Synexin, a Ca^{2+} and phospholipid-binding protein, was one of the proteins identified. Direct interaction between galectin-3 and synexin was confirmed by glutathione-*S*-transferase pull-down assay *in vitro*. It was also shown that galectin-3 failed to translocate to the perinuclear mitochondrial membranes when expression of synexin was downregulated, suggesting a role for synexin in galectin-3 trafficking. Furthermore, downregulation of synexin abolished antiapoptotic activity of galectin-3 [23]. Taken together, these results suggest that synexin mediates galectin-3 translocation to the perinuclear mitochondrial membranes, where it regulates mitochondrial integrity critical for the regulation of apoptosis.

In addition to the two mechanisms above, it was found that galectin-3 might play an antiapoptotic role in human bladder carcinoma J82 cells by enhancing Akt activity. Parental J82 and the control vector-transfected J82 cells exhibit low levels of constitutively active Akt, resulting in sensitivity to TRAIL-induced apoptosis, on the other hand, J82 cells overexpressing galectin-3 cells expressed a high level of constitutively active Akt and were resistant to TRAIL-induced apoptosis. The prote-cting effect of galectin-3 can be blocked by wortmannin and LY-294002, which are phosphatidylinositol 3-kinase (PI3K) inhibitors. Moreover, the blockage of TRAIL-induced apoptosis in J82 cells seemed to be mediated by Akt through the inhibition of BID cleavage [24]. These results suggest that galectin-3 involves Akt as a modulator molecule in protecting bladder carcinoma cells from TRAIL-induced apoptosis. However in human breast cancer BT549 cells, galectin-3 may regulate TRAIL-induced apoptosis in an opposite way. Studies showed that galectin-3 enhanced TRAIL-induced apoptosis in BT549 cells, which is mediated through dephosphorylation of Akt [25].

13.4 GALECTIN-3 AND TUMOR CELL PROLIFERATION

Growth-related expression and localization of galectin-3 suggest roles of this protein in the regulation of cell proliferation. Indeed, significant evidence has shown that galectin-3 is implicated in regulating tumor cell growth either by promoting or suppressing cell proliferation, in various cell lines.

Human leukemia T cells transfected with galectin-3 displayed higher growth rates than control nonexpressing transfectants [26]. Inhibition of galectin-3 gene expression in human pituitary adenoma cell line HP75 by RNA interference decreased cell proliferation [27]. Inhibition of galectin-3 expression led to loss of serum-independent growth, acquisition of growth inhibition properties by cell contact, abrogation of anchorage-independent growth, and a significant suppression of tumor growth in nude mice [28]. A similar method has been applied to a human thyroid papillary carcinoma cell line and results showed that the anchorage-independent growth and saturation density of the clones expressing antisense RNA were significantly suppressed compared with those of control clones [29]. Galectin-3 expression decreases with progression of prostate cancer and in contrast, it exerts functions more like an inhibitory molecule for prostate cancer. Overexpression of galectin-3 in human prostate cancer cell line LNCaP resulted in reduced cell proliferation and tumor growth in nude mice [30].

In most cases, the mechanism behind growth regulation by galectin-3 is not known. However, recent investigation shed light on this problem. One example is hensin, responsible for differentiation promoting activity and present in the extracellular matrix, was identified as a binding partner for galectin-3 and cooperates with galectin-3 to promote epithelial differentiation [31]. It was also found that galectin-3 induces cyclin D1 promoter activity in human breast epithelial cells independent of cell adhesion through multiple *cis*-elements, including the SP1 and CRE sites and that galectin-3

cooperates with pRb for cyclin D1 promoter activation [32]. This study reveals a cell growth promoting activity of galectin-3 through cyclin D1 induction.

13.5 GALECTIN-3 AND TUMOR METASTASIS

In the past, numerous reports have demonstrated that galectin-3 expression is correlated with tumor progression and metastasis in various cancers, but galectin-3 may play an inhibitory or inducing role in cancer metastasis depending on different cancer types and organ systems. For example, in gastric cancer, melanoma, and pancreatic cancer, galectin-3 functions more like a tumor suppressor gene product. Okada et al. found that reduced expression of galectin-3 in gastric cancer was significantly associated with lymph node metastasis, advanced stage, a poor prognosis, and multivariate analysis showed that galectin-3 expression was an independent prognostic factor [33]. In thin primary melanoma lesions, the expression of galectin-3 was higher than that in metastases and seemed to correlate inversely with the aggressiveness as estimated by the Breslow index, which is recognized as the main prognostic factor in melanoma [34]. In pancreatic adenocarcinoma, patient cases in the low galectin-3 expression group had a significant tendency to be at later stages, to have distant metastasis, and to have less differentiated tumors, compared with cases in the high galectin-3 expression group. Postoperative overall survival was worse in the low galectin-3 expression group than in the high galectin-3 expression group, and the risk ratio of prognosis was 2.06 among patients in the low galectin-3 expression group compared with the high galectin-3 expression group [35]. Reversely, in human breast cancer, colon cancer, and human brain tumors, galectin-3 is implicated in tumor progression and metastasis [14,36,37]. Human breast cancer cell line BT549, which is galectin-3 null, was transfected with galectin-3, after intrasplenic injection, galectin-3 transfected BT549 cells formed metastatic colonies in the liver, while galectin-3 null BT549 cells did not, and a single mutation in the NWGR motif of galectin-3 obliterated its metastatic capability [21]. It demonstrated that galectin-3 enhances metastatic potential of the human breast cancer cell line BT549, possibly through its bcl-2-like NWGR motif. In colorectal cancer, it was shown that compared with normal colorectal mucosa, galectin-3 expression in the primary lesions of the cancer was significantly increased, correlating with the progression of clinical stage, liver metastasis, venous invasion, and lymph node metastasis. The group showing strongly positive galectin-3 had a significantly poorer prognosis than the negative or weakly positive group in terms of disease-free survival [38]. Similarly, introduction of galectin-3 antisense into metastatic colon cancer cells (LSLiM6, HM7) resulted in a marked decrease in liver colonization and spontaneous metastasis by LSLiM6 and HM7 cells, whereas upregulation of galectin 3 resulted in increased metastasis by LS174T cells [36]. It was also demonstrated that expression of the endogenous galectin-3 correlates with the malignant potential and metastasis of tumors in the central nervous system. Glioblastomas (Grade 4 astrocytomas) stained strongly for galectin-3, anaplastic astrocytomas (Grade 3) exhibited intermediate expression, whereas low-grade astrocytomas (Grade 2) did not express the endogenous galectin-3. Normal brain tissue and benign tumors did not express galectin-3, whereas metastases to the brain were all positive for galectin-3 expression. Metastases expressed significantly more galectin-3 than the primary tumors from which they were derived [14].

Though it has been indicated that galectin-3 plays an important role in tumor progression and metastasis, the mechanisms underlying the metastasis-associated role of galectin-3 is unclear and needs to be elucidated. Metastasis is a major fatal complication associated with malignancies. In many cases, tumor cells are disseminated through the blood flow. This transvenous metastasis is a complex process, involving many cell–cell and cell–extracellular matrix interactions. The process of metastasis requires the following steps: Growth at the primary site, angiogenesis at the primary site, detachment from the primary site, invasion through the extracellular matrix, dissemination of cells through the blood flow, tumor embolous formation in capillaries, and extravasation and growth at the secondary sites [39,40]. Here, we made a summary of the possible mechanisms through which galectin-3 impacts the cancer metastasis according to these above process.

13.5.1 Galectin-3 and Angiogenesis

Angiogenesis in the microenvironment of cancer is required for tumor growth at the primary site and provides a gateway to the dissemination through the blood flow. Carbohydrate recognition may have a pivotal role in angiogenesis because angiogenic factors like fibroblast growth factor and vascular endothelial growth factor bind initially to the extracellular matrix proteoglycans before binding to their cognate receptors [41]. Overexpression of galectin-3 in LNCaP, a galectin-3-negative human prostate cancer cell line, induced *in vivo* tumor growth and angiogenesis [42]. Galectin-3 stimulates capillary tube formation of human umbilical vein endothelial cells (HUVEC) *in vitro* and angiogenesis *in vivo*, which was inhibited by specific sugars and antibodies. Direct carbohydrate dependent binding of galectin-3 to HUVEC and stimulation of chemotaxis was also demonstrated [41]. To interfere with galectin–carbohydrate interactions during tumor progression, a current challenge is the design of specific galectin inhibitors for therapeutic purposes. Some researchers reported the synthesis of three novel low molecular weight synthetic lactulose amines (SLA): (1) *N*-lactulose-octamethylenediamine (LDO), (2) *N,N′*-dilactulose-octamethylenediamine (D-LDO), and (3) *N,N′*-dilactulose- dodecamethylenediamine (D-LDD). These compounds showed a differential ability to inhibit binding of galectin-3 to the highly glycosylated protein 90 K in solid-phase assays. In addition, each compound demonstrated selective regulatory effect in different events linked to tumor progression including endothelial cell morphogenesis and angiogenesis [43]. A plant polysaccharide, modified citrus pectin (MCP), was shown to inhibit galectin-3-mediated functions and is reported to inhibit tumor growth, angiogenesis, and metastasis [44–46].

13.5.2 Galectin-3 and Cell–Matrix Interaction

Alterations of tumor cell interactions with laminin, a basement membrane glycoprotein, are consistent features of the invasive and metastatic phenotype. Qualitative and quantitative changes in the expression of laminin-binding proteins have been correlated with the ability of cancer cells to cross basement membranes during the metastatic cascade.

Most of the mammalian cells express various cell adhesion molecules, which mediate binding of cells to extracellular matrix, homotypic cell adhesion, and heterotypic adhesion. Galectin-3 is present on the cell surface and in the extracellular environment [47] and it is regarded as a nonintegrin laminin-binding protein [48]. Galectin-3 on the cell surface fails to modulate adhesion of melanoma cells to laminin [49]. Soluble galectin-3 does not alter melanoma cell adhesion to laminin but oligomerized galectin-3 induces melanoma cell spreading on laminin [50]. Galectin-3 is the major laminin-binding protein on human colon carcinoma cells and its surface expression suggests involvement in metastasis. Poorly differentiated cell lines and DLD-2, a signet-ring carcinoma cell line, are characterized as aggressive cell lines because they adhere to and invade through reconstituted basement membrane significantly better than well-differentiated cell lines. Poorly differentiated cell lines and DLD-2 expressed more surface galectin-3 than well-differentiated cell lines [51], which may be responsible for invasion of colon carcinoma cells. Galectin-3 overexpression in human nonsmall cell lung carcinoma cell line DLKP results in enhanced adhesion to extracellular matrix components, cell motility, and *in vitro* invasiveness [52]. Galectin-3 transfected human breast carcinoma cell line BT549 adhered much more rapidly to laminin- and collagen IV-coated wells than the galectin-3 null expressing BT-549 cells. These cells were also able to invade through matrigel-coated polycarbonate filters at approximately three times the rate of BT-549 parental cells. In addition, galectin-3 expression may modulate the surface expression of some of the integrins specific for laminin and collagen IV adhesion and invasion of basement membrane by breast carcinoma cells [53].

13.5.3 Galectin-3 and Homotypic and Heterotypic Cell Adhesion

The formation of secondary tumors by circulating cancer cells correlates with an increased tendency of the cells to form emboli by aggregation with other tumor cells in microcapillaries followed by

extravasation at secondary sites. In the first step of extravasation, cells bind to endothelial cells followed by penetrating through the layers of endothelial cells and basement membrane. This binding between circulating cells and endothelial cells is mediated by specific interactions between cell surface lectins and carbohydrates present on glycoproteins, glycolipids, and glycosaminoglycans.

It was shown that cell surface galectin-3 mediates homotypic cell adhesion by binding to soluble complementary glycoconjugates [54]. Mac-2-binding protein, originally identified as a ligand for galectin-3 [55], is a heavily N-glycosylated secreted protein and induces homotypic aggregation of melanoma cells. This aggregation is inhibited by lactose or antigalectin-3 antibody. This suggests that Mac-2-binding protein interacts with galectin-3 on tumor cell surfaces, resulting in the formation of multicell aggregation [54].

Interactions of metastatic cancer cells with vasculatory endothelium are critical during early stages of cancer metastasis. The Thomsen–Friedenreich antigen (T antigen), which has β-galactose as the terminal residue, is involved in adhesion of tumor cells to endothelium. It was demonstrated that cancer-associated carbohydrate T antigen plays a leading role in docking breast and prostate cancer cells onto endothelium by specifically interacting with cell surface galectin-3. Importantly, T antigen-bearing glycoproteins are also capable of mobilizing galectin-3 to the surface of endothelial cells, thus priming them for harboring metastatic cancer cells [56]. It was found that both galectin-1 and galectin-3 participate in the adhesion of the human breast carcinoma cells MDA-MB-435 to the endothelium. Recombinant galectin-3 binds to T antigen in a dose-dependent manner. Galectin-3 clusters on endothelial cells at the sites of contact with tumor cells, consistent with its possible interaction with T antigen on cancer cells; the galectin-1 signal, however, strongly accumulated at the sites of cell–cell contacts, predominantly on tumor cells. Inhibitory peptides or antibody against T antigen inhibited homotypic aggregation of tumor cells and heterotypic adhesion of tumor cells to endothelial cells [57,58]. These results suggest that galectin-3 on cell the surface mediates metastasis-associated adhesion of tumor cells to endothelial cells by binding to cancer-specific T antigen, which is dependent on carbohydrate-binding domain. Significance of galectin-3 in homotypic and heterotypic cell–cell interactions was also demonstrated by using three-dimensional cocultures of endothelial and epithelial cells [37], which was inhibited by MCP [44].

13.6 CLINICAL SIGNIFICANCE OF GALECTIN-3 EXPRESSION IN TUMORS

A series of clinical evidence has been reported to support correlation between galectin-3 expression and malignant transformation. Consequently, diagnostic and prognostic significance of galectin-3 has been shown for some types of tumors. Metastasis is one of the most important causes of death in patients with malignant tumors, and tumor biomarkers that can predict metastasis and prognosis are important for clinical management of malignancies. Iurisci et al. reported that circulating levels of galectin-3 in the sera of patients with breast, gastrointestinal, lung, or ovarian cancer, melanoma, and Hodgkin's lymphoma were elevated and circulating levels of galectin-3 reflect biological aspects of tumor behavior associated with a metastasizing phenotype [59]. Gillenwater et al. reported that galectin-3 is localized to superficial mucosal layers in head and neck squamous cell carcinoma (HNSCC), and adjacent to keratin pearls in invasive carcinoma [17]. Xu et al. revealed that malignant thyroid tumors of epithelial cell origin and metastatic lymph nodes of papillary carcinomas exhibit high levels of galectin-3 as well as galectin-1, while neither benign thyroid tumors nor normal adjacent tissues express them [60]. Irimura et al. revealed a direct correlation in human colon cancers between galectin-3 expression and tumor progression [61]. In our study on colon cancers, galectin-3 expression elevates, according to Dukes' stage, and its overexpression correlates with decreased survival of patients [53]. In contrast, Castronovo et al. reported that normal breast tissue expressed high levels of galectin-3 by immunoperoxidase staining, whereas galectin-3 was downregulated in breast cancer cells [62], however Shekhar et al. demonstrated elevated levels of galectin-3 mRNA and protein in the luminal epithelial cells of normal

and benign ducts, downregulation in early grades of ductal carcinoma in situ (DCIS) and reexpression in peripheral tumor cells as DCIS lesions progressed to comedo-DCIS and invasive carcinomas [37]. In ovarian cancer, van den Brule et al. reported that galectin-3 expression was downregulated in cancer cells compared with the normal tissues [63]. We have revealed that galectin-3 is also downregulated in prostate cancer. Approximately 60%–70% of the normal tissue examined demonstrated heterogenous expression of galectin-3, while in stage II tumors, there was a dramatic decrease in galectin-3 expression in both PIN and tumor sections, with only 10.5% of these samples expressing this protein [64]. Previous studies also suggested that the pattern of galectin-3 expression changed in the process of prostatic tumorigenesis and progression. It was reported that the profiles of galectin-3 localization in PIN cells and prostate cancer cells were very different from that observed in normal prostatic epithelial cells. Galectin-3 immunostaining in PIN cells and prostate cancer cells were mostly detected in the cytoplasm compared to that of the normal prostatic epithelial cells in both the nucleus and cytoplasm [65]. These results suggest that galectin-3 has tumor progressive or inhibitory effects depending on tumor cell types. The variations in galectin-3 expression *in vivo* may depend on tumor-specific factors. Studies have been performed on the ability of the natural inhibitor of galectin-3, CP, and MCP to inhibit cancer growth and metastasis, and it was reported that rats fed on a 15% CP-enriched diet showed a higher apoptotic index in their colon [66]. Reduced ^{3}H-thymidine incorporation into

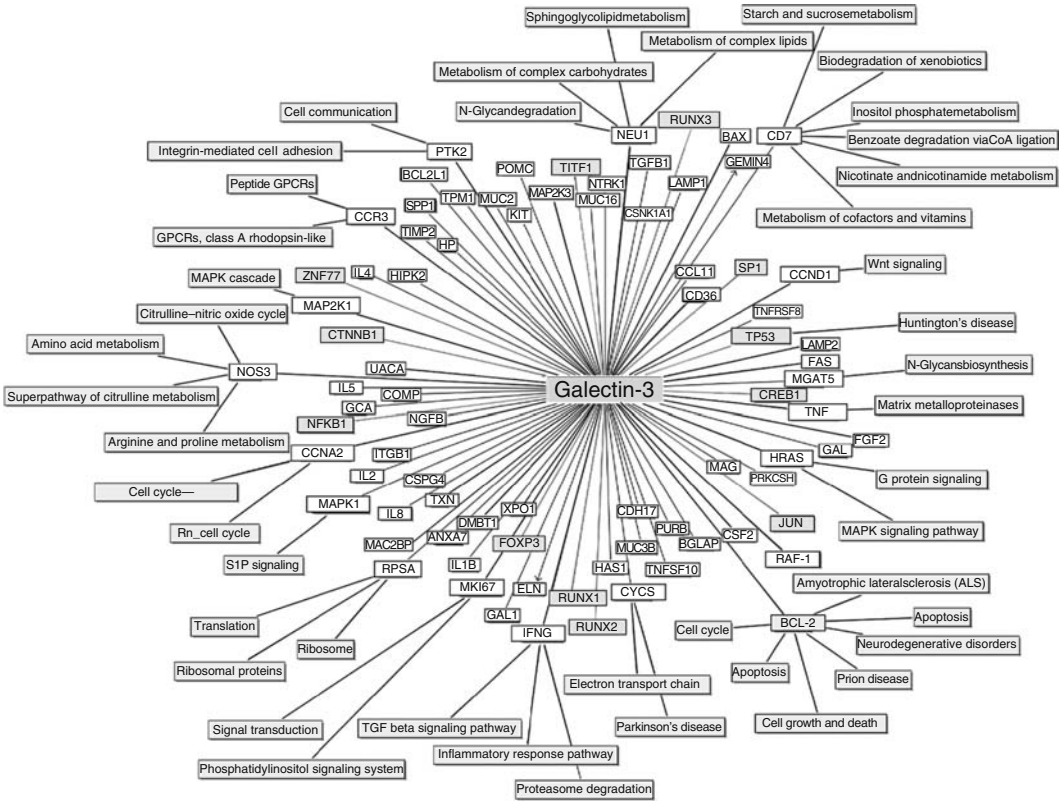


FIGURE 13.2 (See CD for color figure.) Galectin-3 network. The figure created using BiblioSphere software (Genomatix) represents most of the known proteins belonging to the galectin-3 network. Proteins shown in blue boxes are either transcription factors directly regulating the galectin-3 expression, proteins which change transcription activity after interaction with galectin-3, or proteins that have great impact on metabolic pathways. In most cases, the evidence for transcriptional regulation was confirmed by MatInspector software. Other proteins in the figure change their activity after interaction with galectin-3 and have great impact on metabolic pathways, processes, or diseases (yellow boxes).

DNA and, correspondingly, cell growth was reported when human prostatic JCA-1 cells were grown in media containing MCP [46,67]. Daily oral administration of MCP reduced the growth of implanted colon-25 tumor in Balb/c mice and dietary pectin reduced the growth of intramuscularly transplanted mouse tumors [68]. In phase II clinical trials on colorectal carcinoma patients, MCP showed a positive clinical activity, where five out of 23 patients showed tumor stabilization for a period of 2–6 months, and tumor shrinkage in one patient.

13.7 CONCLUSION AND FUTURE DIRECTIONS

To date, many studies have suggested that galectin-3 is a multifunctional protein implicated in various biological events including cell growth, apoptosis, cell adhesion, angiogenesis, and tumor metastasis. Here, we summarized that galectin-3 is involved in regulating apoptosis of tumor cells, but its effect is discrepant depending on its cellular distribution. Abundant research has confirmed the antiapoptotic role of intracellular galectin-3 in various cancers. However, extracellular galectin-3 exerts proapoptotic function in T-cells. More studies need be focused on the role of extracellular galectin-3 in different cell types to investigate if the proapoptotic role of extracellular galectin-3 is a universal phenomenon and to explore the mechanisms. It was shown that most of the intracellular functions of galectin-3 involve protein–protein interactions rather than protein–carbohydrate recognition as expected of a lectin, but most of the extracellular functions of galectin-3 are mediated by protein–carbohydrate interactions. Interactions between galectin-3 and other molecules trigger different signal transduction pathways and subsequent effects. In this review, we also summarized that galectin-3 can either promote or suppress tumor cells proliferation depending on the cell types and that galectin-3 is implicated in cancer metastasis and tumor progression, but in some cancers, it induces the metastasis process, while in other cases it inhibits the process. All of this suggests that galectin-3 can display diverse functions, being tumorigenic in some cases while protecting from cancer progression in other cases. In order to clarify a definitive function of galectin-3, assessment of galectin-3-associated molecules in biological phenomena is required. Previous studies have reported some galectin-3-associated molecules in various cell biological processes (Figure 13.2). To further understand the function of galectin-3, it will become more important to explore the signal transduction pathway triggered by interaction between galectin-3 and associated molecules.

ACKNOWLEDGMENTS

We thank Victor Hogan for his critical reading and editing of the manuscript. We also thank Vivian Powell for manuscript preparation and editing. This manuscript was supported by NIH R37CA46120 to A.R.

REFERENCES

1. Barondes, S.H., et al., Galectins: A family of animal beta-galactoside-binding lectins. *Cell*, 1994. 76(4): 597–598.
2. Hirabayashi, J., et al., Oligosaccharide specificity of galectins: A search by frontal affinity chromatography. *Biochim Biophys Acta*, 2002. 1572(2–3): 232–254.
3. Sacchettini, J.C., L.G. Baum, and C.F. Brewer, Multivalent protein–carbohydrate interactions. A new paradigm for supermolecular assembly and signal transduction. *Biochemistry*, 2001. 40(10): 3009–3015.
4. Barondes, S.H., et al., Galectins. Structure and function of a large family of animal lectins. *J Biol Chem*, 1994. 269(33): 20807–20810.
5. Gong, H.C., et al., The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells. *Cancer Res*, 1999. 59(24): 6239–6245.
6. Herrmann, J., et al., Primary structure of the soluble lactose binding lectin L-29 from rat and dog and interaction of its non-collagenous proline-, glycine-, tyrosine-rich sequence with bacterial and tissue collagenase. *J Biol Chem*, 1993. 268(35): 26704–26711.

7. Nakahara, S., et al., Importin-mediated nuclear translocation of galectin-3. *J Biol Chem*, 2006. 281(51): 39649–39659.
8. Nakahara, S., et al., Characterization of the nuclear import pathways of galectin-3. *Cancer Res*, 2006. 66(20): 9995–10006.
9. Hughes, R.C., Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim Biophys Acta*, 1999. 1473(1): 172–185.
10. Moutsatsos, I.K., et al., Endogenous lectins from cultured cells: Nuclear localization of carbohydrate-binding protein 35 in proliferating 3T3 fibroblasts. *Proc Natl Acad Sci USA*, 1987. 84(18): 6452–6456.
11. Fukumori, T., et al., CD29 and CD7 mediate galectin-3-induced type II T-cell apoptosis. *Cancer Res*, 2003. 63(23): 8302–8311.
12. Akahani, S., et al., Galectin-3: A novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family. *Cancer Res*, 1997. 57(23): 5272–5276.
13. Takenaka, Y., et al., Nuclear export of phosphorylated galectin-3 regulates its antiapoptotic activity in response to chemotherapeutic drugs. *Mol Cell Biol*, 2004. 24(10): 4395–4406.
14. Bresalier, R.S., et al., Expression of the endogenous galactose-binding protein galectin-3 correlates with the malignant potential of tumors in the central nervous system. *Cancer*, 1997. 80(4): 776–787.
15. Konstantinov, K.N., B.A. Robbins, and F.T. Liu, Galectin-3, a beta-galactoside-binding animal lectin, is a marker of anaplastic large-cell lymphoma. *Am J Pathol*, 1996. 148(1): 25–30.
16. Lotan, R., et al., Expression of a 31-kDa lactoside-binding lectin in normal human gastric mucosa and in primary and metastatic gastric carcinomas. *Int J Cancer*, 1994. 56(4): 474–480.
17. Gillenwater, A., et al., Expression of galectins in head and neck squamous cell carcinoma. *Head & Neck*, 1996. 18(5): 422–432.
18. Kim, H.R., et al., Cell cycle arrest and inhibition of anoikis by galectin-3 in human breast epithelial cells. *Cancer Res*, 1999. 59(16): 4148–4154.
19. Lin, H.M., et al., Galectin-3 mediates genistein-induced G(2)/M arrest and inhibits apoptosis. *Carcinogenesis*, 2000. 21(11): 1941–1945.
20. Moon, B.K., et al., Galectin-3 protects human breast carcinoma cells against nitric oxide-induced apoptosis: Implication of galectin-3 function during metastasis. *Am J Pathol*, 2001. 159(3): 1055–1060.
21. Song, Y.K., T.R. Billiar, and Y.J. Lee, Role of galectin-3 in breast cancer metastasis: Involvement of nitric oxide. *Am J Pathol*, 2002. 160(3): 1069–1075.
22. Fukumori, T., et al., Galectin-3 regulates mitochondrial stability and antiapoptotic function in response to anticancer drug in prostate cancer. *Cancer Res*, 2006. 66(6): 3114–3119.
23. Yu, F., et al., Galectin-3 translocates to the perinuclear membranes and inhibits cytochrome *c* release from the mitochondria. A role for synexin in galectin-3 translocation. *J Biol Chem*, 2002. 277(18): 15819–15827.
24. Oka, N., et al., Galectin-3 inhibits tumor necrosis factor-related apoptosis-inducing ligand-induced apoptosis by activating Akt in human bladder carcinoma cells. *Cancer Res*, 2005. 65(17): 7546–7553.
25. Lee, Y.J., et al., Reconstitution of galectin-3 alters glutathione content and potentiates TRAIL-induced cytotoxicity by dephosphorylation of Akt. *Exp Cell Res*, 2003. 288(1): 21–34.
26. Yang, R.Y., D.K. Hsu, and F.T. Liu, Expression of galectin-3 modulates T-cell growth and apoptosis. *Proc Natl Acad Sci USA*, 1996. 93(13): 6737–6742.
27. Riss, D., et al., Differential expression of galectin-3 in pituitary tumors. *Cancer Res*, 2003. 63(9): 2251–2255.
28. Honjo, Y., et al., Down-regulation of galectin-3 suppresses tumorigenicity of human breast carcinoma cells. *Clin Cancer Res*, 2001. 7(3): 661–668.
29. Yoshii, T., et al., Galectin-3 maintains the transformed phenotype of thyroid papillary carcinoma cells. *Int J Oncol*, 2001. 18(4): 787–792.
30. Ellerhorst, J.A., et al., Effects of galectin-3 expression on growth and tumorigenicity of the prostate cancer cell line LNCaP. *Prostate*, 2002. 50(1): 64–70.
31. Hikita, C., et al., Induction of terminal differentiation in epithelial cells requires polymerization of hensin by galectin 3. *J Cell Biol*, 2000. 151(6): 1235–1246.
32. Lin, H.M., et al., Galectin-3 enhances cyclin D(1) promoter activity through SP1 and a cAMP-responsive element in human breast epithelial cells. *Oncogene*, 2002. 21(52): 8001–8010.
33. Okada, K., et al., Reduced galectin-3 expression is an indicator of unfavorable prognosis in gastric cancer. *Anticancer Res*, 2006. 26(2B): 1369–1376.
34. Vereecken, P., et al., Expression of galectin-3 in primary and metastatic melanoma: Immunohistochemical studies on human lesions and nude mice xenograft tumors. *Arch Dermatol Res*, 2005. 296(8): 353–358.

35. Shimamura, T., et al., Clinicopathological significance of galectin-3 expression in ductal adenocarcinoma of the pancreas. *Clin Cancer Res*, 2002. 8(8): 2570–2575.
36. Bresalier, R.S., et al., Metastasis of human colon cancer is altered by modifying expression of the beta-galactoside-binding protein galectin 3. *Gastroenterology*, 1998. 115(2): 287–296.
37. Shekhar, M.P., et al., Alterations in galectin-3 expression and distribution correlate with breast cancer progression: functional analysis of galectin-3 in breast epithelial–endothelial interactions. *Am J Pathol*, 2004. 165(6): 1931–1941.
38. Nakamura, M., et al., Involvement of galectin-3 expression in colorectal cancer progression and metastasis. *Int J Oncol*, 1999. 15(1): 143–148.
39. Kohn, E.C., Development and prevention of metastasis. *Anticancer Res*, 1993. 13(6B): 2553–2559.
40. Raz, A. and R. Lotan, Endogenous galactoside-binding lectins: A new class of functional tumor cell surface molecules related to metastasis. *Cancer Metastasis Rev*, 1987. 6(3): 433–452.
41. Nangia-Makker, P., et al., Galectin-3 induces endothelial cell morphogenesis and angiogenesis. *Am J Pathol*, 2000. 156(3): 899–909.
42. Califice, S., et al., Dual activities of galectin-3 in human prostate cancer: Tumor suppression of nuclear galectin-3 vs tumor promotion of cytoplasmic galectin-3. *Oncogene*, 2004. 23(45): 7527–7536.
43. Rabinovich, G.A., et al., Synthetic lactulose amines: Novel class of anticancer agents that induce tumor-cell apoptosis and inhibit galectin-mediated homotypic cell aggregation and endothelial cell morphogenesis. *Glycobiology*, 2006. 16(3): 210–220.
44. Nangia-Makker, P., et al., Inhibition of human cancer cell growth and metastasis in nude mice by oral intake of modified citrus pectin. *J Natl Cancer Inst*, 2002. 94(24): 1854–1862.
45. Pienta, K.J., et al., Inhibition of spontaneous metastasis in a rat prostate cancer model by oral administration of modified citrus pectin [see comment]. *J Natl Cancer Inst*, 1995. 87(5): 348–353.
46. Platt, D. and A. Raz, Modulation of the lung colonization of B16-F1 melanoma cells by citrus pectin. *J Natl Cancer Inst*, 1992. 84(6): 438–442.
47. Cherayil, B.J., S.J. Weiner, and S. Pillai, The Mac-2 antigen is a galactose-specific lectin that binds IgE. *J Exp Med*, 1989. 170(6): 1959–1972.
48. Woo, H.J., et al., The major non-integrin laminin binding protein of macrophages is identical to carbohydrate binding protein 35 (Mac-2). *J Biol Chem*, 1990. 265(13): 7097–7099.
49. van den Brule, F.A., et al., Galectin-3, a laminin binding protein, fails to modulate adhesion of human melanoma cells to laminin. *Neoplasma*, 1995. 42(5): 215–219.
50. van den Brule, F.A., F.T. Liu, and V. Castronovo, Transglutaminase-mediated oligomerization of galectin-3 modulates human melanoma cell interactions with laminin. *Cell Adhes Commun*, 1998. 5(6): 425–435.
51. Lee, E.C., et al., Carbohydrate-binding protein 35 is the major cell-surface laminin-binding protein in colon carcinoma. *Arch Surg*, 1991. 126(12): 1498–1502.
52. O'Driscoll, L., et al., Galectin-3 expression alters adhesion, motility and invasion in a lung cell line (DLKP), in vitro. *Anticancer Res*, 2002. 22(6A): 3117–3125.
53. Warfield, P.R., et al., Adhesion of human breast carcinoma to extracellular matrix proteins is modulated by galectin-3. *Invasion Metastasis*, 1997. 17(2): 101–112.
54. Inohara, H., et al., Interactions between galectin-3 and Mac-2-binding protein mediate cell–cell adhesion. *Cancer Res*, 1996. 56(19): 4530–4534.
55. Rosenberg, I., et al., Mac-2-binding glycoproteins. Putative ligands for a cytosolic beta-galactoside lectin. *J Biol Chem*, 1991. 266(28): 18731–18736.
56. Glinsky, V.V., et al., The role of Thomsen-Friedenreich antigen in adhesion of human breast and prostate cancer cells to the endothelium. *Cancer Res*, 2001. 61(12): 4851–4857.
57. Glinsky, V.V., et al., Effects of Thomsen–Friedenreich antigen-specific peptide P-30 on beta-galactoside-mediated homotypic aggregation and adhesion to the endothelium of MDA-MB-435 human breast carcinoma cells. *Cancer Res*, 2000. 60(10): 2584–2588.
58. Khaldoyanidi, S.K., et al., MDA-MB-435 human breast carcinoma cell homo- and heterotypic adhesion under flow conditions is mediated in part by Thomsen–Friedenreich antigen–galectin-3 interactions. *J Biol Chem*, 2003. 278(6): 4127–4134.
59. Iurisci, I., et al., Concentrations of galectin-3 in the sera of normal controls and cancer patients. *Clin Cancer Res*, 2000. 6(4): 1389–1393.
60. Xu, X.C., A.K. el-Naggar, and R. Lotan, Differential expression of galectin-1 and galectin-3 in thyroid tumors. Potential diagnostic implications. *Am J Pathol*, 1995. 147(3): 815–822.
61. Irimura, T., et al., Increased content of an endogenous lactose-binding lectin in human colorectal carcinoma progressed to metastatic stages. *Cancer Res*, 1991. 51(1): 387–393.

62. Castronovo, V., et al., Decreased expression of galectin-3 is associated with progression of human breast cancer. *J Pathol*, 1996. 179(1): 43–48.
63. van den Brule, F.A., et al., Expression of the 67-kD laminin receptor, galectin-1, and galectin-3 in advanced human uterine adenocarcinoma. *Hum Pathol*, 1996. 27(11): 1185–1191.
64. Pacis, R.A., et al., Decreased galectin-3 expression in prostate cancer. *Prostate*, 2000. 44(2): 118–123.
65. Van den Brule, F.A., et al., Alteration of the cytoplasmic/nuclear expression pattern of galectin-3 correlates with prostate carcinoma progression. *Int J Cancer*, 2000. 89(4): 361–367.
66. Avivi-Green, C., Z. Madar, and B. Schwartz, Pectin-enriched diet affects distribution and expression of apoptosis-cascade proteins in colonic crypts of dimethylhydrazine-treated rats. *Int J Mol Med*, 2000. 6(6): 689–698.
67. Hsieh, T.C. and J.M. Wu, Changes in cell growth, cyclin/kinase, endogenous phosphoproteins and nm23 gene expression in human prostatic JCA-1 cells treated with modified citrus pectin. *Biochem Mol Biol Int*, 1995. 37(5): 833–841.
68. Taper, H.S., N.M. Delzenne, and M.B. Roberfroid, Growth inhibition of transplantable mouse tumors by non-digestible carbohydrates. *Int J Cancer*, 1997. 71(6): 1109–1112.

14 Myelin-Associated Glycoprotein (Siglec-4): A Nervous System Lectin That Regulates Axon–Myelin Stability and Axon Regeneration

Ronald L. Schnaar and Niraj R. Mehta

CONTENTS

14.1 INTRODUCTION

Myelin-associated glycoprotein (MAG, Siglec-4) is a multifunctional signaling protein that sits at the critical interface between axons—the thin nerve cell extensions that carry action potentials between nerve cells—and the multilamellar myelin wrap that ensheathes them [1,2]. Myelin acts not only to insulate but also to nurture axons, to regulate axon cytoarchitecture, and to help direct the local distribution of ion channels required for rapid nerve conduction. MAG plays key signaling roles that stabilize axon–myelin interactions [3,4], regulate the structure of myelin and axons [5–7], and control axon regeneration after injury [8,9]. Some of MAG's functions depend on its lectin activity, which will be described more fully in a subsequent section.

14.2 MAG STRUCTURE, DISTRIBUTION, AND FUNCTIONS

MAG is a cell surface-expressed member of the immunoglobulin-like (Ig) superfamily, with five extracellular Ig-like domains, a single transmembrane domain, and one of two alternative short cytoplasmic tails [1]. It is heavily glycosylated, with about 30% of its ~100 kDa mass represented by N-linked glycans [10]. MAG is produced only by myelinating glial cells: Oligodendrocytes in the central nervous system (CNS) and Schwann cells in the peripheral nervous system (PNS). Although MAG is a quantitatively minor myelin protein, comprising 1% of CNS and 0.1% of PNS myelin protein, it is not expressed uniformly throughout myelin. Although most of multilayered myelin consists of highly compacted membrane wraps, MAG is not expressed in compacted myelin. In the CNS, MAG is expressed primarily on the inner-most (periaxonal) noncompacted myelin wrap (see Figure 14.1) [11], whereas in the PNS, it is found on periaxonal myelin and other noncompacted myelin (paranodal loops, Schmitt–Lanterman incisures, and the mesaxon) [12].

MAG is coded by a single gene that is highly conserved among vertebrates [13]. Human and rodent MAG are 95% identical at the amino acid level over the entire extracellular expressed domain (five immunoglobulin-like domains, ~500 amino acids); human and fish (*Danio rerio)* MAG share >40% sequence identity. In mammals, MAG has two alternatively spliced forms, referred to as short (S-MAG) and long (L-MAG) [14]. They have identical extracellular Ig-like and transmembrane domains, differing only in the length and C-terminal sequence of their intracellular domains. Whereas L-MAG is predominant in the developing CNS, S-MAG dominates in the PNS throughout development and its relative expression in the CNS increases with development until it is as abundant as L-MAG in the adult [15]. The functional implications of expressing two forms of MAG have not been fully explored, although a selective mutation of L-MAG expression is revealing (see below). After injury, the extracellular domain MAG may also be released into the local environment by a proteolytic cleavage within the juxtamembrane sequence [16].

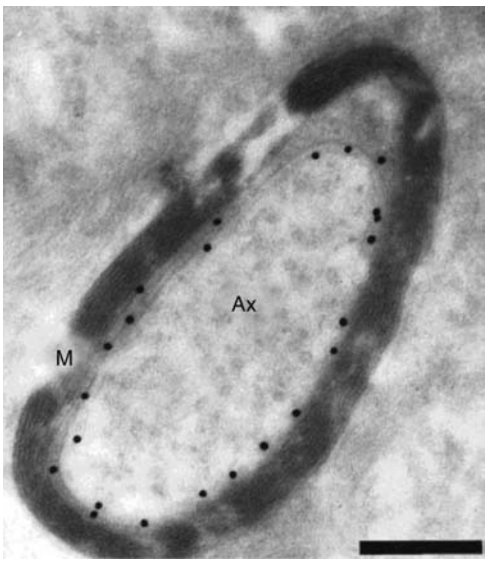


FIGURE 14.1 Ultrastructural location of MAG in the mouse optic nerve. Expression of MAG in a mature cross-section of a myelinated optic nerve axon. MAG expression is restricted to the periaxonal space. Primary anti-MAG antibodies were visualized with gold-conjugated secondary antibody. Ax, axon; M, myelin. Scale bar, 0.2 µm. (Reproduced from Schachner, M. and Bartsch, U., *Glia*, 29, 154, 2000; Bartsch, U., Kirchhoff, F., and Schachner, M., *J. Comp. Neurol*., 284, 451, 1989. With permission.)

Because MAG is expressed on the wrap of myelin directly apposed to the axon surface (Figure 14.1) [11], MAG has long been hypothesized to mediate axon–myelin interactions [1]. MAG functions have been revealed by both *in vitro* cell biological studies and *in vivo* genetic mutations. Mice engineered to carry a disrupted *Mag* gene lack MAG protein expression yet produce abundant myelin, indicating that MAG is not required for the formation of functional compact myelin [5,6]. However, *Mag*-null mice have a reduced number of myelinated axons and an increased number of unmyelinated axons in the CNS compared to wild-type mice [17,18], as well as aberrant forms of myelination including multiple myelinated axons, redundant myelin loops, and local myelin overproduction. The structure of the first myelin loop adjacent to the axon is also altered in *Mag*-null mice, with a significant loss of the characteristic periaxonal cytoplasmic collar [5,6,18]. These data infer that MAG plays a role in regulating myelination, presumably by modulating oligodendrocyte and Schwann cell function upon contact with the axon. Some of MAG's roles in oligodendrocytes (in the CNS) involve its intracellular domain, as evidenced by CNS (but not PNS) dysmyelination phenotypes in mice with normal MAG levels, but engineered to lack the unique intracellular polypeptide extension found on L-MAG. In addition to its effects on myelination and myelin structures, the absence of MAG has profound affects on axons in the CNS and PNS.

As *Mag*-null mice age, they exhibit highly significant axon degeneration [3,4], displaying a late-onset progressive axonal atrophy and increased Wallerian degeneration in both the PNS and CNS [2,3,19,20]. Additional evidence for MAG's role in nurturing axons comes from evaluation of MAG effects on axonal neurofilaments [7]. Myelinated segments of axons have a greater diameter (and thereby greater conductivity) than unmyelinated segments. The increased diameter is due to an increase in phosphorylation of major axonal cytoskeletal components, neurofilaments. Myelinated axons in *Mag*-null mice have ultrastructural characteristics of unmyelinated axons and do not display myelin-mediated neurofilament phosphorylation, increased neurofilament spacing, or increased axon caliber [7,21,22]. Evidence that the effect of MAG on neurofilament phosphorylation is mediated by direct contact between MAG and neurons comes from an experiment in which rat dorsal root ganglion neurons (DRGN) were cultured on MAG-expressing cells *in vitro*, resulting in a fivefold increase in neurofilament phosphorylation [21]. The relevance of this mechanism to human pathology is apparent from the finding of decreased neurofilament spacing in nerve biopsy specimens from patients with anti-MAG-associated neuropathy [23]. These data imply that MAG binds to complementary receptors on the axon surface to induce intraaxonal signals that regulate the axon cytoskeleton and nurture the axon.

A primary function of myelin is to enhance nerve conduction velocity. This function depends on myelin's insulating properties, but also on the highly structured longitudinal gaps in myelin called nodes of Ranvier, where voltage-dependent sodium channels cluster [24]. Maintenance of proper architecture at nodes of Ranvier and adjacent paranodal regions is essential for rapid saltatory nerve conduction. MAG, along with other myelin molecules, participates in defining the distribution of axon molecules at nodes of Ranvier [25]. The distribution of juxtaparanodal ion channels and cell recognition proteins on axons is altered in MAG mutants, compromising nerve conduction.

In addition to helping regulate myelination, stabilizing axon–myelin interactions and nurturing axons, MAG also inhibits axon regeneration [8,9]. Axons in the CNS fail to extend beyond an injury site even after a carefully controlled "microcrush" experimental lesion [26]. However, CNS axons retain the ability to regenerate when exposed to a supportive local environment, such as peripheral nerve bridges [27]. A wealth of research has led to the principle that endogenous factors in the local environment inhibit axon regeneration, thwarting the axon's intrinsic ability to regenerate [28]. The distal end of the injured axon is exposed to diverse inhibitors in the glial scar that forms at the injury site, and on myelin that persists at the injury site [28]. At least four axon regeneration inhibitors (ARIs) have been identified, including MAG [8,9], Nogo [29], oligodendrocyte–myelin glycoprotein [30], and chondroitin sulfate proteoglycans [31], each of which inhibits axon outgrowth *in vitro*. In each case, the inhibitor binds to complementary target ligands on the axon–nerve cell surface to initiate signaling that results in axonal cytoskeletal changes that halt axon outgrowth [28].

Evidence that MAG is an axon regeneration inhibitor came from *in vitro* cell biological studies. When proteins were solubilized from myelin membranes and separated chromatographically, some fractions inhibited neurite outgrowth *in vitro* [9]. The major inhibitory peak coincided with the elution of MAG and purified MAG was found to inhibit axon outgrowth. Neurons seeded on monolayers of fibroblasts engineered to express MAG had reduced neurite outgrowth [8]. Many studies using a variety of neurite and axon outgrowth models *in vitro* subsequently established MAG as an inhibitor and have implicated specific neuronal receptors, signaling pathways, and modulators of MAG's inhibitory activity [32–39] (see the following discussion). Studies *in vivo* demonstrate a modest but significant enhancement of axon regeneration in the absence of MAG, with *Mag*-null mice, but not wild-type mice, extending small numbers of long axons beyond an experimental lesion in the corticospinal tract [40]. Mice lacking MAG also demonstrate an increase in axon regeneration along residual peripheral myelin sheathes in mice that fail to efficiently clear myelin from PNS injury sites [41]. The modest degree of axon regeneration in *Mag*-null mice is indicative of other animal models in which only one of the four known ARIs is genetically ablated or blocked [42–45]. Although the relative role of MAG as an inhibitor of axon regeneration *in vivo* has yet to be established, most current models include MAG as a major contributor to the inhibitory environment of the CNS [28,46,47]. These models primarily invoke an inhibitory role for full-length MAG on residual myelin, although proteolytically released soluble forms of MAG may also be inhibitory [48,49].

Presumably, MAG's effects on axon–myelin stability and axon cytoarchitecture, as well as MAG's inhibition of axon outgrowth after injury, are initiated when MAG binds directly to complementary receptors on the surface of axons or growth cones. Direct binding of MAG to neurons was first demonstrated using MAG reconstituted in fluorescent liposomes, which bound specifically to cultured DRGN and spinal cord neurons *in vitro* [50,51]. When DRGN from newborn mice were cultured on fibroblasts expressing full-length MAG, axon outgrowth was altered [51]. MAG's binding to neurons and its function in altering axon outgrowth were both blocked by a specific anti-MAG monoclonal antibody. These findings first demonstrated functional neuronal receptors for MAG, the full identity of which is still under debate 20 years later.

As a member of the Siglec family of sialic acid-binding lectins, it is reasonable to propose that MAG binding to sialoglycoconjugates on the axon surface is responsible for some, if not most, of MAG's effects. A description of MAG's glycan binding specificity and evidence for the functional role of MAG's glycan binding are addressed in the following sections.

14.3 MAG (SIGLEC-4) GLYCAN-BINDING SPECIFICITY

MAG is a member of the Siglec family of lectins [52,53]. Sialic acid-recognizing Ig-superfamily lectins (Siglecs) are a family of 13 proteins (in humans) evolved to bind to terminal sialic acid residues of cell surface glycans. Each Siglec has specificity for the extended glycan on which the sialic acid is expressed, comprised of the sialic acid, its linkage position to the penultimate sugar, and the other sugar residues in the sialoglycan chain. Initial glycan recognition studies revealed that MAG bound to native but not desialylated human erythrocytes, confirming its ability to bind sialic acids [54]. Reconstitution of erythrocyte sialoglycans with different sialyltransferases revealed that MAG bound selectively to a NeuAcα2–3Galβ1–3GalNAc terminus [54,55]. In the nervous system, this terminus is expressed most abundantly on gangliosides GD1a and GT1b (see structures, Figure 14.2) [56]. The same trisaccharide terminus is also found on O-linked glycoproteins, which are much less abundant than gangliosides in the brain [57].

Studies using a variety of natural and synthetic glycolipids as receptors (Figure 14.3) revealed that MAG binds with high affinity to the major brain gangliosides GD1a and GT1b, and with yet higher affinity to rare "α-series" gangliosides, which bear an additional sialic acid-linked α2–6 to the GalNAc, the simplest being GD1α (NeuAcα2–3Galβ1–3[NeuAcα2–6]GalNAcβ1–4Galβ1–4Glcβ1–1′Cer). Whereas MAG binding required the structural fidelity of the terminal sialic acid

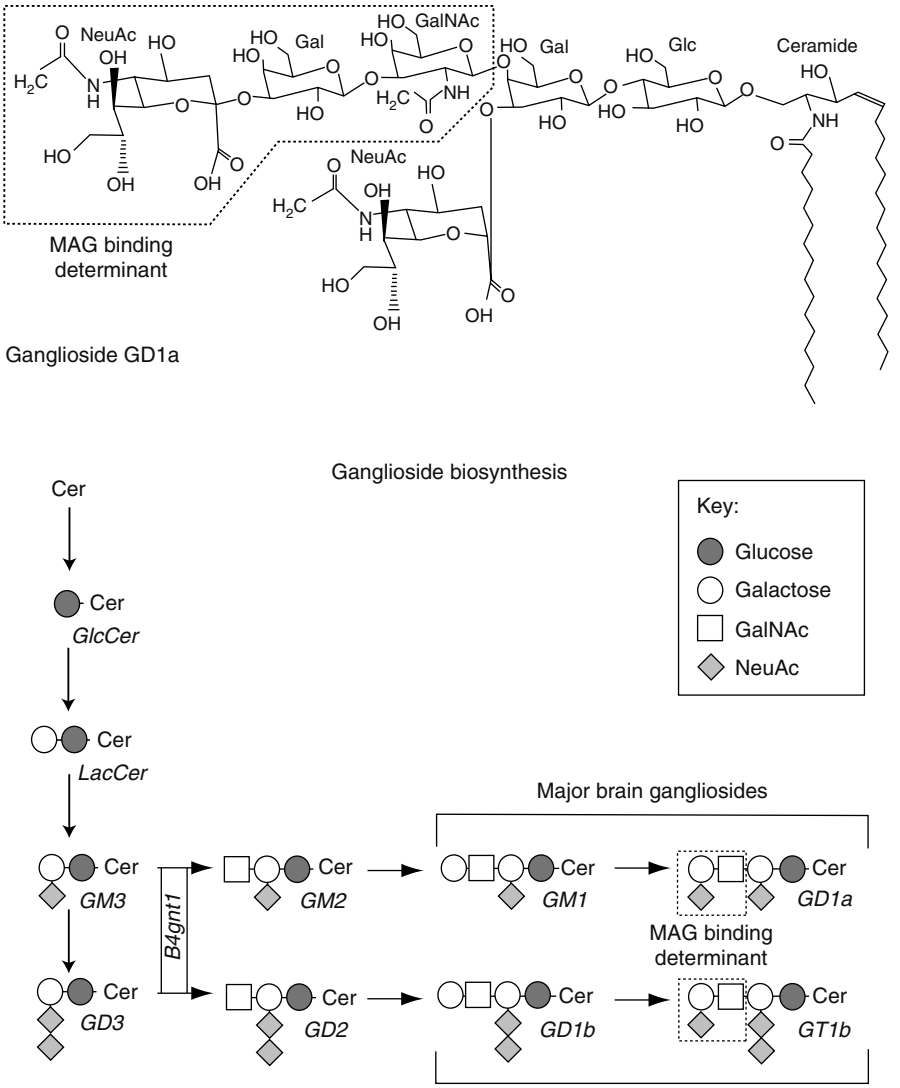


FIGURE 14.2 (See CD for color figure.) MAG-binding sialoglycan determinant on gangliosides. (Upper panel) Major brain ganglioside GD1a. The characteristic neutral core of brain "ganglio-series" gangliosides (Galβ1–3GalNAcβ1–4Galβ1–4Glcβ1–1′ceramide) carries two sialic acids, one on the 3-hydroxyl of each galactose residue. The key MAG-binding determinant, NeuAcα2–3Galβ1–3GalNAc is outlined. (Lower panel) Stepwise biosynthesis of the major brain gangliosides starting from ceramide. Different glycosyltransferases act sequentially, first on ceramide then on the growing glycosphingolipid glycan chain, to create the four major brain gangliosides common to all mammals. The biosynthetic step encoded by the *N*-acetylgalactosaminyltransferase gene *B4galnt1* (previously designated *Galgt1* or *GalNAcT*) is noted. Mice lacking a functional *B4galnt1* gene express the same total amount of brain ganglioside as wild type mice, but as the precursor structures GM3 and GD3. These mice, which lack the MAG-binding determinant on gangliosides, have a similar phenotype to mice lacking MAG.

(see below), the internal sialic acids of GD1a and GD1α could be replaced with sulfate groups without decreasing MAG-binding affinity [58]. These data imply that the primary glycan-binding site of MAG has high structural specificity for NeuAcα2–3Galβ1–3GalNAc, and that adjacent anionic charges on the glycan structure enhance binding significantly, perhaps due to complementary cationic patches on the surface of MAG near the primary sialic acid-binding site. The accuracy of this prediction will have to await enhanced structural data; the crystallographic structure of MAG

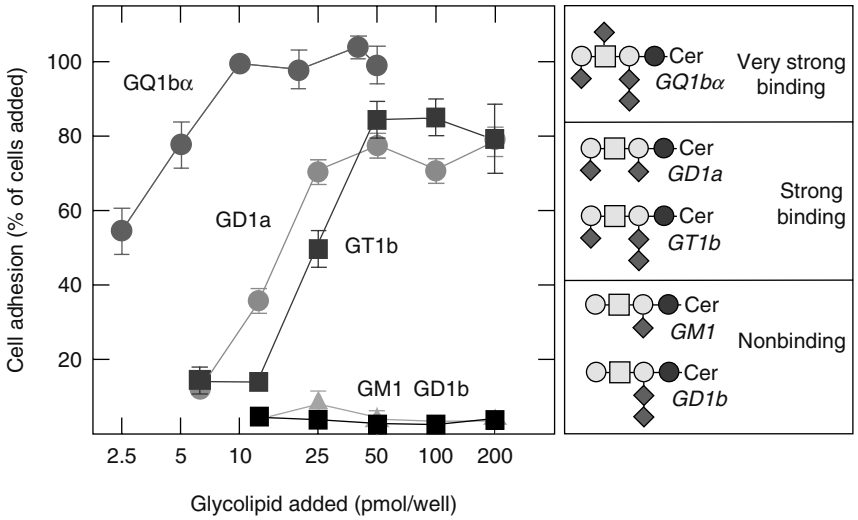


FIGURE 14.3 (See CD for color figure.) MAG-mediated cell adhesion to gangliosides: structural specificity. COS cells (fibroblasts) transfected to express full length MAG were placed in microwells adsorbed with the indicated gangliosides. MAG-mediated cell adhesion is expressed as a percent of the MAG-transfected cells added to each well. Ganglioside structures are shown schematically using the key in Figure 14.2. (From Collins, B.E., Kiso, M., Hasegawa, A., Tropak, M.B., Roder, J.C., Crocker, P.R., and Schnaar, R.L., *J. Biol. Chem.*, 272, 16889, 1997; Yang, L.J.S., Zeller, C.B., Shaper, N.L., Kiso, M., Hasegawa, A., Shapiro, R.E., and Schnaar, R.L., *Proc. Natl Acad. Sci. USA*, 93, 814, 1996. With permission.)

has not been reported to date. However, by analogy with the crystal structure of Siglec-1 [59,60], and based on site-directed mutagenesis [61], the primary sialoglycan-binding site is on the N-terminal Ig-like domain, with the sialic acid carboxylate making a salt bridge to arginine 118, although a secondary sialoglycan-binding site has been implicated based on functional studies [62] (see the following discussion).

In terms of its selectivity and requirements for sialic acid substructure, MAG is among the most stringent sialic acid-binding lectins thus far studied. Any synthetic alteration in the terminal sialic acid, including the absence or acetylation of any of the sialic acid hydroxyl groups, esterification of the carboxyl group, or hydroxylation of the *N*-acetyl group (to *N*-glycolyl) resulted in sharply decreased MAG binding [63–66]. Saturation transfer difference nuclear magnetic resonance (NMR) confirms the close association of the terminal sialic acid residue in a synthetic sialoglycan bound to MAG [66].

The glycan specificity of MAG was confirmed and extended using soluble sialosides as inhibitors of MAG–sialoglycan binding in an enzyme-linked immunosorbent assay (ELISA)-type assay [67]. In a screen of several monovalent sialoglycans, the two with highest affinity for MAG were a disialylated tetrasaccharide akin to the GD1α terminus, NeuAcα2–3Gal(NeuAcα2–6)GalNAc-R (K_i = 0.3 μM, disialyl T-antigen), and a trisaccharide akin to the GD1a/GT1b terminus, NeuAcα2–3Galβ1–3GalNAc-R (K_i = 1.6 μM, sialyl T-antigen), where the aglycone (R) was *N*-acetyl threonine methyl ester. The aglycone of the sialoglycans affected binding as well as the saccharide structure [55,67], raising the possibility that one might design enhanced glycan-mimetic inhibitors of MAG. In this light, although modifications of substituents on the terminal α2–3 linked sialic acid are generally not tolerated [63,64], synthetic sialic acid methyl glycosides with a hydrophobic moiety replacing the C9 hydroxyl of sialic acid had greatly enhanced MAG binding [60].

A summary of glycan binding specificity studies reveals that MAG binds preferentially to glycans with a terminal α2–3-linked NeuAc residue linked to Galβ1–3GalNAc, and that binding is enhanced by additional sialic acid residues on the GalNAc 6-hydroxyl (such as in GD1α) or on the internal galactose of gangliosides such as GD1a and GT1b. Given its high specificity for sialic acid binding

and its high evolutionary conservation, it is reasonable to hypothesize that MAG's sialic acid binding relates to its biological functions. The next section explores the evidence for this hypothesis.

14.4 MAG–GLYCAN BINDING IN MAG FUNCTION

14.4.1 MAG–Glycan Binding Supports Axon–Myelin Stability

The finding that MAG binds selectively to NeuAcα2–3Galβ1–3GalNAc termini provided the means to test whether that terminus on gangliosides, in particular, is related to MAG's functions. Gangliosides are biosynthesized stepwise by the sequential action of glycosyltransferases, many of which are glycolipid-specific (Figure 14.2) [68]. In particular, the *N*-acetylgalactosaminyltransferase coded by the *B4galnt1* gene (previously called *Galgt1* or *GalNAcT*) is the single gene responsible for adding the GalNAc residue to growing ganglio-series gangliosides. In mice engineered to carry a disrupted *B4galnt1* gene, the NeuAcα2–3Galβ1–3GalNAc ganglioside terminus is absent [69,70]. If MAG binding to that terminus on gangliosides is responsible for MAG's biological effects, the phenotype of *B4galnt1*-null mice would be expected to share characteristics with that of *Mag*-null mice. Studies of *B4galnt1*-null mice support this conclusion.

B4galnt1-null mice lack all complex gangliosides, including MAG-binding gangliosides GD1a, GT1b, and the α-series gangliosides. However, they express the same total amount of brain gangliosides (and ganglioside sialic acid) as wild-type mice, in the form of the upstream precursors GM3 and GD3, which are quantitatively minor gangliosides in wild-type adult mammalian brain [70,71]. MAG does not bind to GD3 and binds to GM3 with one-tenth the affinity as to GD1a or GT1b [64]. Despite the gross shift in their brain ganglioside patterns, *B4galnt1*-null mice appear normal at birth and into young adulthood [69]. Notably, at the age of 3 months, they begin to show signs of behavioral neuropathy, with altered gait and diminished reflexes [72]. When they reach 1 year of age, they have notably impaired hindlimb function, walking in short labored steps, and display other motor behavioral deficits. Ultrastructural studies revealed markedly increased axon degeneration, decreased neurofilament spacing, and an increase in unmyelinated and dysmyelinated axons in *B4galnt1*-null mice [73], all deficits that are similar to those that had been noted in *Mag*-null mice [2,7,18,19]. Also similar to *Mag*-null mice, *B4galnt1*-null mice display altered distribution of paranodal and juxtaparanodal markers at nodes of Ranvier [74]. Additional evidence for a relationship between MAG and brain gangliosides comes from the observation that MAG expression is progressively and selectively decreased in the brains of *B4galnt1*-null mice [71]. This occurs without a change in the expression level of the *Mag* gene, suggesting that MAG–ganglioside binding stabilizes MAG expression.

Subsequent experiments compared motor behavior and axon ultrastructure in *B4galnt1*-null, *Mag*-null, and double-null mice in the same experiments and on the same strain background [20]. The phenotypes of the three mutant strains differed from wild-type but were notably similar to one another. All three mutant genotypes had similar marked increases in PNS and CNS axon degeneration, similarly reduced axon diameters and neurofilament spacing, and similar behavioral deficits, including impaired motor coordination and reflexes. Unexpectedly, *Mag*- and *B4galnt1*-null mouse strains were both hyperactive despite their motor deficits. These data indicate the importance of the NeuAcα2–3Galβ1–3GalNAc terminus on gangliosides in the maintenance of long-term axon–myelin stability, and imply that the role of MAG in this stability is primarily via gangliosides and the role of gangliosides primarily via MAG. Thus, MAG is a nervous system lectin designed to bind to axonal gangliosides containing the NeuAcα2–3Galβ1–3GalNAc terminus, regulate axon cytoarchitecture, and mediate long-term axon–myelin stability.

14.4.2 MAG–Glycan Binding in the Regulation of Axon Regeneration

Controversy remains about whether sialoglycan binding is responsible for MAG's inhibition of axon regeneration [39,75]. Based on *in vitro* axon outgrowth studies, strong data both support and refute

the hypothesis that sialoglycan binding is required for MAG inhibition of axon outgrowth [39,61,76]. Nerve cells, typically isolated from newborn rats or mice, have been grown on culture surfaces adsorbed with MAG extracted from myelin [9,39], on living fibroblasts engineered to ectopically express MAG [8,51], or have been treated with engineered soluble chimeric forms of MAG [61]. In all cases, MAG inhibits either the extent or rate of axon outgrowth. In some studies, MAG inhibition is partially or completely reversed by disrupting MAG–sialoglycan interactions [39,61,62,76], whereas in other cases, such disruption is without effect [32,61,77].

At the heart of the controversy is the finding that MAG binds in a glycan-independent manner to Nogo receptor (NgR) [32,77], a glycosylphosphatidylinositol-anchored (GPI-linked) protein on axons and growth cones that also mediates the inhibitory effects of two other ARIs, Nogo and OMgp [78]. Axon outgrowth studies confirmed that MAG inhibits axon outgrowth via NgR binding and that NgR-mediated inhibition is not via MAG's sialic acid-binding site. These data appear to dismiss the lectin activity of MAG as key to its axon outgrowth inhibition, but are at odds with the effects of sialidase in other published studies [39,62,76,79].

A hypothesis consistent with all of the data is that MAG has two independent inhibition domains, one that binds to sialoglycans and another that binds to NgR. The two independent MAG inhibition domains, independently or in concert, may inhibit axon outgrowth from different neuronal cell types (depending on NgR or sialoglycan expression), may mediate different MAG functions (axon cytoarchitecture or axon outgrowth inhibition), or may mediate inhibition by different molecular forms of MAG (full-length membrane-bound MAG or proteolytically released soluble MAG). In this section, the evidence for MAG axon inhibition via its lectin domains is presented, followed by a model of MAG's dual inhibitory domains.

14.4.2.1 Sialidase and Sialoglycans Reverse MAG-Mediated Axon Outgrowth Inhibition *in Vitro*

Neurons, typically isolated from young rats or mice, are cultured *in vitro*, where under appropriate conditions, they extend axon-like projections, referred to as neurites. Early studies found that when DRGN from 1 day old mice were cultured on fibroblasts ectopically expressing MAG, neurite outgrowth was enhanced (rather than inhibited) in a MAG-specific manner (blocked by an anti-MAG monoclonal antibody) [51]. In contrast, subsequent studies demonstrated that neurite extension from several different neuronal cell types, including DRGN, was inhibited when they were plated on MAG-expressing fibroblasts. This apparent discrepancy was resolved when it was discovered that the response of DRGN to MAG is developmentally regulated, with enhancement shifting to inhibition at 3 days after birth in rodents. The effect of MAG on neurite outgrowth from other types of rat and mouse neurons appears to be inhibitory, even at day 1 postnatal. Together, the data emphasize that different neuronal cells respond differently to MAG and that MAG inhibits neurite outgrowth from mature neurons.

In early studies, MAG-mediated inhibition of neurite outgrowth from cerebellar granule neurons (CGN) isolated from newborn rats and grown on MAG-expressing fibroblasts was partially and significantly reversed by sialidase treatment or by addition of high concentrations of sialoglycans [76]. In subsequent studies, inhibition of axon outgrowth from CGN isolated from young rats (4–5 day old) when plated on substrates adsorbed with native MAG was completely reversed by sialidase [39]. Further evidence that sialic acid binding is important to MAG's axon outgrowth inhibition, at least for CGN, came from studies using monovalent sialoglycoside inhibitors. As described above, screening assays revealed that disialyl T-antigen (NeuAcα2–3Gal(NeuAcα2–6)GalNAc-R) and the trisaccharide NeuAcα2–3Galβ1–3GalNAc-R are potent blockers of MAG binding [67]. If sialoglycan binding by MAG is responsible for its inhibition of axon outgrowth, one would predict that adding these glycosides would reverse MAG-mediated inhibition. This was found to be the case using native MAG as an inhibitor of the extensive axon outgrowth from rat CGN (Figure 14.4). Sialoglycosides reversed MAG-mediated inhibition of axon outgrowth with efficacy proportional to their K_i values for MAG binding; sialoglycans with poor MAG binding were without effect [79].

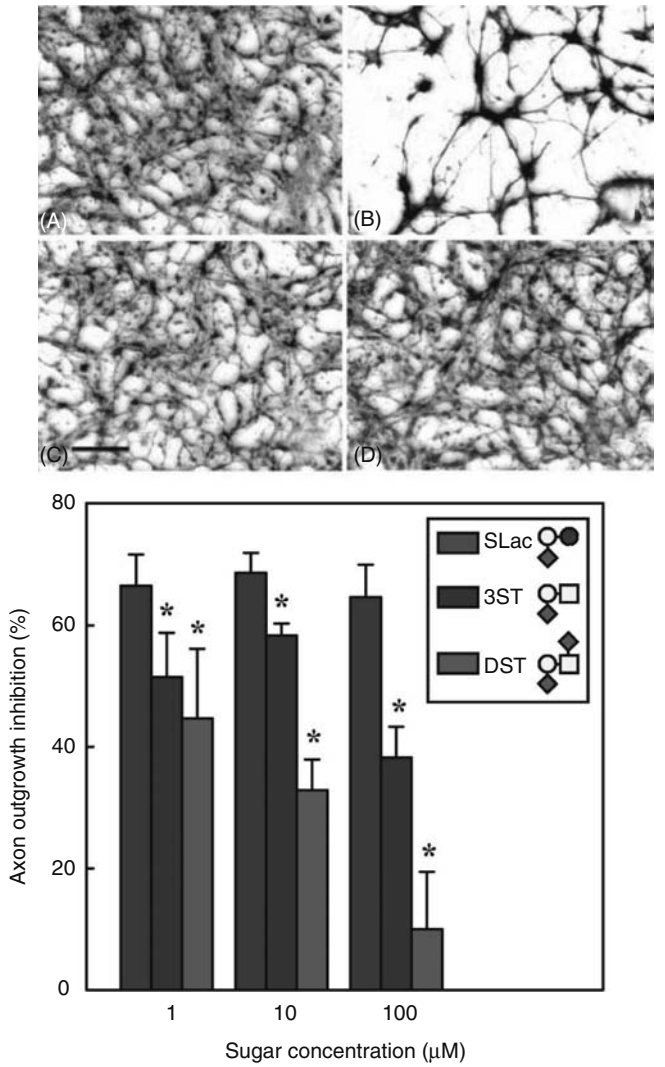


FIGURE 14.4 (See CD for color figure.) Reversal of MAG-mediated axon outgrowth inhibition by monovalent sialosides [79]. (Upper panels) Rat cerebellar granule neurons (CGN) were plated on control substrata (A, C) and on the same substrata adsorbed with myelin, containing MAG (B, D). After 48 h in control medium (A, B) or in the same medium containing 10 μM disialyl T glycoside (C, D), cells were immunostained to detect axons. Fluorescent micrographs are presented as reverse gray scale images. On the control substratum (A) cells extend a lacy network of axons. On the MAG-containing surfaces, axons clump together avoiding the substratum (B). Addition of 10 μM disialyl T glycoside reverses MAG-mediated inhibition (D). Bar = 100 μM. (Lower panel) Monovalent sialosides were added at the indicated concentrations to CGN cultured on control and MAG-containing substrata. After 48 h in culture, the cells were immunostained and axon outgrowth was quantified. Data are expressed as MAG-mediated axon outgrowth inhibition from six to seven independent cultures (mean ± S.E.). Statistical analyses were performed using Student's *t*-test: *, $p < 0.02$. Sialosides reversed MAG-mediated inhibition proportional to their MAG-binding affinity, in the order (best to worst) disialyl T glycoside (DST) > sialyl T glycoside (NeuAcα2–3Galβ1–4GalNAc-R, 3ST) > 3-sialyllactose (SLac), which failed to reverse inhibition. (Reproduced from Vyas, A.A., Blixt, O., Paulson, J.C., and Schnaar, R.L., *J. Biol. Chem.*, 280, 16305, 2005. With permission.)

14.4.2.2 Sialidase Enhances Spinal Axon Outgrowth *in Vivo*

Since sialidase reversed MAG-mediated inhibition, we asked whether sialidase delivery to the injured spinal cord in live animals can improve recovery. As an initial test, we used a brachial plexus injury model [80]. The brachial plexus is the network of nerves that conducts signals between the spinal cord and the shoulder, arm and hand. When the arm is violently pulled away from the body, the brachial plexus nerves may be torn from the spinal cord, leaving the arm without motor or sensory function. To mimic this injury (and a potential therapy) in the rat, the brachial plexus nerves were cut flush to the spinal cord, then a nerve graft was used to bridge the spinal cord to the severed shoulder nerve. In some animals, sialidase was delivered to the graft insertion site. After 1 month to allow repair, the number of spinal motoneurons that extended new axons into the graft was counted using a retrograde dye. Sialidase treatment resulted in a marked, 2.6-fold enhancement of spinal axon outgrowth. These data implicate sialoglycans in the control of axon outgrowth after injury and suggest a therapeutic approach to enhance recovery from traumatic spinal cord injury. Whether enhanced spinal axon outgrowth was due to reversal of MAG-ganglioside binding or another sialic acid-dependent mechanism has not yet been established.

14.4.2.3 Gangliosides are Sialoglycan Receptors for MAG-Mediated Inhibition of Axon Outgrowth

Gangliosides are the most abundant sialoglycans in the brain and the major brain gangliosides GD1a and GT1b carry most of the preferred NeuAcα2–3Galβ1–3GalNAc MAG binding determinant [56]. The role of gangliosides as MAG receptors in axon outgrowth inhibition was tested using MAG inhibition of axon outgrowth from CGN's isolated from young (4–5 day old) rats and mice [39]. Four lines of evidence, taken together, support the conclusion that gangliosides per se are functional receptors for MAG-mediated inhibition of axon outgrowth in this neuronal system: (i) sialidase reverses MAG inhibition; (ii) a metabolic inhibitor of glycolipid biosynthesis, P4 (1-phenyl-2-hexadecanoylamino-3-morpholino-1-propanol) reverses MAG inhibition; (iii) CGN from *B4galnt1*-null mice, which lack the NeuAcα2–3Galβ1–3GalNAc terminus specifically on gangliosides, are less responsive to MAG; and (iv) highly specific IgG-class antibodies to GD1a and GT1b selectively decrease MAG inhibition.

If MAG on myelin binds to gangliosides on the axon to initiate its inhibitory effects, how might this occur? One possibility is that MAG, which is concentrated on the periaxonal myelin membrane, clusters gangliosides on the apposing axon membrane, resulting in coclustering of signaling molecules. This hypothesis is supported by the finding that multivalent forms of anti-GD1a and anti-GT1b antibodies mimic MAG inhibition [39,62]. It is proposed that adapter molecules associate with gangliosides to generate a transmembrane inhibitory signal, although the identity of such adapters is yet to be established. One potential adapter protein is $p75^{ntr}$, the "low-affinity" neurotrophin receptor, which was reported to bind to ganglioside GT1b and is also reported to be part of the NgR receptor complex [38,78]. However, the role of p75 in axon outgrowth inhibition has been questioned based on genetic evidence [81], and signaling via NgR is sialidase-independent in most studies [32,77], making it unlikely that GT1b or any other sialoglycan is required. Another NgR family member, NgR2, binds preferentially to MAG and mediates inhibition, apparently in a sialic acid-dependent manner [82]. Although additional studies will be needed to establish the receptors and mechanisms of MAG's interactions with axons, a working model based on current knowledge follows.

14.5 MAG: A MULTIFUNCTIONAL SIGNALING MOLECULE WITH DUAL RECEPTORS

MAG stabilizes axon–myelin interactions, regulates axonal cytoarchitecture, nurtures axons, and restricts axon regeneration. Some of these functions are mediated by MAG's lectin activity, whereas

others appear to be sialoglycan-independent. For example, MAG is inferred to have a primary sialic acid-binding site in N-terminal Ig-like domain centered at arginine 118 [60]. Mutation of this arginine to an alanine or aspartic acid sharply diminished the ability of a soluble MAG chimera (MAG-Fc) to bind to neurons and reversed the ability of the chimera to inhibit axon outgrowth from cerebellar neurons, supporting a role of MAG's lectin-binding activity in axon outgrowth inhibition [61]. However, in the same study, when the same arginine was mutated on fibroblasts expressing cell surface MAG, inhibition of axon outgrowth from CGN was retained, presumably in the absence of sialic acid-dependent lectin activity. Further complicating the interpretation of these data, a subsequent study indicated that R118-mutated MAG-Fc partially inhibited axon outgrowth from hippocampal neurons, and that the residual activity was sialic acid-dependent, implicating a second sialic acid-binding site on MAG [62]. The second lectin site hypothesis is consistent with enhanced binding of disialyl T-antigen, where both sialic acids contribute to MAG-binding affinity [67]. Unfortunately, crystallization of MAG and resolution of its sialic acid binding sites has not yet been reported.

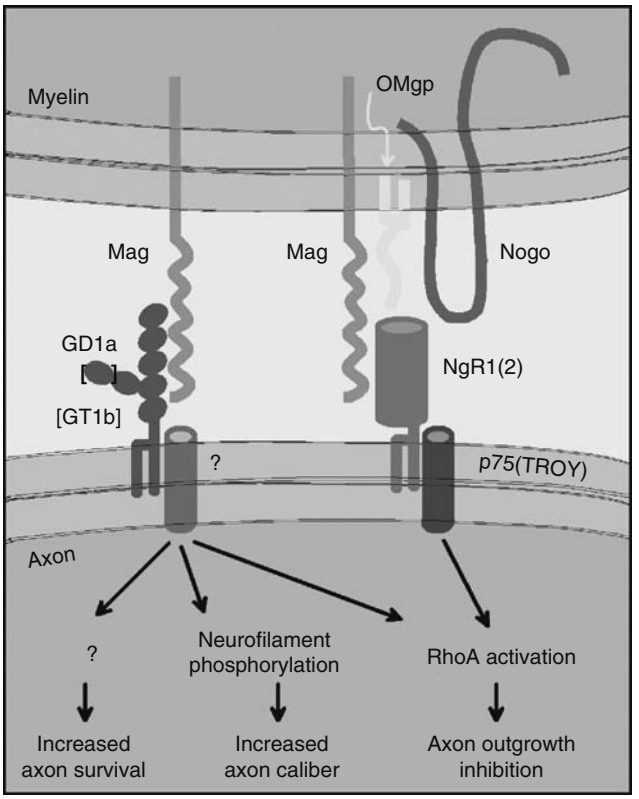


FIGURE 14.5 (See color insert following blank page 170. Also see CD for color figure.) Dual receptor model for MAG function. MAG, on the periaxonal myelin membrane, engages the sialoglycans GD1a and GT1b via its sialic acid binding sites. Via as yet unidentified transmembrane adapter molecules, signals are conveyed to the neuron/axon that enhance axon survival, modulate axon caliber via neurofilament phosphorylation, and inhibit axon outgrowth via activation of the small nonreceptor GTPase, RhoA. Independently, depending on the neuronal cell type, MAG and other axon regeneration inhibitors OMgp and Nogo engage the Nogo receptor (NgR) to activate RhoA and inhibit axon regeneration. Evidence suggests that MAG binding to gangliosides and NgR is independent and via different polypeptide domains. Knowledge of the dual pathways may provide insight into dysmyelination disease outcomes and provide novel targets to enhance axon regeneration after injury. (Modified from Mehta, N.R., Lopez, P.H., Vyas, A.A., and Schnaar, R.L., *J. Biol. Chem*., 282, 27875, 2007. With permission.)

As noted above, MAG binds with high affinity to the Nogo receptor, NgR [32,77]. As a GPI-linked protein, NgR and related family members are susceptible to release from the cell surface by phosphatidylinositol-specific phospholipase C (PIPLC) [83]. Studies using cerebellar and DRGNs demonstrated that MAG inhibition of axon outgrowth was reversed by PIPLC or by other methods that disrupt NgR, and that MAG-NgR interactions were sialic acid-independent [32,77]. Although subsequent studies suggested that NgR and its homologs (NgR2, NgR3) may bind to MAG in a sialic acid-dependent manner [82], the majority of studies suggest that functional MAG–NgR binding is independent of its lectin-binding activity. Recent domain swap studies between sialoadhesin (Siglec-1) and MAG (Siglec-4) indicate that MAG's axon inhibitory activity, at least in some experimental systems, is dependent on MAG's fifth Ig-like domain, which to date has not been reported to contain a sialic acid-binding activity, whereas inhibition was independent of MAG's first Ig-like domain, which is the primary sialic acid-binding site [75].

Although the model of NgR-mediated signaling by different axon regeneration inhibitors (MAG, Nogo, OMgp) has strong experimental support, and is attractive as a unifying hypothesis, recent evidence indicates that it falls short of fitting all of the data, especially for MAG: (i) Mice engineered to lack NgR have limited or no improvement in axon outgrowth after CNS injury [84,85]; (ii) NgR mRNA is not readily detectable in many neurons that respond to axon regeneration inhibitors *in vivo*. This includes many neurons in the adult human spinal cord, adult DRG, and the neostriatum, whether intact or lesioned [28,86,87]; and (iii) Removal of NgR with PIPLC reverses MAG-mediated inhibition in some studies [32,77], but clearly not in others [88], even under conditions that totally reverse Nogo inhibition. It is also notable that treatment with sialidase enhanced spinal axon outgrowth *in vivo*, whereas PIPLC was without therapeutic benefit [80].

14.6 CONCLUSION AND FUTURE DIRECTIONS

Together, current findings are consistent with a model in which MAG has separate receptor-binding domains, a lectin-binding domain specific for sialoglycans and a distinct NgR-binding domain (Figure 14.5). These domains may act independently or in concert to engage the axon and initiate MAG's multiple functions. Genetic data indicate that sialoglycans, and particularly gangliosides, are functional receptors for axon–myelin stabilization by MAG, and at least in part for MAG's regulation of the axon cytoarchitecture and the distribution of ion channels and adhesion molecules at nodes of Ranvier [20,74]. These data imply that MAG's lectin activity is key to those functions. In contrast, MAG inhibition of axon outgrowth appears to involve both the lectin (sialoglycan) and nonlectin (NgR) binding domains, depending on the neuronal cell type and the nature of the presentation between MAG and the axon/growth cone/neuron cell surface [32,39,77]. This knowledge can be used to devise new approaches to reverse the inhibition of axon regeneration caused by MAG and other ARIs with the aim of providing novel therapies to enhance recovery from CNS injury and disease.

REFERENCES

1. Trapp, B.D., Myelin-associated glycoprotein. Location and potential functions, *Ann. N.Y. Acad. Sci*., 605, 29, 1990.
2. Schachner, M. and Bartsch, U., Multiple functions of the myelin-associated glycoprotein MAG (siglec-4a) in formation and maintenance of myelin, *Glia*, 29, 154, 2000.
3. Fruttiger, M., Montag, D., Schachner, M., and Martini, R., Crucial role for the myelin-associated glycoprotein in the maintenance of axon-myelin integrity, *Eur. J. Neurosci*., 7, 511, 1995.
4. Fujita, N., Kemper, A., Dupree, J., Nakayasu, H., Bartsch, U., Schachner, M., Maeda, N., Suzuki, K., and Popko, B., The cytoplasmic domain of the large myelin-associated glycoprotein isoform is needed for proper CNS but not peripheral nervous system myelination., *J. Neurosci*., 18, 1970, 1998.
5. Montag, D., Giese, K.P., Bartsch, U., Martini, R., Lang, Y., Bluthmann, H., Karthingasan, J., Kirschner, D.A., Wintergerst, E.S., Nave, K.-A., Zielasek, J., Toyka, K.V., Lipp, H.-P., and Schachner, M., Mice deficient for the myelin-associated glycoprotein show subtle abnormalities in myelin, *Neuron*, 13, 229, 1994.

6. Li, C., Tropak, M.B., Gerial, R., Clapoff, S., Abramow-Newerly, W., Trapp, B., Peterson, A., and Roder, J., Myelination in the absence of myelin-associated glycoprotein, *Nature*, 369, 747, 1994.
7. Yin, X., Crawford, T.O., Griffin, J.W., Tu, P., Lee, V.M., Li, C., Roder, J., and Trapp, B.D., Myelin-associated glycoprotein is a myelin signal that modulates the caliber of myelinated axons, *J. Neurosci*., 18, 1953, 1998.
8. Mukhopadhyay, G., Doherty, P., Walsh, F.S., Crocker, P.R., and Filbin, M.T., A novel role for myelin-associated glycoprotein as an inhibitor of axonal regeneration, *Neuron*, 13, 757, 1994.
9. McKerracher, L., David, S., Jackson, D.L., Kottis, V., Dunn, R.J., and Braun, P.E., Identification of myelin-associated glycoprotein as a major myelin-derived inhibitor of neurite growth, *Neuron*, 13, 805, 1994.
10. Quarles, R.H., Myelin-associated glycoprotein (MAG): past, present and beyond, *J. Neurochem*., 100, 1431, 2007.
11. Bartsch, U., Kirchhoff, F., and Schachner, M., Immunohistological localization of the adhesion molecules L1, N-CAM, and MAG in the developing and adult optic nerve of mice, *J. Comp. Neurol*., 284, 451, 1989.
12. Trapp, B.D., Andrews, S.B., Cootauco, C., and Quarles, R., The myelin-associated glycoprotein is enriched in multivesicular bodies and periaxonal membranes of actively myelinating oligodendrocytes, *J. Cell Biol*., 109, 2417, 1989.
13. Arquint, M., Roder, J., Chia, L.-S., Down, J., Wilkinson, D., Bayley, H., Braun, P., and Dunn, R., Molecular cloning and primary structure of myelin-associated glycoprotein, *Proc. Natl Acad. Sci. USA*, 84, 600, 1987.
14. Lai, C., Brow, M.A., Nave, K.A., Noronha, A.B., Quarles, R.H., Bloom, F.E., Milner, R.J., and Sutcliffe, J.G., Two forms of 1B236/myelin-associated glycoprotein, a cell adhesion molecule for postnatal neural development, are produced by alternative splicing, *Proc. Natl Acad. Sci. USA*, 84, 4337, 1987.
15. Tropak, M.B., Johnson, P.W., Dunn, R.J., and Roder, J.C., Differential splicing of MAG transcripts during CNS and PNS development, *Brain Res*., 464, 143, 1988.
16. Sato, S., Quarles, R.H., and Brady, R.O., Susceptibility of the myelin-associated glycoprotein and basic protein to a neutral protease in highly purified myelin from human and rat brain, *J. Neurochem*., 39, 97, 1982.
17. Bartsch, S., Montag, D., Schachner, M., and Bartsch, U., Increased number of unmyelinated axons in optic nerves of adult mice deficient in the myelin-associated glycoprotein (MAG), *Brain Res*., 762, 231, 1997.
18. Li, C., Trapp, B., Ludwin, S., Peterson, A., and Roder, J., Myelin associated glycoprotein modulates glia–axon contact in vivo, *J. Neurosci. Res*., 51, 210, 1998.
19. Bjartmar, C., Yin, X., and Trapp, B.D., Axonal pathology in myelin disorders, *J. Neurocytol*., 28, 383, 1999.
20. Pan, B., Fromholt, S.E., Hess, E.J., Crawford, T.O., Griffin, J.W., Sheikh, K.A., and Schnaar, R.L., Myelin-associated glycoprotein and complementary axonal ligands, gangliosides, mediate axon stability in the CNS and PNS: Neuropathology and behavioral deficits in single- and double-null mice, *Exp. Neurol*., 195, 208, 2005.
21. Dashiell, S.M., Tanner, S.L., Pant, H.C., and Quarles, R.H., Myelin-associated glycoprotein modulates expression and phosphorylation of neuronal cytoskeletal elements and their associated kinases, *J. Neurochem*., 81, 1263, 2002.
22. Kumar, S., Yin, X., Trapp, B.D., Paulaitis, M.E., and Hoh, J.H., Role of long-range repulsive forces in organizing axonal neurofilament distributions: evidence from mice deficient in myelin-associated glycoprotein, *J. Neurosci. Res*., 68, 681, 2002.
23. Lunn, M.P., Crawford, T.O., Hughes, R.A., Griffin, J.W., and Sheikh, K.A., Anti-myelin-associated glycoprotein antibodies alter neurofilament spacing, *Brain*, 125, 904, 2002.
24. Poliak, S. and Peles, E., The local differentiation of myelinated axons at nodes of Ranvier, *Nat. Rev. Neurosci*., 4, 968, 2003.
25. Marcus, J., Dupree, J.L., and Popko, B., Myelin-associated glycoprotein and myelin galactolipids stabilize developing axo-glial interactions, *J. Cell Biol*., 156, 567, 2002.
26. Selles-Navarro, I., Ellezam, B., Fajardo, R., Latour, M., and McKerracher, L., Retinal ganglion cell and nonneuronal cell responses to a microcrush lesion of adult rat optic nerve, *Exp. Neurol*., 167, 282, 2001.
27. David, S. and Aguayo, A.J., Axonal elongation into peripheral nervous system "bridges" after central nervous system injury in adult rats, *Science*, 214, 931, 1981.
28. Sandvig, A., Berry, M., Barrett, L.B., Butt, A., and Logan, A., Myelin-, reactive glia-, and scar-derived CNS axon growth inhibitors: Expression, receptor signaling, and correlation with axon regeneration, *Glia*, 46, 225, 2004.
29. Caroni, P. and Schwab, M.E., Two membrane protein fractions from rat central myelin with inhibitory properties for neurite growth and fibroblast spreading, *J. Cell Biol*., 106, 1281, 1988.

30. Wang, K.C., Koprivica, V., Kim, J.A., Sivasankaran, R., Guo, Y., Neve, R.L., and He, Z., Oligodendrocyte-myelin glycoprotein is a Nogo receptor ligand that inhibits neurite outgrowth, *Nature*, 417, 941, 2002.
31. Snow, D.M., Lemmon, V., Carrino, D.A., Caplan, A.I., and Silver, J., Sulfated proteoglycans in astroglial barriers inhibit neurite outgrowth in vitro, *Exp. Neurol*., 109, 111, 1990.
32. Domeniconi, M., Cao, Z., Spencer, T., Sivasankaran, R., Wang, K., Nikulina, E., Kimura, N., Cai, H., Deng, K., Gao, Y., He, Z., and Filbin, M., Myelin-associated glycoprotein interacts with the Nogo66 receptor to inhibit neurite outgrowth, *Neuron*, 35, 283, 2002.
33. Hasegawa, Y., Fujitani, M., Hata, K., Tohyama, M., Yamagishi, S., and Yamashita, T., Promotion of axon regeneration by myelin-associated glycoprotein and Nogo through divergent signals downstream of Gi/G, *J. Neurosci*., 24, 6826, 2004.
34. Shen, Y.J., DeBellard, M.E., Salzer, J.L., Roder, J., and Filbin, M.T., Myelin-associated glycoprotein in myelin and expressed by Schwann cells inhibits axonal regeneration and branching, *Mol. Cell Neurosci*., 12, 79, 1998.
35. Song, H., Ming, G., He, Z., Lehmann, M., Tessier-Lavigne, M., and Poo, M., Conversion of neuronal growth cone responses from repulsion to attraction by cyclic nucleotides, *Science*, 281, 1515, 1998.
36. Wang, K.C., Kim, J.A., Sivasankaran, R., Segal, R., and He, Z., P75 interacts with the Nogo receptor as a co-receptor for Nogo, MAG and OMgp, *Nature*, 420, 74, 2002.
37. Wong, E.V., David, S., Jacob, M.H., and Jay, D.G., Inactivation of myelin-associated glycoprotein enhances optic nerve regeneration, *J. Neurosci*., 23, 3112, 2003.
38. Yamashita, T., Higuchi, H., and Tohyama, M., The p75 receptor transduces the signal from myelin-associated glycoprotein to Rho, *J. Cell Biol*., 157, 565, 2002.
39. Vyas, A.A., Patel, H.V., Fromholt, S.E., Heffer-Lauc, M., Vyas, K.A., Dang, J., Schachner, M., and Schnaar, R.L., Gangliosides are functional nerve cell ligands for myelin-associated glycoprotein (MAG), an inhibitor of nerve regeneration, *Proc. Natl Acad. Sci. USA*, 99, 8412, 2002.
40. Li, M., Shibata, A., Li, C., Braun, P.E., McKerracher, L., Roder, J., Kater, S.B., and David, S., Myelin-associated glycoprotein inhibits neurite/axon growth and causes growth cone collapse, *J. Neurosci. Res*., 46, 404, 1996.
41. Schäfer, M., Fruttiger, M., Montag, D., Schachner, M., and Martini, R., Disruption of the gene for the myelin-associated glycoprotein improves axonal regrowth along myelin in C57BL/Wld^s mice, *Neuron*, 16, 1107, 1996.
42. Kim, J.E., Li, S., GrandPre, T., Qiu, D., and Strittmatter, S.M., Axon regeneration in young adult mice lacking Nogo-A/B, *Neuron*, 38, 187, 2003.
43. Simonen, M., Pedersen, V., Weinmann, O., Schnell, L., Buss, A., Ledermann, B., Christ, F., Sansig, G., van der, P.H., and Schwab, M.E., Systemic deletion of the myelin-associated outgrowth inhibitor Nogo-A improves regenerative and plastic responses after spinal cord injury, *Neuron*, 38, 201, 2003.
44. Zheng, B., Ho, C., Li, S., Keirstead, H., Steward, O., and Tessier-Lavigne, M., Lack of enhanced spinal regeneration in Nogo-deficient mice, *Neuron*, 38, 213, 2003.
45. Bradbury, E.J., Moon, L.D., Popat, R.J., King, V.R., Bennett, G.S., Patel, P.N., Fawcett, J.W., and McMahon, S.B., Chondroitinase ABC promotes functional recovery after spinal cord injury, *Nature*, 416, 636, 2002.
46. He, Z. and Koprivica, V., The Nogo signaling pathway for regeneration block, *Annu. Rev. Neurosci*., 27, 341, 2004.
47. Schwab, M.E., Nogo and axon regeneration, *Curr. Opin. Neurobiol*., 14, 118, 2004.
48. Tang, S., Woodhall, R.W., Shen, Y.J., DeBellard, M.E., Saffell, J.L., Doherty, P., Walsh, F.S., and Filbin, M.T., Soluble myelin-associated glycoprotein (MAG) found in vivo inhibits axonal regeneration, *Mol. Cell. Neurosci*., 9, 333, 1997.
49. Tang, S., Qiu, J., Nikulina, E., and Filbin, M.T., Soluble myelin-associated glycoprotein released from damaged white matter inhibits axonal regeneration, *Mol. Cell Neurosci*., 18, 259, 2001.
50. Poltorak, M., Sadoul, R., Keilhauer, G., Landa, C., Fahrig, T., and Schachner, M., Myelin-associated glycoprotein, a member of the L2/HNK-1 family of neural cell adhesion molecules, is involved in neuron–oligodendrocyte and oligodendrocyte–oligodendrocyte interaction, *J. Cell Biol*., 105, 1893, 1987.
51. Johnson, P.W., Abramow-Newerly, W., Seilheimer, B., Sadoul, R., Tropak, M.B., Arquint, M., Dunn, R.J., Schachner, M., and Roder, J.C., Recombinant myelin-associated glycoprotein confers neural adhesion and neurite outgrowth function, *Neuron*, 3, 377, 1989.
52. Crocker, P.R., Siglecs: Sialic-acid-binding immunoglobulin-like lectins in cell–cell interactions and signalling, *Curr. Opin. Struct. Biol*., 12, 609, 2002.
53. Varki, A. and Angata, T., Siglecs—the major subfamily of I-type lectins, *Glycobiology*, 16, 1R, 2006.

54. Kelm, S., Pelz, A., Schauer, R., Filbin, M.T., Song, T., de Bellard, M.E., Schnaar, R.L., Mahoney, J.A., Hartnell, A., Bradfield, P., and Crocker, P.R., Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of sialic acid-dependent adhesion molecules of the immunoglobulin superfamily, *Curr. Biol*., 4, 965, 1994.
55. Kelm, S., Brossmer, R., Isecke, R., Gross, H.-J., Strenge, K., and Schauer, R., Functional groups of sialic acids involved in binding to siglecs (sialoadhesins) deduced from interactions with synthetic analogues., *Eur. J. Biochem*., 255, 663, 1998.
56. Schnaar, R.L., Glycolipid-mediated cell-cell recognition in inflammation and nerve regeneration, *Arch. Biochem. Biophys*., 426, 163, 2004.
57. Schnaar, R.L., Glycobiology of the nervous system, in: *Carbohydrates in Chemistry and Biology, Part II: Biology of Saccharides*, Ernst, B., Hart, G.W., and Sinaÿ, P., Eds., Wiley-VCH, Weinheim, Germany, 2000, Chapter 62, p. 1013.
58. Collins, B.E., Ito, H., Sawada, N., Ishida, H., Kiso, M., and Schnaar, R.L., Enhanced binding of the neural siglecs, myelin-associated glycoprotein and Schwann cell myelin protein, to Chol-1 (α-series) gangliosides and novel sulfated Chol-1 analogs, *J. Biol. Chem*., 274, 37637, 1999.
59. May, A.P., Robinson, R.C., Vinson, M., Crocker, P.R., and Jones, E.Y., Crystal structure of the N-terminal domain of sialoadhesin in complex with 3′ sialyllactose at 1.85 A resolution, *Mol. Cell*, 1, 719, 1998.
60. Zaccai, N.R., Maenaka, K., Maenaka, T., Crocker, P.R., Brossmer, R., Kelm, S., and Jones, E.Y., Structure-guided design of sialic Acid-based siglec inhibitors and crystallographic analysis in complex with sialoadhesin, *Structure (Camb.)*, 11, 557, 2003.
61. Tang, S., Shen, Y.J., de Bellard, M.E., Mukhopadhyay, G., Salzer, J.L., Crocker, P.R., and Filbin, M.T., Myelin-associated glycoprotein interacts with neurons via a sialic acid binding site at ARG118 and a distinct neurite inhibition site., *J. Cell Biol*., 138, 1355, 1997.
62. Vinson, M., Strijbos, P.J., Rowles, A., Facci, L., Moore, S.E., Simmons, D.L., and Walsh, F.S., Myelin-associated glycoprotein interacts with ganglioside GT1b: A mechanism for neurite outgrowth inhibition, *J. Biol. Chem*., 276, 20280, 2001.
63. Collins, B.E., Yang, L.J.S., Mukhopadhyay, G., Filbin, M.T., Kiso, M., Hasegawa, A., and Schnaar, R.L., Sialic acid specificity of myelin-associated glycoprotein binding, *J. Biol. Chem*., 272, 1248, 1997.
64. Collins, B.E., Kiso, M., Hasegawa, A., Tropak, M.B., Roder, J.C., Crocker, P.R., and Schnaar, R.L., Binding specificities of the sialoadhesin family of I-type lectins. Sialic acid linkage and substructure requirements for binding of myelin-associated glycoprotein, Schwann cell myelin protein, and sialoadhesin, *J. Biol. Chem*., 272, 16889, 1997.
65. Collins, B.E., Fralich, T.J., Itonori, S., Ichikawa, Y., and Schnaar, R.L., Conversion of cellular sialic acid expression from *N*-acetyl- to *N*-glycolylneuraminic acid using a synthetic precursor, *N*-glycolylmannosamine pentaacetate: Inhibition of myelin-associated glycoprotein binding to neural cells, *Glycobiology*, 10, 11, 2000.
66. Neubacher, B., Scheid, S., Kelm, S., Frasch, A.C., Meyer, B., and Thiem, J., Synthesis of Neu5Ac oligosaccharides and analogues by transglycosylation and their binding properties as ligands to MAG, *Chembiochem*., 7, 896, 2006.
67. Blixt, O., Collins, B.E., van, d.N.I., Crocker, P.R., and Paulson, J.C., Sialoside specificity of the siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin-associated glycoprotein, *J. Biol. Chem*., 278, 31007, 2003.
68. Kolter, T., Proia, R.L., and Sandhoff, K., Combinatorial ganglioside biosynthesis, *J. Biol. Chem*., 277, 25859, 2002.
69. Takamiya, K., Yamamoto, A., Furukawa, K., Yamashiro, S., Shin, M., Okada, M., Fukumoto, S., Haraguchi, M., Takeda, N., Fujimura, K., Sakae, M., Kishikawa, M., Shiku, H., and Aizawa, S., Mice with disrupted GM2/GD2 synthase gene lack complex gangliosides but exhibit only subtle defects in their nervous system, *Proc. Natl Acad. Sci. USA*, 93, 10662, 1996.
70. Liu, Y., Wada, R., Kawai, H., Sango, K., Deng, C., Tai, T., McDonald, M.P., Araujo, K., Crawley, J.N., Bierfreund, U., Sandhoff, K., Suzuki, K., and Proia, R.L., A genetic model of substrate deprivation therapy for a glycosphingolipid storage disorder, *J. Clin. Invest*., 103, 497, 1999.
71. Sun, J., Shaper, N.L., Itonori, S., Heffer-Lauc, M., Sheikh, K.A., and Schnaar, R.L., Myelin-associated glycoprotein (Siglec-4) expression is progressively and selectively decreased in the brains of mice lacking complex gangliosides, *Glycobiology*, 14, 851, 2004.
72. Chiavegatto, S., Sun, J., Nelson, R.J., and Schnaar, R.L., A functional role for complex gangliosides: Motor deficits in GM2/GD2 synthase knockout mice, *Exp. Neurol*., 166, 227, 2000.

73. Sheikh, K.A., Sun, J., Liu, Y., Kawai, H., Crawford, T.O., Proia, R.L., Griffin, J.W., and Schnaar, R.L., Mice lacking complex gangliosides develop Wallerian degeneration and myelination defects, *Proc. Natl Acad. Sci. USA*, 96, 7532, 1999.
74. Susuki, K., Baba, H., Tohyama, K., Kanai, K., Kuwabara, S., Hirata, K., Furukawa, K., Furukawa, K., Rasband, M.N., and Yuki, N., Gangliosides contribute to stability of paranodal junctions and ion channel clusters in myelinated nerve fibers, *J. Neurochem*., 99(Suppl. 1), 29, 2006.
75. Filbin, M.T., What is the function of the sialic acid binding properties of myelin-associated glycoprotein (MAG)? *J. Neurochem*., 99(Suppl. 1), 10, 2006.
76. DeBellard, M.-E., Tang, S., Mukhopadhyay, G., Shen, Y.-J., and Filbin, M.T., Myelin-associated glycoprotein inhibits axonal regeneration from a variety of neurons via interaction with a sialoglycoprotein, *Mol. Cell. Neurosci*., 7, 89, 1996.
77. Liu, B.P., Fournier, A., GrandPre, T., and Strittmatter, S.M., Myelin-Associated Glycoprotein as a Functional Ligand for the Nogo-66 Receptor, *Science*, 297, 1190, 2002.
78. Yiu, G. and He, Z., Glial inhibition of CNS axon regeneration, *Nat. Rev. Neurosci*., 7, 617, 2006.
79. Vyas, A.A., Blixt, O., Paulson, J.C., and Schnaar, R.L., Potent glycan inhibitors of myelin-associated glycoprotein enhance axon outgrowth in vitro, *J. Biol. Chem*., 280, 16305, 2005.
80. Yang, L.J., Lorenzini, I., Vajn, K., Mountney, A., Schramm, L.P., and Schnaar, R.L., Sialidase enhances spinal axon outgrowth in vivo, *Proc. Natl Acad. Sci. USA*, 103, 11057, 2006.
81. Song, X.Y., Zhong, J.H., Wang, X., and Zhou, X.F., Suppression of p75NTR does not promote regeneration of injured spinal cord in mice, *J. Neurosci*., 24, 542, 2004.
82. Venkatesh, K., Chivatakarn, O., Lee, H., Joshi, P.S., Kantor, D.B., Newman, B.A., Mage, R., Rader, C., and Giger, R.J., The Nogo-66 receptor homolog NgR2 is a sialic acid-dependent receptor selective for myelin-associated glycoprotein, *J. Neurosci*., 25, 808, 2005.
83. Fournier, A.E., GrandPre, T., and Strittmatter, S.M., Identification of a receptor mediating Nogo-66 inhibition of axonal regeneration, *Nature*, 409, 341, 2001.
84. Zheng, B., Atwal, J., Ho, C., Case, L., He, X.L., Garcia, K.C., Steward, O., and Tessier-Lavigne, M., Genetic deletion of the Nogo receptor does not reduce neurite inhibition in vitro or promote corticospinal tract regeneration in vivo, *Proc. Natl Acad. Sci. USA*, 102, 1205, 2005.
85. Kim, J.E., Liu, B.P., Park, J.H., and Strittmatter, S.M., Nogo-66 receptor prevents raphespinal and rubrospinal axon regeneration and limits functional recovery from spinal cord injury, *Neuron*, 44, 439, 2004.
86. Hunt, D., Coffin, R.S., and Anderson, P.N., The Nogo receptor, its ligands and axonal regeneration in the spinal cord; a review, *J. Neurocytol*., 31, 93, 2002.
87. Josephson, A., Trifunovski, A., Widmer, H.R., Widenfalk, J., Olson, L., and Spenger, C., Nogo-receptor gene activity: cellular localization and developmental regulation of mRNA in mice and humans, *J. Comp. Neurol*., 453, 292, 2002.
88. Niederost, B., Oertle, T., Fritsche, J., McKinney, R.A., and Bandtlow, C.E., Nogo-A and myelin-associated glycoprotein mediate neurite growth inhibition by antagonistic regulation of RhoA and Rac1, *J. Neurosci*., 22, 10368, 2002.
89. Yang, L.J.S., Zeller, C.B., Shaper, N.L., Kiso, M., Hasegawa, A., Shapiro, R.E., and Schnaar, R.L., Gangliosides are neuronal ligands for myelin-associated glycoprotein, *Proc. Natl Acad. Sci. USA*, 93, 814, 1996.
90. Mehta, N.R., Lopez, P.H., Vyas, A.A., and Schnaar, R.L., Gangliosides and Nogo receptors independently mediate myelin-associated glycoprotein inhibition of neurite outgrowth in different nerve cells, *J. Biol. Chem*., 282, 27875, 2007.

15 Hyaluronan-Binding Proteoglycans

Edward N. Harris and Paul H. Weigel

CONTENTS

15.1 INTRODUCTION

Molecular components of the extracellular matrix (ECM) are critical for matrix stabilization, cellular differentiation, and tissue morphogenesis. Once thought to function only as the extracellular glue that primarily binds cells together, proteoglycans in the ECM carry out a diverse array of activities and can be grouped into five functionally based families: (i) hyalectins (lecticans), which contain lectin-like carbohydrate-recognition domains (CRDs) and bind hyaluronan (HA); (ii) heparan sulfate (HS) proteoglycans that sequester growth factors via their HS chains; (iii) small leucine-rich proteoglycans (SLRPs) containing leucine-rich domains that interact with collagens and other ECM proteins; (iv) phosphacans that function primarily as receptor-like protein–tyrosine phosphatases; and (v) part-time proteoglycans that reside both on the cell surface and in the ECM. More than 40 cDNAs encoding proteoglycan core proteins have been discovered thus far [1]. For the sake of brevity, this chapter will focus only on the four lecticans that bind HA (Figure 15.1): Aggrecan, neurocan, brevican, and versican.

All lecticans have two N-terminal Link domains, one immunoglobulin-like module, a central glycosaminoglycan (GAG)-attachment region, and one or more EGF-like domains in the C-terminus. The lecticans are post-translationally modified by O-Linked GAG chains on Ser residues of multiple Ser-Gly dipeptide repeats. Aggrecan, versican, neurocan, and brevican contain 120, 20, 7, and 3 potential sites, respectively. Generally, the CS chains of the lecticans maintain structural integrity of the tissue ECM by serving as barrier functions to limit the growth and migration of cells. For example, the CS chains of neurocan and brevican act as barriers to inhibit neurite outgrowth and migration so that, hypothetically, dendrites are available for longer periods of time in their local environment to make the correct synaptic junctions [2] and set up guidance cues for migration [3,4]. The barrier functions of the lecticans exist on two levels: (I) as mechanical barriers, the proteoglycans form neuronal nets to physically constrain neurite outgrowth [5] and (II) as cell signaling barriers, the proteoglycans can activate intracellular signaling pathways regulated by the Rho-family GTPases to promote the disruption of the actin cytoskeleton and growth cone

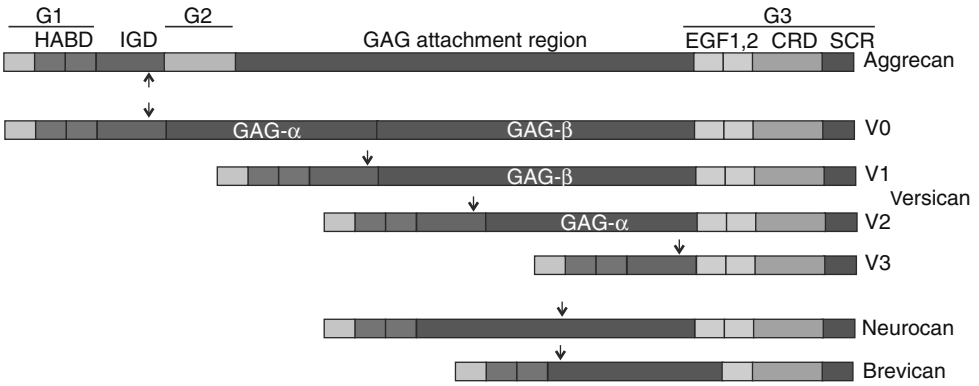


FIGURE 15.1 (See color insert following blank page 170. Also see CD for color figure.) Lectican domain organization. The common domains of the lecticans include two globular domains, G1 and G3, and a CS/KS-attachment region (red). The G1 region contains two Link modules (blue) that bind to HA. The G3 domain contains at least two EGF-like regions (yellow), a carbohydrate recognition domain (CRD; green), and a short complement regulatory region (SRC; black). Glycanated aggrecan is the largest of the lecticans, containing up to 120 CS and KS chains, in addition to an Ig-fold sequence (IGD; brown) and a G2 region (orange). The four splice variants of versican (V0–V3) contain N- and C-terminal regions that are very homologous with aggrecan and differing CS-attachment regions (except for V3) encoded by different exons, which are preferentially spliced depending on the host cell-type. Versican V0 contains the longest core protein of all the lecticans and contains up to 20 CS chains. Neurocan, with seven CS chains, and brevican, with three CS chains, differ in size and glycanation and do not contain an Ig-domain (IGD). Metalloproteolyic cleavage sites for ADAMTS-1/4 or MMP-2 (in neurocan only) are represented by the arrows.

collapse [6]. In addition to the barrier functions, GAGs in close proximity may also serve to sequester water to maintain tissue hydration.

Neurocan and brevican are primarily neural ECM lecticans, whereas aggrecan is found mainly in cartilage and versican is in vascular vessel walls. Aggrecan and versican play minor, albeit important and sometimes transient, functions in central nervous system (CNS) development and ECM composition [7]. Aggrecan, versican, and brevican are proteolytically degraded by ADAMTS-1/4 (a disintegrin and metalloproteinase with thrombospondin motifs) metalloproteases, whereas neurocan is processed by matrix metalloproteinases 2 (MMP-2). Metalloprotease cleavage of the lecticans is part of their normal turnover process and an increase in metalloprotease activity in tissues such as the brain may be a prerequisite for diseases such as Alzheimer's disease and epilepsy [8]. Numerous reviews have described the biology, chemistry, and physiological impact of these lecticans, their splice variants, and their cohorts of interacting partners on matrix and cellular composition in tissues [9–19].

15.2 AGGRECAN

Chondrocytes synthesize and secrete the ECM components of cartilage. The most abundant protein is collagen II, which is up to 60% of the dry weight of the cartilage, followed by aggrecan (35% of dry weight). All other components make up the remaining 5% of the tissue dry weight [20]. The aggrecan core protein is heavily modified in the endoplasmic reticulum (ER)/Golgi by addition of chondroitin sulfate (CS) and keratan sulfate (KS) chains prior to secretion from the cell. These glycosaminoglycan (GAG) modifications greatly increase the molecular mass of the molecule by at least—four- to fivefold over the core protein alone. After aggrecan is secreted from chondrocytes, it interacts with HA and Link proteins to form organized structures—proteoglycan aggregates. The CS

and KS chains are highly sulfated, giving the entire aggrecan molecule a high polyanionic charge, which attracts and binds large amounts of water and cations. Seventy percent of the wet weight of cartilage is due to water sequestered by the high-density negative charge associated with the sulfate and carboxylate groups on the GAGs, giving the tissue a gel-like quality [21]. The viscoelastic behavior of the GAGs in aggrecan–HA aggregates is directly responsible for the unique physical properties of articular cartilage, which includes compressive resilience during joint loading and load distribution over the surface of the tissue.

The aggrecan core protein consists of three disulfide-bonded globular domains, G1, G2, and G3, and a GAG-attachment region, primarily containing CS (Figure 15.1). The N-terminus of the G1 domain contains the HA-binding regions, which consist of two Link domains of ~93 amino acids with homology to the Link domains found in other hyalectins such as the hyaluronic acid receptor for endocytosis (HARE), TSG-6, CD44, and lymphatic vascular endothelial hyaluronan receptor 1 (LYVE-1) (Chapter 12). Another ECM protein called "Link protein" (due to its function of "linking" aggrecan and HA) binds to both HA, with its own Link domain, and the G1 domain of aggrecan via an Ig-fold [22]. This trimeric (HA–Link protein–aggrecan) complex creates a stable and long-lasting three-dimensional structure that is strong, yet flexible enough to withstand the high-impact physical forces constantly experienced by our joints and tissues. Binding studies using surface plasmon resonance show that a fragment of aggrecan comprising G1 and G2 binds HA with a K_d of 226 nM, whereas the Link protein binds HA with a K_d of 89 nM, a higher affinity [23].

The G2 domain shares some structural features with G1, but lacks the ability to bind HA. It is heavily substituted with KS chains and N-terminally adjacent to the CS attachment region. Both G1 and G2 domains inhibit aggrecan monomer secretion until the protein is fully modified and correctly folded. On the other hand, the G3 domain is known to promote secretion of the monomer and progression of the protein along the secretory pathway after GAG addition [24]. The region between the G1 and G2 domains, the interglobular domain (IGD region), is cleaved by proteolysis during normal aggrecan turnover. Conditions in which there is an increased exposure to, or expression of, inflammatory cytokines (e.g., as found in arthritis) promote the degeneration of aggrecan and weakening of the cartilage. The loss of the G2 through G3 portion of aggrecan changes the physicochemical attributes of the HA–aggrecan aggregate structure, which is difficult to repair, since the G1 domain fragment may remain bound to HA and thus prevent the binding of newly synthesized intact aggrecan monomers with HA [25–27].

The G3 or C-terminal domain of aggrecan contains two EGF-like domains (EGF1 and EGF2), a CRD, and a complement regulatory protein B component or short complement repeat (SCR). The primary role of the G3 domain is to aid secretion of the aggrecan monomer from chondrocytes after protein synthesis and GAG assembly in the ER/Golgi. The G3 domain is essentially a cap that prevents degradation, while the individual subdomains within G3 assist with processing, trafficking, and secretion of the heavily glycosylated core protein [28,29]. The G3 domain in articular cartilage is primarily associated with newly synthesized or immature aggrecan, whereas most mature aggrecan core proteins have lost their G3 domains in age-related processes [30].

Assembly of the aggrecan supramolecular structure is a highly ordered process that begins on the chondrocyte surface. In humans, HA is made at the cell surface by any of the three membrane-bound HA synthase isozymes (HAS1, 2, or 3) [31–33]. As HA is extruded from the cell surface, newly synthesized aggrecan and Link proteins may bind with the HA to form aggregates. The Link protein is required for proper alignment and ordering of the aggrecan along the HA [34]. The details for how Link protein is secreted and helps assemble these structures are not known. Furthermore, since cartilage is an acellular tissue, special processes are needed to facilitate the assembly of new ECM components many cell diameters away from the site of synthesis. One facilitation mechanism is that not all immature aggrecan monomers can bind HA, which is advantageous for development, since these aggrecan monomers then have time to travel away from the cell (i.e., the site of synthesis) before they become anchored to a HA molecule. This delayed incorporation of aggrecan may be the result of structural changes in the G1 domain [35,36], or the interaction of the G3 domain with other

ECM structural proteins such as fibulin1/2, tenacin-C and tenacin-R [37]. Unincorporated aggrecan has a short half-life, 24 h or less, whereas aggrecan incorporated into HA aggregates has a half-life of many years [38]. The immature unassociated aggrecan monomer needs to undergo several physical changes, depending on the environment or it will be degraded. Once aggrecan binds with HA and Link protein, the molecule is very stable and can proceed to assemble into supramolecular aggregates with other aggrecan and ECM molecules such as biglycan, decorin, fibulin, and the collagens.

15.3 VERSICAN

Human versican is a "versatile" molecule found in the ECM of connective tissues and the ECMs of a wide variety of "soft" tissues, including brain, arterial and venous vasculature, epithelia, and mesenchyme. The full-length core protein is the longest of the four lecticans, discussed in this chapter, and is encoded by a gene that extends over 100 kb and includes 15 exons. The core protein is organized in a fashion similar to the other lecticans (Figure 15.1) with a HA-binding globular G1 domain at the N-terminus and two GAG attachment domains, termed GAG-α and GAG-β, in the middle region encoded by 3 and 5.3 kb exons, respectively. The C-terminal portion of the protein contains two EGF-like domains, a CRD and a complement regulatory region. Alternative splicing of versican results in three variants (termed V1, V2, and V3), as well as full-length (V0), that are encoded by 3, 9, 6.5, and 12 kb transcripts, respectively. The molecular masses for these variants (V0–V3) are 370, 262, 180, and 72 kDa. V1 does not include GAG-α, V2 does not contain GAG-β, and V3 lacks both GAG attachment regions and is, therefore, technically not a proteoglycan [19].

Versican is the major proteoglycan in most arteries and veins, where it interacts with HA and Link proteins to form stable high molecular mass aggregates that fill spaces not occupied by other proteins such as collagen and elastin. The interaction of versican with other proteins and HA provide the blood vessels with walls that are both firm and elastic, thus able to contract or dilate in response to blood pressure, temperature changes, etc. In addition to normal turnover, inflammation and aging promote the degradation of versican, a process in which the core protein is cleaved by members of the ADAMS-TS metalloproteases family, in addition to matrilysin and plasmin. The loss of versican in the blood vessel walls may contribute to a weakening of the overall structure, giving rise to aneurysms and pseudoaneurysms [39,40].

Arterial smooth muscle cells (ASMC) are the primary source of versican in both arteries and blood vessels and express the transcripts to produce V0, V1, and V3 [41,42]. Platelet-derived growth factor, through the MAP kinase pathway, stimulates an increase in versican core protein transcription and CS chain elongation [43]. Conversely, IL-1β produced by inflammatory-stimulated macrophages decreases versican expression, suggesting that macrophage binding to vasculature lesions may contribute to the versican breakdown associated with this pathology [44]. Versican enriched with CS-6 and CS-4 chains is involved in the advanced stages of atherosclerotic lesions. Versican is prominent near the edges of necrotic cores and in close proximity to lipoproteins [45]. *In vitro* binding studies demonstrate that multiple low-density lipoprotein (LDL) particles can bind to single CS chains with an affinity of 23 nM, and an increase in chain length directly correlates with increased binding of LDL particles with the same affinity [18,46]. Treatment of ASMCs with growth factors such as EGF and TGF-β also cause elongation of CS chains on versican and may serve as a positive feedback stimulus in the early stages of atherogenesis [47,48].

The increased expression of versican V2, a product of oligodendrocytes, occurs in the periphery of brain lesions. *In vitro* cultures of oligodendrocytes indicate that the expressed versican acts as a linker between HA and tenascin-R, forming a rudimentary ECM. Axon regeneration failure is induced due to this injury response, since versican mediates growth-cone collapse and growth inhibition through its CS chains and core protein sequence [49]. Since all of the lecticans, including versican, present in the brain are growth-inhibitory, several methods to eliminate versican have been devised in order to promote axon regeneration. These include

the use of antiversican antibodies to eliminate versican expressing cells via complement-killing in the rat [50] and the infusion of antimitotic agents to prevent oligodendrocyte recruitment to the lesion [51].

15.4 NEUROCAN

The rat brain proteoglycan, neurocan, was identified and characterized using monoclonal antibodies [52]. During embryonic development, neurocan is commonly found as a 250 kDa monomer, and then later during maturation, it occurs as two protein isoforms of 130 and 150 kDa due to proteolysis at a MMP-2 cleavage site [53]. The N-terminal 130 kDa segment is the major HA-binding component and contains two CS chains, while the C-terminal 150 kDa segment binds with other ECM proteins and calcium ions [54]. Expression of neurocan is confined to the CNS, although it is present in the developing sensory nerves of the optical and olfactory systems in lower animals [55,56]. In mature brain, the expression of neurocan is very low, but is markedly increased in areas undergoing matrix remodeling. In fact, expression of neurocan may be required for remodeling of hippocampal synaptic connections during learning and memory-induced changes in synaptic transmissions [57].

Neurocan interacts with other molecules via the core protein, its GAG chains, or both. Interacting partners of neurocan can be classified into three different groups. First, the matrix proteins tenascin-C and tenascin-R are important for ECM aggregation during brain development when neurocan and heparin, which enhances the activity, are localized to perineuronal nets (PNNs) [58,59]. Second, growth factors associated with mobility such as FGF-2 [60], HB-GAM, and amphoterin [61] interact with neurocan and also bind heparin. Binding with amphoterin and HB-GAM is CS-chain dependent in contrast to FGF-2, which binds to the core protein. These growth factors support neurite outgrowth and extension in the developing brain, whereas neurocan serves as a modulator for these growth factors. Third, neural cell adhesion molecules such as N-CAM, L1/Ng-CAM, and TAG-1/axonin-1 are involved with neurite outgrowth and navigation. Binding of these cell adhesion molecules is CS-dependent, although there is weak binding to the C-terminal portion of the core protein.

Although neurocan interacts with a wide variety of adhesion, growth, and matrix molecules, the neurocan knockout mouse does not exhibit any gross anatomical brain pathologies [62]. Despite the involvement of neurocan in modeling of the CNS matrix, the proteoglycan is not required for the final outcome of how the brain structure develops. Thus far, the phenotypes of neurocan deficiency have only been observed in mice and give no indication of the possible affects on higher-brain functions such as thought and reasoning. Interestingly, an L1/Ng-CAM deficiency causes mental retardation in humans, but causes no physical or mental manifestation in mice, proving that not all animal models reflect the human condition [63].

15.5 BREVICAN

Expression of brevican or brain-enriched hyaluronan-binding protein (BEHAB) is restricted to the brain and is particularly enriched in neuroglial sheaths of velate protoplasmic astrocytes in the cerebellar granular layer [64] and in PNNs of large neurons [65]. In the rat, brevican is expressed at high levels in the neuroaxis of the ventricular zones and may be important for the development of glial cells. Brevican was first discovered in rats and cats as a partial protein with a molecular mass of ~38 kDa [66]. The full-length 160 kDa core protein is now known to be proteolytically cleaved, resulting in two physiologically relevant fragments of 100 and 60 kDa [67].

The ECM of the brain contains cross-linking proteins called the tenascins, which serve not only as part of the lattice that supports the neuronal network of the brain, but also as coregulators of morphogenesis and remodeling associated with injury, learning, and growth. Neurocan binds strongly with tenascin-C, and both proteins increase in expression during tissue regeneration [68]. On the other hand, brevican binds with tenascin-R via its fibronectin III domain in a CS-independent manner [69]. Tenascin-R, HA,

and the lecticans, especially brevican and neurocan, are basic components of the PNNs first described by Camillo Golgi and Santiago Ramon y Cajal in the 1890s. Neurons and glial cells secrete the materials that compose the PNN, which is a reticular net-like feature covering these cells that forms the organized lattice material in the intercellular spaces [65]. The PNN, which is structurally similar in some ways to the aggrecan–HA proteoglycan aggregates of cartilage, provides a tougher shield to protect the enclosed neurons and glial cells from movement within an otherwise vulnerable soft tissue matrix.

Brevican is a unique marker for glioma, a type of cancer that originates within the glial cells of the brain and metastasizes very extensively within the cranial ECM, but does not extend outside the CNS boundaries [70]. One of the main features of a high-grade glioma is the presence of an underglycosylated form of brevican ($B/b_{\Delta g}$), which migrates at 150 kDa on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Other forms of brevican are also present, such as a differentially glycosylated form that migrates slightly larger than 160 kDa and the proteolytic by-products of ADAMTS-4-induced cleavage of the native protein [71–73]. These features are not present in other neuropathologies such as Alzheimer's disease, nonglial tumors, or Parkinson's disease. Like the other lecticans, brevican exerts its inhibitory control on neurite outgrowth and cell migration via it CS chains. Brevican ($B/b_{\Delta g}$), a likely agent for glioma aggressiveness, may tip the balance in the favor of growth and migration of neurons.

15.6 CONCLUSION AND FUTURE DIRECTIONS

Once thought to be ECM components that only hold tissues and organs in place, the lecticans are increasing recognized as dynamic molecules that influence cell behaviors through effector protein/GAG interactions and signaling pathways. The underlying causes of many intrinsic human diseases, such as Alzheimer's disease and stenosis, can be traced to a "malfunction" of one or more lecticans. The binding partners and matrix associations of this class of molecules are fairly well understood. More exciting future directions will focus on gene regulation and the development of transgenic mice with reporter genes inserted within the promoter regions of versican, brevican, and neurocan. Since little is known about the *cis* or *trans* regulatory elements of how these genes are turned on or off, the expression of reporter genes in conjunction with lectican expression will provide answers to questions such as how protein expression corresponds with injury or disease and what regulates that expression. Since the versican knockout is lethal in mice, reporter genes in this model and others will aid our understanding of gene expression during development, disease, and injury and may ultimately provide targets for gene therapeutic interventions.

ACKNOWLEDGMENT

We acknowledge support by National Institute of General Medical Sciences grants GM69961, GM35978, and F32-GM070262 from the National Institutes of Health.

REFERENCES

1. N.B. Schwartz, E.W. Pirok, J.R. Mensch, and M.S. Domowicz, Domain organization, genomic structure, evolution, and regulation of expression of the aggrecan gene family, *Prog Nucleic Acid Res Mol Biol* 62, 177–225, 1999.
2. P.A. Brittis, D.R. Canning, and J. Silver, Chondroitin sulfate as a regulator of neuronal patterning in the retina, *Science* 255, 733–736, 1992.
3. D.M. Snow, D.A. Steindler, and J. Silver, Molecular and cellular characterization of the glial roof plate of the spinal cord and optic tectum: A possible role for a proteoglycan in the development of an axon barrier, *Dev Biol* 138, 359–376, 1990.
4. D.M. Snow, M. Watanabe, P.C. Letourneau, and J. Silver, A chondroitin sulfate proteoglycan may influence the direction of retinal ganglion cell outgrowth, *Development* 113, 1473–1485, 1991.

5. T. Murakami and A. Ohtsuka, Perisynaptic barrier of proteoglycans in the mature brain and spinal cord, *Arch Histol Cytol* 66, 195–207, 2003.
6. A. Sandvig, M. Berry, L.B. Barrett, A. Butt, and A. Logan, Myelin-, reactive glia-, and scar-derived CNS axon growth inhibitors: Expression, receptor signaling, and correlation with axon regeneration, *Glia* 46, 225–251, 2004.
7. D. Carulli, K.E. Rhodes, D.J. Brown, T.P. Bonnert, S.J. Pollack, K. Oliver, P. Strata, and J.W. Fawcett, Composition of perineuronal nets in the adult rat cerebellum and the cellular origin of their components, *J Comp Neurol* 494, 559–577, 2006.
8. K. Satoh, N. Suzuki, and H. Yokota, ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin motifs) is transcriptionally induced in beta-amyloid treated rat astrocytes, *Neurosci Lett* 289, 177–180, 2000.
9. Y. Yamaguchi, Lecticans: Organizers of the rain extracellular matrix, *Cell Mol Life Sci* 57, 276–289, 2000.
10. J. Dudhia, Aggrecan, aging and assembly in articular cartilage, *Cell Mol Life Sci* 62, 2241–2256, 2005.
11. U. Rauch, Extracellular matrix components associated with remodeling processes in brain, *Cell Mol Life Sci* 61, 2031–2045, 2004.
12. N.B. Schwartz and M. Domowicz, Proteoglycans in brain development, *Glycoconj J* 21, 329–341, 2004.
13. U. Rauch, K. Feng, and X.H. Zhou, Neurocan: A brain chondroitin sulfate proteoglycan, *Cell Mol Life Sci* 58, 1842–1856, 2001.
14. U. Hartmann and P. Maurer, Proteoglycans in the nervous system—The quest for functional roles in vivo, *Matrix Biol* 20, 23–35, 2001.
15. A. Oohira, F. Matsui, Y. Tokita, S. Yamauchi, and S. Aono, Molecular interactions of neural chondroitin sulfate proteoglycans in the brain development, *Arch Biochem Biophys* 374, 24–34, 2000.
16. C.E. Bandtlow and D.R. Zimmermann, Proteoglycans in the developing brain: New conceptual insights for old proteins, *Physiol Rev* 80, 1267–1290, 2000.
17. C.L. Nutt, C.A. Zerillo, G.M. Kelly, and S. Hockfield, Brain enriched hyaluronan binding (BEHAB)/brevican increases aggressiveness of CNS-1 gliomas in Lewis rats, *Cancer Res* 61, 7056–7059, 2001.
18. T.N. Wight and M.J. Merrilees, Proteoglycans in atherosclerosis and restenosis: Key roles for versican, *Circ Res* 94, 1158–1167, 2004.
19. M. Rahmani, B.W. Wong, L. Ang, C.C. Cheung, J.M. Carthy, H. Walinski, and B.M. McManus, Versican: Signaling to transcriptional control pathways, *Can J Physiol Pharmacol* 84, 77–92, 2006.
20. M. Venn and A. Maroudas, Chemical composition and swelling of normal and osteoarthrotic femoral head cartilage. I. Chemical composition, *Ann Rheum Dis* 36, 121–129, 1977.
21. A. Maroudas, H. Muir, and J. Wingham, The correlation of fixed negative charge with glycosaminoglycan content of human articular cartilage, *Biochim Biophys Acta* 177, 492–500, 1969.
22. P.F. Goetinck, N.S. Stirpe, P.A. Tsonis, and D. Carlone, The tandemly repeated sequences of cartilage link protein contain the sites for interaction with hyaluronic acid, *J Cell Biol* 105, 2403–2408, 1987.
23. H. Watanabe, S.C. Cheung, N. Itano, K. Kimata, and Y. Yamada, Identification of hyaluronan-binding domains of aggrecan, *J Biol Chem* 272, 28057–28065, 1997.
24. W. Luo, T.S. Kuwada, L. Chandrasekaran, J. Zheng, and M.L. Tanzer, Divergent secretory behavior of the opposite ends of aggrecan, *J Biol Chem* 271, 16447–16450, 1996.
25. M.Z. Ilic, C.J. Handley, H.C. Robinson, and M.T. Mok, Mechanism of catabolism of aggrecan by articular cartilage, *Arch Biochem Biophys* 294, 115–122, 1992.
26. P. Loulakis, A. Shrikhande, G. Davis, and C.A. Maniglia, N-terminal sequence of proteoglycan fragments isolated from medium of interleukin-1-treated articular-cartilage cultures. Putative site(s) of enzymic cleavage, *Biochem J* 284 (Pt 2), 589–593, 1992.
27. A. Ratcliffe, J.A. Tyler, and T.E. Hardingham, Articular cartilage cultured with interleukin 1-increased release of link protein, hyaluronate-binding region and other proteoglycan fragments, *Biochem J* 238, 571–580, 1986.
28. J.M. Day, A.D. Murdoch, and T.E. Hardingham, The folded protein modules of the C-terminal G3 domain of aggrecan can each facilitate the translocation and secretion of the extended chondroitin sulfate attachment sequence, *J Biol Chem* 274, 38107–38111, 1999.
29. M.S. Domowicz, E.W. Pirok, III, T.E. Novak, and N.B. Schwartz, Role of the C-terminal G3 domain in sorting and secretion of aggrecan core protein and ubiquitin-mediated degradation of accumulated mutant precursors, *J Biol Chem* 275, 35098–35105, 2000.
30. J. Dudhia, C.M. Davidson, T.M. Wells, D.H. Vynios, T.E. Hardingham, and M.T. Bayliss, Age-related changes in the content of the C-terminal region of aggrecan in human articular cartilage, Biochem J 313 (Pt 3), 933–940, 1996.

31. P.H. Weigel, Functional characteristics and catalytic mechanisms of the bacterial hyaluronan synthases, *Int Union Biochem Mol Biol* 54, 201–210, 2002.
32. P.H. Weigel, V.C. Hascall, and M. Tammi, Hyaluronan synthases, *J Biol Chem* 272, 13997–14000, 1997.
33. N. Itano and K. Kimata, Mammalian hyaluronan synthases, *IUBMB Life* 54, 195–199, 2002.
34. M. Morgelin, D. Heinegard, J. Engel, and M. Paulsson, The cartilage proteoglycan aggregate: Assembly through combined protein–carbohydrate and protein–protein interactions, *Biophys Chem* 50, 113–128, 1994.
35. L.I. Melching and P.J. Roughley, Studies on the interaction of newly secreted proteoglycan subunits with hyaluronate in human articular cartilage, *Biochim Biophys Acta* 1035, 20–28, 1990.
36. J.A. Buckwalter, J.C. Pita, F.J. Muller, and J. Nessler, Structural differences between two populations of articular cartilage proteoglycan aggregates, *J Orthop Res* 12, 144–148, 1994.
37. A. Lundell, A.I. Olin, M. Morgelin, S. al Karadaghi, A. Aspberg, and D.T. Logan, Structural basis for interactions between tenascins and lectican C-type lectin domains: Evidence for a crosslinking role for tenascins, *Structure* 12, 1495–1506, 2004.
38. M.T. Bayliss, S. Howat, C. Davidson, and J. Dudhia, The organization of aggrecan in human articular cartilage. Evidence for age-related changes in the rate of aggregation of newly synthesized molecules, *J. Biol. Chem.* 275, 6321–6327, 2000.
39. A.D. Theocharis, I. Tsolakis, A. Hjerpe, and N.K. Karamanos, Human abdominal aortic aneurysm is characterized by decreased versican concentration and specific downregulation of versican isoform V(0), *Atherosclerosis* 154, 367–376, 2001.
40. A.P. Burke, H. Jarvelainen, F.D. Kolodgie, A. Goel, T.N. Wight, and R. Virmani, Superficial pseudoaneurysms: Clinicopathologic aspects and involvement of extracellular matrix proteoglycans, *Mod Pathol* 17, 482–488, 2004.
41. T.N. Wight and V.C. Hascall, Proteoglycans in primate arteries. III. Characterization of the proteoglycans synthesized by arterial smooth muscle cells in culture, *J Cell Biol* 96, 167–176, 1983.
42. V. Asundi, K. Cowan, D. Matzura, W. Wagner, and K.L. Dreher, Characterization of extracellular matrix proteoglycan transcripts expressed by vascular smooth muscle cells, *Eur J Cell Biol* 52, 98–104, 1990.
43. E. Schonherr, H.T. Jarvelainen, L.J. Sandell, and T.N. Wight, Effects of platelet-derived growth factor and transforming growth factor-beta 1 on the synthesis of a large versican-like chondroitin sulfate proteoglycan by arterial smooth muscle cells, *J Biol Chem* 266, 17640–17647, 1991.
44. I.J. Edwards, W.D. Wagner, and R.T. Owens, Macrophage secretory products selectively stimulate dermatan sulfate proteoglycan production in cultured arterial smooth muscle cells, *Am J Pathol* 136, 609–621, 1990.
45. I. Halpert, U.I. Sires, J.D. Roby, S. Potter-Perigo, T.N. Wight, S.D. Shapiro, H.G. Welgus, S.A. Wickline, and W.C. Parks, Matrilysin is expressed by lipid-laden macrophages at sites of potential rupture in atherosclerotic lesions and localizes to areas of versican deposition, a proteoglycan substrate for theáenzyme, *Proc Natl Acad Sci USA* 93, 9748–9753, 1996.
46. K.L. Olin, S. Potter-Perigo, P.H. Barrett, T.N. Wight, and A. Chait, Lipoprotein lipase enhances the binding of native and oxidized low density lipoproteins to versican and biglycan synthesized by cultured arterial smooth muscle cells, *J Biol Chem* 274, 34629–34636, 1999.
47. P.J. Little, L. Tannock, K.L. Olin, A. Chait, and T.N. Wight, Proteoglycans synthesized by arterial smooth muscle cells in the presence of transforming growth factor-{beta}1 exhibit increased binding to LDLs, *Arterioscler Thromb Vasc Biol* 22, 55–60, 2002.
48. S. Potter-Perigo, C. Baker, C. Tsoi, K.R. Braun, S. Isenhath, G.M. Altman, L.C. Altman, and T.N. Wight, Regulation of proteoglycan synthesis by leukotriene D4 and epidermal growth factor in bronchial smooth muscle cells, *Am J Respir Cell Mol Biol* 30, 101–108, 2004.
49. R.A. Asher, D.A. Morgenstern, M.C. Shearer, K.H. Adcock, P. Pesheva, and J.W. Fawcett, Versican is upregulated in CNS injury and is a product of oligodendrocyte lineage cells, *J Neurosci* 22, 2225–2236, 2002.
50. J.K. Dyer, J.A. Bourque, and J.D. Steeves, Regeneration of brainstem-spinal axons after lesion and immunological disruption of myelin in adult rat, *Exp Neurol* 154, 12–22, 1998.
51. K.E. Rhodes, L.D. Moon, and J.W. Fawcett, Inhibiting cell proliferation during formation of the glial scar: Effects on axon regeneration in the CNS, *Neuroscience* 120, 41–56, 2003.
52. U. Rauch, P. Gao, A. Janetzko, A. Flaccus, L. Hilgenberg, H. Tekotte, R.K. Margolis, and R.U. Margolis, Isolation and characterization of developmentally regulated chondroitin sulfate and chondroitin/keratan sulfate proteoglycans of brain identified with monoclonal antibodies, *J Biol Chem* 266, 14785–14801, 1991.
53. F. Matsui, E. Watanabe, and A. Oohira, Immunological identification of two proteoglycan fragments derived from neurocan, a brain-specific chondroitin sulfate proteoglycan, *Neurochem Int* 25, 425–431, 1994.

54. J. Stenflo, Y. Stenberg, and A. Muranyi, Calcium-binding EGF-like modules in coagulation proteinases: Function of the calcium ion in module interactions, *Biochim Biophys Acta* 1477, 51–63, 2000.
55. H. Li, T.C. Leung, S. Hoffman, J. Balsamo, and J. Lilien, Coordinate regulation of cadherin and integrin function by the chondroitin sulfate proteoglycan neurocan, *J Cell Biol* 149, 1275–1288, 2000.
56. H.J. Clarris, U. Rauch, and B. Key, Dynamic spatiotemporal expression patterns of neurocan and phosphacan indicate diverse roles in the developing and adult mouse olfactory system, *J Comp Neurol* 423, 99–111, 2000.
57. V.Y. Bolshakov, H. Golan, E.R. Kandel, and S.A. Siegelbaum, Recruitment of new sites of synaptic transmission during the cAMP-dependent late phase of LTP at CA3-CA1 synapses in the hippocampus, *Neuron* 19, 635–651, 1997.
58. K. Feng, I. Arnold-Ammer, and U. Rauch, Neurocan is a heparin binding proteoglycan, *Biochem Biophys Res Commun* 272, 449–455, 2000.
59. G. Bruckner, J. Grosche, S. Schmidt, W. Hartig, R.U. Margolis, B. Delpech, C.I. Seidenbecher, R. Czaniera, and M. Schachner, Postnatal development of perineuronal nets in wild-type mice and in a mutant deficient in tenascin-R, *J Comp Neurol* 428, 616–629, 2000.
60. P. Milev, P. Maurel, A. Chiba, M. Mevissen, S. Popp, Y. Yamaguchi, R.K. Margolis, and R.U. Margolis, Differential regulation of expression of hyaluronan-binding proteoglycans in developing brain: Aggrecan, versican, neurocan, and brevican, *Biochem Biophys Res Commun* 247, 207–212, 1998.
61. P. Milev, A. Chiba, M. Haring, H. Rauvala, M. Schachner, B. Ranscht, R.K. Margolis, and R.U. Margolis, High affinity binding and overlapping localization of neurocan and phosphacan/protein-tyrosine phosphatase-zeta/beta with tenascin-R, amphoterin, and the heparin-binding growth-associated molecule, *J Biol Chem* 273, 6998–7005, 1998.
62. X.H. Zhou, C. Brakebusch, H. Matthies, T. Oohashi, E. Hirsch, M. Moser, M. Krug, C.I. Seidenbecher, T.M. Boeckers, U. Rauch, R. Buettner, E.D. Gundelfinger, and R. Fassler, Neurocan is dispensable for brain development, *Mol Cell Biol* 21, 5970–5978, 2001.
63. T. Brummendorf, S. Kenwrick, and F.G. Rathjen, Neural cell recognition molecule L1: From cell biology to human hereditary brain malformations, *Curr Opin Neurobiol* 8, 87–97, 1998.
64. H. Yamada, B. Fredette, K. Shitara, K. Hagihara, R. Miura, B. Ranscht, W.B. Stallcup, and Y. Yamaguchi, The brain chondroitin sulfate proteoglycan brevican associates with astrocytes ensheathing cerebellar glomeruli and inhibits neurite outgrowth from granule neurons, *J Neurosci* 17, 7784–7795, 1997.
65. K. Hagihara, R. Miura, R. Kosaki, E. Berglund, B. Ranscht, and Y. Yamaguchi, Immunohistochemical evidence for the brevican-tenascin-R interaction: Colocalization in perineuronal nets suggests a physiological role for the interaction in the adult rat brain, *J Comp Neurol* 410, 256–264, 1999.
66. D.M. Jaworski, G.M. Kelly, and S. Hockfield, BEHAB, a new member of the proteoglycan tandem repeat family of hyaluronan-binding proteins that is restricted to the brain [published erratum appears in *J Cell Biol* 1997;137(2):521], *J Cell Biol* 125, 495–509, 1994.
67. H. Yamada, K. Watanabe, M. Shimonaka, M. Yamasaki, and Y. Yamaguchi, cDNA cloning and the identification of an aggrecanase-like cleavage site in rat brevican, *Biochem Biophys Res Commun* 216, 957–963, 1995.
68. P.L. Jones and F.S. Jones, Tenascin-C in development and disease: Gene regulation and cell function, *Matrix Biol* 19, 581–596, 2000.
69. A. Aspberg, R. Miura, S. Bourdoulous, M. Shimonaka, D. Heinegård, M. Schachner, E. Ruoslahti, and Y. Yamaguchi, The C-type lectin domains of lecticans, a family of aggregating chondroitin sulfate proteoglycans, bind tenascin-R by protein-protein interactions independent of carbohydrate moiety, *Proc Natl Acad Sci USA* 94, 10116–10121, 1997.
70. F. Thorsen and B.B. Tysnes, Brain tumor cell invasion, anatomical and biological considerations, *Anticancer Res* 17, 4121–4126, 1997.
71. D.M. Jaworski, G.M. Kelly, and S. Hockfield, The CNS-specific hyaluronan binding protein, BEHAB, is expressed during periods of glial cell generation and motility, *Semin Neurosci* 8, 391–396, 1996.
72. S.C. Gary, G.M. Kelly, and S. Hockfield, BEHAB/brevican: A brain-specific lectican implicated in gliomas and glial cell motility, *Curr Opin Neurobiol* 8, 576–581, 1998.
73. M.S. Viapiano, W.L. Bi, J. Piepmeier, S. Hockfield, and R.T. Matthews, Novel tumor-specific isoforms of behab/brevican identified in human malignant gliomas, *Cancer Res* 65, 6726–6733, 2005.

16 Roles of Coral Lectins in Morphological Change of Zooxanthellae

Hisao Kamiya, Mitsuru Jimbo, Kazuhiko Koike, and Ryuichi Sakai

CONTENTS

16.1 INTRODUCTION

In marine invertebrates, lectins, especially those in the hemolymph, are believed to contribute as nonself recognition factors in their defense mechanisms [1,2]. Moreover, there is an increasing number of evidence supporting that the marine invertebrate lectins are involved also in various endogenous processes such as embryonic developments [3] and biomineralization [4] together with self and nonself recognition. Lectins are also thought to be involved in extracellular biological events as well. Some lectins of marine invertebrates such as giant clams [5], sponges [6], and tunicates [7] are proposed to mediate the interaction between invertebrate hosts and symbiotic bacteria or microalgae. Because of the property of lectins that can recognize specific surface carbohydrate chains, it is feasible to postulate that lectins function as chemical factors in recognition of specific microorganisms by a host. In fact, an idea that lectins mediate the recognition between legumes and nitrogen-fixing bacteria has been reinforced by recent molecular biological evidence [8]. It is thus tempting to assume that lectin-mediated symbioses between marine invertebrate hosts and microorganisms are rather general phenomena.

If this assumption was true, lectins may be present in corals, which rely their nutrient supply largely on symbionts; that is, photosynthetic dinoflagellate of zooxanthellae (*Symbiodinium* spp.). It has also been reported that surface glycoproteins or glycoconjugates of symbiotic microalgae are important when recruitment of the algae by hosts, since glycosidase or lectin treatment of the symbiont surface attenuated the rate of symbiont cell acquisition [9–11]. These observations strongly suggested the presence of chemical factors that recognize microalgal cells, and lectins seem to be

the most plausible candidates. Some strains of *Symbiodinium* cells in culture transform once a day from flagellated swimming form (motile form) to nonswimming form (coccoid form). Cells divide only when they are in the coccoid form. This peculiar cycle called diel cycle is unique to the free living cells, and when the cells are harbored in the coral tissue, they are maintained in the coccid form [12–14]. Taking advantage of this unique phenomenon, we screened for marine coral extracts that regulate the diel cycle and found that D-galactose-binding lectins from the octocoral *Sinularia lochmodes* induces the morphological change of *Symbiodinium* spp. from the free swimming motile form into nonmotile coccoid form without affecting their viability [15]. This finding indicated that lectins take some part in the chemical communication between coral and symbiotic microalgae, and may provide a "chemistry way" of investigating one of the most intriguing biological phenomena in shallow tropical sea. In this chapter, we briefly summarize previous findings regarding coral-derived lectins and describe the chemical properties and physiological activity of *S. lochmodes* lectins to *Symbiodinium* cells and discuss their possible contribution in the establishment of symbiosis between hosts and symbiotic microalgae.

16.2 LECTINS IN CORALS: AN OVERVIEW

In spite of increasing amount of studies on the animal lectins, those in corals are limited in both number and scope to date. In octocorals, the first example is an D-galactan-binding lectin from the gorgonian *Eunicella cavolinii* [16,17]. *E. cavolinii* lectin inhibited self-aggregation of the gorgonian cell line and was assumed to control cell–cell interactions as a negative modulator [16]. A D-galactose-binding lectin was also isolated from an unidentified *Sinularia* sp. collected in Fiji [18]. The purified lectin, named Sinularian, induced the binding of macrophages to tumor cells and also inhibited the cleavage of sea urchin egg irreversibly at the stage of blastula. Another species of Okinawan ocotocoral, *Lobohphytum variatum*, contained mucin type I binding lectins. Two lectins isolated (LVL-1 and LVL-2) were glycoproteins composed of subunits having a *M*r of 53 kDa [19]. Ecological or physiological importance of the lectins was left to be studied.

In hexacorals, isolation of the lectin was first reported from the zoanthid *Gerardia savaglia* [20]. The D-mannose-binding lectin showed plural bioactivity including a mitogenic activity to lipopolysaccharide-treated mice spleen lymphocytes, inhibition of the nuclear envelop mRNA translocation system of rat liver, and inhibition of human immunodeficiency virus type I (HIV-1, strain HTLV-IIIB) infection on H9 lymphocytes [21]. The lectin also inhibited syncytium formation in the HTLV-IIIB/H9-Jurkat cell system and the inhibition was abolished in the presence of D-mannose. However, the role of *Gerardia* lectin in the lectin-producing animal has not been reported.

16.3 SLL-2, A LECTIN FROM THE OCTOCORAL *SINULARIA LOCHMODES*: A PUTATIVE CHEMICAL FACTOR IN THE CORAL–ZOOXANTHELLAE SYMBIOSIS

16.3.1 ISOLATION AND CHEMICAL PROPERTIES OF *S. LOCHMODES* LECTIN, SLL-2

During the screening for the lectins in corals, we found that the octocoral *S. lochmodes* of Okinawan waters possessed the lectin with a binding specificity to D-galactose and related carbohydrates. The major lectin, SLL-2, was isolated by affinity chromatography on the acid-treated Sepharose 4B gels and gel filtration [22]. The color of the collected octocorals or their extract rapidly turned black and the purified lectin preparation was also brownish in color, suggesting the melanization of lectin molecule or the formation of melanin–lectin complex. The "blackening" or "melanization" is frequently accompanied by insolubilization of the extract. Additions of kojic acid, a tyrosinase inhibitor, and L-ascorbic acid, an antioxidant, to the extract was effective in preventing the blackening of the extract and insolubilization. A combination of affinity chromatography using D-galactosamine-bound HiTrap column and gel filtration improved the yield of SLL-2 up to 30 mg

from a 100 g of raw specimen, as much as three times higher than the purification procedure without the tyrosinase inhibitor and the antioxidant [23].

To examine the sugar-binding properties, frontal affinity chromatography of SLL-2 using 135 different carbohydrates was carried out by Dr. S. Nakamura at the National Institute of Advanced Industrial Science and Technology, Japan, to whom we are indebted for the analysis. SLL-2 was found to bind specifically to carbohydrate constituents derived from glycolipids but not from N-liked carbohydrates: that is, SLL-2 showed high affinity to Forssman antigen and certain kinds of carbohydrate having a Galα1–3Gal structure. This is of our particular interest since some glycosphingolipids, having carbohydrate moiety similar to Forssman antigen, were known to construct a microdomain on cell surface and concerned in cell surface signal transduction such as kinase activation [24,25], leaving a possibility that SLL-2 recognizes some signaling-related carbohydrate chains on the zooxanthella cell surface.

16.3.2 Cloning of cDNA-Encoding SLL-2

In sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), SLL-2 gave a single protein band at a *M*r of 16.5 kDa with or without a reducing reagent, while the molecular weight of intact SLL-2 was estimated to be 122 kDa by gel filtration [22]. In 2D-PAGE, however, the component of 16.5 kDa was further separated into nine spots with p*I* values ranging from 3 to 5, indicating the heterogeneity of SLL-2 subunits. In cloning of SLL-2 cDNA based on the amino acid sequence of amino-terminal regions, three closely related cDNA named SLL-2a, b, and c (98 amino acids residues each) were obtained. The deduced amino acid sequences showed more than 78% identity to each other. Together with the hypothetical proteins encoded by an open reading frame of ATP synthase operon in *Rhodobacter capsulatus* and others, SLL-2 had 33% similarity to the C-terminal region of discoidin I, a D-galactose-binding lectin from slime mold *Dictyostelium discoideum* [26] expressed in a developmentally regulated manner [27]. Recently, an antimicrobial snail lectin HPA from *Helix pomata* was reported to show 23–25% identity (depending on different SLL-2 subunits) to SLL-2 [28].

16.3.3 Localization of SLL-2 in *S. lochmodes*

The immunohistochemical localization of SLL-2 in the coral tissue was investigated using polyclonal anti-SLL-2 antibody (polyclonal IgG) raised in rabbit. On the cross-section of a coral polyp, numerous *Symbiodinium* cells in their coccoid form were observed in the internal mesenteries (Figure 16.1). On the immunostained section treated with the antibody, SLL-2 immunoreactivity was distributed in whole tissue of the coral, including mesogloea, and densely on the interfacial area of the symbionts but not within the algal cells (Figure 16.1B). At the level of transmission electron microscopy, above results were confirmed; particles of antibody-conjugated colloidal gold (indicated by black dots in Figure 16.2) distributed within the host cell but most densely on the cell wall surface of the *Symbiodinium* cell. Some dense labeling was also observed within the secretory granules of the host cell (Figure 16.2B) and the nematocyte (Figure 16.2C).

16.3.4 Activity of SLL-2: In-Culture Experiment against *Symbiodinium*

As mentioned, cultured (free living) *Symbiodinium* cells are in a diel cycle, in that cell transforms its shape once a day. This morphological change was of our particular interest since the algae maintained its coccoid form entirely in its host. We thus constructed a model system using three strains of culturable *Symbiodinium* spp. P083-2, JCUCS-1, and CS-156, since the symbiont from *S. lochmodes* was not culturable. Note that these strains belong to different clades, namely A, B, and F, respectively according to recent classification of *Symbiodinium* spp. based on several loci of rRNA gene [29,30]. Among these strains we considered, CS-156 would be the most representative strain for

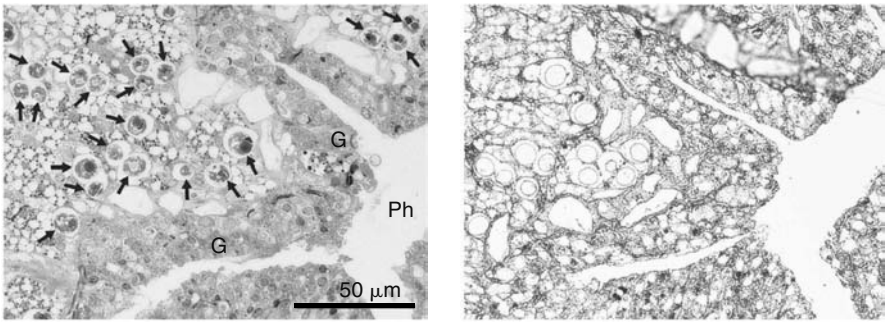


FIGURE 16.1 Light micrographs of a cross-section of *Sinularia lochmodes* polyp. (A) Close-up of a tentacle counterstained with toluidine-blue. Numerous *Symbiodinium* cells are present (arrows). (B) Same field to the left micrograph but stained with anti-SLL-2 IgG (rabbit-poly) and labeled with secondary IgG (antirabbit goat) conjugated with colloidal gold and then followed by silver enhancement. Silver-enhanced area distributes to the whole host tissue and on the surface of the *Symbiodinium* cells. G: gastrodermis; Ph: pharynx.

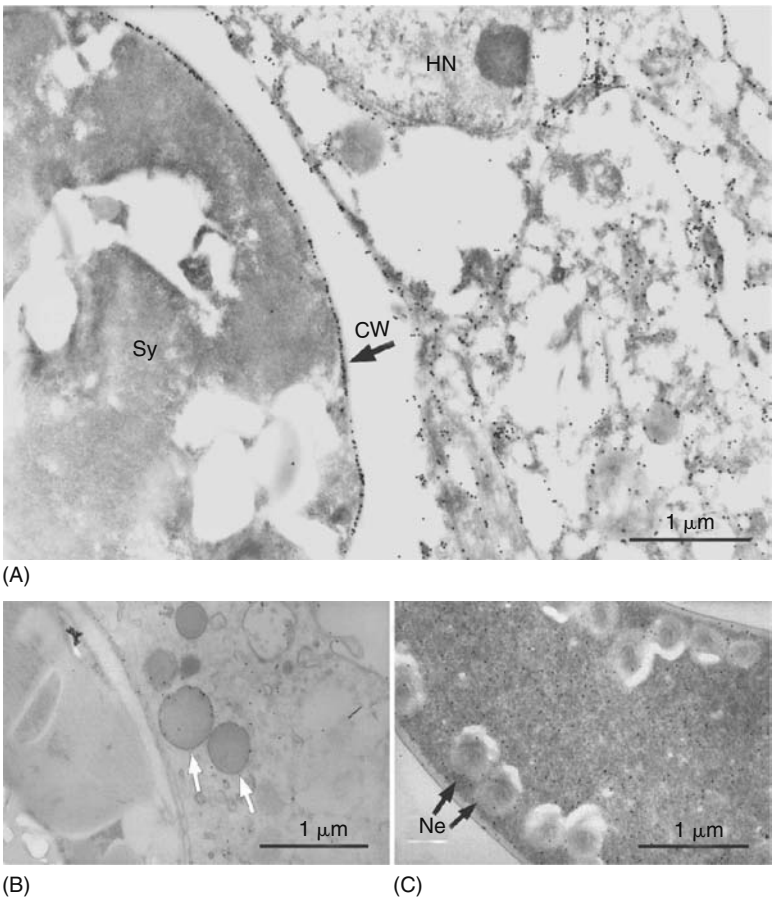


FIGURE 16.2 Transmission electron micrographs of *Sinularia lochmodes* reacted with anti-SLL-2 IgG, and labeled with secondary antibody (antirabbit goat IgG) conjugated with colloidal gold (15 nm). (A) A section showing a *Symbiodinium* cell (Sy) and a host cell. Particles of the colloidal gold distribute densely on the surface of *Symbiodinium* cell wall (CW) and sparsely within the host cytoplasm. (B) The particles seem to localize to secretory granules of the host cell (white arrows). (C) Internal fluid of nematocyte is also densely labeled. HN: host nuclei; Ne: nematocysts.

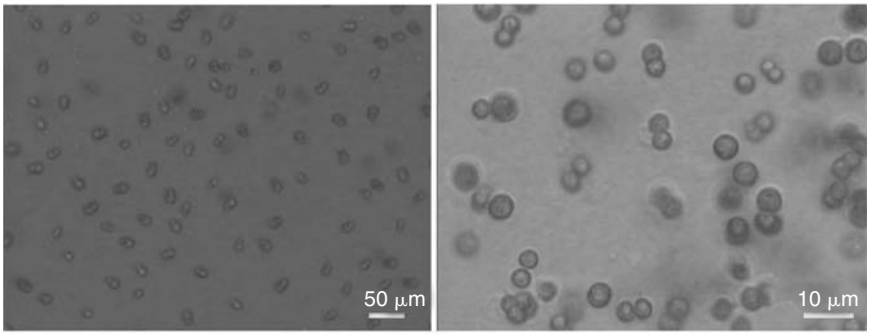


FIGURE 16.3 Video-captured images of *Symbiodinium* (strain CS-156) after SLL-2 treatment. Actively swimming cells in the control experiment (without SLL-2), showing characteristic motile cell form (left). Cells with SLL-2 treatment (day 3, 100 μg/ml) showing only coccoid-form cells and some dividing cells are present (right).

SLL-2 interaction that can occur in *S. lochmodes*, since *S. lochmodes* harbors mainly clade C [22], which is a sister taxon of clade F.

The assay was performed by adding test compound to the acclimated culture of each strain. Cells were in synchronized regular diel alternation between motile and nonmotile (coccoid) forms, and more than 70% of the cells were in the motile form at the point of sample addition (9–10 AM). The cells aggregated to form clump within 1 h after the addition of SLL-2 (100 μg/mL) and then started to transform into stationary coccoid form within 3 days (Figure 16.3). Most of cells were arrested at the coccoid form for the whole experimental period of 1 week, but cells seemed to be healthy with regard to their cell division rate. At the lower concentration (10 μg/mL), differential effect of SLL-2 to strains emerged. At these concentrations (10 or 100 μg/mL), the growth of CS-156 and JCUCS-1 was not affected, whereas that of P083-2 gradually became inhibited, even at the lower concentration tested [15].

The effects of SLL-2 on the *Symbiodinium* culture should be mediated by binding of the lectin to some carbohydrate moiety of the cell surface since addition of 0.2 M melibiose, the most effective inhibitor of the hemagglutination activity of SLL-2 [22], to the culture media completely inhibited its effect. Of note, other D-galactose-binding lectin such as *Arachis hypogaea* (peanut) lectin did not show any activity even at higher concentrations (~300 μg/mL) [15].

When the effects of SLL-2 on nonsymbiotic dinoflagellates were tested, cultures of *Gymnodinium catenatum* and *Prorocentrum micans* resulted in death after the cells inflated and burst within 24 h at 100 μg/mL. A chlorophyte, *Tetraselmis* sp., aggregated and formed cell clumps with complete loss of cell motility and significant inhibition of growth. Immunostaining of the SLL-2-treated microalgae using the anti-SLL-2 antibody indicate that the lectin localized on the microalgal cell surface (Figure 16.4) [15].

16.4 CONCLUSION AND FUTURE DIRECTIONS

Symbioses between macro- and microorganisms provide in general a metabolic diversity to the hosts, and microorganisms gain stable inhabitant in return. In oligotrophic tropical and subtropical milieu, symbiotic associations between photosynthetic zooxanthellae (*Symbiodinium* spp.) and invertebrates such as corals are crucial for the survival of host animals since they depend mainly on photosynthetic products translocated from *Symbiodinium* spp. In the case of corals, however, the evacuation of *Symbiodinium* from hosts, known as "coral bleaching," often resulting in the death of host animals, has become serious problem in tropical region. Causes of this phenomenon are largely unknown but there is clearly an unequal rate between acquisition and evacuation of the algae. At present, the host animals are postulated to acquire *Symbiodinium* either by vertical transmission

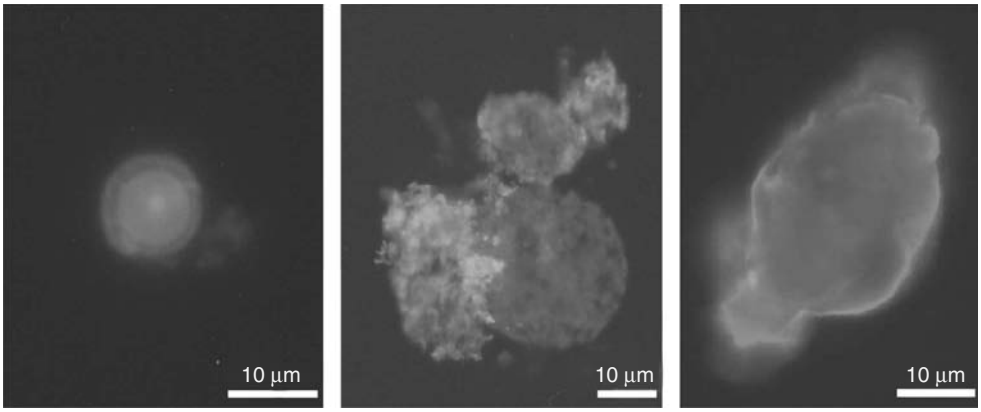


FIGURE 16.4 (See color insert following blank page 170. Also see CD for color figure.) Fluorescent micrographs of immunostained cells at day 6 after treatment of SLL-2 (100μg/mL). Binding of anti-SLL-2 antibody is shown by green fluorescence come from FITC-secondary antibody. Red fluorescence indicates the presence of chlorophyll *a* in the plastids. (A) CS-156; (B) destroyed cell of *Gymnodinium catenatum* with stained debris and (C) burst cell of *Prorocentrum micans* also with stained debris. (Modified from Koike et al., *Biol. Bull*., 207, 80, 2004. With permission.)

(acquisition from the mother) or by horizontal transmission (acquisition from the environment by either larval or adult stages) [31]. In the latter case, the animals are supposed to acquire *Symbiodinium* cells suited to the environment where they have settled. For the initial acquisition or infection of *Symbiodinium*, two main mechanisms have been suggested: Chemotactic attraction of free swimming motile *Symbiodinium* cells to the host's mouth and gastric cavity [32,33] and the capturing *Symbiodinium* cells while feeding [33,34]. *Symbiodinium* spp. must be selected, regardless of the mechanisms for acquisition, from nonsymbiotic microalgae at some point by the host animals. Several papers have published on the mechanism of selective acquisition of the algae [11,35–40].

On the basis of observation in our model, following scenario can be proposed for the early event of symbiosis: Planktonic microalgae introduced to the gastric cavity of the coral are subject to chemical selection due to their differential responses to lectins; that is, the lectins induce nonsymbiotic algae to burst helping digestion. On the other hand, the lectins turn physiology of *Symbiodinium* into the mode that may be favorable to the symbiosis between the algae and the coral. Our observation of the discrete effects of SLL-2 on different types of microalgae suggests that further selection may occur among the different clades of *Symbiodinium*. Actually, in gorgonians, infection by several clades of *Symbiodinium* occur at the onset of symbiosis, however, the population eventually converge into certain clade of microalgae [36].

Recently, Wood-Charlson et al. [41] reported that at the onset of symbiosis between *Fungia scutaria* larvae and their endosymbiotic dinoflagellate algae, the recognition of algal cell surface was key to successful infection. They identified α-mannose/α-glucose and α-galactose residues as potential recognition ligands on the algal cell surface. Thus they postulated that lectins/glycans could be concerned deeply with the infection of *Symbiodinium* in the hexacoral *Fungia scutaria*. In our survey for the lectins of scleractian coral that is capable of inducing the morphological change of *Symbiodinium* cells, we have found a calcium-dependent D-galactose-binding lectin in *F. echinata*. The lectin transformed the swimming cells of the *Symbiodinium* strain of *Fungia* origin to the coccoid forms. These results suggested that the regulation of *Symbiodinium* morphology is common among coral lectins. Details of *Fungia* lectin will be published elsewhere.

Investigation of the physiological and ecological aspects of these symbioses especially focused on the chemical compounds, which select and maintain the symbiotic dinoflagellates, might allow us to better understand the environmental and ecological problems such as coral bleaching, which is triggered by the evacuation of *Symbiodinium* cells from the host.

REFERENCES

1. Kilpatrick, D.C., Animal lectins: A historical introduction and overview, *Biochim. Biophys. Acta,* 1572, 187, 2002.
2. Vasta, G.R. and Lambris, J.D., Innate immunity in the Aegean: Ancient pathways for today's survival, *Dev. Comp. Immunol.,* 26, 217, 2002.
3. Alliegro, M.C. and Alliegro, M.A., The structure and activities of echinonectin: A developmentally regulated cell adhesion glycoprotein with galactose-specific lectin activity, *Glycobiology*, 1, 253, 1991.
4. Muramoto, K. et al., Inhibition of the growth of calcium carbonate crystals by multiple lectins in the coelomic fluid of the acorn barnacle *Megabalanus rosa, Comp. Biochem. Physiol*., 107B, 401, 1994.
5. Uhlenbruck, G. and Steinhausen, G., Tridacnins: Symbiosis-profit or defense-purpose? *Dev. Comp. Immunol.,* 1, 183, 1977.
6. Müller, W.E.G. et al., Lectin, a possible basis for symbiosis between bacteria and sponges, *J. Bacteriol*., 145, 548, 1981.
7. Müller, W.E.G. et al., Biochemical basis for the symbiotic relationship *Didemnum–Prochloron* (Prochlorophyta), *Biol. Cell,* 51, 381, 1984.
8. Hirsch, A.M., Lum, M.R., and Downie, J.A., What makes the rhizobia-legume symbiosis so special? *Plant Physiol.,* 127, 1484, 2001.
9. Weis, D., Correlation of infectivity and concanavalin A agglutinability of algae exosymbiotic from *Paramecium bursaria, J. Protozool.,* 26, 117, 1978.
10. Reisser, W., Radunz, A., and Wiessner, W., Participation of algal surface structures in the cell recognition process during infection of aposymbiotic *Paramecium bursaria* with symbiotic chlorellae, *Cytobios,* 33, 39, 1982.
11. Lin, K.L., Wang, J.T., and Fang, L.S., Participation of glycoproteins on zooxanthella cell walls in the establishment of a symbiotic relationship with the sea anemone, *Aiptasia pulchella*, *Zoolog. Studies*, 39, 172, 2000.
12. Fitt, W.K. and Trench, R.K., The relation of diel patterns of cell division to diel patterns of motility in the symbiotic dinoflagellate *Symbiodinium microadriaticum* Freudenthal in culture, *New Phytol*., 94, 421, 1983.
13. Trench, R.K., Microalgal-invertebrate symbioses: A review, *Endocytobiosis Cell Res*., 9, 135, 1993.
14. Yacobovitch, T., Benayahu, Y., and Weis, V.M., Motility of zooxanthellae isolated from the Red Sea soft coral *Heteroxenia fuscescens* (Cnidaria), *J. Exp. Mar. Biol. Ecol*., 298, 35, 2004.
15. Koike, K. et al., Octocoral chemical signaling selects and controls dinoflagellate symbionts, *Biol. Bull*., 207, 80, 2004.
16. Müller, W.E.G. et al., Cell–cell recognition system in gorgonians: Description of the basic mechanism, *Mar. Biol*., 76, 1, 1983.
17. Kljajic, Z. et al., Possible control mechanism of cell motility in the gorgonian *Eucella cavolinii*. *Mar. Biol*., 84, 225, 1985.
18. Goto, R. et al., Purification and characterization of an agglutinin of the soft coral *Sinularia* species, *Develop. Comp. Immunol*., 16, 9, 1992.
19. Goto-Nance, R. et al., Purification and characterization of the lectins of the soft coral *Lobophytum variatum*, *Fish. Sci*., 62, 297, 1996.
20. Kljajic, Z. et al., A D-mannose-specific lectin from *Gerardia savaglia* that inhibits nucleocytoplasmic transport of mRNA, *Eur. J. Biochem*., 169, 97, 1987.
21. Müller, W.E.G. et al., The D-mannose-specific lectin from *Gerardia savaglia* blocks binding of human immunodeficiency virus type I to H9 cells and human lymphocytes in vitro, *J. Acquir. Immun. Defic. Syndr*., 1, 453, 1988.
22. Jimbo, M. et al., The D-galactose-binding lectin of the octocoral *Sinularia lochmodes*: Characterization and possible relationship to the symbiotic dinoflagellates, *Comp. Biochem. Physiol*., 125B, 227, 2000.
23. Jimbo, M. et al., Cloning and characterization of a lectin from the octocoral *Sinularia lochmodes, Biochem. Biophys. Res. Commun*., 330, 157, 2005.

24. Tillack, T.W. et al., Localization of globoside and Forssman glycolipids on erythrocyte membranes, *Biochim. Biophys. Acta*, 733, 15, 1983.
25. Sorice, M. et al., Evidence for the existence of ganglioside-enriched plasma membrane domains in human peripheral lymphocytes, *J. Lipid Res.*, 38, 969, 1997.
26. Barondes, S.H., Cooper, D.N., and Haywood-Reid, P.L., Discoidin I and discoidin II are localized differently in developing *Dictyostelium discoideum*, *J. Cell. Biol.*, 96, 291, 1983.
27. Springer, W.R., Cooper, D.N., and Barondes, S.H., Discoidin I is implicated in cell–substratum attachment and ordered cell migration of *Dictyostelium discoideum* and resembles fibronectin, *Cell*, 39, 557, 1984.
28. Sanchez, J.F. et al., Biochemical and structural analysis of *Helix pomatia* agglutinin. A hexameric lectin with a novel fold, *J. Biol. Chem.*, 281, 20171, 2006.
29. Carlos, A.A. et al., Phylogenetic position of Symbiodinium (Dinophyceae) isolates from tridacnids (Bivalvia), cardiids (Bivalvia), a sponge (Porifera), a soft coral (Anthozoa), and a free-living strain, *J. Phycol.*, 35, 1054, 1999.
30. LaJeunesse, T.C., Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: In search of a "species" level marker, *J. Phycol.*, 37, 866, 2001.
31. Trench, R.K., Dinoflagellates in non-parasitic symbioses, in *The Biology of Dinoflagellates*, Blackwell, Oxford, 1987, 530.
32. Kinzie, R.A., Experimental infection of aposymbiotic gorgonian polyps with zooxanthellae, *J. Exp. Mar. Biol. Ecol.*, 15, 335, 1974.
33. Fitt, W.K., The role of chemosensory behavior of *Symbiodinium microadriaticum*, intermediate hosts, and host behavior in the infection of coelenterates and mollusks with zooxanthellae, *Mar. Biol.*, 81, 9, 1984.
34. Riggs, L.L., Feeding behavior in *Aiptasia tagetes* (Duschassaing and Michelotti) planulae: A plausible mechanism for zooxanthellae infection of aposymbiotic planktotrophic planulae, *Caribb. J. Sci.*, 24, 201, 1988.
35. Reynolds, W.S., Schwarz, J.A., and Weis, V.M., Symbiosis-enhanced gene expression in cnidarian–algal associations: Cloning and characterization of a cDNA, sym32, encoding a possible cell adhesion protein, *Comp. Biochem. Physiol.*, 126 A, 33, 2000.
36. Coffroth, M.A., Santos, S.R., and Goulet, T.L., Early ontogenetic expression of specificity in a cnidarian–algal symbiosis. *Mar. Ecol. Prog. Ser.*, 222, 85, 2001.
37. Schwarz, J.A., Weis, V.M., and Potts, D.C., Feeding behavior and acquisition of zooxanthellae by planula larvae of the sea anemone *Anthopleura elegantissima*, *Mar. Biol.*, 140, 471, 2002.
38. Rodriguez-Lanetty, M., Chang, S.J., and Song, J.I., Specificity of two temperate dinoflagellates–anthozoan associations from the north-western Pacific Ocean, *Mar. Biol.*, 143, 501, 2003.
39. Schwarz, J.A. and Weis, V.M., Localization of a symbiosis-related protein, sym32, in the *Anthopleura elegantissima–Symbiodinium muscatinei* association, *Biol. Bull.*, 205, 339, 2003.
40. Rodriguez-Lanetty, M. et al., Temporal and spatial infection dynamics indicate recognition events in the early hours of a dinoflagellate/coral symbiosis, *Biol. Bull.*, 149, 713, 2006.
41. Wood-Charlson, E.M. et al., Lectin/glycan interactions play a role in recognition in a coral/dinoflagellate symbiosis, *Cell. Microbiol.*, 8, 1985, 2006.

Part V

Cell–Cell Interactions, Signaling, and Transport

17 Siglecs: Roles in Cell–Cell Interactions and Signaling

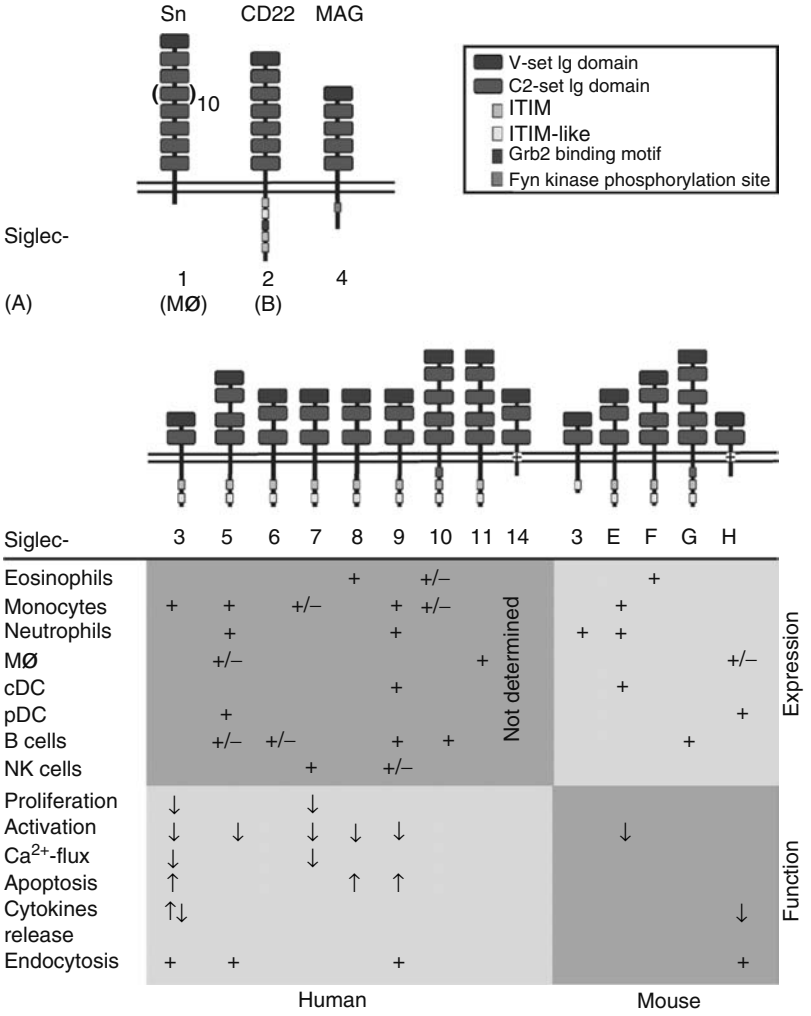
Cornelia Oetke and Paul R. Crocker

CONTENTS

17.1 INTRODUCTION

The siglecs (sialic acid–binding Ig-type lectins) belong to the Ig superfamily and form the main subpopulation of the Ig-type (I-type) lectins. They are type 1 membrane proteins, characterized by an amino-terminal V-set Ig domain, containing the sialic acid–binding site, followed by various numbers of C2-set Ig domains. Siglecs can be further subdivided into two subsets, the evolutionary conserved, but distantly related (~25%–30% sequence identity) group containing sialoadhesin (Sn, CD169, Siglec-1), CD22 (Siglec-2), and myelin-associated glycoprotein (MAG, Siglec-4) (Figure 17.1A), and the rapidly evolving group of CD33-related siglecs that share high sequence similarity (50%–99% sequence identity) (Figure 17.1B). Sn, CD22, and MAG have been found in all mammals so far examined and MAG additionally in birds [1] and fish [2]. In contrast, the repertoire of CD33-related siglecs varies considerably between species, with nine known in humans and only five in mice (Figure 17.1).

With the exception of MAG and Siglec-6, siglecs are primarily expressed by cells of the hematopoietic system (Figure 17.1). While some have a very restricted expression pattern, others are more widely expressed. For example, Sn is mainly expressed by macrophages, CD22 by B cells, Siglec-8 by eosinophils, and Siglec-7 by natural killer (NK) cells whereas Siglec-9 can be found on neutrophils, monocytes, dendritic cells, and NK cells. Since they also vary in their preference for sialylated ligands, each siglec probably has a distinct function. Most CD33-related siglecs and CD22 are inhibitory signaling receptors, having one or more immunoreceptor tyrosine-based motifs (ITIMs) in their cytoplasmic tails, whereas Sn seems to be involved in cell–cell interactions. This chapter will focus mainly on Sn and the CD33-related siglecs, while CD22 and MAG are discussed in more detail in Chapters 19 and 14, respectively.



(B) CD33-related Siglecs

FIGURE 17.1 (See color insert following blank page 170. Also see CD for color figure.) Schematic presentation of Siglecs, their expression and function. Shown are the two subgroups of Siglecs, (A) Sialoadhesin, CD22, and MAG; and (B) the CD33-related siglecs. Each box represents one Ig-domain, but note that 10 domains have been merged for Sn. Signaling motifs in the cytoplasmic tail are depicted by colored areas. Expression patterns of various siglecs on hematopoietic cells are given below. The cellular expression pattern for Siglec-14 is not determined, MAG is only expressed by glial cells and Siglec-6 is primarily expressed in the placenta. (B) The lower panel summarizes the effects of siglecs on various cellular functions. The arrows indicate an (↑) enhancing or (↓) reducing effect of antibody cross-linked siglecs on cellular functions and siglec involvement in endocytosis is shown as (+).

17.2 SIGLEC NOMENCLATURE

Given the complexity of siglec repertoires across species, it has been important to develop a universally accepted nomenclature [3]. The original siglecs, Sn, CD22, CD33, and MAG, were designated Siglecs-1, -2, -3, and -4, respectively, but are usually described by their original names. Additional siglecs were named in the order of their discovery. Roman numerals are used for siglec-like sequences with arginine mutations resulting in loss of their sialic acid–binding ability. Given the uncertainty over the orthologous relationships between human and mouse CD33-related siglecs, the murine proteins

were assigned a letter suffix (E–H) (Figure 17.1) [3,4]. A prefix can be used to distinguish species, for example, between human (h) and mouse (m).

17.3 SIGLECS AND THEIR SIALYLATED LIGANDS

By definition, all siglecs have in common the ability to bind sialic acids, a family of acidic amino sugars, with *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), and 9(7)-*O*-acetylated Neu5Ac being predominant in mammals [5]. Sialic acids are usually found at the terminal, nonreducing ends of oligosaccharide chains, in α2–3, α2–6, and α2–8 linkages on glycans of both glycolipids and glycoproteins. Similarly to many lectins, siglecs bind to sialylated oligosaccharides with a rather low affinity (K_d = 0.1–3 mM) and therefore depend on multivalent clustering for high-avidity binding. Each siglec has its own unique binding profile that varies not only in preference for linkage or sialic acid modification, but also in selectivity. Although sialic acid is the primary binding determinant, underlying glycans can contribute to individual siglec binding [6]. While hCD22 has a very strong preference for Neu5Ac(Gc)2-6Gal generated by the ST6GalI sialyltransferase, other siglecs, like mSiglec-E, bind to a wide variety of oligosaccharides, including Neu5Ac2-6Gal1-4GlcNAc, Neu5Gc2-6Gal1-4GlcNAc, Neu5Ac2-3Gal1-4GlcNAc, Neu5Ac2-8NeuAc2-3Gal1-4GlcNAc, Neu5Ac2-6Gal[6S]1-4GlcNAc(Fuc), and Neu5Ac2-6Gal1-4GlcNAc[6S](Fuc). Identification of the sugar ligands of various siglecs has advanced rapidly due to the development of glycan microarrays. hSiglecs-8, -9 and mSiglec-F, for example, were unexpectedly found to prefer specific types of sulfated sugars, but still had an absolute requirement for sialic acid [7–9].

Molecular insights into the interactions between siglecs and their ligands have been provided by siglec crystal structures, first of Sn in complex with 3′ sialyllactose [10] and more recently of Siglec-7 in the absence and presence of various sialylated sugars, including α2–8 disialylated GT1b oligosaccharide [6,11]. All siglecs have a conserved arginine residue on the F β-strand of the V-set Ig domain that anchors the carboxyl group of the terminal sialic acid inside the binding pocket [6,10,12]. Mutation of this residue has been shown to result in loss of binding of most siglecs. Although the V-set domain of Sn and Siglec-7 share a very similar fold and structural template for sialic acid recognition, there are also clear differences. The C–C′ loop in particular appears to be important in determining the fine ligand specificity of siglecs, and together with the C′–D loop, undergoes a major conformational change upon ligand binding [6].

For lectins, the combination of oligosaccharide ligands presented on a protein or lipid carrier is defined as a "counterreceptor" [13]. The membrane mucins CD43, PSGL-1, and MUC-1 [14,15] have been described as counterreceptors for Sn. CD22 has been shown to bind to itself, IgM, CD45, and members of the Ly-6 family [16–19], but the biological relavance of these interactions remains to be shown. The CD33-related siglecs are less well investigated in this respect. Siglec-7 expressed on NK cells can bind b-series gangliosides on target cells [6,20], while for Siglec-8 some 6′-sulfo-sLe^x presenting polypeptides like GlyCAM-1 have been suggested [7]. However, it remains to be seen whether siglec counterreceptors are important in mediating the biological function of siglecs or whether simply the overall density and presentation of appropriate sialylated ligands, irrespective of their carriers, are important.

17.4 *CIS*- AND *TRANS*-LIGANDS

The glycocalyx of cells is usually rich in sialic acids. Consequently, ligands for siglecs are not only presented on opposing cells (in *trans*), but also on the same cell surface (in *cis*), the latter resulting in masking of siglecs [21]. Sn appears to be an exception to the rule; due to its extended structure, it is thought to protrude from the cell surface, minimizing possible *cis*-interactions. To detect sugar-binding activity of siglecs on cell surfaces, using, for example, synthetic sialylated probes, cells usually need to be desialylated. However, it has been shown for CD22 that synthetic high-affinity sialosides can outcompete *cis*-interactions [22] and that CD22 is able to engage in *trans*-interactions

without unmasking first [23], demonstrating that *cis*-ligands downregulate, but do not preclude, binding of ligands in *trans*. Recent results indicate that *cis*-interactions of CD22 are mainly homotypic in nature, allowing the formation of homomultimeric complexes on the B cell surface [19].

The biological significance of *cis*- versus *trans*-interactions has been investigated for two siglecs, Siglec-7 and CD22. Siglec-7 is expressed on NK cells [24] and preferentially binds to α2,8-disialylated ligands [25]. In cytotoxicity assays, expression of Siglec-7 ligands on target cells (in *trans*) slightly reduced killing and this became more pronounced after removal of *cis*-interactions on NK cells by desialylation [20]. Similar for CD22, B cell activation was reduced when the cell displaying the antigen also expressed α2,6-linked sialic acids [26]. In contrast, B cells from ST6Gal-I-deficient mice do not express any *cis*-ligands for CD22 and have a hypoactive phenotype [27]. While *trans*-interactions recruit the inhibitory siglecs to the site of contact, resulting in dampened signaling, *cis*-interactions tether them away, allowing strong signaling despite the expression of inhibitory receptors. To what extent *cis*-interaction of siglecs are regulated remains to be shown, but there is some evidence that siglecs can become unmasked following cellular activation [21] and striking changes in expression of glycosyltransferases are well-known to occur during differentiation and activation of leukocytes.

17.5 SIALOADHESIN

Sn was first discovered as a sheep erythrocyte receptor on macrophages and was the first siglec to be characterized as a sialic acid–binding lectin [28]. In certain respects, it is an atypical siglec. It is by far the largest member of the siglec family with 17 Ig domains [29] and unlike most CD33-related siglecs and CD22, it lacks tyrosine-based signaling motifs in its cytoplasmic tail [30]. Although the protein is very well conserved between mammals (69%–78% between human, mouse, and pig) the amino acid identity of the cytoplasmic tail is low [4,31,32]. These characteristics argue for an adhesive, rather than signaling function.

Sn expression is restricted to subsets of resident and some inflammatory macrophages. In mice, the strongest expression can be found on metallophilic macrophages in spleen, supcapsular sinus, and medullary macrophages in lymph nodes, as well as on the majority of resident bone marrow macrophages. Alveolar macrophages, red pulp macrophages in spleen, and Kupffer cells express Sn at lower levels [33,34]. Under inflammatory conditions, infiltrating macrophages were often found to express Sn, as for example in experimental autoimmune uveoretinitis (EAU) [35], glomerulonephritis [36], rheumatoid arthritis [32], breast cancer [15], inherited demyelination [37,38], and in a graft versus leukemia model [39]. It has also been reported that Sn can be expressed by blood monocytes in HIV-infected patients [40] and by monocyte-derived dendritic cells cocultured with human rhinovirus [41]. Although Sn is a nonphagocytic receptor, it can facilitate macrophage uptake of sialylated *Neisseria meningitidis* strains *in vitro* [42] and mediate endocytic uptake of the porcine reproductive and respiratory syndrome virus (PRRSV) [31,43]. However, pig Sn has a potential internalization motif in its cytoplasmic tail that is absent in humans and mice [31]. It remains to be investigated whether the uptake of sialylated particles by Sn is relevant *in vivo*.

As mentioned above, CD43, PSGL-1, and MUC-1 [14,15] have been described as potential sialic acid–dependent counterreceptors for Sn on leukocytes and tumor cells. Additionally, other lectins were found to bind to Sn in a sialic acid–independent manner, like the mouse macrophage galactose-type C-type lectin 1 [44] and the cysteine-rich domain of the mannose receptor [45]. *In vivo*, Sn has been found in contact zones of macrophages and developing granulocytes in bone marrow [46], but so far no defect in granulocyte development in Sn-/- mice has been seen [47]. Under resting conditions, Sn-/- mice show only subtle alterations in the hematopoietic system, but the absence of Sn resulted in a delayed disease onset in a model of EAU, which was consistent with a reduced antigen-dependent proliferation of T cells from the draining lymph nodes [48]. Even more striking was the beneficial effect of Sn deficiency in two independent, genetically determined, demyelination models similar to inherited neuropathy [37] and multiple sclerosis in

humans [38]. Both macrophage and $CD8^+$ T cell infiltration were reduced in the absence of Sn, resulting in attenuated myelin degeneration and improved nerve conduction properties. In all three systems, macrophages in concert with T cells contribute to the disease progression, with EAU being $CD4^+$ T cell-mediated as opposed to the demyelination models which are $CD8^+$ T cell-mediated. In conclusion, the studies of Sn-deficient mice have revealed that Sn is important in promoting T cell-dependent autoimmune inflammatory responses, but the mechanisms remain elusive at this stage.

There is, however, contradictory evidence that Sn might fulfill inhibitory roles under certain circumstances. Coculture of dendritic cells (DCs) with human rhinovirus 14 (HRV), the causative agent of the common cold, induced expression of the inhibitory receptor B7-H1 (PD-L1), and Sn [41]. T cells, when cocultured with these DCs became anergic, a process that was partially overcome by Sn-blocking antibodies. Since HRV-stimulated DCs bound directly to T cells, it is possible that Sn interferes with synapse formation and T cell activation when expressed on DCs. Sn might fulfill different tasks, when expressed on macrophages, like affecting migration or phagocytosis. It is also possible that macrophage–T cell interactions are influenced by Sn, resulting in promotion of T cell activation under certain circumstances [49].

17.6 CD33-RELATED SIGLECS

The CD33-related siglecs share high sequence similarity and most are encoded within a gene cluster on chromosome 19 in humans and chromosome 7 in mice [4]. This region also contains many other members of the Ig superfamily expressed by leukocytes and is known as the leukocyte receptor complex [50]. CD33 was the first siglec of this group to be discovered, leading to the name CD33-related siglecs. This siglec subgroup has undergone rapid evolution, especially in primates, involving gene deletion or duplication, pseudogenization, exon shuffling or loss, and specific amino acid changes [4], resulting in extensive species differences. Most human and mouse CD33-related siglecs contain two cytoplasmic tyrosine-based motifs, a membrane proximal ITIM (S/I/V/LxYxxI/V/L) and a membrane distal ITIM-like motif (E/DYxEV/IR/K), and function as inhibitory receptors [51–55]. The exceptions are hSiglec-14 (no tyrosine-based motif), mCD33 (only an ITIM-like motif), and mSiglec-H (no tyrosine-based motif) (Figure 17.1B).

Typically, ligand engagement of inhibitory receptors results in tyrosine phosphorylation of the ITIM by Src family tyrosine kinases and recruitment of Src homology 2 (SH2)-domain containing phosphatases (SHPs), such as SHP-1 and SHP-2. SHP recruitment can result in decreased tyrosine phosphorylation of relevant substrate in the vicinity and regulate cellular activation [56]. Mutagenesis studies of CD33-related siglecs have shown that the membrane-proximal ITIM is both necessary and sufficient to recruit both SHP-1 and SHP-2. The membrane distal ITIM-like motif, however, is necessary for optimal binding of SHP-1, but is not required for SHP-2 recruitment [51–55,57]. For Siglec-5, Siglec-7, and Siglec-9, it was also shown that the membrane-proximal ITIM is dominant in mediating the inhibitory signal [51,55]. The membrane-distal ITIM-like motifs in some siglecs (TEYSEI/VK/R, e.g., in CD33, Siglec-5, -6, and -9) is similar to the tyrosine-based motif in SLAM-related receptors, which is recognized by the SH2 domain adaptors SAP and EAT2. Recent data suggest, however, that these adaptors require the consensus sequence TIYxxV/I [58] and there is currently no evidence that they interact with CD33-related siglecs.

Interestingly, mutation of the tyrosine within the membrane-proximal ITIM also increased sialic acid–dependent binding of CD33, Siglec-5, -7, and -9 [51,55,57]. Suppression of adhesion by the intact ITIM appeared to be SHP-2 dependent, because binding activity correlated inversely with the ability to recruit SHP-2, and cotransfection of catalytically inactive, dominant-negative SHP-2 mutants with wt Siglec-5 led to increased binding [55]. SHP-2 has been implicated in actin reorganization [59] and it is therefore plausible that SHP-2 recruitment could negatively regulate siglec clustering required for efficient ligand binding.

Recently, it has been shown that suppressor of cytokine signaling 3 (SOCS 3) and SHP-1/2 share a similar ligand-binding specificity and that it binds to the phosphorylated ITIM of CD33 [60] and Siglec-7 [61]. Following binding, SOCS 3 recruits a complex that acts as an E3 ligase and targets both the siglec and SOCS 3 itself for proteasomal degradation. Furthermore, Orr et al. could show that SOCS 3 blocked the inhibitory effect of hCD33 and Siglec-7 on cytokine-induced proliferation [60,61]. SOCS 3 seems to counteract hCD33 and Siglec-7 signaling in two ways, by competing with SHP-1/2 for binding sites and by degradation. SOCS proteins are negative regulators of cytokine signaling by inhibiting the JAK/STAT signal transduction pathway. Their basal levels are low, but they become induced by various cytokines and hormones in a wide variety of cells [62]. SOCS 3 has been associated with inflammatory diseases like inflammatory bowel disease, Crohn's disease, and rheumatoid arthritis and selective ablation of SOCS 3 in macrophages resulted in hyperactivation of STAT3 following IL-6 activation [63]. To date, too little it known about siglecs as well as SOCS 3 to predict the consequences of their interplay *in vivo*.

Finally, there is growing evidence that ITIM-bearing siglecs are internalized by endocytosis and that the tyrosine residues are crucial for this [64–68]. The consensus sequence YxxΦ (where Φ stands for amino acid with bulky hydrophobic side chain) required for interaction with the adaptor protein AP-2 is in accordance with the ITIM consensus sequence [69]. For siglecs, as with other inhibitory receptors, the degree of tyrosine phosphorylation is likely to be the critical factor in regulating which function predominates, as endocytosis might dominate in the absence of phosphorylation whereas high-level recruitment of SH2 domain-containing effectors like SHP-1, SHP-2, and SOCS 3 depend on tyrosine phosphorylation.

17.7 FUNCTIONS OF CD33-RELATED SIGLECS

CD33-related siglecs are expressed by cells of the hematopoietic system [24,29,64,70–75] with the exception of Siglec-6 (OB-BP1), which was found to be strongly expressed by cyto- and syncytiotropoblasts in the placenta [76]. Although it is difficult to assign direct orthologs between human and mouse, the CD33-related siglecs cover the same range of cells, with at least one or more being expressed on neutrophils, eosinophils, monocytes, B cells, macrophages, DCs, and pDCs (Figure 17.1). Resting T cells usually do not express any siglecs in mouse and humans, but do in chimpanzees and it has been argued that this may to some extent explain differing reactivities of T cells between great apes and humans [77]. Some progress has been made in understanding the biological function of these inhibitory receptors, but there are still more questions than answers. CD33 has long been of interest to hematologists, because of its expression on committed myeloid progenitors and absence from hematopoietic stem cells. It is also expressed on most acute myeloid leukemias (AML) cells and has been exploited as a target for antibody-based therapy via Myelotarg, a humanized anti-CD33 antibody coupled to the toxin calicheamicin-γ_1. Although cross-linking of hCD33 has been shown to inhibit proliferation and induce apoptosis of AML cells [78] (Figure 17.1B), the beneficial effect of Mylotarg is thought to depend on endocytic uptake of the antibody and release of the toxin to induce apoptosis [79]. More recently, other siglecs have been found on AML cells, with Siglec-9 being the most strongly expressed, both in terms of percentage and expression levels [65,68]. Since endocytosis upon antibody binding appears to be a common feature of CD33-related siglecs, targeting multiple siglecs for AML treatment might be a promising avenue.

Siglec-8 and -9 have both been implicated in mediating apoptosis in eosinophils and neutrophils (Figure 17.1B), respectively, upon antibody cross-linking [80–82]. They do not contain any death domains, but rather seem to mediate apoptosis via formation of reactive oxygen species (ROS), which trigger mitochondrial release of proapoptotic proteins, such as cytochrome C [83]. Surprisingly, in both cases the siglec-mediated death was enhanced in the presence of normally antiapoptotic cytokines, including IL-5 and GM-CSF for eosinophils [82] and GM-CSF and interferons for neutrophils [80]. Von Gunten et al. hypothesize that siglec-mediated cell death might play a role in resolving an inflammatory response or that siglec ligands might be expressed at high levels in some

organs like the brain to maintain an immune privileged environment [83]. The same authors have also shown that human intravenous immunoglobulin (IVIg) preparations, used in autoimmune disease treatment, contain anti-Siglec-9 autoantibodies, which can cause neutrophil death [84]. Whether these contribute to the beneficial or occasionally adverse effects observed during IVIg treatment requires more investigation.

Despite the differences in CD33-related siglecs in humans and mice, knockout mice will be important tools for studying siglec function. CD33-deficient mice were reported to have little phenotype [85], but mCD33 is unusual in lacking the membrane proximal ITIM [85] and being expressed on neutrophils in mice rather than on monocytes in humans. mSiglec-E is the likely ortholog of Siglec-9 expressed on myeloid cells and mSiglec-G is the ortholog of hSiglec-10 expressed by B cells. Analysis of Siglec-E- and -G-deficient mice is likely to reveal the regulatory roles of CD33-related siglecs expressed by these cell types. Similarly, Siglec-F-deficient mice are expected to provide insights into the functions of Siglec-8 on eosinophils, since these two siglecs appear to be functionally convergent paralogs [7].

17.8 SIGLEC-H AND OTHER DAP12-ASSOCIATED SIGLECS

The CD33-related Siglec-H is expressed on mouse plasmacytoid dendritic cells (pDC), a cell type associated with high secretion of type I interferons upon viral stimulation [86,87]. Although mSiglec-H has all the characteristic features of siglecs, including the arginine essential for sialic acid recognition, glycan binding could not be demonstrated for Siglec-H and interestingly, the crucial arginine residue is mutated in the rat ortholog. It therefore remains to be proven whether Siglec-H is strictly speaking a siglec or whether it has other, as yet uncharacterized, ligands. Siglec-H also lacks tyrosine-based signaling motifs in the cytoplasmic tail. Instead, it associates with the adaptor protein DAP12 via a charged lysine residue in its transmembrane region. Experiments with DAP12-deficient mice have shown that Siglec-H surface expression relies on the presence of DAP12 [87]. DAP12 is an activating molecule that contains cytoplasmic immunoreceptor tyrosine-based activation motifs (ITAMs). It is therefore surprising that antibody cross-linking of Siglec-H on pDCs inhibited IFN-α secretion in response to TLR9 stimulation [87,88]. Siglec-H also mediated endocytosis and cross-presentation of antigens [86] and might be involved in internalization and presentation of pathogens to endosomal TLR7 and 9 [88].

Certain other siglecs also carry a positively charged amino acid in their transmembrane region and associate with DAP12, namely chimpanzee Siglec-13 [87] and the newly discovered human Siglec-14 [75]. Siglec-14 is unique in that its first two domains are virtually identical to Siglec-5, while there is no similarity in the third domain and beyond. Since the mRNA expression patterns of Siglec-5 and Siglec-14 in tissues are very similar, this might be the first example for paired activating and inhibitory receptors within the siglec family, a feature that is commonly seen, however, within the Ig superfamily [89]. Supporting this hypothesis, partial sequence identity between Siglec-5 and -14 was also found in gorilla, orangutan, chimpanzee, and bonobo [75]. In the great apes, the essential arginine residue required for sialic acid recognition is mutated in both Siglec-5 and Siglec-14, suggesting that these receptors function in concert to balance activating and inhibitory signaling.

17.9 SIGLECS AND PATHOGENS

Sialic acid has emerged rather late in evolution, starting with the deuterostomes and is therefore normally absent in potential pathogens like bacteria and protozoa [5]. Some of these pathogens have, however, acquired the ability to present sialic acids on their surfaces as a form of disguise, in order to trick the innate immune system and prevent their recognition as “foreign” organisms [90]. In comparison, most mammalian viruses are normally sialylated as they are produced by the host’s translational machinery. It is possible that pathogens not only use sialic acids as a protection

shield, but also as a mediator for cell entry or immune response modulation by exploiting the sialic acid–binding siglecs. Sn, for example, serves as receptor for the PRRSV and mediates uptake into alveolar macrophages in a sialic acid–dependent manner [31,43]. Sn has further been shown to facilitate uptake of *N. meningitidis* [42] and *Trypanosoma cruzi* [91] by macrophages *in vitro*. There is also growing evidence that sialylated pathogens like *Campylobacter jejuni, N. meningitidis*, and group B *Streptococci* interact with CD33-related siglecs [42,92,93]. Apart from facilitated uptake, engagement of an inhibitory receptor might also be beneficial to the pathogen in dampening the immune response. Although studies are underway to investigate the role of siglecs during infection, *in vivo* evidence for either a beneficial or detrimental role in the host–pathogen interactions is missing so far.

17.10 CONCLUSION AND FUTURE DIRECTIONS

In terms of discovery, siglecs are a relatively young family of surface receptors and our current understanding of their biological functions is mostly based on *in vitro* experiments. The present challenge will be to transfer this knowledge to *in vivo* studies, in order to learn more about their role in regulating immune responses. The CD33-related siglecs represent one of the largest subgroups of inhibitory receptors in the immune system. Their recent characterization has led to a wider interest and siglecs are now studied not only by a small group of glycobiologist, but also by a growing number of immunologists. Most of the siglec knockout mice are now available and provide new tools for studying this group of proteins. We can therefore anticipate rapid developments in our knowledge of siglec functions in the coming years.

REFERENCES

1. Dulac, C. et al., Molecular characterization of the Schwann cell myelin protein, SMP: Structural similarities within the immunoglobulin superfamily, *Neuron*, 8, 323, 1992.
2. Lehmann, F. et al., Evolution of sialic acid-binding proteins: Molecular cloning and expression of fish siglec-4, *Glycobiology*, 14, 959, 2004.
3. Crocker, P. R. et al., Siglecs: A family of sialic-acid binding lectins, *Glycobiology*, 8, v, 1998.
4. Angata, T. et al., Large-scale sequencing of the CD33-related Siglec gene cluster in five mammalian species reveals rapid evolution by multiple mechanisms, *Proc Natl Acad Sci USA*, 101, 13251, 2004.
5. Angata, T. and Varki, A., Chemical diversity in the sialic acids and related alpha-keto acids: An evolutionary perspective, *Chem Rev*, 102, 439, 2002.
6. Attrill, H. et al., Siglec-7 undergoes a major conformational change when complexed with the alpha(2,8)-disialylganglioside GT1b, *J Biol Chem*, 281, 32774, 2006.
7. Tateno, H., Crocker, P. R., and Paulson, J. C., Mouse Siglec-F and human Siglec-8 are functionally convergent paralogs that are selectively expressed on eosinophils and recognize 6′-sulfo-sialyl Lewis X as a preferred glycan ligand, *Glycobiology*, 15, 1125, 2005.
8. Rapoport, E. M. et al., Probing sialic acid binding Ig-like lectins (siglecs) with sulfated oligosaccharides, *Biochemistry (Mosc)*, 71, 496, 2006.
9. Campanero-Rhodes, M. A. et al., Carbohydrate microarrays reveal sulphation as a modulator of siglec binding, *Biochem Biophys Res Commun*, 344, 1141, 2006.
10. May, A. P. et al., Crystal structure of the N-terminal domain of sialoadhesin in complex with 3′ sialyllactose at 1.85 Å resolution, *Mol Cell*, 1, 719, 1998.
11. Alphey, M. S. et al., High resolution crystal structures of Siglec-7. Insights into ligand specificity in the Siglec family, *J Biol Chem*, 278, 3372, 2003.
12. Vinson, M. et al., Characterization of the sialic acid-binding site in sialoadhesin by site-directed mutagenesis, *J Biol Chem*, 271, 9267, 1996.
13. Crocker, P. R. and Feizi, T., Carbohydrate recognition systems: Functional triads in cell–cell interactions, *Curr Opin Struct Biol*, 6, 679, 1996.
14. van den Berg, T. K. et al., Cutting edge: CD43 functions as a T cell counterreceptor for the macrophage adhesion receptor sialoadhesin (Siglec-1), *J Immunol*, 166, 3637, 2001.

15. Nath, D. et al., Macrophage–tumour cell interactions: Identification of MUC1 on breast cancer cells as a potential counter-receptor for the macrophage-restricted receptor, sialoadhesin, *Immunology*, 98, 213, 1999.
16. Stamenkovic, I. et al., The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and alpha 2–6 sialyltransferase, CD75, on B cells, *Cell*, 66, 1133, 1991.
17. Leprince, C. et al., CD22 associates with the human surface IgM-B-cell antigen receptor complex, *Proc Natl Acad Sci USA*, 90, 3236, 1993.
18. Pflugh, D. L., Maher, S. E., and Bothwell, A. L., Ly-6 superfamily members Ly-6A/E, Ly-6C, and Ly-6I recognize two potential ligands expressed by B lymphocytes, *J Immunol*, 169, 5130, 2002.
19. Han, S. et al., Homomultimeric complexes of CD22 in B cells revealed by protein–glycan cross-linking, *Nat Chem Biol*, 1, 93, 2005.
20. Nicoll, G. et al., Ganglioside GD3 expression on target cells can modulate NK cell cytotoxicity via siglec-7-dependent and -independent mechanisms, *Eur J Immunol*, 33, 1642, 2003.
21. Razi, N. and Varki, A., Masking and unmasking of the sialic acid-binding lectin activity of CD22 (Siglec-2) on B lymphocytes, *Proc Natl Acad Sci USA*, 95, 7469, 1998.
22. Collins, B. E. et al., High-affinity ligand probes of CD22 overcome the threshold set by *cis* ligands to allow for binding, endocytosis, and killing of B cells, *J Immunol*, 177, 2994, 2006.
23. Collins, B. E. et al., Masking of CD22 by cis ligands does not prevent redistribution of CD22 to sites of cell contact, *Proc Natl Acad Sci USA*, 101, 6104, 2004.
24. Falco, M. et al., Identification and molecular cloning of p75/AIRM1, a novel member of the sialoadhesin family that functions as an inhibitory receptor in human natural killer cells, *J Exp Med*, 190, 793, 1999.
25. Avril, T. et al., Probing the *cis* interactions of the inhibitory receptor Siglec-7 with alpha2,8-disialylated ligands on natural killer cells and other leukocytes using glycan-specific antibodies and by analysis of alpha2,8-sialyltransferase gene expression, *J Leukoc Biol*, 80, 787, 2006.
26. Lanoue, A. et al., Interaction of CD22 with alpha2,6-linked sialoglycoconjugates: Innate recognition of self to dampen B cell autoreactivity? *Eur J Immunol*, 32, 348, 2002.
27. Hennet, T. et al., Immune regulation by the ST6Gal sialyltransferase, *Proc Natl Acad Sci USA*, 95, 4504, 1998.
28. Crocker, P. R. and Gordon, S., Properties and distribution of a lectin-like hemagglutinin differentially expressed by murine stromal tissue macrophages, *J Exp Med*, 164, 1862, 1986.
29. Munday, J., Floyd, H., and Crocker, P. R., Sialic acid binding receptors (siglecs) expressed by macrophages, *J Leukoc Biol*, 66, 705, 1999.
30. Crocker, P. R., Siglecs: Sialic-acid-binding immunoglobulin-like lectins in cell–cell interactions and signalling, *Curr Opin Struct Biol*, 12, 609, 2002.
31. Vanderheijden, N. et al., Involvement of sialoadhesin in entry of porcine reproductive and respiratory syndrome virus into porcine alveolar macrophages, *J Virol*, 77, 8207, 2003.
32. Hartnell, A. et al., Characterization of human sialoadhesin, a sialic acid binding receptor expressed by resident and inflammatory macrophage populations, *Blood*, 97, 288, 2001.
33. Bilyk, N. and Holt, P. G., The surface phenotypic characterization of lung macrophages in C3H/HeJ mice, *Immunology*, 74, 645, 1991.
34. Crocker, P. R. and Gordon, S., Mouse macrophage hemagglutinin (sheep erythrocyte receptor) with specificity for sialylated glycoconjugates characterized by a monoclonal antibody, *J Exp Med*, 169, 1333, 1989.
35. Jiang, H. R., Lumsden, L., and Forrester, J. V., Macrophages and dendritic cells in IRBP-induced experimental autoimmune uveoretinitis in B10RIII mice, *Invest Ophthalmol Vis Sci*, 40, 3177, 1999.
36. Ikezumi, Y. et al., The sialoadhesin (CD169) expressing a macrophage subset in human proliferative glomerulonephritis, *Nephrol Dial Transplant*, 20, 2704, 2005.
37. Kobsar, I. et al., Attenuated demyelination in the absence of the macrophage-restricted adhesion molecule sialoadhesin (Siglec-1) in mice heterozygously deficient in P0, *Mol Cell Neurosci*, 31, 685, 2006.
38. Ip, C. W. et al., Immune cells contribute to myelin degeneration and axonopathic changes in mice overexpressing proteolipid protein in oligodendrocytes, *J Neurosci*, 26, 8206, 2006.
39. Muerkoster, S. et al., Sialoadhesin-positive host macrophages play an essential role in graft-versus-leukemia reactivity in mice, *Blood*, 93, 4375, 1999.
40. Pulliam, L., Sun, B., and Rempel, H., Invasive chronic inflammatory monocyte phenotype in subjects with high HIV-1 viral load, *J Neuroimmunol*, 157, 93, 2004.
41. Kirchberger, S. et al., Human rhinoviruses inhibit the accessory function of dendritic cells by inducing sialoadhesin and B7-H1 expression, *J Immunol*, 175, 1145, 2005.
42. Jones, C., Virji, M., and Crocker, P. R., Recognition of sialylated meningococcal lipopolysaccharide by siglecs expressed on myeloid cells leads to enhanced bacterial uptake, *Mol Microbiol*, 49, 1213, 2003.

43. Delputte, P. L. and Nauwynck, H. J., Porcine arterivirus infection of alveolar macrophages is mediated by sialic acid on the virus, *J Virol*, 78, 8094, 2004.
44. Kumamoto, Y. et al., Identification of sialoadhesin as a dominant lymph node counter-receptor for mouse macrophage galactose-type C-type lectin 1, *J Biol Chem*, 279, 49274, 2004.
45. Martinez-Pomares, L. et al., Cell-specific glycoforms of sialoadhesin and CD45 are counter-receptors for the cysteine-rich domain of the mannose receptor, *J Biol Chem*, 274, 35211, 1999.
46. Crocker, P. R. et al., Ultrastructural localization of a macrophage-restricted sialic acid binding hemagglutinin, SER, in macrophage-hematopoietic cell clusters, *Blood*, 76, 1131, 1990.
47. Oetke, C. et al., Sialoadhesin-deficient mice exhibit subtle changes in B- and T-cell populations and reduced immunoglobulin M levels, *Mol Cell Biol*, 26, 1549, 2006.
48. Jiang, H. R. et al., Sialoadhesin promotes the inflammatory response in experimental autoimmune uveoretinitis, *J Immunol*, 177, 2258, 2006.
49. Pozzi, L. A., Maciaszek, J. W., and Rock, K. L., Both dendritic cells and macrophages can stimulate naive CD8 T cells in vivo to proliferate, develop effector function, and differentiate into memory cells, *J Immunol*, 175, 2071, 2005.
50. Kelley, J., Walter, L., and Trowsdale, J., Comparative genomics of natural killer cell receptor gene clusters, *PLoS Genet*, 1, 129, 2005.
51. Avril, T. et al., The membrane-proximal immunoreceptor tyrosine-based inhibitory motif is critical for the inhibitory signaling mediated by Siglecs-7 and -9, CD33-related Siglecs expressed on human monocytes and NK cells, *J Immunol*, 173, 6841, 2004.
52. Yu, Z. et al., mSiglec-E, a novel mouse CD33-related siglec (sialic acid-binding immunoglobulin-like lectin) that recruits Src homology 2 (SH2)-domain-containing protein tyrosine phosphatases SHP-1 and SHP-2, *Biochem J*, 353, 483, 2001.
53. Ulyanova, T., Shah, D. D., and Thomas, M. L., Molecular cloning of MIS, a myeloid inhibitory siglec, that binds protein-tyrosine phosphatases SHP-1 and SHP-2, *J Biol Chem*, 276, 14451, 2001.
54. Paul, S. P. et al., Myeloid specific human CD33 is an inhibitory receptor with differential ITIM function in recruiting the phosphatases SHP-1 and SHP-2, *Blood*, 96, 483, 2000.
55. Avril, T. et al., Siglec-5 (CD170) can mediate inhibitory signaling in the absence of immunoreceptor tyrosine-based inhibitory motif phosphorylation, *J Biol Chem*, 280, 19843, 2005.
56. Barrow, A. D. and Trowsdale, J., You say ITAM and I say ITIM, let's call the whole thing off: The ambiguity of immunoreceptor signalling, *Eur J Immunol*, 36, 1646, 2006.
57. Taylor, V. C. et al., The myeloid-specific sialic acid-binding receptor, CD33, associates with the protein–tyrosine phosphatases, SHP-1 and SHP-2, *J Biol Chem*, 274, 11505, 1999.
58. Veillette, A., Immune regulation by SLAM family receptors and SAP-related adaptors, *Nat Rev Immunol*, 6, 56, 2006.
59. Schoenwaelder, S. M. et al., The protein tyrosine phosphatase Shp-2 regulates RhoA activity, *Curr Biol*, 10, 1523, 2000.
60. Orr, S. J. et al., CD33 responses are blocked by SOCS3 through accelerated proteasomal-mediated turnover, *Blood*, 282, 3418, 2006.
61. Orr, S. J. et al., SOCS3 targets Siglec 7 for proteasomal degradation and blocks Siglec 7-mediated responses, *J Biol Chem*, 2006.
62. Fujimoto, M. and Naka, T., Regulation of cytokine signaling by SOCS family molecules, *Trends Immunol*, 24, 659, 2003.
63. Yasukawa, H. et al., IL-6 induces an anti-inflammatory response in the absence of SOCS3 in macrophages, *Nat Immunol*, 4, 551, 2003.
64. Lock, K. et al., Expression of CD33-related siglecs on human mononuclear phagocytes, monocyte-derived dendritic cells and plasmacytoid dendritic cells, *Immunobiology*, 209, 199, 2004.
65. Biedermann, B. et al., Analysis of the CD33-related siglec family reveals that Siglec-9 is an endocytic receptor expressed on subsets of acute myeloid leukemia cells and absent from normal hematopoietic progenitors, *Leuk Res*, 3, 211, 2006.
66. Walter, R. B. et al., Influence of CD33 expression levels and ITIM-dependent internalization on gemtuzumab ozogamicin-induced cytotoxicity, *Blood*, 105, 1295, 2005.
67. John, B. et al., The B cell coreceptor CD22 associates with AP50, a clathrin-coated pit adapter protein, via tyrosine-dependent interaction, *J Immunol*, 170, 3534, 2003.
68. Nguyen, D. H., Ball, E. D., and Varki, A., Myeloid precursors and acute myeloid leukemia cells express multiple CD33-related Siglecs, *Exp Hematol*, 34, 728, 2006.
69. Bonifacino, J. S. and Traub, L. M., Signals for sorting of transmembrane proteins to endosomes and lysosomes, *Annu Rev Biochem*, 72, 395, 2003.

70. Zhang, J. Q. et al., The murine inhibitory receptor mSiglec-E is expressed broadly on cells of the innate immune system whereas mSiglec-F is restricted to eosinophils, *Eur J Immunol*, 34, 1175, 2004.
71. Zhang, J. Q. et al., Siglec-9, a novel sialic acid binding member of the immunoglobulin superfamily expressed broadly on human blood leukocytes, *J Biol Chem*, 275, 22121, 2000.
72. Angata, T. and Varki, A., Cloning, characterization, and phylogenetic analysis of siglec-9, a new member of the CD33-related group of siglecs. Evidence for co-evolution with sialic acid synthesis pathways, *J Biol Chem*, 275, 22127, 2000.
73. Angata, T. and Varki, A., Siglec-7: A sialic acid-binding lectin of the immunoglobulin superfamily, *Glycobiology*, 10, 431, 2000.
74. Angata, T. et al., Cloning and characterization of human Siglec-11. A recently evolved signaling that can interact with SHP-1 and SHP-2 and is expressed by tissue macrophages, including brain microglia, *J Biol Chem*, 277, 24466, 2002.
75. Angata, T. et al., Discovery of Siglec-14, a novel sialic acid receptor undergoing concerted evolution with Siglec-5 in primates, *FASEB J*, 20, 1964, 2006.
76. Patel, N. et al., OB-BP1/Siglec-6, a leptin- and sialic acid-binding protein of the immunoglobulin superfamily, *J Biol Chem*, 274, 22729, 1999.
77. Nguyen, D. H. et al., Loss of Siglec expression on T lymphocytes during human evolution, *Proc Natl Acad Sci USA*, 103, 7765, 2006.
78. Vitale, C. et al., Surface expression and function of p75/AIRM-1 or CD33 in acute myeloid leukemias: Engagement of CD33 induces apoptosis of leukemic cells, *Proc Natl Acad Sci USA*, 98, 5764, 2001.
79. Linenberger, M. L., CD33-directed therapy with gemtuzumab ozogamicin in acute myeloid leukemia: Progress in understanding cytotoxicity and potential mechanisms of drug resistance, *Leukemia*, 19, 176, 2005.
80. von Gunten, S. et al., Siglec-9 transduces apoptotic and nonapoptotic death signals into neutrophils depending on the proinflammatory cytokine environment, *Blood*, 106, 1423, 2005.
81. Nutku, E., Hudson, S. A., and Bochner, B. S., Mechanism of Siglec-8-induced human eosinophil apoptosis: Role of caspases and mitochondrial injury, *Biochem Biophys Res Commun*, 336, 918, 2005.
82. Nutku, E. et al., Ligation of Siglec-8: A selective mechanism for induction of human eosinophil apoptosis, *Blood*, 101, 5014, 2003.
83. von Gunten, S. and Simon, H. U., Sialic acid binding immunoglobulin-like lectins may regulate innate immune responses by modulating the life span of granulocytes, *Faseb J*, 20, 601, 2006.
84. von Gunten, S. et al., Immunologic and functional evidence for anti-Siglec-9 autoantibodies in intravenous immunoglobulin preparations, *Blood*, 108, 4255, 2006.
85. Brinkman-Van der Linden, E. C. et al., CD33/Siglec-3 binding specificity, expression pattern, and consequences of gene deletion in mice, *Mol Cell Biol*, 23, 4199, 2003.
86. Zhang, J. et al., Characterization of Siglec-H as a novel endocytic receptor expressed on murine plasmacytoid dendritic cell precursors, *Blood*, 107, 3600, 2006.
87. Blasius, A. L. et al., Siglec-H is an IPC-specific receptor that modulates type I IFN secretion through DAP12, *Blood*, 107, 2474, 2006.
88. Blasius, A. L. and Colonna, M., Sampling and signaling in plasmacytoid dendritic cells: The potential roles of Siglec-H, *Trends Immunol*, 27, 255, 2006.
89. Takai, T. and Ono, M., Activating and inhibitory nature of the murine paired immunoglobulin-like receptor family, *Immunol Rev*, 181, 215, 2001.
90. Vimr, E. R. et al., Diversity of microbial sialic acid metabolism, *Microbiol Mol Biol Rev*, 68, 132, 2004.
91. Monteiro, V. G. et al., Increased association of *Trypanosoma cruzi* with sialoadhesin positive mice macrophages, *Parasitol Res*, 97, 380, 2005.
92. Carlin, A. F. et al., Group B streptococcal capsular sialic acids interact with siglecs (immunoglobulin-like lectins) on human leukocytes, *J Bacteriol*, 189, 1231, 2006.
93. Avril, T. et al., Sialic acid-binding immunoglobulin-like lectin 7 mediates selective recognition of sialylated glycans expressed on *Campylobacter jejuni* lipooligosaccharides, *Infect Immun*, 74, 4133, 2006.

18 Signaling through the Fungal β-Glucan Receptor Dectin-1

Ann M. Kerrigan, Kevin M. Dennehy, and Gordon D. Brown

CONTENTS

18.1 INTRODUCTION

The innate immune response of vertebrates is triggered if a pathogen breaches the physical barriers of the host. It provides a rapid but nonspecific response and is the first line of defense. The functions of innate immunity are based on a variety of germline-encoded receptors, the pattern recognition receptors (PRRs) [1]. PRRs are proteins expressed by cells of the innate immune system that recognize conserved molecules associated with microbial pathogens. These conserved microbial structures are known as pathogen-associated molecular patterns (PAMPs). The recognition of PAMPs enables the immune system to distinguish foreign organisms from cells of the host and initiates a rapid inflammatory response.

Toll-like receptors (TLRs) are a family of PRRs involved in the recognition of a broad range of microbial structures and they are the major molecules through which mammals sense infection and activate inflammatory responses. For example, TLR2 plays an important role in recognition and responses to fungal pathogens [2,3] and TLR4 is critical for detection and clearance of Gram-negative bacteria [4]. There are also several other non-TLR receptors that have roles in microbial detection, including complement receptors, pentraxins, scavenger receptors, classical and nonclassical C-type lectins and collectins. In general, however, the exact function of the non-TLR PRRs in the control of infection remains unclear.

An exception to this is the nonclassical C-type lectin receptor, Dectin-1. Dectin-1 is involved in the innate immune response to fungal pathogens and has been shown to bind many fungal species, including *Saccharomyces* [2], *Candida* [5], *Coccidioides* [6], *Pneumocystis* [7,8], and *Aspergillus* [9–11]. It recognizes β-glucan components of the fungal cell wall [12] and *in vitro* has been shown

to mediate fungal binding, uptake, and killing [13,14], and the production of numerous cytokines and chemokines [2,7,15,16]. Studies using Dectin-1-deficient mice have established the essential role of β-glucan recognition in the development of an antifungal inflammatory response and the control of fungal dissemination *in vivo* [5,8]. In fact, to date, Dectin-1 is the only example of a non-TLR PRR that can mediate its own intracellular signaling to induce a protective immune response. This chapter will focus on independent and cooperative signaling by Dectin-1 and how the current knowledge fits into the overall paradigm of immunoreceptor signaling and pattern recognition.

18.2 DECTIN-1: STRUCTURE, EXPRESSION, AND LIGANDS

Dectin-1 was originally identified as a dendritic cell specific receptor that modulates T cell function by recognizing an unidentified ligand on lymphocytes [17,18]. It was reidentified as a β-glucan receptor following a screen of a murine macrophage cDNA expression library with zymosan, a β-glucan-rich extract of *Saccharomyces cerevisiae* [12]. β-Glucans are carbohydrate polymers with well-characterized immunostimulatory activity that are found primarily in the cell walls of fungi, but also in plants and some bacteria [19]. Other receptors such as complement receptor 3 (CR3) have been implicated in β-glucan and fungal recognition [20]. However, studies using specific antagonists, blocking monoclonal antibodies, and knockout mice have clearly shown that Dectin-1 is the major β-glucan receptor on primary macrophages, dendritic cells, and neutrophils [5,8,21].

Dectin-1 is a member of the nonclassical C-type lectin family. It is a type II transmembrane protein consisting of a single extracellular lectin-like domain connected to a single-pass transmembrane region by a stalk and a cytoplasmic tail, which contains an immunoreceptor tyrosine-based activation (ITAM)-like motif (Figure 18.1). It has a similar structure to other members of the nonclassical C-type lectin family, with two significant exceptions. Firstly, there are no cysteine residues in its stalk region, indicating that it does not dimerize. Secondly, the cytoplasmic ITAM-like motif is unusual as most other known activating nonclassical C-type lectin receptors, such as Ly49D (lymphocyte antigen 49D), do not contain cytoplasmic ITAM-like motifs, but rather associate with signaling molecules, such as DAP12, to carry out cellular activation. However, C-type lectin-like receptor 2 (CLEC-2), another

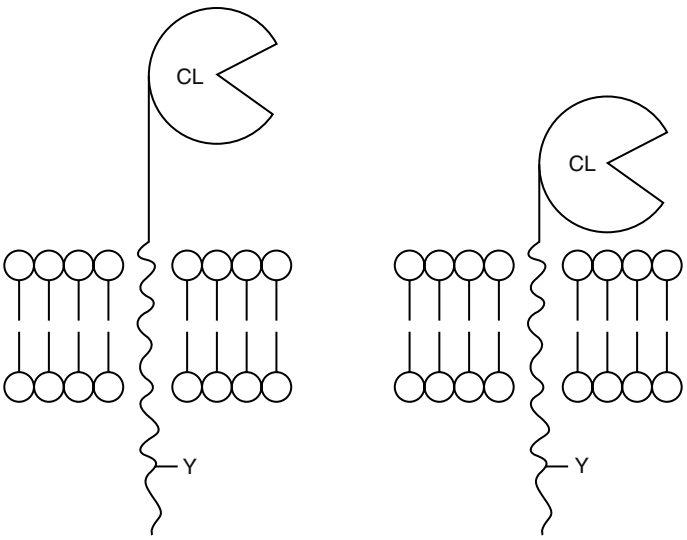


FIGURE 18.1 Structure of the β-glucan receptor Dectin-1. Dectin-1 contains an extracellular C-type lectin-like carbohydrate recognition domain (CL), stalk region, transmembrane domain, and cytoplasmic tail that contains an ITAM-like sequence. Alternative splicing can generate an isoform of Dectin-1 that lacks a stalk.

receptor that is encoded within the same gene cluster as Dectin-1, does have a similar ITAM-like motif to Dectin-1, and has also been shown to induce cellular activation [22].

Oligosaccharide microarray technology has shown that Dectin-1 specifically binds β-1,3-linked glucose oligomers [23]. Other members of the nonclassical C-type lectins do not generally recognize carbohydrates, as they lack the conserved residues involved in the ligation of calcium that coordinates carbohydrate binding. Dectin-1 also lacks these conserved residues and its Ca^{2+}-independent mechanism for recognition of carbohydrates is still unclear. However, a study, which involved binding analysis of 32 point mutants, revealed that at least two residues, Trp221 and His223 in the extracellular region of Dectin-1, are crucial for β-glucan binding [24].

The murine form of the receptor is expressed by many cell types, including macrophages, dendritic cells, monocytes, neutrophils, and a subset of splenic T cells [25]. There are at least two isoforms of murine Dectin-1, encoding the full-length receptor and a stalkless version. Studies have shown that zymosan binding by the smaller isoform is significantly weakened at lower temperatures and there are also differences in the levels of cytokine production induced by the two isoforms. These studies indicate that Dectin-1 isoforms are functionally distinct and may represent a means whereby cellular responses to fungal infection can be regulated [26]. Structurally and functionally similar human homologs have been identified. The expression of human Dectin-1 differs slightly in that it is also expressed on B cells and eosinophils. Human Dectin-1 mRNA is also alternatively spliced, resulting in eight isoforms, of which only two are functional for β-glucan binding [27]. Of the two functional human isoforms, the most predominant lacks the stalk region [27] (Figure 18.1), and similar to the mouse, there are also temperature-dependent differences in zymosan binding [26]. The two major human isoforms are expressed differently in various cell types although the significance of this remains unclear. It has also been shown that the level of Dectin-1 expression in murine macrophages can be influenced by various cytokines and microbial factors. For example IL-4, IL-13, and GM-CSF cause Dectin-1 expression to be highly upregulated. In contrast IL-10, LPS and dexamethasone cause down-regulation of Dectin-1 expression [28].

18.3 SIGNALING THROUGH ITAM-CONTAINING RECEPTORS

Studies of molecules containing ITAM motifs have contributed largely to the elucidation of patterns of immunoreceptor signaling and are helpful when considering signaling through the ITAM-like motif of Dectin-1. The T cell receptor (TCR)–CD3 complex, the B cell receptor and Ig Fc receptors (FcR) all signal via src and Syk family tyrosine kinases, phosphatases, and a host of adapter proteins that serve as molecular links to downstream signaling pathways.

The activatory Fcγ receptors are characterized by the presence of an ITAM motif in an accessory chain that associates with the receptor. When Fc receptors interact with their immunoglobulin ligands, the extracellular domains of the receptors aggregate. This results in tyrosine phosphorylation of the ITAM motif by members of the Src-kinase family and subsequent recruitment of SH2-containing kinases including members of the Syk-kinase family. These early events trigger the initiation of downstream signaling cascades that ultimately lead to cellular activation [29]. Such events include phosphorylation and activation of phospholipase C (PLC), which leads to Ca^{2+} flux and activation of nuclear factor of activated T cells (NFAT) transcription factor. In addition, activation of protein kinase C (PKC) isoforms couple ITAM-containing receptors to the caspase recruitment domain (CARD) proteins such as CARD 11. Such proteins form complexes with the adaptors MALT1 and Bcl10, leading to the recruitment and activation of TNF receptor associated factor 6 (TRAF6), which in turn couples to inhibitor of KappaB kinases (IKK). Activation of the IKK complex results in phosphorylation of an inhibitory protein IκB. IκB is bound in a complex with inactive NFκB. Once IκB is phosphorylated, it dissociates from NFκB, allowing the transcription factor to translocate to the nucleus and induce various genes [30]. The TLRs also activate NFκB, but utilize separate pathways either through the adaptor MyD88 or Toll/IL-1R domain-containing adapter inducing IFN-beta (TRIF).

The functional TCR complex is equipped with 10 ITAMs that are phosphorylated by Lck, a Src family kinase, on receptor engagement. This allows recruitment of ZAP70, a Syk family kinase, and phosphorylation of the adaptor molecules LAT and SLP-76, as well as the enzymes Vav and PLCγ1 [31]. PKCθ is subsequently recruited to the membrane and activated. This event appears to be a specific link to the NKκB pathway. PKCθ activates CARD 11 by phosphorylating specific serine residues [32]. This in turn leads to recruitment of Bcl10 and MALT1 and activation of NFκB.

The traditional ITAM motif is characterized by a consensus sequence that includes two tyrosines, usually 10–12 amino acids apart: $YxxI/Lx_{(6-12)}YxxI/L$. The signaling motif of Dectin-1 is only slightly different from other activatory receptors and consists of $YxxxI/Lx_7YxxL/I$. Given this likeness, it was originally thought that Dectin-1 may signal in a similar manner to ITAM-containing receptors such as those described above. Studies have since shown that while parallels can indeed be drawn between signaling by Dectin-1 and other activatory receptors, such as FcγR and TCR, there are also notable differences (Figure 18.2). As previously mentioned, *in vitro* studies have shown that ligand binding by Dectin-1 can induce many cellular responses, such as phagocytosis, phospholipase A2, COX2, the respiratory burst and cytokine production. These various aspects of Dectin-1 function will be discussed in more detail in the subsequent sections. Signaling from Dectin-1 alone appears to be sufficient for many of these responses, but others, such as the induction of TNF, require cooperation through TLR2 and signaling via MyD88 [2,15,16]. While many activatory receptors are known to associate with Syk kinase, Dectin-1 is presently the sole non-TLR PRR shown to signal via Syk. Furthermore, the association between Dectin-1 and Syk occurs via a novel interaction involving only one tyrosine [15]. However, some cellular responses induced by Dectin-1, such as phagocytosis, do not require Syk. The presence of TLR and Syk-dependent and -independent pathways highlights the complexity of Dectin-1 signaling. The elucidation of the signaling mechanisms of Dectin-1 has potential implications for other non-TLR PRRs that have similar cytoplasmic motifs [33].

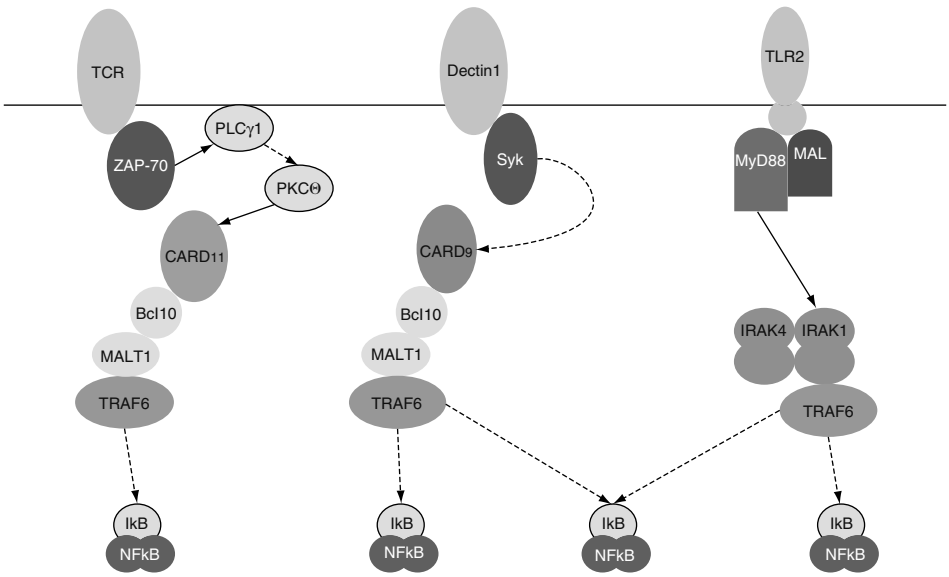


FIGURE 18.2 (See CD for color figure.) Comparison of NFκB activation pathways through Dectin-1, TLR2, and the TCR. Dectin-1 and the TCR use Syk family kinases to couple to the CARD9 and CARD11 adaptors, respectively, leading to assembly of the TRAF6 complex, degradation of IκB and NFκB activation. TLR2 utilizes the adaptors MyD88 and MAL, and IRAK kinases to couple to the TRAF6 complex and NFκB activation. Dectin-1 and TLR2 collaborative responses may occur by signal integration. Dashed lines represent pathways with multiple steps.

18.4 PHAGOCYTOSIS

Dectin-1 triggers cytoskeletal rearrangements leading to particle internalization. Figure 18.3 illustrates a Dectin-1 expressing cell internalizing *Candida albicans*. This phagocytic activity of Dectin-1 is dependent on the ITAM-like motif in the cytoplasmic tail [13]. As previously mentioned, the signaling motif of Dectin-1 differs only slightly from the traditional ITAM motif of activatory receptors. However, this difference appears to be sufficient to warrant an alternative mechanism of internalization. It is widely accepted that the Src-ITAM-Syk signaling pathway results in phagocytosis by FcγR [34]. Studies using macrophages deficient in various Src kinases suggest that there is a high level of redundancy for these enzymes and that all of them are required during activation of phagocytosis by FcγR [29,35,36]. The phagocytic capacity of Dectin-1 depends only on its membrane proximal tyrosine residue within the ITAM-like motif [13]. Dectin-1 becomes tyrosine phosphorylated at this residue upon ligand binding. However, in contrast to FcγR, src kinases appear to be only partially responsible, as was shown by using PP2, a src kinase inhibitor whose inclusion in internalization assays did not cause complete inhibition of uptake [13]. This suggests that other kinases may be involved. Furthermore, in macrophages, Dectin-1-mediated phagocytosis does not require Syk. This indicates the existence of a novel signaling pathway for Dectin-1-mediated phagocytosis in macrophages [13,15]. It has also been shown that phosphoinositide-3 (PI-3) kinase is not essential for Dectin-1-mediated phagocytosis, whereas PKC, ras-related C3 botulinum toxin substrate 1 (Rac-1), and Cdc42 are required [13]. On the other hand, Syk does have some influence on Dectin-1-mediated phagocytosis in dendritic cells where the kinase is recruited to the membrane proximal tyrosine of Dectin-1 [16].

Dectin-1 is also involved in the uptake of opsonized zymosan. Pentraxin-3 (PTX-3) binds to zymosan and it has been shown that internalization of PTX-3 opsonized zymosan is dependent on Dectin-1 [37]. Macrophages actually express higher levels of PTX3 mRNA in the presence of zymosan and it has been suggested that during fungal infection, the PTX3 gene is activated and secreted PTX3 may enhance the clearance of the pathogen [37]. Dectin-1 can also recognize exposed

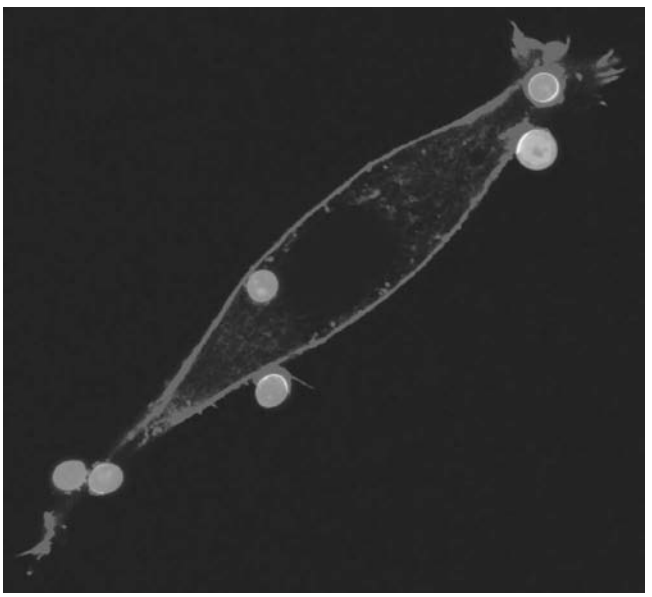


FIGURE 18.3 (see color insert following blank page 170. Also see CD for color figure.) Dectin-1 mediates phagocytosis of *Candida albicans.* A NIH-3T3 cell expressing Dectin-1 phagocytosing FITC-labeled *C. albicans.* Actin filaments and phagocytic cups are stained red. (From Brown, G.D. and Gordon, S., *Nature*, 413, 36, 2001. With permission.)

β-glucans on complement opsonized zymosan and it has been shown that Dectin-1 is necessary for inflammatory responses to opsonized particles [38].

Other molecules may also associate with Dectin-1 during internalization of yeast. For example, it has been reported that the tetraspanin CD63 is internalized after yeast phagocytosis, and immunoprecipitation experiments have shown that it associates with Dectin-1 [39]. Although the functional significance of a Dectin-1-CD63 interaction has not yet been elucidated, it may represent part of a signaling complex. A functional interaction between another tetraspanin, CD37, and Dectin-1 has also been reported. This study demonstrated that CD37 inhibits signaling pathways leading to Dectin-1-mediated cytokine production [40]. As cells of the immune system express several PRRs, signals from the various receptors must be integrated to produce an appropriate immune response. This can be achieved by clustering immunoreceptors and signaling molecules into complexes using tetraspanins. Tetraspanins can laterally interact with immunoreceptors to create "tetraspanin microdomains" and consequently regulate downstream signaling [41–43]. The Dectin-1-CD63 or CD37 complexes are therefore quite likely to form part of a larger complex that integrates signaling from other receptors [40].

18.5 CYTOSOLIC PHOSPHOLIPASE A2 ACTIVATION AND CYCLOOXYGENASE 2 EXPRESSION

Zymosan has long been used to investigate arachidonic acid release and metabolism [44,45]. Arachidonic acid is released from cell membranes by cytosolic phospholipase A2 ($cPLA_2$), after which it is available for the production of the inflammatory mediators, leukotrienes, and prostaglandins. Metabolism of free arachidonic acid by the 5-lipoxygenase (5-LO) pathway, or the cyclooxygnase (COX) pathway, results in the production of leukotrienes or prostaglandins, respectively. Recent work has shown that Dectin-1 has a role in $cPLA_2$ activation by zymosan and *C. albicans* [46,47]. Ligation of Dectin-1 by zymosan and *C. albicans* also increases expression of COX2 and consequently prostaglandin production. The cytoplasmic tail of Dectin-1 is required for zymosan-induced arachidonic acid release although internalization is not strictly necessary. Arachidonic acid release and leukotriene production stimulated through Dectin-1 ligation is TLR2-independent, but requires Syk, Src kinases, and PI3 kinase [46,47]. In contrast, COX2 expression and consequent prostaglandin production are Syk-dependent and partially TLR2-dependent [46].

18.6 RESPIRATORY BURST

The signaling cascade initiated by FcγR-mediated phagocytosis results in activation of the respiratory burst. The respiratory burst involves production of reactive oxygen species (ROS) that combat invading microorganisms, and it is initiated by the NADPH oxidase enzyme. FcR signaling in neutrophils causes migration of the various subunits of NADPH oxidase to the phagosome membranes and subsequent activation of the complex [48]. Various protein kinases have been reported to activate NADPH oxidase, but the initial signaling events take place along the Src-ITAM-Syk pathway [29]. The generation of intracellular signals by Dectin-1 following ligand binding also result in the activation of NADPH oxidase and production of ROS [3,15]. It also seems that Dectin-1 can cooperate in some way with TLRs in the production of ROS. For example, ROS production induced by zymosan can be enhanced when cells are prestimulated with TLR ligands, indicating that the Dectin-1-triggered response can be primed by TLR activation [15]. There is also a requirement for Syk in ROS production by macrophages following Dectin-1 ligation [14]. However, this respiratory burst does not occur in all macrophages. Evidence suggests that macrophages can exist in two states, one that is permissive to Syk activation and ROS production and one that is not. These cells that initiate a respiratory burst represent a specialized subpopulation of macrophages that are somehow primed for production of reactive oxygen intermediates [14].

18.7 CYTOKINES

Activation of Dectin-1 results in the production of numerous chemokines and cytokines, such as TNF, IL-10, IL-2, IL-12, IL-6, IL-1α, IL-1β, MCP-1, and MIP1α [2,5,9,15,49]. The most well studied in relation to Dectin-1 induced production are described below. Cytokine production and other responses induced by Dectin-1 are summarized in Figure 18.4.

18.7.1 TNF

Tumor necrosis factor (TNF) is an essential proinflammatory cytokine required for the successful control of many fungal pathogens [50–53]. Therefore, it is not surprising that Dectin-1 ligation by zymosan and live fungal pathogens induces TNF production [2,15]. In fact, Dectin-1 is the first non-TLR receptor identified as being involved in the generation of a proinflammatory response to fungal pathogens. Studies using cell lines expressing a truncated version of Dectin-1 lacking the cytoplasmic tail demonstrated that the Dectin-1 cytoplasmic tail is required to induce TNF production in response to zymosan [2,15]. However, particle internalization is not required, as cell lines expressing Dectin-1 still show elevated levels of TNF production when phagocytosis is inhibited [2]. TNF production induced by Dectin-1 also requires the recognition of another component of zymosan by TLR2 and signaling through the MyD88 pathway [2,15]. It is well established that stimulation of TLRs leads to activation of NFκB and production of proinflammatory cytokines such as TNF. However, the specificity of these responses is not clearly understood and it is believed that other immunoreceptors contribute to their induction. This appears to be the case for TLR2 and Dectin-1, as the simultaneous engagement of both these receptors greatly enhances TLR2-mediated TNF production [2,15]. Gross et al. have recently shown that bone marrow derived dendritic cells (BMDCs) from CARD 9-deficient mice displayed severely defective TNF production induced by zymosan [54]. CARD 9 is a CARD protein that is related to CARD 11 which is mentioned previously. In the same study, inclusion of specific inhibitors of Syk abolished zymosan-induced production of TNF by wild-type BMDCs. This work shows that zymosan-induced production of TNF is highly dependent on Syk and CARD 9. Gross et al. also provide evidence to support the idea that CARD 9 operates upstream of Bcl10 and Malt1 to transduce signals for NFκB activation and cytokine production [54].

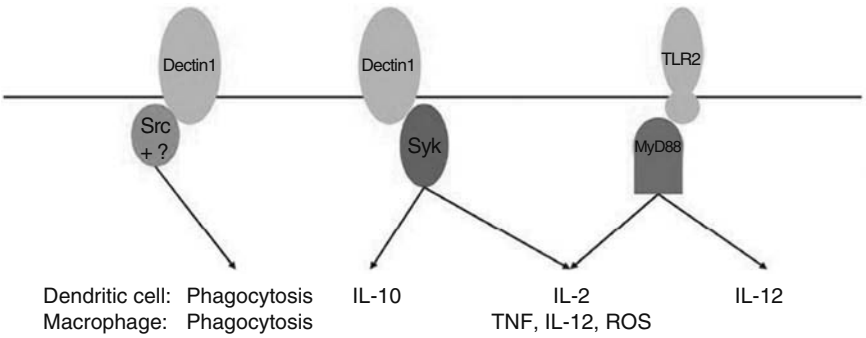


FIGURE 18.4 (See CD for color figure.) Responses induced by Dectin-1 and TLR2 signaling pathways. Fungal stimulation of dendritic cells induces IL-10 through the Dectin-1-Syk pathway, and IL-12 through the TLR2-MyD88 pathway. Dectin-1 and TLR2 collaborate to induce IL-2 in dendritic cells and TNF, IL-12 and ROS in macrophages. Phagocytosis in both dendritic cells and macrophages is largely Syk-independent.

18.7.2 IL-12

Dectin-1 also enhances IL-12 production induced by TLR2 stimulation [15]. IL-12 plays a critical role in Th1-based responses, which have been linked to resistance to fungal infection [55]. It has been suggested that Dectin-1 recognition of β-glucans in fungi may serve to focus Th1-type responses through production of IL-12 [15]. Unlike TNF, the production of IL-12 is CARD 9-independent [54] and only partly dependent on Syk, as shown using cells from CARD 9-deficient mice and Syk-deficient mice [16,54].

18.7.3 IL-10 AND IL-2

Zymosan also induces significant levels of Dectin-1-mediated IL-10 and IL-2 production [16,49]. The production of IL-10 is somewhat surprising given that IL-10 counteracts the effects of IL-12. IL-10 production has advantages and disadvantages for the host during fungal infection. It impairs the antifungal effector mechanisms of phagocytes and the secretion of proinflammatory cytokines. However, its anti-inflammatory action may be beneficial late in the course of infection to aid in the resolution of the inflammatory response [55]. The production of IL-10 by dendritic cells is TLR-independent [16], but requires the cytoplasmic tail of Dectin-1. Like TNF, Dectin-1-mediated IL-10 production is dependent on Syk and CARD 9 [54]. IL-2 has an essential role in the development and peripheral activity of regulatory T cells [56–58], which are reported to have an essential role in protective memory immunity to fungal infections [34,59]. Production of IL-2 requires the cytoplasmic tail of Dectin-1 and is dependent on Syk and CARD 9 [54]. Maximal IL-2 synthesis also requires signaling via MyD88 [16].

18.8 CONCLUSION AND FUTURE DIRECTIONS

The discovery of Dectin-1 as a receptor for fungal pathogens has changed the way that we think about non-TLR PRRs and opened up a new facet of innate immunology. The TLRs are clearly key components of the innate immune system that recognize a wide range of microbial pathogens. However, the emergence of Dectin-1 as the first example of a signaling non-TLR PRR has shown that TLRs do not function alone. Furthermore, the collaborative signaling between Dectin-1 and TLR2 has provided the skeleton of what may be a valuable model for other receptors.

The immune response is launched in response to microbial recognition by a complex network of receptors. The specific cellular response to this recognition is dependent on functional cooperation of downstream signals initiated by receptor activation. Dectin-1 collaborates with TLR2 to tailor the immune response and induce production of TNF, IL-12, and ROS via a novel interaction involving Syk and other unknown molecules. Future research into the nature of Dectin-1 and TLR2 collaboration will define this functional cooperation and provide insight into collaboration between other immunoreceptors. Dectin-1 is therefore likely to be the first of a number of non-TLR, but Syk-coupled PRRs that are critical for the control of various infectious agents.

ACKNOWLEDGMENTS

We thank the Wellcome Trust (UK), National Institutes of Health, CANSA (South Africa), and National Research Foundation (South Africa) for funding. GDB is a Wellcome Trust Senior Research Fellow.

REFERENCES

1. Janeway, C. A., Jr., The immune system evolved to discriminate infectious nonself from noninfectious self, *Immunol Today*, 13, 11, 1992.
2. Brown, G. D., et al., Dectin-1 mediates the biological effects of beta-glucans, *J Exp Med*, 197, 1119, 2003.

3. Gantner, B. N., Simmons, R. M., and Underhill, D. M., Dectin-1 mediates macrophage recognition of Candida albicans yeast but not filaments, *EMBO J*, 24, 1277, 2005.
4. Poltorak, A., et al., Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: Mutations in Tlr4 gene, *Science*, 282, 2085, 1998.
5. Taylor, P. R., et al., Dectin-1 is required for beta-glucan recognition and control of fungal infection, *Nat Immunol*, 8, 31, 2007.
6. Viriyakosol, S., et al., Innate immunity to the pathogenic fungus *Coccidioides posadasii* is dependent on Toll-like receptor 2 and Dectin-1, *Infect Immun*, 73, 1553, 2005.
7. Steele, C., et al., Alveolar macrophage-mediated killing of *Pneumocystis carinii* f. sp. muris involves molecular recognition by the Dectin-1 beta-glucan receptor, *J Exp Med*, 198, 1677, 2003.
8. Saijo, S., et al., Dectin-1 is required for host defense against *Pneumocystis carinii* but not against *Candida albicans*, *Nat Immunol*, 8, 39, 2007.
9. Steele, C., et al., The beta-glucan receptor dectin-1 recognizes specific morphologies of *Aspergillus fumigatus*, *PLoS Pathog*, 1, e42, 2005.
10. Hohl, T. M., et al., *Aspergillus fumigatus* triggers inflammatory responses by stage-specific beta-glucan display, *PLoS Pathog*, 1, e30, 2005.
11. Gersuk, G. M., et al., Dectin-1 and TLRs permit macrophages to distinguish between different Aspergillus fumigatus cellular states, *J Immunol*, 176, 3717, 2006.
12. Brown, G. D. and Gordon, S., Immune recognition. A new receptor for beta-glucans, *Nature*, 413, 36, 2001.
13. Herre, J., et al., Dectin-1 uses novel mechanisms for yeast phagocytosis in macrophages, *Blood*, 104, 4038, 2004.
14. Underhill, D. M., et al., Dectin-1 activates Syk tyrosine kinase in a dynamic subset of macrophages for reactive oxygen production, *Blood*, 106, 2543, 2005.
15. Gantner, B. N., et al., Collaborative induction of inflammatory responses by dectin-1 and Toll-like receptor 2, *J Exp Med*, 197, 1107, 2003.
16. Rogers, N. C., et al., Syk-dependent cytokine induction by Dectin-1 reveals a novel pattern recognition pathway for C type lectins, *Immunity*, 22, 507, 2005.
17. Ariizumi, K., et al., Identification of a novel, dendritic cell-associated molecule, dectin-1, by subtractive cDNA cloning, *J Biol Chem*, 275, 20157, 2000.
18. Grunebach, F., et al., Molecular and functional characterization of human Dectin-1, *Exp Hematol*, 30, 1309, 2002.
19. Brown, G. D. and Gordon, S., Fungal beta-glucans and mammalian immunity, *Immunity*, 19, 311, 2003.
20. Ross, G. D., et al., Specificity of membrane complement receptor type three (CR3) for beta-glucans, *Complement*, 4, 61, 1987.
21. Brown, G. D., et al., Dectin-1 is a major beta-glucan receptor on macrophages, *J Exp Med*, 196, 407, 2002.
22. Suzuki-Inoue, K., et al., A novel Syk-dependent mechanism of platelet activation by the C-type lectin receptor CLEC-2, *Blood*, 107, 542, 2006.
23. Palma, A. S., et al., Ligands for the beta-glucan receptor, Dectin-1, assigned using "designer" microarrays of oligosaccharide probes (neoglycolipids) generated from glucan polysaccharides, *J Biol Chem*, 281, 5771, 2006.
24. Adachi, Y., et al., Characterization of beta-glucan recognition site on C-type lectin, dectin 1, *Infect Immun*, 72, 4159, 2004.
25. Taylor, P. R., et al., The beta-glucan receptor, dectin-1, is predominantly expressed on the surface of cells of the monocyte/macrophage and neutrophil lineages, *J Immunol*, 169, 3876, 2002.
26. Heinsbroek, S. E., et al., Expression of functionally different dectin-1 isoforms by murine macrophages, *J Immunol*, 176, 5513, 2006.
27. Willment, J. A., Gordon, S., and Brown, G. D., Characterization of the human beta-glucan receptor and its alternatively spliced isoforms, *J Biol Chem*, 276, 43818, 2001.
28. Willment, J. A., et al., Dectin-1 expression and function are enhanced on alternatively activated and GM-CSF-treated macrophages and are negatively regulated by IL-10, dexamethasone, and lipopolysaccharide, *J Immunol*, 171, 4569, 2003.
29. Garcia-Garcia, E. and Rosales, C., Signal transduction during Fc receptor-mediated phagocytosis, *J Leukoc Biol*, 72, 1092, 2002.
30. Karin, M. and Delhase, M., The I kappa B kinase (IKK) and NF-kappa B: Key elements of proinflammatory signalling, *Semin Immunol*, 12, 85, 2000.
31. Pitcher, L. A. and van Oers, N. S., T-cell receptor signal transmission: Who gives an ITAM?, *Trends Immunol*, 24, 554, 2003.

32. Rawlings, D. J., Sommer, K., and Moreno-Garcia, M. E., The CARMA1 signalosome links the signalling machinery of adaptive and innate immunity in lymphocytes, *Nat Rev Immunol*, 6, 799, 2006.
33. Brown, G. D., Dectin-1: A signalling non-TLR pattern-recognition receptor, *Nat Rev Immunol*, 6, 33, 2006.
34. Crowley, M. T., et al., A critical role for Syk in signal transduction and phagocytosis mediated by Fcgamma receptors on macrophages, *J Exp Med*, 186, 1027, 1997.
35. Fitzer-Attas, C. J., et al., Fcgamma receptor-mediated phagocytosis in macrophages lacking the Src family tyrosine kinases Hck, Fgr, and Lyn, *J Exp Med*, 191, 669, 2000.
36. Majeed, M., et al., Role of Src kinases and Syk in Fcgamma receptor-mediated phagocytosis and phagosome-lysosome fusion, *J Leukoc Biol*, 70, 801, 2001.
37. Diniz, S. N., et al., PTX3 function as an opsonin for the dectin-1-dependent internalization of zymosan by macrophages, *J Leukoc Biol*, 75, 649, 2004.
38. Taylor, P. R., Tsoni, V., Willment, J. A., Dennehy, K. M., Rosas, M., Findon, H., Haynes, K., Steele, C., Botto, M., Gordon, S., and Brown, G. D., A critical role for beta-gluan recognition in the control of fungal infection, *Nat Immunol*, 8(1), 31, 2007.
39. Mantegazza, A. R., et al., CD63 tetraspanin slows down cell migration and translocates to the endosomal-lysosomal-MIICs route after extracellular stimuli in human immature dendritic cells, *Blood*, 104, 1183, 2004.
40. Meyer-Wentrup, F., et al., Dectin-1 interaction with tetraspanin CD37 inhibits IL-6 production, *J Immunol*, 178, 154, 2007.
41. Boucheix, C. and Rubinstein, E., Tetraspanins, *Cell Mol Life Sci*, 58, 1189, 2001.
42. Hemler, M. E., Tetraspanin proteins mediate cellular penetration, invasion, and fusion events and define a novel type of membrane microdomain, *Annu Rev Cell Dev Biol*, 19, 397, 2003.
43. Wright, M. D., Moseley, G. W., and van Spriel, A. B., Tetraspanin microdomains in immune cell signalling and malignant disease, *Tissue Antigens*, 64, 533, 2004.
44. Wijkander, J. and Sundler, R., A role for protein kinase C-mediated phosphorylation in the mobilization of arachidonic acid in mouse macrophages, *Biochim Biophys Acta*, 1010, 78, 1989.
45. Raulf, M. and Konig, W., Modulation of leukotriene release from human polymorphonuclear leucocytes by PMA and arachidonic acid, *Immunology*, 64, 51, 1988.
46. Suram, S., et al., Regulation of cytosolic phospholipase A2 activation and cyclooxygenase 2 expression in macrophages by the beta-glucan receptor, *J Biol Chem*, 281, 5506, 2006.
47. Olsson, S. and Sundler, R., The macrophage beta-glucan receptor mediates arachidonate release induced by zymosan: Essential role for Src family kinases, *Mol Immunol*, 2006.
48. Babior, B. M., Phagocytes and oxidative stress, *Am J Med*, 109, 33, 2000.
49. Granucci, F., et al., Early IL-2 production by mouse dendritic cells is the result of microbial-induced priming, *J Immunol*, 170, 5075, 2003.
50. Kawakami, K., et al., Contribution of tumour necrosis factor-alpha (TNF-alpha) in host defence mechanism against Cryptococcus neoformans, *Clin Exp Immunol*, 106, 468, 1996.
51. Mehrad, B., Strieter, R. M., and Standiford, T. J., Role of TNF-alpha in pulmonary host defense in murine invasive aspergillosis, *J Immunol*, 162, 1633, 1999.
52. Netea, M. G., et al., Increased susceptibility of TNF-alpha lymphotoxin-alpha double knockout mice to systemic candidiasis through impaired recruitment of neutrophils and phagocytosis of Candida albicans, *J Immunol*, 163, 1498, 1999.
53. Allendoerfer, R. and Deepe, G. S., Jr., Blockade of endogenous TNF-alpha exacerbates primary and secondary pulmonary histoplasmosis by differential mechanisms, *J Immunol*, 160, 6072, 1998.
54. Gross, O., et al., Card9 controls a non-TLR signalling pathway for innate anti-fungal immunity, *Nature*, 442, 651, 2006.
55. Romani, L., Immunity to fungal infections, *Nat Rev Immunol*, 4, 1, 2004.
56. Almeida, A. R., et al., Homeostasis of peripheral CD4+ T cells: IL-2R alpha and IL-2 shape a population of regulatory cells that controls CD4+ T cell numbers, *J Immunol*, 169, 4850, 2002.
57. Klebb, G., et al., Interleukin-2 is indispensable for development of immunological self-tolerance, *Clin Immunol Immunopathol*, 81, 282, 1996.
58. Furtado, G. C., et al., Interleukin 2 signaling is required for CD4(+) regulatory T cell function, *J Exp Med*, 196, 851, 2002.
59. Montagnoli, C., et al., B7/CD28-dependent CD4+ CD25+ regulatory T cells are essential components of the memory-protective immunity to Candida albicans, *J Immunol*, 169, 6298, 2002.

19 CD22: A Regulator of B Cell Survival and Signal Transduction

Susan H. Smith and Thomas F. Tedder

CONTENTS

19.1 INTRODUCTION

B lymphocytes are the central mediators of humoral immunity. They differentiate through highly regulated pathways before becoming mature plasma cells that secrete Ag-specific antibody. B cells depend on cues from their extracellular microenvironment for development, homeostasis, activation, proliferation, and effector function. These functions are regulated through cell-surface molecules that generate transmembrane signals, regulate intercellular communication, and direct lymphocyte localization within tissues. These events depend on signals generated by the B cell Ag receptor (BCR) that is composed of membrane immunoglobulin (Ig) noncovalently associated with disulfide-linked CD79a/CD79b (Ig-α/Ig-β) heterodimers.

In addition to the BCR, multiple other signaling molecules provide important functional links between the cell surface and intracellular signaling (Tedder 1998). CD22 and CD19 represent two specialized cell surface molecules (Buhl and Cambier 1997; Cyster and Goodnow 1997; O'Rourke et al. 1997; Tedder et al. 1997; Nitschke and Tsubata 2004) that also function as "response regulators" (Tedder 1998) to modulate the intensity, quality, and duration of signals generated for homeostatic survival as well as Ag-induced responses (Fujimoto et al. 1998; Sato et al. 1998). Response regulators establish intrinsic signaling thresholds that provide a context for other transmembrane and cytoplasmic signals.

CD22 expression is developmentally regulated. Cytoplasmic CD22 protein appears during early B cell development, while initial CD22 surface expression is found at the immature B cell stage in the bone marrow (Erickson et al. 1996). CD22 surface levels increase as B cells mature with the highest levels of CD22 expression found in resting mature B cells. Newly formed immature B cells are subject to multiple survival checkpoints where the function and reactivity of their BCRs are

assessed. Proper expression and normal BCR signal transduction is essential for the continued survival of both immature and mature B lymphocytes (Lam et al. 1997). The acquisition of surface CD22 concurrent with surface Ig places it in a unique position to regulate BCR-mediated signal transduction during critical selection steps.

19.2 CD22 DOMAIN ORGANIZATION AND INTERACTION PARTNERS

CD22 is a type I transmembrane protein comprised of a single amino-terminal V-set Ig domain and six C-2-set Ig domains (Wilson et al. 1991; Torres et al. 1992; Engel et al. 1995). Although smaller CD22 mRNA isoforms exist that lack one or more Ig-like domains due to alternative splicing, cell-surface CD22 is almost exclusively comprised of 847 amino acids, including a 141 amino acid cytoplasmic tail (Wilson et al. 1991; Engel et al. 1993). CD22 is highly conserved among mammals. Mouse CD22 is roughly 62% identical in overall amino acid sequence to human CD22 (Torres et al. 1992). The highest level of conservation (70%) is found between the seventh Ig domain, the cytoplasmic tail, and the transmembrane domains (Wilson et al. 1991; Torres et al. 1992).

CD22 recognizes specific N-linked oligosaccharides containing α2,6-linked sialic acids (Engel et al. 1993; Powell et al. 1993; Sgroi et al. 1993; Powell and Varki 1994). As depicted in Figure 19.1, ligand binding activity of CD22 resides exclusively in Ig-like domains 1 and 2 based on the localization of a large panel of function-blocking anti-CD22 mAbs (Engel et al. 1995; Law et al. 1995). Using these antibodies as well as domain-swap and domain-deletion mutants of hCD22, the two amino-terminal Ig-like domains were determined to embody the functional unit mediating CD22 adhesion with ligands found on lymphocytes, neutrophils, monocytes, and erythrocytes (Engel et al. 1995). This is in contrast to sialoadhesin, which requires only the amino-terminal V-set Ig domain (domain 1) for ligand binding, and may implicate the adjacent C2-set domain of CD22 (domain 2) in the proper folding of a sialic acid–binding site on domain 1 (Nath et al. 1995).

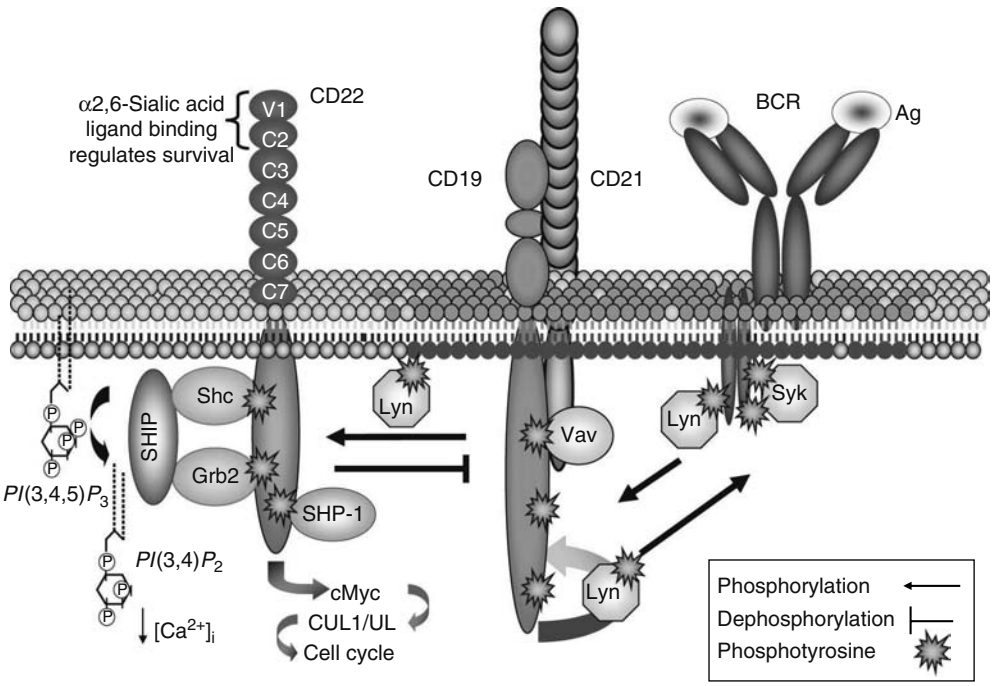


FIGURE 19.1 (See color insert following blank page 170. Also see CD for color figure.) CD22 modulates signal transduction and survival in B cells.

Potential CD22 ligands identified *in vitro* include structurally diverse sialic acid–bearing molecules expressed by lymphocytes, monocytes, neutrophils, and erythrocytes as well as nonhematopoietic cells (Tedder et al. 1997). Specific examples include CD22 itself, all isoforms of CD45, soluble IgM pentamers, haptoglobin, Ly-6 proteins, and a variety of other diverse proteins present on leukocyte cell surfaces (Stamenkovic et al. 1991; Hanasaki et al. 1995; Pflugh et al. 2002). CD22 is also reported to bind cell surface Ig (Leprince et al. 1993; Peaker and Neuberger 1993), although it is not known whether this association is direct through the extracellular ligand-binding or intracellular CD22 domains or is accomplished via intermediate bridging molecules. CD22 has been shown to bind sialic acid–bearing glycoproteins on B cells in *cis*, as well as on adjacent cells in *trans. Cis* binding has been referred to as "masking" since it blocks CD22 binding to synthetic sialoside probes (Razi and Varki 1998). Importantly, CD22 masking does not impede its binding to *trans* ligands on adjacent cells (Collins et al. 2004).

The ~140 amino acid cytoplasmic domain of CD22 contains six tyrosine residues that are potential targets for phosphorylation following surface BCR or CD22 ligation (Wilson et al. 1991; Torres et al. 1992). Three tyrosine residues reside in immunoreceptor tyrosine-based inhibitory motifs (ITIMs) and are implicated in recruitment of the src-homology-2 (SH2) domain-containing phosphatase 1 (SHP-1). Two cytoplasmic tyrosine residues also fall within an immunoreceptor tyrosine-based activation motif (ITAM; Leprince et al. 1993), a region found in many activating receptors of the immune system, including the BCR, T cell antigen receptor (TCR), and activating Fc receptors (Reth 1989). Through these motifs, CD22 provides binding sites for SHP1, PLC-γ2, phosphatidylinositol (PI) 3-kinase, Grb2, Syk, Shc, and SHIP (Figure 19.1) (Blasioli et al. 1999; Yohannan et al. 1999; O'Keefe et al. 1996; Poe et al. 2000).

19.3 CD22 SIGNALING REGULATES B CELL ACTIVATION

CD22 has earned a reputation as a negative regulator of B cell receptor signaling, which is largely based on the recruitment of SHP-1 by phosphorylated tyrosine residues in the CD22 cytoplasmic tail. In fact, of all phosphotyrosine-containing proteins, CD22 is the dominant binding partner for SHP-1 in B cells (Sato et al. 1997). CD19 is also a major target for CD22 regulation (see Figure 19.1) (Poe et al. 2000), as CD19 is hyperphosphoryated in $CD22^{-/-}$ B cells. In fact, $CD19^{-/-}$ and $CD22^{-/-}/CD19^{-/-}$ double-deficient mice bear strikingly similar phenotypes, suggesting that the inhibitory effects of CD22 inhibition act primarily via CD19 (Fujimoto et al. 1999). However, given the array of known binding partners for CD22, the mechanism by which CD22 limits BCR signal strength may not be solely dependent upon SHP-1 and the regulation of CD19. Rather, CD22 participates in a tetrameric complex with SHIP, Grb-2, and Shc (Poe et al. 2000), which may also play a direct role in altering calcium mobilization following BCR engagement (Figure 19.1) (O'Keefe et al. 1996).

To tease out the importance of CD22 *in vivo*, four independent lines of CD22-deficient mice have been characterized (O'Keefe et al. 1996; Otipoby et al. 1996; Sato et al. 1996; Nitschke et al. 1997). B cells from $CD22^{-/-}$ mice exhibit an activated phenotype, including decreased surface IgM and increased major histocompatibility complex (MHC) class II levels. Proliferative responses to BCR cross-linking are reduced depending on the genetic background of the mice being examined (Poe et al. 2004a), although normal proliferation in $CD22^{-/-}$ B cells is observed after lipopolysaccharide (LPS) stimulation or CD38 cross-linking (Sato et al. 1996). In contrast, $CD22^{-/-}$ B cells are hyperresponsive to CD40 stimulation (Poe et al. 2004b). That CD22 contains both positive and negative signaling elements in its intracellular domain may also explain why both positive and negative signaling roles for CD22 are observed during different *in vitro* experiments. The most consistent observation, however, is that calcium mobilization in response to BCR stimulation is heightened in $CD22^{-/-}$ B cells, which contributes to its reputation as a negative regulator of BCR signals.

19.4 CD22–LIGAND INTERACTIONS REGULATE B CELL SURVIVAL

Mice deficient in CD22 have relatively normal peripheral B cell numbers, although marginal zone B cells are absent and the number of mature recirculating and marrow B cells are reduced

(O'Keefe et al. 1996; Otipoby et al. 1996; Sato et al. 1996; Nitschke et al. 1997). CD22$^{-/-}$ B cells also exhibit higher turnover, based on 5-bromo-deoxyuridine (BrdU)–labeling studies, which reveals a shorter average lifespan *in vivo* (Otipoby et al. 1996; Nitschke et al. 1997; Haas et al. 2006). In addition, T-cell-independent immune responses are decreased in CD22$^{-/-}$ mice (Otipoby et al. 1996; Sato et al. 1996; Nitschke et al. 1997), although this may reflect the lack of a functional marginal zone B cell population, which are the primary responders to T-independent Ag (Sato et al. 1996; Nitschke et al. 1997; Samardzic et al. 2002).

Studies conducted with CD22$^{-/-}$ mice have been unable to address the importance of CD22 ligand binding in regulating CD22 function. Therefore, to directly assess the relevance of CD22 ligand binding in B cell activation and survival, two lines of mice were generated that express mutant CD22 receptors incapable of ligand binding, CD22Δ1–2 and CD22AA (Poe et al. 2004b). CD22Δ1–2 mice express a truncated form of CD22 that lacks the two amino terminal Ig domains required for ligand binding. CD22AA mice harbor two point mutations (R130 to A130, R137 to A137) in the amino terminal Ig domain of CD22 that results in the inability of CD22 to bind sialic acid–bearing ligands (van der Merwe et al. 1996; Poe et al. 2004b). CD22Δ1–2 and CD22AA B cells regulate BCR-induced calcium responses normally. In addition, CD22 tyrosine phosphorylation and SHP-1 recruitment are also normal in CD22Δ1–2 and CD22AA B cells after BCR ligation. However, CD22$^{-/-}$, CD22Δ1–2, and CD22AA B cells are similarly hypoproliferative to BCR signals yet hyperproliferative to CD40 stimulation *in vitro* (Poe et al. 2004b). These data indicate that some signal transducing elements of CD22 function independent of ligand binding, while normal B cell proliferation in response to various stimuli may require CD22–ligand interactions. That CD22 ligand binding is required for optimal proliferation responses may relate to the observation that CD22 regulates Myc transcription factor–induced expression of the ubiquitin ligase complex protein Cullin 1 (CUL1) after BCR ligation, which promotes ubiquitin-dependent cell cycle progression (Poe et al. 2004a).

Despite the complexity of B cell signal transduction, CD22–ligand interactions are vital for normal B cell physiology. There is a significant reduction in the number of marginal zone and circulating B cells in CD22$^{-/-}$, CD22Δ1–2, and CD22AA mice. In addition, mature B cells have high *in vivo* turnover rates in CD22Δ1–2, CD22AA, and CD22$^{-/-}$ mice (Poe et al. 2004b). CD22 was originally identified as a potent adhesion molecule and was implicated in B cell migration *in vivo* (Floyd et al. 2000). Consistent with this notion, CD22 ligands have been identified on bone marrow sinusoidal endothelium (Nitschke et al. 1999). However, the direct comparison of migration *in vivo* by wild type, CD22$^{-/-}$, CD22Δ1–2, and CD22AA B cells at early time points following adoptive transfer indicate that survival, and not differential migration, is responsible for the observed decrease in mature B cell numbers in the bone marrow of CD22$^{-/-}$ mice. While CD22 may regulate the threshold for B cell activation in a ligand-independent manner, high turnover rates and reductions in specific B cell populations in CD22$^{-/-}$, CD22Δ1–2, and CD22AA mice indicate that normal B cell survival *in vivo* is dependent on CD22–ligand interaction.

To further understand the role of ligand binding in CD22-mediated B cell survival, CD22–ligand interactions have been disrupted *in vivo* using a panel of mouse antimouse monoclonal antibodies (mAbs) that inhibit CD22 ligand binding to varying degrees. Within this panel, antibodies termed MB22-9, -10, and -11 respectively block ligand binding to increasing degrees, whereas MB22-8 and Cy34.1 mAbs efficiently bind to CD22 but do not inhibit ligand binding (Haas et al. 2006). Mice treated with ligand-blocking CD22 mAbs exhibit dramatically reduced (75%–85%) circulating B cell numbers whereas mAb MB22-8, which does not block ligand binding, does not have a significant effect (Haas et al. 2006). Some B cell subsets are particularly dependent on CD22–ligand interactions as mAb treatment resulted in dramatic reductions in marginal zone and mature B cells in the bone marrow, whereas mature splenic B cells are not significantly altered. Lymph node (LN) B cell numbers are only slightly reduced upon treatment with the MB22-9 and MB22-10 mAbs, while treatment with MB22-11 mAb causes significant reductions. Since blood, marginal zone, and LN B cell subsets represent a minor proportion of the peripheral B cell population, the observed

reductions do not alter total peripheral B cell numbers. Nevertheless, BrdU-labeling studies reveal striking alterations in B cell turnover rates in MB22-10- and MB22-11-mAb treated mice. Increased turnover indicates a change in proliferation, survival, or both. However, mAb treatment following adoptive transfer of Carboxyfluorescein succinimidyl ester (CFSE)-labeled B cells indicates that survival and not proliferation is the primary contributor to enhanced B cell turnover *in vivo* (Haas et al. 2006).

The physiology of CD22 ligand binding *in vivo* has also been addressed by targeting the sialyltransferase ST6Gal-I, which is responsible for producing α2–6-linked sialic acids on oligosaccharides, recognized as CD22 ligands. Mice deficient in ST6Gal-I exhibit reduced serum IgM levels, although B cell development and peripheral numbers appear normal (Hennet et al. 1998). B cells from ST6Gal-I-deficient mice express reduced levels of IgM and CD22 on the cell surface. In contrast to $CD22^{-/-}$ B cells, ST6Gal-I-deficient B cells do not exhibit an "activated" phenotype, including increased surface MHC class II expression. Furthermore, B cells from ST6Gal-I-deficient mice have impaired proliferative responses to both BCR stimulation and CD40 signaling and reduced antibody production after challenge (Hennet et al. 1998). Thus, ST6Gal-I-deficient mice show gross abnormalities in immune function, but differ from $CD22^{-/-}$ mice. These differences also support the observation that CD22 can act independent of ligand binding. However, additional roles for α2,6-linked oligosaccharides in the immune system, as well as functional redundancy for CD22–ligands may still exist. Efforts to resolve these possibilities have led to the production of CD22/ST6Gal-I double-deficient mice (Collins et al. 2006). B cells from CD22/ST6Gal-I double-deficient mice largely resemble $CD22^{-/-}$ B cells with regard to calcium mobilization, upregulation of the activation marker CD86, and proliferation after BCR cross-linking, and further support a role for CD22 in altering B cell signal transduction independent of ligand interactions.

Taken together, data from $CD22^{-/-}$, CD22Δ1–2, CD22AA, ST6Gal-I-deficient, and CD22 mAb-treated mice decisively demonstrate a role for CD22–ligand interactions in the survival of peripheral B cells in addition to the ligand-independent effects of CD22 on B cell signal transduction. Since CD22 provides an important functional link between B cell survival and regulation of intracellular signaling, it is not surprising that CD22-directed therapies are being developed and tested for both oncology and autoimmunity applications.

19.5 CD22 AS A THERAPEUTIC TARGET

Therapy for B cell malignancies and severe autoimmune disease has primarily relied upon broadly immunosuppressive agents, such as cyclophosphamide, methotrexate, or cyclosporine A, which have significant toxicities. Therapies directed at specifically reducing B cell numbers have therefore acquired a great deal of attention and enthusiasm. For example, a chimerized CD20 mAb (Rituximab) effectively reduces normal and malignant B cell numbers without significant toxicity and is currently being used widely in patients with lymphoma and autoimmune disease.

19.6 CD22 THERAPEUTIC APPLICATIONS IN ONCOLOGY

CD22 is expressed by 75%–80% of B cell lymphomas and leukemias (Press et al. 2001). Antibody binding to CD22 results in rapid internalization although its degradation kinetics are unchanged (Shan and Press 1995). In light of these observations, CD22 has emerged as a pertinent target for immunotherapy of B cell malignancies. Radiolabeled CD22 mAbs (Stein et al. 1993) have demonstrated high sensitivities in the diagnosis and the staging of B cell lymphomas, while therapeutic ^{131}I- or ^{90}Y-labeled CD22 mAbs (LL2) that bind to the third extracellular Ig-like domain of CD22 have generated partial to complete remissions (Juweid et al. 1999; Sieling et al. 1999; Coleman et al. 2003). Naked CD22 mAb (Epratuzumab) mediates antibody- and complement-dependent cytotoxicity *in vitro*, with preclinical and early clinical studies providing a rationale for its use in immunotherapy (Leonard and Link 2002; Carnahan et al. 2003). Because malignant B cells rapidly internalize CD22 and CD22 mAb binding leads to efficient intracellular routing of the CD22–mAb

complex to intracellular compartments (Press et al. 1994), the use of CD22 mAbs to deliver immunotoxins to non-Hodgkin's lymphoma and chronic lymphocytic leukemia cells has also received considerable attention (Press et al. 1989; Sieber et al. 2003).

The importance of CD22 ligand binding for normal B cell survival suggests that blocking this interaction may influence malignant B cell survival. In fact, ligand-blocking mAbs have potent *in vivo* effects on lymphoma growth. This concept was first validated using a Raji lymphoma xenograft model (Tuscano et al. 2003b) and the HB22-7 mAb (Engel et al. 1993). Surprisingly, mice treated with naked HB22-7 mAb alone have impressive tumor volume reductions, with superior cure and survival rates when compared to other treatment groups (Tuscano et al. 2003b). Furthermore, using ligand-blocking mouse antimouse mAb, A20 mouse lymphoma cells transferred into Rag1-deficient syngeneic hosts were found to require CD22 ligand binding for survival as well. Mice treated with MB22-10 mAb following adoptive transfer contain approximately 90% fewer A20 lymphoma cells than control mice (Haas et al. 2006). These findings demonstrate that both normal and malignant B cells require CD22 ligand binding for survival, validating CD22 as a rational target for treating B cell malignancies.

19.7 CD22-DIRECTED THERAPIES IN AUTOIMMUNITY

Autoimmune diseases represent complex disorders characterized by adaptive immune responses that are inappropriately directed against self-antigens. Despite the inherent complexity, B cells are important contributors in the pathogenesis of autoimmunity beyond their ability to produce autoantibodies (Madaio 1998; Chan et al. 1999a,b; Tuscano et al. 2003a). This has provided new incentives to identify the multiple checkpoints that control positive and negative selection of B cells, both centrally in the bone marrow and in peripheral lymphoid tissues. A complex series of checks and balances throughout B cell development is necessary to allow the production of a large and diverse population of B cells capable of generating high-affinity effector antibodies that have been purged of pathological autoreactivity. As a result, perturbations in select regulatory pathways that affect B cell function or selection processes may lead to autoimmune disease. Two broad categories of defects that lead to autoimmunity have been identified: those that alter signal transduction thresholds for cellular activation and those that alter B cell longevity. Since CD22 regulates both signal transduction thresholds and B cell longevity, it is not surprising that CD22 has been implicated in the development of autoimmunity, at least in mouse models of disease.

That the extracellular ligand binding domains of CD22 are divergent in autoimmune strains of mice suggests that CD22 may directly influence autoantibody production or may contribute to the severity of autoimmune disorders (Mary et al. 2000). First, at least three alleles of the mouse Cd22 gene have been identified that encode proteins that are ~3% different in amino acid sequence (Law et al. 1993). One of these, the Cd22a allele, is uniquely found in some autoimmune-prone strains of mice (DBA/1, DBA/2J, NZB, NZW, NZC, PL/J, and AKR/J) (Nadler et al. 1997; Lajaunias et al. 1999). Although not proven, multiple investigators postulate that CD22 allelic differences contribute to autoimmunity. In addition, NZW mice synthesize aberrant CD22 mRNAs, some of which interfere with upregulated cell surface CD22 expression on NZW B cells following LPS exposure (Mary et al. 2000). Thus, CD22 genetic polymorphisms may contribute to autoimmunity.

A second line of evidence that CD22 influences the development of autoimmunity is that one $CD22^{-/-}$ mouse line has increased autoantibody production with age, including antibodies against double-stranded (ds) DNA and dsDNA–histone complexes (O'Keefe et al. 1996, 1999). Heterozygous $CD22^{+/-}$ C57BL/6 mice carrying the autoimmune acceleration gene, Yaa, also have markedly increased production of IgG anti-DNA autoantibodies (Mary et al. 2000). The hyperactivated status of $CD22^{-/-}$ B cells is consistent with susceptibility to autoimmunity such that CD22 may augment autoimmune disease once induced or when combined with increasing age, other disease susceptibility loci, or pathogens. Alternatively, it has been proposed that CD22 ligand engagement prevents the development of autoimmunity (Lanoue et al. 2002). By contrast, our line of $CD22^{-/-}$ mice develops

a hyperlgMphenotype with age, with autoantibody levels merely paralleling increased serum Ig levels (unpublished observations). Thus, in this line of mice, CD22 deficiency appears to result in overall B cell hyperactivity rather than preferentially supporting the expansion of autoreactive clones. Furthermore, estrogen administration, which accelerates and exacerbates autoimmunity in mice (Grimaldi et al. 2001) specifically upregulates CD22 and SHP1 levels in B cells (Grimaldi et al. 2002). These authors suggest that CD22 overexpression may reduce the occurrence of BCR-induced apoptosis in autoreactive cells and promote their survival. Thus, there is evidence for CD22 deficiency as well as CD22 overexpression in influencing autoimmunity.

The importance of B cell survival to autoimmune disease is well appreciated, leading to the pursuit of therapies that influence this potent biological outcome. By interfering with B cell survival and enhancing B cell turnover, it is envisioned that B cells or B cell subsets can be specifically targeted without the widespread toxicity of conventional treatments including systemic corticosteroid administration. Moreover, it may be possible to effect therapeutic benefit by attenuating B cell responses to transmembrane signals without eliminating B cells that give rise to protective humoral immunity. Accelerated B cell apoptosis and turnover by therapeutic mAbs that block CD22 ligand engagement may therefore have considerable benefit for the treatment of autoimmunity (Tuscano et al. 2003a).

19.8 CONCLUSION AND FUTURE DIRECTIONS

CD22 expressed by mature B-lineage cells binds ligands *in vivo* to regulate BCR-mediated signal transduction and provide essential survival signals. CD22 ligand-dependent and independent pathways may function independently or converge to ultimately regulate the composition of the preimmune B cell repertoire and the generation of physiologically relevant responses to foreign or self-Ags. Through its effects on both B cell survival and signal transduction, CD22 likely plays an important role in normal B cell selection, activation, and differentiation during immune responses. A better understanding of CD22 function and regulation may also afford mechanisms by which humoral immunity can be modulated, leading to the development of new strategies to augment antimicrobial defense and acquired immune responses, and for the development of novel approaches aimed at regulating chronic infections and inflammatory disorders. Since B cells contribute substantially to many human autoimmune diseases and hematologic malignancies, and these diseases are often correlated with altered signal transduction or unregulated B cell survival, new approaches targeting CD22 directly or interfering with CD22–ligand binding and function represent a valid and exciting new arena for therapeutic advancement.

REFERENCES

Blasioli, J., S. Paust, et al. 1999. Definition of the sites of interaction between the protein tyrosine phosphatase SHP-1 and CD22. *J. Biol. Chem*. 274: 2303–2307.

Buhl, A. M. and J. C. Cambier 1997. Co-receptor and accessory regulation of B-cell antigen receptor signal transduction. *Immunol. Rev.* 160: 127–138.

Carnahan, J., P. Wang, et al. 2003. Epratuzumab, a humanized monoclonal antibody targeting CD22: Characterization of in vitro properties. *Clin. Cancer Res.* 9: 3982S–3990S.

Chan, O. T., L. G. Hannum, et al. 1999a. A novel mouse with B cells but lacking serum antibody reveals an antibody-independent role for B cells in murine lupus. *J. Exp. Med*. 189: 1639–1648.

Chan, O. T., M. P. Madaio, et al. 1999b. The central and multiple roles of B cells in lupus pathogenesis. *Immunol. Rev.* 169: 107–121.

Coleman, M., D. M. Goldenberg, et al. 2003. Epratuzumab: Targeting B-cell malignancies through CD22. *Clin. Cancer Res.* 9: 3991S–3994S.

Collins, B. E., O. Blixt, et al. 2004. Masking of CD22 by *cis* ligands does not prevent redistribution of CD22 to sites of cell contact. *Proc. Natl Acad. Sci. USA* 101: 6104–6109.

Collins, B. E., B. A. Smith, et al. 2006. Ablation of CD22 in ligand-deficient mice restores B cell receptor signaling. *Nat. Immunol*. 7(2): 199–206.

Cyster, J. G. and C. C. Goodnow 1997. Tuning antigen receptor signaling by CD22: Integrating cues from antigens and the microenvironment. *Immunity* 6: 509–517.

Engel, P., Y. Nojima, et al. 1993. The same epitope on CD22 of B lymphocytes mediates the adhesion of erythrocytes, T and B lymphocytes, neutrophils and monocytes. *J. Immunol*. 150: 4719–4732.

Engel, P., N. Wagner, et al. 1995. Identification of the ligand binding domains of CD22, a member of the immunoglobulin superfamily that uniquely binds a sialic acid-dependent ligand. *J. Exp. Med*. 181: 1581–1586.

Erickson, L. D., L. T. Tygrett, et al. 1996. Differential expression of CD22 (Lyb8) on murine B cells. *Int. Immunol*. 8: 1121–1129.

Floyd, H., L. Nitschke, et al. 2000. A novel subset of murine B cells that expresses unmasked forms of CD22 is enriched in the bone marrow: Implications for B-cell homing to the bone marrow. *Immunology* 101: 342–347.

Fujimoto, M., A. P. Bradney, et al. 1999. Modulation of B lymphocyte antigen receptor signal transduction by a CD19/CD22 regulatory loop. *Immunity* 11: 191–200.

Fujimoto, M., J. C. Poe, et al. 1998. CD19 regulates B lymphocyte responses to transmembrane signals. *Semin. Immunol*. 10: 267–277.

Grimaldi, C. M., J. Cleary, et al. 2002. Estrogen alters thresholds for B cell apoptosis and activation. *J. Clin. Invest*. 109(12): 1625–1633.

Grimaldi, C. M., D. J. Michael, et al. 2001. Cutting edge: Expansion and activation of a population of autoreactive marginal zone B cells in a model of estrogen-induced lupus. *J. Immunol*. 167: 1886–1890.

Haas, K. M., S. Sen, et al. 2006. CD22 ligand binding regulates normal and malignant B lymphocyte survival in vivo. *J. Immunol*. 177: 3063–3073.

Hanasaki, K., L. D. Powell, et al. 1995. Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22: selective recognition of immunoglobulin M and haptoglobin. *J. Biol. Chem*. 270: 7543–7550.

Hennet, T., D. Chui, et al. 1998. Immune regulation by the ST6Gal sialyltransferase. *Proc. Natl Acad. Sci. USA* 95: 4504–4509.

Juweid, M. E., E. Stadtmauer, et al. 1999. Pharmacokinetics, dosimetry, and initial therapeutic results with ^{131}I- and ^{111}In-/^{90}Y-labeled humanized LL2 anti-CD22 monoclonal antibody in patients with relapsed, refractory non-Hodgkin's lymphoma. *Clin. Cancer Res*. 5: 3292–3302.

Lajaunias, F., N. Ibnou-Zekri, et al. 1999. Polymorphisms in the *Cd22* gene of inbred mouse strains. *Immunogenetics* 49: 991–995.

Lam, K. P., R. Kuhn, et al. 1997. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell* 90: 1073–1083.

Lanoue, A., F. D. Batista, et al. 2002. Interaction of CD22 with α2,6-linked sialoglycoconjugates: Innate recognition of self to dampen B cell autoreactivity? *Eur. J. Immunol*. 32: 348–355.

Law, C. -L., A. Aruffo, et al. 1995. Ig domains 1 and 2 of murine CD22 constitute the ligand-binding domain and bind multiple sialylated ligands expressed on B and T cells. *J. Immunol*. 155: 3368–3376.

Law, C. -L., R. M. Torres, et al. 1993. Organization of the murine *Cd22* locus. Mapping to chromosome 7 and characterization of two alleles. *J. Immunol*. 151: 175–187.

Leonard, J. P. and B. K. Link 2002. Immunotherapy of non-Hodgkin's lymphoma with hLL2 (epratuzumab, an anti-CD22 monoclonal antibody) and Hu1D10 (apolizumab). *Semin. Oncol*. 29(1 Suppl 2): 81–86.

Leprince, C., K. E. Draves, et al. 1993. CD22 associates with the human surface IgM-B cell antigen receptor complex. *Proc. Natl Acad. Sci. USA* 90: 3236–3240.

Madaio, M. P. 1998. B cells and autoantibodies in the pathogenesis of lupus nephritis. *Immunol. Res*. 17: 123–132.

Mary, C., C. Laporte, et al. 2000. Dysregulated expression of the *Cd22* gene as a results of a short interspersed nucleotide element insertion in *Cd22^a* lupus-prone mice. *J. Immunol*. 165: 2987–2996.

Nadler, M. J. S., P. A. McLean, et al. 1997. B cell antigen receptor-evoked calcium influx is enhanced in CD22-deficient B cell lines. *J. Immunol*. 159: 4233–4243.

Nath, D., P. A. van der Merwe, et al. 1995. The amino-terminal immunoglobulin-like domain of sialoadhesin contains the sialic acid binding site. Comparison with CD22. *J. Biol. Chem*. 270: 26184–26191.

Nitschke, L., R. Carsetti, et al. 1997. CD22 is a negative regulator of B-cell receptor signaling. *Curr. Biol*. 7: 133–143.

Nitschke, L., H. Floyd, et al. 1999. Identification of CD22 ligands on bone marrow sinusoidal endothelium implicated in CD22-dependent homing of recirculating B cells. *J. Exp. Med*. 189: 1513–1518.

Nitschke, L. and T. Tsubata 2004. Molecular interactions regulate BCR signal inhibition by CD22 and CD72. *Trends Immunol*. 25(10): 543–550.

O'Keefe, T. L., G. T. Williams, et al. 1999. Deficiency in CD22, a B cell-specific inhibitory receptor, is sufficient to predispose to development of high affinity autoantibodies. *J. Exp. Med*. 189: 1307–1313.

O'Keefe, T. L., G. T. Williams, et al. 1996. Hyperresponsive B cells in CD22-deficient mice. *Science* 274: 798–801.

O'Rourke, L., R. Tooze, et al. 1997. Co-receptors of B lymphocytes. *Curr. Opin. Immunol*. 9(3): 324–329.

Otipoby, K. L., K. B. Andersson, et al. 1996. CD22 regulates thymus-independent responses and the lifespan of B cells. *Nature* 384: 634–637.

Peaker, C. J. G. and M. S. Neuberger 1993. Association of CD22 with the B cell antigen receptor. *Eur. J. Immunol*. 23: 1358–1363.

Pflugh, D. L., S. E. Maher, et al. 2002. Ly-6 superfamily members Ly-6A/E, Ly-6C, and Ly-6I recognize two potential ligands expressed by B lymphocytes. *J. Immunol*. 169: 5130–5136.

Poe, J. C., M. Fujimoto, et al. 2000. CD22 forms a quaternary complex with SHIP, Grb2 and Shc. A pathway for regulation of B lymphocyte antigen receptor-induced calcium flux. *J. Biol. Chem*. 275: 17420–17427.

Poe, J. C., K. M. Haas, et al. 2004a. Severely-impaired B lymphocyte proliferation, survival and induction of the c-Myc:Cullin 1 ubiquitin ligase pathway resulting from CD22 deficiency on the C57BL/6 genetic background. *J. Immunol*. 172: 2100–2110.

Poe, J. C., Y. Fujimoto, et al. 2004b. CD22 regulates B lymphocyte function in vivo through both ligand-dependent and ligand-independent mechanisms. *Nat. Immunol*. 5: 1078–1087.

Powell, L. D., D. Sgroi, et al. 1993. Natural ligands of the B cell adhesion molecule CD22β carry N-linked oligosaccharides with α-2,6-linked sialic acids that are required for recognition. *J. Biol. Chem*. 268: 7019–7027.

Powell, L. D. and A. Varki 1994. The oligosaccharide binding specificities of CD22-β, a sialic acid-specific lectin of B-cells. *J. Biol. Chem*. 269: 10628–10636.

Press, O. W., A. G. Farr, et al. 1989. Endocytosis and degradation of monoclonal antibodies targeting human B-cell malignancies. *Cancer Res*. 49: 4906–4912.

Press, O. W., J. Howell-Clark, et al. 1994. Retention of B-cell-specific monoclonal antibodies by human lymphoma cells. *Blood* 83: 1390–1397.

Press, O. W., J. P. Leonard, et al. 2001. Immunotherapy of non-Hodgkin's lymphomas. *Hematology* 2001: 221–240.

Razi, N. and A. Varki 1998. Masking and unmasking of the sialic acid-binding lectin activity of CD22 (Siglec-2) on B lymphocytes. *Proc. Natl Acad. Sci. USA* 95: 7469–7474.

Reth, M. 1989. Antigen receptor tail clue. *Nature* 338: 383–384.

Samardzic, T., D. Marinkovic, et al. 2002. Reduction of marginal zone B cells in CD22-deficient mice. *Eur. J. Immunol*. 32: 561–567.

Sato, S., P. J. Jansen, et al. 1997. CD19 and CD22 reciprocally regulate Vav tyrosine phosphorylation during B lymphocyte signaling. *Proc. Natl Acad. Sci., USA* 94: 13158–13162.

Sato, S., A. S. Miller, et al. 1996. CD22 is both a positive and negative regulator of B lymphocyte antigen receptor signal transduction: altered signaling in CD22-deficient mice. *Immunity* 5: 551–562.

Sato, S., J. M. Tuscano, et al. 1998. CD22 negatively and positively regulates signal transduction through the B lymphocyte antigen receptor. *Semin. Immunol*. 10: 287–297.

Sgroi, D., A. Varki, et al. 1993. CD22, a B cell-specific immunoglobulin superfamily member, is a sialic acid-binding lectin. *J. Biol. Chem*. 268: 7011–7018.

Shan, D. and O. W. Press 1995. Constitutive endocytosis and degradation of CD22 by human B cells. *J. Immunol*. 154: 4466–4475.

Sieber, T., D. Schoeler, et al. 2003. Selective internalization of monoclonal antibodies by B-cell chronic lymphocytic leukaemia cells. *Br. J. Hematol*. 121: 458–461.

Sieling, P. A., D. Jullien, et al. 1999. CD1 expression by dendritic cells in human leprosy lesions: Correlation with effective host immunity. *J. Immunol*. 162: 1851–1858.

Stamenkovic, I., D. Sgroi, et al. 1991. The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and α2,6 sialyltransferase, CD75, on B cells. *Cell* 66: 1133–1144.

Stein, R., E. Belisle, et al. 1993. Epitope specificity of the anti-(B cell lymphoma) monoclonal antibody, LL2. *Cancer Immunol. Immunother.* 37(5): 293–298.

Tedder, T. F. 1998. Response-regulators of B lymphocyte signaling thresholds provide a context for antigen receptor signal transduction. *Semin. Immunol*. 10: 259–265.

Tedder, T. F., J. Tuscano, et al. 1997. CD22, a B lymphocyte-specific adhesion molecule that regulates antigen receptor signaling. *Annu. Rev. Immunol*. 15: 481–504.

Torres, R. M., C. -L. Law, et al. 1992. Identification and characterization of the murine homologue of CD22, a B lymphocyte-restricted adhesion molecule. *J. Immunol*. 149: 2641–2649.

Tuscano, J. M., G. Harris, et al. 2003a. B lymphocytes contribute to autoimmune disease pathogenesis: Current trends and clinical implications. *Autoimmunity Rev.* 2: 101–108.

Tuscano, J. M., R. T. O'Donnell, et al. 2003b. The anti-CD22 ligand blocking antibody, HB22.7, has independent lymphomacidal properties, and augments the efficacy of ^{90}Y-DOTA-peptide-Lym-1 in lymphoma xenografts. *Blood* 101: 3641–3647.

van der Merwe, P. A., P. R. Crocker, et al. 1996. Localization of the putative sialic acid-binding site on the immunoglobulin superfamily cell-surface molecule CD22. *J. Biol. Chem*. 271: 9273–9280.

Wilson, G. L., C. H. Fox, et al. 1991. cDNA cloning of the B cell membrane protein CD22: A mediator of B–B cell interactions. *J. Exp. Med*. 173: 137–146.

Yohannan, J., J. Wienands, et al. 1999. Analysis of tyrosine phosphorylation-dependent interactions between stimulatory effector proteins and the B cell co-receptor CD22. *J. Biol. Chem*. 274(26): 18769–18776.

20 Galectins and Integrins in Pre-B Cell Development

M. Espeli, L. Gauthier, S.J.C. Mancini, F. Mourcin, B. Rossi, and C. Schiff

CONTENTS

20.1 INTRODUCTION

The most striking feature of the immune system in vertebrates is its ability to recognize a large diversity of pathogens by T and B lymphocytes that express on their surface specific receptors, the T cell receptor (TCR) and the B cell receptor (BCR), respectively. These cells derive from hematopoietic stem cells through discrete differentiation pathways that take place in the liver during embryonic and fetal life and in the bone marrow (BM) after birth. T cells mature in the thymus whereas B cells continue to differentiate in the BM. Lympho-stromal interactions in multiple microenvironments within the thymus and the BM have a crucial role in the regulation of lymphocyte development and selection. The genes encoding TCR and BCR antigen-binding domains are assembled during T and B cell differentiation by a series of gene rearrangements. Mechanisms of allelic exclusion prevent expression of multiple functional receptors on a given T or B cell.

In this chapter, we will report on the early steps of B cell development in the BM by focusing on the specific interactions between early B cell precursors and the BM microenvironment. We will describe how early B cells interact with BM stromal cells and the functional consequences of such interactions for B cell development. More precisely, we will concentrate on the mode of activation of the receptor expressed by early pre-B cells, called the pre-BCR, that depends on the secretion of galectin-1 by BM stromal cells and on specific interactions between the pre-BCR, galectin-1, and integrin members.

20.2 B CELL DEVELOPMENT IN THE BONE MARROW

B cell development in BM proceeds throughout life and generates immature B cells of widely diversified specificities, able to recognize foreign antigens but tolerant to self. The balance between cell autonomous mechanisms and signals from the microenvironment controls the various steps of B cell differentiation.

20.2.1 Mechanisms Leading to the Generation of B Cell Diversity

B cells develop from uncommitted progenitors through the combinatorial action of specific transcription factors including PU-1, E2A, EBF, and PAX5, which activate the B-cell gene expression program and the rearrangements of immunoglobulin (Ig) genes. B cell differentiation is characterized by the sequential expression of various cell surface markers and by the ordered rearrangement of IgH and IgL gene segments. Schematically, four successive stages have been described: Pro-B, pre-BI, pre-BII, and immature B (Figure 20.1A). Completion of V_H to D_H–J_H rearrangements at the

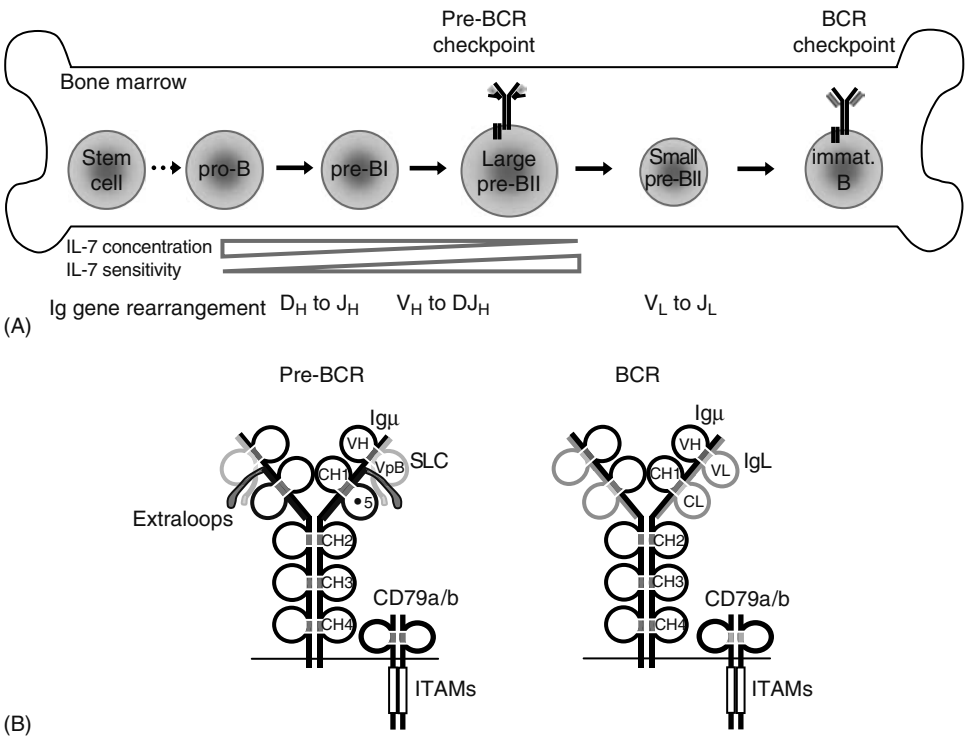


FIGURE 20.1 (See CD for color figure.) B cell differentiation in the bone marrow. (A) Developmental stages of murine B cell differentiation in the BM. The two checkpoints at which B cells are screened for functional Ig gene rearrangements are indicated. (Nomenclature used is from Melchers, F., ten Boekel, E., Seidl, T., Kong, X.C., Yamagami, T., Onishi, K., Shimizu, T., Rolink, A.G., and Andersson, J., *Immunol. Rev.*, 175, 33, 2000.) (B) Schematic structure of the pre-BCR and BCR complexes. The pre-BCR (left) is composed of two Igμ chains associated with two SLC and the CD79a/CD79b transducing module. The SLC is composed of the λ-like/λ5 and the VpreB chains. The COOH-terminal region of VpreB and the NH_2-terminal portion of λ-like, constitute the extraloop regions. (From Guelpa-Fonlupt, V., Bossy, D., Alzari, P., Fumoux, F., Fougereau, M., and Schiff, C., *Mol. Immunol.*, 31, 1099, 1994; Lanig, H., Bradl, H., and Jack, H.M., *Mol. Immunol.*, 40, 1263, 2004.) The BCR (right) is composed of two Igμ chains associated with two IgL chains and the CD79a/CD79b transducing module.

pre-BI stage leads to the production of the Igmu (Igμ) heavy chain. Some of these Igμ chains associate with the surrogate light chain (SLC), composed of the λ-like (also called λ5) and VpreB proteins, and with the signaling molecules CD79a and CD79b (Figure 20.1B), to form the pre-B cell receptor (pre-BCR). Once a functional V_H–D_H–J_H gene is obtained, the resulting Igμ heavy chain exerts a negative feedback, called allelic exclusion, on further rearrangement of the IgH locus, so that there is only one IgH chain that is expressed in a given cell [1]. In addition, large pre-BII cells expressing the pre-BCR proliferate and differentiate toward small pre-BII cells. At this step, the SLC expression, and consequently that of the pre-BCR, is downregulated, leading to the arrest of proliferation and to the initiation of IgL gene rearrangements. Finally, cells start to express the BCR, composed of the Igμ and IgL chains associated with CD79a and CD79b (Figure 20.1B) and become immature B cells. The recombination process for IgL chains is regulated in the same manner as for the IgH chains, so that only one IgL chain is expressed in a given cell (either κ or λ). Before emigrating from the BM to peripheral lymphoid organs, immature B cell are selected on the basis of their BCR specificity. Potential self-reactive B clones are eliminated by apoptosis (clonal deletion), modified by secondary rearrangements (receptor editing, V_H replacement), or rendered hyporesponsive (anergy) (reviewed in Ref. [2]). Thus, by ensuring the selection of functional rearrangements at the IgH and IgL loci, the pre-BCR and the BCR are crucial checkpoints of B cell differentiation.

20.2.2 Role of the Bone Marrow Microenvironment in B Cell Development

The microenvironment in the BM also plays a fundamental role in B cell development. Hematopoietic cells and lymphoid progenitors are located in cellular niches, in close contact with a highly organized three-dimensional microenvironment indispensable for their differentiation. Stromal cells are key constituents of the environment and they form a network in the intersinusoidal spaces of the bone cavity that extends from the subendosteal region to the sinusoids. Precursor B cells differentiate and proliferate in the interstices of this network in close contact with long processes of stromal cells [3]. Adhesion molecules, including CD44, selectins, and integrins control the interaction between B cell progenitors, the extracellular matrix (ECM) components, and the stromal cells. ECM components provide also a network in which soluble factors secreted by stromal cells and crucial for B cell development, such as CXCL12 (SDF-1) and IL-7, can be trapped.

CXCL12 sequestration into the fibronectin lattice activates a number of integrins, such as lymphocyte function-associated antigen (LFA-1), very late antigen (VLA-4), and VLA-5 [4] that mediate adhesion, migration, survival, and differentiation of the cells [5]. Integrins, which are heterodimeric transmembrane molecules consisting of α and β subunits, bind ECM components such as fibronectin and laminin, but also cellular receptors such as vascular cell adhesion molecule (VCAM-1). Integrin–ligand interactions participate in the adhesion of B cell progenitors to the BM environment and are critical for sustaining B lymphopoiesis [6,7]. VLA-4 and VLA-5 are particularly important in promoting pro-B cell adhesion and proliferation on BM stromal cells [8,9]. Moreover, α4, in association with β1 or β7 integrins, plays an important role in maintaining normal hematopoiesis [10]. The IL7 cytokine is also very important for mouse B cell development and was originally cloned from a stromal cell line as a factor supporting the growth of pre-B cells *in vitro* [11]. Latter, IL7 was shown not only to promote proliferation and survival of early B cell precursors in mice, but also to function as a differentiation factor to induce B cell development of hematopoietic precursors [12].

20.3 ROLE OF THE PRE-BCR DURING B CELL DEVELOPMENT IN THE BONE MARROW

As previously mentioned, the pre-BCR is composed of the Igμ, SLC, and CD79a/CD79b chains (Figure 20.1B). The 3-D structure of the pre-BCR has not been solved so far. As the SLC resemble a regular IgL chain, VpreB and λ-like/λ5 proteins are predicted to be comparable to variable and constant IgL domains, respectively [13–15]. Moreover, it was proposed that major interactions occur

between λ-like/λ5 and the IgCμ1 domain on one hand and between VpreB and the V_H domain on the other hand. Finally, the COOH-terminal region of VpreB (20 amino acid residues) and the NH_2-terminal portion of λ-like (50 amino acid residues), depicted as the extraloop (EL) regions (Figure 20.1B), are thought to "loop" out from the main SLC domain structure [14,15].

In mice lacking the transmembrane portion of Igμ chain (μMT mice), B cell development is blocked at the pre-BII cell stage [16]. V_H to D_H–J_H rearrangements can occur, but in the absence of membrane-bound pre-BCR, and consequently in absence of cell surface expression of the pre-BCR, cells do not proliferate and Igμ chains are not allelically excluded [16,17]. Surprisingly, in mice lacking the SLC, B cell development is not totally blocked although pre-BII cells do not proliferate and Igμ chains are still allelically excluded [18–20]. These data suggest that the pre-BCR mediates two separate functions, i.e., Igμ chain allelic exclusion and pre-BII cell differentiation and expansion. The allelic exclusion process is totally dependent upon Igμ membrane deposition but does not involve the SLC. In contrast, the SLC seems to be essential for pre-B cells to enter into cell cycle and to differentiate. In addition, Igμ chains that fail to pair with SLC lose the ability to be expressed at the cell surface and are counterselected [21], indicating that the SLC is also implicated in the selection of an early Igμ repertoire.

These functions require the generation of pre-BCR intracellular signals involving the phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) localized on the cytoplasmic tails of the CD79a/CD79b molecules, by tyrosine kinases such as Lyn and Syk. In response to this phosphorylation, two main signaling pathways are activated. The first is dependent on the tyrosine kinase Btk and the adaptor protein SLP-65 and gives rise to the downregulation of the SLC, the IL7 receptor, and the RAG genes. This pathway stops the proliferation of pre-BII cells and drives their development toward immature B cells. The second pathway is Btk/SLP-65-independent and is implicated in the proliferation and the survival of pre-BII cells through the Bcl-2 and ERK pathways [22]. The pre-BCR induces high proliferation signals, implicating a fine regulation of its expression at the pre-BII cell surface and a tight control of the signals that are delivered within these cells. In line with this, Btk and SLP-65 have been shown to act as tumor suppressor molecules [23].

20.4 PRE-BCR ACTIVATION IS TRIGGERED BY DIRECT BINDING TO GALECTIN-1 ANCHORED TO INTEGRINS

Although the signal pathways downstream of the pre-BCR are well characterized, the mechanisms implicated in pre-BCR activation are not completely understood and both constitutive [24,25] and ligand-induced [26–28] signaling have been described. It has been shown that a fraction of the pre-BCR is associated with raft structures, leading to its constitutive activation, and that pre-BCR engagement enhances this association, resulting in calcium flux [29] and changes in protein tyrosine phosphorylation [24]. In support of the existence of external pre-BCR ligands, it was demonstrated that the murine pre-BCR could specifically interact with stromal cell-associated heparan sulfates (HS) [27]. Moreover, we reported the identification of the S-type lectin galectin-1 (GAL1), expressed by stromal cells, as a human pre-BCR ligand [26].

Galectins are a large family of calcium-independent S-type lectins, widely conserved in animals, plants, and microorganisms and with specificity for galactose derivatives. Typically, galectins bind to type I Galβ1,3 GlcNAc or type II Galβ1,4 GlcNAc units, often with preference for one of these structures and have increased affinity for lactosamine polymers. To date, 15 galectins have been identified in mammals [30,31]. They share at least one carbohydrate recognition domain (CRD) of about 135 amino acids. GAL1 consists of a single CRD and occurs naturally as a noncovalent homodimer formed by extended β-sheet interactions across the two monomeric subunits [32]. In adult tissues, GAL1 is abundantly expressed in many cell types such as skeletal, smooth, and cardiac muscle and from other cells of mesenchymal origin. High GAL1 expression is observed in sensory, motor, and olfactory neurons and in stromal cells of hematopoietic and lymphoid organs [33]. Functional

studies have implicated galectins in cell growth, differentiation, and apoptosis, in addition to cell adhesion, chemoattraction, and cell migration. Galectins can either promote or suppress cell growth, depending on the cells, the type of receptors and the doses used [34].

20.4.1 Pre-BCR and GAL1 Interact by a Direct Protein–Protein Contact

We have first demonstrated that soluble pre-BCRs, or the SLC alone, are able to interact specifically with different stromal cell lines derived from normal tissues that sustain B cell differentiation. The molecule involved in this interaction was identified by preparative biochemistry followed by mass spectrometry, as GAL1, produced by stromal cell lines [26]. By surface plasmon resonance analysis, SLC was shown to interact with recombinant GAL1 by direct protein–protein interactions, with an affinity constant of $2 \times 10^6\ M^{-1}$. It was also determined that one SLC molecule binds to one GAL1 homodimer. Moreover, preincubation of the SLC with an antiserum against the 50 NH_2-terminal amino acids of λ-like blocks the binding of SLC to the immobilized GAL1, suggesting that the NH_2 extra loop (EL) of λ-like is the major GAL1-binding site on SLC. These contacts have been confirmed by isothermal titration calorimetry (ITC), as we demonstrated that the 50 amino acid–long λ-like EL peptide specifically binds to GAL1. The EL of λ-like contains seven positively charged residues conserved between human and mice, which confer very high basic properties to this region. By contrast, determination of the electrostatic potential of the GAL1 homodimer shows that GAL1 is largely acidic, with the exception of the basic β-galactoside-binding sites. Thus, we postulate that the conserved positively charged λ-like EL amino acids could interact with the negatively charged regions of the GAL1 homodimer, independently of the two sugar pockets that account, as presented below, for the binding of GAL1 to glycosylated integrin counterreceptors [35,36]. The direct SLC binding to GAL1, with a binding site independent of the CRD carbohydrate pocket, is an interesting observation since extracellular galectins most frequently bind to their targets in a sugar-dependent manner. However, direct protein–protein interactions have been reported for various intracellularly expressed galectins [37,38].

20.4.2 Formation of a Lattice between Pre-BCRs, GAL1, and Integrins when Pre-B Cells Interact with Stromal Cells

SLC/GAL1 interactions have also been analyzed in the context of pre-B and stromal cells contact. During pre-B/stromal cell cocultures, pre-B cells establish close interactions with stromal cells and a clear and remarkable translocation of the pre-BCR is observed at the contact zone between the two cells. Labeling of GAL1 shows a patchy distribution all over the stromal cell surface, whereas GAL1 is homogeneously distributed at the cell surface of pre-B cells not in contact with stromal cells and concentrates at the contact zone between interacting pre-B and stromal cells. Thus, the stromal cell-derived GAL1 is mainly polarized at the contact zone between pre-B and stromal cells, indicating that almost all GAL1 counterreceptors are relocalized in this area. Double GAL1 and SLC labeling shows that GAL1 and the pre-BCR colocalize and are polarized at the contact area between the two cells, with a molecular surface organization presenting the characteristics of a synapse. Pre-BCR relocalization is dependent upon GAL1 anchoring to glycosylated counterreceptors expressed by stromal and pre-B cells and we showed that pre-BCR-GAL1–counterreceptor complexes completely relocalize at the contact zone between the two cells. This "immune developmental synapse" is accompanied by the initiation of intracellular tyrosine kinase activity and signal transduction from the pre-BCR, as revealed by the phosphorylation of CD79a ITAMs [26].

Recently, we identified α4β1 (VLA-4), α5β1 (VLA-5), and α4β7 integrins by biochemical and confocal analysis as major GAL1 glycosylated counterreceptors involved in synapse formation. Integrins in the synapse are present in their active conformation, as revealed using a conformational-specific anti-β1 integrin mAb. GAL1 has been already shown to bind via its CRD to members of the

integrin family, such as α7β1 [39] or β1 integrin [40]. We also identified ADAM15, fibronectin, and possibly VCAM1 as integrin ligands expressed by stromal cells and present at the intercellular junction between pre-B and stromal cells. Confocal microscopy analysis revealed that pre-B cell integrins and their stromal cell ligands, together with the pre-BCR and GAL1, form a homogeneous lattice at the contact area between pre-B and stromal cells. Finally, we demonstrated that cross-linking of β1 integrins expressed by pre-B cells, in the presence of GAL1, is necessary and sufficient for driving pre-BCR recruitment into the synapse. Pre-BCR clustering is followed by Lyn recruitment to CD79a and by Lyn and CD79a phosphorylation, leading to initiation of pre-BCR signaling [35].

20.5 FUNCTIONAL ROLE OF THE PRE-BCR/STROMAL CELL SYNAPSE

We proposed that formation of the pre-B/stromal cell synapse proceeds in two steps (Figure 20.2): (1) the stromal cell-derived GAL1 binds to glycosylated pre-B cell integrins and to pre-BCRs and (2) the active relocalization of pre-B cell integrins, mediated by their interactions with stromal cell ligands, drives pre-BCRs relocalization into the synapse. Our data suggest that the lattice generated by integrin-GAL1–pre-BCR interactions could reinforce the adhesion of B cell precursors on stromal cells and may regroup the different pre-BCRs into the synapse in order to reach the activation

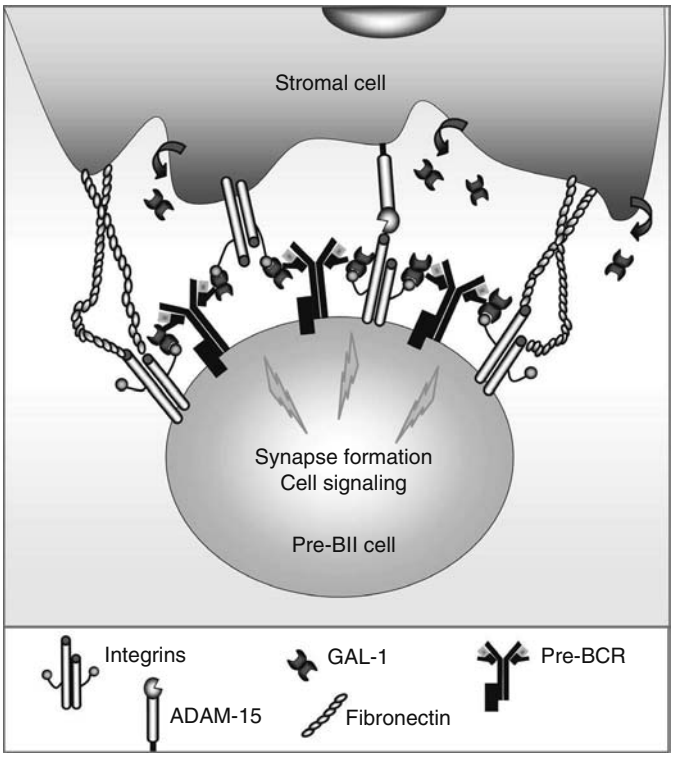


FIGURE 20.2 (See color insert following blank page 170. Also see CD for color figure.) Model of pre-BCR signaling. GAL1 secreted by stromal cells is captured by relocalizing pre-B cell integrins interacting with their ligands on stromal cells and by pre-BCRs. The active integrin relocalization in the presence of GAL1 drives the pre-BCR into the pre-B/stromal cell synapse, leading to formation of a homogeneous lattice and initiation of pre-BCR signaling. Synapse formation and pre-BCR clustering are necessary for pre-B cell differentiation and proliferation. (From Gauthier, L., Rossi, B., Roux, F., Termine, E., and Schiff, C., *Proc. Natl Acad. Sci. USA*, 99, 13014, 2002; Rossi, B., Espeli, M., Schiff, C., and Gauthier, L., *J. Immunol*., 177, 796, 2006.)

threshold necessary to initiate pre-B cell development programs. Indeed, in SLC mutant mice, the pre-BII population is deeply decreased [20], suggesting that pre-BCR clustering and activation could be essential for the cycling initiation of large pre-BII cells and for promoting large to small pre-BII cell differentiation.

In humans and mice, pre-BCR expression and activation are essential for both pre-B cell differentiation and proliferation, but differences are observed between the two species. In the mouse, as mentioned above, Jäck and coworkers demonstrated that HS interact with the pre-BCR [27] but in humans, our preliminary experiments indicate that HS are not pre-BCR ligands (unpublished observations). In mice, the importance of IL-7 for pre-B cell proliferation leads to the hypothesis that HS, which bind to the pre-BCR and also to IL-7, may participate in the recruitment of cells expressing the pre-BCR into specialized IL-7-expressing BM niches. In this species, synapse formation between pre-B and stromal cells is also dependent of the presence of GAL1 and of glycosylated integrins, as we showed using mouse pre-B cell lines but also pre-BII cells isolated from normal BM ([36] and unpublished observations). Thus in mice, HS and GAL1 could act synergistically to promote pre-B cell differentiation and proliferation through the pre-BCR and the IL-7R, respectively. As human pre-B cells are not dependent on IL-7 for B cell differentiation/proliferation, requirement of HS could be less important for their development. Synapse formation, which leads to pre-BCRs clustering and activation, probably conditions both pre-B cell differentiation and proliferation [36].

20.6 STROMAL CELL MICROENVIRONMENT FOR PRE-B CELL DIFFERENTIATION AND PROLIFERATION

In addition to providing a physical support, BM stromal cells secrete soluble factors (IL-7, CXCL12, SCF, Flt-3L), which regulate precursor B cell growth, maturation, and survival [11,41,42]. A recent report by Tokoyoda et al. has shown that the spatial repartition of stromal and early B cell progenitors follows a strict and coordinated regulation [43]. CXCL12-expressing cells represent a small population of stromal cells, scattered throughout the BM and located at some distance from the stromal cells expressing IL-7. Hematopoietic progenitors are attached to the processes of CXCL12-expressing cells and early B cell progenitors adjoin their cellular bodies. When pre-BI cells, which require IL-7 for cell proliferation develop, they move away from CXCL12-producing cells and are found adjacent to IL-7-expressing cells. It remains to be determined if IL-7-expressing stromal cells secrete also GAL1 or if neighboring IL-7 and GAL1-expressing cells coexist. As mentioned before, at least for the mouse system, IL-7 and GAL1-expressing cells may constitute a favorable BM niche indispensable for pre-BII cell proliferation and selection.

For B cell precursors, the stimulation and the transducing mechanisms that control integrin activation are poorly understood. Integrins could favor the activation of the pre-BCR by inducing its clustering, but also by decreasing the threshold of pre-BCR activation. Indeed, for mature B cells expressing a functional BCR, it was shown that recruitment of the LFA1 integrin into the synapse formed between B cells and antigen-presenting cells facilitated BCR signaling by diminishing the threshold of BCR-dependent signaling [44]. Pre-BII and stromal cell adhesion could also be favored by the specialized cellular environment. Indeed, BCR stimulation was reported to activate integrins, leading to enhanced adhesion by an inside-out phenomenon dependent on Btk and PLCγ2 expression [45]. As integrin recruitment in the pre-B/stromal cell synapse in the presence of GAL1 conditions pre-BCR activation, we postulate that such activation could, in return, favor integrin activation, thus contributing to enhance the adhesion of pre-BII cells into specialized BM stromal cell niches.

20.7 CONCLUSION AND FUTURE DIRECTIONS

When pre-B cells adhere to BM stromal cells, GAL1 creates a link between pre-B cell integrins and pre-BCRs, leading to lattice formation, pre-BCR clustering, and activation. The precise role of such

activation and the implication of GAL1 in this process *in vivo* remain to be determined. The identification of BM-derived pre-BCR ligand highlights the role played by stromal cells in the early steps of B cell differentiation. Moreover, as it was recently demonstrated for hematopoietic stem cells, cellular niches within the BM deliver essential cytokines/chemokines and concentrate molecules essential to induce close contacts with B cell progenitors and to sustain their development.

For the evolutionary point of view, Pelanda et al. have proposed that the pre-BCR could have evolved from an autoreactive BCR able to generate a cellular compartment devoted to efficient IgL rearrangements, an equivalent of the present receptor editing compartment, in order to improve the efficiency of antibody diversification [46]. The identification of GAL1, a molecule widely expressed and highly conserved between species, as a pre-BCR ligand could be in agreement with this hypothesis. Interestingly, the human IgL lambda, VpreB, λ-like, and GAL1 genes are located on the same genetic region, on chromosome 22 between q11 and q13 bands [47,48] and thus could have been subjected to concomitant evolutionary events.

REFERENCES

1. Alt, F.W., T.K. Blackwell, and G.D. Yancopoulos. 1987. Development of the primary antibody repertoire. *Science* 238:1079–1087.
2. Meffre, E., R. Casellas, and M.C. Nussenzweig. 2000. Antibody regulation of B cell development. *Nature Immunol* 1:379–385.
3. Jacobsen, K., J. Tepper, and D.G. Osmond. 1990. Early B-lymphocyte precursor cells in mouse bone marrow: Subosteal localization of B220+ cells during postirradiation regeneration. *Exp Hematol* 18:304–310.
4. Peled, A., O. Kollet, T. Ponomaryov, I. Petit, S. Franitza, V. Grabovsky, M.M. Slav, A. Nagler, O. Lider, R. Alon, D. Zipori, and T. Lapidot. 2000. The chemokine SDF-1 activates the integrins LFA-1, VLA-4, and VLA-5 on immature human CD34(+) cells: Role in transendothelial/stromal migration and engraftment of NOD/SCID mice. *Blood* 95:3289–3296.
5. Hynes, R.O. 1992. Integrins: Versatility, modulation, and signaling in cell adhesion. *Cell* 69:11–25.
6. McGinnes, K., V. Quesniaux, J. Hitzler, and C. Paige. 1991. Human B-lymphopoiesis is supported by bone marrow-derived stromal cells. *Exp Hematol* 19:294–303.
7. Jacobsen, K. and D.G. Osmond. 1990. Microenvironmental organization and stromal cell associations of B lymphocyte precursor cells in mouse bone marrow. *Eur J Immunol* 20:2395–2404.
8. Qiang, Y.W., M. Kitagawa, M. Higashi, G. Ishii, C. Morimoto, and K. Harigaya. 2000. Activation of mitogen-activated protein kinase through alpha5/beta1 integrin is required for cell cycle progression of B progenitor cell line, Reh, on human marrow stromal cells. *Exp Hematol* 28:1147–1157.
9. Hahn, B.K., D. Piktel, L.F. Gibson, and K.S. Landreth. 2000. Hematopoiesis: The role of stromal integrin interactions in pro-B cell proliferation. *Hematology* 5:153–160.
10. Arroyo, A.G., J.T. Yang, H. Rayburn, and R.O. Hynes. 1999. Alpha4 integrins regulate the proliferation/differentiation balance of multilineage hematopoietic progenitors in vivo. *Immunity* 11:555–566.
11. Sudo, T., M. Ito, Y. Ogawa, M. Iizuka, H. Kodama, T. Kunisada, S. Hayashi, M. Ogawa, K. Sakai, and S. Nishikawa. 1989. Interleukin 7 production and function in stromal cell-dependent B cell development. *J Exp Med* 170:333–338.
12. von Freeden-Jeffry, U., P. Vieira, L.A. Lucian, T. McNeil, S.E. Burdach, and R. Murray. 1995. Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *J Exp Med* 181: 1519–1526.
13. Melchers, F. 1999. Fit for life in the immune system? Surrogate L chain tests H chains that test L chains. *Proc Natl Acad Sci USA* 96:2571–2573.
14. Guelpa-Fonlupt, V., D. Bossy, P. Alzari, F. Fumoux, M. Fougereau, and C. Schiff. 1994. The human pre-B cell receptor: Structural constraints for a tentative model of the pseudo-light (psi L) chain. *Mol Immunol* 31:1099–1108.
15. Lanig, H., H. Bradl, and H.M. Jack. 2004. Three-dimensional modeling of a pre-B-cell receptor. *Mol Immunol* 40:1263–1272.
16. Kitamura, D., J. Roes, R. Kuhn, and K. Rajewsky. 1991. A B cell deficient mouse by targeted disruption of the membrane exons of the immunoglobulin μ chain gene. *Nature* 350:423–426.
17. Kitamura, D. and K. Rajewsky. 1992. Targeted disruption of mu chain membrane exon causes loss of heavy-chain allelic exclusion. *Nature* 356:154–156.

18. Mundt, C., S. Licence, T. Shimizu, F. Melchers, and I.L. Martensson. 2001. Loss of precursor B cell expansion but not allelic exclusion in VpreB1/VpreB2 double-deficient mice. *J Exp Med* 193:435–446.
19. ten Boekel, E., F. Melchers, and A.G. Rolink. 1998. Precursor B cells showing H chain allelic inclusion display allelic exclusion at the level of pre-B cell receptor surface expression. *Immunity* 8:199–207.
20. Martensson, I.L., A. Rolink, F. Melchers, C. Mundt, S. Licence, and T. Shimizu. 2002. The pre-B cell receptor and its role in proliferation and Ig heavy chain allelic exclusion. *Semin Immunol* 14:335–342.
21. ten Boekel, E., F. Melchers, and A.G. Rolink. 1997. Changes in the V(H) gene repertoire of developing precursor B lymphocytes in mouse bone marrow mediated by the pre-B cell receptor. *Immunity* 7:357–368.
22. Hendriks, R.W., and R. Kersseboom. 2006. Involvement of SLP-65 and Btk in tumor suppression and malignant transformation of pre-B cells. *Semin Immunol* 18:67–76.
23. Flemming, A., T. Brummer, M. Reth, and H. Jumaa. 2003. The adaptor protein SLP-65 acts as a tumor suppressor that limits pre-B cell expansion. *Nat Immunol* 4:38–43.
24. Guo, B., R.M. Kato, M. Garcia-Lloret, M.I. Wahl, and D.J. Rawlings. 2000. Engagement of the human pre-B cell receptor generates a lipid raft-dependent calcium signaling complex. *Immunity* 13:243–253.
25. Rolink, A.G., T. Winkler, F. Melchers, and J. Andersson. 2000. Precursor B cell receptor-dependent B cell proliferation and differentiation does not require the bone marrow or fetal liver environment. *J Exp Med* 191:23–32.
26. Gauthier, L., B. Rossi, F. Roux, E. Termine, and C. Schiff. 2002. Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. *Proc Natl Acad Sci USA* 99:13014–13019.
27. Bradl, H., J. Wittmann, D. Milius, C. Vettermann, and H.M. Jack. 2003. Interaction of murine precursor B cell receptor with stroma cells is controlled by the unique tail of lambda 5 and stroma cell-associated heparan sulfate. *J Immunol* 171:2338–2348.
28. Ohnishi, K. and F. Melchers. 2003. The nonimmunoglobulin portion of lambda5 mediates cell-autonomous pre-B cell receptor signaling. *Nat Immunol* 4:849–856.
29. Bossy, D., J. Salamero, D. Olive, M. Fougereau, and C. Schiff. 1993. Structure, biosynthesis, and transduction properties of the human mu-psi L complex: Similar behavior of preB and intermediate preB-B cells in transducing ability. *Int Immunol* 5:467–478.
30. Hughes, R.C. 1997. The galectin family of mammalian carbohydrate-binding molecules. *Biochem Soc Trans* 25:1194–1198.
31. Houzelstein, D., I.R. Goncalves, A.J. Fadden, S.S. Sidhu, D.N. Cooper, K. Drickamer, H. Leffler, and F. Poirier. 2004. Phylogenetic analysis of the vertebrate galectin family. *Mol Biol Evol* 21:1177–1187.
32. Bourne, Y., B. Bolgiano, D.I. Liao, G. Strecker, P. Cantau, O. Herzberg, T. Feizi, and C. Cambillau. 1994. Crosslinking of mammalian lectin (galectin-1) by complex biantennary saccharides. *Nat Struct Biol* 1:863–870.
33. Chiariotti, L., P. Salvatore, G. Benvenuto, and C.B. Bruni. 1999. Control of galectin gene expression. *Biochimie* 81:381–388.
34. Yang, R.Y. and F.T. Liu. 2003. Galectins in cell growth and apoptosis. *Cell Mol Life Sci* 60:267–276.
35. Rossi, B., M. Espeli, C. Schiff, and L. Gauthier. 2006. Clustering of pre-B cell integrins induces galectin-1-dependent pre-B cell receptor relocalization and activation. *J Immunol* 177:796–803.
36. Espeli, M., B. Rossi, S.J. Mancini, P. Roche, L. Gauthier, and C. Schiff. 2006. Initiation of pre-B cell receptor signaling: Common and distinctive features in human and mouse. *Semin Immunol* 18:56–66.
37. Cooper, D.N. 2002. Galectinomics: finding themes in complexity. *Biochim Biophys Acta* 1572:209–231.
38. Hsu, D.K. and F.T. Liu. 2004. Regulation of cellular homeostasis by galectins. *Glycoconj J* 19: 507–515.
39. Gu, M., W. Wang, W.K. Song, D.N. Cooper, and S.J. Kaufman. 1994. Selective modulation of the interaction of alpha 7 beta 1 integrin with fibronectin and laminin by L-14 lectin during skeletal muscle differentiation. *J Cell Sci* 107(Pt 1):175–181.
40. Moiseeva, E.P., B. Williams, A.H. Goodall, and N.J. Samani. 2003. Galectin-1 interacts with beta-1 subunit of integrin. *Biochem Biophys Res Commun* 310:1010–1016.
41. Egawa, T., K. Kawabata, H. Kawamoto, K. Amada, R. Okamoto, N. Fujii, T. Kishimoto, Y. Katsura, and T. Nagasawa. 2001. The earliest stages of B cell development require a chemokine stromal cell-derived factor/pre-B cell growth-stimulating factor. *Immunity* 15:323–334.
42. Lyman, S.D. and S.E. Jacobsen. 1998. c-kit ligand and Flt3 ligand: Stem/progenitor cell factors with overlapping yet distinct activities. *Blood* 91:1101–1134.
43. Tokoyoda, K., T. Egawa, T. Sugiyama, B.I. Choi, and T. Nagasawa. 2004. Cellular niches controlling B lymphocyte behavior within bone marrow during development. *Immunity* 20:707–718.

44. Carrasco, Y.R., S.J. Fleire, T. Cameron, M.L. Dustin, and F.D. Batista. 2004. LFA-1/ICAM-1 interaction lowers the threshold of B cell activation by facilitating B cell adhesion and synapse formation. *Immunity* 20:589–599.
45. Spaargaren, M., E.A. Beuling, M.L. Rurup, H.P. Meijer, M.D. Klok, S. Middendorp, R.W. Hendriks, and S.T. Pals. 2003. The B cell antigen receptor controls integrin activity through Btk and PLCgamma2. *J Exp Med* 198:1539–1550.
46. Pelanda, R., S. Schwers, E. Sonoda, R.M. Torres, D. Nemazee, and K. Rajewsky. 1997. Receptor editing in a transgenic mouse model: Site, efficiency, and role in B cell tolerance and antibody diversification. *Immunity* 7:765–775.
47. Mattei, M.G., F. Fumoux, N. Roeckel, M. Fougereau, and C. Schiff. 1991. The human pre-B-specific lambda-like cluster is located in the 22q11.2-22q12.3 region, distal to the IgC lambda locus. *Genomics* 9:544–546.
48. Baldini, A., T. Gress, K. Patel, R. Muresu, L. Chiariotti, P. Williamson, Y. Boyd, I. Casciano, V. Wells, C.B. Bruni, et al. 1993. Mapping on human and mouse chromosomes of the gene for the β-galactoside-binding protein, an autocrine-negative growth factor. *Genomics* 15:216–218.
49. Melchers, F., E. ten Boekel, T. Seidl, X.C. Kong, T. Yamagami, K. Onishi, T. Shimizu, A.G. Rolink, and J. Andersson. 2000. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunol Rev* 175:33–46.

21 Golgi N-Glycan Processing and Galectin Functions

Ken S. Lau, Ivan R. Nabi, Michael Demetriou, and James W. Dennis

CONTENTS

21.1 INTRODUCTION

Galectins (Gal) are a family of proteins characterized by a ~130 amino acid carbohydrate recognition domain (CRD) with binding specificity for β-galactosides [1]. The galectins are widely expressed in vertebrate tissues, and galectin family expansion with the evolution of multicellular eukaryotes is also accompanied by increases in complex N-glycan structural diversity [2,3]. Mammalian galectins are classified into three groups: The single CRD group (Gal-1, -2, -5, -7, -10, -11, and -13), the tandem-repeat group (Gal -4, -6, -8, -9, and -12), and the chimeric group that includes only Gal-3 with a single CRD and a ~120 amino acid N-terminal proline- and glycine-rich domain with homology to collagen α1 [4].

21.2 GALECTINS BIND *N*-ACETYLLACTOSAMINE, COMMON TO MANY GLYCOPROTEINS

Galectin CRDs bind to the minimal *N*-acetyllactosamine disaccharide, requiring the 4-OH and 6-OH of galactose, and 3-OH of GlcNAc [5]. Binding affinities for *N*-acetyllactosamine range from 30 μM to >1 mM for different galectins [6]. Binding of Gal-1, -3, and -8 to the different CHO glycosylation mutants revealed that complex N-glycans on glycoproteins at the cell surface are the major ligands for these galectins [7]. Addition of the *N*-acetyllactosamine units begins with GlcNAc branching, catalyzed in the medial Golgi by *N*-acetylglucosaminyltransferases (encoded by Mgat1, Mgat2, Mgat3, Mgat4a/b, and Mgat5) [8]. With the exception of the product of GlcNAc-TIII (Mgat3), the GlcNAc antennae are extended with β-1,4 galactose to form *N*-acetyllactosamine

antennae. *N*-acetyllactosamine-branched N-glycans are further elongated to varying degrees with poly(*N*-acetyllactosamines) in the *trans* Golgi, producing a series of complex N-glycans with differing affinities for galectins [6]. Gal-3, Gal-8, and Gal-9 display increasing affinities for N-glycans as a function of *N*-acetyllactosamine branching and extension, reaching <1 μM. However, increases in *N*-acetyllactosamine content of N-glycans only modestly enhance binding to Gal-1 [6].

Gal-3 forms multimers with positive cooperativity upon binding to synthetically tethered glycans, suggesting that glycan densities and affinities on glycoproteins may be the major driving force of Gal-3 oligomerization [9,10]. The N-terminal domain of Gal-3 is susceptible to proteolysis by metalloproteinases [11], which has been shown to interfere with multimer formation. Gal-3 is subjected to additional levels of regulation, notably phosphorylation at Ser^6 reduces binding to ligands [12]. Gal-3 secretion occurs by vesicular release from the cytoplasm, independent of the endoplasmic reticulum (ER) and Golgi secretory networks [13]. This suggests that Gal-3 binding to glycoprotein occurs after their export to the cell surface by physically independent pathways. In polarized epithelial cells, Gal-3 promotes apical localization of glycoproteins in a lipid raft-independent manner [14].

The addition of recombinant galectins to cells in culture has been shown to perturb many phenotypes, which in many cases, are dependent on CRD function as shown by the negation of these effects by competition with lactose [15]. Galectins have been implicated in cell cycle progression, apoptosis, T cell activation, signal transduction, tumor progression, and cell migration [15]. Galectin-8-coated substratum stimulates integrin-mediated spreading, Erk/PI3K signaling, and F-actin reorganization in a similar fashion to that of cells on fibronectin, while soluble Gal-8 binds to β1 integrin and antagonizes adhesion and signaling [16]. Exogenous Gal-8, requiring its CRD, induces neutrophil adhesion, metalloproteinase-9, and superoxide production comparable to stimulation with fMLP peptide [17]. Consistent with the stoichiometry required for multimer formation, Gal-3 added to mammary carcinoma cells promotes focal adhesion signaling and cell motility with a tight concentration optimum [18]. Galectins bind a variety of glycoproteins on the surface of various cell types including integrin β1, laminin, fibronectin, lysosomal associated membrane glycoprotein [LAMPs], and signaling receptors [13,19]. Although *N*-acetyllactosamine is a common structure in N- and O-linked glycans on glycoproteins, galectin family members show differences in affinities for glycans and capacity to multimerize, which may result in qualitatively different multimeric complexes with distinct glycoprotein constituents [20]. In this regard, exogenous Gal-1 and Gal-3 show binding preferences for different glycoproteins and can induce different death pathways in T cells. CD45 and CD71, but not CD29 and CD43, are implicated in galectin-3-induced T cell death, while CD7 is required for galectin-1-induced death [21]. The mechanism by which glycoprotein binding specificity is determined by different galectins requires further investigation.

21.3 DYNAMICS OF GALECTIN–GLYCOPROTEIN INTERACTIONS AT THE CELL SURFACE

Since complex N-glycans are found on most cell surface glycoproteins, their relative affinities for galectins may be a critical upstream regulator of galectin functions. Galectins with 2 CRDs or single CRD and self-association domains induce formation of molecular lattices by interacting with multivalent ligands [20]. Transmembrane glycoproteins commonly have multiple N-glycans (e.g., epithelial growth factor receptor [EGFR] has 8/11 sites occupied), and therefore, can act as multivalent ligands for galectins [22,23]. N-glycan processing generates a mixed population of N-glycans at each N-X-S/T position due to the inherent inefficiency of the Golgi pathway. The GlcNAc-branching enzymes display decreasing affinities for common donor UDP-GlcNAc and lower concentration levels stepwise down the pathway (Figure 21.1). Secondly, transit of glycoproteins through the Golgi limits their exposure time to the enzymes, and thus, reactions are incomplete. As a result, the kinetic and physical properties of GlcNAc branching and extension are major sources of N-glycan structural heterogeneity on cell surface glycoproteins. Glycoprotein isoforms with different combinations of N-glycans at each site, termed glycoforms, have distinct binding kinetics for galectin.

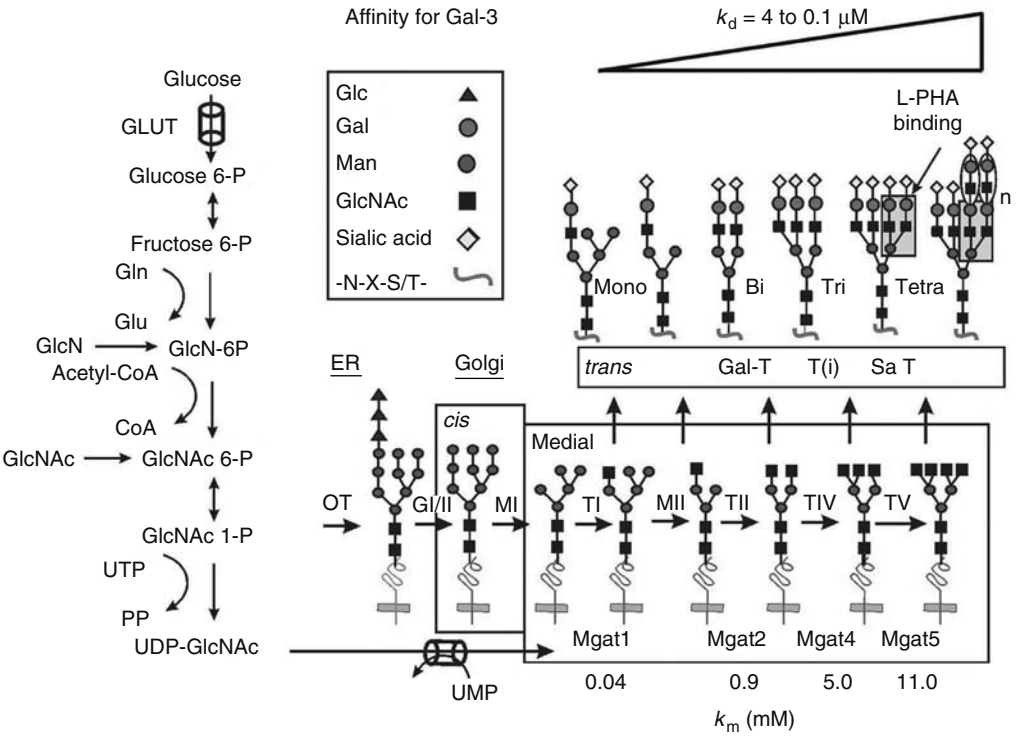


FIGURE 21.1 (See CD for color figure.) The hexosamine pathway and N-glycan branching Metabolite flux: glucose (Glc), glutamine (Gln), acetyl-CoA, glucosamine (GlcN), and GlcNAc through the hexosamine pathway supplies UDP-GlcNAc to the Golgi. $Glc_3Man_9GlcNAc_2$-pp-dolichol is the donor for the initial substitution of N-X-S/T motifs in the ER. The medial Golgi *N*-acetylglucosaminyltransferases, designated here by their gene names: Mgat1, Mgat2, Mgat4, Mgat5 generate branched N-glycans that display a range of affinities for galectins. The K_m values for UDP-GlcNAc reported for Mgat1, Mgat2, Mgat4, and Mgat5 are indicated. The product of Mgat3 is not extended with Gal (not a ligand for galectin), and has been omitted from the diagram. The gray box indicates the minimal galectin requirement for high-affinity galectin binding. (Adapted from Lau, K., Partridge, E.A., Silvescu, C.I., Grigorian, A., Pawling, J., Reinhold, V.N., Demetriou, M., and Dennis, J.W., *Cell*, 129, 123, 2007.)

Due to the lack of allosteric interaction between individual N-glycans, the on-rate of glycoprotein-to-lattice binding can be estimated as proportional to Σk_{on} of all N-glycans on the glycoprotein. Once a glycoprotein is bound to a highly multivalent lattice, the effective off-rate ($k_{off\text{-}effective}$) is lower than expected from monomeric interactions in solution (Figure 21.2A) [3]. Allovalency, also termed avidity in the context of antibody–antigen interactions, occurs because of a high local concentration of potential interaction sites available to recently unbound ligands. The high amount of local interaction sites results in higher probabilities of rebinding than can be expected from independent monovalent interactions with the same number of binding sites. Thus, the overall avidities ($k_{on}/k_{off\text{-}effective}$) of glycoforms for lattice binding are enhanced many fold by multivalent galectin interaction with multivalent glycoproteins.

Complex N-glycans facing the solvent display rotational freedom about the glycosidic linkages of the trimannosyl-core and are generally not restrained by the protein [24], which allows a range of geometries for galectin cross-linking. Transmembrane glycoproteins are confined in a planar manner, enhancing the amount of ligand available for high-density cross-linking by galectins. At critical concentrations of multimeric ligand and optimal cross-linking by Gal-3 results in Gal-3 pentamers, a nonsymmetrical arrangement that permits diverse connections and lattice geometries [10]. Considering the situation where a newly exported glycoprotein encounters an existing galectin lattice on the cell surface, affinity (k_{on}) and avidity (affinity enhancement due to rebinding $k_{off\text{-}effective}$) are dependent on

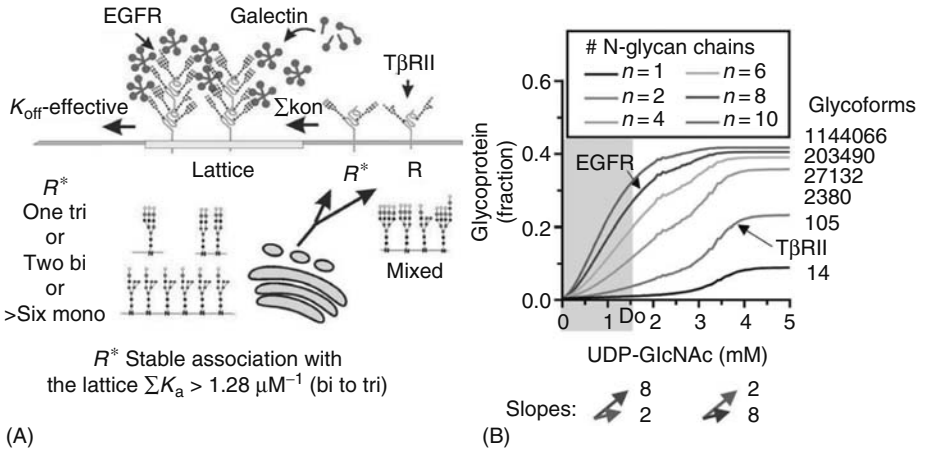


FIGURE 21.2 (See CD for color figure.) N-glycans, glycoprotein glycoforms, and association with the galectin lattice (A) N-glycan branching of a glycoprotein cooperates with N-glycan multiplicity to form stable association with the galectin lattice. Threshold for stable association to the galectin lattice are obtained with either a single high-affinity N-glycan (i.e., single triantennary) or multiple low-affinity N-glycans (i.e., two biantennary or six monoantennary). (B) Simulation of surface glycoprotein fractions in Mgat5$^{-/-}$ cells as a function of Golgi UDP-GlcNAc concentration and N-glycan multiplicities (n). The glycoforms column indicates the number of glycoforms generated for each N-glycan multiplicity using 14 possible N-glycan structures. Slopes indicate the change in surface high n (n = 8, i.e., EGFR) relative to low n (n = 2, i.e., TβR) glycoproteins early and late in the UDP-GlcNAc titration. (Adapted from Lau, K., Partridge, E.A., Silvescu, C.I., Grigorian, A., Pawling, J., Reinhold, V.N., Demetriou, M., and Dennis, J.W., *Cell*, 129, 123, 2007.)

the number of N-glycans (n), which is a sequence encoded feature of each glycoprotein defined by N-X-S/T motifs, their substitution efficiencies by ER oligosaccharidyltransferase, as well as GlcNAc-branching and poly(*N*-acetyllactosamine) extensions in the Golgi [3].

21.4 SIGNALING RECEPTORS, Gal-3 AND Mgat5, IN CANCER CELLS

Transformation-dependent increases in Mgat5 gene expression [25,26] and corresponding β1,6GlcNAc-branching of N-glycans are associated with poor prognosis in cancer patients [27,28]. Similarly, Gal-3 levels increase with transformation and is associated with tumor cell metastasis [29]. Overexpression of either Gal-3 or Mgat5 in epithelial cells confers tumorigenicity and metastatic potential in nude mice [30]. Interestingly, microarray analysis indicated that a variety of subsystems are affected by Gal-3 overexpression including oxidative stress, cell cycle regulation, and cytoskeleton remodeling [31].

Consistent with a pleiotropic effect of Gal-3 and N-glycan ligands, we have shown that Mgat5-deficient mammary tumors in the polyomavirus middle T (PyMT) transgenic model are less responsive to IGF, EGF, PDGF, FGF, and TGF-β due to a shift in receptors from the cell surface to the endosomes [19]. The PyMT oncoprotein is a lipid-linked adaptor protein that recruits p85 and Shc, thus activating phosphoinositide 3-kinase (PI3K) and Ras/Erk signaling [32]. PI3K signaling stimulates cytoskeleton remodelling, endocytosis, and metabolism, and Ras/Erk promotes cell proliferation. Redistribution of receptors to the endosomes and caveolae can be induced in the wild-type tumor cells by disrupting galectin binding with lactose, but not sucrose [19]. Sensitivity to cytokines is completely rescued by blocking coated-pit and caveolae endocytosis with K^+ depletion and nystatin, respectively. PyMT Mgat5$^{-/-}$ mammary carcinoma cells display reduced metastasis *in vivo* and retain an epithelial morphology in tissue culture, while PyMT Mgat5$^{+/+}$ tumor cells appear mesenchymal. Tumor cells remain dependent on extracellular cues to polarize the cytoskeleton and promote the epithelial-to-mesenchymal transition (EMT) phenotype of invasive carcinomas [33].

Notably, EMT requires a balance of both membrane proximal and nuclear signaling by receptor tyrosine kinase (RTK)/Erk/PI3K and TβR/Smad2 pathways, as described further below [34,35]. Expression of Mgat5 from a retroviral vector in PyMT Mgat5$^{-/-}$ cells rescued both sensitivity to cytokines and EMT, suggesting that Mgat5 is necessary to induce the invasive phenotype in PyMT-transformed mammary epithelial cells [19].

Although primary mammary tumor formation is delayed in Mgat5$^{-/-}$ PyMT transgenic mice, Mgat5$^{+/+}$ and Mgat5$^{-/-}$ tumor cells grow at similar rates in tissue culture and as solid tumors when injected into mice [36]. This suggests that secondary genetic or epigenetic changes in Mgat5$^{-/-}$ tumor may compensate for growth factor insensitivity. In this regard, caveolin-1 (Cav-1) expression is commonly decreased in advanced Mgat5$^{-/-}$ PyMT tumors, but not in Mgat5-expressing PyMT tumors [37]. Loss of caveolin-1 in Mgat5$^{-/-}$ PyMT tumor cells was shown to be associated with improved sensitivity to EGF. Cav-1 functions as a negative regulator of RTK signaling and is concentrated in cholesterol lipid rafts [38]. The caveolin scaffolding domain (residues 82–101) mediates its own oligomerization and binding to a conserved caveolin-binding motif in many signaling proteins [39]. Cav-1-deficient mice show enhanced progression of PyMT mammary tumors, as well as increases in mammary and intestinal stem cell proliferation [40,41]. Ectopic expression of Cav-1 in Mgat5$^{-/-}$ PyMT tumors suppressed sensitivity to EGF, but overexpression in Mgat5$^{+/+}$ PyMT tumor cells had no effect [37].

Fluorescence recovery after photobleaching (FRAP) was used to compare the effect of the galectin lattice and Cav-1 on the lateral mobility of EGFR–YFP fusion protein [37]. An area of the cell surface is photobleached and the time to recover the bleached area is quantified. In wild-type cells, first-order diffusion rates of EGFR–YFP mobility is enhanced by lactose pretreatment, consistent with the role of the galectin lattice in restricting glycoprotein diffusion. Comparison of Mgat5$^{+/+}$ and Mgat5$^{-/-}$ tumor cells indicates that expression of β1,6GlcNAc-branched N-glycans slows EGFR–YFP mobility by several fold. However, Cav-1 overexpression in Mgat5$^{-/-}$ cells increased the immobile fraction, while first-order diffusion rates remained similar, indicating that Cav-1 sequesters EGFR in a stable microdomain, distinct in this regard from the galectin lattice. Mgat5$^{-/-}$ (9) cells, which display markedly suppressed EGFR signaling [19] express intermediate levels of endogenous Cav-1, while Mgat5$^{-/-}$ (10) cell line derived from a rapidly progressed tumor are completely deficient in Cav-1 and display responsiveness to EGF. Expression of Cav-1 in Mgat5$^{-/-}$ (10) cells suppressed EGFR signaling and increased the proportion of the FRAP immobile fraction, generating a phenotype similar to that of the Mgat5$^{-/-}$ (9) cells. Moreover, loss of Cav-1 expression correlates with primary tumor size in PyMT Mgat5$^{-/-}$ mice but not in PyMT Mgat5$^{+/-}$ littermates [37]. These studies suggest that Cav1 loss is advantageous to tumor autonomy only when β1,6GlcNAc-branched N-glycans are below a threshold for optimal lattice formation, possibly a situation that occurs early in tumor formation before Mgat5 gene expression is upregulated by Ras pathway activation [37]. Reduced mobility of receptors conferred by galectin cross-linking decreases the fraction of receptors lost to coated-pit endocytosis and Cav-1-positive rafts. The lattice appears to be a significant factor for maintaining sensitivity to cytokines in proliferating and motile cells, including tumor cells, macrophage, and activated T cells [19]. These FRAP studies combined with surface cross-linking and downstream signaling experiments suggest that galectins can sequester receptors in a ligand-responsive state, in what appears to be a distinct membrane microenvironment.

21.5 ENDOCYTOSIS, N-GLYCANS, AND RECEPTOR REGULATION

Glycoprotein levels at the cell surface are dependent on endocytosis rates, vesicular trafficking, and recycling back to the cell surface [19]. Structural features of transmembrane glycoproteins promote internalization from the surface, such as binding sites in the cytosolic tail for β-arrestin, ubiqitination, and AP1/AP2 clathrin adaptor proteins [42–44]. TGF-β receptor II (TβRII), cytotoxic T lymphocyte-associated antigen-4 (CTLA-4), and glucose transporter-4 (GLUT4) are examples of glycoproteins with

AP1/AP2 adaptor-binding motifs that promote rapid constitutive endocytosis and increased residency in the endosomes [45–47]. Constitutive endocytosis is stimulated by RTK signaling and oncogenic activation of Src and PI3K family members, which have a tendency to reduce receptors at the cells surface [48–50]. Strengthening of the galectin lattice opposes the loss of the receptors, an effect that may be driven by PI3K/Akt- and Ras/Erk-dependent increases in Mgat5 gene expression and metabolic flux through the hexosamine pathway [25,26,51]. However, stimulation of endocytosis occurs within 5–15 min by RTK, while turnover of glycoproteins to alter the N-glycan composition at the cell surface is likely to require hours [52]. Thus, strengthening the galectin lattice is protracted relative to upregulation of endocytosis, and initially, this allows glycoproteins with low galectin binding affinity (low number of N-glycans $n < 5$) to be preferentially depleted from the surface relative to those that have more N-glycans ($n \geq 5$). Feedback to Mgat5 gene expression and metabolite flux to the hexosamine pathway enhances tri- and tetraantennary N-glycans, eventually restoring low n receptors to the surface. Considering functional differences between glycoproteins with different number of N-glycans, this mechanism can establish a time-dependent order of activation of functionally distinct pathways [3].

Remarkably, receptor kinases that stimulate cell proliferation and promote oncogenesis (e.g., EGFR, PDGFR, FGFR, IGFR, etc.) tend to have high numbers and densities of Asn-X-Ser/Thr (X ≠ Pro) motifs, compared to receptor kinases mediating tissue morphogenesis and arrest (TβR, EPHR, TIE1, etc.) [3]. Thus, high n receptors (growth/anabolic) are more readily retained at the cell surface in the lattice, while low n receptors (arrest/differentiation) require either low constitutive endocytosis rates or upregulation of Golgi N-glycan branching or both for optimal levels of surface retention. Therefore, initial increases in endocytosis following RTK stimulation of quiescent cells depletes glycoproteins critical to negative growth regulation, but with increasing feedback to the N-glycan processing and hexosamine pathways, arrest glycoproteins and homeostasis are restored (Figure 21.3).

The degree of N-glycans branching pathway is limited by UDP-GlcNAc concentrations in the medial Golgi [53,54]. UDP-GlcNAc, the common donor substrate for all branching GlcNAc-Ts, is generated by the hexosamine pathway from glucose, glutamine, and acetyl-CoA, key intermediates of glycolysis, fatty acid, and nitrogen metabolism [55]. UDP-GlcNAc is synthesized in the cytosol and transported into the Golgi reaching millimolar concentrations [56]. The N-glycan branching GlcNAc-Ts (Mgat1, 2, 4, 5) display stepwise declining affinities for UDP-GlcNAc (0.04–10 mM) as well as declining enzyme concentrations (Figure 21.1). The transit of glycoproteins through the Golgi limits their exposure time to the resident enzymes. Together, these features result in multistep ultrasensitivity to UDP-GlcNAc for tri- and tetraantennary N-glycan production [3]. An ultrasensitive response is characterized by a sigmoid-shaped curve, notably a delay than a sharp rise in output over a narrow range of stimulus, where the steepness of the rise as defined by the Hill coefficient (nH) is >> 1 [57]. In contrast, the common Michaelis–Menten relationship for simple one-step substrate–product relationship produces a hyperbolic response curve (nH~1). Ultrasensitive pathways deliver switch-like (all-or-none) responses as observed in other biological systems, notably cell cycle progression induced by Raf-Mek-Erk signaling in *Xenopus oocytes* [58], and multisite phosphorylation of Sic1 driving mitotic progression in yeast [59].

Tri- and tetraantennary N-glycan production exhibit switch-like responses to increasing UDP-GlcNAc in nontransformed epithelial cells, where GlcNAc-TIV and -TV activities are limiting relative to GlcNAc-TI and -TII. For glycoproteins with one or a few N-glycans ($n < 5$), the fractional increase in glycoforms with these structures mirrors the Golgi pathway response profile (i.e., ultrasensitive). In turn, this fraction has a higher affinity for the galectin lattice, and increases at the surface with similar ultrasensitivity (Figure 21.2B). For example, GLUT4 ($n = 1$) is recruited from endosomes to the cell surface in a switch-like manner with increasing hexosamine flux, but recruitment is blocked by mutation of the single N-X-S/T motif [3]. However, for glycoproteins with high numbers of N-glycans ($n \geq 5$), Golgi processing increases glycoform numbers exponentially, producing more glycoforms with wider ranges of avidities for the galectin lattice. For these high n

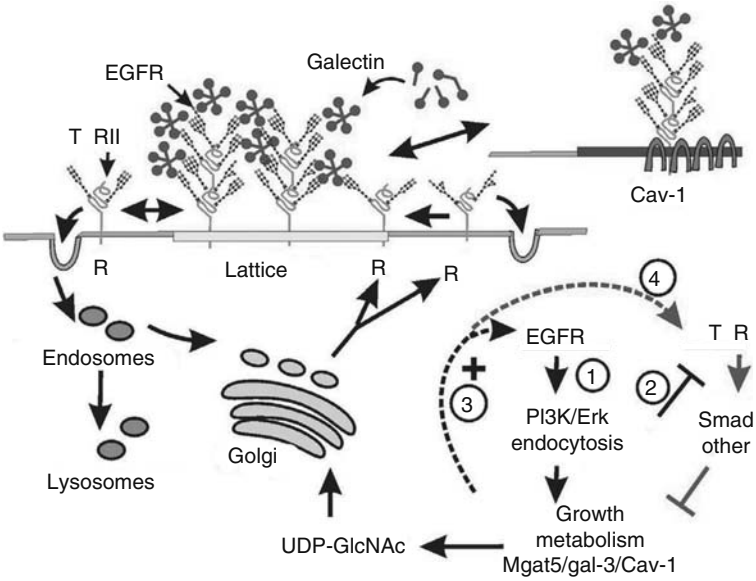


FIGURE 21.3 (See color insert following blank page 170. Also see CD for color figure.) Model of galectin lattice and Cav-1 microdomains to regulate surface glycoproteins residencies and downstream pathways. Growth factor receptors (RTKs) have high numbers of N-glycans, and therefore, high affinities for surface galectins, while TGF-β receptors I and II have low multiplicities (n = 1 and 2). Glycoforms generated in the Golgi are above (R^*) or below (R) the affinity threshold for stable association with the galectin lattice. Glycoproteins maintained in the lattice are protected from loss into the endocytic compartments or Cav-1 positive lipid rafts. In Mgat5$^{-/-}$ cells, insufficient positive feedback to growth signaling (black) results in a predominance of arrest signaling (red). In wild-type cells, (1) stimulation of RTKs in quiescent cells, (2) increases PI3K signaling, and promotes f-actin remodeling and preferential internalization of low n receptors. (3) This enhances positive feedback to the hexosamine pathway/N-glycan processing, which leads ultimately to (4) increasing TβR association with galectins and autocrine arrest signaling. (Adapted from Lau, K., Partridge, E.A., Silvescu, C.I., Grigorian, A., Pawling, J., Reinhold, V.N., Demetriou, M., and Dennis, J.W., *Cell*, 129, 123, 2007.)

glycoproteins, more glycoforms have minimal avidity and associate with the lattice at basal concentration of UDP-GlcNAc (~1.5 mM), whereas the remaining unassociated molecules are then recruited to the lattice in a graded manner with increasing UDP-GlcNAc (Figure 21.2A and B). Thus, ultrasensitivity to UDP-GlcNAc for high-affinity N-glycan production is effectively suppressed for surface recruitment of glycoproteins with high n [3].

Experimental results and computational modeling with high n and low n receptors indicate that hexosamine flux to the galectin/glycoprotein lattice has the potential to regulate transitions between growth and arrest signaling. TGF-β is widely expressed and has picomolar affinity for TβRI/II receptor complex. Hence, changes in receptor numbers over a narrow range can control threshold levels of Smad2/3 activation. TβRI/II is rapidly recycled between the surface and endosomes, independent of their activation by ligand. Hence, downstream TGF-β signaling is particular sensitive to receptor regulation by trafficking [60]. For TβR (n = 2) and CTLA-4 (n = 2) in T cells, surface receptors can be increased to levels that are functionally dominant for cell arrest by lattice upregulation (Figure 21.2B) [3]. At the same time, surface levels of growth-promoting high n RTKs are also maintained by the lattice, which may serve to prime cells for a shift to more favorable growth conditions later. However, increased Mgat4/5 expression, common in oncogene-transformed cells, enhances overall N-glycan affinities for galectins and increases lattice association of high n RTKs, possibly to levels that cannot be suppressed by TβR and other negative regulators.

Regulation of growth and arrest by hexosamine flux to the galectin/glycoprotein lattice may serve as an intermediate for adaptation of cellular responses to the nutrient environment *in vivo* [3]. In keeping with this idea, primary Mgat5$^{-/-}$ mouse embryonic fibroblasts (MEFs) and stem cells display less RTK/PI3K/Erk and more TGF-β/Smad signaling, an imbalance favoring growth arrest [61]. Glucose uptake by Mgat5$^{-/-}$ MEFs is impaired, particularly under low serum or low glucose growth conditions. Moreover, Mgat5-deficient mice are resistant to weight-gain on a calorie-enriched diet, while calorie intake and physical activity are the same as Mgat5$^{+/+}$ littermates. Mgat5-deficient mice are hypersensitive to fasting, mildly hypoglycemic, and display increased oxidative respiration. In these mice, self-renewal of muscle satellite cells and bone marrow osteogenesis are limited, and this is associated with early osteoporosis and other aspects of aging. Imbalances in stem cell and tissue renewal are features of mice deficient in Rho GTPase Rac1, Pten phosphatase [62–64], and p53 [65–67]. Thus, β1,6GlcNAc-branched N-glycans appear to promote anabolic metabolism and tissue renewal *in vivo*, possibly by maintaining homeostasis in growth and arrest signaling [3,61].

In addition to cell-autonomous effects of Mgat5 deficiency on glucose homeostasis, systemic regulation by insulin, glucagons, or other peptide hormones may contribute to the Mgat5$^{-/-}$ metabolic phenotype. In this regard, the Mgat4a$^{-/-}$ mice display a predisposition to type II diabetes, and their pancreatic β cells are deficient in glucose-stimulated insulin secretion due to a reduction in galectin binding and surface retention of GLUT2 [68]. We have shown that hexosamine flux enhances surface expression of GLUT4 in HEK293T cells in an N-glycan-dependent manner, suggesting the possibility that other nutrient and ion transporters may be dependent on metabolic flux to N-glycan branching [3]. N-glycan branching changes with the turnover of glycoproteins and occurs on a longer timescale than changes in ionic and metabolic fluxes. More generally, regulation of metabolic flux, ion channel function, and signaling occurs on small timescales, but must be integrated with long-term cellular responsiveness to growth and arrest cues, which we suggest, depends on the continual finetuning of N-glycan modifications.

Pathologically high glucose conditions increase UDP-GlcNAc levels, as observed in the muscle of GLUT1 transgenic mice [69]. Glutamine:fructose-6-phosphate amidotransferase (GFAT) activity, a rate-limiting enzyme in the hexosamine pathway, is also increased in these mice. However, adaptation to starvation conditions is a more important consideration from an evolutionary point of view. In cell culture, low glucose conditions enhances UDP-GlcNAc levels and β1,6GlcNAc-branched glycans at the cell surface, possibly due to a slowing of membrane turnover and glycoprotein transit through the Golgi. During short-term glucose starvation, increased oxidative metabolism and reactive oxygen species (ROS) production leads to oxidation of glyceraldehyde-3-phosphate dehydrogenase, diverting fructose-6-phosphate from glycolysis into the hexosamine pathway [70]. Also, a shift from anabolic to catabolic metabolism increases acetyl-CoA and glutamine availability to GFAT, which are both limiting substrates for UDP-GlcNAc biosynthesis. Low-glucose conditions also reduces dolichol-pp-oligosaccharide levels, and induces the PERK ER stress response, which slows protein synthesis and thereby restores balance between protein synthesis and glycosylation [71]. However, in prolonged low-glucose conditions, the efficiency of N-X-S/T occupancy is reduced along with surface residency of EGFR [72]. Therefore, cotranslational glycosylation in the ER and subsequent Golgi processing depends ultimately on metabolite flux- to sugar–nucleotide pools. Differential use of N-X-S/T sites can be expected to alter Golgi-dependent regulation, but should maintain the relationship between low and high *n* glycoproteins.

Galectins family members and their complex N-glycan ligands are widely expressed in mammalian tissues [13,15], and given that they interact, deficiencies in either may produce similar phenotypes. Single-gene deficiencies in Gal-1 and Gal-3 as well as Gal-1/Gal-3 compound mutant mice are viable and fertile [73], but display subtle phenotypes revealed under stress and with age. In Gal-3$^{-/-}$ mice, reepithelialization of wounds is slower [74], diabetic autoimmune glomerulopathy is accelerated [75], and peritoneal inflammation, macrophage spreading, and motility are reduced [76]. Liver injury in Gal-3$^{-/-}$ induces the usual inflammatory cytokines, but hepatic satellite cells are less sensitive to TGF-β correlating with less hepatic fibrosis in the mice [77]. Both

Gal-3$^{-/-}$ and Mgat5$^{-/-}$ mice display age-related skeletal defects and failure of renewal. Chondrocyte survival is deficient in Gal-3$^{-/-}$ mice [78], while bone marrow osteoprogenitor cells are decreased in Mgat5$^{-/-}$ mice, accompanied by early osteoporosis [51]. These results are consistent with the notion that β1,6GlcNAc-branching and Gal-3 interact *in vivo*, and both are partially limiting for stem cell survival and tissue renewal. Mgat5$^{-/-}$ mice also have a metabolic defect with increased oxidative respiration, and low-grade inflammation associated with increased sensitivity to autoimmune disease, which may contribute to early aging [61]. Gal-3 and Mgat5 compound mutant mice might display enhancement for some defects, although this remains to be further examined.

21.6 REGULATION OF EPITHELIAL-TO-MESENCHYMAL TRANSITION IN CANCER CELLS BY N-GLYCAN PROCESSING

Oncogenes such as viral PyMT, constitutively activated Ras, and loss of a Pten allele generally act downstream of the cell surface and reduce the need for extracellular growth factors for proliferation. However, responses to growth factors and TGF-β are still necessary to promote EMT, the phenotype that characterizes invasive carcinomas [33]. Importantly, EMT in cancer cells requires both pathways in balance to polarize the cytoskeleton and to sustain cell proliferation [33]. In focal adhesions, RTKs and integrin clustering stimulate CDC42/Rac GTPases, while TGF-β binding to TβRII promotes degradation of RhoA [35]. However, ligand binding to TβRII also stimulates phosphorylation of TβRI and in turn Smad2/3, promoting binding to Smad4 and translocation of the complex to the nucleus. The Smad complex associates with other transcription factors and decreases c-Myc and increases $p21^{Cip1}$ and $p15^{INK4a}$ gene expression, promoting cell cycle arrest [79,80]. In opposition to TGF-β signaling, RTK/PI3K/Erk activation of cyclinD/Cdc4 kinase phosphorylates Rb, inducing growth, with Smad2/3 blocking its activity [81]. EMT was rescued in Mgat5$^{-/-}$ tumor cells by GlcNAc supplementation to the hexosamine/Golgi pathways as well as Mgat5 expression, but not with exogenous TGF-β [3,19].

The Mgat5$^{-/-}$ mutation suppresses tumor progression and growth signaling in MEFs, while the Pten$^{+/-}$ mutation displays the opposite phenotype. Pten, the phosphoinositide 3-phosphatase, is a potent tumor suppressor mutated in many human tumors, and loss of one allele of Pten enhances PI3K/Akt signaling. The survival of compound Pten$^{+/-}$ Mgat5$^{-/-}$ and Pten$^{+/-}$ Mgat5$^{+/-}$ mice is enhanced compared to Pten$^{+/-}$ alone due to an apparent delay in cancer development [51]. However, survival of compound mutant mice is intermediate (with [Pten$^{+/-}$] < [Pten$^{+/-}$Mgat5$^{-/-}$ and Pten$^{+/-}$Mgat5$^{+/-}$] < [Mgat5$^{-/-}$] < [Mgat5$^{+/-}$ and Mgat5$^{+/+}$]). β1,6GlcNAc-branched N-glycans are increased in Pten $^{+/-}$ MEFs, indicating positive feedback from PI3K/Akt to N-glycan branching [51,82]. Although the Mgat5 mutation was epistatic for Pten$^{+/-}$ proliferation and Akt activation by substratum adhesion, Pten$^{+/-}$Mgat5$^{-/-}$ double mutant MEFs were not rescued for planar asymmetry. This is consistent with the idea that extracellular cues are required to drive the asymmetry of signaling and microfilament remodeling associated with cell motility.

21.7 T CELL ACTIVATION IS DEPENDENT ON N-GLYCAN BRANCHING AND GALECTINS

In naïve T cells, galectin-3 binding to T cell receptors (TCR) impedes TCR mobility. However, endocytosis rates are very low, and rather than preventing surface loss, galectin binding impedes TCR clustering in response to low-dose antigen, which increases the threshold for T cell activation [83]. The TCRα and β chains have seven N-glycans, and CD3 γ and δ each has one. In Mgat5$^{-/-}$ T cells, TCR mobilization and clustering in response to TCR agonist coated beads were greatly enhanced. TCR-dependent tyrosine phosphorylation, actin microfilament reorganization, and Ca^{2+} mobilization were greater in Mgat5$^{-/-}$ T cells 1–10 min after exposure to ligand [83]. Responses to phorbol esters, a cell-permeable stimulator of signaling downstream of TCR was normal, indicating

that the defect in Mgat5$^{-/-}$ T cells occurs at the cell surface. TCR affinity for peptide–MHC that activates T cells is 10^{-4} to 10^{-5} M, and presumably can compete favorably with the TCR–galectin lattice for immune synapse formation [84]. Mgat5$^{-/-}$ mice on the 129/sv background show late-onset spontaneous kidney autoimmune disease and are hypersensitive to EAE, while Mgat5$^{+/-}$ mice are intermediate [83]. Similarly, mice deficient in α-mannosidase II, an earlier enzyme in the β1,6GlcNAc-branching pathway, also develop late onset spontaneous kidney autoimmunity [85].

Following activation, T cells become more sensitive to antigen stimulation even though TCR levels and the intrinsic affinity of TCR for peptide–MHC does not change [86,87]. Receptor-associated kinase lck increases, and signaling sensitivity downstream of TCR appears to be enhanced in activated cells [87]. Mgat5 transcription, enzyme activity, and N-glycan products are increased following activation [88,89], but the influence of β1,6GlcNAc branched N-glycans may shift from TCR to integrins. In this regard, leukocyte migration into the peritoneal cavity in response to an injection of thioglycolate was delayed in Mgat5$^{-/-}$ mice, and the cells were more adhesive on fibronectin [19]. Activated T cells undergo multiple rounds of cell division, differentiate into effector T cells, and finally growth arrest. TCR ($n = 9$) activation in naïve T cells induces proliferation, IL-2 and PI3K signaling [89], increased glucose metabolism [90] and UDP-GlcNAc biosynthesis, Mgat5 gene expression, and ultimately recruits CTLA-4 ($n = 2$) to the lattice from endosomes to the cell surface (~twofold increase), resulting in growth arrest [3]. Pathogen elimination requires rapid expansion of reactive T cells, followed by abrupt suppression to limit inflammation and avoid autoimmune responses. Hexosamine flux to the Golgi pathway negatively regulates naïve T cells by increasing the threshold for antigen–MHC stimulation, and later by promoting surface retention of CTLA-4 in T cell blasts. Several autoimmune-prone strains of mice display deficiencies in β1,6GlcNAc-branching (PL/J, NOD) comparable to that of Mgat5$^{+/-}$ mice on strain backgrounds that are more resistant to autoimmune diseases (129/sv, C57BL6) [54,83].

21.8 BEHAVIORAL PHENOTYPE IN Mgat5-DEFICIENT MICE

Mgat5-deficient mice are normal in sensorimotor tests, with intact function of the hypothalamus-pituitary-adrenal axis, but display resistance to stress-induced depression and modest impairment in the hidden version of Morris water maze task attributed to a cognitive performance deficit [91]. It is possible that certain neurotransmitter receptors or reuptake transporters are dependent on β1,6GlcNAc-branching for activity or surface expression. β1,6GlcNAc-branched N-glycans are extensively expressed in the central nervous system (CNS) from E9.5 to adulthood in brain areas such as cortex, hypothalamus, and hippocampus [82], regions associated with depressive and stress-related responses [92,93]. N-glycan processing may be required for monoaminergic (i.e., serotonergic and noradrenergic) or other neurotransmitter systems (i.e., dopaminergic, GABAergic, and glutamatergic) that underlie mood changes. In this regard, comparison of the serotonin transporter (SERT) expressed in wild-type and Mgat5-deficient CHO cells revealed that β1,6GlcNAc-branched N-glycans contribute to SERT function, homooligomerization, and binding to myosin IIA [94]. SERT regulates the serotonergic system through the reuptake and clearance of serotonin released from the nerve terminal and is blocked by selective serotonin reuptake inhibitor (SSRI) antidepressants [95]. N-glycosylation of β2-adrenergic receptors regulates receptor trafficking to the lysosomes after prolonged agonist exposure [96]. Moreover, mice deficient in serotonin or noradrenaline transporter show similar antidepressant phenotype as Mgat5$^{-/-}$ mice [97,98]. It is possible that β1,6GlcNAc branching of N-glycans and galectin binding retains certain receptors and reuptake transporters in the synaptic cleft, playing a role in depression.

21.9 CONCLUSION AND FUTURE DIRECTIONS

Although *Caenorhabditis elegans* express functional Mgat1, Mgat2, and Mgat5 enzymes (not Mgat4) [99], complex N-glycans are present in very low amounts, and mutations of these genes have

no developmental phenotype under normal laboratory conditions [100]. However, minimal branching by Mgat1 is required for mammalian embryogenesis [101,102]. The Mgat2 deficiency is postpartum lethal with morphogenic defects, similar to the homologous deficiency in human type II congenital disorder of glycosylation [103]. In mammals, N-glycan sites are reduced in number and density on receptor kinases that mediate arrest/differentiation, while receptors in *C. elegans* and *Drosophila melanogaster* do not segregate in this manner and generally have relatively high N-X-S/T multiplicities. This suggests that coevolution of N-glycan branching and N-glycan multiplicity in receptor kinases and other critical glycoproteins was driven by a need for greater conditional regulation of the immune system and metabolism in mammals. The Mgat5$^{-/-}$ mice appear normal at birth but adult phenotypes involve hypersensitivity to various stressors, indicating a postnatal role of N-glycosylation in robustness. Computational modeling indicates that expression of Mgat1 and Mgat2 alone provides insufficient affinity for galectin binding by low *n* glycoproteins, and in this regard, is consistent with developmental defects observed in mouse mutants. In conclusion, a better understanding of the metabolic and conditional regulators of N-glycan processing and galectin binding may lead to new treatments for autoimmune diseases, congenital disorders of glycosylation, and a host of conditions associated with aging and cancer.

ACKNOWLEDGMENTS

This research was supported by grants from CIHR and Genome Canada through the OGI to J.W.D., funding from NIH to M.D., and K.S.L. was funded by a NSERC PGSD scholarship.

REFERENCES

1. Cooper, D.N. 2002. Galectinomics: Finding themes in complexity. *Biochim Biophys Acta* 1572:209–231.
2. Drickamer, K. and Fadden, A.J. 2002. Genomic analysis of C-type lectins. *Biochem Soc Symp* 69: 59–72.
3. Lau, K., Partridge, E.A., Silvescu, C.I., Grigorian, A., Pawling, J., Reinhold, V.N., Demetriou, M., and Dennis, J.W. 2007. Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* 129:123–124.
4. Herrmann, J., Turck, C.W., Atchison, R.E., Huflejt, M.E., Poulter, L., Gitt, M.A., Burlingame, A.L., Barondes, S.H., and Leffler, H. 1993. Primary structure of the soluble lactose binding lectin L-29 from rat and dog and interaction of its non-collagenous proline-, glycine-, tyrosine-rich sequence with bacterial and tissue collagenase. *J Biol Chem* 268:26704–26711.
5. Leffler, H. and Barondes, S.H. 1986. Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. *J Biol Chem* 261:10119–10126.
6. Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W.E., et al. 2002. Oligosaccharide specificity of galectins: A search by frontal affinity chromatography. *Biochim Biophys Acta* 1572:232–254.
7. Patnaik, S.K., Potvin, B., Carlsson, S., Sturm, D., Leffler, H., and Stanley, P. 2006. Complex N-glycans are the major ligands for galectin-1, -3, and -8 on Chinese hamster ovary cells. *Glycobiology* 16:305–317.
8. Schachter, H. 1986. Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides. *Biochem Cell Biol* 64:163–181.
9. Lee, R.T. and Lee, Y.C. 2000. Affinity enhancement by multivalent lectin–carbohydrate interaction. *Glycoconj J* 17:543–551.
10. Ahmad, N., Gabius, H.J., Andre, S., Kaltner, H., Sabesan, S., Roy, R., Liu, B., Macaluso, F., and Brewer, C.F. 2003. Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *J Biol Chem* 279:10841–10847.
11. Ochieng, J., Fridman, R., Nangia-Makker, P., Kleiner, D.E., Liotta, L.A., Stetler-Stevenson, W.G., and Raz, A. 1994. Galectin-3 is a novel substrate for human matrix metalloproteinases-2 and -9. *Biochemistry* 33:14109–14114.
12. Mazurek, N., Conklin, J., Byrd, J.C., Raz, A., and Bresalier, R.S. 2000. Phosphorylation of the beta-galactoside-binding protein galectin-3 modulates binding to its ligands. *J Biol Chem* 275:36311–36315.
13. Hughes, R.C. 1999. Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim Biophys Acta* 1473:172–185.

14. Delacour, D., Greb, C., Koch, A., Salomonsson, E., Leffler, H., Le Bivic, A., and Jacob, R. 2007. Apical sorting by galectin-3-dependent glycoprotein clustering. *Traffic* 8:379–388.
15. Dumic, J., Dabelic, S., and Flogel, M. 2006. Galectin-3: An open-ended story. *Biochim Biophys Acta* 1760:616–635.
16. Levy, Y., Ronen, D., Bershadsky, A.D., and Zick, Y. 2003. Sustained induction of ERK, protein kinase B, and p70 S6 kinase regulates cell spreading and formation of F-actin microspikes upon ligation of integrins by galectin-8, a mammalian lectin. *J Biol Chem* 278:14533–14542.
17. Nishi, N., Shoji, H., Seki, M., Itoh, A., Miyanaka, H., Yuube, K., Hirashima, M., and Nakamura, T. 2003. Galectin-8 modulates neutrophil function via interaction with integrin alphaM. *Glycobiology* 13:755–763.
18. Lagana, A., Goetz, J.G., Cheung, P., Raz, A., Dennis, J.W., and Nabi, I.R. 2006. Galectin binding to Mgat5-modified N-glycans regulates fibronectin matrix remodeling in tumor cells. *Mol Cell Biol* 26:3181–3193.
19. Partridge, E.A., Le Roy, C., Di Guglielmo, G.M., Pawling, J., Cheung, P., Granovsky, M., Nabi, I.R., Wrana, J.L., and Dennis, J.W. 2004. Regulation of cytokine receptors by Golgi N-glycan processing and endocytosis. *Science* 306:120–124.
20. Brewer, C.F., Miceli, M.C., and Baum, L.G. 2002. Clusters, bundles, arrays and lattices: Novel mechanisms for lectin–saccharide-mediated cellular interactions. *Curr Opin Struct Biol* 12:616–623.
21. Stillman, B.N., Hsu, D.K., Pang, M., Brewer, C.F., Johnson, P., Liu, F.T., and Baum, L.G. 2006. Galectin-3 and galectin-1 bind distinct cell surface glycoprotein receptors to induce T cell death. *J Immunol* 176:778–789.
22. Stroop, C.J., Weber, W., Gerwig, G.J., Nimtz, M., Kamerling, J.P., and Vliegenthart, J.F.G. 2000. Characterization of the carbohydrate chains of the secreted form of the human epidermal growth factor receptor. *Glycobiology* 10:901–917.
23. Zhen, Y., Caprioli, R.M., and Staros, J.V. 2003. Characterization of glycosylation sites of the epidermal growth factor receptor. *Biochemistry* 42:5478–5492.
24. Rudd, P.M., Wormald, M.R., Stanfield, R.L., Huang, M., Mattsson, N., Speir, J.A., DiGennaro, J.A., Fetrow, J.S., Dwek, R.A., and Wilson, I.A. 1999. Roles for glycosylation of cell surface receptors involved in cellular immune recognition. *J Mol Biol* 293:351–366.
25. Kang, R., Saito, H., Ihara, Y., Miyoshi, E., Koyama, N., Sheng, Y., and Taniguchi, N. 1996. Transcriptional regulation of the *N*-acetylglucosaminyltranserase V gene in human bile duct carcinoma cells (HuCC-T1) is mediated by Ets-1. *J Biol Chem* 271:26706–26712.
26. Buckhaults, P., Chen, L., Fregien, N., and Pierce, M. 1997. Transcriptional regulation of *N*-acetylglucosaminyltransferase V by the src Oncogene. *J Biol Chem* 272:19575–19581.
27. Fernandes, B., Sagman, U., Auger, M., Demetriou, M., and Dennis, J.W. 1991. Beta1–6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. *Cancer Res* 51:718–723.
28. Seelentag, W.K., Li, W.P., Schmitz, S.F., Metzger, U., Aeberhard, P., Heitz, P.U., and Roth, J. 1998. Prognostic value of beta1,6-branched oligosaccharides in human colorectal carcinoma. *Cancer Res* 58:5559–5564.
29. Gong, H.C., Honjo, Y., Nangia-Makker, P., Hogan, V., Mazurak, N., Bresalier, R.S., and Raz, A. 1999. The NH2 terminus of galectin-3 governs cellular compartmentalization and functions in cancer cells. *Cancer Res* 59:6239–6245.
30. Demetriou, M., Nabi, I.R., Coppolino, M., Dedhar, S., and Dennis, J.W. 1995. Reduced contact-inhibition and substratum adhesion in epithelial cells expressing GlcNAc-transferase V. *J Cell Biol* 130:383–392.
31. Mazurek, N., Sun, Y.J., Price, J.E., Ramdas, L., Schober, W., Nangia-Makker, P., Byrd, J.C., Raz, A., and Bresalier, R.S. 2005. Phosphorylation of galectin-3 contributes to malignant transformation of human epithelial cells via modulation of unique sets of genes. *Cancer Res* 65:10767–10775.
32. Webster, M.A., Hutchinson, J.N., Rauh, M.J., Muthuswamy, S.K., Anton, M., Tortorice, C.G., Cardiff, R.D., Graham, F.L., Hassell, J.A., and Muller, W.J. 1998. Requirement for both Shc and phosphatidylinositol 3′ kinase signaling pathways in polyomavirus middle T-mediated mammary tumorigenesis. *Mol Cell Biol* 18:2344–2359.
33. Thiery, J.P. 2003. Epithelial–mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol* 15:740–746.
34. Oft, M., Akhurst, R.J., and Balmain, A. 2002. Metastasis is driven by sequential elevation of H-ras and Smad2 levels. *Nat Cell Biol* 4:487–494.

35. Ozdamar, B., Bose, R., Barrios-Rodiles, M., Wang, H.R., Zhang, Y., and Wrana, J.L. 2005. Regulation of the polarity protein Par6 by TGFbeta receptors controls epithelial cell plasticity. *Science* 307:1603–1609.
36. Mendelsohn, R., Cheung, P., Berger, L., Partridge, E.A., Lau, K., Pawling, J., and Dennis, J.W. 2007. Control of tumor metabolism and growth by N-glycan processing, *Cancer Res* 67:9771–9780.
37. Lajoie P., Partridge, E.A., Guay, G.S.N., Goetz, J.G., Pawling, J., Lagana, A., Dennis, J.W., and Nabi, I.R. 2007. Plasma membrane domain organisation regulates EGFR signaling in tumor cells. *J Cell Biol* 179:341–356.
38. Lu, Z., Ghosh, S., Wang, Z., and Hunter, T. 2003. Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. *Cancer Cell* 4:499–515.
39. Okamoto, T., Schlegel, A., Scherer, P.E., and Lisanti, M.P. 1998. Caveolins, a family of scaffolding proteins for organizing "preassembled signaling complexes" at the plasma membrane. *J Biol Chem* 273:5419–5422.
40. Williams, T.M., Medina, F., Badano, I., Hazan, R.B., Hutchinson, J., Muller, W.J., Chopra, N.G., Scherer, P.E., Pestell, R.G., and Lisanti, M.P. 2004. Caveolin-1 gene disruption promotes mammary tumorigenesis and dramatically enhances lung metastasis in vivo. Role of Cav-1 in cell invasiveness and matrix metalloproteinase (MMP-2/9) secretion. *J Biol Chem* 279:51630–51646.
41. Li, J., Hassan, G.S., Williams, T.M., Minetti, C., Pestell, R.G., Tanowitz, H.B., Frank, P.G., Sotgia, F., and Lisanti, M.P. 2005. Loss of caveolin-1 causes the hyper-proliferation of intestinal crypt stem cells, with increased sensitivity to whole body gamma-radiation. *Cell Cycle* 4:1817–1825.
42. Soubeyran, P., Kowanetz, K., Szymkiewicz, I., Langdon, W.Y., and Dikic, I. 2002. Cbl-CIN85–endophilin complex mediates ligand-induced downregulation of EGF receptors. *Nature* 416:183–186.
43. Petrelli, A., Gilestro, G.F., Lanzardo, S., Comoglio, P.M., Migone, N., and Giordano, S. 2002. The endophilin-CIN85-Cbl complex mediates ligand-dependent downregulation of c-Met. *Nature* 416:187–190.
44. Polo, S., Sigismund, S., Faretta, M., Guidi, M., Capua, M.R., Bossi, G., Chen, H., De Camilli, P., and Di Fiore, P.P. 2002. A single motif responsible for ubiguitin recognition and monoubiquitination in endocytic proteins. *Nature* 416:451–455.
45. Ehrlich, M., Shmuely, A., and Henis, Y.I. 2001. A single internalization signal from the di-leucine family is critical for constitutive endocytosis of the type II TGF-beta receptor. *J Cell Sci* 114:1777–1786.
46. Bradshaw, J.D., Lu, P., Leytze, G., Rodgers, J., Schieven, G.L., Bennett, K.L., Linsley, P.S., and Kurtz, S.E. 1997. Interaction of the cytoplasmic tail of CTLA-4 (CD152) with a clathrin-associated protein is negatively regulated by tyrosine phosphorylation. *Biochemistry* 36:15975–15982.
47. Al-Hasani, H., Kunamneni, R.K., Dawson, K., Hinck, C.S., Muller-Wieland, D., and Cushman, S.W. 2002. Roles of the N- and C-termini of GLUT4 in endocytosis. *J Cell Sci* 115:131–140.
48. Ware, M.F., Tice, D.A., Parsons, S.J., and Lauffenburger, D.A. 1997. Overexpression of cellular Src in fibroblasts enhances endocytic internalization of epidermal growth factor receptor. *J Biol Chem* 272:30185–30190.
49. Raiborg, C., Gronvold Bache, K., Mehlum, A., Stang, E., and Stenmark, H. 2001. Hrs recruits clathrin to early endosomes. *EMBO J* 20:5008–5021.
50. Lanzetti, L., Palamidessi, A., Areces, L., Scita, G., and Di Fiore, P.P. 2004. Rab5 is a signalling GTPase involved in actin remodelling by receptor tyrosine kinases. *Nature* 429:309–314.
51. Cheung, P. and Dennis, J.W. 2007. Mgat5 and Pten interact to regulate cell growth and polarity. *Glycobiology* 17:767–773.
52. Helgemo, S.L. and Hart, G.W. 1998. Recycling cell surface glycoproteins undergo limited oligosaccharide reprocessing in LEC1 mutant Chinese hamster ovary cells. *Glycobiology* 8:1173–1182.
53. Sasai, K., Ikeda, Y., Fujii, T., Tsuda, T., and Taniguchi, N. 2002. UDP-GlcNAc concentration is an important factor in the biosynthesis of beta1,6-branched oligosaccharides: regulation based on the kinetic properties of *N*-acetylglucosaminyltransferase V. *Glycobiology* 12:119–127.
54. Grigorian, A., Lee, S.-U., Tian, W., Chen, I.-J., Gao, G., Mendelsohn, R., Dennis, J.W., and Demetriou, M. 2007. Control of T cell mediated autoimmunity by metabolite flux to N-glycan biosynthesis. *J Biol Chem* 282:20027–20035.
55. Broschat, K.O., Gorka, C., Page, J.D., Martin-Berger, C.L., Davies, M.S., Huang Hc, H.C., Gulve, E.A., Salsgiver, W.J., and Kasten, T.P. 2002. Kinetic characterization of human glutamine-fructose-6-phosphate amidotransferase I: potent feedback inhibition by glucosamine 6-phosphate. *J Biol Chem* 277:14764–14770.

56. Waldman, B.C. and Rudnick, G. 1990. UDP-GlcNAc transport across the Golgi membrane: Electroneutral exchange for dianionic UMP. *Biochemistry* 29:44–52.
57. Koshland, D.E., Jr., Goldbeter, A., and Stock, J.B. 1982. Amplification and adaptation in regulatory and sensory systems. *Science* 217:220–225.
58. Ferrell, J.E.J. and Machleder, E.M. 1998. The biochemical basis of an all-or-none cell fate switch in *Xenopus oocytes. Science* 280:895–898.
59. Klein, P., Pawson, T., and Tyers, M. 2003. Mathematical modeling suggests cooperative interactions between a disordered polyvalent ligand and a single receptor site. *Curr Biol* 13:1669–1678.
60. Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F., and Wrana, J.L. 2003. Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nat Cell Biol* 5:410–421.
61. Cheung, P., Pawling, J., Partridge, E.A., Sukhu, B., Grynpas, M., and Dennis, J.W. 2007. Metabolic homeostasis and tissue renewal are dependent on beta 1, 6GlcNAc-branched N-glycans. *Glycobiology* 17:828–837.
62. Benitah, S.A., Frye, M., Glogauer, M., and Watt, F.M. 2005. Stem cell depletion through epidermal deletion of Rac1. *Science* 309:933–935.
63. Zhang, J., Grindley, J.C., Yin, T., Jayasinghe, S., He, X.C., Ross, J.T., Haug, J.S., Rupp, D., Porter-Westpfahl, K.S., Wiedemann, L.M., et al. 2006. PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention. *Nature* 441:518–522.
64. Yilmaz, O.H., Valdez, R., Theisen, B.K., Guo, W., Ferguson, D.O., Wu, H., and Morrison, S.J. 2006. Pten dependence distinguishes haematopoietic stem cells from leukaemia-initiating cells. *Nature* 441:475–482.
65. Tyner, S.D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., et al. 2002. p53 mutant mice that display early ageing-associated phenotypes. *Nature* 415:45–53.
66. van Heemst, D., Mooijaart, S.P., Beekman, M., Schreuder, J., de Craen, A.J., Brandt, B.W., Slagboom, P.E., and Westendorp, R.G. 2005. Variation in the human TP53 gene affects old age survival and cancer mortality. *Exp Gerontol* 40:11–15.
67. Matoba, S., Kang, J.G., Patino, W.D., Wragg, A., Boehm, M., Gavrilova, O., Hurley, P.J., Bunz, F., and Hwang, P.M. 2006. p53 regulates mitochondrial respiration. *Science* 312:1650–1653.
68. Ohtsubo, K., Takamatsu, S., Minowa, M.T., Yoshida, A., Takeuchi, M., and Marth, J.D. 2005. Dietary and genetic control of glucose transporter 2 glycosylation promotes insulin secretion in suppressing diabetes. *Cell* 123:1307–1321.
69. Buse, M.G., Robinson, K.A., Marshall, B.A., and Mueckler, M. 1996. Differential effects of GLUT1 or GLUT4 overexpression on hexosamine biosynthesis by muscles of transgenic mice. *J Biol Chem* 271:23197–23202.
70. Du, X.L., Edelstein, D., Rossetti, L., Fantus, I.G., Goldberg, H., Ziyadeh, F., Wu, J., and Brownlee, M. 2000. Hyperglycemia-induced mitochondrial superoxide overproduction activates the hexosamine pathway and induces plasminogen activator inhibitor-1 expression by increasing Sp1 glycosylation. *Proc Natl Acad Sci USA* 97:12222–12226.
71. Shang, J., Gao, N., Kaufman, R.J., Ron, D., Harding, H.P., and Lehrman, M.A. 2007. Translation attenuation by PERK balances ER glycoprotein synthesis with lipid-linked oligosaccharide flux. *J Cell Biol* 176:605–616.
72. Konishi, A. and Berk, B.C. 2003. Epidermal growth factor receptor transactivation is regulated by glucose in vascular smooth muscle cells. *J Biol Chem* 278:35049–35056.
73. Colnot, C., Fowlis, D., Ripoche, M.A., Bouchaert, I., and Poirier, F. 1998. Embryonic implantation in galectin 1/galectin 3 double mutant mice. *Dev Dyn* 211:306–313.
74. Cao, Z., Said, N., Amin, S., Wu, H.K., Bruce, A., Garate, M., Hsu, D.K., Kuwabara, I., Liu, F.T., and Panjwani, N. 2002. Galectins-3 and -7, but not galectin-1, play a role in re-epithelialization of wounds. *J Biol Chem* 277:42299–42305.
75. Pugliese, G., Pricci, F., Iacobini, C., Leto, G., Amadio, L., Barsotti, P., Frigeri, L., Hsu, D.K., Vlassara, H., Liu, F.T., et al. 2001. Accelerated diabetic glomerulopathy in galectin-3/AGE receptor 3 knockout mice. *FASEB J* 15:2471–2479.
76. Hsu, D.K., Yang, R.Y., Pan, Z., Yu, L., Salomon, D.R., Fung-Leung, W.P., and Liu, F.T. 2000. Targeted disruption of the galectin-3 gene results in attenuated peritoneal inflammatory responses. *Am J Pathol* 156:1073–1083.
77. Henderson, N.C., Mackinnon, A.C., Farnworth, S.L., Poirier, F., Russo, F.P., Iredale, J.P., Haslett, C., Simpson, K.J., and Sethi, T. 2006. Galectin-3 regulates myofibroblast activation and hepatic fibrosis. *Proc Natl Acad Sci USA* 103:5060–5065.

78. Colnot, C., Sidhu, S.S., Balmain, N., and Poirier, F. 2001. Uncoupling of chondrocyte death and vascular invasion in mouse galectin 3 null mutant bones. *Dev Biol* 229:203–214.
79. Siegel, P.M. and Massague, J. 2003. Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer* 3:807–821.
80. Seoane, J., Le, H.V., Shen, L., Anderson, S.A., and Massague, J. 2004. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* 117:211–223.
81. Matsuura, I., Denissova, N.G., Wang, G., He, D., Long, J., and Liu, F. 2004. Cyclin-dependent kinases regulate the antiproliferative function of Smads. *Nature* 430:226–231.
82. Granovsky, M., Fata, J., Pawling, J., Muller, W.J., Khokha, R., and Dennis, J.W. 2000. Suppression of tumor growth and metastasis in Mgat5-deficient mice. *Nat Med* 6:306–312.
83. Demetriou, M., Granovsky, M., Quaggin, S., and Dennis, J.W. 2001. Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. *Nature* 409:733–739.
84. Grakoui, A., Bromley, S.K., Sumen, C., Davis, M.M., Shaw, A.S., Allen, P.M., and Dustin, M.L. 1999. The immunological synapse: a molecular machine controlling T cell activation. *Science* 285:221–227.
85. Chui, D., Oh-Eda, M., Liao, Y.-F., Panneerselvam, K., Lal, A., Marek, K.W., Freeze, H.H., Moremen, K.W., Fukuda, M.N., and Marth, J.D. 1997. Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell* 90:157–167.
86. Fahmy, T.M., Bieler, J.G., Edidin, M., and Schneck, J.P. 2001. Increased TCR avidity after T cell activation: A mechanism for sensing low-density antigen. *Immunity* 2:135–143.
87. Slifka, M.K. and Whitton, J.L. 2001. Functional avidity maturation of CD8(+) T cells without selection of higher affinity TCR. *Nat Immunol* 2:711–717.
88. Lemaire, S., Derappe, C., Michalski, J.C., Aubery, M., and Neel, D. 1994. Expression of beta1–6-branched N-linked oligosaccharides is associated with activation in human T4 and T8 cell populations. *J Biol Chem* 269:8069–8074.
89. Morgan, R., Gao, G., Pawling, J., Dennis, J.W., Demetriou, M., and Li, B. 2004. *N*-acetylglucosaminyltransferase V (Mgat5)-mediated N-glycosylation negatively regulates Th1 cytokine production by T cells. *J Immunol* 173:7200–7208.
90. Frauwirth, K.A., Riley, J.L., Harris, M.H., Parry, R.V., Rathmell, J.C., Plas, D.R., Elstrom, R.L., June, C.H., and Thompson, C.B. 2002. The CD28 signaling pathway regulates glucose metabolism. *Immunity* 16:769–777.
91. Soleimani, L., Roder, J.C., Dennis, J.W., and Lipina, T. (2008). Beta *N*-acetylglucosaminyltransferase V (Mgat5) deficiency reduces the depression-like phenotype in mice. *Genes Brain Behav* 7:334–343.
92. Manji, H.K., Drevets, W.C., and Charney, D.S. 2001. The cellular neurobiology of depression. *Nat Med* 7:541–547.
93. Davidson, R.J., Pizzagalli, D., Nitschke, J.B., and Putnam, K. 2002. Depression: Perspectives from affective neuroscience. *Annu Rev Psychol* 53:545–574.
94. Ozaslan, D., Wang, S., Ahmed, B.A., Kocabas, A.M., McCastlain, J.C., Bene, A., and Kilic, F. 2003. Glycosyl modification facilitates homo- and hetero-oligomerization of the serotonin transporter. A specific role for sialic acid residues. *J Biol Chem* 278:43991–44000.
95. Horschitz, S., Hummerich, R., and Schloss, P. 2001. Down-regulation of the rat serotonin transporter upon exposure to a selective serotonin reuptake inhibitor. *Neuroreport* 12:2181–2184.
96. Mialet-Perez, J., Green, S.A., Miller, W.E., and Liggett, S.B. 2004. A primate-dominant third glycosylation site of the beta2-adrenergic receptor routes receptors to degradation during agonist regulation. *J Biol Chem* 279:38603–38607.
97. Holmes, A., Yang, R.J., Murphy, D.L., and Crawley, J.N. 2002. Evaluation of antidepressant-related behavioral responses in mice lacking the serotonin transporter. *Neuropsychopharmacology* 27:914–923.
98. Haller, J., Bakos, N., Rodriguiz, R.M., Caron, M.G., Wetsel, W.C., and Liposits, Z. 2002. Behavioral responses to social stress in noradrenaline transporter knockout mice: Effects on social behavior and depression. *Brain Res Bull* 58:279–284.
99. Warren, C.E., Krizus, A., Roy, P.J., Culotti, J.G., and Dennis, J.W. 2002. The *Caenorhabditis elegans* gene, gly-2, can rescue the *N*-acetylglucosaminyltransferaseV mutation of Lec4 cells. *J Biol Chem* 277:22829–22838.
100. Cipollo, J.F., Awad, A.M., Costello, C.E., and Hirschberg, C.B. 2005. N-glycans of *Caenorhabditis elegans* are specific to developmental stages. *J Biol Chem* 280:26063–26072.
101. Boehmelt, G., Wakeham, A., Elia, A., Sasaki, T., Plyte, S., Potter, J., Yang, Y., Tsang, E., Ruland, J., Iscove, N.N., et al. 2000. Decreased UDP-GlcNAc levels abrogate proliferation control in EMeg32-deficient cells. *EMBO J* 19:5092–5104.

102. Ioffe, E. and Stanley, P. 1994. Mice lacking *N*-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proc Natl Acad Sci USA* 91:728–732.
103. Wang, Y., Tan, J., Sutton-Smith, M., Ditto, D., Panico, M., Campbell, R.M., Varki, N.M., Long, J.M., Jaeken, J., Levinson, S.R., et al. 2001. Modeling human congenital disorder of glycosylation type IIa in the mouse: conservation of asparagine-linked glycan-dependent functions in mammalian physiology and insights into disease pathogenesis. *Glycobiology* 11:1051–1070.

Part VI

Recognition and Effector Functions in Innate Immunity

22 Mannan-Binding Lectin Polymorphisms and Infectious Diseases

*Mette Møller-Kristensen, Steffen Thiel, and Jens Chr. Jensenius**

CONTENTS

22.1 INTRODUCTION

Experiments indicate that mannan-binding lectin (MBL) functions as a complement activating protein in the innate immune defense. This is supported by clinical observations implicating its role in the defense against severe infections in children and adults [1–3]. The binding of MBL to micro-

* *Conflicts of interest*: S.T. and J.C.J. have financial interests in NatImmune A/S, established for the production of recombinant MBL.

organisms has been intensively studied and include many human pathogens [4–6]. In the following sections, we have attempted to collect and discuss most of the published studies on MBL deficiency and susceptibility to infections. For discussions of association of MBL deficiency with other diseases, we refer to other recent reviews [7–10].

In 1978, a mannan-binding protein found in rabbit liver was isolated and characterized by Kawasaki et al. [11], and 5 years later, human MBL was isolated and characterized by the same group [12]. MBL appears to be present ubiquitously in mammalian serum as well as in chickens. The protein was originally named mannan-binding protein. Later the term mannan- or mannose-binding lectin was adopted.

MBL is a plasma protein primarily produced by hepatocytes, but some extrahepatic transcription in small intestine and testis has also been reported [13]. In mice expression, an additional tissue has been found [14]. As a member of the human collectin family, MBL is characterized by the presence of a collagen-like region and a calcium-dependent carbohydrate recognition domain (CRD) [15,16]. The N-terminal region of human MBL contains three cysteines, which are responsible for cross-linking three identical polypeptides into a structural subunit. The following collagen-like region imposes the trimeric structure, which is a prerequisite for the assembly of the range of high-order oligomers appearing in serum.

MBL-mediated complement activation was first described by Ikeda et al. [17], who demonstrated that mannan-coated erythrocytes opsonized with rat MBL were susceptible to serum hemolysis. It is now evident that the lectin pathway can be initiated by both MBL [17,18] and by ficolins [19,20] in complex with MBL-associated serine proteases (MASP). A range of microorganisms, including bacteria, viruses, fungi, and parasites, present pathogen-associated molecular patterns (PAMPs), for example, mannose and *N*-acetyl-glucosamine-rich polysaccharides for MBL binding or acetylated carbohydrates for ficolin binding [15,21]. Upon ligand binding, MBL and ficolins presumably undergo conformational changes allowing the associated MASPs to be activated. In view of its biological activity, its seems obvious to speculate that MBL will express a role *in vivo* in the antimicrobial defense, and indeed, a correlation between low MBL levels and recurrent infections was reported in a seminal paper from Super et al. in 1989 [22].

22.2 GENOTYPES ASSOCIATED WITH MBL DEFICIENCY

Interindividual variation of MBL levels range from a few ng/mL to 10 μg/mL [23] (Figure 22.1), and MBL deficiency is recognized as the most common human immunodeficiency state. MBL deficiency is caused by single nucleotide polymorphisms (SNPs) within the promoter region and in exon 1 of the *mbl2* gene (*mbl1* is a pseudogene in humans).

In 1991, Sumiya et al. sequenced the *mbl2* gene from three unrelated British children with opsonic defects, recurrent infections, and low MBL levels [24]. All three children presented a point mutation in codon 54, which was later found also in three healthy Chinese adults with low MBL levels [25]. Screening a group of Gambian individuals with MBL deficiency next revealed a point mutation in codon 57 [26]. These mutations, codon 54 and codon 57, are found almost exclusively in non-Africans and Africans, respectively. Some Caucasians and Africans with low MBL levels showed none of these mutations, but revealed a mutation in codon 52 [27]. The three mutations in exon 1 at codon 54, 57, and 52 are also known as B, C, and D allotypes (collectively, O allotypes), while A indicates the wild-type MBL allele. In the B and C variants, the substitution of a glycine residue in the collagen-like region leads to the disruption of the Gly-X-Y triplets, suggested to interfere with the assembly of higher molecular forms of MBL. The D variant introduces a cysteine that might disrupt the oligomer formation by additional disulfide bonds. Rodents and other animals have two MBL isotypes, MBL-A and MBL-C. Mutations made in rat MBL-A showed, for all three substitutions, an impaired formation of higher oligomers and a decreased ability to interact with complement components, leading to loss of activity [28]. Serum from individuals homozygous for any of the three structural polymorphisms, or compound homozygous, e.g., B/C, O/O, does not contain higher order MBL oligomers [29].

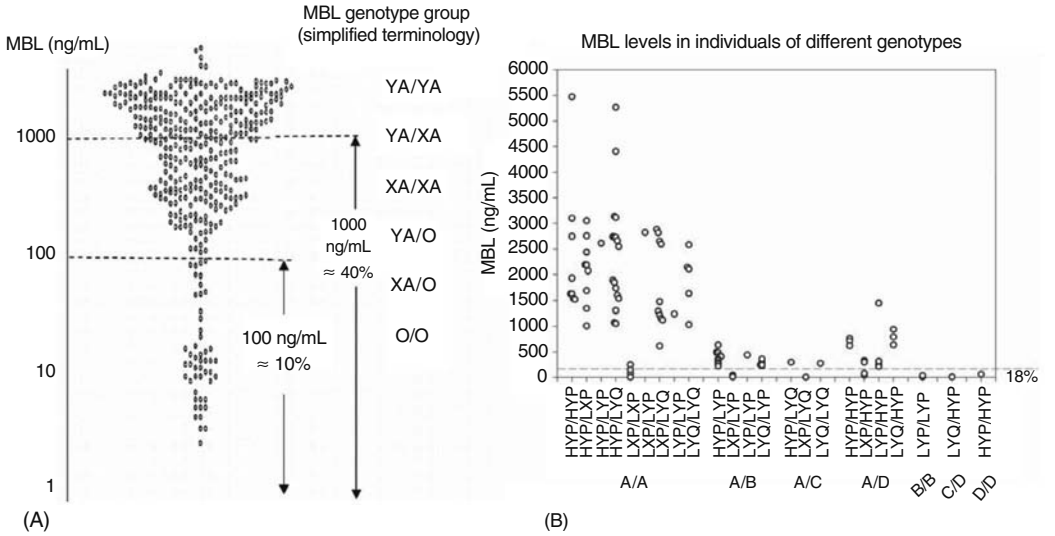


FIGURE 22.1 Association between promoter and structural polymorphisms in the human *mbl2* gene and the serum levels of MBL. (A) The MBL concentration in serum is linked to the MBL genotype. Each dot represents a healthy Danish blood donor. Three structural polymorphisms, termed B, C, and D (simplified by O), with A referring to the wild-type MBL, are found in exon 1 and X is a promoter variant. The genotypes are situated next to the MBL levels in order of influence on MBL levels. (Modified from Gadjeva, M., Takahashi, K., and Thiel. S., *Mol. Immunol*., 41, 113, 2004.) (B) MBL levels in individuals of different MBL genotypes. H/L, X/Y, and Q/P are promoter variants. Eighteen percent of the individuals are below the dashed line at 200 ng MBL/mL. (Modified from Steffensen, R., Thiel, S., Varming, K., Jersild, C., and Jensenius, J.C., *J. Immunol. Methods*, 241, 33, 2000.)

Identification of promoter polymorphisms showed MBL levels to be modulated also at the transcriptional level. Two polymorphisms positioned at -550 (H/L variant) and at -221 (X/Y variant) in the promoter region of *mbl2* [30] are found in linkage disequilibrium, resulting in the three promoter haplotypes, HY, LY, and LX, associated with high, medium, and low MBL concentration in serum, respectively. The three exon 1 mutations are in strong linkage disequilibrium with the promoter haplotypes and only a few of the theoretically possible MBL haplotypes are reported: i.e., seven common haplotypes, HYPA, LYPA, LYQA, LXPA, LYPB, LYQC, and HYPD [30,31]; and two rare haplotypes, HXPA [32] and LYPD [33]. Heterozygocity for one of the structural variants B, C, or D results in decreased MBL levels, while homozygocity or compound homozygocity of structural variant allotypes, or one of these together with the LXA haplotype, e.g., LXA/B, results in MBL deficiency [23,30] (Figure 22.1). Other MBL polymorphisms have recently been discovered [34] but the importance of these remains to be elucidated.

22.3 MBL AND DISEASE

Consensus is not yet been reached on the definition of MBL deficiency, but MBL concentrations below 100 ng/mL are in general regarded as pronounced MBL deficiency. Approximately 10% of Caucasians are in this group (Figure 22.1). Clinical studies indicate that MBL deficiency is linked to increased susceptibility to infections, may be present at MBL levels below 0.5 µg/mL or even up to 1 µg/mL [35,36]. Worldwide, individuals homozygous for structural variants (B, C, or D) or heterozygous for a structural variant combined with the LXA promoter comprise 5–20% of the population.

In 1989, Super et al. reported that defects in phagocytosis were associated with low levels of MBL [22]. Since then, several diseases have been found to be associated with MBL deficiency,

including infections, autoimmune diseases, recurrent miscarriages, and cystic fibrosis (reviewed in Refs. [7,37–40]). In other clinical situations such as transplantations and ischemia/reperfusion injuries, high MBL levels may be aggravating the pathology [41–43]. The high prevalence of MBL deficiency suggests that MBL may promote, rather than limit, infections with some pathogens.

22.3.1 Bacterial Infections

22.3.1.1 SIRS and Sepsis

As defined by Bone et al., a patient is diagnosed with systemic inflammatory response syndrome (SIRS) when at least two out of four signs are present: fever, elevated heart rate (tachycardia), elevated respiratory rate (tachypnea), or an abnormal white blood cell count [44]. Sepsis, an acute systemic response generally caused by bacterial endotoxins, is clinically defined as SIRS together with an infection. The response to infections varies between individuals, and even when given optimal medication, a small number of patients develop severe sepsis that may result in multiple organ failure and death. In general, a septic response occurs when the immune defense fails to contain the invading pathogen, and an unbalanced immune system is suggested to play a role in development of sepsis (reviewed in Ref. [45]). Some data indicate that MBL may be a player, a modifier, in this mechanism (Table 22.1). Thus in pediatric patients admitted to the intensive care unit (ICU), Fidler et al. [46] found significant association between variant MBL genotype, low MBL levels, and SIRS. A strong association was also observed between low MBL levels and the development of sepsis: 12 out of 16 patients with severe sepsis were MBL-deficient. Contrary, no correlation between MBL deficiency and SIRS was seen in a group of adult patients admitted to the ICU [47]. However, also in this study, MBL deficiency was significantly associated with the development of sepsis, severe sepsis, and septic shock.

In a prospective study of critically ill Caucasian patients diagnosed with SIRS, the prevalence of infections at the admission to the ICU was found to correlate with MBL-variant allotypes [48], but patients with MBL variants were not found to have significantly increased rates of sepsis, septic shock, or death. In another investigation of Caucasians diagnosed with sepsis, a significantly increased frequency of hetero- or homozygocity for MBL variants was found in patients as compared with healthy controls [49]. In this study, immunosuppressed patients, patients diagnosed with human immunodeficiency virus (HIV), hematological malignancy, neutropenia, chronic liver failure, or patients receiving chemotherapy were excluded from the study, thus demonstrating that MBL deficiency may influence the susceptibility to sepsis in nonimmunocompromised adults [49]. Acute respiratory distress syndrome (ARDS) is often caused by sepsis and Gong et al. recently reported a significantly increase of septic shock in ARDS patients homozygote for the MBL B variant [50].

Association between MBL polymorphisms and the development of neonatal sepsis was found in a German cohort of very low birth-weight infants [51]. One-third of the septic patients (16 out of 50) were found to be either hetero- or homozygous for the MBL variant B, C, or D. Recently, Eisen et al. [52] reported no difference in the MBL levels when comparing sepsis patients with healthy controls, however, the MBL/MASP function determined by C4b deposition was significantly lower in the patients ($p < 0.001$). The described studies involve different patient populations, which may be one of the reasons for the disagreement on the influence of MBL on the development of SIRS and the severity of sepsis. While further investigations are certainly called for, our interpretation is that the studies quoted point to a significant role of MBL in the susceptibility to infection.

22.3.1.2 *Neisseria meningitidis*

Potential pathogenic strains of *N. meningitidis* are carried as commensals in a small group of individuals, but only few individuals develop meningococcal diseases. Complement deficiencies, in particular, in the components of the membrane attack complex (C5 to C9), are found at a high frequency in these patients [53]. MBL binding to *N. meningitidis* has been investigated intensively [4,54–58]. Recently,

TABLE 22.1
MBL Variants and MBL Levels in SIRS and Sepsis

Patients	Controls	*p*-Value	Age	Comments	Reference
SIRS					
272	250	NS[a]	Adults	DF ↑ MBL variants →	[47]
100[b]	Internal	<0.0001	Children	DF ↑ MBL variants ↑	[46]
		<0.0001		DF ↑ MBL levels ↓	
50[c]	Internal	0.001	Children	DF ↑ MBL variants ↑	[46]
		0.0006		DF ↑ MBL levels ↓	
Sepsis					
197	250	<0.001[d]	Adults	DF ↑ MBL variants ↑	[47]
50[e]	Internal	0.002	Children	DF ↑ MBL variants ↑	[46]
		<0.0001		DF ↑ MBL levels ↓	
50	306	NS[f]	Infants	DF ↑ MBL variants ↑	[51]
222		NS[g]	Adults	DF ↑ MBL variants →	[48]
174	353	0.001[h]	Adults	DF ↑ MBL variants ↑	[49]
		0.05		DF ↑ MBL levels ↓	
195	236	0.27[i]	Adults	DF ↑ MBL levels →	[52]
212	442	0.04	Adults	DF ↑ MBL variants ↑	[50]

Note: NS, Not significant; DF, disease frequency; MBL variants ↑, increasing frequency of the MBL variants; MBL levels ↓, decreasing MBL levels in serum or plasma; →, no change.

[a] MBL deficiency was significantly associated with sepsis, severe sepsis and septic shock but not SIRS.

[b] Pediatric ICU patients with ($n = 59$) and without ($n = 41$) SIRS.

[c] Pediatric ICU patients with infection, with ($n = 35$) and without ($n = 15$) SIRS.

[d] Sepsis patients vs. controls. A decrease of serum MBL concentration was seen with increased severity of sepsis.

[e] Pediatric ICU patients with infection ($n = 50$): sepsis ($n = 19$) or septic shock ($n = 16$).

[f] Seventeen out of 50 sepsis patients were MBL-deficient.

[g] Compared with previously reported Caucasian control groups. Patients with MBL variants had significantly increased prevalence of positive bacterial cultures at admission ($p < 0.02$), but not significantly increased rates of sepsis.

[h] Significant association between MBL variant and susceptibility to sepsis, severe sepsis, and septic shock.

[i] Functional MBL deficiency (determined by C4b deposition assay) was more common in sepsis patients than controls ($p < 0.001$).

Jack et al. [59] showed that MBL was not involved in the initial attachment of *N. meningitidis* to macrophages, but it acts as a "dependent opsonin," accelerating the uptake and killing of *N. meningitidis*. Hibberd et al. [60] genotyped patients diagnosed with meningococcal disease for MBL polymorphisms and compared with children admitted to the hospital with other diseases than infections. Increased frequency of MBL mutations was found in the patients with meningococcal disease (Table 22.2). A similar association was found when comparing a group of meningitis survivors with unrelated controls from the same community. Contrary and unexpectedly when focusing on the group of children with meningitis, MBL deficiency appeared to be associated with less severe disease manifestation. MBL deficiency was also reported to be associated with *N. meningitidis* infections in two family studies. Bax et al. [61] isolated *N. meningitidis* serotype C from a young patient with a family history of meningitis. This patient and two relatives previously diagnosed with meningitis were all found to have low levels of MBL (below 200 ng/mL) and were later genotyped as B/B, B/D, and LXA/B, respectively [62]. In

TABLE 22.2
MBL Variants and Meningitis

Patients	Controls	*p*-Value	Age	Comments	Reference
99[a]	40	NS[b]	Teenagers	DF ↑ MBL variants →	[64]
194	272	0.0006[c]	Children	DF ↑ MBL variants ↑	[60]
72	110	0.06[d]	Children + adults	DF ↑ MBL variants ↑	[60]

Note: NS, Not significant; DF, disease frequency.

[a] Patients who survived severe systemic disease caused by serogroup B or C meningococci.

[b] MBL variants were significantly more frequent in patients younger than 16 years ($p = 0.06$) [65].

[c] Frequency of homozygote variants in a hospital-based study.

[d] Frequency of homozygote variants in a community-based study.

a Danish family, a combination of properdin and MBL deficiency was found to associate with *N. meningitidis* infection. Two out of six properdin-deficient males were found to be MBL-deficient and both were diagnosed with meningitis [63]. MBL levels in Norwegian teenagers who survived severe systemic meningococcal disease caused by either serotype B or C meningococci did not differ from healthy control [64]. However, the frequency of the MBL variants was increased in the patients younger than 16 years [65] (Table 22.2).

The combined information indicates that MBL may be a determinant of susceptibility to meningococcal disease, and vaccination of MBL-deficient individuals may be considered. On the other hand, once infected MBL may aggravate the disease.

22.3.1.3 Pneumonia and Otitis Media

Streptococcus pneumoniae is often harmlessly colonizing the mucosa of the upper respiratory tract, but invasion by *S. pneumoniae* is a major cause of pneumonia and acute otitis media. The incidence of invasive pneumococcal disease is highest among young children and elderly, and is often associated with immunosuppressed conditions. Serotype 14 is a frequent cause of pneumococcal infectious disease and presents repeating *N*-acetyl-glucosamine residues on the surface of the polysaccharide capsule. The presence of this MBL-reactive sugar suggests that MBL-deficient individuals could be more susceptible to pneumococcal pneumonia. However, there are conflicting data on MBL binding to *S. pneumoniae* [5,66].

In a case–control study of patients with invasive pneumococcal disease and matched controls, Roy et al. [67] found a higher frequency of homozygocity for MBL variants in patients than in healthy controls (Table 22.3). Pneumonia was the most frequent diagnosis in the patient group followed by streptococcal meningitis. Another study comparing patients admitted with *S. pneumoniae* bacteremia (the majority with pneumonia) with controls found no association between MBL deficiency and susceptibility to invasive pneumococcal infections [68]. A meta-analysis of the two studies showed a small, but significantly increased risk of invasive pneumococcal disease with MBL deficiency [69]. In HIV-positive patients with *S. pneumoniae* infection, an increased incidence of MBL deficiency was found [70]. A recent Caucasian study showed no association between invasive pneumococcal disease and variant MBL genotypes [71]. The authors conducted a meta-analysis revealing clear correlation between *S. pneumoniae* infection and MBL variants [71]. MBL deficiency did not correlate with recurrent pneumococcal infections in a small group of Colombian children [72].

Studies on the association of otitis media with MBL deficiency (Table 22.3) have also yielded conflicting results. In a study from Turkey, children with recurrent otitis media had a defect in opsonization [73], likely due to MBL deficiency [22]. A tendency of developing atopy or acute otitis media within the first year was observed in a group of babies with the same opsonizing defect [74].

TABLE 22.3
MBL Variant and MBL Levels in Pneumonia and Otitis Media

Patients	Controls	*p*-Value	Age	Comments	Reference
Pneumonia					
229	353	0.002[a]	Children + adults	DF ↑ MBL variants ↑	[67]
108	679	0.046[b]	Children + adults	DF ↑ MBL variants ↑	[67]
140	250	0.083	Adults	DF ↑ MBL variants →	[68]
		$p < 0.0001$[c]	Children + adults	DF ↑ MBL variants ↑	[69]
11		NS	Children	DF ↑ MBL level ↑[d]	[72]
63	162	0.113	Children + adults	DF ↑ MBL variants →	[71]
432	765	0.0001[e]		DF ↑ MBL variants ↑	[71]
Otitis media					
89	123	0.65	Children/adults	DF ↑ MBL variants →	[77]
51	103	<0.001	Children	DF ↑ SOD ↑[f]	[73]
74[g]	Internal	NS	Children	DF ↑ MBL variants →	[81]
252[h]	Internal	<0.001[i]	Children	DF ↑ MBL variants ↑	[80]
44	2060	0.920[j]	Children	DF ↑ MBL levels →	[76]
21	39	0.06	Children[k]	DF ↑ MBL levels ↓	[79]
34	61	NS	Children[l]	DF ↑ MBL levels ↓	[79]
90[m]	Internal	0.29	Children	DF ↑ MBL levels →	[82]
17	172	OR 14.35[n]	Children/adults	DF ↑ MBL variants ↑	[75]
113	internal	0.027[o]	Children	DF ↑ MBL variants ↑	[34]

Note: NS, Not significant; DF, disease frequency; SOD, serum opsonization defect.

[a] Homozygocity for MBL variants, but not heterozygocity or promoter polymorphisms were associated with susceptibility.

[b] Confirmatory study.

[c] Combined analysis of data from Refs. [67,69].

[d] None of the patients had MBL levels below 1 µg/mL.

[e] Combined analysis of data from Refs. [67,69,71].

[f] Defective opsonization of yeast particles was observed in 13.7% of patients with acute purulent otitis media (OM) compared with 2.9% in healthy controls. Five out of 23 children with recurrent purulent OM had defective opsonization ($p < 0.001$).

[g] Community-based study in Greenland.

[h] Population-based, prospective study in Greenland.

[i] The association between MBL deficiency and increased risk of acute respiratory tract infections was primarily restricted to children aged 6 to 17 months.

[j] In a retrospective study of 2104 children, 44 were hospitalized with OM. MBL serum levels below 120 ng/mL were not more frequent in patients with OM.

[k] Children followed from birth to 2 years. Children with recurrent OM had decreased MBL levels.

[l] Children between the age of 2 and 4 years. Children with recurrent OM had decreased serum MBL levels.

[m] Children receiving bilateral tympanostomy tubes for OM with effusion (OME) ($n = 90$), children developing recurrent bilateral OME ($n = 56$), recurrent unilateral OME ($n = 17$) compared with no recurrent OME ($n = 17$).

[n] The MBL XA/O genotype was found in 40% of the OM patients compared with less than 5% of the controls.

[o] A significant association between MBL variants and acute OM was found in children of the age of 12–24 month, but not in children older than 24 months.

An increased frequency of the MBL B variant was found in patients with recurrent otitis media, and combined with the LXPA haplotype, a significantly higher frequency was found compared with healthy controls [75]. In a study of children hospitalized for infections during early childhood, low MBL levels did not correlate with otitis media [76]. In Copenhagen, no correlation with MBL deficiency was found in children with recurrent otitis media [77], but severe recurrent otitis media was

found in a group of MBL-deficient children with non-HIV-related immune deficiencies [78]. In a cohort of children, followed during their first 4 years, MBL deficiency was not found to predispose to otitis media, however a subgroup of children aged 2 to 4 years with recurrent otitis media tended to have lower MBL levels [79]. Children aged 12 to 24 months with variant MBL genotypes were found to have significantly increased number of otitis media episodes [34]. An increased risk of acute respiratory tract infection, and also acute otitis media, was reported in MBL-deficient children from West Greenland [80]. However, a different community study of Greenlandic children found no support for an association between MBL deficiency and otitis media [81]. A recent study of children with recurrent otitis media who received bilateral tympanostomy tubes showed a nonsignificant trend of lower serum MBL levels in cases of recurrence within the follow-up period [82]. The conflicts between the outcomes of the studies might be due to exposure to different *S. pneumoniae* strains with various MBL-binding capabilities, but could also be due to other streptococci or *Hemophilus influenzae*. One might repeat that MBL should be regarded as a disease modifier rather that the sole agent responsible for the clinical picture.

22.3.1.4 Tuberculosis

Immune molecules are used (misused) by pathogens to promote their entrance into the intracellular space and a popular hypothesis proposes to explain the high frequency of MBL variants worldwide, by hypothesizing that low MBL levels might protect against intracellular infections. Inhalation of *Mycobacterium tuberculosis*, an obligate intracellular pathogen, is responsible for most cases of tuberculosis. *M. tuberculosis* express ligands for MBL, i.e., mannose residues, in the lipoarabinomannan membrane, and is reported to take advantage of complement activation during uptake in phagocytes [83–86].

In non-HIV positive tuberculosis, patients from Tanzania significantly higher MBL levels ($p = 0.0067$) were detected when comparing with controls [87] (Table 22.4). Increased serum MBL levels in tuberculosis patients were reported in two other studies [88,89]. Mombo et al. [90] showed a correlation between the frequencies of the MBL deficiency alleles and the incidence of tuberculosis in sub-Saharan Africa. Two other studies found the MBL-variant genes significantly less frequent in patients with tuberculosis or tubercular meningitis [91,92], supporting the hypothesis that variant MBL may be associated with protection against tuberculosis. A similar correlation was found looking at western European population with both HIV and tuberculosis [93]. That high MBL levels may predispose to tuberculosis (or the infection itself is causing the increase in MBL) was not confirmed in a larger study using genotypes to indicate MBL deficiency [94]. Also no differences in MBL genotypes were found in Caucasian or Hispanic tuberculosis patients compared with controls [95]. The use of genetic markers negates the problem of clinical-mediated effect on MBL levels. In contrast, a significant increase of homozygocity for MBL variants was found in Indian tuberculosis patients when compared with a high-risk group of healthy spouses, indicating a protective role of MBL. Curiously, heterozygote patients with inactive tuberculosis showed protection against relapse [96]. No correlation with death rate due to tuberculosis was found. Turner et al. found almost no MBL variants in two groups of indigenous Australians, which they suggest might in part explain the fatal consequences of the introduction of tuberculosis by the European settlers in the nineteenth century [97]. Although there is no consensus, the largest studies could not confirm that MBL contributes to the infections of tubercular bacteria.

22.3.2 Virus Infections

22.3.2.1 Hepatitis

Hepatitis viruses are the major causes of chronic liver disease, including cirrhosis and hepatocellular carcinoma (HCC) caused by hepatitis B virus (HBV) or hepatitis C virus (HCV). Most adults recover

TABLE 22.4
MBL Variants and MBL Levels in Tuberculosis

Patients	Controls	*p*-Value	Age	Comments	Reference
94	113	0.0067[a]	Adults	DF ↑ MBL levels ↑	[87]
397[b]	422	0.037	Adults	DF ↑ MBL C variant ↓	[92]
202[c]	109	0.008	Adults	DF ↑ MBL variant (O/O)↑	[96]
91[d]	79	0.017	Adults	DF ↑ MBL B variant ↓	[91]
64[e]	91	0.002	Children	DF ↑ MBL B variant ↓	[91]
67[f]	187	0.004	Adults	DF ↑ MBL levels ↑	[91]
31[g]	187	<0.001	Children	DF ↑ MBL levels ↑	[91]
109[h]	250	0.032	Adults	DF ↑ MBL levels ↑	[88]
487[i]	232	NS	Adults	DF ↑ MBL variants →	[95]
176[j]	71	<0.01	Adults	DF ↑ MBL B variant ↑	[95]
48	58	0.0085	Adults	DF ↑ MBL levels ↑[k]	[89]
127[l]	344	NS	Adults	DF ↑ MBL variants →	[93]
443	423	NS	Adults	DF ↑ MBL variants →	[94]

Note: NS, Not significant; DF, disease frequency.

[a] The frequency of undetectable MBL in tuberculosis ($n = 191$) patients (Tanzania) did not differ from controls ($p = 0.5$).

[b] Gambian tuberculosis patients.

[c] Indian tuberculosis patients, active ($n = 56$) and inactive ($n = 146$). 10.9% of the patients were homozygote for MBL variants compared to 1.8% of healthy controls.

[d] South African patients with pulmonary TB.

[e] South African patients with tubercular meningitis.

[f] Individuals who recovered from tuberculosis.

[g] Individuals who recovered from tubercular meningitis.

[h] Patients including Caucasians, African-Americans, Asians, and Inuits, compared with Caucasian controls.

[i] Patients and controls include African-Americans, Caucasians, and Hispanics.

[j] African-American TB patients only.

[k] Indian tuberculosis patients had increased levels of serum MBL but less frequent WT genotype (A/A) (50%) than controls (63.8%). A negative correlation between MBL serum level and phagocytosis was found in both patients and controls ($p = 0.019$).

[l] Caucasian non-HIV infected tuberculosis patients.

from an acute infection and it is not fully understood why some develop chronic hepatitis. Numerous studies have examined associations between hepatitis, virus progression, and MBL deficiency, but the results are conflicting (Table 22.5).

Thomas et al. reported the MBL D variant to be associated with chronic HBV infection in Caucasians [98]. However, the frequency of the D variant in the control group of this study was low compared to other studies of Caucasians. In a large Gambian study ($n = 990$), more than 50% of the 337 patients found positive for HBV suffered from chronic HBV. The study failed to detect an association between MBL deficiency and chronic HBV [92]. Similarly, a Caucasian study showed no association between MBL deficiency and patients with chronic HBV infections or in patients who spontaneously recovered from acute infection [99]. In Vietnamese patients with acute HBV infection, a significant increased frequency of the MBL B variant was found, but not in the chronic HBV patients [100]. Furthermore, the MBL B variant was associated with high viral loads and elevated transaminase levels as indications of liver damage [100]. However, two other studies found association, but not significantly, of the B allotype with decreased HBV clearance or disease progression [101,102]. The frequency of the MBL B variant was significantly

TABLE 22.5
MBL Variant and MBL Levels in Hepatitis B and C

Patients	Controls	*p*-Value	Age	Comments	Reference
HBV					
33[a]	98	0.0004[b]	Adults	DF ↑ MBL D variant ↑	[98]
337	653	0.37[c]	Adults + children	DF ↑ MBL C variant →	[92]
61[a]	60	NS	Adults	DF ↑ MBL D variant →	[99]
190[d]	117	NS	Adults	DF ↑ MBL variant →	[102]
		NS		DF ↑ MBL level →	
45[e]	117	0.007	Adults	DF ↑ MBL B variant ↑	[102]
		0.0084		DF ↑ MBL level ↓	
14[f]	117	0.0026	Adults	DF ↑ MBL B variant ↑	[102]
		0.0013		DF ↑ MBL level ↓	
43[g]	260	NS	Adults	DF ↑ MBL B variant →	[105]
123[h]	112	0.079	Adults	DF ↑ MBL B variant ↑	[100]
31[i]	112	0.01	Adults	DF ↑ MBL B variant ↑	[100]
372[j]	126	0.081	Adults	DF ↑ MBL B variant ↑	[101]
320[k]	484	NS	Adults	DF ↑ MBL variant →	[103]
		NS		DF ↑ MBL level →	
199	320	0.002[l]	Adults	DF ↑ MBL variant ↑	[103]
		0.001		DF ↑ MBL level ↓	
189	338	0.04[m]	Adults	DF ↑ MBL variant ↑	[104]
HCV					
93	218	NS	Adults	DF ↑ MBL B variant →	[106]
52[n]	50	NS	Adults	DF ↑ MBL B variant →	[107]
16[o]	22	$p < 0.01$	Adults	DF ↑ MBL levels ↑	[109]
66[p]	50	NS	Adults	DF ↑ MBL levels →	[112]
180	566	$p < 0.0001$	Adults	DF ↑ MBL levels ↑	[108]

Note: NS, Not significant; DF, disease frequency; HBV, hepatitis B virus; HCV, hepatitis C virus.

[a] Caucasian patients with chronic hepatitis B.

[b] A significant difference in the frequency of the MBL D variant was found in Caucasian patients with chronic HBV ($n = 7$) vs. controls ($n = 4$) but not in Asians. No association was found with acute HBV patients. No association with HBV was found for MBL B or C variants.

[c] No association with MBL C or MBL D variants in African HBV patients.

[d] Chinese patients including chronic carriers of hepatitis B or C, as well as asymptomatic HBV or HCV patients, symptomatic HBV-related cirrhosis, patients with spontaneously bacterial peritonitis and HBV-related hepatocellular carcinoma.

[e] Patients with symptomatic cirrhosis.

[f] HBV patients with spontaneous bacterial peritonitis, 64.3% had the MBL B variant.

[g] Japanese patients with fulminant hepatic failure caused by HBV infection.

[h] Vietnamese patients including all HBV infections (acute and chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma).

[i] Patients (Vietnamese patients including all HBV infections [acute and chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma) diagnosed with acute hepatitis B.

[j] Korean patients with HBV persistence including inactive healthy carrier, chronic hepatitis, liver cirrhosis.

[k] Nonprogressed Hong Kong Chinese carriers positive for HBsAg but without clinical symptoms.

[l] Hong Kong Chinese, progressed carriers with hepatocellular carcinoma or cirrhosis vs nonprogressed carriers.

[m] US Patients with persistent HBV had significantly higher frequency of the low producing genotypes (XA/O and O/O) than individuals naturally recovered from HBV infection. Furthermore the group of patients recovered from HBV had significantly higher frequency of the high producing genotype (YA/YA) ($p = 0.005$).

[n] Japanese HCV patients including patients with chronic or active hepatitis and patients with liver cirrhosis.

[o] Japanese HCV patients (negative for HBV) with cryoglobulinemic glomerulonephritis.

[p] Patients with chronic HCV.

increased in chronic HBV patients with cirrhosis and in cirrhosis patients with spontaneous bacterial peritonitis. Measurements of serum MBL levels confirmed the results.

MBL, which is produced by the hepatocytes, could possibly cause an increase in viral load in the liver and thus predispose to subsequent liver damage. In a study of more than 1000 Hong Kong Chinese, MBL deficiency levels and genotyping was associated with HBV-related cirrhosis and also HCC, however, not with asymptomatic chronic HBV infection [103]. This indicates no involvement of MBL deficiency in the development of chronic HBV, but MBL may be a factor in some disease-associated complications.

Thio et al. [104] reported a significant increase of the low-producing MBL genotypes (XA/O, O/O) in the chronic HBV group, whereas the frequency of the high-producing MBL genotype (YA/YA) was significantly increased in the group of patients that recovered from HBV. In a study of HBV patients progressing into fulminant hepatic failure, no difference in MBL genotype was found when compared with healthy controls [105]. However, when dividing the patients into survivors and nonsurvivors, the MBL B variant was found significantly increased in nonsurvivors ($p = 0.043$), and in parallel, the high-producing promoter variant HY and the MBL levels were significantly decreased ($p = 0.0048$).

In contrast to HBV viremia, which is primarily observed in individuals infected early in life, HCV viremia tends to persist even after infection acquired in later life. HCV inhibits the protein kinase PKR, a mediator of interferon activity, and subsequently impairs the ability of interferon to protect the host cells from viral infection. In a Japanese study of 93 patients with chronic HCV infection, half of the patients failed to respond to interferon treatment [106]. The lack of treatment response was correlated with an increased frequency of the MBL B variant ($p = 0.039$), and MBL replacement therapy together with interferon treatment was suggested. Another Japanese study found no difference in MBL genotype between HCV patients and healthy controls [107]. However, all patients who were hetero- or homozygous for the MBL B variant had active hepatitis or liver cirrhosis.

Some studies have found elevated MBL levels in patients with chronic HCV compared with healthy controls [108,109]. The development of cirrhosis was slightly, but nonsignificantly, decreased in the group of low MBL levels, indicating an association between MBL deficiency and less severe hepatitis [108]. Ohsawa et al. [109] reported significantly increased MBL levels in 16 patients with HCV infection and cryoglobulinemia, an extrahepatic complication in HCV, which may result in glomerulonephritis. The high MBL levels could, however, be due to steroid hormone therapy, since growth hormone [110] and more pronounced thyroid hormone [111] induce increase in MBL levels. Increased MBL levels were not observed in another study of HCV patients [112]. Despite the many investigations on hepatitis virus and MBL, a simple conclusion is not yet possible.

22.3.2.2 Human Immunodeficiency Virus

A possible connection between MBL and infection with HIV was first indicated by *in vitro* studies showing that MBL could inhibit infection of target cells. The assay conditions were note quite physiological since HIV was preincubated with purified MBL before adding to a T cell line in culture with heat-inactivated serum [113]. Inhibition of HIV was not confirmed by others [114]. Investigation on MBL ligands on HIV showed that MBL binds to purified envelope protein from in virus-infected T cells, and to recombinant gp120 and gp110 produced in insect cells, but not to the nonglycosylated gp120 expressed in *Escherichia coli* [115]. Removal of the MBL ligands, high mannose, but not removal of non-MBL ligands, complex N-linked carbohydrates, from gp120 with endoglycosidases, abolished MBL binding. MBL binding and HIV neutralization was enhanced by treatment with neuraminidase, which revealed MBL ligands by removal of terminal sialic acid residues from the carbohydrate side-chains, and was also enhanced by the presence of mannosidase I inhibitor under culturing conditions, which results in increase in high-mannose structures [116,117]. Dendritic cells recognize and internalize

TABLE 22.6
MBL Variants and MBL Levels in HIV

Patients	Controls	*p*-Value	Age	Comments	Reference
80	123	0.027	Adults	DF ↑ MBL levels ↓[a]	[124]
173	113	0.009	Adults	DF ↑ MBL variants ↑	[87]
		NS		DF ↑ MBL levels →[b]	
96	123	0.005	Adults	DF ↑ MBL variant (O/O)↑[c]	[70]
111	194	0.01	Adults	DF ↑ MBL variants ↑	[119]
52	27	NS[d]	Children	DF ↑ MBL variants →	[123]
138	140	NS	Adults	DF ↑ MBL variants →	[121]
68	120	NS[e]	Adults	DF ↑ MBL variants →	[90]
114	165	0.0155	Children	DF ↑ MBL variants ↑	[125]
90	120	NS	Infants	DF ↑ MBL variants →	[122]
414	344	0.042[f]	Adults	DF ↑ MBL variants ↓	[93]
145	99	0.0125	Adults	DF ↑ MBL variants ↑	[120]

Note: NS, Not significant; DF, disease frequency.

[a] More persons with MBL levels below 100 ng/mL was found in HIV patients (10%) compared to healthy controls (2.4%). No overall association between MBL serum levels and HIV or progression to AIDS was found.

[b] No difference in MBL serum concentration.

[c] MBL variants were not associated with the progression to AIDS.

[d] HIV-positive children vs. HIV-negative children born from HIV positive mothers.

[e] All individuals homozygous for C/C ($n = 2$) or compound homozygous B/C ($n = 2$) were HIV positive ($p = 0.019$).

[f] HIV patients without tuberculosis vs. controls.

pathogens for presentation to T cells. This may proceed via the "dendritic cell-mediated intercellular adhesion molecule 3-grabbing nonintergrin" (DC-SIGN), which is used by the virus to infect T cells. MBL was found to block the interaction between HIV and DC-SIGN, thus protecting T cells from infection [118].

Although the studies performed *in vitro* indicate a role for MBL in HIV infections, the published clinical studies have shown ambiguous results (Table 22.6). In a study of HIV-infected individuals from Tanzania, 85% with concomitant tuberculosis, MBL deficiency was not more frequent in HIV-infected individuals compared with HIV-negative patients with tuberculosis and healthy controls; however, variant MBL allotypes were present at higher frequency in HIV positive individuals [87]. This result is supported by studies finding homozygocity for MBL variants or compound homozygocity significantly increased in HIV infected patients [90,119,120]. Other studies did not find such a connection [93,121–123]. The first study investigating this question found no overall difference in MBL serum levels in HIV patients compared with a control group. However, a significant higher number of HIV patients had MBL levels below 100 ng/mL as compared with controls [124]. Interestingly, the presence of MBL variants was recently reported to result in early HIV diagnosis, but no association with increased mortality was found [93].

Some groups have addressed the possibility that MBL might play a role in vertical transmission of HIV. In an Italian study, the frequency of homozygocity for MBL variants was higher in HIV-positive than in HIV-negative offspring, however not significantly. (It was significantly higher in HIV-positive mothers than in healthy controls [122].) In a cohort of Brazilian children, the frequency of the homozygous wild-type MBL genotype A/A was significantly lower in HIV-infected children compared with exposed but uninfected children, and the opposite was found

when comparing the frequency of the mutant alleles [125]. By comparing HIV-infected and uninfected children from HIV-positive mothers, a nonsignificant trend of increased MBL B variant [123] and a significant increased frequency of the promoter variant HY was found [126]. Both variants were found associated with rapid progression to acquire immunodeficiency syndrome (AIDS), and recently MBL deficiency was linked to more severe disease in children [127]. Kuhn et al. [128] have recently reported that MBL-deficient infants have a significantly increased risk of acquiring HIV infection from their HIV-positive mothers. However, in a cohort given general immunostimulatory vitamin A supplement, this increased susceptibility vanished, demonstrating the modifying effect of MBL in a group of vitamin A–deficient and thus partly immunodeficient individuals.

The progression of HIV to AIDS was analyzed in a Danish study including HIV-positive and HIV-negative homosexual men [70]. No association was found between MBL levels and the time from a positive HIV test to progression of AIDS, but MBL-deficient patients showed a significant decrease in survival time. In particular, decreased survival is found in MBL-deficient patients with pneumonia, herpes zoster virus, oral candidiasis, or *Cryptosporidium parvum* [70,129]. Another study showed a slower progression to AIDS in HIV patients with MBL variants, but revealed no association between MBL genotypes and survival. The mean survival time was decreased in MBL-deficient AIDS patients with opportunistic infections [130]. Three studies indicate that the infection or the treatment may influence the MBL levels. Thus higher MBL levels were reported in patients with AIDS compared with asymptomatic HIV patients [131,132], and treatment with the highly active antiretroviral therapy (HAART) resulted in an increased MBL serum level that did not reflect viral replication as defined by the HIV RNA levels in plasma [133]. One study found MBL-deficient HIV patients to differ significantly in susceptibility to infections. Fifteen percent of the HIV patients with MBL levels below 650 ng/mL had bacterial pneumonia compared with 8% of those with higher MBL levels ($p = 0.036$) [134]. However, another study found an increased bacterial infection in HIV patients with MBL levels above 100 ng/mL [135].

As presented above, there are many conflicting results, likely partly due to the differences in patient groups and methodology (e.g., few studies consider the pronounced effect of the LXA haplotype).

22.3.2.3 Severe Acute Respiratory Syndrome

Severe acute respiratory syndrome (SARS) originated in southern China in 2002 and spread to Hong Kong and other countries during 2003. It is caused by the SARS-associated coronavirus (SARS-Cov). While SARS is described as a viral pneumonia, the virus also seems to affect the intestines, kidneys, and the liver, and the disease can progress into multiple organ damage [136]. A surface spike protein of SARS-Cov has a high-mannose structure and MBL binding to the virus leads to complement activation [137]. MBL was found to inhibit the infectivity of the virus on rhesus monkey kidney cells, and a large case–control study involving 569 Hong Kong Chinese SARS patients and 1188 healthy controls showed a significantly increased frequency of the MBL B variant, as well as significantly lower MBL levels in SARS patients [137] (Table 22.7). In northern China, susceptibility to SARS was also reported to be associated with the MBL B variant, however, no association was found with disease severity or death [138]. No difference between MBL genotypes was reported in Hong Kong Chinese patients between those recovering from SARS and the critically ill SARS patients or patients who died from SARS [139]. The studies indicate that MBL-deficient individuals are more susceptible to SARS (Table 22.7), but the MBL status did not seem to affect the outcome of the disease.

One might conclude that despite convincing *in vitro* effect of MBL on HIV and SARS, the clinical data do not allow for an easy interpretation. It appears that one needs *in vivo* experiments to resolve the issue of MBL's effect on viral infections.

TABLE 22.7
MBL Variants and MBL Levels in SARS

Patients	Controls	*p*-Value	Age	Comments	Reference
352	392	0.00086	Adults	DF ↑ MBL B variant ↑[a]	[138]
180	200	NS	Adults	DF ↑ MBL variant →	[139]
569	1188	<0.001	Adults	DF ↑ MBL variants ↑[a]	[137]
				DF ↑ MBL levels ↓	

Note: NS, Not significant; DF, disease frequency.

[a] The MBL B variant was associated with susceptibility to SARS but not with disease severity.

22.3.3 Fungal Infections

Fungal infections are usually caused by *Candida* species, in particular *Candida albicans*, which may be found in the commensal flora at the mucous membrane of the mouth and vagina. This opportunist pathogen mainly infects immunocompromised patients, and oral candidiasis is often found in AIDS patients. MBL binds strongly to mannan residues on *C. albicans* [5], promotes killing through complement activation, and is also found to inhibit growth of *C. albicans* [140]. An *in vivo* role of MBL in innate responses to *Candida* infections was suggested when mice transgenic for human MBL were subjected to *C. albicans* infection [141].

22.3.3.1 Candidiasis

Vulvovaginal candidiasis (VVC), an opportunistic mucosal disease, is often caused by *C. albicans*. Recently it was reported that the frequency of the MBL B variant was significantly increased in Chinese women diagnosed with VVC compared with the control group [142] (Table 22.8). Remarkably, the MBL concentration in cervicovaginal lavage was increased more than 10 times in the VVC group. The study included six women with recurrent VVC, defined as more than four incidences per year, and five of these women were heterozygous for the MBL B variant. In a group of Latvian women with recurrent VVC, of which 90% were diagnosed with *C. albicans* infection, 68% of the patients positive for *C. albicans* were either heterozygote or homozygote for the MBL B variant, which was significantly different from the healthy control group [143]. Another study found a nonsignificant tendency of increased MBL deficiency in women who had a history of three or more episodes of VVC [144]. However, binding of MBL to *C. albicans* could not be detected in the presence of cervicovaginal

TABLE 22.8
MBL Variants and MBL Levels in Vulvovaginal Candidiasis

Patients	Controls	*p*-Value	Age	Comments	Reference
42	43	<0.0001	Adults	DF ↑ MBL B variant ↑[a]	[143]
25	23	NS	Adults	DF ↑ MBL CVL levels →	[144]
51	54	0.038	Adults	DF ↑ MBL variants ↑	[142]
		<0.001		DF ↑ MBL CVL levels ↑[b]	

Note: NS, Not significant; DF, disease frequency; CVL, cervicovaginal lavage.

[a] Patients with recurrent vulvovaginal candidiasis (rVVC).

[b] CVL MBL concentration were lower in women with rVVC ($n = 6$) than controls, but the frequency of MBL variants were higher ($p < 0.01$).

lavage from a VVC patient, whereas C3 deposition was detected. This could be due to conditions of low pH, low levels of MBL, MBL consumption by the vaginal yeast, or mutation of the yeast. The authors did not examine if MBL could bind in the presence of cervicovaginal lavage from a healthy woman. The data suggest that MBL may play a role on this mucosal surface. Importance of MBL on intestinal mucosa might be inferred from the association of severe diarrhea with MBL deficiency. Further studies are clearly needed to resolve this question.

22.3.4 Parasite Infections

22.3.4.1 Malaria

Plasmodium falciparum, the parasite responsible for most of the fatal malaria cases, enters the blood stream after multiplying in liver cells and penetrates erythrocytes. MBL is capable of recognizing proteins of *P. falciparum*-infected erythrocytes, but addition of MBL to *P. falciparum* cultures *in vitro* did not inhibit parasite growth [145]. By flow cytometry, a calcium-dependent MBL binding to the membrane of *P. falciparum*-infected erythrocytes was found, and the binding was inhibited by mannose [146]. Curiously, *N*-acetyl-glycosamine, but not mannose or glucose, was found to completely block the *P. falciparum* invasion of erythrocytes [147].

Studies of African children with severe or mild malaria and controls, revealed no difference in the frequency of MBL genotypes [92,146] (Table 22.9). However, a significant increase of parasites in MBL-deficient malaria patients with complicated disease was found, and MBL could be a possible regulator of *P. falciparum* parasitemia [146]. In African children with severe malaria, defined as more than 250,000 *P. falciparum*/μL blood, a significant higher frequency of patients were found to have MBL levels below 200 ng/mL as compared with children with mild malaria [148]. In a small

TABLE 22.9
MBL Variants and MBL Levels in Parasite Infection

Patients	Controls	*p*-Value	Age	Comments	Reference
Malaria					
504	426	NS	Children	DF ↑ MBL variant →[a]	[92]
100	100	$p = 0.04$	Children	DF ↑ MBL variant ↑[b]	[148]
		$p = 0.02$		DF ↑ MBL levels ↓	
35				DF ↑ MBL levels ↑[c]	[149]
323	228	$p = 0.32$	Children	DF ↑ MBL variant →[a, d]	[146]
131	136[e]	$p = 0.01$	Children	DF ↑ MBL variants ↑[f]	[150]
Leishmaniasis					
34[g]	55[h]	$p < 0.003$	Adults	DF ↑ MBL levels ↑	[152]
		NS		DF ↑ MBL variants ↓	

Note: NS, Not significant; DF, disease frequency.

[a] MBL deficiency was not associated with clinical malaria or the progression to severe malaria.

[b] Patients with severe malaria vs. mild malaria.

[c] Thirty-four out of 35 patients with mild or asymptomatic malaria had the MBL genotype A/A.

[d] MBL-deficient patients with complicated malaria had significantly higher ($p = 0.02$) parasite counts.

[e] Children with uncomplicated malaria.

[f] Association with MBL deficiency and susceptibility to severe malaria.

[g] Patients infected with *L. chagasi*.

[h] Infected asymptomatic individuals.

study, including 35 patients from Papua New Guinea with mild or asymptomatic malaria, more than 90% where found to have wild-type MBL genotype, some with very high MBL levels as compared with a healthy control group [149], possibly indicating a controlling effect of MBL. Another study revealed association between MBL deficiency variants and severe malaria, but not with mild malaria and the author suggest that malaria could have been one of the evolutionary driving forces shaping the high levels of MBL polymorphisms in the African population [150]. While the literature on involvement of MBL in malaria is still inconclusive (Table 22.9), the findings appear to indicate that MBL has some protective relevance rather than MBL deficiency being advantageous.

22.3.4.2 Leishmaniasis

Leishmania are obligate intracellular parasites replicating in the host macrophages. It is transmitted by sand flies and causes several distinct types of diseases of which visceral or cutaneous lesions dominate. The surface of *Leishmania* parasites is rich in mannose residues, which are thought to play a role for the entrance into the macrophage through the macrophage mannose receptor. MBL was found to bind to different Leishmania strains, to live parasites as well as to lipophosphoglycans purified from the promastogote state, and the binding could be inhibited by mannose [151]. In a Brazilian case–control study, patients with visceral leishmaniasis caused by *Leishmania chagasi* had significantly higher MBL levels compared with infected but asymptomatic patients or healthy controls [152]. An increased frequency of the MBL A allotype was found among the patients, and MBL levels were found to correlate with the probability of disease development (Table 22.9). Possible explanations for these findings are that increased MBL levels enhance the targeting of the organism to the host phagocytes, as postulated for tuberculosis. Paradoxically, the parasite might use MBL to escape from complement killing and hide within phagocytes. Such a mechanism was recently indicated by studies *in vitro* with *Leishmania braziliensis*, which is causing mucocutaneous leishmaniasis. Binding of MBL to the parasite was shown, but MBL was not necessary for complement lysis and MBL alone did not lead to damage of the parasite [153].

22.4 MBL AND INFECTIONS IN IMMUNOCOMPROMISED PATIENTS

22.4.1 Neutropenia

A serious side effect of chemotherapy is neutropenia. In accordance with the view of MBL as a disease modifier, MBL deficiency or MBL insufficiency is associated with increased susceptibility to infections during or after chemotherapy. However, not all studies reveal this connection (Table 22.10). The first investigation in this area found that among adults undergoing chemotherapy for leukemia, patients who

TABLE 22.10
MBL Variants and MBL Levels in Immunocompromised Patients with and without Infections

Patients	Controls	*p*-Value	Age	Comments	Reference
54[a]	Internal	<0.0001[b]	Adults	CSI ↑ MBL level ↓	[35]
100[c]	Internal	0.014[d]	Children	DFN ↑ MBL variants ↑	[36]
		0.012		DFN ↑ MBL level ↓	
97[e]	Internal	0.007	Adults	MI ↑ MBL level ↓	[162]
107[f]	Internal	NS	Adults	MI ↑ MBL levels →	[167]
80[g]	Internal	NS	Adults	IF ↑ MBL levels→	[157]
		NS		DF ↑ MBL levels →	

TABLE 22.10 (continued)
MBL Variants and MBL Levels in Immunocompromised Patients with and without Infections

Patients	Controls	*p*-Value	Age	Comments	Reference
128[h]	Internal	<0.05[i]	Adults	MI ↑ MBL level ↓	[159]
		NS		DF ↑ MBL levels →	
113[j]	Internal	0.0016	Adults	MI ↑ MBL variants ↑	[168]
49[k]	Internal	0.01[l]	Adults	MI ↑ MBL variants ↑	[169]
225[m]	Internal	NS	Adults	DF ↑ MBL levels →	[158]
225	Internal	0.008	Adults	CSI ↑ MBL level ↓	[158]
66[n]	44	NS	Children	FN ↑ MBL variants →	[154]
		NS[o]		FN ↑ MBL level →	
137[p]	Internal	NS	Children	IF ↑ MBL variants →	[156]
106[q]	Internal	NS[r]	Adults	MI ↑ MBL variants →	[164]
133[s]	Internal	NS	Adults	CSI ↑ MBL variants →	[160]
113[t]	Internal	0.02	Adults	CSI ↑ MBL variants ↑	[161]
94[u]	Internal	0.003[v]	Children	CSI + FN ↑ MBL level ↓	[155]

Note: NS, Not significant; CSI, clinical significant infections; DFN, days of febrile neutropenia; FN, febrile neutropenia; IF, infection frequency; DF, days of fever; MI, major infection; AML, acute myeloid leukemia, ALL, acute lymphoid leukemia.

[a] Patients including multiple myeloma (n = 18), non-Hodgkin's lymphoma (n = 13), AML (n = 7), chronic lymphocytic leukemia or Hodgkin's lymphoma (n = 4), chronic myeloid leukemia (n = 2), Burkitt's lymphoma (n = 1), Waldenström's macroglobulinemia (n = 1), ALL (n = 1), aplastic anemia (n = 1).

[b] Patients with less than 500 ng MBL/mL were more susceptible to CSI.

[c] Patients including ALL (n = 55), AML (n = 12), non-Hodgkin's lymphoma (n = 18), neuroblastoma (n = 9), severe aplastic anemia (n = 4), and myelodysplastic syndrome (n = 2).

[d] Patients with MBL variant (A/O, O/O) and MBL levels below 1000 ng MBL/mL had significant higher numbers of days of febrile neutropenia.

[e] Patients receiving allogenic hemopoietic stem cell transplantation.

[f] Patients with acute (n = 39) or chronic leukemia (n = 68).

[g] Patients with AML.

[h] Patients including non-Hodgkin's lymphoma (n = 38), AML (n = 29), multiple myeloma (n = 25), ALL (n = 13), Hodgkin's disease (n = 12), chronic myeloid leukemia (n = 3), chronic lymphocytic leukemia (n = 2), myelodysplastic syndrome (n = 2), amyloidosis (n = 1), myelofibrosis (n = 1), and germ cell malignancy (n = 1).

[i] Patients with MBL levels below 100 ng/mL had significantly more major infections than the group with no infections.

[j] Patients including non-Hodgkin's lymphoma (n = 66), AML (n = 25), ALL (n = 10), Hodgkin's lymphoma (n = 6), multiple myeloma (n = 4), neuroblastoma (n = 1), synovial sarcoma (n = 1) all treated with high-dose chemotherapy and autologous peripheral blood stem cell transplantation.

[k] Patients undergoing liver transplantation.

[l] MI in patients receiving a liver from donors with MBL-variant genotype (O/O) compared with MBL genotype (A/A).

[m] Patients with hemotological malignancy.

[n] Patients (n = 110) including lymphoma (n = 17), hematological malignancies (n = 41, 37 with ALL), solid tumors (n = 52). After chemotherapy the patients were divided into febrile neutropenic (n = 66) and nonfebrile (n = 44).

[o] Comparing patients with MBL levels below and above 1000 ng MBL/mL.

[p] Patients with non-B ALL.

[q] Patients receiving allogenic hemopoietic stem cell transplantation.

[r] No association between variant MBL genotypes (or MASP-2 genotypes) and bacterial or viral infections. Low donor MBL levels were associated with the risk of invasive fungal infection (p = 0.018).

[s] Patients with multiple myeloma.

[t] Patients receiving high-dose chemotherapy and autologous stem cell transplantation.

[u] Patients diagnosed with a malignancy.

[v] Increased susceptibility to febrile neutropenia with severe bacterial infections in patients with MBL levels below 100 ng/mL compared with patients with MBL levels between 100 and 999 ng/mL.

acquired clinically significant infections, defined as bacteremia, pneumonia or both, presented lower MBL levels than patients who did not get serious infections [35]. Patients with MBL levels below 0.5 μg/mL were at particular risk for severe infections. Significant association between low MBL concentration and severe infections related to chemotherapy was at the same time reported in a study of children with various types of cancers [36]. The patients had at least one episode of febrile neutropenia independent of the MBL genotype. However, twice as many days of febrile neutropenia was observed in patients with MBL levels below 1 μg/mL as compared with children having higher levels. Further, it was found that three out of four children admitted to the ICU with severe infections presented variant MBL allotypes. A later study of pediatric febrile neutropenia found no association between MBL deficiency and the severity of infection [154]. The neutropenia was more severe in the children of this study, and only one febrile episode per patient was studied. Significantly, in this study, all children admitted to the ICU with septic shock had MBL concentrations below 1 μg/mL. Recently, Schlapbach et al. [155] reported that MBL deficiency (defined as MBL levels below 100 ng/mL) was associated with an increased susceptibility to febrile neutropenia with severe bacterial infection in pediatric cancer patients. A study focused on children with acute lymphoblastic leukemia saw no association between MBL deficiency and infection during the first 50 days of chemotherapy [156]. But, it was noted that two out of four patients who died because of sepsis were MBL-deficient. In adults undergoing chemotherapy for acute myeloid leukemia, no association between MBL levels and the incidence or duration of fever, the occurrence of sepsis, or survival was found [157]. In a study of adult hematological cancer patients, the rate of severe infections was significantly increased in MBL-deficient patients, while MBL did not seem to influence the number of febrile days [158]. A Scottish study on a mixed population of adults receiving chemotherapy as treatment for hematological malignancies and as preparation for bone marrow transplantation, showed no overall correlation of MBL levels on rates or the severity of infections [159]. However, all patients with MBL levels below 0.1 μg/mL suffered from infections, and a significant higher incidence of major infections. A recent study on MBL genotypes in patients treated for multiple myeloma revealed no apparent protective effect of MBL against severe infections in general, although a protective effect against septicemia was indicated [160]. However, when focusing on patients treated with high-dose melphalan and autologous stem cell transplantation (ASCT) multiple myeloma patients with MBL variants had a significantly higher risk of septicemia during the ASCT procedure [161].

Bone marrow transplantation is the only curative therapy for a number of hematological diseases, and neutropenia caused by the aggressive chemotherapy often results in severe infections. MBL variants were reported to be associated with major infections after allogeneic hemopoietic stem cell transplantation [162] and another study found the MBL genotype HYA (associated with high MBL levels) apparently protective against infections [163]. Surprisingly, also the MBL genotype of the donor was reported to be important with transplantation from an MBL-deficient donor resulting in higher risk of infection [162,164]. MBL synthesis by the transplanted cells was suggested and this supposition appears to be supported by findings of MBL synthesis and expression on the cell surface of human monocytes and dendritic cells [165]. However, donor cell MBL production was not confirmed in a study of hematopoietic stem cell transplantation from MBL sufficient donors to MBL-deficient patients [166]. In a study of MBL genotyping of 107 patients with acute or chronic leukemia and their respective donors, no association between MBL deficiency and posttransplant infections was found [167]. Another study reported an increased risk of major bacterial infections in MBL-deficient patients when treated with high-dose chemotherapy followed by autologous peripheral blood stem cell transplantation [168]. Major infections, such as sepsis, pneumonia, and bacterial meningitis, was found in four out of 12 MBL-deficient patients, compared with only six out of 101 MBL-sufficient patients.

In a study of a different patient population, Bouwman et al. [169] found a strong association between donor MBL genotype and risk of severe infections in patients receiving standard immunosuppressive therapy followed by orthotopic liver transplantation. One year after transplantation, four out of six patients who received a liver from a donor with homozygous-variant MBL genotype (O/O) had severe infections compared with three out of 25 patients receiving a liver from a donor with the MBL wild-type genotype (A/A) ($p = 0.01$).

TABLE 22.11
MBL Variants and MBL Levels in Patients with and without Postoperative Infections

Patients	Controls	*p*-Value	Age	Comments	Reference
172[a]	Internal	0.013	Adults	POI ↑ MBL level ↓	[172]
611[b]	Internal	NS	Adults	POI ↑ MBL levels →	[174]
87[c]	524	0.01	Adults	POP ↑ MBL level ↓	[174]

Note: NS, Not significant; POI, postoperative infections; POP, postoperative pneumonia.

[a] Patients undergoing major elective gastrointestinal surgery for malignant disease.

[b] Patients undergoing major elective surgery for colorectal cancer.

[c] Patients (those undergoing major elective surgery for colorectal cancer) developing POP had low MBL serum levels, and decreased survival ($p = 0.003$).

Recently, neutropenic MBL knockout mice were found to be susceptible to i.v. infection with *Staphylococcus aureus* [170]. When the bacteria were administrated i.p., an effect of MBL was seen only after rendering the mice neutropenic by administration of cyclophosphamide.

While conflicting results have thus been reported, it is our contention that the reports on MBL deficiency as a major modifier in neutropenia outweigh the lack of association seen in other studies. It is possible that difference in patient groups with respect to age, disease, and chemotherapy regimen may account for the contradicting results [171].

22.4.2 Major Surgery

Major surgery is another setting where the immune system is suppressed, and postoperative infections have been associated with low MBL levels (Table 22.11). A significantly increased frequency of low MBL levels was determined in patients who developed postoperative infections after elective gastrointestinal surgery for malignant disease [172]. Furthermore, low preoperative MBL levels, and high procalcitonin levels, a sensitive marker for bacterial infection in the early postoperative phase may predict the occurrence of postoperative infections [173]. Another study encompassing a large cohort found significant association of MBL deficiency and the development of postoperative pneumonia, and postoperative pneumonia resulted in an increased recurrence rate and a poorer survival [174]. For infections other than pneumonia, the association with MBL levels was not significant in this study.

22.4.3 Burn Injury

Using a mouse model, we have recently found association between MBL deficiency and susceptibility to infections with *Pseudomonas aeruginosa* after burn injury [175]. Burns result in immune suppression and, in addition, disrupt the skin barrier, which is a large area of nonspecific defense. This greatly increases the risk of infection by providing microorganisms ready access to the interior of the body, and *P. aeruginosa* infections occur frequently among burn victims.

22.5 MBL REPLACEMENT THERAPY

The idea of MBL replacement therapy is based on early findings of a familial plasma-associated defect of phagocytosis [176]. Infusion of fresh plasma to patients with this opsonic defect improved the conditions [176–179]. In 1998, the first studies of reconstitution with human plasma derived–MBL in MBL-deficient individuals was reported [180], and plasma-derived MBL was used in a

safety and pharmacokinetic phase I study also in Iceland [181]. The mean half-life of the transfused MBL was 3 days (varied from 1 to 7 days between individuals), and to maintain an MBL level of at least 1 μg/mL, the infusions of 0.1 mg MBL/kg body weight is required.

Concerns of possible viral or prion contamination and the high prevalence of MBL deficiency prompted the production of recombinant MBL (rMBL) in a human embryonic kidney (HEK) cell line [182]. The HEK cells expressed a wide range of rMBL oligomers as seen also by other investigators [183]. However, affinity chromatography on mannose-derived TSK beads allowed the purification of MBL oligomers corresponding to the plasma-derived MBL used for therapy [184,185]. The rMBL was safely administrated in healthy individuals with MBL levels below 0.5 μg/mL. A single i.v. dose of 0.5 mg MBL/kg body weight increased the MBL level to approximately 10 μg/mL with a half life of 30 h [186].

Clinical investigations point to patient groups that may benefit from MBL therapy. Under normal conditions most MBL-deficient individuals show no clinical symptoms, which indicates that an additional immune disorder needs to be present for manifest increased susceptibility to infections. Positive effects of rhMBL have already been reported in different animal models [170,175,187], suggesting that MBL therapy may be effective as prophylaxis or treatment particularly in immunocompromised patients. MBL-deficient patients with acute or recurrent infections may also benefit from MBL therapy. While the review has focused on infection, other studies indicate that rMBL might be envisaged as a disease-modifying drug in autoimmune or chronic diseases.

22.6 CONCLUSION AND FUTURE DIRECTIONS

It can be quite bewildering trying to get a grasp of the literature on MBL and disease even when focusing on infections. In this chapter, we have tried to give a comprehensive overview of clinical investigations on MBL and infections without too much discrimination with regard to number of cases or the quality of the reports. Comparing results from different groups are difficult on several levels: Patient definitions and treatments not only vary, but also the definitions of MBL deficiency in terms of MBL levels in serum or plasma and in terms of deficiency-associated allotypes or genotypes vary. MBL deficiency was in early reports simply determined as MBL levels below detection limit. With accumulating clinical information, it has become customary rather to use levels of somewhere between 200 and 1000 ng/mL as cut-off. However, one would be well advised to regard these levels as only possibly useful guidelines, since the level at which MBL becomes clinically relevant may well vary with the disease studies. This means that nonparametric analyses of the data may be pertinent. One should certainly never use the means of MBL concentrations, as this is forbidden when analyzing for components, which, like MBL, show clearly non-Gaussian distributions. How MBL is estimated is another variable, and unfortunately, no internationally accepted standard exists. Most sandwich assays appear to measure about the same levels as obtained by assays employing a mannan-coat as the first layer [188]. This is fortunate since poorly oligomerized MBL can be found in sera from O/O individuals, mostly below 0.5 μg/mL, but reportedly even higher. This aberrant MBL has no known biological effector function. In terms of genotyping, one should be aware that only some of the more recent papers include analysis for the XL promotor allotype. This is quite critical, since the gene frequency of this variant allotype (always linked to structural allotype A) is high in Caucasians (about 20%), and in conjunction with a variant coding allotype on the other chromosome always result in very low MBL levels.

Inherited MBL deficiency is the most common congenital immunodeficiency with a frequency of 10%–20% of Caucasians and even higher in other populations [189]. Clearly, the vast majority of these individuals suffer no problems from this deficiency. The immune defence is beautifully redundant. MBL, like other complement components, should be regarded as a disease modifier, rather than as the determining cause of the observed clinical symptoms. This also means that epidemiological studies focused on nonselected populations face difficulties in detecting correlations between diseases and

MBL genotypes or levels, as was indeed revealed when investigating the so-called Copenhagen City Heart study encompassing 9245 healthy, adult Danes [190]. As detailed above, numerous studies focusing on patient groups admitted to hospital, generally tertiary referral units, have shown convincing associations between MBL deficiency and infections. Such studies enjoy the advantage of concentrating patients with special diagnosis from very large populations and including only severe cases. Thus, when a group of a few hundred patients is analyzed, they may be derived from a population of millions.

Investigations on MBL knockout mice illustrate that MBL has an effect on experimental infections, not least the reconstitution of the wild-type phenotype by infusion with recombinant MBL appears convincing. Caution is of cause always called for when generalizing from murine experiments.

When considering the studies presented in this chapter, it may appear perplexing to arrive at any solid conclusions as the biological effects and clinical meaning of MBL. For those who believe that MBL replacement therapy will show beneficial effects for patients suffering from infections in one or another clinical situation, it is to be hoped that one can narrow down small groups of indications with maximal possibility of effect of such treatment in order that the development of such treatment may proceed.

REFERENCES

1. Summerfield, J. A., M. Sumiya, M. Levin, and M. W. Turner. 1997. Association of mutations in mannose binding protein gene with childhood infection in consecutive hospital series. *BMJ 314*:1229.
2. Summerfield, J. A., S. Ryder, M. Sumiya, M. Thursz, A. Gorchein, M. A. Monteil, and M. W. Turner. 1995. Mannose binding protein gene mutations associated with unusual and severe infections in adults. *Lancet 345*:886.
3. Kakkanaiah, V. N., G. Q. Shen, E. A. Ojo-Amaize, and J. B. Peter. 1998. Association of low concentrations of serum mannose-binding protein with recurrent infections in adults. *Clin Diagn Lab Immunol 5*:319.
4. van Emmerik, L. C., E. J. Kuijper, C. A. Fijen, J. Dankert, and S. Thiel. 1994. Binding of mannan-binding protein to various bacterial pathogens of meningitis. *Clin Exp Immunol 97*:411.
5. Neth, O., D. L. Jack, A. W. Dodds, H. Holzel, N. J. Klein, and M. W. Turner. 2000. Mannose-binding lectin binds to a range of clinically relevant microorganisms and promotes complement deposition. *Infect Immun 68*:688.
6. Jack, D. L. and M. W. Turner. 2003. Anti-microbial activities of mannose-binding lectin. *Biochem Soc Trans 31*:753.
7. Thiel, S., P. D. Frederiksen, and J. C. Jensenius. 2006. Clinical manifestations of mannan-binding lectin deficiency. *Mol Immunol 43*:86.
8. Dommett, R. M., N. Klein, and M. W. Turner. 2006. Mannose-binding lectin in innate immunity: Past, present and future. *Tissue Antigens 68*:193.
9. Bouwman, L. H., B. O. Roep, and A. Roos. 2006. Mannose-binding lectin: Clinical implications for infection, transplantation, and autoimmunity. *Hum Immunol 67*:247.
10. Worthley, D. L., P. G. Bardy, D. L. Gordon, and C. G. Mullighan. 2006. Mannose-binding lectin and maladies of the bowel and liver. *World J Gastroenterol 12*:6420.
11. Kawasaki, T., R. Etoh, and I. Yamashina. 1978. Isolation and characterization of a mannan-binding protein from rabbit liver. *Biochem Biophys Res Commun 81*:1018.
12. Kawasaki, N., T. Kawasaki, and I. Yamashina. 1983. Isolation and characterization of a mannan-binding protein from human serum. *J Biochem (Tokyo) 94*:937.
13. Seyfarth, J., P. Garred, and H. O. Madsen. 2006. Extra-hepatic transcription of the human mannose-binding lectin gene (mbl2) and the MBL-associated serine protease 1–3 genes. *Mol Immunol 43*:962.
14. Uemura, K., M. Saka, T. Nakagawa, N. Kawasaki, S. Thiel, J. C. Jensenius, and T. Kawasaki. 2002. L-MBP is expressed in epithelial cells of mouse small intestine. *J Immunol 169*:6945.
15. Holmskov, U., S. Thiel, and J. C. Jensenius. 2003. Collectins and ficolins: Humoral lectins of the innate immune defense. *Annu Rev Immunol 21*:547.
16. Hakansson, K. and K. B. Reid. 2000. Collectin structure: A review. *Protein Sci 9*:1607.
17. Ikeda, K., T. Sannoh, N. Kawasaki, T. Kawasaki, and I. Yamashina. 1987. Serum lectin with known structure activates complement through the classical pathway. *J Biol Chem 262*:7451.
18. Thiel, S., T. Vorup-Jensen, C. M. Stover, W. Schwaeble, S. B. Laursen, K. Poulsen, A. C. Willis, P. Eggleton, S. Hansen, U. Holmskov, K. B. Reid, and J. C. Jensenius. 1997. A second serine protease associated with mannan-binding lectin that activates complement. *Nature 386*:506.

19. Matsushita, M., Y. Endo, N. Hamasaki, and T. Fujita. 2001. Activation of the lectin complement pathway by ficolins. *Int Immunopharmacol 1*:359.
20. Frederiksen, P. D., S. Thiel, C. B. Larsen, and J. C. Jensenius. 2005. M-ficolin, an innate immune defence molecule, binds patterns of acetyl groups and activates complement. *Scand J Immunol 62*:462.
21. Holmskov, U., R. Malhotra, R. B. Sim, and J. C. Jensenius. 1994. Collectins: Collagenous C-type lectins of the innate immune defense system. *Immunol Today 15*:67.
22. Super, M., S. Thiel, J. Lu, R. J. Levinsky, and M. W. Turner. 1989. Association of low levels of mannan-binding protein with a common defect of opsonisation. *Lancet 2*:1236.
23. Steffensen, R., S. Thiel, K. Varming, C. Jersild, and J. C. Jensenius. 2000. Detection of structural gene mutations and promoter polymorphisms in the mannan-binding lectin (MBL) gene by polymerase chain reaction with sequence-specific primers. *J Immunol Methods 241*:33.
24. Sumiya, M., M. Super, P. Tabona, R. J. Levinsky, T. Arai, M. W. Turner, and J. A. Summerfield. 1991. Molecular basis of opsonic defect in immunodeficient children. *Lancet 337*:1569.
25. Lipscombe, R. J., Y. L. Lau, R. J. Levinsky, M. Sumiya, J. A. Summerfield, and M. W. Turner. 1992. Identical point mutation leading to low levels of mannose binding protein and poor C3b mediated opsonisation in Chinese and Caucasian populations. *Immunol Lett 32*:253.
26. Lipscombe, R. J., M. Sumiya, A. V. Hill, Y. L. Lau, R. J. Levinsky, J. A. Summerfield, and M. W. Turner. 1992. High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. *Hum Mol Genet 1*:709.
27. Madsen, H. O., P. Garred, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, S. Thiel, and A. Svejgaard. 1994. A new frequent allele is the missing link in the structural polymorphism of the human mannan-binding protein. *Immunogenetics 40*:37.
28. Wallis, R. and J. Y. Cheng. 1999. Molecular defects in variant forms of mannose-binding protein associated with immunodeficiency. *J Immunol 163*:4953.
29. Garred, P., F. Larsen, H. O. Madsen, and C. Koch. 2003. Mannose-binding lectin deficiency revisited. *Mol Immunol 40*:73.
30. Madsen, H. O., P. Garred, S. Thiel, J. A. Kurtzhals, L. U. Lamm, L. P. Ryder, and A. Svejgaard. 1995. Interplay between promoter and structural gene variants control basal serum level of mannan-binding protein. *J Immunol 155*:3013.
31. Madsen, H. O., M. L. Satz, B. Hogh, A. Svejgaard, and P. Garred. 1998. Different molecular events result in low protein levels of mannan-binding lectin in populations from southeast Africa and South America. *J Immunol 161*:3169.
32. Sullivan, K. E., C. Wooten, D. Goldman, and M. Petri. 1996. Mannose-binding protein genetic polymorphisms in black patients with systemic lupus erythematosus. *Arthritis Rheum 39*:2046.
33. Skalnikova, H., T. Freiberger, J. Chumchalova, H. Grombirikova, and A. Sediva. 2004. Cost-effective genotyping of human MBL2 gene mutations using multiplex PCR. *J Immunol Methods 295*:139.
34. Wiertsema, S. P., B. L. Herpers, R. H. Veenhoven, M. M. Salimans, H. J. Ruven, E. A. Sanders, and G. T. Rijkers. 2006. Functional polymorphisms in the mannan-binding lectin 2 gene: effect on MBL levels and otitis media. *J Allergy Clin Immunol 117*:1344.
35. Peterslund, N. A., C. Koch, J. C. Jensenius, and S. Thiel. 2001. Association between deficiency of mannose-binding lectin and severe infections after chemotherapy. *Lancet 358*:637.
36. Neth, O., I. Hann, M. W. Turner, and N. J. Klein. 2001. Deficiency of mannose-binding lectin and burden of infection in children with malignancy: A prospective study. *Lancet 358*:614.
37. Turner, M. W. and R. M. Hamvas. 2000. Mannose-binding lectin: Structure, function, genetics and disease associations. *Rev Immunogenet 2*:305.
38. Petersen, S. V., S. Thiel, and J. C. Jensenius. 2001. The mannan-binding lectin pathway of complement activation: Biology and disease association. *Mol Immunol 38*:133.
39. Eisen, D. P. and R. M. Minchinton. 2003. Impact of mannose-binding lectin on susceptibility to infectious diseases. *Clin Infect Dis 37*:1496.
40. Nuytinck, L. and F. Shapiro. 2004. Mannose-binding lectin: Laying the stepping stone from clinical research to personalized medicine. *Future Medicine 1*:35.
41. Fiane, A. E., V. Videm, P. S. Lingaas, L. Heggelund, E. W. Nielsen, O. R. Geiran, M. Fung, and T. E. Mollnes. 2003. Mechanism of complement activation and its role in the inflammatory response after thoracoabdominal aortic aneurysm repair. *Circulation 108*:849.
42. Moller-Kristensen, M., W. Wang, M. Ruseva, S. Thiel, S. Nielsen, K. Takahashi, L. Shi, A. Ezekowitz, J. C. Jensenius, and M. Gadjeva. 2005. Mannan-binding lectin recognizes structures on ischaemic reperfused mouse kidneys and is implicated in tissue injury. *Scand J Immunol 61*:426.

43. Walsh, M. C., T. Bourcier, K. Takahashi, L. Shi, M. N. Busche, R. P. Rother, S. D. Solomon, R. A. Ezekowitz, and G. L. Stahl. 2005. Mannose-binding lectin is a regulator of inflammation that accompanies myocardial ischemia and reperfusion injury. *J Immunol 175*:541.
44. Bone, R. C., R. A. Balk, F. B. Cerra, R. P. Dellinger, A. M. Fein, W. A. Knaus, R. M. Schein, and W. J. Sibbald. 1992. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. *Chest 101*:1644.
45. Dahmer, M. K., A. Randolph, S. Vitali, and M. W. Quasney. 2005. Genetic polymorphisms in sepsis. *Pediatr Crit Care Med 6*:S61.
46. Fidler, K. J., P. Wilson, J. C. Davies, M. W. Turner, M. J. Peters, and N. J. Klein. 2004. Increased incidence and severity of the systemic inflammatory response syndrome in patients deficient in mannose-binding lectin. *Intensive Care Med 30*:1438.
47. Garred, P., J. S. J, L. Quist, E. Taaning, and H. O. Madsen. 2003. Association of mannose-binding lectin polymorphisms with sepsis and fatal outcome, in patients with systemic inflammatory response syndrome. *J Infect Dis 188*:1394.
48. Sutherland, A. M., K. R. Walley, and J. A. Russell. 2005. Polymorphisms in CD14, mannose-binding lectin, and Toll-like receptor-2 are associated with increased prevalence of infection in critically ill adults. *Crit Care Med 33*:638.
49. Gordon, A. C., U. Waheed, T. K. Hansen, G. A. Hitman, C. S. Garrard, M. W. Turner, N. J. Klein, S. J. Brett, and C. J. Hinds. 2006. Mannose-binding lectin polymorphisms in severe sepsis: Relationship to levels, incidence, and outcome. *Shock 25*:88.
50. Gong, M. N., W. Zhou, P. L. Williams, B. T. Thompson, L. Pothier, and D. C. Christiani. 2006. Polymorphisms in the mannose binding lectin-2 gene and acute respiratory distress syndrome. *Crit Care Med 35*:48.
51. Ahrens, P., E. Kattner, B. Kohler, C. Hartel, J. Seidenberg, H. Segerer, J. Moller, and W. Gopel. 2004. Mutations of genes involved in the innate immune system as predictors of sepsis in very low birth weight infants. *Pediatr Res 55*:652.
52. Eisen, D. P., M. M. Dean, P. Thomas, P. Marshall, N. Gerns, S. Heatley, J. Quinn, R. M. Minchinton, and J. Lipman. 2006. Low mannose-binding lectin function is associated with sepsis in adult patients. *FEMS Immunol Med Microbiol 48*:274.
53. Figueroa, J. E. and P. Densen. 1991. Infectious diseases associated with complement deficiencies. *Clin Microbiol Rev 4*:359.
54. Jack, D. L., A. W. Dodds, N. Anwar, C. A. Ison, A. Law, M. Frosch, M. W. Turner, and N. J. Klein. 1998. Activation of complement by mannose-binding lectin on isogenic mutants of *Neisseria meningitidis* serogroup B. *J Immunol 160*:1346.
55. Jack, D. L., R. C. Read, A. J. Tenner, M. Frosch, M. W. Turner, and N. J. Klein. 2001. Mannose-binding lectin regulates the inflammatory response of human professional phagocytes to *Neisseria meningitidis* serogroup B. *J Infect Dis 184*:1152.
56. Estabrook, M. M., D. L. Jack, N. J. Klein, and G. A. Jarvis. 2004. Mannose-binding lectin binds to two major outer membrane proteins, opacity protein and porin, of *Neisseria meningitidis*. *J Immunol 172*:3784.
57. Drogari-Apiranthitou, M., C. A. Fijen, S. Thiel, A. Platonov, L. Jensen, J. Dankert, and E. J. Kuijper. 1997. The effect of mannan-binding lectin on opsonophagocytosis of *Neisseria meningitidis*. *Immunopharmacology 38*:93.
58. Kuipers, S., P. C. Aerts, and H. van Dijk. 2003. Differential microorganism-induced mannose-binding lectin activation. *FEMS Immunol Med Microbiol 36*:33.
59. Jack, D. L., M. E. Lee, M. W. Turner, N. J. Klein, and R. C. Read. 2005. Mannose-binding lectin enhances phagocytosis and killing of *Neisseria meningitidis* by human macrophages. *J Leukoc Biol 77*:328.
60. Hibberd, M. L., M. Sumiya, J. A. Summerfield, R. Booy, and M. Levin. 1999. Association of variants of the gene for mannose-binding lectin with susceptibility to meningococcal disease. Meningococcal Research Group. *Lancet 353*:1049.
61. Bax, W. A., O. J. Cluysenaer, A. K. Bartelink, P. C. Aerts, R. A. Ezekowitz, and H. van Dijk. 1999. Association of familial deficiency of mannose-binding lectin and meningococcal disease. *Lancet 354*:1094.
62. Salimans, M. M., W. A. Bax, F. Stegeman, M. van Deuren, A. K. Bartelink, and H. van Dijk. 2004. Association between familial deficiency of mannose-binding lectin and mutations in the corresponding gene and promoter region. *Clin Diagn Lab Immunol 11*:806.

63. Bathum, L., H. Hansen, B. Teisner, C. Koch, P. Garred, K. Rasmussen, and P. Wang. 2006. Association between combined properdin and mannose-binding lectin deficiency and infection with *Neisseria meningitidis. Mol Immunol 43*:473.
64. Garred, P., T. E. Michaelsen, G. Bjune, S. Thiel, and A. Svejgaard. 1993. A low serum concentration of mannan-binding protein is not associated with serogroup B or C meningococcal disease. *Scand J Immunol 37*:468.
65. Garred, P., H. O. Madsen, A. Svejgaard, and T. E. Michaelsen. 1999. Mannose-binding lectin and meningococcal disease. *Lancet 354*:336.
66. Krarup, A., U. B. Sorensen, M. Matsushita, J. C. Jensenius, and S. Thiel. 2005. Effect of capsulation of opportunistic pathogenic bacteria on binding of the pattern recognition molecules mannan-binding lectin, L-ficolin, and H-ficolin. *Infect Immun 73*:1052.
67. Roy, S., K. Knox, S. Segal, D. Griffiths, C. E. Moore, K. I. Welsh, A. Smarason, N. P. Day, W. L. McPheat, D. W. Crook, and A. V. Hill. 2002. MBL genotype and risk of invasive pneumococcal disease: A case–control study. *Lancet 359*:1569.
68. Kronborg, G., N. Weis, H. O. Madsen, S. S. Pedersen, C. Wejse, H. Nielsen, P. Skinhoj, and P. Garred. 2002. Variant mannose-binding lectin alleles are not associated with susceptibility to or outcome of invasive pneumococcal infection in randomly included patients. *J Infect Dis 185*:1517.
69. Kronborg, G. and P. Garred. 2002. Mannose-binding lectin genotype as a risk factor for invasive pneumococcal infection. *Lancet 360*:1176.
70. Garred, P., H. O. Madsen, U. Balslev, B. Hofmann, C. Pedersen, J. Gerstoft, and A. Svejgaard. 1997. Susceptibility to HIV infection and progression of AIDS in relation to variant alleles of mannose-binding lectin. *Lancet 349*:236.
71. Moens, L., E. Van Hoeyveld, W. E. Peetermans, C. De Boeck, J. Verhaegen, and X. Bossuyt. 2006. Mannose-binding lectin genotype and invasive pneumococcal infection. *Hum Immunol 67*:605.
72. Rodriguez, J. A., G. D. Salgado, and A. Saiz. 2004. MBL deficiency does not associate with recurrent pneumococcal infections in a group of Colombian children. *J Trop Pediatr 50*:379.
73. Tezcan, I., Y. Yilmaz, F. Oner, L. Yel, O. Sanal, F. Ersoy, M. Onerci, and A. I. Berkel. 1997. Defective serum opsonization activity in children aged 6–48 months having acute purulent otitis media. *Turk J Pediatr 39*:453.
74. Richardson, V. F., V. F. Larcher, and J. F. Price. 1983. A common congenital immunodeficiency predisposing to infection and atopy in infancy. *Arch Dis Child 58*:799.
75. Nuytinck, L., E. De Meester, M. Van Thielen, and P. Govaerts. 2006. Role of mannose-binding lectin (MBL2) genotyping in predicting the risk of recurrent otitis media (rOM). *Adv Exp Med Biol 586*:281.
76. Kielgast, S., S. Thiel, T. B. Henriksen, T. Bjerke, J. Olsen, and J. C. Jensenius. 2003. Umbilical cord mannan-binding lectin and infections in early childhood. *Scand J Immunol 57*:167.
77. Garred, P., K. Brygge, C. H. Sorensen, H. O. Madsen, S. Thiel, and A. Svejgaard. 1993. Mannan-binding protein levels in plasma and upper-airways secretions and frequency of genotypes in children with recurrence of otitis media. *Clin Exp Immunol 94*:99.
78. Garred, P., H. O. Madsen, B. Hofmann, and A. Svejgaard. 1995. Increased frequency of homozygosity of abnormal mannan-binding-protein alleles in patients with suspected immunodeficiency. *Lancet 346*:941.
79. Thorarinsdottir, H. K., B. R. Ludviksson, T. Vikingsdottir, M. O. Leopoldsdottir, B. Ardal, T. Jonsson, H. Valdimarsson, and G. J. Arason. 2005. Childhood levels of immunoglobulins and mannan-binding lectin in relation to infections and allergy. *Scand J Immunol 61*:466.
80. Koch, A., M. Melbye, P. Sorensen, P. Homoe, H. O. Madsen, K. Molbak, C. H. Hansen, L. H. Andersen, G. W. Hahn, and P. Garred. 2001. Acute respiratory tract infections and mannose-binding lectin insufficiency during early childhood. *JAMA 285*:1316.
81. Homoe, P., H. O. Madsen, K. Sandvej, A. Koch, and P. Garred. 1999. Lack of association between mannose-binding lectin, acute otitis media and early Epstein–Barr virus infection among children in Greenland. *Scand J Infect Dis 31*:363.
82. Straetemans, M., S. P. Wiertsema, E. A. Sanders, G. T. Rijkers, K. Graamans, B. van der Baan, and G. A. Zielhuis. 2005. Immunological status in the aetiology of recurrent otitis media with effusion: Serum immunoglobulin levels, functional mannose-binding lectin and Fc receptor polymorphisms for IgG. *J Clin Immunol 25*:78.
83. Garred, P., M. Harboe, T. Oettinger, C. Koch, and A. Svejgaard. 1994. Dual role of mannan-binding protein in infections: Another case of heterosis? *Eur J Immunogenet 21*:125.

84. Schlesinger, L. S. 1993. Macrophage phagocytosis of virulent but not attenuated strains of *Mycobacterium tuberculosis* is mediated by mannose receptors in addition to complement receptors. *J Immunol 150*:2920.
85. Schlesinger, L. S., T. M. Kaufman, S. Iyer, S. R. Hull, and L. K. Marchiando. 1996. Differences in mannose receptor-mediated uptake of lipoarabinomannan from virulent and attenuated strains of *Mycobacterium tuberculosis* by human macrophages. *J Immunol 157*:4568.
86. Hoppe, H. C., B. J. de Wet, C. Cywes, M. Daffe, and M. R. Ehlers. 1997. Identification of phosphatidylinositol mannoside as a mycobacterial adhesin mediating both direct and opsonic binding to nonphagocytic mammalian cells. *Infect Immun 65*:3896.
87. Garred, P., C. Richter, A. B. Andersen, H. O. Madsen, I. Mtoni, A. Svejgaard, and J. Shao. 1997. Mannan-binding lectin in the sub-Saharan HIV and tuberculosis epidemics. *Scand J Immunol 46*:204.
88. Soborg, C., H. O. Madsen, A. B. Andersen, T. Lillebaek, A. Kok-Jensen, and P. Garred. 2003. Mannose-binding lectin polymorphisms in clinical tuberculosis. *J Infect Dis 188*:777.
89. Selvaraj, P., M. S. Jawahar, D. N. Rajeswari, K. Alagarasu, M. Vidyarani, and P. R. Narayanan. 2006. Role of mannose binding lectin gene variants on its protein levels and macrophage phagocytosis with live *Mycobacterium tuberculosis* in pulmonary tuberculosis. *FEMS Immunol Med Microbiol 46*:433.
90. Mombo, L. E., C. Y. Lu, S. Ossari, I. Bedjabaga, L. Sica, R. Krishnamoorthy, and C. Lapoumeroulie. 2003. Mannose-binding lectin alleles in sub-Saharan Africans and relation with susceptibility to infections. *Genes Immun 4*:362.
91. Hoal-Van Helden, E. G., J. Epstein, T. C. Victor, D. Hon, L. A. Lewis, N. Beyers, D. Zurakowski, A. B. Ezekowitz, and P. D. Van Helden. 1999. Mannose-binding protein B allele confers protection against tuberculous meningitis. *Pediatr Res 45*:459.
92. Bellamy, R., C. Ruwende, K. P. McAdam, M. Thursz, M. Sumiya, J. Summerfield, S. C. Gilbert, T. Corrah, D. Kwiatkowski, H. C. Whittle, and A. V. Hill. 1998. Mannose binding protein deficiency is not associated with malaria, hepatitis B carriage nor tuberculosis in Africans. *QJM 91*:13.
93. Garcia-Laorden, M. I., M. J. Pena, J. A. Caminero, A. Garcia-Saavedra, M. I. Campos-Herrero, A. Caballero, and C. Rodriguez-Gallego. 2006. Influence of mannose-binding lectin on HIV infection and tuberculosis in a Western European population. *Mol Immunol 43*:2143.
94. Soborg, C., A. B. Andersen, N. Range, W. Malenganisho, H. Friis, P. Magnussen, M. M. Temu, J. Changalucha, H. O. Madsen, and P. Garred. 2007. Influence of candidate susceptibility genes on tuberculosis in a high endemic region. *Mol Immunol 44*:2213.
95. El Sahly, H. M., R. A. Reich, S. J. Dou, J. M. Musser, and E. A. Graviss. 2004. The effect of mannose binding lectin gene polymorphisms on susceptibility to tuberculosis in different ethnic groups. *Scand J Infect Dis 36*:106.
96. Selvaraj, P., P. R. Narayanan, and A. M. Reetha. 1999. Association of functional mutant homozygotes of the mannose binding protein gene with susceptibility to pulmonary tuberculosis in India. *Tuber Lung Dis 79*:221.
97. Turner, M. W., L. Dinan, S. Heatley, D. L. Jack, B. Boettcher, S. Lester, J. McCluskey, and D. Roberton. 2000. Restricted polymorphism of the mannose-binding lectin gene of indigenous Australians. *Hum Mol Genet 9*:1481.
98. Thomas, H. C., G. R. Foster, M. Sumiya, D. McIntosh, D. L. Jack, M. W. Turner, and J. A. Summerfield. 1996. Mutation of gene of mannose-binding protein associated with chronic hepatitis B viral infection. *Lancet 348*:1417.
99. Hohler, T., M. Wunschel, G. Gerken, P. M. Schneider, K. H. Meyer zum Buschenfelde, and C. Rittner. 1998. No association between mannose-binding lectin alleles and susceptibility to chronic hepatitis B virus infection in German patients. *Exp Clin Immunogenet 15*:130.
100. Song le, H., V. Q. Binh, D. N. Duy, S. Juliger, T. C. Bock, A. J. Luty, P. G. Kremsner, and J. F. Kun. 2003. Mannose-binding lectin gene polymorphisms and hepatitis B virus infection in Vietnamese patients. *Mutat Res 522*:119.
101. Cheong, J. Y., S. W. Cho, S. K. Lim, H. Shin do, S. K. Yoon, J. E. Lee, K. B. Hahm, and J. H. Kim. 2005. Lack of association between hepatitis B virus infection and polymorphism of mannose-binding lectin gene in Korean population. *J Korean Med Sci 20*:65.
102. Yuen, M. F., C. S. Lau, Y. L. Lau, W. M. Wong, C. C. Cheng, and C. L. Lai. 1999. Mannose binding lectin gene mutations are associated with progression of liver disease in chronic hepatitis B infection. *Hepatology 29*:1248.
103. Chong, W. P., Y. F. To, W. K. Ip, M. F. Yuen, T. P. Poon, W. H. Wong, C. L. Lai, and Y. L. Lau. 2005. Mannose-binding lectin in chronic hepatitis B virus infection. *Hepatology 42*:1037.

104. Thio, C. L., T. Mosbruger, J. Astemborski, S. Greer, G. D. Kirk, S. J. O'Brien, and D. L. Thomas. 2005. Mannose binding lectin genotypes influence recovery from hepatitis B virus infection. *J Virol 79*:9192.
105. Hakozaki, Y., M. Yoshiba, K. Sekiyama, E. Seike, J. Iwamoto, K. Mitani, M. Mine, T. Morizane, K. Ohtani, Y. Suzuki, and N. Wakamiya. 2002. Mannose-binding lectin and the prognosis of fulminant hepatic failure caused by HBV infection. *Liver 22*:29.
106. Matsushita, M., M. Hijikata, Y. Ohta, K. Iwata, M. Matsumoto, K. Nakao, K. Kanai, N. Yoshida, K. Baba, and S. Mishiro. 1998. Hepatitis C virus infection and mutations of mannose-binding lectin gene MBL. *Arch Virol 143*:645.
107. Sasaki, K., A. Tsutsumi, N. Wakamiya, K. Ohtani, Y. Suzuki, Y. Watanabe, N. Nakayama, and T. Koike. 2000. Mannose-binding lectin polymorphisms in patients with hepatitis C virus infection. *Scand J Gastroenterol 35*:960.
108. Kilpatrick, D. C., T. E. Delahooke, C. Koch, M. L. Turner, and P. C. Hayes. 2003. Mannan-binding lectin and hepatitis C infection. *Clin Exp Immunol 132*:92.
109. Ohsawa, I., H. Ohi, M. Tamano, M. Endo, T. Fujita, A. Satomura, M. Hidaka, Y. Fuke, and M. Matsushita. 2001. Cryoprecipitate of patients with cryoglobulinemic glomerulonephritis contains molecules of the lectin complement pathway. *Clin Immunol 101*:59.
110. Hansen, T. K., S. Thiel, R. Dall, A. M. Rosenfalck, P. Trainer, A. Flyvbjerg, J. O. Jorgensen, and J. S. Christiansen. 2001. GH strongly affects serum concentrations of mannan-binding lectin: Evidence for a new IGF-I independent immunomodulatory effect of GH. *J Clin Endocrinol Metab 86*:5383.
111. Riis, A. L., T. K. Hansen, S. Thiel, C. H. Gravholt, S. Gjedde, L. C. Gormsen, J. O. Jorgensen, J. Weeke, and N. Moller. 2005. Thyroid hormone increases mannan-binding lectin levels. *Eur J Endocrinol 153*:643.
112. Dumestre-Perard, C., D. Ponard, C. Drouet, V. Leroy, J. P. Zarski, N. Dutertre, and M. G. Colomb. 2002. Complement C4 monitoring in the follow-up of chronic hepatitis C treatment. *Clin Exp Immunol 127*:131.
113. Ezekowitz, R. A., M. Kuhlman, J. E. Groopman, and R. A. Byrn. 1989. A human serum mannose-binding protein inhibits in vitro infection by the human immunodeficiency virus. *J Exp Med 169*:185.
114. Larkin, M., R. A. Childs, T. J. Matthews, S. Thiel, T. Mizuochi, A. M. Lawson, J. S. Savill, C. Haslett, R. Diaz, and T. Feizi. 1989. Oligosaccharide-mediated interactions of the envelope glycoprotein gp120 of HIV-1 that are independent of CD4 recognition. *AIDS 3*:793.
115. Haurum, J. S., S. Thiel, I. M. Jones, P. B. Fischer, S. B. Laursen, and J. C. Jensenius. 1993. Complement activation upon binding of mannan-binding protein to HIV envelope glycoproteins. *AIDS 7*:1307.
116. Hart, M. L., M. Saifuddin, K. Uemura, E. G. Bremer, B. Hooker, T. Kawasaki, and G. T. Spear. 2002. High mannose glycans and sialic acid on gp120 regulate binding of mannose-binding lectin (MBL) to HIV type 1. *AIDS Res Hum Retroviruses 18*:1311.
117. Hart, M. L., M. Saifuddin, and G. T. Spear. 2003. Glycosylation inhibitors and neuraminidase enhance human immunodeficiency virus type 1 binding and neutralization by mannose-binding lectin. *J Gen Virol 84*:353.
118. Spear, G. T., M. R. Zariffard, J. Xin, and M. Saifuddin. 2003. Inhibition of DC-SIGN-mediated *trans* infection of T cells by mannose-binding lectin. *Immunology 110*:80.
119. Pastinen, T., K. Liitsola, P. Niini, M. Salminen, and A. C. Syvanen. 1998. Contribution of the CCR5 and MBL genes to susceptibility to HIV type 1 infection in the Finnish population. *AIDS Res Hum Retroviruses 14*:695.
120. Vallinoto, A. C., M. R. Menezes-Costa, A. E. Alves, L. F. Machado, V. N. Azevedo, L. L. Souza, M. D. Ishak, and R. Ishak. 2006. Mannose-binding lectin gene polymorphism and its impact on human immunodeficiency virus 1 infection. *Mol Immunol 43*:1358.
121. Malik, S., M. Arias, C. Di Flumeri, L. F. Garcia, and E. Schurr. 2003. Absence of association between mannose-binding lectin gene polymorphisms and HIV-1 infection in a Colombian population. *Immunogenetics 55*:49.
122. Crovella, S., M. Bernardon, L. Braida, M. Boniotto, S. Guaschino, E. Ferrazzi, P. Martinelli, and S. Alberico. 2005. Italian multicentric pilot study on MBL2 genetic polymorphisms in HIV positive pregnant women and their children. *J Matern Fetal Neonatal Med. 17*:253.
123. Amoroso, A., M. Berrino, M. Boniotto, S. Crovella, E. Palomba, G. Scarlatti, C. Serra, P. A. Tovo, and S. Vatta. 1999. Polymorphism at codon 54 of mannose-binding protein gene influences AIDS progression but not HIV infection in exposed children. *AIDS 13*:863.

124. Nielsen, S. L., P. L. Andersen, C. Koch, J. C. Jensenius, and S. Thiel. 1995. The level of the serum opsonin, mannan-binding protein in HIV-1 antibody-positive patients. *Clin Exp Immunol 100*:219.
125. Boniotto, M., L. Braida, D. Pirulli, L. Arraes, A. Amoroso, and S. Crovella. 2003. MBL2 polymorphisms are involved in HIV-1 infection in Brazilian perinatally infected children. *AIDS 17*:779.
126. Boniotto, M., S. Crovella, D. Pirulli, G. Scarlatti, A. Spano, L. Vatta, S. Zezlina, P. A. Tovo, E. Palomba, and A. Amoroso. 2000. Polymorphisms in the MBL2 promoter correlated with risk of HIV-1 vertical transmission and AIDS progression. *Genes Immun 1*:346.
127. Dzwonek, A., V. Novelli, M. Bajaj-Elliott, M. Turner, M. Clapson, and N. Klein. 2006. Mannose-binding lectin in susceptibility and progression of HIV-1 infection in children. *Antivir Ther 11*:499.
128. Kuhn, L., A. Coutsoudis, D. Trabattoni, D. Archary, T. Rossi, L. Segat, M. Clerici, and S. Crovella. 2006. Synergy between mannose-binding lectin gene polymorphisms and supplementation with vitamin A influences susceptibility to HIV infection in infants born to HIV-positive mothers. *Am J Clin Nutr 84*:610.
129. Kelly, P., D. L. Jack, A. Naeem, B. Mandanda, R. C. Pollok, N. J. Klein, M. W. Turner, and M. J. Farthing. 2000. Mannose-binding lectin is a component of innate mucosal defense against *Cryptosporidium parvum* in AIDS. *Gastroenterology 119*:1236.
130. Maas, J., A. M. de Roda Husman, M. Brouwer, A. Krol, R. Coutinho, I. Keet, R. van Leeuwen, and H. Schuitemaker. 1998. Presence of the variant mannose-binding lectin alleles associated with slower progression to AIDS. *Amsterdam Cohort Study. AIDS 12*:2275.
131. Prohaszka, Z., S. Thiel, E. Ujhelyi, J. Szlavik, D. Banhegyi, and G. Fust. 1997. Mannan-binding lectin serum concentrations in HIV-infected patients are influenced by the stage of disease. *Immunol Lett 58*:171.
132. Lian, Y. C., M. Della-Negra, R. Rutz, V. Ferriani, D. de Moraes Vasconcelos, A. J. da Silva Duarte, M. Kirschfink, and A. S. Grumach. 2004. Immunological analysis in paediatric HIV patients at different stages of the disease. *Scand J Immunol 60*:615.
133. Heggelund, L., T. E. Mollnes, T. Ueland, B. Christophersen, P. Aukrust, and S. S. Froland. 2003. Mannose-binding lectin in HIV infection: relation to disease progression and highly active antiretroviral therapy. *J Acquir Immune Defic Syndr 32*:354.
134. Hundt, M., H. Heiken, and R. E. Schmidt. 2000. Association of low mannose-binding lectin serum concentrations and bacterial pneumonia in HIV infection. *Aids 14*:1853.
135. McBride, M. O., P. B. Fischer, M. Sumiya, M. O. McClure, M. W. Turner, C. J. Skinner, J. N. Weber, and J. A. Summerfield. 1998. Mannose-binding protein in HIV-seropositive patients does not contribute to disease progression or bacterial infections. *Int J STD AIDS 9*:683.
136. Lau, Y. L. and J. S. Peiris. 2005. Pathogenesis of severe acute respiratory syndrome. *Curr Opin Immunol 17*:404.
137. Ip, W. K., K. H. Chan, H. K. Law, G. H. Tso, E. K. Kong, W. H. Wong, Y. F. To, R. W. Yung, E. Y. Chow, K. L. Au, E. Y. Chan, W. Lim, J. C. Jensenius, M. W. Turner, J. S. Peiris, and Y. L. Lau. 2005. Mannose-binding lectin in severe acute respiratory syndrome coronavirus infection. *J Infect Dis 191*:1697.
138. Zhang, H., G. Zhou, L. Zhi, H. Yang, Y. Zhai, X. Dong, X. Zhang, X. Gao, Y. Zhu, and F. He. 2005. Association between mannose-binding lectin gene polymorphisms and susceptibility to severe acute respiratory syndrome coronavirus infection. *J Infect Dis 192*:1355.
139. Yuan, F. F., J. Tanner, P. K. Chan, S. Biffin, W. B. Dyer, A. F. Geczy, J. W. Tang, D. S. Hui, J. J. Sung, and J. S. Sullivan. 2005. Influence of FcgammaRIIA and MBL polymorphisms on severe acute respiratory syndrome. *Tissue Antigens 66*:291.
140. Ip, W. K., and Y. L. Lau. 2004. Role of mannose-binding lectin in the innate defense against Candida albicans: enhancement of complement activation, but lack of opsonic function, in phagocytosis by human dendritic cells. *J Infect Dis 190*:632.
141. Tabona, P., A. Mellor, and J. A. Summerfield. 1995. Mannose binding protein is involved in first-line host defence: Evidence from transgenic mice. *Immunology 85*:153.
142. Liu, F., Q. Liao, and Z. Liu. 2006. Mannose-binding lectin and vulvovaginal candidiasis. *Int J Gynaecol Obstet 92*:43.
143. Babula, O., G. Lazdane, J. Kroica, W. J. Ledger, and S. S. Witkin. 2003. Relation between recurrent vulvovaginal candidiasis, vaginal concentrations of mannose-binding lectin, and a mannose-binding lectin gene polymorphism in Latvian women. *Clin Infect Dis 37*:733.
144. Pellis, V., F. De Seta, S. Crovella, F. Bossi, R. Bulla, S. Guaschino, O. Radillo, P. Garred, and F. Tedesco. 2005. Mannose binding lectin and C3 act as recognition molecules for infectious agents in the vagina. *Clin Exp Immunol 139*:120.

145. Klabunde, J., A. C. Uhlemann, A. E. Tebo, J. Kimmel, R. T. Schwarz, P. G. Kremsner, and J. F. Kun. 2002. Recognition of plasmodium falciparum proteins by mannan-binding lectin, a component of the human innate immune system. *Parasitol Res 88*:113.
146. Garred, P., M. A. Nielsen, J. A. Kurtzhals, R. Malhotra, H. O. Madsen, B. Q. Goka, B. D. Akanmori, R. B. Sim, and L. Hviid. 2003. Mannose-binding lectin is a disease modifier in clinical malaria and may function as opsonin for *Plasmodium falciparum*-infected erythrocytes. *Infect Immun 71*:5245.
147. Jungery, M., G. Pasvol, C. I. Newbold, and D. J. Weatherall. 1983. A lectin-like receptor is involved in invasion of erythrocytes by *Plasmodium falciparum*. *Proc Natl Acad Sci USA 80*:1018.
148. Luty, A. J., J. F. Kun, and P. G. Kremsner. 1998. Mannose-binding lectin plasma levels and gene polymorphisms in *Plasmodium falciparum* malaria. *J Infect Dis 178*:1221.
149. Juliger, S., P. G. Kremsner, M. P. Alpers, J. C. Reeder, and J. F. Kun. 2002. Restricted polymorphisms of the mannose-binding lectin gene in a population of Papua New Guinea. *Mutat Res 505*:87.
150. Boldt, A. B., A. Luty, M. P. Grobusch, K. Dietz, A. Dzeing, M. Kombila, P. G. Kremsner, and J. F. Kun. 2006. Association of a new mannose-binding lectin variant with severe malaria in Gabonese children. *Genes Immun 7*:393.
151. Green, P. J., T. Feizi, M. S. Stoll, S. Thiel, A. Prescott, and M. J. McConville. 1994. Recognition of the major cell surface glycoconjugates of Leishmania parasites by the human serum mannan-binding protein. *Mol Biochem Parasitol 66*:319.
152. Santos, I. K., C. H. Costa, H. Krieger, M. F. Feitosa, D. Zurakowski, B. Fardin, R. B. Gomes, D. L. Weiner, D. A. Harn, R. A. Ezekowitz, and J. E. Epstein. 2001. Mannan-binding lectin enhances susceptibility to visceral leishmaniasis. *Infect Immun 69*:5212.
153. Ambrosio, A. R. and I. J. De Messias-Reason. 2005. Leishmania (Viannia) braziliensis: Interaction of mannose-binding lectin with surface glycoconjugates and complement activation. *An antibody-independent defence mechanism. Parasite Immunol 27*:333.
154. Frakking, F. N., M. D. van de Wetering, N. Brouwer, K. M. Dolman, J. Geissler, B. Lemkes, H. N. Caron, and T. W. Kuijpers. 2006. The role of mannose-binding lectin (MBL) in paediatric oncology patients with febrile neutropenia. *Eur J Cancer 42*:909.
155. Schlapbach, L. J., C. Aebi, M. Otth, A. R. Luethy, K. Leibundgut, A. Hirt, and R. A. Ammann. 2006. Serum levels of mannose-binding lectin and the risk of fever in neutropenia pediatric cancer patients. *Pediatr Blood Cancer* 49:11.
156. Lausen, B., K. Schmiegelow, B. Andreassen, H. O. Madsen, and P. Garred. 2006. Infections during induction therapy of childhood acute lymphoblastic leukemia—No association to mannose-binding lectin deficiency. *Eur J Haematol 76*:481.
157. Bergmann, O. J., M. Christiansen, I. Laursen, P. Bang, N. E. Hansen, J. Ellegaard, C. Koch, and V. Andersen. 2003. Low levels of mannose-binding lectin do not affect occurrence of severe infections or duration of fever in acute myeloid leukaemia during remission induction therapy. *Eur J Haematol 70*:91.
158. Vekemans, M., A. Georgala, C. Heymans, F. Muanza, M. Paesmans, J. Klastersky, M. Barette, N. Meuleman, F. Huet, O. J. Robinson, O. Marchetti, T. Calandra, S. Costantini, A. Ferrant, K. Petersen, M. Axelsen, and M. Aoun. 2005. Influence of mannan binding lectin serum levels on the risk of infection during chemotherapy-induced neutropenia in adult haematological cancer patients. *Clin Microbiol Infect 11*:20.
159. Kilpatrick, D. C., L. A. McLintock, E. K. Allan, M. Copland, T. Fujita, N. E. Jordanides, C. Koch, M. Matsushita, H. Shiraki, K. Stewart, M. Tsujimura, M. L. Turner, I. M. Franklin, and T. L. Holyoake. 2003. No strong relationship between mannan binding lectin or plasma ficolins and chemotherapy-related infections. *Clin Exp Immunol 134*:279.
160. Molle, I., R. Steffensen, S. Thiel, and N. A. Peterslund. 2006. Chemotherapy-related infections in patients with multiple myeloma: associations with mannan-binding lectin genotypes. *Eur J Haematol 77*:19.
161. Molle, I., N. A. Peterslund, S. Thiel, and R. Steffensen. 2006. MBL2 polymorphism and risk of severe infections in multiple myeloma patients receiving high-dose melphalan and autologous stem cell transplantation. *Bone Marrow Trans 38*:555.
162. Mullighan, C. G., S. Heatley, K. Doherty, F. Szabo, A. Grigg, T. P. Hughes, A. P. Schwarer, J. Szer, B. D. Tait, L. Bik To, and P. G. Bardy. 2002. Mannose-binding lectin gene polymorphisms are associated with major infection following allogeneic hemopoietic stem cell transplantation. *Blood 99*:3524.
163. Mullighan, C. G. and P. G. Bardy. 2004. Mannose-binding lectin after infection following allogeneic hemopoietic stem cell transplantation. *Leukemia Lymphoma 45*:247.

164. Granell, M., A. Urbano-Ispizua, B. Suarez, M. Rovira, F. Fernandez-Aviles, C. Martinez, M. Ortega, C. Uriburu, A. Gaya, J. M. Roncero, A. Navarro, E. Carreras, J. Mensa, J. Vives, C. Rozman, E. Montserrat, and F. Lozano. 2006. Mannan-binding lectin pathway deficiencies and invasive fungal infections following allogeneic stem cell transplantation. *Exp Hematol 34*:1435.
165. Downing, I., C. Koch, and D. C. Kilpatrick. 2003. Immature dendritic cells possess a sugar-sensitive receptor for human mannan-binding lectin. *Immunology 109*:360.
166. Kilpatrick, D. C., K. Stewart, E. K. Allan, L. A. McLintock, T. L. Holyoake, and M. L. Turner. 2005. Successful haemopoietic stem cell transplantation does not correct mannan-binding lectin deficiency. *Bone Marrow Trans 35*:179.
167. Rocha, V., R. F. Franco, R. Porcher, H. Bittencourt, W. A. Silva, A. Latouche, A. Devergie, H. Esperou, P. Ribaud, G. Socie, M. A. Zago, and E. Gluckman. 2002. Host defense and inflammatory gene polymorphisms are associated with outcomes after HLA-identical sibling bone marrow transplantation. *Blood 100*:3908.
168. Horiuchi, T., H. Gondo, H. Miyagawa, J. Otsuka, S. Inaba, K. Nagafuji, K. Takase, H. Tsukamoto, T. Koyama, H. Mitoma, Y. Tamimoto, Y. Miyagi, T. Tahira, K. Hayashi, C. Hashimura, S. Okamura, and M. Harada. 2005. Association of MBL gene polymorphisms with major bacterial infection in patients treated with high-dose chemotherapy and autologous PBSCT. *Genes Immun 6*:162.
169. Bouwman, L. H., A. Roos, O. T. Terpstra, P. de Knijff, B. van Hoek, H. W. Verspaget, S. P. Berger, M. R. Daha, M. Frolich, A. R. van der Slik, Doxiadis, II, B. O. Roep, and A. F. Schaapherder. 2005. Mannose binding lectin gene polymorphisms confer a major risk for severe infections after liver transplantation. *Gastroenterology 129*:408.
170. Shi, L., K. Takahashi, J. Dundee, S. Shahroor-Karni, S. Thiel, J. C. Jensenius, F. Gad, M. R. Hamblin, K. N. Sastry, and R. A. Ezekowitz. 2004. Mannose-binding lectin-deficient mice are susceptible to infection with *Staphylococcus aureus*. *J Exp Med 199*:1379.
171. Klein, N. J. and D. C. Kilpatrick. 2004. Is there a role for mannan/mannose-binding lectin (MBL) in defence against infection following chemotherapy for cancer? *Clin Exp Immunol 138*:202.
172. Siassi, M., W. Hohenberger, and J. Riese. 2003. Mannan-binding lectin (MBL) serum levels and postoperative infections. *Biochem Soc Trans 31*:774.
173. Siassi, M., J. Riese, R. Steffensen, M. Meisner, S. Thiel, W. Hohenberger, and J. Schmidt. 2005. Mannan-binding lectin and procalcitonin measurement for prediction of postoperative infection. *Crit Care 9*:R483.
174. Ytting, H., I. J. Christensen, J. C. Jensenius, S. Thiel, and H. J. Nielsen. 2005. Preoperative mannan-binding lectin pathway and prognosis in colorectal cancer. *Cancer Immunol Immunother 54*:265.
175. Moller-Kristensen, M., W. K. Ip, L. Shi, L. D. Gowda, M. R. Hamblin, S. Thiel, J. C. Jensenius, R. A. Ezekowitz, and K. Takahashi. 2006. Deficiency of mannose-binding lectin greatly increases susceptibility to postburn infection with *Pseudomonas aeruginosa*. *J Immunol 176*:1769.
176. Miller, M. E., J. Seals, R. Kaye, and L. C. Levitsky. 1968. A familial, plasma-associated defect of phagocytosis. *Lancet* 292:60.
177. Scott, H., E. J. Moynahan, R. A. Risdon, B. A. Harvey, and J. F. Soothill. 1975. Familial opsonization defect associated with fatal infantile dermatitis, infections, and histiocytosis. *Arch Dis Child 50*:311.
178. Candy, D. C., V. F. Larcher, J. H. Tripp, J. T. Harries, B. A. Harvey, and J. F. Soothill. 1980. Yeast opsonisation in children with chronic diarrhoeal states. *Arch Dis Child 55*:189.
179. Valdimarsson, H. 2003. Infusion of plasma-derived mannan-binding lectin (MBL) into MBL-deficient humans. *Biochem Soc Trans 31*:768.
180. Valdimarsson, H., M. Stefansson, T. Vikingsdottir, G. J. Arason, C. Koch, S. Thiel, and J. C. Jensenius. 1998. Reconstitution of opsonizing activity by infusion of mannan-binding lectin (MBL) to MBL-deficient humans. *Scand J Immunol 48*:116.
181. Valdimarsson, H., T. Vikingsdottir, P. Bang, S. Saevarsdottir, J. E. Gudjonsson, O. Oskarsson, M. Christiansen, L. Blou, I. Laursen, and C. Koch. 2004. Human plasma-derived mannose-binding lectin: A phase I safety and pharmacokinetic study. *Scand J Immunol 59*:97.
182. Vorup-Jensen, T., E. S. Sorensen, U. B. Jensen, W. Schwaeble, T. Kawasaki, Y. Ma, K. Uemura, N. Wakamiya, Y. Suzuki, T. G. Jensen, K. Takahashi, R. A. Ezekowitz, S. Thiel, and J. C. Jensenius. 2001. Recombinant expression of human mannan-binding lectin. *Int Immunopharmacol 1*:677.
183. Zerefos, P. G., P. C. Ioannou, J. Traeger-Synodinos, G. Dimissianos, E. Kanavakis, and T. K. Christopoulos. 2006. Photoprotein aequorin as a novel reporter for SNP genotyping by primer extension-application to the variants of mannose-binding lectin gene. *Hum Mutat 27*:279.
184. Jensenius, J. C., P. H. Jensen, K. McGuire, J. L. Larsen, and S. Thiel. 2003. Recombinant mannan-binding lectin (MBL) for therapy. *Biochem Soc Trans 31*:763.

185. Jensen, P. H., D. Weilguny, F. Matthiesen, K. A. McGuire, L. Shi, and P. Hojrup. 2005. Characterization of the oligomer structure of recombinant human mannan-binding lectin. *J Biol Chem 280*:11043.
186. Petersen, K. A., F. Matthiesen, T. Agger, L. Kongerslev, S. Thiel, K. Cornelissen, and M. Axelsen. 2006. Phase I safety, tolerability, and pharmacokinetic study of recombinant human mannan-binding lectin. *J Clin Immunol 26*:465.
187. Gadjeva, M., S. R. Paludan, S. Thiel, V. Slavov, M. Ruseva, K. Eriksson, G. B. Lowhagen, L. Shi, K. Takahashi, A. Ezekowitz, and J. C. Jensenius. 2004. Mannan-binding lectin modulates the response to HSV-2 infection. *Clin Exp Immunol 138*:304.
188. Frederiksen, P. D., S. Thiel, L. Jensen, A. G. Hansen, F. Matthiesen, and J. C. Jensenius. 2006. Quantification of mannan-binding lectin. *J Immunol Methods 315*:49.
189. Sjoholm, A. G., G. Jonsson, J. H. Braconier, G. Sturfelt, and L. Truedsson. 2006. Complement deficiency and disease: An update. *Mol Immunol 43*:78.
190. Dahl, M., A. Tybjaerg-Hansen, P. Schnohr, and B. G. Nordestgaard. 2004. A population-based study of morbidity and mortality in mannose-binding lectin deficiency. *J Exp Med 199*:1391.
191. Gadjeva, M., K. Takahashi, and S. Thiel. 2004. Mannan-binding lectin—A soluble pattern recognition molecule. *Mol Immunol 41*:113.

23 Immunoregulatory Roles of Lung Surfactant Proteins A and D

Nades Palaniyar, Grith L. Sorensen, and Uffe Holmskov

CONTENTS

23.1 INTRODUCTION

Surfactant proteins A (SP-A) and D (SP-D) are soluble host defense pattern recognition collectins that recognize carbohydrate arrays present on broad spectrum of microbes [1–3]. These proteins and serum collectin mannan-binding lectins (MBL) preferentially bind carbohydrate and lipid moieties

via their globular domains to facilitate immune functions such as agglutination of bacteria, neutralization of viruses, and clearance of bacteria and fungi. SP-A and SP-D also interact with immunoglobulins, complement component C1q, and receptors present on adaptive and innate immune cells that results in enhanced microbial clearance and reduced lung inflammation [4,5]. Both of these collectins also recognize apoptotic immune cells and their components to mediate their clearance. These proteins, particularly SP-D, regulate lipid homeostasis in the lung [6–8]. Although SP-A is primarily confined to the respiratory system, SP-D is present in several extrapulmonary tissues and at mucosal surfaces. Notably, concentrations of SP-A and SP-D are significantly altered in lung washing as well as in serum during several disease conditions. Furthermore, certain genetic polymorphisms of SP-D (Met11Thr) are known to alter the oligomeric assembly of the protein. Therefore, these collectins may be used as valuable disease markers.

23.2 TISSUE DISTRIBUTION OF SP-A AND SP-D

SP-A and SP-D were originally discovered from the pulmonary surfactant [9–12], a proteolipid complex that overlies the epithelial layer of the lung. These collectins constitute more than 90% (w/w) of the proteins found associated with lung surfactant lipids. Therefore, SP-A and SP-D are ideally located and situated for effectively opsonizing microbes and carbohydrate-coated molecules that reach the enormously large inner surface of the lungs (~100 m^2). Although SP-A is primarily expressed in the respiratory and urogenital tracts [13], SP-D is expressed in several other mucosal surfaces and extrapulmonary tissues including gut, skin, and ovary [13–17]. Both SP-A and SP-D are detectable in the blood, and considered as biomarkers for many lung diseases.

23.3 STRUCTURE OF SP-A AND SP-D

Similar to other collectins, both SP-A and SP-D contain a short interchain disulfide bond-forming N-terminal segment, a collagen-like region with Gly-Xaa-Yaa repeats (where X is any amino acid and Y is often hydroxyproline or hydroxylysine), a hydrophobic neck region [18,19], and a C-terminal globular domain [20,21]. Trimeric subunits of SP-A (3 × ~26–32 kDa) assemble as an octadecamer (540 kDa) by forming a common collagenous stem. These oligomers are further stabilized by covalent interchain disulfide bond formation at the short N-terminal segment [22,23]. The ~20 nm-long quaternary structural assembly orients all the carbohydrate recognition domains (CRDs) to a similar plane [24,25]. Further oligomerization of SP-A results in the formation of multi- or unidirectional protein fibers, ranging up to a few micrometers in length [25,26] (Figures 23.1 and 23.2). Trimeric SP-D subunits (3 × 43 kDa) assemble as higher order oligomers via the interactions at the N-terminal segments and only at the proximal parts of the collagen-like regions [27,28]. Dodecameric SP-D molecules appear as X-shaped ~100 nm-long assemblies (516 kDa) whereas higher order multimers appear as "asterisks or fuzzy balls" [21,27].

SP-D has several common polymorphic forms. Recent studies show that native and recombinant form of SP-D with Thr/Thr11 variant predominantly assembles as trimers. It lacks the usual higher order multimeric form observed for other polymorphic forms of SP-D [29]. The Thr/Thr11 variant occurs in approximately 17% of the Danish population, and is associated with lower serum levels of SP-D. Additional studies in other populations have found a similar frequency of the Thr/Thr11 variant of SP-D [30], suggesting that the polymorphism has been conserved in many populations. The trimeric form of SP-D had reduced binding to mannan, several types of bacteria and influenza A virus (IAV), as compared to binding by the multimeric Met/Met11 form of SP-D. Interestingly, despite reduced binding to these microbes, the trimeric form of SP-D had equivalent binding to lipopolysaccharide (LPS) [29] and to human neutrophils defensins [31]. Binding of this SP-D to LPS or human neutrophils defensins was not mediated by the calcium-dependent lectin activity of the protein (in contrast to binding to bacteria or IAV). Recently, it was demonstrated that natural SP-D in the trimeric form not only has reduced ability to bind IAV, but has less viral neutralizing activity, supporting the hypothesis that the Thr/Thr11 form of SP-D may be associated with reduced innate defense against IAV [32]. The

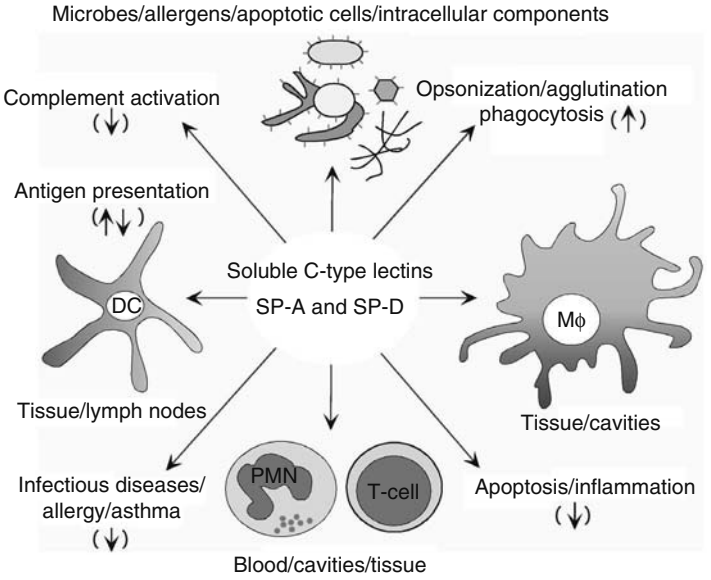


FIGURE 23.1 (See CD for color figure.) Functions of SP-A and SP-D. SP-A and SP-D bind to a variety of microbes, allergens, immunoglobulins, apoptotic cells, and DNA and thereby function as opsonins to enhance uptake of these cells and particles to phagocytes (Mϕ and PMN) including dendritic cells (DC). The binding of SP-A and SP-D to pathogens occurs by different mechanisms. Some pathogens are aggregated or directly lysed after binding and the collectins also have direct effects on immune cells and modulate the production of cutokines and other inflammatory mediators.

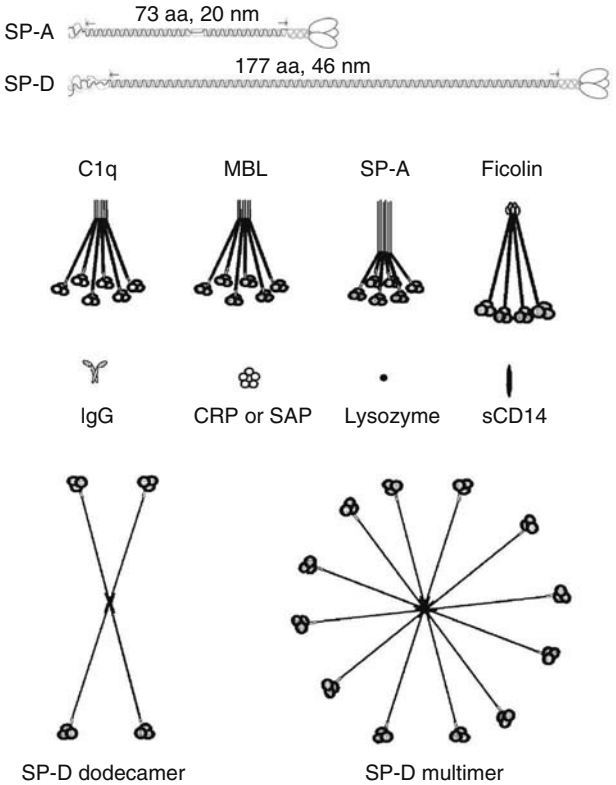


FIGURE 23.2 Subunit and oligomeric structure of SP-A and SP-D. The oligomeric structures are compared to related proteins. The molecules are drawn to show approximate relative dimensions.

Thr/Thr11 variant is associated with increased risk of *Mycobacterium tuberculosis* infection [33] and the Met/Met11 form is associated with increased severity of respiratory syncytial virus (RSV) infection [34].

The globular domains of SP-A and SP-D are primarily responsible for their interactions with many macromolecules (Figure 23.3). SP-A preferentially binds to dipalmitoylphosphatidylcholine, the major surface tension reducing lipid component (~40%–50%, w/w) of the pulmonary surfactant [35,36], via its globular domain [35,37–39]. The phospholipid-binding domain of SP-A partially overlaps with that of the carbohydrate-binding domain [37,40–42]. In contrast, globular domains of SP-D bind phosphatidylinositol, a minor lipid component (~2%, w/w of the lipid) of the surfactant lipid [43,44]. This calcium-dependent phospholipid binding ability of SP-D is similar to that of the serum collectin MBL [40,45]. Carbohydrate-binding pockets of SP-D trimers are spaced as a triangle with ~51 Å to each side (Figure 23.3) [18]. These sites bind hydroxyl groups present on the terminal glucose and related saccharide moieties via calcium-coordinated bonds [46]. It was recently demonstrated that extended carbohydrates like *p*-nitro-phenyl-a-D-maltoside binds with nearly a log-fold higher affinity than maltose, the prototypical carbohydrate competitor for CRDs of SP-D,

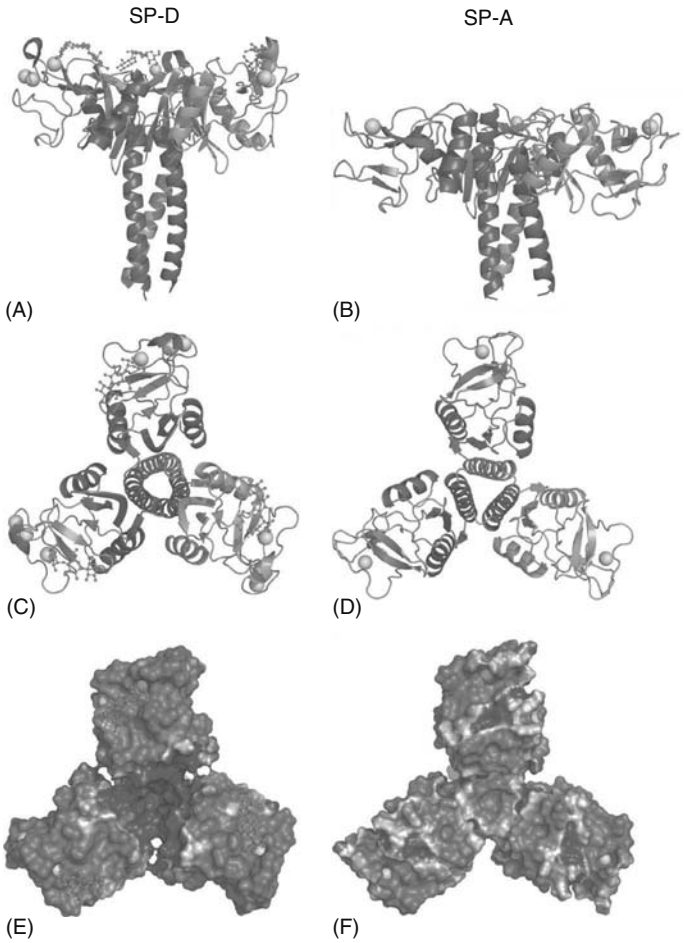


FIGURE 23.3 (See color insert following blank page 170. Also see CD for color figure.) Crystal structure of trimeric nCRD SP-A with calcium (From Head, J.F., Mealy, T.R., McCormack, F.X., and Seaton, B.A., *J. Biol. Chem.*, 278, 43254, 2003) bound and crystal structure of nCRD SP-D bound to maltotriose (From Crouch, E., et al., *J. Biol. Chem.*, 281, 18008, 2006). The figure is made using PyMOL and the atomic coordinates and structure factors (code 1R13 [SP-A] and code 2GGU [SP-D]) are obtained from www.scsb.org. The ribbon presentations are shown as site view and top views while the electrostatic surface representations are shown as top views.

and that residues outside the calcium-binding pocket was responsible for this binding [47]. The center of the three globular domains of one SP-D subunit further has a positively charged surface, suggesting that it may interact effectively with negatively charged components such as DNA and LPS via charge interactions [18,48].

23.4 FUNCTIONS OF SP-A AND SP-D

Since these collectins bind to a variety of ligands, they are suggested to modulate several physiological and pathophysiological functions (Figure 23.1) in addition to their roles in innate immunity [49,50]. The functions attributed to these proteins include the alleviation of allergic responses [51–55], maintenance of surfactant homeostasis [6,7] and surface activity of surfactant in the presence of inhibitors [22,38], prevention of lipid oxidation [56,57], opsonization of apoptotic cells [58,59] and their debris [60] for removal, enhancement of pathogen clearance, and maintenance of inflammation-free lung and other tissues [61–63].

23.4.1 SP-A and SP-D Are Pattern Recognition Molecules

SP-A and SP-D recognize repeating patterns present on microbes and dying cells. They recognize bacterial (LPS, PGN, LTA, LAM), fungal (mannan), and viral (G- and F-protein from RSV, hemagglutinin of influenza, gp-120 of HIV) and mycoplasma (lipids) surface molecules efficiently. Carbohydrate targets present on the host cells are primarily sialic acid based sugars, and hence, are not suitable targets for collectins. However, when the host cells are altered in the apoptotic process, intracellular contents are brought onto the cell surface or released and these molecules become target for collectins. Some of these molecules include nucleic acids (DNA, RNA) [60,64] and glycosaminoglycan [65], and SP-D binds to these components. Recent studies indicate that SP-A and SP-D also bind antibody molecules (IgG, IgM, sIgA, IgE) and complement component C1q [3,66,67]. Therefore, SP-A and SP-D could exert immune functions by multiple mechanisms.

23.4.2 Direct Killing of Microbes

Binding of collectins to the microbial surface increases the permeability of the membrane [68,69]. Therefore, like several other bacteriocidal peptides, collectins also participate in direct killing of microbial pathogens, including Gram-negative bacteria, Gram-positive bacteria, and fungi [70,71]. Interestingly SP-D was recently demonstrated to bind directly to human neutrophil defensins, and in some instances, a cooperative antimicrobial effect between SP-D and the defensins were noted [31].

23.4.3 Microbial Agglutination

It is well known that antibodies (10–20 nm) will bind their target antigens and form immune complexes and form aggregates. In line with this, collectins with multiple subunits effectively agglutinate their targets. SP-D, which has "X" or "asterisk"-like oligomers orienting the trimeric CRDs in multiple directions, agglutinate microorganisms effectively [72]. The agglutination results in the arrest in microbial growth and increases ability of the macrophages to phagocyte the immune complexes.

23.4.4 Chemotaxis of Phagocytes

Both SP-A and SP-D act as chemotactic agents so that the phagocytes are recruited to the site where collectin concentrations are high [73–76]. Binding of collectins appears to induce directed actin polymerization for phagocytes migration [77]. This ability is useful for identifying the killed or aggregated pathogens at the site of infections.

23.4.5 Phagocytosis

Collectin-coated microbes are phagocytosed more effectively than unopsonized ones. Although the mechanistic details are not clearly established, collectins appear to affect both the microbes and immune cells during phagocytosis. It was recently demonstrated that SP-A and SP-D enhance the phagocytosis of *Mycobaterium avium* through increased activity of the macrophage mannose receptor [78] and that SP-A augments the phagocytosis of *Streptococcus pneumonia* by increasing the cell surface expression of the classical scavenger receptor A [79]. Several candidate receptors for SP-A and SP-D have been suggested to exist on phagocytes such as alveolar type II cells [76,80,81], macrophages [82,83], and neutrophils [84].

23.4.5.1 Alveolar Type II Cells

Alveolar type II cells are secretory and phagocytic cells [85,86]. SP-A interacts with receptors on these cells via the collagen-like region [87] or CRDs [88] and is internalized by the coated-pit pathway [89]. Both SP-A and SP-D are internalized and resecreted by type II cells, but only SP-A appears to have the ability to regulate lipid recycling [80]. Although the effect of SP-A on type II cell and lipid recycling has been well-established *in vitro* [22], SP-A knockout mouse models showed no apparent defect in lipid homeostasis [63,90].

23.4.5.2 Macrophages and Monocytes

SP-A and SP-D bind to monocytic cells or cell lines and alveolar macrophages [91,92]. Binding of SP-D to alveolar macrophages [92] and although the receptors for these molecules are not clearly identified, these collectins are internalized and degraded by phagocytes [93,94]. Mouse models show that SP-D, but not SP-A, deficiency leads to the accumulation of macrophages and lipids in the mouse lungs [6,7]. These defects could be corrected by treatment with native SP-D [62] or recombinant fragments of SP-D, or SP-D(n/CRD) [58].

23.4.5.3 Neutrophils

Microbes opsonized by both SP-A and SP-D are preferentially phagocytosed by polymorphonuclear leukocytes (PMNs or PML) [75,84]. LPS-mediated inflammatory changes particularly increase the amount of SP-D associated with PMNs and lung tissue [76]. Like alveolar macrophages, PMNs also phagocytose the lipids present in the lung [95], and lipid concentration may be altered by PMNs during lung injury and inflammation.

23.4.5.4 Lymphocytes

The CD8 cytotoxic T-cells directly kill infected cells by inducing apoptosis. In contrast, T_H1-type CD4 cells secrete IFN-γ and IL-2, which are important for activating macrophages and preventing granulomatous disease such as tuberculosis. The T_H2-type CD4 cells secrete IL-4, IL-5, IL-10, and IL-13, which are important to induce antibody generation and secretion from B-lymphocytes. Current paradigm is that the shift from T_H2 to T_H1 type CD4 cells and cytokine profiles is essential to minimize pulmonary allergy and asthma [96,97]. SP-A and SP-D alleviate allergy and modulate adaptive immune response [54,97–99]. SP-A and SP-D appear to interact with lymphocytes via the collagen-like region of the protein [99,100] and the IL-4 and IL-13 seems to form a negative feedback circuit with surfactant protein-D in the allergic airway response [101].

23.4.6 Intracellular Killing

SP-A and SP-D-mediated uptake of immune complexes or binding collectins to phagocytes increases the oxidative bust in immune cells [102,103]. The exact molecular mechanisms, however, have been not determined.

23.4.7 Clearance of Apoptotic Cells

SP-D deficiency results in the accumulation of apoptotic alveolar macrophages in the lungs [58]. Treatment of these mice with a 60 kDa SP-D (n/CRD) recombinant fragment that contains the proximal collagen-like $(Gly\text{-}X\text{-}Y)_8$ region, neck, and CRDs, or native SP-D corrects accumulation of apoptotic alveolar macrophages [58]. SP-A and SP-D recognize apoptotic cells [59], however, the molecules that they recognize on the surface of these cells are uncertain. SP-D, but not SP-A, effectively binds nucleic acids (DNA and RNA) via both its collagen-like regions and globular CRDs (Figure 23.3) [104]. This binding appears to be important *in vivo* since SP-D knockout mice accumulate DNA in the lung and anti-DNA antibodies in the sera [104].

23.4.8 Antigen Presentation via Dendritic Cells

SP-A and SP-D are involved in dendritic cell (DC)-mediated uptake and processing of antigens [105,106]. It is known that immature DCs act only as phagocytes, whereas the mature DCs process foreign antigen, and present them to the adaptive immune system. SP-A inhibits DC differentiation and subsequent DC-mediated T-cell activation, whereas SP-D promotes these functions. This finding represents a significant understanding of the role of SP-A and SP-D in interlinking innate and adaptive immunity [49].

23.5 LINKING INNATE AND ADAPTIVE IMMUNE SYSTEMS

Antibodies and complement proteins are important molecules in executing the functions of adaptive immunity. Both SP-A [66] and SP-D [3] recognize antibody molecules via their CRDs. Since SP-D effectively recognizes various antibody molecules including IgG, secretory IgA, IgE, and IgM, and enhances IgG-coated model immune complexes [3,66], it could utilize the adaptive immune molecules for effective clearance of immune complexes. SP-A and SP-D also interact with complement protein C1q [67]. SP-A–C1q interactions have been proposed to downregulate complement activation and subsequent inflammatory response [107].

Interestingly, collectins are evolutionarily more ancient molecules than antibodies. For example, tunicates express collectins whereas antibodies are only detected in younger organisms such as lamprey [108]. Therefore, higher organisms have the advantage that innate immune molecules such as collectins recognize carbohydrate arrays on microbes and are furthermore able to utilize antibody molecules to effectively defend the host against microbial pathogens.

Toll-like receptors (TLRs) are transmembrane proteins that contain extracellular leucine-rich repeat (LRR) domains. They are present on phagocytes, bind microbial cell-surface components, and regulate signaling pathways to initiate an immune response. Both SP-A and SP-D interact with TLRs. SP-A directly interacts with TLR-2 and reduces the interaction between soluble ligands and the receptor, and suppresses TNF-α production by U937 monocytic cells [109]. SP-A also interacts with TLR-4 and modulates inflammation [110]. SP-D likewise binds decorin, another LRR domain-containing protein [65], and TLR-2 and TLR-4 [111]. Therefore, SP-A and SP-D could interlink innate and adaptive immune systems via more than one mechanism.

23.6 RECEPTOR INTERACTIONS OF THE PULMONARY COLLECTINS

SP-A and SP-D mediate multiple functions via different receptors. These collectins interact with proteins/receptors, including decorin [65], glycoprotein-340 (gp-340) [112,113], and microfibril-associated protein 4 (MFAP-4) [114,115]. When collectins bind apoptotic cells, they interact with CD91/calreticulin receptor complex [59,82,116]. It has been proposed that collectins may act either as pro- or anti-inflammatory proteins depending on their interactions with CD91/calreticulin and

signal inhibitory regulatory protein α. The orientation of collectins with respect to the ligands and phagocytes appears to determine the signaling pathways [82]. CD14 is another receptor involved in apoptotic cell clearance. Both SP-A and SP-D interact with this receptor, hence, such binding may be important for microbial and apoptotic cell clearance [117,118]. SP-A and SP-D has been implicated to interact with several other receptors or binding molecules [119].

23.7 OTHER FUNCTIONS

SP-A is essential for reorganization of lung surfactant into intricate tubular arrays known as tubular myelin (TM) [26,120,121]. TM has previously been considered as an essential intermediate in the conversion of secreted lamellar body into the surfactant film. SP-A knockout mice lack TM but form stable surfactant film [90,122,123], therefore, TM is not essential for surface-active film formation. However, the SP-A–TM interaction may be important under altered physiological conditions such as hyperventilation [124].

Gene-knockout mice models show that SP-D plays an important function in lipid homeostasis *in vivo* [6,7]. Lipid-related defects include an age-dependent accumulation of surfactant lipids in the alveolus and alveolar macrophages. SP-D knockout mice have chronically inflamed lungs, increased matrix metalloproteinase expression, and progressive development of emphysema and fibrosis. Globular domains of these proteins inhibit the oxidation of lipids [56], and SP-A prevents the inactivation of surfactant film by oxidized lipids *in vitro* [57]. Moreover, enhanced bodyweight and fat percentage characterize SP-D-deficient mice [125].

23.8 DISEASES

Although human disease resulting from one or more mutations or deletions in SP-A or SP-D has not been identified, SP-A and SP-D are considered to be markers of or contributors to the pathogenesis of clinical conditions characterized by increased mucosal inflammation or lung parenchyma damage. In addition, accumulating evidence in mouse models of infection and inflammation indicates that recombinant forms of the surfactant proteins are biologically active *in vivo* and may have therapeutic potential in controlling pulmonary inflammatory disease. The presence of the surfactant collectins, especially SP-D, in nonpulmonary tissues, such as the gastrointestinal tract and genital organs, suggest additional actions located to other mucosal surfaces. In the following sections, we highlight the roles of SP-A and SP-D in infectious diseases, asthma, and cystic fibrosis (CF). We refer the readers to Sorensen et al. [126] for a detailed review on the effects of SP-A and SP-D in human pulmonary disease.

23.8.1 Bacterial Infections

SP-A and SP-D bind Gram-positive and Gram-negative bacteria as well as mycobacteria. These interactions result in differential biological response. The major cell wall component of Gram-negative bacteria is LPS, which is a ligand for both SP-A and SP-D, and the concentrations of both of these proteins increase in response to intratracheal instillation of LPS [127]. SP-A binds and aggregates lipid A and LPS from rough strains, but not LPS from smooth strains [128], whereas SP-D binds to LPS core polysaccharides. SP-A, but not SP-D, increases the uptake and degradation of LPS by alveolar macrophages [129]. A possible mechanism is that SP-A binds directly to CD14 and inhibits LPS binding and cellular responses to LPS [128]. However, both SP-D and MBL also bind CD14 [117,130] but by different mechanism. CD14 binds complexes of LPS and LPS-binding protein and then forms a complex with TLR4 and MD-2. This leads to stimulation of cells by the induction of NF-κB. Recombinant SP-D administration improved survival of newborn lambs exposed to intratracheal LPS indicating its involvement in the resolution of endotoxin induced lung injury [131]. SP-A binds to *Hemophilus influenzae* type A via P2 outer membrane protein, but not to the highly virulent type B strain [132].

SP-A and SP-D also bind to Gram-positive bacteria where the ligand for SP-D is lipoteichoic acid and peptidoglycan [133]. The major cell wall components of Gram-positive bacteria are peptidoglycan and lipoteichoic acid. Although SP-A does not bind peptidoglycan, it significantly reduces peptidoglycan-elicited TNF-α secretion from alveolar macrophages by a mechanism that involves direct binding of SP-A to TLR2 [109]. SP-D binds both of the major ligands present on Gram-positive bacteria [133].

SP-A binds mycobacteria and increases the phagocytosis of *M. tuberculosis* by alveolar macrophages [134]. In contrast, SP-D inhibits uptake of *M. tuberculosis* by alveolar macrophages, a process that is independent of bacterial agglutination [135]. Increased attachment of *M. tuberculosis* to the alveolar macrophage often increases the pathogenesis of pulmonary tuberculosis. Both SP-A and SP-D bind lipoarabinomannan of mycobacteria [136,137]. Although the recognition of LAMs by SP-A requires lipid components, binding of SP-D to LAMs requires lectin–carbohydrate type interactions. These proteins appear to work in concert with mannose receptor to exert their biological functions in the killing and clearance of mycobacteria [138].

Studies conducted in *Spa* (−/−) and *Spd* (−/−) mice suggest that these proteins promote the clearance of microbial cell wall components from the lung and protect the lung from inflammatory injury [139].

23.8.2 Fungal Infection

SP-A and SP-D recognize several fungi and the major interactions occur via cell wall mannan or other glycoproteins [1]. *Pneumocystis carinii* is a common cause of life-threatening pneumonia in immunocompromised patients. SP-A and SP-D bind specifically to the surface of *P. carinii*, reacting with the glycoprotein gpA. SP-A and SP-D are found on the surface of *P. carinii* in fresh isolates and both molecules enhance the binding of *P. carinii* to alveolar macrophages [140]. These collectins bind uncapsulated *Cryptococcus neoformans* [141], but only SP-D agglutinates the fungus. *Aspergillus fumigatus* is an opportunistic pathogen and causes pulmonary infections mainly in immunosuppressed hosts. Both SP-A and SP-D bind carbohydrate moieties present on *A. fumigatus* conidia in a calcium-dependent manner [142]. They enhance the phagocytosis and killing of *A. fumigatus* conidia by human alveolar macrophages and circulating neutrophils [75]. Interestingly, intranasal administration of SP-D increased survival from 0% to 80% in a murine model of invasive pulmonary aspergillosis and this treatment was as efficient as treatment with amphotericin B [75].

23.8.3 Viral Infections

SP-A and SP-D bind glycoproteins present on enveloped viruses, including influenza virus, human immunodeficiency virus (HIV), and herpes simplex virus. They also bind to the nonenveloped rotavirus. In general, the binding is due to the interaction of the collectin carbohydrate-binding site with viral envelope glycoprotein, but the binding of SP-A to influenza A virus and to herpes simplex virus 1 is through the interaction of a viral lectin with the N-linked carbohydrate on the CRD of SP-A. Both inhibition and promotion of viral infection by the collectins have been reported.

The interaction between collectins and influenza A virus has been extensively examined both *in vitro* and *in vivo* [4]. SP-D [143,144] and SP-A [143,145] inhibit influenza virus hemagglutination and infectivity. SP-D strongly agglutinates virus, which could lead to mucociliary clearance of the virus. SP-D also enhances the binding and uptake of influenza A virus by neutrophil granulocytes [143,146], which protects the neutrophils against influenza A virus-mediated depression of functional activity. Instead, an enhanced respiratory burst response to influenza A virus is observed when SP-D is present [144]. SP-A also mediates the binding of influenza A to neutrophils, but this does not protect the neutrophils against the virus [143]. In contrast, SP-A opsonizes influenza A virus for phagocytosis by alveolar macrophages, an effect not shown by SP-D [147].

23.8.4 Allergy and Asthma

SP-A and SP-D are significantly altered in allergy and related conditions [54,98,148,149]. These collectins appear to modulate multiple mechanisms: They block the interaction between IgE and allergens [98,150] and block the binding of allergens to IgE and subsequent histamine release by basophils [98] and mononucleated cells [99]. SP-A inhibits the secretion of IL-8 by eosinophils [151], suggesting that it may dampen the inflammatory response. Animal models show that both SP-A and SP-D reduce IgE, blood eosinophilia, allergic responses, and airway hypersensitivity [51,54]. Interestingly, treatment of allergic mice with SP-A and SP-D suppresses subsequent allergic reactions. Collectin-treated mice continued to maintain low concentrations of IgE and eosinophils in the serum whereas allergen exposure leads to an immediate and dramatic (~10-fold) upregulation of these collectin, particularly SP-D, concentrations in the lung [52,53,55,149,152,153]. Overexpression of allergic type (T_H2) cytokines such as IL-5 [152] and IL-13 [154] also results in the upregulation of SP-D.

SP-A and SP-D concentrations increase in alveolar lavage of asthmatic patients [155]. The baseline levels of SP-D were also significantly elevated in serum of allergic asthmatic patients. Serum SP-D was predictive of the late asthmatic response and for eosinophil cationic protein concentrations after allergen challenge [156]. A variety of small clinical studies and case studies have nevertheless indicated raised serum SP-D levels associated to allergy [157–161]. Despite the somewhat contradictory data regarding the relations between SP-A/SP-D levels and asthma/allergy, mouse models, explants investigations, and cell culture studies together propose essential roles for SP-A and SP-D in immune regulation involved in the pathogenesis of allergic asthma.

23.8.5 Cystic Fibrosis

CF is caused by a mutation in an ion transporter protein, however, the morbidity and mortality of CF patients are attributable to lung infections. SP-A concentration is high in the bronchoalveolar lavage fluid (BALF) of infants and young children irrespective of the presence or absence of pulmonary inflammation [162]. However, in older children and adults with CF, the SP-A concentration is low in BALF [163–166]. The SP-D concentration in the BALF of young patients with CF is low [167] as well as in patients with CF with more advanced lung disease [163]. However, in patients who were clinically stable, the SP-D (in contrast to SP-A) was not significantly decreased unless infection was present [166].

Proteases in CF lungs degrade SP-A and SP-D [2,168–173]. SP-A [174–177] and SP-D [166,178,179] enhance the phagocytosis of CF associated pathogenic microorganism *Pseudomonas aeruginosa*. A recent study has further demonstrated partial restoration of phagocytosis and CD14 expression by SP-A treatment of CF bronchoalveolar cells ex vivo, indicating that some defects may be reversible in CF patients [180]. The clearance of apoptotic cells, which is further of critical importance in the control of inflammation, is defective in CF and may be affected by both SP-A and SP-D [59]. Serum SP-D was markedly increased in CF patients. Serum SP-D levels were suitable for monitoring pulmonary function, but not suitable for diagnosis of the infection [181]. On the basis of such data, therapy based on pulmonary administration of recombinant SP-D or guidance of therapy based on proteomic evaluation of pulmonary protein composition have previously been suggested [169,182].

23.9 CONCLUSION AND FUTURE DIRECTIONS

On the basis of the activity *in vitro*, their patterns of expression and gene regulation, and their involvement in pathways of innate immunity, SP-A and SP-D are suggested to serve as host defence molecules both by direct antimicrobial activities and immunoregulatory functions. Moreover, mouse models of gene deficiency have supported specific antimicrobial effects of the molecules and a role

in regulation of inflammation. SP-D was further recognized as a regulator of lipid homeostasis. The two collectins have thus emerged as effector molecules of the innate immune system involving activities not only as endogenous antibiotics but also as mediators of inflammation. They also involve in the cross-talk between innate and adaptive immunity through interactions with complement system via antibody molecules, and DCs via TLRs.

Future perspectives in collectin research include a more elaborate description of the immunoregulatory functions of SP-A and SP-D. The signaling has primarily been recognized in terms of its boosting effect on cellular responsiveness in host defence systems. Further comprehensive understanding of this signaling may contribute to the development of potential therapeutic applications. Despite suggestive benefits of replacement therapy and the value of SP-A and SP-D in predicting the susceptibility of certain diseases or the extent of disease like infection, asthma, and CF, for now, investigation of the collectin status remains primarily a research tool. Future studies should include a more rigorous examination of collectin status by both measurement of systemic levels and genotyping in large patient cohorts. This may help to identify the most important disease associations and identify those clinical settings in which replacement therapy is most likely to be beneficial.

ACKNOWLEDGMENT

We thank Jesper B. Møller for help with Figure 23.3.

REFERENCES

1. Holmskov, U., Thiel, S., and Jensenius, J.C. Collections and ficolins: Humoral lectins of the innate immune defense. *Annu Rev Immunol* 21, 547–578, 2003.
2. Beatty, A.L., Malloy, J.L., and Wright, J.R. Pseudomonas aeruginosa degrades pulmonary surfactant and increases conversion in vitro. *Am J Respir Cell Mol Biol* 32, 128–134, 2005.
3. Nadesalingam, J., Reid, K.B., and Palaniyar, N. Collectin surfactant protein D binds antibodies and interlinks innate and adaptive immune systems. *FEBS Lett* 579, 4449–4453, 2005.
4. Crouch, E., Hartshorn, K., and Ofek, I. Collectins and pulmonary innate immunity. *Immunol Rev* 173, 52–65, 2000.
5. Wright, J.R. Immunoregulatory functions of surfactant proteins. *Nat Rev Immunol* 5, 58–68, 2005.
6. Botas, C., et al. Altered surfactant homeostasis and alveolar type II cell morphology in mice lacking surfactant protein D. *Proc Natl Acad Sci U S A* 95, 11869–11874, 1998.
7. Korfhagen, T.R., et al. Surfactant protein-D regulates surfactant phospholipid homeostasis in vivo. *J Biol Chem* 273, 28438–28443, 1998.
8. Hawgood, S. and Poulain, F.R. The pulmonary collectins and surfactant metabolism. *Annu Rev Physiol* 63, 495–519, 2001.
9. King, R.J. and Clements, J.A. Surface active materials from dog lung. II. Composition and physiological correlations. *Am J Physiol* 223, 715–726, 1972.
10. Akatsuka, K., et al. Rat zona pellucida glycoproteins: Molecular cloning and characterization of the three major components. *Mol Reprod Dev* 51, 454–467, 1998.
11. Persson, A., et al. Purification and biochemical characterization of CP4 (SP-D), a collagenous surfactant-associated protein. *Biochemistry* 28, 6361–6367, 1989.
12. Possmayer, F. A proposed nomenclature for pulmonary surfactant-associated proteins. *Am Rev Respir Dis* 138, 990–998, 1988.
13. Akiyama, J., et al. Tissue distribution of surfactant proteins A and D in the mouse. *J Histochem Cytochem* 50, 993–996, 2002.
14. Fisher, J.H. and Mason, R. Expression of pulmonary surfactant protein D in rat gastric mucosa. *Am J Respir Cell Mol Biol* 12, 13–18, 1995.
15. Madsen, J., et al. Localization of lung surfactant protein D on mucosal surfaces in human tissues. *J Immunol* 164, 5866–5870, 2000.
16. van Rozendaal, B.A., van Golde, L.M., and Haagsman, H.P. Localization and functions of SP-A and SP-D at mucosal surfaces. *Pediatr Pathol Mol Med* 20, 319–339, 2001.
17. Stahlman, M.T., Gray, M.E., Hull, W.M., and Whitsett, J.A. Immunolocalization of surfactant protein-D (SP-D) in human fetal, newborn, and adult tissues. *J Histochem Cytochem* 50, 651–660, 2002.

18. Hakansson, K., Lim, N.K., Hoppe, H.J., and Reid, K.B. Crystal structure of the trimeric alpha-helical coiled-coil and the three lectin domains of human lung surfactant protein D. *Structure Fold Des* 7, 255–264, 1999.
19. Kovacs, H., et al. Solution structure of the coiled-coil trimerization domain from lung surfactant protein D. *J Biomol NMR* 24, 89–102, 2002.
20. Crouch, E., Persson, A., Chang, D., and Heuser, J. Molecular structure of pulmonary surfactant protein D (SP-D). *J Biol Chem* 269, 17311–17319, 1994.
21. Crouch, E., Chang, D., Rust, K., Persson, A., and Heuser, J. Recombinant pulmonary surfactant protein D. Post-translational modification and molecular assembly. *J Biol Chem* 269, 15808–15813, 1994.
22. McCormack, F.X. Functional mapping of surfactant protein A. *Pediatr Pathol Mol Med* 20, 293–318, 2001.
23. Palaniyar, N., Ikegami, M., Korfhagen, T., Whitsett, J., and McCormack, F.X. Domains of surfactant protein A that affect protein oligomerization, lipid structure and surface tension. *Comp Biochem Physiol A Mol Integr Physiol* 129, 109–127, 2001.
24. Voss, T., Eistetter, H., Schafer, K.P., and Engel, J. Macromolecular organization of natural and recombinant lung surfactant protein SP 28-36. Structural homology with the complement factor C1q. *J Mol Biol* 201, 219–227, 1988.
25. Palaniyar, N., Ridsdale, R.A., Possmayer, F., and Harauz, G. Surfactant protein A (SP-A) forms a novel supraquaternary structure in the form of fibers. *Biochem Biophys Res Commun* 250, 131–136, 1998.
26. Palaniyar, N., et al. Filaments of surfactant protein A specifically interact with corrugated surfaces of phospholipid membranes. *Am J Physiol* 276, L631–641, 1999.
27. Lu, J., et al. Structural similarity between lung surfactant protein D and conglutinin. Two distinct, C-type lectins containing collagen-like sequences. *Eur J Biochem* 215, 793–799, 1993.
28. Palaniyar, N., et al. The role of pulmonary collectin N-terminal domains in surfactant structure, function, and homeostasis in vivo. *J Biol Chem* 277, 26971–26979, 2002.
29. Leth-Larsen, R., et al. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 174, 1532–1538, 2005.
30. Heidinger, K., et al. Polymorphisms in the human surfactant protein-D (SFTPD) gene: Strong evidence that serum levels of surfactant protein-D (SP-D) are genetically influenced. *Immunogenetics* 57, 1–7, 2005.
31. Hartshorn, K.L., White, M.R., Tecle, T., Holmskov, U., and Crouch, E.C. Innate defense against influenza A virus: Activity of human neutrophil defensins and interactions of defensins with surfactant protein D. *J Immunol* 176, 6962–6972, 2006.
32. Hartshorn, K.L., et al. Reduced influenza viral neutralizing activity of natural human trimers of surfactant protein D. *Respir Res* 8, 9, 2007.
33. Floros, J., et al. Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population. *J Infect Dis* 182, 1473–1478, 2000.
34. Lahti, M., et al. Surfactant protein D gene polymorphism associated with severe respiratory syncytial virus infection. *Pediatr Res* 51, 696–699, 2002.
35. Kuroki, Y. and Akino, T. Pulmonary surfactant protein A (SP-A) specifically binds dipalmitoylphosphatidylcholine. *J Biol Chem* 266, 3068–3073, 1991.
36. Yu, S.H. and Possmayer, F. Dipalmitoylphosphatidylcholine and cholesterol in monolayers spread from adsorbed films of pulmonary surfactant. *J Lipid Res* 42, 1421–1429, 2001.
37. Palaniyar, N., McCormack, F.X., Possmayer, F., and Harauz, G. Three-dimensional structure of rat surfactant protein A trimers in association with phospholipid monolayers. *Biochemistry* 39, 6310–6316, 2000.
38. Ikegami, M., et al. The collagen-like region of surfactant protein A (SP-A) is required for correction of surfactant structural and functional defects in the SP-A null mouse. *J Biol Chem* 276, 38542–38548, 2001.
39. Blanco, O. and Catala, A. Surfactant protein A inhibits the non-enzymatic lipid peroxidation of porcine lung surfactant. *Prostaglandins Leukot Essent Fatty Acids* 65, 185–190, 2001.
40. Chiba, H., et al. Introduction of mannose binding protein-type phosphatidylinositol recognition into pulmonary surfactant protein A. *Biochemistry* 38, 7321–7331, 1999.
41. Sano, H., et al. Analysis of chimeric proteins identifies the regions in the carbohydrate recognition domains of rat lung collectins that are essential for interactions with phospholipids, glycolipids, and alveolar type II cells. *J Biol Chem* 273, 4783–4789, 1998.
42. Head, J.F., Mealy, T.R., McCormack, F.X., and Seaton, B.A. Crystal structure of trimeric carbohydrate recognition and neck domains of surfactant protein A. *J Biol Chem* 278, 43254–43260, 2003.
43. Ogasawara, Y., Kuroki, Y., and Akino, T. Pulmonary surfactant protein D specifically binds to phosphatidylinositol. *J Biol Chem* 267, 21244–21249, 1992.

44. Persson, A.V., Gibbons, B.J., Shoemaker, J.D., Moxley, M.A., and Longmore, W.J. The major glycolipid recognized by SP-D in surfactant is phosphatidylinositol. *Biochemistry* 31, 12183–12189, 1992.
45. Kuroki, Y., et al. A novel type of binding specificity to phospholipids for rat mannose- binding proteins isolated from serum and liver. *FEBS Lett* 414, 387–392, 1997.
46. Weis, W.I., Drickamer, K., and Hendrickson, W.A. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* 360, 127–134, 1992.
47. Crouch, E., et al. Contributions of phenylalanine 335 to ligand recognition by human surfactant protein D: Ring interactions with SP-D ligands. *J Biol Chem* 281, 18008–18014, 2006.
48. Hakansson, K. and Reid, K.B. Collectin structure: A review. *Protein Sci* 9, 1607–1617, 2000.
49. Palaniyar, N., Nadesalingam, J., and Reid, K.B. Pulmonary innate immune proteins and receptors that interact with gram-positive bacterial ligands. *Immunobiology* 205, 575–594, 2002.
50. Reid, K.B., Colomb, M., Petry, F., and Loos, M. Complement component C1 and the collectins–first-line defense molecules in innate and acquired immunity. *Trends Immunol* 23, 115–117, 2002.
51. Madan, T., et al. Surfactant proteins A and D protect mice against pulmonary hypersensitivity induced by *Aspergillus fumigatus* antigens and allergens. *J Clin Invest* 107, 467–475, 2001.
52. Haczku, A., et al. *Aspergillus fumigatus*-induced allergic airway inflammation alters surfactant homeostasis and lung function in BALB/c mice. *Am J Respir Cell Mol Biol* 25, 45–50, 2001.
53. Wang, J.Y., Shieh, C.C., Yu, C.K., and Lei, H.Y. Allergen-induced bronchial inflammation is associated with decreased levels of surfactant proteins A and D in a murine model of asthma. *Clin Exp Allergy* 31, 652–662, 2001.
54. Strong, P., Reid, K.B., and Clark, H. Intranasal delivery of a truncated recombinant human SP-D is effective at down-regulating allergic hypersensitivity in mice sensitized to allergens of *Aspergillus fumigatus*. *Clin Exp Immunol* 130, 19–24, 2002.
55. Haley, K.J., et al. Alterations in lung collectins in an adaptive allergic immune response. *Am J Physiol Lung Cell Mol Physiol* 282, L573–L584, 2002.
56. Bridges, J.P., et al. Pulmonary surfactant proteins A and D are potent endogenous inhibitors of lipid peroxidation and oxidative cellular injury. *J Biol Chem* 275, 38848–38855, 2000.
57. Rodriguez Capote, K., McCormack, F.X., and Possmayer, F. Pulmonary surfactant protein-A (SP-A) restores the surface properties of surfactant after oxidation by a mechanism that requires the Cys6 interchain disulfide bond and the phospholipid binding domain. *J Biol Chem* 278, 20461–20474, 2003.
58. Clark, H., et al. Surfactant protein D reduces alveolar macrophage apoptosis in vivo. *J Immunol* 169, 2892–2899, 2002.
59. Vandivier, R.W., et al. Role of surfactant proteins A, D, and C1q in the clearance of apoptotic cells in vivo and in vitro: Calreticulin and CD91 as a common collectin receptor complex. *J Immunol* 169, 3978–3986, 2002.
60. Palaniyar, N., Clark, H., Nadesalingam, J., Hawgood, S., and Reid, K.B. Surfactant protein D binds genomic DNA and apoptotic cells, and enhances their clearance, in vivo. *Ann N Y Acad Sci* 1010, 471–475, 2003.
61. LeVine, A.M. and Whitsett, J.A. Pulmonary collectins and innate host defense of the lung. *Microbes Infect* 3, 161–166, 2001.
62. Zhang, L., Ikegami, M., Dey, C.R., Korfhagen, T.R., and Whitsett, J.A. Reversibility of pulmonary abnormalities by conditional replacement of surfactant protein D (SP-D) in vivo. *J Biol Chem* 277, 38709–38713, 2002.
63. Li, G., et al. Surfactant protein-A–deficient mice display an exaggerated early inflammatory response to a beta-resistant strain of influenza A virus. *Am J Respir Cell Mol Biol* 26, 277–282, 2002.
64. Palaniyar, N., et al. Nucleic acid is a novel ligand for innate, immune pattern recognition collectins surfactant proteins A and D and mannose-binding lectin. *J Biol Chem* 279, 32728–32736, 2004.
65. Nadesalingam, J., et al. Identification and characterization of a novel interaction between pulmonary surfactant protein D and decorin. *J Biol Chem* 278, 25678–25687, 2003.
66. Lin, P.M. and Wright, J.R. Surfactant protein A binds to IgG and enhances phagocytosis of IgG-opsonized erythrocytes. *Am J Physiol Lung Cell Mol Physiol* 291, L1199–L1206, 2006.
67. Watford, W.T., Smithers, M.B., Frank, M.M., and Wright, J.R. Surfactant protein A enhances the phagocytosis of C1q-coated particles by alveolar macrophages. *Am J Physiol Lung Cell Mol Physiol* 283, L1011–L1022, 2002.
68. Wu, H., et al. Surfactant proteins A and D inhibit the growth of Gram-negative bacteria by increasing membrane permeability. *J Clin Invest* 111, 1589–1602, 2003.
69. McCormack, F.X., et al. Macrophage-independent fungicidal action of the pulmonary collectins. *J Biol Chem* 278, 36250–36256, 2003.

70. McCormack, F.X. New concepts in collectin-mediated host defense at the air-liquid interface of the lung. *Respirology* 11 Suppl, S7–10, 2006.
71. Kingma, P.S., Zhang, L., Ikegami, M., Hartshorn, K., McCormack, F.X., and Whitsett, J.A. Correction of pulmonary abnormalities in Sftpd–/– mice requires the collagenous domain of surfactant protein D. *J Biol Chem* 281, 24496–24505, 2006.
72. Kuan, S.F., Rust, K., and Crouch, E. Interactions of surfactant protein D with bacterial lipopolysaccharides. Surfactant protein D is an *Escherichia coli*-binding protein in bronchoalveolar lavage. *J Clin Invest* 90, 97–106, 1992.
73. Cai, G.Z., Griffin, G.L., Senior, R.M., Longmore, W.J., and Moxley, M.A. Recombinant SP-D carbohydrate recognition domain is a chemoattractant for human neutrophils. *Am J Physiol* 276, L131–L136, 1999.
74. Crouch, E.C., Persson, A., Griffin, G.L., Chang, D., and Senior, R.M. Interactions of pulmonary surfactant protein D (SP-D) with human blood leukocytes. *Am J Respir Cell Mol Biol* 12, 410–415, 1995.
75. Madan, T., et al. Binding of pulmonary surfactant proteins A and D to *Aspergillus fumigatus* conidia enhances phagocytosis and killing by human neutrophils and alveolar macrophages. *Infect Immun* 65, 3171–3179, 1997.
76. Herbein, J.F. and Wright, J.R. Enhanced clearance of surfactant protein D during LPS-induced acute inflammation in rat lung. *Am J Physiol Lung Cell Mol Physiol* 281, L268–L277, 2001.
77. Tino, M.J. and Wright, J.R. Surfactant proteins A and D specifically stimulate directed actin-based responses in alveolar macrophages. *Am J Physiol* 276, L164–L174, 1999.
78. Kudo, K., et al. Pulmonary collectins enhance phagocytosis of *Mycobacterium avium* through increased activity of mannose receptor. *J Immunol* 172, 7592–7602, 2004.
79. Kuronuma, K., et al. Pulmonary surfactant protein A augments the phagocytosis of *Streptococcus pneumoniae* by alveolar macrophages through a casein kinase 2-dependent increase of cell surface localization of scavenger receptor A. *J Biol Chem* 279, 21421–21430, 2004.
80. Williams, M.C. Uptake of lectins by pulmonary alveolar type II cells: Subsequent deposition into lamellar bodies. *Proc Natl Acad Sci USA* 81, 6383–6387, 1984.
81. Herbein, J.F., Savov, J., and Wright, J.R. Binding and uptake of surfactant protein D by freshly isolated rat alveolar type II cells. *Am J Physiol Lung Cell Mol Physiol* 278, L830–L839, 2000.
82. Gardai, S.J., et al. By binding SIRPalpha or calreticulin/CD91, lung collectins act as dual function surveillance molecules to suppress or enhance inflammation. *Cell* 115, 13–23, 2003.
83. Yang, Q.W., et al. Role of Toll-like receptor 4/NF-kappaB pathway in monocyte-endothelial adhesion induced by low shear stress and ox-LDL. *Biorheology* 42, 225–236, 2005.
84. Hartshorn, K.L., et al. Pulmonary surfactant proteins A and D enhance neutrophil uptake of bacteria. *Am J Physiol* 274, L958–L969, 1998.
85. Fehrenbach, H. Alveolar epithelial type II cell: Defender of the alveolus revisited. *Respir Res* 2, 33–46, 2001.
86. Mason, R.J. and Williams, M.C. Type II alveolar cell. Defender of the alveolus. *Am Rev Respir Dis* 115, 81–91, 1977.
87. Kresch, M.J., Christian, C., and Lu, H. Isolation and partial characterization of a receptor to surfactant protein A expressed by rat type II pneumocytes. *Am J Respir Cell Mol Biol* 19, 216–225, 1998.
88. Wright, J.R., Borchelt, J.D., and Hawgood, S. Lung surfactant apoprotein SP-A (26–36 kDa) binds with high affinity to isolated alveolar type II cells. *Proc Natl Acad Sci USA* 86, 5410–5414, 1989.
89. Stevens, P.A., Wissel, H., Zastrow, S., Sieger, D., and Zimmer, K.P. Surfactant protein A and lipid are internalized via the coated-pit pathway by type II pneumocytes. *Am J Physiol Lung Cell Mol Physiol* 280, L141–L151, 2001.
90. Korfhagen, T.R., et al. Altered surfactant function and structure in SP-A gene targeted mice. *Proc Natl Acad Sci USA* 93, 9594–9599, 1996.
91. Pison, U., Wright, J.R., and Hawgood, S. Specific binding of surfactant apoprotein SP-A to rat alveolar macrophages. *Am J Physiol* 262, L412–417, 1992.
92. Miyamura, K., et al. Surfactant protein D binding to alveolar macrophages. *Biochem J* 300, 237–242, 1994.
93. Dong, Q. and Wright, J.R. Degradation of surfactant protein D by alveolar macrophages. *Am J Physiol* 274, L97–L105, 1998.
94. Wright, J.R. and Youmans, D.C. Degradation of surfactant lipids and surfactant protein A by alveolar macrophages in vitro. *Am J Physiol* 268, L772–L780, 1995.
95. Quintero, O.A. and Wright, J.R. Clearance of surfactant lipids by neutrophils and macrophages isolated from the acutely inflamed lung. *Am J Physiol Lung Cell Mol Physiol* 282, L330–L339, 2002.

96. Horner, A.A., Van Uden, J.H., Zubeldia, J.M., Broide, D., and Raz, E. DNA-based immunotherapeutics for the treatment of allergic disease. *Immunol Rev* 179, 102–118, 2001.
97. Clark, H. and Reid, K.B. Structural requirements for SP-D function in vitro and in vivo: Therapeutic potential of recombinant SP-D. *Immunobiology* 205, 619–631, 2002.
98. Madan, T., et al. Lung surfactant proteins A and D can inhibit specific IgE binding to the allergens of *Aspergillus fumigatus* and block allergen-induced histamine release from human basophils. *Clin Exp Immunol* 110, 241–249, 1997.
99. Wang, J.Y., Shieh, C.C., You, P.F., Lei, H.Y., and Reid, K.B. Inhibitory effect of pulmonary surfactant proteins A and D on allergen- induced lymphocyte proliferation and histamine release in children with asthma. *Am J Respir Crit Care Med* 158, 510–518, 1998.
100. Borron, P., et al. Surfactant protein A inhibits T cell proliferation via its collagen-like tail and a 210-kDa receptor. *Am J Physiol* 275, L679–L686, 1998.
101. Haczku, A., et al. IL-4 and IL-13 form a negative feedback circuit with surfactant protein-D in the allergic airway response. *J Immunol* 176, 3557–3565, 2006.
102. Atochina, E.N., Beers, M.F., Hawgood, S., Poulain, F., Davis, C., Fusaro, T., and Gow, A.J. Surfactant protein-D, a mediator of innate lung immunity, alters the products of NO metabolism. *Am J Respir Cell Mol Biol* 30, 271–279, 2004.
103. Hickman-Davis, J.M., et al. Lung surfactant and reactive oxygen-nitrogen species: Antimicrobial activity and host-pathogen interactions. *Am J Physiol Lung Cell Mol Physiol* 281, L517–L523 2001.
104. Palaniyar, N., et al. Innate immune collectin surfactant protein D enhances the clearance of DNA by macrophages and minimizes anti-DNA antibody generation. *J Immunol* 174, 7352–7358, 2005.
105. Brinker, K.G., et al. Surfactant protein D enhances bacterial antigen presentation by bone marrow-derived dendritic cells. *Am J Physiol Lung Cell Mol Physiol* 281, L1453–L1463, 2001.
106. Brinker, K.G., Garner, H., and Wright, J.R. Surfactant protein A modulates the differentiation of murine bone marrow-derived dendritic cells. *Am J Physiol Lung Cell Mol Physiol* 284, L232–L241, 2003.
107. Watford, W.T., Wright, J.R., Hester, C.G., Jiang, H., and Frank, M.M. Surfactant protein A regulates complement activation. *J Immunol* 167, 6593–6600, 2001.
108. Fujita, T. Evolution of the lectin-complement pathway and its role in innate immunity. *Nat Rev Immunol* 2, 346–353, 2002.
109. Murakami, S., et al. Surfactant protein A inhibits peptidoglycan-induced tumor necrosis factor-alpha secretion in U937 cells and alveolar macrophages by direct interaction with toll-like receptor 2. *J Biol Chem* 277, 6830–6837, 2002.
110. Guillot, L., et al. Cutting edge: The immunostimulatory activity of the lung surfactant protein-A involves Toll-like receptor 4. *J Immunol* 168, 5989–5992, 2002.
111. Ohya, M., et al. Human pulmonary surfactant protein D binds the extracellular domains of Toll-like receptors 2 and 4 through the carbohydrate recognition domain by a mechanism different from its binding to phosphatidylinositol and lipopolysaccharide. *Biochemistry* 45, 8657–8664, 2006.
112. Holmskov, U., et al. Isolation and characterization of a new member of the scavenger receptor superfamily, glycoprotein-340 (gp-340), as a lung surfactant protein-D binding molecule. *J Biol Chem* 272, 13743–13749, 1997.
113. Holmskov, U., et al. Cloning of gp-340, a putative opsonin receptor for lung surfactant protein D. *Proc Natl Acad Sci USA* 96, 10794–10799, 1999.
114. Lausen, M., et al. Microfibril-associated protein 4 is present in lung washings and binds to the collagen region of lung surfactant protein D. *J Biol Chem* 274, 32234–32240, 1999.
115. Schlosser, A., et al. Microfibril-associated protein 4 binds to surfactant protein A (SP-A) and colocalizes with SP-A in the extracellular matrix of the lung. *Scand J Immunol* 64, 104–116, 2006.
116. Ogden, C.A., et al. C1q and mannose binding lectin engagement of cell surface calreticulin and CD91 initiates macropinocytosis and uptake of apoptotic cells. *J Exp Med* 194, 781–795, 2001.
117. Sano, H., et al. Surfactant proteins A and D bind CD14 by different mechanisms. *J Biol Chem* 275, 22442–22451, 2000.
118. Takahashi, H., Sano, H., Chiba, H., and Kuroki, Y. Pulmonary surfactant proteins A and D: Innate immune functions and biomarkers for lung diseases. *Curr Pharm Des* 12, 589–598, 2006.
119. Holmskov, U.L. Collectins and collectin receptors in innate immunity. *APMIS Suppl* 100, 1–59, 2000.
120. Williams, M.C. Conversion of lamellar body membranes into tubular myelin in alveoli of fetal rat lungs. *J Cell Biol* 72, 260–277, 1977.
121. Savov, J., Wright, J.R., and Young, S.L. Incorporation of biotinylated SP-A into rat lung surfactant layer, type II cells, and clara cells. *Am J Physiol Lung Cell Mol Physiol* 279, L118–L126, 2000.

122. Ikegami, M., et al. Characteristics of surfactant from SP-A-deficient mice. *Am J Physiol* 275, L247–L254, 1998.
123. Ikegami, M., Korfhagen, T.R., Bruno, M.D., Whitsett, J.A., and Jobe, A.H. Surfactant metabolism in surfactant protein A-deficient mice. *Am J Physiol* 272, L479–L485, 1997.
124. Jain, D., et al. SP-A is necessary for increased clearance of alveolar DPPC with hyperventilation or secretagogues. *Am J Physiol Lung Cell Mol Physiol* 284, L759–L765, 2003.
125. Sorensen, G.L., et al. Surfactant protein D of the innate immune defence is inversely associated with human obesity and SP-D deficiency infers increased body weight in mice. *Scand J Immunol* 64, 633–638, 2006.
126. Sorensen, G.L., Husby, S., and Holmskov, U. Surfactant protein A and surfactant protein D variation in pulmonary disease. *Immunobiology* 212, 381–416, 2007.
127. McIntosh, J.C., Swyers, A.H., Fisher, J.H., and Wright, J.R. Surfactant proteins A and D increase in response to intratracheal lipopolysaccharide. *Am J Respir Cell Mol Biol* 15, 509–519, 1996.
128. Sano, H., et al. Pulmonary surfactant protein A modulates the cellular response to smooth and rough lipopolysaccharides by interaction with CD14. *J Immunol* 163, 387–395, 1999.
129. Borron, P., et al. Surfactant-associated protein A inhibits LPS-induced cytokine and nitric oxide production in vivo. *Am J Physiol Lung Cell Mol Physiol* 278, L840–L847, 2000.
130. Chiba, H., et al. Rat mannose-binding protein a binds CD14. *Infect Immun* 69, 1587–1592, 2001.
131. Ikegami, M., Carter, K., Bishop, K., Yadav, A., Masterjohn, E., Brondyk, W., Scheule, R.K., and Whitsett, J.A. Intratracheal recombinant surfactant protein D prevents endotoxin shock in the newborn preterm lamb. *Am J Respir Crit Care Med* 173, 1342–1347, 2006.
132. McNeely, T.B. and Coonrod, J.D. Aggregation and opsonization of type A but not type B *Hemophilus influenzae* by surfactant protein A. *Am J Respir Cell Mol Biol* 11, 114–122, 1994.
133. van de Wetering, J.K., et al. Characteristics of surfactant protein A and D binding to lipoteichoic acid and peptidoglycan, 2 major cell wall components of gram-positive bacteria. *J Infect Dis* 184, 1143–1151, 2001.
134. Gaynor, C.D., McCormack, F.X., Voelker, D.R., McGowan, S.E., and Schlesinger, L.S. Pulmonary surfactant protein A mediates enhanced phagocytosis of *Mycobacterium tuberculosis* by a direct interaction with human macrophages. *J Immunol* 155, 5343–5351, 1995.
135. Ferguson, J.S., Voelker, D.R., Ufnar, J.A., Dawson, A.J., and Schlesinger, L.S. Surfactant protein D inhibition of human macrophage uptake of *Mycobacterium tuberculosis* is independent of bacterial agglutination. *J Immunol* 168, 1309–1314, 2002.
136. Ferguson, J.S., Voelker, D.R., McCormack, F.X., and Schlesinger, L.S. Surfactant protein D binds to *Mycobacterium tuberculosis* bacilli and lipoarabinomannan via carbohydrate-lectin interactions resulting in reduced phagocytosis of the bacteria by macrophages. *J Immunol* 163, 312–321, 1999.
137. Sidobre, S., Nigou, J., Puzo, G., and Riviere, M. Lipoglycans are putative ligands for the human pulmonary surfactant protein A attachment to mycobacteria. Critical role of the lipids for lectin–carbohydrate recognition. *J Biol Chem* 275, 2415–2422, 2000.
138. Beharka, A.A., et al. Pulmonary surfactant protein A up-regulates activity of the mannose receptor, a pattern recognition receptor expressed on human macrophages. *J Immunol* 169, 3565–3573, 2002.
139. Whitsett, J.A. Surfactant proteins in innate host defense of the lung. *Biol Neonate* 88, 175–180, 2005.
140. Zimmerman, P.E., Voelker, D.R., McCormack, F.X., Paulsrud, J.R., and Martin, W.J., 2nd. 120 kDa surface glycoprotein of *Pneumocystis carinii* is a ligand for surfactant protein A. *J Clin Invest* 89, 143–149, 1992.
141. O'Riordan, D.M., et al. Surfactant protein D interacts with *Pneumocystis carinii* and mediates organism adherence to alveolar macrophages. *J Clin Invest* 95, 2699–2710, 1995.
142. Schelenz, S., Malhotra, R., Sim, R.B., Holmskov, U., and Bancroft, G.J. Binding of host collectins to the pathogenic yeast *Cryptococcus neoformans*: Human surfactant protein D acts as an agglutinin for acapsular yeast cells. *Infect Immun* 63, 3360–3366, 1995.
143. Hartshorn, K.L., et al. Mechanisms of anti-influenza activity of surfactant proteins A and D: Comparison with serum collectins. *Am J Physiol* 273, L1156–L1166, 1997.
144. Hartshorn, K., et al. Interactions of recombinant human pulmonary surfactant protein D and SP-D multimers with influenza A. *Am J Physiol* 271, L753–L762, 1996.
145. Benne, C.A., et al. Interactions of surfactant protein A with influenza A viruses: Binding and neutralization. *J Infect Dis* 171, 335–341, 1995.
146. Hartshorn, K.L., et al. Neutrophil deactivation by influenza A viruses: Mechanisms of protection after viral opsonization with collectins and hemagglutination-inhibiting antibodies. *Blood* 87, 3450–3461, 1996.

147. Benne, C.A., Benaissa-Trouw, B., van Strijp, J.A., Kraaijeveld, C.A., and van Iwaarden, J.F. Surfactant protein A, but not surfactant protein D, is an opsonin for influenza A virus phagocytosis by rat alveolar macrophages. *Eur J Immunol* 27, 886–890, 1997.
148. Madan, T., et al. Protective role of lung surfactant protein D in a murine model of invasive pulmonary aspergillosis. *Infect Immun* 69, 2728–2731, 2001.
149. Singh, M., et al. Protective effects of a recombinant fragment of human surfactant protein D in a murine model of pulmonary hypersensitivity induced by dust mite allergens. *Immunol Lett* 86, 299–307, 2003.
150. Wang, J.Y., Kishore, U., Lim, B.L., Strong, P., and Reid, K.B. Interaction of human lung surfactant proteins A and D with mite (*Dermatophagoides pteronyssinus*) allergens. *Clin Exp Immunol* 106, 367–373, 1996.
151. Cheng, G., et al. Surfactant protein A exhibits inhibitory effect on eosinophils IL-8 production. *Biochem Biophys Res Commun* 270, 831–835, 2000.
152. Mishra, A., Weaver, T.E., Beck, D.C., and Rothenberg, M.E. Interleukin-5-mediated allergic airway inflammation inhibits the human surfactant protein C promoter in transgenic mice. *J Biol Chem* 276, 8453–8459, 2001.
153. Kasper, M., et al. Increased surfactant protein D in rat airway goblet and Clara cells during ovalbumin-induced allergic airway inflammation. *Clin Exp Allergy* 32, 1251–1258, 2002.
154. Homer, R.J., et al. Pulmonary type II cell hypertrophy and pulmonary lipoproteinosis are features of chronic IL-13 exposure. *Am J Physiol Lung Cell Mol Physiol* 283, L52–L59, 2002.
155. Cheng, G., et al. Increased levels of surfactant protein A and D in bronchoalveolar lavage fluids in patients with bronchial asthma. *Eur Respir J* 16, 831–835, 2000.
156. Koopmans, J.G., et al. Serum surfactant protein D is elevated in allergic patients. *Clin Exp Allergy* 34, 1827–1833, 2004.
157. Inase, N., et al. A clinical study of hypersensitivity pneumonitis presumably caused by feather duvets. *Ann Allergy Asthma Immunol* 96, 98–104, 2006.
158. Higashi, A., et al. Involvement of eicosanoids and surfactant protein D in extrinsic allergic alveolitis. *Eur Respir J* 26, 1069–1073, 2005.
159. Saikai, T., et al. Hypersensitivity pneumonitis induced by the spore of *Pleurotus eryngii* (Eringi). *Intern Med* 41, 571–573, 2002.
160. Tanaka, H., et al. Mushroom worker's lung caused by spores of *Hypsizigus marmoreus* (Bunashimeji): Elevated serum surfactant protein D levels. *Chest* 118, 1506–1509, 2000.
161. Tsushima, K., et al. Hypersensitivity pneumonitis due to Bunashimeji mushrooms in the mushroom industry. *Int Arch Allergy Immunol* 137, 241–248, 2005.
162. Hull, J., South, M., Phelan, P., and Grimwood, K. Surfactant composition in infants and young children with cystic fibrosis. *Am J Respir Crit Care Med* 156, 161–165, 1997.
163. Postle, A.D., et al. Deficient hydrophilic lung surfactant proteins A and D with normal surfactant phospholipid molecular species in cystic fibrosis. *Am J Respir Cell Mol Biol* 20, 90–98, 1999.
164. Griese, M., Birrer, P., and Demirsoy, A. Pulmonary surfactant in cystic fibrosis. *Eur Respir J* 10, 1983–1988, 1997.
165. Meyer, K.C., et al. Function and composition of pulmonary surfactant and surfactant-derived fatty acid profiles are altered in young adults with cystic fibrosis. *Chest* 118, 164–174, 2000.
166. Griese, M., et al. Pulmonary surfactant, lung function, and endobronchial inflammation in cystic fibrosis. *Am J Respir Crit Care Med* 170, 1000–1005, 2004.
167. Noah, T.L., et al. Bronchoalveolar lavage fluid surfactant protein-A and surfactant protein-D are inversely related to inflammation in early cystic fibrosis. *Am J Respir Crit Care Med* 168, 685–691, 2003.
168. Rubio, F., Cooley, J., Accurso, F.J., and Remold-O'Donnell, E. Linkage of neutrophil serine proteases and decreased surfactant protein-A (SP-A) levels in inflammatory lung disease. *Thorax* 59, 318–323, 2004.
169. von Bredow, C., Birrer, P., and Griese, M. Surfactant protein A and other bronchoalveolar lavage fluid proteins are altered in cystic fibrosis. *Eur Respir J* 17, 716–722, 2001.
170. von Bredow, C., Wiesener, A., and Griese, M. Proteolysis of surfactant protein D by cystic fibrosis relevant proteases. *Lung* 181, 79–88, 2003.
171. Griese, M., von Bredow, C., and Birrer, P. Reduced proteolysis of surfactant protein A and changes of the bronchoalveolar lavage fluid proteome by inhaled alpha 1-protease inhibitor in cystic fibrosis. *Electrophoresis* 22, 165–171, 2001.
172. Mariencheck, W.I., Alcorn, J.F., Palmer, S.M., and Wright, J.R. *Pseudomonas aeruginosa* elastase degrades surfactant proteins A and D. *Am J Respir Cell Mol Biol* 28, 528–537, 2003.

173. Malloy, J.L., Veldhuizen, R.A., Thibodeaux, B.A., O'Callaghan, R.J., and Wright, J.R. *Pseudomonas aeruginosa* protease IV degrades surfactant proteins and inhibits surfactant host defense and biophysical functions. *Am J Physiol Lung Cell Mol Physiol* 288, L409–L418, 2005.
174. Khubchandani, K.R., Oberley, R.E., and Snyder, J.M. Effects of surfactant protein A and NaCl concentration on the uptake of *Pseudomonas aeruginosa* by THP-1 cells. *Am J Respir Cell Mol Biol* 25, 699–706, 2001.
175. Mariencheck, W.I., Savov, J., Dong, Q., Tino, M.J., and Wright, J.R. Surfactant protein A enhances alveolar macrophage phagocytosis of a live, mucoid strain of *P. aeruginosa*. *Am J Physiol* 277, L777–L786, 1999.
176. Manz-Keinke, H., Plattner, H., and Schlepper-Schafer, J. Lung surfactant protein A (SP-A) enhances serum-independent phagocytosis of bacteria by alveolar macrophages. *Eur J Cell Biol* 57, 95–100, 1992.
177. Tino, M.J. and Wright, J.R. Surfactant protein A stimulates phagocytosis of specific pulmonary pathogens by alveolar macrophages. *Am J Physiol* 270, L677–L688, 1996.
178. Restrepo, C.I., Dong, Q., Savov, J., Mariencheck, W.I., and Wright, J.R. Surfactant protein D stimulates phagocytosis of *Pseudomonas aeruginosa* by alveolar macrophages. *Am J Respir Cell Mol Biol* 21, 576–585, 1999.
179. Bufler, P., et al. Surfactant protein A and D differently regulate the immune response to nonmucoid Pseudomonas aeruginosa and its lipopolysaccharide. *Am J Respir Cell Mol Biol* 28, 249–256, 2003.
180. Alexis, N.E., Muhlebach, M.S., Peden, D.B., and Noah, T.L. Attenuation of host defense function of lung phagocytes in young cystic fibrosis patients. *J Cyst Fibros* 5, 17–25, 2006.
181. Krane, M. and Griese, M. Surfactant protein D in serum from patients with allergic bronchopulmonary aspergillosis. *Eur Respir J* 22, 592–595, 2003.
182. Clark, H. and Reid, K. The potential of recombinant surfactant protein D therapy to reduce inflammation in neonatal chronic lung disease, cystic fibrosis, and emphysema. *Arch Dis Child* 88, 981–984, 2003.

24 C-Type Lectin Receptors on Dendritic Cells

Yvette van Kooyk

CONTENTS

24.1 INTRODUCTION: THE IMMUNE RESPONSE

One of the most fascinating tasks of the immune system is to detect the presence of infectious agents and remove or inactivate the invader without destroying the host self-tissues. This process needs to be regulated with enormous precision given the enormous molecular variety of pathogens and their high replication and mutation rates. The first immune cells to come in contact with invading pathogens are dendritic cells (DCs), the sentinels of the immune system. DCs recognize microbial structures and through cell–cell interactions with T cells, an immune response is elicited. Sometimes the immune system fails to induce a proper response, resulting in immunodeficiency disorders, autoimmune disease, allergic reactions, and cancer.

24.2 DENDRITIC CELLS, SENTINELS OF THE IMMUNE SYSTEM

DCs are professional antigen-presenting cells (APC), which are strategically positioned at the boundaries between the inner and the outside world [1]. The development of DCs is considered to occur in distinct stages. Hematopoietic pluripotent stem cells, under as yet unknown influences, constantly generate DC progenitors in the bone marrow, which give rise to circulating precursors in the blood [2]. These DC precursors are scattered throughout the body in virtually all nonlymphoid tissues, such as skin and mucosal tissue via the blood circulation and function as a continuous surveillance patrol for incoming foreign antigens. DCs positioned in the peripheral tissues are called immature DCs. Immature DCs have the unique capacity to efficiently capture antigens and process these antigens endogenously for presentation in cell-surface expressed major histocompatibility

complex (MHC) class I and II molecules. To become licensed to activate naïve T helper cells, DCs must undergo a maturation process during the migration from the peripheral tissues to the lymph nodes. In the lymph node, the captured antigen is presented to the naïve T cells. Maturation of DCs is characterized by a decreased Ag-uptake capacity and an increased cell surface expression of major MHC and costimulatory molecules that are involved in T cell activation [1]. In addition, rearrangement of cytoskeleton, adhesion molecules, and cytokine receptors upon maturation allow DCs to migrate from peripheral tissues to secondary lymphoid organs [3].

DCs play a pivotal role in the immune response by providing signals that direct naïve T helper (Th) cells to proliferate and differentiate into Th1, Th2, or T regulatory cells. Three signals are delivered by an APC that are thought to determine the fate of naïve T cells (Figure 24.1) [4]. Recently it has been shown that the signals delivered by the APC to the T cell determine its polarization into a specific type of effector cell, referred to as signal 3 (Figure 24.1). T helper 1 cells activate cytotoxic T cells and stimulate the microbicidal properties of macrophages (cell-mediated immune response). T helper 2 cells initiate the humoral immune response by activating naïve antigen-specific B cells to produce antibodies. The polarizing signals are mediated by various soluble and membrane-bound factors, such as interleukin-12 (IL-12) or CC-chemokine ligand 2 (CCL2) [5].

For pathogen recognition, DCs express specialized receptors so-called pattern recognition receptors (PRRs) expressed on the cell surface of DCs. Two main classes of PRRs expressed by DCs are Toll-like receptors (TLRs) and C-type lectin receptors (CLRs). Depending on their tissue localization and differentiation state, DCs express different sets of CLR and TLR to allow a specialized response to specific microbes. In this way, DCs are professional APCs that bridge the innate and adaptive immunity that brings pathogen information through DC to instruct specific T cells.

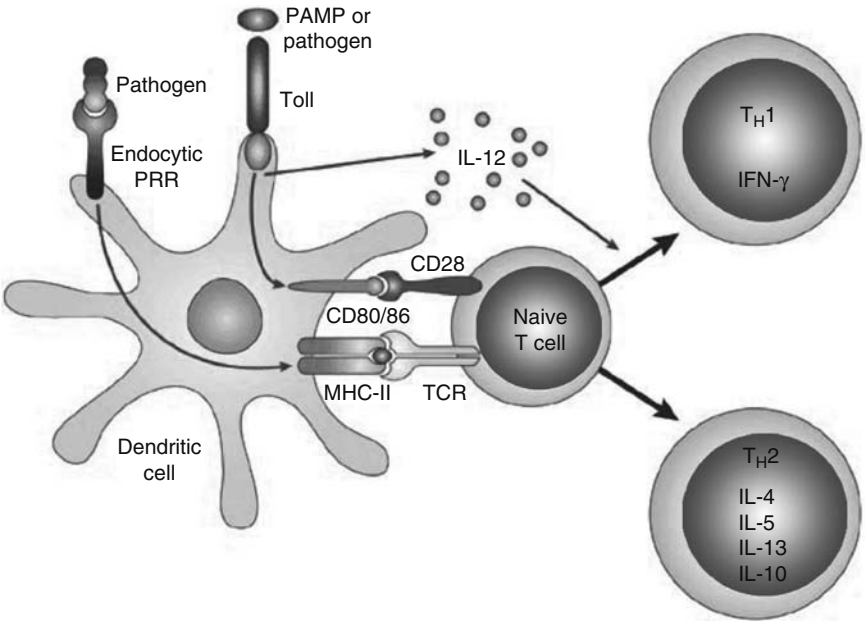


FIGURE 24.1 (See color insert following blank page 170. Also see CD for color figure.) T cell polarization requires three dendritic cell-derived signals. Signal 1 is an antigen-specific signal that is mediated through T cell receptor (TCR) and MHC class II-associated peptides. Signal 2 is the costimulatory signal, mediated by triggering of CD28 by CD80 and CD86. Signal 3 is the polarizing signal mediated by soluble or membrane bound factors like cytokines such as IL-12 (Th1) or chemokines such as CCL2 (Th2).

24.3 TOLL-LIKE RECEPTORS

Human TLRs are type 1 transmembrane glycoproteins characterized by an extracellular domain that contains leucine-rich repeat units and cytoplasmic domain of IL-1 receptor, designated as the Toll/IL-1 receptor (TIR) domain. At least 11 human TLRs have been identified [6], which are expressed as homodimers or heterodimers (TLR-2 with TLR-1 and TLR-6 with TLR2, respectively). TLRs are broadly expressed on cells of the immune system and are the best-studied immune sensors of invading pathogens [7,8].

Their ligands, bacterial products such as nucleic acids, peptidoglycans, lipo-polysaccharide (LPS), flagelin, ds-RNA, or CpG DNA, require internalization to the endosome before signaling is possible [9]. Activation of most TLRs induces cellular responses associated with acute and chronic inflammation. When TLR ligands interact with their specific TLRs, intracellular adaptor proteins transduce signals that lead to enhanced expression of genes encoding cytokines and other inflammatory mediators.

All members of the TLR superfamily signal in a rather similar manner owing to the presence of the TIR domain, which activates common signaling pathways, most notably those leading to activation of the transcription factor nuclear factor κB (NF-κB) and stress-activated protein kinases. These pathways are defined as the MyD-88-dependent and -independent pathway [10]. TLR ligands form a combinatorial code by which DCs discriminate pathogens and this implicates a new strategy for priming an immune response.

24.4 C-TYPE LECTINS

The term "C-type lectin receptor" (CLR) designates carbohydrate-binding molecules that bind their ligands in a Ca^{2+}-dependent manner through a carbohydrate recognition domain (CRD). The CRD contains a well-preserved consensus sequence of approximately 115–130 amino acids, which are involved in calcium binding and sugar specificity. One calcium ion is essential for proper positioning of the binding site. The second calcium ion is situated in the primary binding site and coordinates ligand binding [11]. In addition, the sugar binding to a lectin depends on subtle differences in the arrangements of carbohydrate residues and their branching. Furthermore, the secondary binding site within the CRD was shown to play an important role in sugar binding by mutation studies [12,13].

CLRs have diverse functions [14]. They recognize endogenous ligands and thereby mediate cell–cell interactions during immune responses. CLRs bind to soluble self-antigens, leading to immune tolerance and maintenance of endogenous glycoprotein homeostasis. CLRs function also as pathogen recognition receptors through the recognition of pathogen-associated molecular pattern (PAMPs). Furthermore, cooperation between TLRs and CLRs has been demonstrated and it seems that appropriate immune responses rely on the interaction of many different antigen sensing and sampling mechanisms [15,16].

APCs express CLRs that are predominantly type II transmembrane receptors with a single CRD, such as DC-specific intracellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin (DC-SIGN), MGL, dectin-1, blood DC antigen 2 (BDCA2), and langerin (Figure 24.2). In addition, there are also type I transmembrane receptors, such as the mannose receptor and DEC-205, which have multiple CRDs, although not all of these act as functional CRDs (Figure 24.2) [14]. True Ca^{2+}-dependent CLRs fall into two broad categories, those recognizing mannose- or fucose- or *N*-acetylglucosamine-type ligands and those recognizing galactose- or *N*-acetylgalactosamine-type ligands, which can be defined by the presence of a distinctive triplet of amino acids within the CRD: EPN for mannose-type receptors and QPD for galactose-type receptors [11].

The capacity of CLRs to detect microorganisms depends on the degree of oligomerization of the lectin receptor, as well as on the PAMP present on the microbial surface. By creating multimers, in which several CRDs point toward a common direction, the binding valency increases and allows interactions with the dense carbohydrate expression on microbial surfaces. To increase their binding

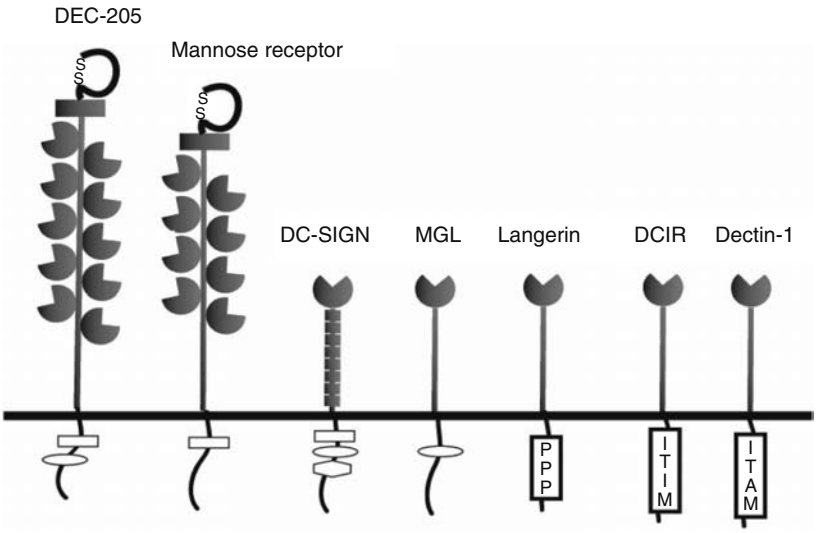


FIGURE 24.2 (See CD for color figure.) Two types of C-type lectins are present on dendritic cells. Type I transmembrane receptors like DEC-205 and MR contain 8–10 carbohydrate recognition domains (CRDs) at their extracellular amino-terminus. Type II transmembrane receptors contain only one CRD at their carboxy-terminal extracellular domain. Cytoplasmic tails can contain ITIM or ITAM motifs, proline rich domains, or tyrosine-based motifs, triacidic cluster or a dileucine motif.

strength to PAMPs, transmembrane CLRs have developed several strategies. Proteomic analysis of membrane-purified DC-SIGN complexes showed that DC-SIGN exists as tetramers on the surface of immature monocyte-derived DCs and that this assembly is required for high binding of glycoproteins such as HIV-1 gp120 [17–19]. Biochemical and hydrodynamic studies on truncated DC-SIGN demonstrated that the neck portion of each molecule adjacent to the CRD is sufficient to mediate the formation of dimers, whereas the neck regions near the amino-terminal are required to stabilize tetramers [19]. The CRDs are flexibly linked to the neck regions, which project CRDs from the cell surface and enable DC-SIGN to bind to various glycans on microbial surfaces [19]. Higher levels of organization were shown for DC-SIGN due to clustering on the cell surface of DCs [18]. The organization of DC-SIGN in so-called "microdomains" on the plasma membrane is important for the binding and internalization of virus particles, suggesting that these multimolecular assemblies of DC-SIGN act as a docking site for pathogens such as HIV-1 to invade the host.

24.5 DC-SIGN

DC-SIGN (CD209) is a type II membrane C-type lectin with a short amino terminal cytoplasmic tail and a single carboxyl terminal CRD [20]. Reports on the carbohydrate specificity and function within the immune system of lectins are emerging the last few years. By using glycan microarray technology, the carbohydrate specificity of lectins can easily be identified. Glycan microarray studies confirmed that DC-SIGN recognizes two classes of glycans, mannose-containing glycans terminated with manα1–2 residues, and various fucosylated glycans with the Fucα1-3GlcNAc- and Fucα1-4GlcNAc-containing glycans found as terminal sequences of N- and O-linked glycans [21,22].

The cytoplasmic tail of DC-SIGN contains three internalization motives, namely a tyrosine-based motif of the sequence YKSL, a dileucine motif, and a triacidic cluster EEE. Engering et al. showed that the tyrosine motif and the dileucine motif are involved in the ligand-induced internalization of DC-SIGN [23]. As a cell adhesion receptor DC-SIGN mediates interactions between DCs and resting T cells through binding to ICAM-3 [20] and supports transendothelial migration through

interaction with ICAM-2 [24], however, the endogenous carbohydrate ligand still remained inconclusive. Recently Lewis antigen has been identified as ligands on ICAM-3 [25] and carcino-embryonic antigen (CEA) [26].

Many studies have described the function of DC-SIGN as a pathogen receptor. DC-SIGN binds HIV-1 gp120 and facilitates the transport of HIV from mucosal sites to draining lymph nodes where infection of T-lymphocytes occurs [27]. The primary structure of the CRD contains conserved residues consistent with classical mannose-specific CRDs. The CRD of DC-SIGN recognizes both internal branched mannose residues as well as terminal dimannoses [22,28,29]. Since the identification of DC-SIGN as HIV-1 receptor on DCs, interactions with other pathogens have been reported including *Mycobacterium tuberculosis* [30], *Candida albicans* [31], *Helicobacter pylori* [32], *Schistosoma mansoni* [33], and *Neisseria meningitidis* [34], but also viruses like hepatitis C virus (HCV) [35–37], Ebola [38,39], cytomegalovirus (CMV) [40], and dengue [41]. These interactions are due to the mannose- or fucose-type glycans that these pathogens express on their surface or that are present in their secretion products.

The function of C-type lectins such as DC-SIGN is to deliver antigens to intracellular compartments resembling late endosomes/lysosomes for degradation and these antigens are subsequently presented to T-cells *in vitro* [23,42]. While data suggest a role for DC-SIGN in host defense, evidence is now accumulating that some pathogens may preferentially interact with DC-SIGN as part of a pathogenic strategy and exploit the CLR-mediated uptake to their own benefit, namely to evade the host immune defenses.

24.6 L-SIGN

Liver/lymph node-specific ICAM-3-grabbing nonintegrin (L-SIGN/CD209L/DC-SIGN-R) is a human homolog of DC-SIGN [43,44]. L-SIGN shares 77% amino acid sequence identity with DC-SIGN and has functional similarity by recognizing ICAM-2, ICAM-3, HIV-1 gp120, and Ebola. Within the cytoplasmic tail, the most differences between DC-SIGN and L-SIGN are found. Like DC-SIGN, L-SIGN contains a dileucine motif and a triacidic cluster EED. L-SIGN lacks the tyrosine-based motif [43]. In addition, elucidation of the crystal structures of the CRDs of both DC-SIGN and L-SIGN, in combination with binding studies, revealed that both receptors recognize high-mannose oligosaccharides [28,29]. Whereas for DC-SIGN increasing evidence indicates that its major ligands are α3/α4-fucosylated glycans, no data so far indicated a role for L-SIGN in the binding of fucosylated oligosaccharides. However, recently we have demonstrated that L-SIGN like DC-SIGN can bind to several Lewis antigens, with the exception of Lewis x [13]. L-SIGN is not expressed by DCs, but is expressed by endothelial cells present in lymph node sinuses and liver sinusoidal endothelial cells (LSECs). LSECs function as a liver-resident APC population on which L-SIGN could play a role in antigen clearance as well as LSEC-leukocyte adhesion [45,46]. Recent reports indicate that L-SIGN represents a liver-specific receptor for HCV and L-SIGN might play an important role in HCV infection and immunity [35,42]. L-SIGN also binds to *S. mansoni* soluble egg antigens (SEA) [33].

24.7 MGL

The macrophage galactose binding lectin (MGL) (CD301), also called DC-asialoglycoprotein receptor (DC-ASGP-R) or human macrophage lectin (HML) is the only reported galactose-type C-type lectin on human DCs. MGL consists of one CRD domain and contains an YENF internalization motif for endocytosis via clathrin-coated pits. MGL is expressed as a trimer on human DCs and macrophages at an intermediate stage of differentiation from monocytes [47]. MGL-positive immature DC and macrophages are most abundantly expressed in the dermis and the skin. MGL expression is also found on DC-like cells in the T cell areas of human lymph node and tonsil [48]. By using glycan microarray profiling, the carbohydrate specificity of MGL was

characterized [49]. MGL has an exclusive specificity for terminal α- and β-linked GalNAc residues that naturally occur as parts of glycoproteins and glycosphingolipids. Up till now, terminal GalNAc moieties expressed by tumor antigens as well as the human parasite *S. mansoni* and a subset of gangliosides [49], are identified as ligands for MGL. Recently the interaction of MGL and CD45 on effector T cells has been described [48].

24.8 MANNOSE RECEPTOR

The mannose receptor (MR, CD206) is a type I transmembrane domain that consists of eight CRDs, a fibronectin type II domain and a cysteine rich N-terminal domain. The mannose receptor is expressed on macrophage subsets, *in vitro* cultured DCs, and certain DC populations from inflamed skin as well as lymphatic and hepatic sinusoidal endothelium [50,51]. The MR has been shown to recycle constitutively while releasing its cargo [52]. It has been implicated in homeostatic processes, such as clearance of endogenous products, cell adhesion, as well as pathogen recognition and antigen presentation [53,54].

Recombinant domain deletion studies have demonstrated that CRD 4 and to a lesser extent CRD 5 are the only domains showing true affinity for *N*-acetylglucosamine, mannose, and fucose-terminating oligosaccharides [54]. Although all eight CRDs contain conserved residues responsible for formation of the hydrophobic fold, only CRD 4 and 5 retain residues needed for Ca^{2+} coordination and consequent carbohydrate binding. The role of the MR as a PRR has been well studied and it has been reported to bind to a wide variety of pathogens like HIV-1 [55], *M. tuberculosis* [56], *C. albicans*, *Trypanosoma cruzi*, *Pneumocystis carinii*, and *Streptococcus pneumoniae* [57]. It preferentially targets mannose-containing structures and differs from DC-SIGN, as it does not interact with Lewis x, despite having affinity for fucose [58]. This can be explained by the differences in arrangement of CRDs, as DC-SIGN has one CRD and needs tetramerization for multivalent binding of its ligands, whereas the MR has eight CRDs. The cysteine rich N-terminal was recently shown to be involved in the binding of sulfated oligosaccharides, while the fibronectin type II domain mediates the activity of the MR to bind collagen [59].

24.9 LANGERIN

Langerhans cells (LC) are present as a network of APCs in the epidermal layers of skin or at the mucosal tissues. Similar as DC, they have a strong potency to induce T cell responses, however, they may differentially induce immunoresponse than DC. LC are morphological and phenotypically different from DC [60]. They are highly extended cells with specific cytoplasmic organelles, the Birbeck granules, which are involved in antigen degradation and processing. LC express different cell surface markers such as CD1a and the C-type lectin Langerin [61,62]. In particular, Langerin expression has been shown to regulate Birbeck granule formation in LC. Different from other C-type lectins, Langerin has in its cytoplasmic tail a proline-rich motif, which has been suggested to affect signaling and internalization properties of Langerin, Langerin has also been shown to bind mannose structures, fucose structures, and GlcNAc structures [63]. Although the carbohydrate specificity of Langerin shows some overlap with that of DC-SIGN, it is not identical to DC-SIGN. For instance, Langerin does not interact with Lewis x, whereas DC-SGIN does. Langerin has been shown to function as a pathogen recognition receptors having affinity for HIV-1 and *M. leprae*. Very recently, it has been shown that Langerin differently handles pathogens compared to DC-SIGN. LC have a strong potency to efficiently degrade pathogens intracellularly while DC-SIGN has shown incomplete degradation of pathogens allowing pathogens such as HIV-1 to escape antigen processing, favoring effective transmission of virions to T cells. Also other pathogens that target DC-SIGN have found ways to modulate T cell responses by modulating cytokine induction or maturation of DC (see below) [27,64]. LC have also been shown to express different sets of TLR receptors indicating that DC and LC express a different repertoire of TLR and CLR, that subsequently may induce different immune responses [65].

24.10 DC EXPRESSED C-TYPE LECTINS AND THEIR MODULATORY SIGNALING PROPERTIES; DECTIN-1, DCIR, DCAR, BDCA2 DC-SIGN, AND DCAL1

Pathogens have evolved strategies to evade the hosts' immune response by subverting the function of PRRs. Certain PAMPs are recognized by TLRs and CLRs simultaneously and although the signal transduction route upon TLR activation is well documented, CLR may also signal and thus influence DC functions. C-type lectin stimulation can either enhance or inhibit TLR signaling. Although C-type lectins do not seem to directly stimulate the immune system, there are new indications that they certainly have signaling properties that affect the balance between immunity and tolerance by influencing TLR signaling [16,66].

Some C-type lectin-like molecules display Ca^{2+}-indepentdent carbohydrate recognition such as Dectin-1. Dectin-1 is expressed on DC but also LCs, macrophages, and some B and T cell subsets. Dectin-1 recognizes β-glucans in a cation-independent manner, and thus has been reported to be a receptor for yeast. Dectin-1 has gained a lot of interest as it is involved in several innate immune responses against fungal infections and its involvement in signaling. Dectin-1 contains an immuno receptor tyrosine-based activatory motif (ITAM) motif and engagement triggers the release of proinflammatory cytokines and chemokines. Combined triggering of Dectin-1 and TLR2 by fungal pathogens leads to the synergized production of TNFα, IL-12, and reactive oxygen species (ROS). Clearly Dectin-1 is the dendritic C-type lectin whose intracellular signaling has been best characterized. β-Glycan interaction with Dectin-1 leads to a Syk- and CARD9-dependent IL-10 release [67,68]. The dectin related ITAM-Syk-Card9 route ITAM-Syk-Card9 route has been suggested to play a general signaling route in DC C-type lectins, however, this does not hold true for the ITAM-1 like domain in the C-type lectin DC-SIGN that has been demonstrated to regulate Syk-independent signaling processes in DC.

The C-type lectin DCIR and DCAR share substantial sequence homology in their extracellular domain. However, whereas DCAR associates with the ITAM-bearing FcR γ chain, DCIR contains and immuno receptor tyrosine-based inhibitory motif (ITIM) and recruits the phosphatases SHP-1 and SHP-2. Cross-linking of DCIR has been shown to result in inhibitory signals. Other C-type lectin that can have inhibitory functions are BDCA-2 and DCAL-1, of which cross-linking have shown to diminish release of interferon α/β by plasmacytoid DC [69].

M. tuberculosis is a potent inducer of T helper 1 (Th1)-polarized immune response, and mycobacterial components have often been shown to stimulate expression of costimulatory molecules and IL-12 production in DCs through TLR2 and TLR4 triggering [15]. *M. tuberculosis* strongly binds to DC-SIGN through their mannose-capped cell wall component lipoarabinomannan (ManLAM). ManLAM targets CLRs to alter the immune response though cross-talk between TLRs and CLRs. In particular, binding of the mycobacterial component ManLAM to immature DCs inhibited LPS-induced maturation and subsequent LPS-mediated IL-12 induction. Inhibitory anti-DC-SIGN antibodies restore maturation of DCs in the presence of ManLAM. The inhibition of DC maturation and the induction of IL-10 may contribute to the virulence of mycobacteria [15]. Thus pathogen recognition by DC-SIGN may modulate DC-induced immune responses, shifting the balance from immune activation toward impairment of the immune response, which would be beneficial to pathogen survival. How DC-SIGN signals are propagated within DCs is not clear. Recently, Caparros and coworkers revealed the first clues on signaling transduction pathways upon DC-SIGN ligation. They showed that DC-SIGN induces the phosphorylation of ERK and Akt, without the concomitant p38MAPK activation [70]. More recently a new signaling route upon DC-SIGN triggering through ManLAM structures has been identified in which raf activation plays an essential role in the signaling cascade [71].

24.11 CONCLUSION AND FUTURE DIRECTIONS

The studies described above demonstrate that the collaborative recognition of distinct microbial components or a pathogen by different classes of innate immune receptors (TLRs and CLRs) is

crucial for orchestrating inflammatory or inhibitory responses. The balance between CLR binding and TLR stimulation may fine-tune regulatory mechanisms to allow appropriate immune responses. To maintain homeostatic control and silence immune activation, pathogens have evolved their survival strategy within the host environment. CLRs can induce unique intracellular signals that modulate DC function although the precise signaling pathways need further investigation. Signals derived from CLRs and TLRs will determine whether pathogens escape immunity or are eliminated.

REFERENCES

1. Banchereau, J. and Steinman, R. M. 1998. Dendritic cells and the control of immunity. *Nature* 392:245–252.
2. Romani, N., Gruner, S., Brang, D., Kampgen, E., Lenz, A., Trockenbacher, B., Konwalinka, G., Fritsch, P. O., Steinman, R. M., and Schuler, G. 1994. Proliferating dendritic cell progenitors in human blood. *J Exp Med* 180:83–93.
3. Sallusto, F. and Lanzavecchia, A. 1999. Mobilizing dendritic cells for tolerance, priming, and chronic inflammation. *J Exp Med* 189:611–614.
4. Kapsenberg, M. L. 2003. Dendritic-cell control of pathogen-driven T-cell polarization. *Nat Rev Immunol* 3:984–993.
5. de Jong, E. C., Vieira, P. L., Kalinski, P., Schuitemaker, J. H., Tanaka, Y., Wierenga, E. A., Yazdanbakhsh, M., and Kapsenberg, M. L. 2002. Microbial compounds selectively induce Th1 cell-promoting or Th2 cell-promoting dendritic cells in vitro with diverse Th cell-polarizing signals. *J Immunol* 168:1704–1709.
6. Janeway, C. A. and Medzhitov, R. 2002. Innate immune recognition. *Annu Rev Immunol* 20:197–216.
7. Medzhitov, R. 2001. Toll-like receptors and innate immunity. *Nat Rev Immunol* 1:135–145.
8. Blander, J. M. and Medzhitov, R. 2006. Toll-dependent selection of microbial antigens for presentation by dendritic cells. *Nature* 440:808–812.
9. Akira, S., Takeda, K., and Kaisho, T. 2001. Toll-like receptors: Critical proteins linking innate and acquired immunity. *Nat Immunol* 2:675–680.
10. Beutler, B. 2004. Inferences, questions and possibilities in toll-like receptor signalling. *Nature* 430:257–263.
11. Drickamer, K. 1999. C-type lectin-like domains. *Curr Opin Struct Biol* 9:585–590.
12. Geijtenbeek, T. B., van Duijnhoven, G. C., van Vliet, S. J., Krieger, E., Vriend, G., Figdor, C. G., and van Kooyk, Y. 2002. Identification of different binding sites in the dendritic cell-specific receptor DC-SIGN for intercellular adhesion molecule 3 and HIV-1. *J Biol Chem* 277:11314–11320.
13. van Liempt, E., Imberty, A., Bank, C. M., van Vliet, S. J., van Kooyk, Y., Geijtenbeek, T. B., and Van, D., I 2004. Molecular basis of the differences in binding properties of the highly related C-type lectins DC-SIGN and L-SIGN to Lewis X trisaccharide and *Schistosoma mansoni* egg antigens. *J Biol Chem* 279:33161–33167.
14. Figdor, C. G., van Kooyk, Y., and Adema, G. J. 2002. C-type lectin receptors on dendritic cells and Langerhans cells. *Nat Rev Immunol* 2:77–84.
15. Geijtenbeek, T. B., van Vliet, S. J., Koppel, E. A., Sanchez-Hernandez, M., Vandenbroucke-Grauls, C. M., Appelmelk, B., and van Kooyk, Y. 2003. Mycobacteria target DC-SIGN to suppress dendritic cell function. *J Exp Med* 197:7–17.
16. Geijtenbeek, T. B., van Vliet, S. J., Engering, A., 't hart, B. A., and van Kooyk, Y. 2004. Self- and nonself-recognition by c-type lectins on dendritic cells. *Annu Rev Immunol* 22:33–54.
17. Bernhard, O. K., Lai, J., Wilkinson, J., Sheil, M. M., and Cunningham, A. L. 2004. Proteomic analysis of DC-SIGN on dendritic cells detects tetramers required for ligand binding but no association with CD4. *J Biol Chem* 279:51828–51835.
18. Cambi, A., de Lange, F., van Maarseveen, N. M., Nijhuis, M., Joosten, B., van Dijk, E. M., de Bakker, B. I., Fransen, J. A., Bovee-Geurts, P. H., van Leeuwen, F. N., van Hulst, N. F., and Figdor, C. G. 2004. Microdomains of the C-type lectin DC-SIGN are portals for virus entry into dendritic cells. *J Cell Biol* 164:145–155.
19. Feinberg, H., Guo, Y., Mitchell, D. A., Drickamer, K., and Weis, W. I. 2005. Extended neck regions stabilize tetramers of the receptors DC-SIGN and DC-SIGNR. *J Biol Chem* 280:1327–1335.
20. Geijtenbeek, T. B. H., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Adema, G. J., van Kooyk, Y., and Figdor, C. G. 2000. Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell* 100:575–585.

21. Guo, Y., Feinberg, H., Conroy, E., Mitchell, D. A., Alvarez, R., Blixt, O., Taylor, M. E., Weis, W. I., and Drickamer, K. 2004. Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol* 11:591–598.
22. Appelmelk, B. J., Van, D., I, van Vliet, S. J., Vandenbroucke-Grauls, C. M., Geijtenbeek, T. B., and van Kooyk, Y. 2003. Cutting edge: Carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells. *J Immunol* 170:1635–1639.
23. Engering, A., Geijtenbeek, T. B., van Vliet, S. J., Wijers, M., van Liempt, E., Demaurex, N., Lanzavecchia, A., Fransen, J., Figdor, C. G., Piguet, V., and van Kooyk, Y. 2002. The dendritic cell-specific adhesion receptor DC-SIGN internalizes antigen for presentation to T cells. *J Immunol* 168:2118–2126.
24. Geijtenbeek, T. B., Krooshoop, D. J., Bleijs, D. A., van Vliet, S. J., van Duijnhoven, G. C., Grabovsky, V., Alon, R., Figdor, C. G., and van Kooyk, Y. 2000. DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. *Nat Immunol* 1:353–357.
25. Bogoevska, V., Nollau, P., Lucka, L., Grunow, D., Klampe, B., Uotila, L. M., Samsen, A., Gahmberg, C. G., and Wagener, C. 2007. DC-SIGN binds ICAM-3 isolated from peripheral human leukocytes through Lewis x residues. *Glycobiology* 17:324–333.
26. van Gisbergen, K. P., Aarnoudse, C. A., Meijer, G. A., Geijtenbeek, T. B., and van Kooyk, Y. 2005. Dendritic cells recognize tumor-specific glycosylation of carcinoembryonic antigen on colorectal cancer cells through dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin. *Cancer Res* 65:5935–5944.
27. Geijtenbeek, T. B. H., Kwon, D. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C. F., Middel, J., Cornelissen, I. L., Nottet, H. S., KewalRamani, V. N., Littman, D. R., Figdor, C. G., and van Kooyk, Y. 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* 100:587–597.
28. Feinberg, H., Mitchell, D. A., Drickamer, K., and Weis, W. I. 2001. Structural basis for selective recognition of oligosaccharides by DC- SIGN and DC-SIGNR. *Science* 294:2163–2166.
29. Mitchell, D. A., Fadden, A. J., and Drickamer, K. 2001. A novel mechanism of carbohydrate recognition by the C-type lectins DC-SIGN and DC-SIGNR. Subunit organization and binding to multivalent ligands. *J Biol Chem* 276:28939–28945.
30. Maeda, N., Nigou, J., Herrmann, J. L., Jackson, M., Amara, A., Lagrange, P. H., Puzo, G., Gicquel, B., and Neyrolles, O. 2003. The cell surface receptor DC-SIGN discriminates between *Mycobacterium* species through selective recognition of the mannose caps on lipoarabinomannan. *J Biol Chem* 278:5513–5516.
31. Cambi, A., Gijzen, K., de Vries, J. M., Torensma, R., Joosten, B., Adema, G. J., Netea, M. G., Kullberg, B. J., Romani, L., and Figdor, C. G. 2003. The C-type lectin DC-SIGN (CD209) is an antigen-uptake receptor for *Candida albicans* on dendritic cells. *Eur J Immunol* 33:532–538.
32. Bergman, M. P., Engering, A., Smits, H. H., van Vliet, S. J., van Bodegraven, A. A., Wirth, H. P., Kapsenberg, M. L., Vandenbroucke-Grauls, C. M., van Kooyk, Y., and Appelmelk, B. J. 2004. *Helicobacter pylori* modulates the T helper cell 1/T helper cell 2 balance through phase-variable interaction between lipopolysaccharide and DC-SIGN. *J Exp Med* 200:979–990.
33. Van Die, I., van Vliet, S. J., Nyame, A. K., Cummings, R. D., Bank, C. M., Appelmelk, B., Geijtenbeek, T. B., and van Kooyk, Y. 2003. The dendritic cell-specific C-type lectin DC-SIGN is a receptor for Schistosoma mansoni egg antigens and recognizes the glycan antigen Lewis x. *Glycobiology* 13:471–478.
34. Steeghs, L., van Vliet, S. J., Uronen-Hansson, H., van Mourik, A., Engering, A., Sanchez-Hernandez, M., Klein, N., Callard, R., van Putten, J. P., van der, L. P., van Kooyk, Y., and van de Winkel, J. G. 2006. *Neisseria meningitidis* expressing lgtB lipopolysaccharide targets DC-SIGN and modulates dendritic cell function. *Cell Microbiol* 8:316–325.
35. Gardner, J. P., Durso, R. J., Arrigale, R. R., Donovan, G. P., Maddon, P. J., Dragic, T., and Olson, W. C. 2003. L-SIGN (CD 209L) is a liver-specific capture receptor for hepatitis C virus. *Proc Natl Acad Sci USA* 100:4498–4503.
36. Lozach, P. Y., Lortat-Jacob, H., De Lacroix De Lavalette, A., Staropoli, I., Foung, S., Amara, A., Houles, C., Fieschi, F., Schwartz, O., Virelizier, J. L., Arenzana-Seisdedos, F., and Altmeyer, R. 2003. DC-SIGN and L-SIGN are high-affinity binding receptors for hepatitis C Virus glycoprotein E2. *J Biol Chem* 278:20358–20366.
37. Pohlmann, S., Zhang, J., Baribaud, F., Chen, Z., Leslie, G. J., Lin, G., Granelli-Piperno, A., Doms, R. W., Rice, C. M., and McKeating, J. A. 2003. Hepatitis C virus glycoproteins interact with DC-SIGN and DC-SIGNR. *J Virol* 77:4070–4080.

38. Alvarez, C. P., Lasala, F., Carrillo, J., Muniz, O., Corbi, A. L., and Delgado, R. 2002. C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in *cis* and in *trans*. *J Virol* 76:6841–6844.
39. Simmons, G., Reeves, J. D., Grogan, C. C., Vandenberghe, L. H., Baribaud, F., Whitbeck, J. C., Burke, E., Buchmeier, M. J., Soilleux, E. J., Riley, J. L., Doms, R. W., Bates, P., and Pohlmann, S. 2003. DC-SIGN and DC-SIGNR bind ebola glycoproteins and enhance infection of macrophages and endothelial cells. *Virology* 305:115–123.
40. Halary, F., Amara, A., Lortat-Jacob, H., Messerle, M., Delaunay, T., Houles, C., Fieschi, F., Arenzana-Seisdedos, F., Moreau, J. F., and Dechanet-Merville, J. 2002. Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection. *Immunity* 17:653–664.
41. Navarro-Sanchez, E., Altmeyer, R., Amara, A., Schwartz, O., Fieshi, F., Virelizier, J. L., Arenzana-Seisdedos, F., and P. Despres. 2003. Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. *EMBO Rep.* 4:723–728.
42. Ludwig, I. S., Lekkerkerker, A. N., Depla, E., Bosman, F., Musters, R. J., Depraetere, S., van Kooyk, Y., and Geijtenbeek, T. B. 2004. Hepatitis C virus targets DC-SIGN and L-SIGN to escape lysosomal degradation. *J Virol* 78:8322–8332.
43. Bashirova, A. A., Geijtenbeek, T. B. H., van Duijnhoven, G. C. F., van Vliet, S. J., Eilering, J. B., Martin, M. P., Wu, L., Martin, T. D., Viebig, N., Knolle, P. A., KewalRamani, V. N., van Kooyk, Y., and Carrington, M. 2001. A dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin (dc-sign)-related protein is highly expressed on human liver sinusoidal endothelial cells and promotes hiv-1 infection. *J Exp Med* 193:671–678.
44. Soilleux, E. J., Barten, R., and Trowsdale, J. 2000. DC-SIGN; a related gene, DC-SIGNR; and CD23 form a cluster on 19p13. *J Immunol* 165:2937–2942.
45. Limmer, A., Ohl, J., Kurts, C., Ljunggren, H. G., Reiss, Y., Groettrup, M., Momburg, F., Arnold, B., and Knolle, P. A. 2000. Efficient presentation of exogenous antigen by liver endothelial cells to CD8+ T cells results in antigen-specific T-cell tolerance. *Nat Med* 6:1348–1354.
46. Knolle, P. A., Uhrig, A., Hegenbarth, S., Loser, E., Schmitt, E., Gerken, G., and Lohse, A. W. 1998. IL-10 down-regulates T cell activation by antigen-presenting liver sinusoidal endothelial cells through decreased antigen uptake via the mannose receptor and lowered surface expression of accessory molecules. *Clin Exp Immunol* 114:427–433.
47. van Vliet, S. J., van Liempt, E., Geijtenbeek, T. B., and van Kooyk, Y. 2006. Differential regulation of C-type lectin expression on tolerogenic dendritic cell subsets. *Immunobiology* 211:577–585.
48. van Vliet, S. J., Gringhuis, S. I., Geijtenbeek, T. B., and van Kooyk, Y. 2006. Regulation of effector T cells by antigen-presenting cells via interaction of the C-type lectin MGL with CD45. *Nat Immunol* 7:1200–1208.
49. van Vliet, S. J., van Liempt, E., Saeland, E., Aarnoudse, C. A., Appelmelk, B., Irimura, T., Geijtenbeek, T. B., Blixt, O., Alvarez, R., Van, D., I, and van Kooyk, Y. 2005. Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *Int Immunol* 17:661–669.
50. Engering, A. J., Cella, M., Fluitsma, D., Brockhaus, M., Hoefsmit, E. C., Lanzavecchia, A., and Pieters, J. 1997. The mannose receptor functions as a high capacity and broad specificity antigen receptor in human dendritic cells. *Eur J Immunol* 27:2417–2425.
51. Linehan, S. A., Coulson, P. S., Wilson, R. A., Mountford, A. P., Brombacher, F., Martinez-Pomares, L., and Gordon, S. 2003. IL-4 receptor signaling is required for mannose receptor expression by macrophages recruited to granulomata but not resident cells in mice infected with *Schistosoma mansoni*. *Lab Invest* 83:1223–1231.
52. Stahl, P., Schlesinger, P. H., Sigardson, E., Rodman, J. S., and Lee, Y. C. 1980. Receptor-mediated pinocytosis of mannose glycoconjugates by macrophages: Characterization and evidence for receptor recycling. *Cell* 19:207–215.
53. Martinez-Pomares, L., Linehan, S. A., Taylor, P. R., and Gordon, S. 2001. Binding properties of the mannose receptor. *Immunobiology* 204:527–535.
54. Taylor, M. E., Bezouska, K., and Drickamer, K. 1992. Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor. *J Biol Chem* 267:1719–1726.
55. Nguyen, D. G. and Hildreth, J. E. 2003. Involvement of macrophage mannose receptor in the binding and transmission of HIV by macrophages. *Eur J Immunol* 33:483–493.
56. Schlesinger, L. S., Kaufman, T. M., Iyer, S., Hull, S. R., and Marchiando, L. K. 1996. Differences in mannose receptor-mediated uptake of lipoarabinomannan from virulent and attenuated strains of *Mycobacterium tuberculosis* by human macrophages. *J Immunol* 157:4568–4575.

57. van Kooyk, Y. and Geijtenbeek, T. B. 2003. DC-SIGN: Escape mechanism for pathogens. *Nat Rev Immunol* 3:697–709.
58. Frison, N., Taylor, M. E., Soilleux, E., Bousser, M. T., Mayer, R., Monsigny, M., Drickamer, K., and Roche, A. C. 2003. Oligolysine-based oligosaccharide clusters: Selective recognition and endocytosis by the mannose receptor and DC-SIGN. *J Biol Chem* 278:23922–23929.
59. Napper, C. E., Drickamer, K., and Taylor, M. E. 2006. Collagen binding by the mannose receptor mediated through the fibronectin type II domain. *Biochem J* 395:579–586.
60. Hunger, R. E., Sieling, P. A., Ochoa, M. T., Sugaya, M., Burdick, A. E., Rea, T. H., Brennan, P. J., Belisle, J. T., Blauvelt, A., Porcelli, S. A., and Modlin, R. L. 2004. Langerhans cells utilize CD1a and langerin to efficiently present nonpeptide antigens to T cells. *J Clin Invest* 113:701–708.
61. Valladeau, J., Ravel, O., Dezutter-Dambuyant, C., Moore, K., Kleijmeer, M., Liu, Y., Duvert-Frances, V., Vincent, C., Schmitt, D., Davoust, J., Caux, C., Lebecque, S., and Saeland, S. 2000. Langerin, a novel C-type lectin specific to Langerhans cells, is an endocytic receptor that induces the formation of Birbeck granules. *Immunity* 12:71–81.
62. Galustian, C., Park, C. G., Chai, W., Kiso, M., Bruening, S. A., Kang, Y. S., Steinman, R. M., and Feizi, T. 2004. High and low affinity carbohydrate ligands revealed for murine SIGN-R1 by carbohydrate array and cell binding approaches, and differing specificities for SIGN-R3 and langerin. *Int Immunol* 16:853–866.
63. Stambach, N. S. and Taylor, M. E. 2003. Characterization of carbohydrate recognition by langerin, a C-type lectin of Langerhans cells. *Glycobiology* 13:401–410.
64. de Witte, L., Nabatov, A., Pion, M., Fluitsma, D., de Jong, M. A., de Gruijl, T., Piguet, V., van Kooyk, Y., and Geijtenbeek, T. B. 2007. Langerin is a natural barrier to HIV-1 transmission by Langerhans cells. *Nat Med* 13:367–371.
65. van der Aar, A. M., Sylva-Steenland, R. M., Bos, J. D., Kapsenberg, M. L., de Jong, E. C., and Teunissen, M. B. 2007. Loss of TLR2, TLR4, and TLR5 on Langerhans cells abolishes bacterial recognition. *J Immunol* 178:1986–1990.
66. Underhill, D. M. and Ozinsky, A. 2002. Toll-like receptors: Key mediators of microbe detection. *Curr Opin Immunol* 14:103–110.
67. Gross, O., Gewies, A., Finger, K., Schafer, M., Sparwasser, T., Peschel, C., Forster, I., and Ruland, J. 2006. Card9 controls a non-TLR signalling pathway for innate anti-fungal immunity. *Nature* 442:651–656.
68. Rogers, N. C., Slack, E. C., Edwards, A. D., Nolte, M. A., Schulz, O., Schweighoffer, E., Williams, D. L., Gordon, S., Tybulewicz, V. L., Brown, G. D., and Reis e Sousa, C. 2005. Syk-dependent cytokine induction by Dectin-1 reveals a novel pattern recognition pathway for C type lectins. *Immunity* 22:507–517.
69. Dzionek, A., Sohma, Y., Nagafune, J., Cella, M., Colonna, M., Facchetti, F., Gunther, G., Johnston, I., Lanzavecchia, A., Nagasaka, T., Okada, T., Vermi, W., Winkels, G., Yamamoto, T., Zysk, M., Yamaguchi, Y., and Schmitz, J. 2001. BDCA-2, a novel plasmacytoid dendritic cell-specific type II C-type lectin, mediates antigen capture and is a potent inhibitor of interferon alpha/beta induction. *J Exp Med* 194:1823–1834.
70. Caparros, E., Munoz, P., Sierra-Filardi, E., Serrano-Gomez, D., Puig-Kroger, A., Rodriguez-Fernandez, J. L., Mellado, M., Sancho, J., Zubiaur, M., and Corbi, A. L. 2006. DC-SIGN ligation on dendritic cells results in ERK and PI3K activation and modulates cytokine production. *Blood* 107:3950–3958.
71. Gringhuis, S.I., Den Dunnen, J., Litjens, M., van het Hof, B., van Kooyk, Y., and Geijtenbeek, T.B. 2007. Raf-1 kinase-dependent acetylation of transcription factor NF-kappaB. *Immunity* 26:605–616.

25 Structural and Functional Roles of C-Type Lectin Receptors on Natural Killer Cells

Nazzareno Dimasi

CONTENTS

25.1 INTRODUCTION

A crucial function in the innate immune defense against tumors, bacteria, virus, and parasites is mediated by a specialized class of large granulated lymphocytes termed natural killer (NK) cells (Trinchieri 1989; Lanier 2005). To sense their environment, NK cells are equipped with specific cell surface receptors. These are known as NK receptor domains (NKD). A fine balance of inhibitory and activating stimuli coordinates the cytolyic activity of the NK cells.

In normal conditions, the dominant signal received by an NK cell through its interaction with the major histocompatibility complex (MHC) class I ligand, expressed on the target cell, is inhibitory (Figure 25.1A). In the event that the normal expression of MHC Class I ligands is decreased or missing, because of tumor or viral infection, the inhibitory signal is then displaced and the NK cell acquires the capacity to kill the target cell. Thus, NK cells sense the absence of self-MHC class I molecules as a way to discriminate normal cells from cells in distress (Figure 25.1B) (Kärre 1985; Ljunggren and Kärre 1985). The activation of NK cells leads to the production of cytokines and chemokines and to an instantaneous secretion of cytotoxic granules that are needed to kill the target cells.

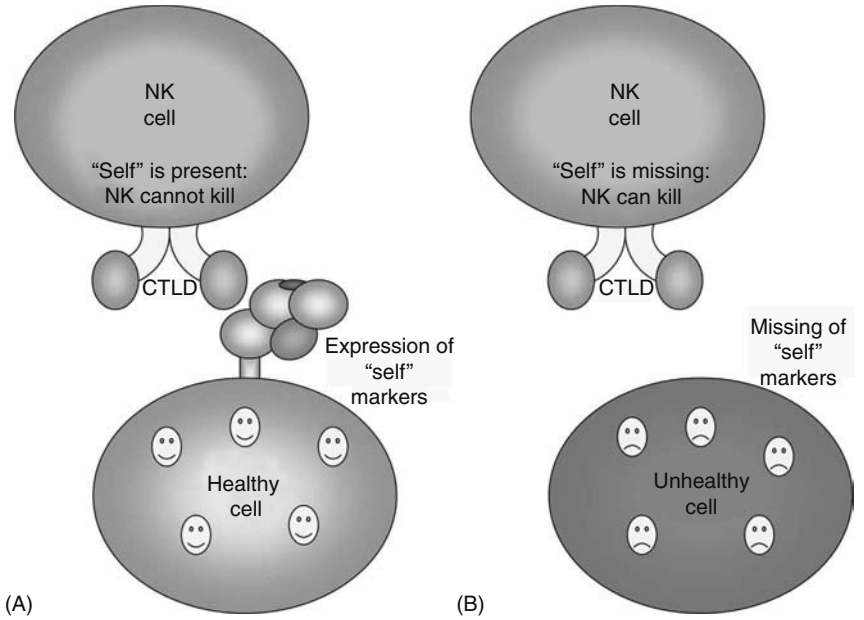


FIGURE 25.1 (See CD for color figure.) Schematic illustration of the NK cells function. This is a basic example of how a mouse NK cell distinguishes healthy versus unhealthy cells. This example is based on the "missing-self recognition" hypothesis introduced by Klas Kärre. (From Kärre, K., in: Callewaert, D. and Herberman, R.B., eds. *Mechanisms of Cytotoxicity by NK Cells*, Academic Press, San Diego, CA, 1985, p. 81; Ljunggren, H.G. and Kärre, K., *J. Exp. Med*., 162, 1745, 1985.) The attack of NK cells (A) on healthy cells is actively suppressed by the recognition of self-MHC class I molecules by NK inhibitory receptors. Consequently (B), when the missing self-markers are absent or their expression is abnormal due to tumor transformation or viral infection, the NK cells are then more likely to be activated. This will boost a sequence of molecular events culminating with the lyses of the unhealthy cells.

Two different structural families of NK cell receptors regulate the NK cell activity: the immunoglobulin-like superfamily (IgSF: killer immunoglobulin-like receptor (KIRs), leukocyte immunoglobulin-like receptors (LIRs), and natural cytotoxicity receptors (NCRs) and the C-type lectin-like domain superfamily (CTLD: Ly49s, NKG2D, CD94/NKG2, NKR-P1, and CD69) (Sawicki 2001; Natarajan et al. 2002; Dimasi et al. 2004; Dimasi and Biassoni 2005; Lanier 2005). KIR receptors in humans recognize the $\alpha 1$ and $\alpha 2$ domains of HLA-A, -B, and -C complexed with the bound antigenic peptide. The LIR in humans recognize the nonpolymorphic $\alpha 3$ domain of classical and nonclassical HLA molecules (Natarajan et al. 2002; Lanier 2005). Members of the CTLD superfamily include the heterodimer CD94/NKG2A, which bind the nonclassical MHC class I HLA-E and Qa-1 in human and mouse, respectively. NKG2D recognizes the MHC class I-like molecules MICA, ULBP, H60, RAE-1β, and MULT-1 (Lanier 2005). The mouse Ly49 recognizes classical MHC class I molecules and, interestingly, MHC class I-like molecules expressed by the cytomegalovirus (Lanier 2005).

Both types of IgSF and CTLD include inhibitory and activating receptors (Figure 25.2). A marked difference between inhibitory and activating receptors is the fact that inhibitory receptors are able upon ligand engagement to mediate intracellular signaling transduction directly (Figure 25.2A). On the contrary, all the activating receptors, with the exception of 2B4, need to associate with a specialized class of adaptor molecules (DAP-10, DAP-12, FcRγ, and CD3ζ) in order to carry out their intracellular signaling transduction (Figure 25.2B).

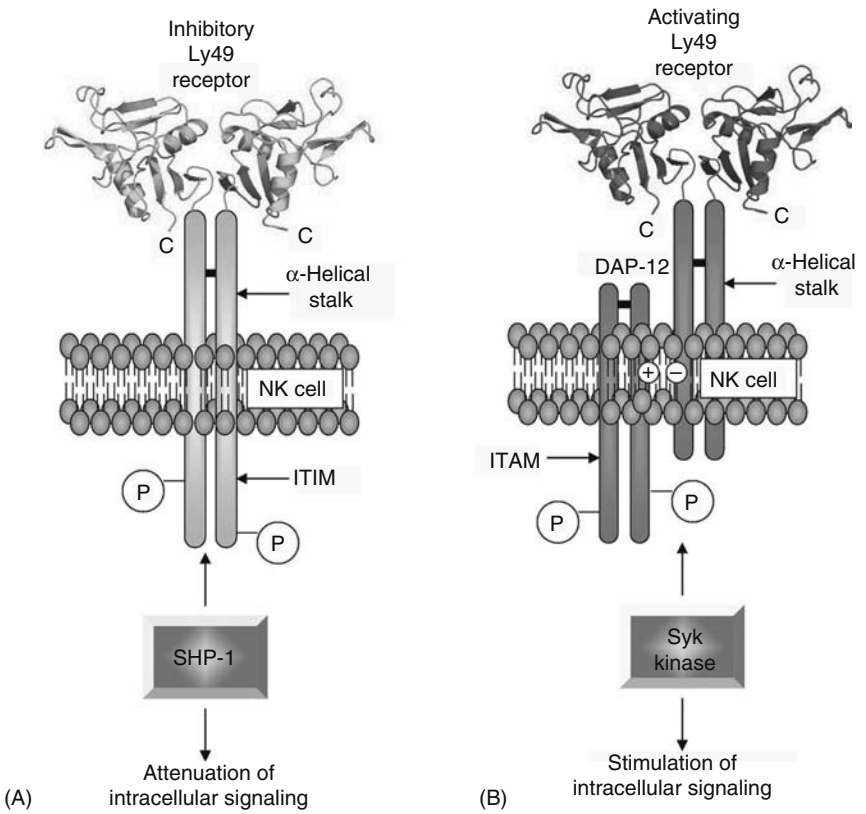


FIGURE 25.2 (See color insert following blank page 170. Also see CD for color figure.) Inhibitory and activating mouse Ly49 NK cell receptors. The Ly49 NK cell receptors are type II disulfide-linked homodimeric (i.e., the carboxy-terminal part of the protein chain is extracellular) transmembrane proteins. These receptors consist of a globular CTLD domain connected to the cell membrane by a long and flexible stalk region of approximately 70 amino acids. In structural terms, inhibitory and activating receptors are very similar. However, in functional terms, inhibitory and activating receptors transduce their intracellular signaling in a very different way. In fact, (A) inhibitory receptors are colinear with their intracellular transducing chains whereas (B) activating receptors, in order to transduce their intracellular signals, need to associate with specialized transducing polypeptides, such as the DAP-12 protein. Upon ligand engagement, the intracellular chains of the inhibitory Ly49 receptors became phosphorylated at specific tyrosine residues within the ITIM consensus. This phosphorylation event is required for the recruitment of adaptor proteins (e.g., SHP-1 kinase) that in turn will initiate a cascade of events culminating with the inhibition of the NK cell activation. Vice versa, ligand engagement of Ly49 activating receptors results in the ITAM phosphorylation of DAP-12 with subsequent recruitment and activation of Syk kinase. This event initiates the NK cell activation. Activating Ly49, associate with DAP-12 through the formation of a specific hydrogen bond in their transmembrane region.

Within the past 10 years, many functional and structural data have been accumulating at an overwhelming pace. These studies revealed the multiplicity of molecular solutions employed by NK cells in their constant battle against invading pathogens and tumors. In addition, these studies uncovered the economy of the molecular evolution; where a robust scaffold, the C-type lectin-domain, is preserved even though the capacity to bind carbohydrate structures in a calcium (Ca^{2+})-dependent manner is lacking on this specialized class of receptors. In this chapter, I focus on the structural and functional roles of C-type lectin receptors on NK cells.

25.2 C-TYPE LECTIN-LIKE NK CELL RECEPTORS

The C-type lectins are calcium-dependent carbohydrate binding proteins (Zelensky and Gready 2005). An example of these classes of proteins is the mannose-binding protein (MBP), E-selectins, tetranectins, and lithostatin. In the NK cells, examples are the mouse Ly49 receptors, NKG2D, CD94, and CD69. These proteins are made up of an intracellular part and an extracellular domain. The extracellular domain has a C-type lectin-like domain (CTLD) of approximately 120 amino acids, and a relatively long (about 70 amino acids) stalks region (Figure 25.2). Despite the diversity of ligands recognized by the NK cell receptors of the CTLD superfamily, they exhibit a highly conserved core structure (Figure 25.3A). This core consists of two antiparallel β-sheets and in most cases two α-helixes (α1 and α2) (Figure 25.3B). Exception to this conserved folding core is the CTLD of CD94, in which the α2-helix is replaced by a long unstructured loop (Boyington et al. 1999: 75). Two invariant disulfide bridges together with a set of conserved hydrophobic residues are responsible for maintaining and stabilizing the CTLD core (Figure 25.3).

A unique characteristic of the CTLD of the NK receptors is that, although they are C-type lectins, they do not appear to have conserved the residues that are important for binding the calcium (Figure 25.3A). As consequence, their ligands are not carbohydrates but are proteins (Natarajan et al. 2002; Lanier 2005).

25.3 LY49 RECEPTORS

The mouse Ly49 NK receptors of the CTLD are a multigene family that maps on chromosome 6 in a region known as the NK gene complex (NKC) (Takei et al. 1997; Yokoyama and Plougastel 2003; Kelley et al. 2005; Ortaldo and Young 2005). At least 23 genes exist within the NKC, Ly49A through W, which are the major MHC class I patrolling receptors on mouse NK cells (Wilhelm et al. 2002; Dimasi and Biassoni 2005; Deng and Mariuzza 2006). Ly49L is the only human member of the Ly49 multigene family identified (Westgaard et al. 1998).

The Ly49 receptors are expressed at the NK cell surface as type-II transmembrane disulfide-linked proteins (Figure 25.2). Each Ly49 chain has a CTLD domain that is connected to the cell membrane by a long unstructured region (Figure 25.2). Of the 23 Ly49 NK receptors, 13 are inhibitory based on functional data or by the presence of inhibitory signature sequences immunoreceptor tyrosine-inhibitory motifs (ITIM) in their intracytoplasmic domains. These are Ly49A, B, C, E, F, G, I, J, O, Q, S, T, and V. The remaining receptors, Ly49D, H, L, M, P, R, U, and W, are believed to have an activating function. Even though the Ly49 proteins have evolved from the C-type lectins, ligand specificity depends principally on protein–protein interaction although carbohydrates could have some role in the recognition events (Lian et al. 1998).

The cellular ligands of the Ly49 NK receptors are classical MHC class I molecules. Ly49 binding to MHC class I requires the presence of any bound antigenic peptide (Correa and Raulet 1995; Orihuela et al. 1996). Nevertheless, some degree of peptide specificity has been reported for Ly49I (Michaelsson et al. 2000) and Ly49C (Hanke et al. 1999). However, these evidences are not supported by structural studies, in which a direct binding of the Ly49 receptors with the specific peptides was not observed (Tormo et al. 1999; Dam et al. 2003).

Recently, a functional role for the Ly49 NK receptors in viral immunity has been revealed by the identification of a MHC class I-like molecule, the murine cytomegalovirus m157 glycoprotein, being the ligand for an inhibitory Ly49 receptor (Ly49I) in a susceptible mouse, and for an activating Ly49 receptor (Ly49H) in a resistant mouse (Arase et al. 2002; Smith et al. 2002; Voigt et al. 2003; Lanier 2005).

25.3.1 3D Structure of Monomeric Ly49

At present, Ly49 NK cell receptors for which the three-dimensional structures are known are Ly49A bound to MHC class I H-2D^d (Tormo et al. 1999; Dam et al. 2006) (Figure 25.4A), Ly49C bound to

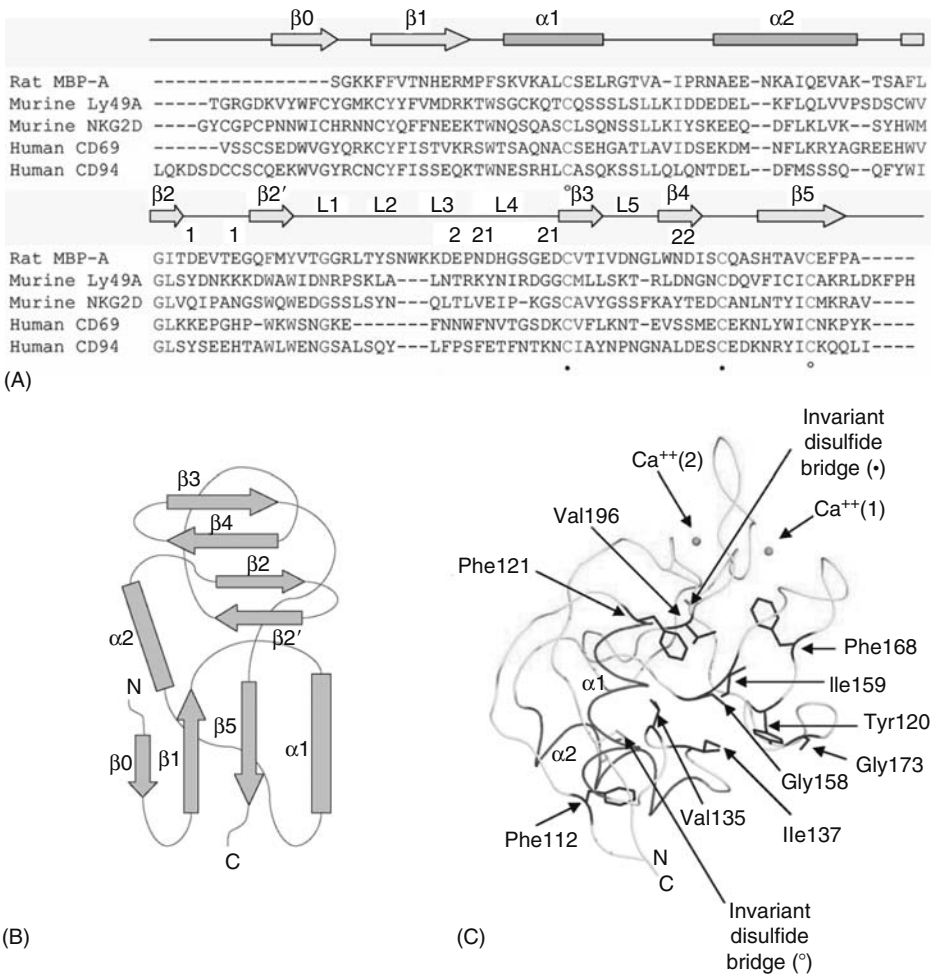


FIGURE 25.3 (See color insert following blank page 170. Also see CD for color figure.) Structure-based sequence alignment, topology diagram, and core structure of the NKD of the CTLD superfamily and of the MBP-A. (A) Structure-based sequence alignment of the NKD belonging to the CTLD fold for which the 3D-structure is known, compared with the classical CTLD of the MBP-A. The secondary structure elements are shown above the alignment and are denoted by arrows (β-strands) and rectangles (α-helixes). The numerations of the secondary structural elements are labeled sequentially according to the numeration of the MPB-A, the superfamily prototype. (From Weis, W.I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W.A., *Science*, 254, 1608, 1991.) The cysteines forming the conserved disulfide bridges in the CTLD core are shown in red and are paired with filled and opened circles. The residues colored in blue (together with the paired cysteines in red) correspond to the residues that are the signature of the CTLD core. The residues that ligate the Ca^{2+} are labeled as 1 or 2 if they bind the Ca^{2+} at site 1 or 2, respectively. The PDB accession numbers for the structures used to generate the alignment are rat MBP-A (1YTT), murine Ly49A (1QO3), murine NKG2D (1HQ8), human CD69 (1E87), and human CD94 (1B6E). The sequence alignments were performed with the program CLUSTAL W. (From Higgins, D., Thompson, J., Gibson, T., Thompson, J., Higgins, D.G., and Gibson, T.J., *Nucleic Acids Res.*, 22, 4673, 1994.) (B) Topology diagram of the CTLD of the NKD receptor superfamily. This secondary structural organization (loops, β-strands, and α-helixes) is well conserved among the Ly49, NKG2D, and CD69 receptors. Exception to this fold is CD94 in which an extended loop replaces the α2-helix. C and N denote the carboxyl- and N-terminus of the protein chain. β-Strands are shown as arrows and α-helixes as rectangles. (C) α-Carbon trace of the CTLD fold characteristic of the NKD superfamily. The side chains for the residues that define the CTLD core of the NKD are shown together with the two invariant disulfide bridges. These residues correspond to the highlighted residues (blue and red) in the sequence alignment shown in (A). These residues are numbered sequentially according to the numeration of the MPB-A. (From Weis, W.I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W.A., *Science*, 254, 1608, 1991.) The locations of the two Ca^{2+} sites are shown as violet dots. The two invariant disulfide bridges (open and closed circles in the sequence alignment) are shown.

MHC Class I H-2 K^b (Dam et al. 2003) (Figure 25.4B), and Ly49I in free form (Dimasi et al. 2002) (Figure 25.4C). As shown in Figure 25.4, the Ly49 CTLD structure consists of two α-helixes (α1 and α2) and two antiparallel β-sheets. The two β-sheets are formed by β-strands β0, β1, and β5 and, β2, β2′, β3, and β4, respectively.

Four intrachain disulfide bonds are conserved among the Ly49 receptors (Figure 25.4). The disulfide bond Cys167–Cys253 and Cys232–Cys245 (the numeration is based upon the full-length mature Ly49A protein) are the invariant disulfides found in all C-type animal lectins (Figure 25.3A and C). The disulfide bridge Cys145–Cys150 is the invariant disulfide present in the long form of the C-type lectin proteins (Weis et al. 1991; Day 1994). The fourth disulfide Cys163–Cys251 is unique in the Ly49 proteins and links the N-terminus of strand β5 to the helix α1. The loops that connect the secondary structural elements constitute regions that are variable in the Ly49 receptors.

25.3.2 3D Structure of Dimeric Ly49

The Ly49 NK receptors exist at the cell surface as disulfide-linked homodimers (Figure 25.2) (Yokoyama et al. 1989). Detailed structural information is available for only the CTLD domains of the Ly49 (Figure 25.5) without any structural information for the stalk region. X-ray crystallography studies revealed that the Ly49 receptors could adopt a different dimerization mode (Dimasi and Biassoni 2005). In fact, the Ly49A homodimer (Tormo et al. 1999; Dam et al. 2006) is different from the Ly49C (Dam et al. 2003) and Ly49I (Dimasi et al. 2002) homodimers.

To form the Ly49 CTLD homodimer, the monomers are linked through the strand β0, which create in the dimer an extended antiparallel β-sheet (Figure 25.5). In the Ly49A homodimer, the C-terminal parts of the α2-helixes pack against one another. This creates a "closed dimer" (Figure 25.5A). Contrarily, in the Ly49C and Ly49I homodimers, the α2-helixes are apart from one another, opening up these dimers (Figure 25.5B and C).

An important consequence of the variability in the Ly49 dimerization mode is to direct the manner in which these homodimers can engage the MHC class I ligands (Tormo et al. 1999; Dam et al. 2003), or MHC class I-like molecules (Kielczewska et al. 2006).

25.3.3 Structural Basis for the MHC Class I Recognition by Ly49 Receptors

The structure of the C-type lectin-like domain of Ly49A in complex with its ligand MHC class I H-2D^d (Tormo et al. 1999) and the complex between Ly49C and its ligand MHC class I H-2 K^b (Dam et al. 2003), determined up to now by x-ray crystallography, are shown in Figure 25.6. These complex structures revealed that the Ly49 molecules could their target through protein–protein interaction, confirming that these NK receptors are Ca-independent lectins. The Ly49A/H-2D^d was the first high-resolution structural example of MHC class I recognition by NK cells. As shown in Figure 25.6A, the Ly49A dimer makes contacts with the H-2D^d ligand asymmetrically through two spatially different sites, termed "site 1" and "site 2," which do not overlap with the TCR-binding area on the H-2D^d allele (Chung et al. 1999). One monomer of Ly49A contacts the H-2D^d ligand ("site 1;" Figure 25.6A) at one end of the peptide-binding platform of the MHC. This region is near one of the glycosylation site of the MHC (Parham 1996, 2000; Mandal and Mukhopadhyay 2001a,b) and is the region of the MHC that is highly polymorphic (Beck and Trowsdale 2000; Nikolich-Zugich et al. 2004; Bontrop 2006). "Site 2" (Figure 25.6A) resides in a cavity under the peptide-binding platform of the H-2D^d, in a region that in part overlaps the CD8 (Gao et al. 1997) and the LIR-1 (Willcox et al. 2003) binding site on the MHC molecule. Here the Ly49A monomer makes extensive contacts with the α1, α2, α3, and β2-microglobulin (β2 m) domains of the MHC (Figure 25.6A). In this complex structure, none of

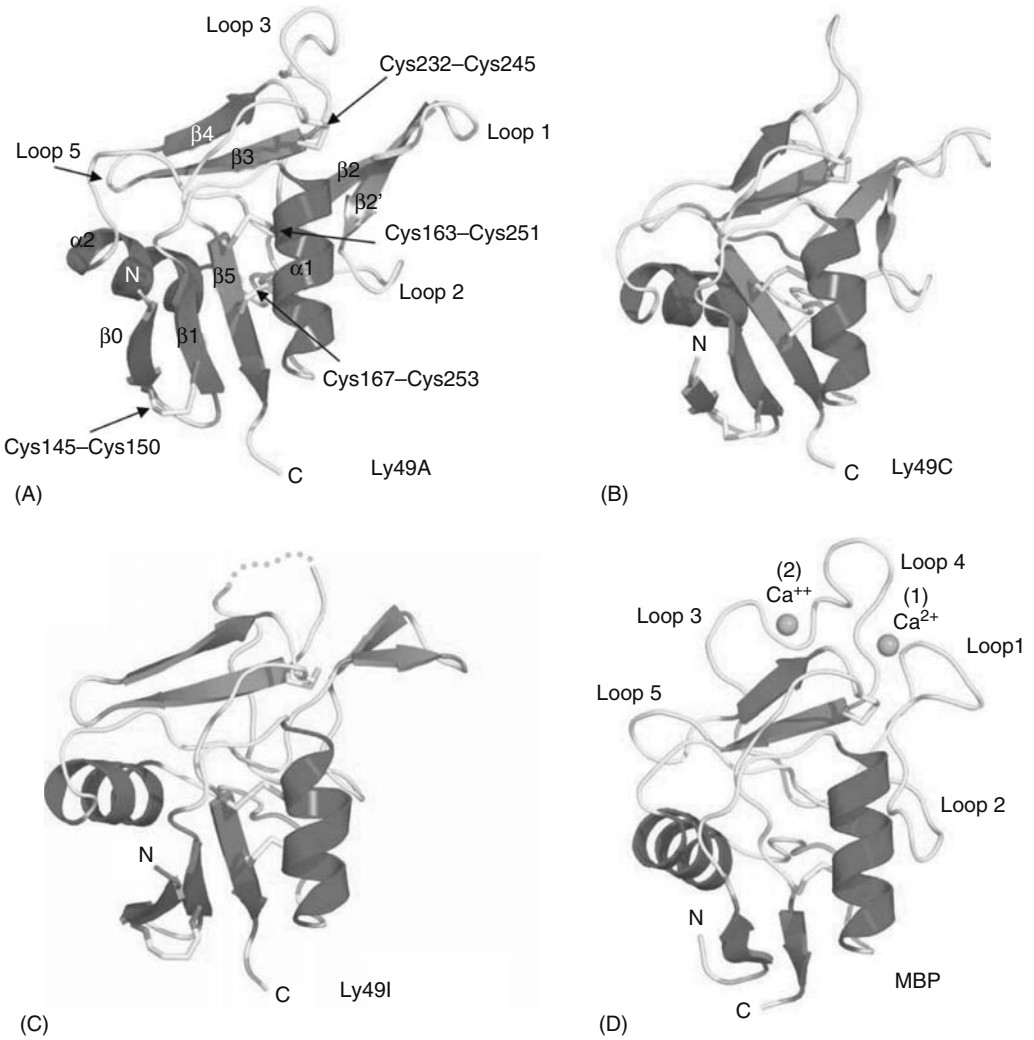


FIGURE 25.4 (See CD for color figure.) Anatomy of CTLD domains for Ly49. As seen in this cartoon representation, the 3D structure of the CTLD of Ly49A (A) (PDB entry code 1QO3, chain C; From Tormo, J., Natarajan, K., Margulies, D.H., and Mariuzza, R.A., *Nature*, 402, 623, 1999), Ly49C (B) (PDB entry code 1P1Z, Chain C; From Dam, J., Guan, R., Natarajan, K., Dimasi, N., Chlewicki, L.K., Kranz, D.M., Schuck, P., Margulies, D.H., and Mariuzza, R.A., *Nat. Immunol*., 4, 1213, 2003), Ly49I (C) (PDB entry code 1JA3, chain A; From Dimasi, N., Sawicki, M.W., Reineck, L.A., Li, Y., Natarajan, K., Margulies, D.H., and Mariuzza, R,A., *J. Mol. Biol*., 320, 573, 2002), and rat MBP-A (D) (PDB entry code 1YTT, chain A; From Weis, W.I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W.A., *Science*, 254, 1608, 1991) are highly conserved. The CTLD fold consists of two α-helixes (α1 and α2) and two antiparallel β-sheets, formed by β-strand β0-β1-β5 and β3-β4. The strand β2 forms a short β-hairpin with strand β2′. In this representation, the disulfide bonds are shown as sticks and are labeled using their corresponding cysteine residues. These disulfide bridges are conserved between the Ly49 NKD. Two of these disulfide bridges (Cys167–Cys253 and Cys232–Cys245) are also conserved with the invariant disulfide bridges found in the MBP-A (see also Figure 25.2A and C). (From Weis, W.I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W.A., *Science*, 254, 1608, 1991.) The secondary structural element α-helixes and β-strands, N and C termini, and the disulfide bridges are shown only in Ly49A. The locations of the loop regions are shown for Ly49A and MBP-A. The two calcium ions at site 1 and site 2, bound to the MBP-A, are drawn as violet spheres. The figure was prepared using the program Pymol. (From DeLano, W.L., *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA, 2002.)

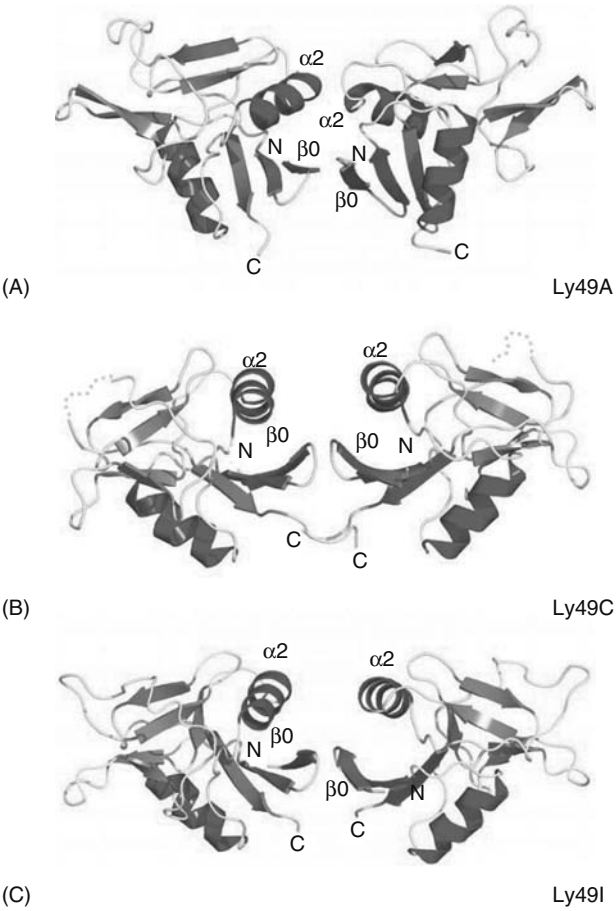


FIGURE 25.5 (See CD for color figure.) Ribbon models showing the different dimerization mode within the Ly49 NK cell receptors. (A) Dimeric structure of the CTLD of Ly49A (PDB entry code 1QO3, chains C and D; From Tormo, J., Natarajan, K., Margulies, D.H., and Mariuzza, R.A., *Nature*, 402, 623, 1999), (B) Ly49C (PDB entry code 1P1Z, chain C and D; From Dam, J., Guan, R., Natarajan, K., Dimasi, N., Chlewicki, L.K., Kranz, D.M., Schuck, P., Margulies, D.H., and Mariuzza, R.A., *Nat. Immunol*., 4, 1213, 2003), and (C) Ly49I (PDB entry code 1JA3; From Dimasi, N., Sawicki, M.W., Reineck, L.A., Li, Y., Natarajan, K., Margulies, D.H., and Mariuzza, R.A., *J. Mol. Biol*., 320, 573, 2002). The secondary structural elements (strand β0 and helix α2) involved in the dimer formation are labeled. This illustration emphasizes the different variability within the dimerization mode of the Ly49 CTLD receptors. The Ly49A dimer is more elongated (55 Å) than the Ly49C or Ly49I dimers (49 Å). C and N denote the carboxy- and the amino-termini of the CTLD chains. This figure was prepared using the program Pymol. (From DeLano, W.L., *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, CA, 2002.)

the two Ly49A monomers make direct contact with the bound peptide and this could explain the lack of peptide specificity for the Ly49A/H-2D^d interaction (Waldenstrom et al. 2002; Wang et al. 2002).

The discovery of these two distinct binding sites raised the question of which of the two sites is the functional binding site. A first clue to answer this question comes from mutagenesis studies of the Ly49A, β2m, and the H-2D^d (Matsumoto et al. 2001a,b; Wang et al. 2002). These studies identified "site 2" as the primary interacting surface responsible for the functional interaction

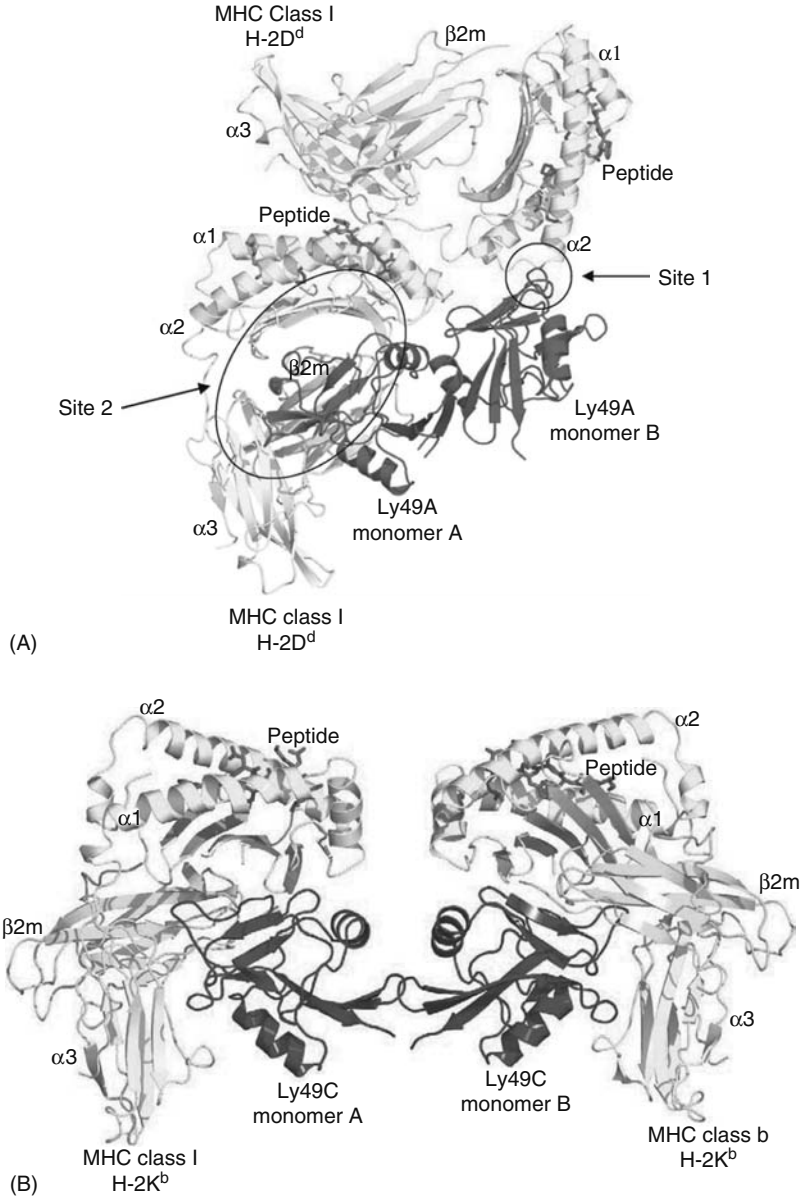


FIGURE 25.6 (See color insert following blank page 170. Also see CD for color figure.) 3D structure of the complex between Ly49A/H-2D^d and Ly49C/H-2 K^b. The different domains composing the MHC class I molecule (α1, α2, α3, β2 m, and the peptide) and the Ly49 monomers are schematically labeled in both panels. (A) The complex structure between Ly49A and H-2D^d (PDB entry code 1QO3) revealed two distinct sites of interaction (termed "site 1" and "site 2") for Ly49 to MHC class I. (From Tormo, J., Natarajan, K., Margulies, D.H., and Mariuzza, R.A., *Nature*, 402, 623, 1999.) "Site 1" is located at the N-terminal end of the α1 helix of the MHC class I molecule and is largely dominated by electrostatic interactions. "Site 2" is located beneath the peptide binding platform and involves interaction with all the MHC Class I domains, including the β2 m. This interface is primarily formed by hydrogen bonds interactions. (B) X-ray crystal structure of Ly49C in complex with its MHC class I ligand H-2 K^b (PDB entry code 1P1Z). (From Dam, J., Guan, R., Natarajan, K., Dimasi, N., Chlewicki, L.K., Kranz, D.M., Schuck, P., Margulies, D.H., and Mariuzza, R.A., *Nat. Immunol*., 4, 1213, 2003.) As seen in this complex structure, MHC class I engagement by Ly49C is different from that of Ly49A. Ly49C engages symmetrically the MHC class I molecules using "site 2." In both complex structures, the Ly49 receptors do not make contacts with the bound peptides.

between Ly49A and H-2D^d. Subsequently, the determination of the complex structure of the Ly49C/H-2 K^b (Dam et al. 2003) pointed out that Ly49 receptors use "site 2" to interact with the MHC ligand (Figure 25.6B). In fact, this complex structure revealed that the Ly49C homodimer engages symmetrically with two H-2 K^b molecules. Each Ly49C monomer makes identical interactions with MHC class I at a location that is virtually superimposable with "site 2" location in the complex Ly49A/H-2D^d. These two complex structures and the crystal structure of the unbound Ly49I (Dimasi et al. 2002), revealed a different dimerization mode for the Ly49 receptors. In fact, the Ly49A dimer observed in the Ly49A/H-2D^d complex is compact (Figure 25.5A), whereas the dimer observed for the Ly49C/H-2 K^b complex is open (Figure 25.5B), similarly to the open conformation observed for the unbound Ly49I (Figure 25.5C). This suggests that the Ly49 receptors have the ability to adopt the open or the closed conformation as a means to discriminate between functional or nonfunctional interaction with their respective ligands (Dam et al. 2006).

25.4 CD94/NKG2 RECEPTORS

CD94/NKG2 are type-II disulfide-linked heterodimers with a lectin-like fold expressed on NK and a small subset of T cells. This complex could function as an inhibitor or an activator depending on which isoform of NKG2 is expressed. The *CD94/NKG2* genes are present in humans (Houchins et al. 1991; Chambers et al. 1993; Yabe et al. 1993; Lazetic et al. 1996; Houchins et al. 1997; Lieto et al. 2006), mouse (Vance et al. 1998, 1999; Lohwasser et al. 1999, 2000, 2001) and rat (Berg et al. 1998; Dissen et al. 1997). These genes are located within the NKC. Specifically they map on human chromosome 12 and in mouse on chromosome 6. These genes encode receptors that recognize nonclassical MHC class I molecules, such as HLA-E in human and Qa-1b in mouse (Lanier 2005) that bind leader peptides from other MHC class I molecules.

Unlike the *Ly49* genes, the *CD94/NKG2* genes have very limited polymorphism. CD94 is encoded by a single gene (Chang et al. 1995) while NKG2 constitutes a small multigenic family that encodes for five different proteins designed NKG2A, NKG2B, NKG2C, NKG2D, and NKG2E. CD94 proteins can be expressed at the NK cell surface as heterodimers with NKG2A and NKG2C (Lazetic et al. 1996; Braud et al. 1998; Cantoni et al. 1998; Vance et al. 1999).

The extracellular domain of CD94 and NKG2A, -B, -C, -D, and -E share high degree of sequence and structure identity (Natarajan et al. 2002). An interesting feature of the CD94/NKG2 system is the fact that depending on which NKG2 chain associates with the CD94 chain, the resulting heterodimeric complex could have inhibitory or activating function. For example, NKG2A has an ITIM motif in its intracellular domain. Thus, the heterodimer CD94/NKG2A functions as an inhibitory receptor on NK cells (Houchins et al. 1997; Palmieri et al. 1999; Guma et al. 2006; Kusumi et al. 2006). On the other hand, the heterodimer CD94/NKG2C serves an activating function because this receptor system in order to transduce intracellular signals need to associate with ITAM-containing signaling molecule DAP-12 (Lanier et al. 1998; Tomasello et al. 2000; Wilson et al. 2000; Sjolin et al. 2002). An exception of this heterodimeric association is NKG2D whose C-type lectin-like protein chains show little homology to other NKG2 proteins and does not pair with CD94. CD94 could be expressed at the cell surface as a type-II disulfide-linked heterodimer (Carretero et al. 1997; Lanier et al. 1998; Vance et al. 1998, 1999). However, it appears that the CD94 homodimer is not functional because CD94 lacks a cytoplasmatic domain and does not have the capacity to bind other signaling transducing polypeptides.

25.4.1 3D Structure of CD94

The structure of CD94 (Figure 25.7A and B) (Boyington et al. 1999) revealed that the CTLD fold is conserved in this receptor. However, distinct features characterize the CTLD of CD94 from the classic CTLD of the MBP-A (Weis et al. 1991). First, CD94 lack the conserved amino acids that are responsible for the coordination of the Ca ions that are essential for carbohydrate binding. Therefore, CD94 function does not involve carbohydrate binding. Second, CD94 (Figure 25.7A and B) lacks the conserved α2-helix, which is observed in MBP-A, Ly49, and NKG2D (Figures 25.4 and 25.7). In CD94, a long unstructured loop region replaces the α2-helix. As it is for the α2-helix in Ly49 and NKG2D dimers (Figures 25.4 and 25.7), this loop makes extensive intramolecular contacts with its symmetry-related counterpart. Four disulfide bridges are present in the CD94 (Figure 25.7A). Three of these disulfide bonds (Cys61–Cys72, Cys89–Cys174, and Cys152–Cys166) represent the invariant disulfides found in the long-form C-type lectin (Day 1994; Sawichi et al. 2001; Zelensky and Gready 2005). The fourth disulfide bridge (Cys61–Cys72) is unique and is required for the correct folding of the N-terminal loop (Figure 25.7A).

As previously discussed (see Section 25.2), CD94 exists at the NK cell surface as disulfide-linked heterodimeric complex with NKG2A, NKG2B, NKG2C, or NKG2E (Lazetic et al. 1996; Braud et al. 1998; Cantoni et al. 1998; Vance et al. 1999). Although nonfunctional, CD94 homodimer was shown to be present at the surface of certain cell lines in which the expression of any NKG2 is absent (Carretero et al. 1997; Lanier et al. 1998; Vance et al. 1998, 1999). In the crystal (Boyington et al. 1999), CD94 monomers form a twofold crystallographic dimer (Figure 25.7A), in which the two monomers are linked through extensive hydrogen bond interaction between their respective first β-strands. In addition, a significant role in maintaining the interface is mediated by loop 3 (Figure 25.7A and B), which replaces the canonical α2-helix observed in other CTLD structures (Figures 25.4 and 25.7).

After the publication of the CD94 structure, a lot of work was put forward in order to crystallize the heterodimer CD94/NKG2. Despite many efforts, at present, there is no available structural data. Nevertheless, an interesting use of the crystallographic CD94 dimer is to model the functional heterodimer CD94/NKG2 (Boyington et al. 1999). If the postulated CD94/NKG2 complex is similar to the CD94 dimer observed in the crystal, then NKG2 monomer needs to have the same loop-to-helix transformation as was observed in the CD94 structure (Figure 25.7). However, this was shown not to be true by the determination of the three-dimensional structure of NKG2D (Figure 25.7C and D; Wolan et al. 2001).

25.4.2 Function and 3D Structure of NKG2D

The discovery that NKG2D plays a pivotal role during tumor eradication has generated a terrific interest in the biology of this NK cell receptor (Lanier 2005). NKG2D is not related to the other NKG2 NK receptors and shares only 20% of amino acid sequence identity with the other NKG2 members (Ho et al. 1998; Wu et al. 1999). As described before, all the NKG2 members form disulfide-linked heterodimers at the NK cell surface with CD94. In contrast, NKG2D does not associate with CD94, but is able to form disulfide-linked homodimers at the cell membrane of the NK cells (Wu et al. 1999). NKG2D activates the NK cell function by associating with a signaling adaptor molecule DAP-10 through a charged transmembrane residue (Wu et al. 1999). Several orthologs of NKG2D were found in chimpanzee, rhesus monkey, cattle, rat, pig, and mice (Lanier 2005).

The 3D-structure of intact murine NKG2D determined at 1.95Å (Wolan et al. 2001) (Figure 25.7C) revealed that that NKG2D preserves the CTLD fold similar to Ly49, CD94, CD69, and MBP-A (Figures 25.4 and 25.7; Natarajan et al. 2002). A characteristic feature of NKG2D structure is that unlike Ly49, CD69, and MBP-A, which retains the α2-helix, the α2-helix in NKG2D is stretched and short. Indeed, this α2-helix is only a one-turn helix (Figure 25.7C and D). As for Ly49, CD69, and CD94, NKG2D lacks the classic Ca-binding residues found in MBP-A (Weis et al. 1991;

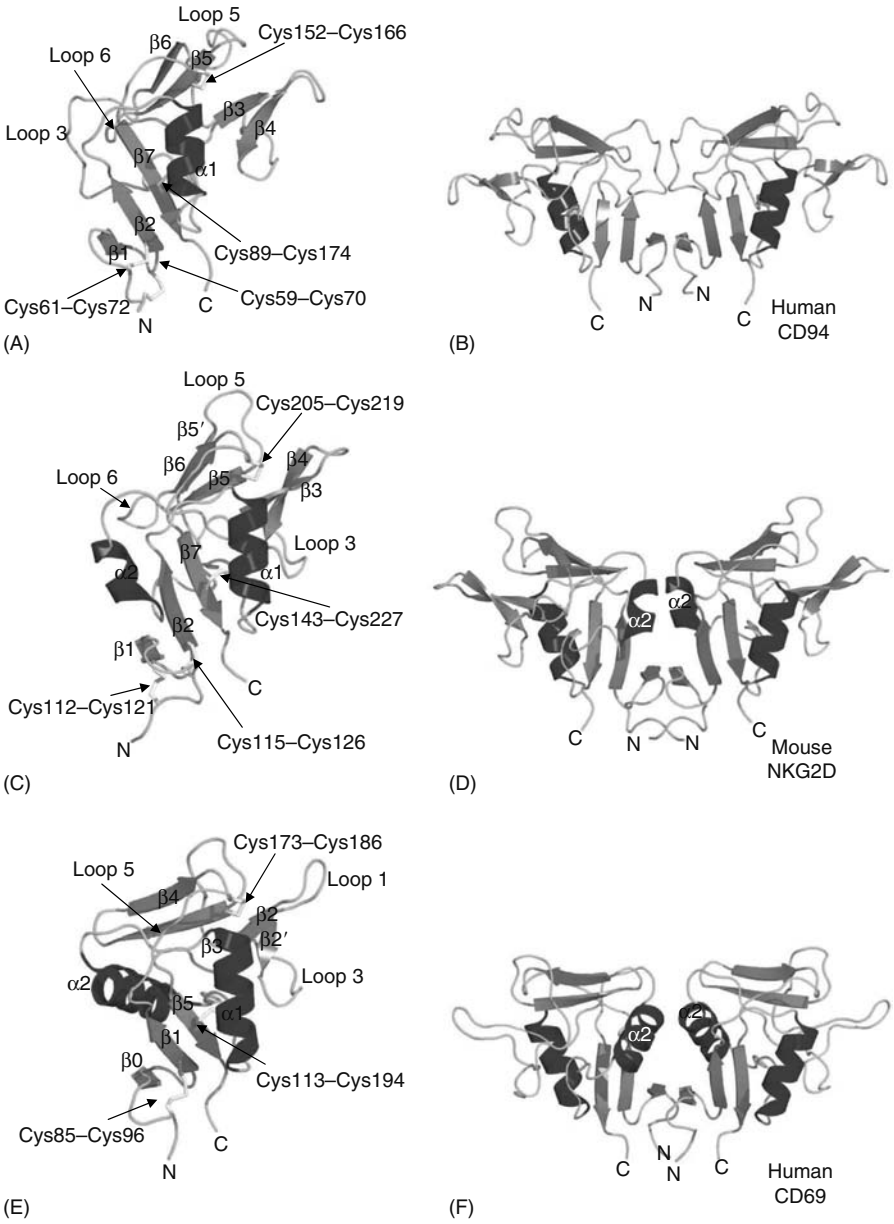


FIGURE 25.7 (See color insert following blank page 170. Also see CD for color figure.) Ribbon representation of human CD94, mouse NKG2D, and human CD69. Monomeric (A) and dimeric (B) structures of human CD94 (PDB entry code 1BE6; From Boyington, J.C., Riaz, A.N., Patamawenu, A., Coligan, J.E., Brooks, A.G., and Sun, P.D., *Immunity*, 1, 75, 1999); monomeric (C) and dimeric (D) structure of mouse NKG2D (PDB entry code 1HQ8; From Wolan, D.W., Teyton, L., Rudolph, M.G., Villmow, B., Bauer, S., Busch, D.H., and Wilson, I.A., *Nat. Immunol*., 3, 248, 2001); and monomeric (E) and dimeric (F) structure of human CD69 (PDB entry code 1E87; From Natarajan, K., Sawicki, M.W., Margulies, D.H., and Mariuzza, R.A., *Biochemistry*, 48, 14779, 2000; Llera, A.S., Viedma, F., Sanchez-Madrid, F., and Tormo, J., *J. Biol. Chem*., 10, 7312, 2001). The secondary structural element (loops, α-helixes, and β-strands), the N and C termini, and the disulfide bridges are shown for all the monomers. The α2 helixes that are involved in the dimer formation in NKG2D and CD69 are labeled. In CD94, the α2 helix is replaced by a long unstructured loop. This loop is buried with its symmetry mate at the dimer interface.

Zelensky and Gready 2005). Four disulfide bridges are present in the NKG2D (Figure 25.7C). Two of these disulfide bridges (Cys143–Cys227 and Cys205–Cys209) are conserved among the CTLD core of Ly49, CD69, and CD94. The disulfide bridge Cys115–Cys126 represents the characteristic disulfide found in the long-form C-type lectin (Day 1994; Sawichi et al. 2001; Zelensky and Gready 2005). The last disulfide bridge, Cys112–Cys121, is a unique feature of NKG2D (Figure 25.7C). This disulfide bridge stabilizes the N-terminal part of the NKG2D through a covalent link between this region with the β-sheet formed by β-strands β1, β2, and β7 (Figure 25.7C).

The NKG2D dimer is very similar to the Ly49A and CD94 dimers (Figures 25.4A, 25.5A, and 25.7D). In this dimer, the homodimer is formed by extensive hydrogen and van der Waals contacts between the β1 strands of both monomers. The α2-helix of both monomers packs at the homodimer interface. The two α2-helixes interact primarily with hydrogen bonds and salt bridges.

25.4.2.1 Ligand Recognition by NKG2D

Different from the CD94/NKG2 heterodimeric complex that recognizes HLA-E, homodimeric NKG2D recognizes cell surface glycoproteins structurally related to MHC class I molecules (Natarajan et al. 2002; Raulet 2003; Lanier 2005). These include MICA, MICB, ULBP1, ULBP2, ULBP3, and ULBP4 in humans and RAE-1α, RAE-1β, RAE-1γ, RAE-1ε, H60, and MULT-1 in mouse (Raulet 2003; Bahram et al. 2005; Lanier 2005; Gonzalez et al. 2006). Human ULBP1, ULBP2, and ULBP3, and mouse RAE-1α, RAE-1β, RAE-1γ, and RAE-1ε are attached to the extracellular site of the cell membrane by a phosphatidylinositol glycan anchor whereas human MICA, MICB, and ULBP4, and mouse H60 and MULT-1 have a membrane spanning the transmembrane domain (Raulet 2003). The expression of these NKG2D ligands is stringently associated with cellular distress, viral infections, and tumors.

Up to now, crystal structures have been reported for human NKG2D in complex with MICA (Figure 25.8A; Li et al. 2001), human NKG2D in complex with ULBP3 (Figure 25.8B; Radaev et al. 2001), and mouse NKG2D in complex with RAE-1β (Figure 25.8C; Li et al. 2002). These complex structures revealed the molecular basis for the remarkable similarity mode used by NKG2D to bind its ligands (Figure 25.8). NKG2D uses its residues at both C- and N-termini of the homodimer interface to bind the α1/α2 platform of its ligands (Figure 25.8). It is of note that although not supported by structural data, functional binding data indicate that this mode of binding could be employed by the Ly49 C-type lectin-like receptors to bind an MHC class I-like molecule (m157) expressed by the cytomegalovirus (Arase et al. 2002; Smith et al. 2002; Voigt et al. 2003; Lanier 2005; Kielczewska et al. 2006).

The overall binding mode of NKG2D/ligands resembles the mode of binding used by KIR-human killer Ig-like receptors (Colonna and Samaridis 1995) to bind the HLA and largely the binding mode between the TCR and their MHC ligands (Rudolph et al. 2006). However, this binding mode is distinct from that used by the Ly49 to bind their MHC ligands (Figure 25.6). In fact, NKG2D uses very similar surfaces of the two monomers to bind two very dissimilar surfaces (α1 and α2) of its ligands, which are distantly related in terms of sequence identity but are similar in their overall 3D structure (Figure 25.8).

NKG2D/complex structures are a remarkable example of molecular recognition via "rigid adaptation" (McFarland and Strong 2003; McFarland et al. 2003). Contrary to the classic "induce fit" mechanism, the rigid adaptation facilitates a receptor like NKG2D to undergo different interactions with distinct ligands without significant conformational changes in the receptor.

25.5 FUNCTION AND 3D STRUCTURE OF CD69

CD69 is one of the first described members of the NK receptors (Hamann et al. 1993; Schnittger et al. 1993; Ziegler et al. 1994). CD69 is expressed as disulfide-linked homodimer at the cell membrane. An extracellular CTLD domain, a stalk region, a transmembrane domain, and an intracellular domain characterize CD69. Contrary to Ly49, CD94, NKG2, and NKG2D whose expression is restricted to NK cells, CD69 expression has been identified on T-lymphocytes, B-lymphocytes, NK cells,

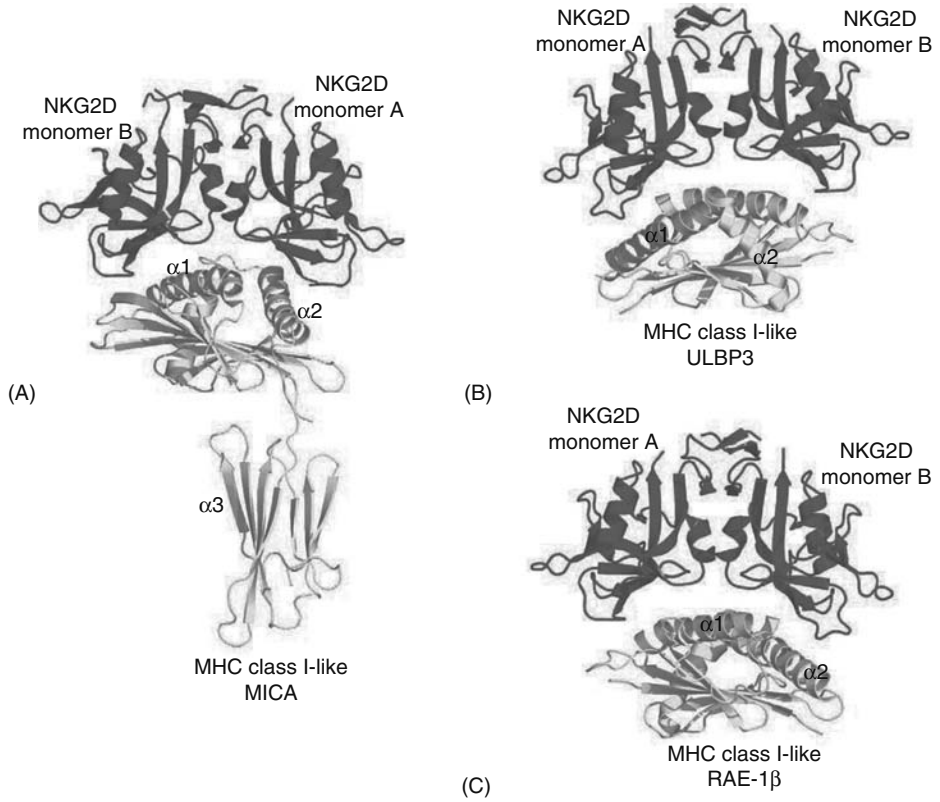


FIGURE 25.8 (See color insert following blank page 170. Also see CD for color figure.) Ribbon diagram showing NKG2D complexed with MHC Class I-like molecules. (A) NKG2D in complex with MICA (PDB entry code 1HYR; From Li, P., Morris, D.L., Willcox, B.E., Steinle, A., Spies, T., and Strong, R.K., *Nat. Immunol*., 5, 443, 2001); (B) with ULBP (PDB entry code 1 KCG; From Radaev, S., Rostro, B., Brooks, A.G., Colonna, M., and Sun, P.D., *Immunity*, 6, 1039, 2001); and (C) with RAE-1β (PDB entry code 1JSK; From Li, P., McDermott, G., and Strong, R.K., *Immunity,* 1, 77, 2002). The two NKG2D monomers are labeled. The secondary structural element α1, α2, and α3 in MICA, the α1 and α2 in ULBP and RAE-1β are schematically labeled. These complex structures unveiled that NKG2D engages similarly its ligands.

macrophages, neutrophils, and eosinophils. The ligand of CD69 is unknown. However, CD69's engagement with monoclonal antibodies at the cell membrane induces rise of intracellular calcium concentration, synthesis of different cytokines with subsequent cellular proliferation and if this occurs on NK cells, lyses of the target cells (Lanier et al. 1988). This implies that CD69 is a functional molecule whose role is the activation of the NK cell function.

Despite relatively low amino acid sequence identity between CD69 and the other CTLD receptors on NK cells, the overall 3D structure of CD69 is very similar to Ly49A, CD94, and NKG2D (Figures 25.4, 25.5, and 25.7; Natarajan et al. 2000; Lleira 2001). The CD69 fold consists of two antiparallel β-sheets, formed by β-strands β0, β1, β5, and β2, β3, β4, respectively, and two α-helixes (Figure 25.7). CD69, like Ly49A, CD94, and NKG2D, lacks the residues that bind calcium in the classic CTLD. Thus, CD69 functions as a noncalcium-dependent lectin.

There are three disulfide bonds in CD69 (Figure 25.7E), two of which (Cys113–Cys194 and Cys173–Cys186) correspond to the characteristic invariant disulfides found in all CTLD. The third disulfide bond, Cys85–Cys96, as for CD94 and NKG2D connects the N-terminus of the protein chain with the strand β0. In the crystal, CD69 forms a compact dimer very similar to the dimer observed for CD94 (Figure 25.7B and F). In the dimer, the N-terminal β0 strands pair through polar

and hydrogen bond interactions (Figure 25.7F). The dimer is further stabilized through the juxtaposition of the C-terminal part of helix α2 (Figure 25.7F), which form a compact hydrophobic core.

25.6 CONCLUSION AND FUTURE DIRECTIONS

The function of NK cells is to recognize and kill tumor or virally infected cells through physical association with ligands expressed at the surface of the target cells. NK cells have evolved a dual receptor system, inhibitory or activating, to sense and ultimately respond to the environment they patrol. An interesting feature of this dual receptor system is the fact that the humans and the mouse use two disparate fold. Human NK cells have NK receptors belonging to the immunoglobulin-like fold, whereas those of the mouse have NK receptors belonging to the C-type lectin-like fold. Both families have inhibitory and activating receptors.

Much effort has been put into dissecting the molecular strategies used by the NK cells to recognize the MHC class I or class I-like scaffolds. These studies revealed that the NK cell receptors of the C-type lectin family interact with several nonoverlapping faces of the class I molecules.

Of special interest between the C-type lectin receptors reviewed in this chapter is NKG2D. This activating receptor has captured a lot of attention since the discovery that the MIC proteins, expressed in transformed cells or in viral infected cells, are the ligands of NKG2D. However, the diversity of the NKG2D ligands presented a major dilemma. Why does a single protein have the ability to recognize diverse ligands? Perhaps NK cells had "captured" a stable scaffold, the C-type lectin, to promptly respond to tumor developments and more importantly to prevent possible pathogen-induced diseases.

ACKNOWLEDGMENT

I apologize to colleagues whose work could not be mentioned in this chapter owing to space limitations and to my inadvertent omission.

REFERENCES

Arase H, Mocarski ES, Campbell AE, Hill AB, and Lanier LL. Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors. *Science* 2002, 296, 1323–1326.

Bahram S, Inoko H, Shiina T, and Radosavljevic M. MIC and other NKG2D ligands: From none to too many. *Curr Opin Immunol* 2005, 5, 505–509.

Beck S and Trowsdale J. The human major histocompatibility complex: Lessons from the DNA sequence. *Annu Rev Genomics Hum Genet* 2000, 1, 117–237.

Berg SF, Dissen E, Westgaard IH, and Fossum S. Two genes in the rat homologous to human NKG2. *Eur J Immunol* 1998, 28, 444–450.

Bontrop RE. Comparative genetics of MHC polymorphisms in different primate species: Duplications and deletions. *Hum Immunol* 2006, 67, 388–397.

Boyington JC, Riaz AN, Patamawenu A, Coligan JE, Brooks AG, and Sun PD. Structure of CD94 reveals a novel C-type lectin fold: Implications for the NK cell-associated CD94/NKG2 receptors. *Immunity* 1999, 1, 75–82.

Braud VM, Allan DS, O'Callaghan CA, Soderstrom K, D'Andrea A, Ogg GS, Lazetic S, Young NT, Bell JI, Phillips JH, Lanier LL, and McMichael AJ. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* 1998, 391, 795–799.

Cantoni C, Biassoni R, Pende D, Sivori S, Accame L, Pareti L, Semenzato G,Moretta L, Moretta A, and Bottino C. The activating form of CD94 receptor complex: CD94 covalently associates with the Kp39 protein that represents the product of the NKG2-C gene. *Eur J Immunol* 1998, 1, 327–338.

Carretero M, Cantoni C, Bellon T, Bottino C, Biassoni R, Rodriguez A, Perez-Villar JJ, Moretta L, Moretta A, and Lopez-Botet M. The CD94 and NKG2-A C-type lectins covalently assemble to form a natural killer cell inhibitory receptor for HLA class I molecules. *Eur J Immunol* 1997, 2, 563–567.

Chambers WH, Adamkiewicz T, and Houchins JP. Type II integral membrane proteins with characteristics of C-type animal lectins expressed by natural killer (NK) cells. *Glycobiology* 1993, 3, 9–14.

Chang C, Rodriguez A, Carretero M, Lopez-Botet M, Phillips JH, and Lanier LL. Molecular characterization of human CD94: A type II membrane glycoprotein related to the C-type lectin superfamily. *Eur J Immunol* 1995, 25, 2433–2437.

Chung DH, Dorfman J, Plaksin D, Natarajan K, Belyakov IM, Hunziker R, Berzofsky JA, Yokoyama WM, Mage MG, and Margulies DH. NK and CTL recognition of a single chain H-2Dd molecule: Distinct sites of H-2Dd interact with NK and TCR. *J Immunol* 1999, 163, 3699–3708.

Colonna M and Samaridis J. Cloning of immunoglobulin-superfamily members associated with HLA-C and HLA-B recognition by human natural killer cells. *Science* 1995, 268, 405–408.

Correa I and Raulet DH. Binding of diverse peptides to MHC class I molecules inhibits target cell lysis by activated natural killer cells. *Immunity* 1995, 2, 61–71.

Dam J, Guan R, Natarajan K, Dimasi N, Chlewicki LK, Kranz DM, Schuck P, Margulies DH, and Mariuzza RA. Variable MHC class I engagement by Ly49 natural killer cell receptors demonstrated by the crystal structure of Ly49C bound to H-2 K(b). *Nat Immunol* 2003, 4, 1213–1222.

Dam J, Baber J, Grishaev A, Malchiodi EL, Schuck P, Bax A, and Mariuzza RA. Variable dimerization of the Ly49A natural killer cell receptor results in differential engagement of its MHC Class I ligand. *J Mol Biol* 2006, 362, 102–113.

Day AJ. The C-type carbohydrate recognition domain (CRD) superfamily. *Biochem Soc Trans* 1994, 1, 83–88.

DeLano WL. *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA, 2002, http://www.pymol.org.

Deng L and Mariuzza RA. Structural basis for recognition of MHC and MHC-like ligands by natural killer cell receptors. *Semin Immunol* 2006, 18, 159–166.

Dimasi N and Biassoni R. Structural and functional aspects of the Ly49 natural killer cell receptors. *Immunol Cell Biol* 2005, 1, 1–8.

Dimasi N, Moretta L, and Biassoni R. Structure of the Ly49 family of natural killer (NK) cell receptors and their interaction with MHC class I molecules. *Immunol Res* 2004, 1, 95–104.

Dimasi N, Sawicki MW, Reineck LA, Li Y, Natarajan K, Margulies DH, and Mariuzza RA. Crystal structure of the Ly49I natural killer cell receptor reveals variability in dimerization mode within the Ly49 family. *J Mol Biol* 2002, 320, 573–585.

Dissen E, Berg SF, Westgaard IH, and Fossum S. Molecular characterization of a gene in the rat homologous to human CD94. *Eur J Immunol* 1997, 27, 2080–2086.

Gao GF, Tormo J, Gerth UC, Wyer JR, McMichael AJ, Stuart DI, Bell JI, Jones EY, and Jakobsen BK. Crystal structure of the complex between human CD8alpha(alpha) and HLA-A2. *Nature* 1997, 387, 630–634.

Gonzalez S, Groh V, and Spies T. Immunobiology of human NKG2D and its ligands. *Curr Top Microbiol Immunol* 2006, 298, 121–138.

Guma M, Budt M, Saez A, Brckalo T, Hengel H, Angulo A, and Lopez-Botet M. Expansion of CD94/NKG2C + NK cells in response to human cytomegalovirus-infected fibroblasts. *Blood* 2006, 107, 3624–3631.

Hamann J, Fiebig H, and Strauss M. Expression cloning of the early activation antigen CD69, a type II integral membrane protein with a C-type lectin domain. *J Immunol* 1993, 150, 4920–4927.

Hanke T, Takizawa H, McMahon CW, Busch DH, Pamer EG, Miller JD, Altman JD, Liu Y, Cado D, Lemonnier FA, Bjorkman PJ, and Raulet DH. Direct assessment of MHC class I binding by seven Ly49 inhibitory NK cell receptors. *Immunity* 1999, 11, 67–77.

Higgins D, Thompson J, Gibson T, Thompson J, Higgins DG, and Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994, 22, 4673–4680.

Ho EL, Heusel JW, Brown MG, Matsumoto K, Scalzo AA, and Yokoyama WM. Murine Nkg2d and Cd94 are clustered within the natural killer complex and are expressed independently in natural killer cells. Proc *Natl Acad Sci USA* 1998, 11, 6320–6325.

Houchins JP, Yabe T, McSherry C, and Bach FH. DNA sequence analysis of NKG2, a family of related cDNA clones encoding type II integral membrane proteins on human natural killer cells. *J Exp Med* 1991, 173, 1017–1020.

Houchins JP, Lanier LL, Niemi EC, Phillips JH, and Ryan JC. Natural killer cell cytolytic activity is inhibited by NKG2-A and activated by NKG2-C. *J Immunol* 1997, 158, 3603–3609.

Kärre K. Role of the target histocompatibility antigens in regulation of natural killer activity: A reevaluation and a hypothesis. In: Callewaert D, Herberman RB, eds. *Mechanisms of Cytotoxicity by NK Cells*. Academic Press, San Diego, CA, 1985, pp. 81–91.

Kelley J, Walter L, and Trowsdale J. Comparative genomics of natural killer cell receptor gene clusters. *PLoS Genet* 2005, 1, 129–139.

Kielczewska A, Kim HS, Lanier LL, Dimasi N, and Vidal SM. Critical residues at the Ly49 Natural Killer receptor's homodimer interface determine functional recognition of m157, a mouse cytomegalovirus MHC class I-like protein. *J Immunol* 2007, 178, 369–377.

Kusumi M, Yamashita T, Fujii T, Nagamatsu T, Kozuma S, and Taketani Y. Expression patterns of lectin-like natural killer receptors, inhibitory CD94/NKG2A, and activating CD94/NKG2C on decidual CD56bright natural killer cells differ from those on peripheral CD56dim natural killer cells. *J Reprod Immunol* 2006, 70, 33–42.

Lanier LL. NK cell recognition. *Annu Rev Immunol* 2005, 23, 225–274.

Lanier LL, Buck DW, Rhodes L, Ding A, Evans E, Barney C, and Phillips JH. Interleukin 2 activation of natural killer cells rapidly induces the expression and phosphorylation of the Leu-23 activation antigen. *J Exp Med* 1988, 5, 1572–1585.

Lanier LL, Corliss B, Wu J, and Phillips JH. Association of DAP12 with activating CD94/NKG2C NK cell receptors. *Immunity* 1998, 8, 693–701.

Lazetic S, Chang C, Houchins JP, Lanier LL, and Phillips JH. Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits. *J Immunol* 1996, 157, 4741–4745.

Li P, McDermott G, and Strong RK. Crystal structures of RAE-1beta and its complex with the activating immunoreceptor NKG2D. *Immunity* 2002, 1, 77–86.

Li P, Morris DL, Willcox BE, Steinle A, Spies T, and Strong RK. Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. *Nat Immunol* 2001, 5, 443–451.

Lian RH, Freeman JD, Mager DL, and Takei F. Role of conserved glycosylation site unique to murine class I MHC in recognition by Ly-49 NK cell receptor. *J Immunol* 1998, 161, 2301–2306.

Lieto LD, Maasho K, West D, Borrego F, and Coligan JE. The human CD94 gene encodes multiple, expressible transcripts including a new partner of NKG2A/B. *Genes Immun* 2006, 7, 36–43.

Ljunggren HG and Kärre K. Host resistance directed selectively against H-2-deficient lymphoma variants. Analysis of the mechanism. *J Exp Med* 1985, 162, 1745–1759.

Llera AS, Viedma F, Sanchez-Madrid F, and Tormo J. Crystal structure of the C-type lectin-like domain from the human hematopoietic cell receptor CD69. *J Biol Chem* 2001, 10, 7312–7319.

Lohwasser S, Hande P, Mager DL, and Takei F. Cloning of murine NKG2A, B and C: Second family of C-type lectin receptors on murine NK cells. *Eur J Immunol* 1999, 29, 755–761.

Lohwasser S, Wilhelm B, Mager DL, and Takei F. The genomic organization of the mouse CD94 C-type lectin gene. *Eur J Immunogenet* 2000, 27, 149–151.

Lohwasser S, Kubota A, Salcedo M, Lian RH, and Takei F. The non-classical MHC class I molecule Qa-1(b) inhibits classical MHC class I-restricted cytotoxicity of cytotoxic T lymphocytes. *Int Immunol* 2001, 13, 321–327.

McFarland BJ and Strong RK. Thermodynamic analysis of degenerate recognition by the NKG2D immunoreceptor: Not induced fit but rigid adaptation. *Immunity* 2003, 19, 803–812.

McFarland BJ, Kortemme T, Yu SF, Baker D, and Strong RK. Symmetry recognizing asymmetry: Analysis of the interactions between the C-type lectin-like immunoreceptor NKG2D and MHC class I-like ligands. *Structure* 2003, 11, 411–422.

Mandal TK and Mukhopadhyay C. Molecular modelling of MHC class I carbohydrates. *Indian J Biochem Biophys* 2001a, 38, 96–103.

Mandal TK and Mukhopadhyay C. Effect of glycosylation on structure and dynamics of MHC class I glycoprotein: A molecular dynamics study. *Biopolymers* 2001b, 59, 11–23.

Matsumoto N, Yokoyama WM, Kojima S, and Yamamoto K. The NK cell MHC class I receptor Ly49A detects mutations on H-2Dd inside and outside of the peptide binding groove. *J Immunol* 2001a, 166, 4422–4428.

Matsumoto N, Mitsuki M, Tajima K, Yokoyama WM, and Yamamoto K. The functional binding site for the C-type lectin-like natural killer cell receptor Ly49A spans three domains of its major histocompatibility complex class I ligand. *J Exp Med* 2001b, 193, 147–158.

Michaelsson J, Achour A, Salcedo M, Kase-Sjostrom A, Sundback J, Harris RA, and Karre K. Visualization of inhibitory Ly49 receptor specificity with soluble major histocompatibility complex class I tetramers. *Eur J Immunol* 2000, 30, 300–307.

Natarajan K, Sawicki MW, Margulies DH, and Mariuzza RA. Crystal structure of human CD69: A C-type lectin-like activation marker of hematopoietic cells. *Biochemistry* 2000, 48, 14779–14786.

Natarajan K, Dimasi N, Wang J, Mariuzza RA, and Margulies DH. Structure and function of natural killer (NK) cell receptors: Multiple molecular solutions to self, non-self discrimination. *Annu Rev Immunol* 2002, 20, 853–885.

Nikolich-Zugich J, Fremont DH, Miley MJ, and Messaoudi I. The role of MHC polymorphism in anti-microbial resistance. *Microbes Infect* 2004, 6, 501–512.

Orihuela M, Margulies DH, and Yokoyama WM. The natural killer cell receptor Ly-49A recognizes a peptide-induced conformational determinant on its major histocompatibility complex class I ligand. *Proc Natl Acad Sci USA* 1996, 93, 11792–11797.

Ortaldo JR and Young HA. Mouse Ly49 NK receptors: Balancing activation and inhibition. *Mol Immunol* 2005, 42, 445–450.

Palmieri G, Tullio V, Zingoni A, Piccoli M, Frati L, Lopez-Botet M, and Santoni A. CD94/NKG2-A inhibitory complex blocks CD16-triggered Syk and extracellular regulated kinase activation, leading to cytotoxic function of human NK cells. *J Immunol* 1999, 162, 7181–718.

Parham P. Functions for MHC class I carbohydrates inside and outside the cell. *Trends Biochem Sci* 1996, 21, 427–433.

Parham P. NK cell receptors: Of missing sugar and missing self. *Curr Biol* 2000, 10, 195–197.

Radaev S, Rostro B, Brooks AG, Colonna M, and Sun PD. Conformational plasticity revealed by the cocrystal structure of NKG2D and its class I MHC-like ligand ULBP3. *Immunity* 2001, 6, 1039–1049.

Raulet DH. Roles of the NKG2D immunoreceptor and its ligands. *Nat Rev Immunol* 2003, 10, 781–790.

Rudolph MG, Stanfield RL, and Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* 2006, 24, 419–466.

Sawichi MW, Dimasi N, Natarajan K, Wang J, Margulies DH, and Mariuzza RA. Structural basis of MHC class I recognition by natural killer cell receptors. *Immunol Rev* 2001, 181, 52–65.

Schnittger S, Hamann J, Dannenberg C, Fiebig H, Strauss M, and Fonatsch C. Regional sublocalization of the human CD69 gene to chromosome bands 12p12.3–p13.2, the predicted region of the human natural killer cell gene complex. *Eur J Immunol* 1993, 23, 2771–2773.

Smith HR, Heusel JW, Mehta IK, Kim S, Dorner BG, Naidenko OV, Iizuka K, Furukawa H, Beckman DL, Pingel JT, Scalzo AA, Fremont DH, and Yokoyama WM. Recognition of a virus-encoded ligand by a natural killer cell activation receptor. *Proc Natl Acad Sci USA* 2002, 99, 8826–8831.

Sjolin H, Tomasello E, Mousavi-Jazi M, Bartolazzi A, Karre K, Vivier E, and Cerboni C. Pivotal role of KARAP/DAP12 adaptor molecule in the natural killer cell-mediated resistance to murine cytomegalovirus infection. *J Exp Med* 2002, 195, 825–834.

Takei F, Brennan J, and Mager DL. The Ly49 family: Genes, proteins and regulation of class I MHC. *Immunol Rev* 1997, 155, 67–77.

Tomasello E, Cant C, Buhring HJ, Vely F, Andre P, Seiffert M, Ullrich A, and Vivier E. Association of signal-regulatory proteins beta with KARAP/DAP-12. *Eur J Immunol* 2000, 30, 2147–2156.

Tormo J, Natarajan K, Margulies DH, and Mariuzza RA. Crystal structure of a lectin-like natural killer cell receptor bound to its MHC class I ligand. *Nature* 1999, 402, 623–631.

Trinchieri G. Biology of natural killer cells. *Adv Immunol* 1989, 47, 187–376.

Vance RE, Kraft JR, Altman JD, Jensen PE, and Raulet DH. Mouse CD94/NKG2A is a natural killer cell receptor for the nonclassical major histocompatibility complex (MHC) class I molecule Qa-1(b). *J Exp Med* 1998, 188, 1841–1848.

Vance RE, Jamieson AM, and Raulet DH. Recognition of the class Ib molecule Qa-1(b) by putative activating receptors CD94/NKG2C and CD94/NKG2E on mouse natural killer cells. *J Exp Med* 1999, 190, 1801–1812.

Voigt V, Forbes CA, Tonkin JN, Degli-Esposti MA, Smith HR, Yokoyama WM, and Scalzo AA. Murine cytomegalovirus m157 mutation and variation leads to immune evasion of natural killer cells. *Proc Natl Acad Sci USA* 2003, 100, 13483–13488.

Waldenstrom M, Achour A, Michaelsson J, Rolle A, and Karre K. The role of an exposed loop in the alpha2 domain in the mouse MHC class I H-2Dd molecule for recognition by the monoclonal antibody 34-5-8S and the NK-cell receptor Ly49A. *Scand J Immunol* 2002, 55, 129–139.

Wang J, Whitman MC, Natarajan K, Tormo J, Mariuzza RA, and Margulies DH. Binding of the natural killer cell inhibitory receptor Ly49A to its major histocompatibility complex class I ligand. Crucial contacts include both H-2Dd and beta 2-microglobulin. *J Biol Chem* 2002, 277, 1433–1442.

Weis WI, Kahn R, Fourme R, Drickamer K, and Hendrickson WA. Structure of the calcium-dependent lectin domain from a rat mannose-binding protein determined by MAD phasing. *Science* 1991, 254, 1608–1615.

Westgaard IH, Berg SF, Orstavik S, Fossum S, and Dissen E. Identification of a human member of the Ly-49 multigene family. *Eur J Immunol* 1998, 28, 1839–1846.

Wilhelm BT, Gagnier L, and Mager DL. Sequence analysis of the ly49 cluster in C57BL/6 mice: A rapidly evolving multigene family in the immune system. *Genomics* 2002, 80, 646–661.

Willcox BE, Thomas LM, and Bjorkman PJ. Crystal structure of HLA-A2 bound to LIR-1, a host and viral major histocompatibility complex receptor. *Nat Immunol* 2003, 4, 913–919.

Wilson MJ, Lindquist JA, and Trowsdale J. DAP12 and KAP10 (DAP10)-novel transmembrane adaptor proteins of the CD3zeta family. *Immunol Res* 2000, 22, 21–42.

Wolan DW, Teyton L, Rudolph MG, Villmow B, Bauer S, Busch DH, and Wilson IA. Crystal structure of the murine NK cell-activating receptor NKG2D at 1.95 Å. *Nat Immunol* 2001, 3, 248–254.

Wu J, Song Y, Bakker AB, Bauer S, Spies T, Lanier LL, and Phillips JH. An activating immunoreceptor complex formed by NKG2D and DAP10. *Science* 1999, 285, 730–732.

Yabe T, McSherry C, Bach FH, Fisch P, Schall RP, Sondel PM, and Houchins JP. A multigene family on human chromosome 12 encodes natural killer-cell lectins. *Immunogenetics* 1993, 37, 455–460.

Yokoyama WM and Plougastel BF. Immune functions encoded by the natural killer gene complex. *Nat Rev Immunol* 2003, 3, 304–316.

Yokoyama WM, Jacobs LB, Kanagawa O, Shevach EM, and Cohen DI. A murine T lymphocyte antigen belongs to a supergene family of type II integral membrane proteins. *J Immunol* 1989, 143, 1379–1386.

Zelensky AN and Gready JE. The C-type lectin-like domain superfamily. *FEBS J* 2005, 272, 6179–6217.

Ziegler SF, Levin SD, Johnson L, Copeland NG, Gilbert DJ, Jenkins NA, Baker E, Sutherland GR, Feldhaus AL, and Ramsdell F. The mouse CD69 gene. Structure, expression, and mapping to the NK gene complex. *J Immunol* 1994, 152, 1228–1236.

26 Activation of Lepidopteran Insect Innate Immune Responses by C-Type Immulectins

Xiao-Qiang Yu and Michael R. Kanost

CONTENTS

26.1 INTRODUCTION

C-type lectins can be classified into groups based on their domain architectures and structural features (Drickamer, 1993; Drickamer and Dodd, 1999; Dodd and Drickamer, 2001; Drickamer and Fadden, 2002; Zelensky and Gready, 2004). Most vertebrate C-type lectins contain a single carbohydrate recognition domain (CRD) connected to other types of domains or motifs (Drickamer, 1993). For example, mannose-binding lectin (MBL), also called mannose-binding protein (MBP), a member of the collectin subgroup, contains a single CRD connected to a collagen-like domain (Holmskov et al., 2003). Some mammalian C-type lectins, such as the macrophage mannose receptor, contain multiple CRDs, but not all of the CRDs are functional for carbohydrate binding (East and Isacke, 2002). In invertebrates, including insects, multiple genes encoding C-type lectin like domains (CTLDs) have been identified in the nematode *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster*, and the mosquito *Anopheles gambiae* (Drickamer and Dodd, 1999; Dodd and Drickamer, 2001; Christophides et al., 2002). C-type lectins have also been isolated from other insect and arthropod species. Invertebrate C-type CRDs have an overall low similarity in amino acid sequence (less than 30% identity) to mammalian C-type

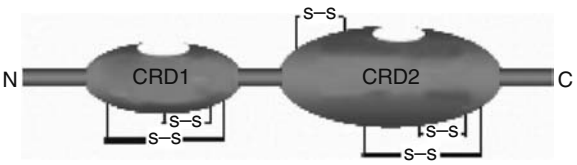


FIGURE 26.1 (See CD for color figure.) Schematic diagram of an IML. An IML contains tandem CRDs. Each CRD may bind different carbohydrates. The amino-terminal short CRD1 is stabilized by two pairs of disulfide bonds, while the carboxyl-terminal long CRD2 has three pairs of disulfide bonds.

CRDs. A group of C-type lectins from arthropods containing dual or tandem CRDs has been named immulectins (IMLs). These dual-CRD lectins have been found in several Lepidopteran species, including *Manduca sexta* (Yu et al., 1999, 2005, 2006; Yu and Kanost, 2000), *Bombyx mori* (accession number: AY297159; Koizumi et al., 1999; Watanabe et al., 2006), *Hyphantria cunea* (Shin et al., 1998), *Helicoverpa armigera* (accession number: ABF83203), *Lonomia oblique* (accession number: AAV91450), and the Chinese shrimp *Fenneropenaeus chinensis* (Liu et al., 2007). The unique dual-CRD architecture has not been found in Dipteran insects such as *D. melanogaster* and *A. gambiae*, or mammals. But genes encoding dual-CRD C-type lectins are present in *C. elegans* and the fish *Fugu rubripes* (Drickamer and Dodd, 1999; Zelensky and Gready, 2004). Members of the lepidopteran immulectin family contain an amino-terminal short form CRD1, which is structurally similar to the CRD of MBL and stabilized by two pairs of disulfide bonds, and a carboxyl-terminal long form CRD2 that is more similar to CRD4 of the macrophage mannose receptor and contains three pairs of disulfide bonds (Yu and Kanost, 2000) (Figure 26.1).

Innate immunity is mediated by pattern recognition receptors (PRRs) encoded by germline cells (Janeway and Medzhitov, 2002). Some mammalian C-type lectins serve as PRRs (Weis et al., 1998; Vasta et al., 1999; Holmskov et al., 2003). For example, MBL is a major PRR in the mammalian innate immune system (Holmskov et al., 2003), because it can bind to high-mannose carbohydrates present on the surface of many microorganisms, including bacteria, fungi, and viruses, to directly enhance phagocytosis or activate the complement system through the lectin pathway (Jack et al., 2001; Holmskov et al., 2003; Jack and Turner, 2003). In insects, β-1,3-glucan recognition proteins (βGRPs) (Kim et al., 2000; Ma and Kanost, 2000; Ochiai and Ashida, 2000; Jiang et al., 2004), thioester containing proteins (TEPs) (Blandin and Levashina, 2004), peptidoglycan recognition proteins (PGRPs) (Steiner, 2004), and C-type lectins (Yu et al., 2002; Kanost et al., 2004) are major PRRs in innate immune recognition. In this chapter, we summarize the activation of immune responses by *M. sexta* immulectins.

26.2 *MANDUCA SEXTA* IMMULECTINS IN INNATE IMMUNITY

26.2.1 Properties of Immulectins

M. sexta immulectins are a group of 32–38 kDa glycoproteins, which are secreted into the hemolymph (Table 26.1). Two isoforms of immulectin-1 (IML-1) are present in hemolymph, with calculated masses (based on native polyacrylamide gel electrophoresis [PAGE] analysis) of 36 and 37.5 kDa (Yu et al., 1999). Analysis of recombinant IML-1 by gel filtration chromatography indicates that it forms oligomers. Two isoforms of IML-2 are also identified in hemolymph with molecular masses of 35,381 and 36,240 Da as determined by mass spectrometry (Yu and Kanost, 2000). The amino-terminal sequences of the two IML-2 isoforms are identical (up to 10 residues

TABLE 26.1
Properties of *M. sexta* Immulectins

	Number of Residues in Mature Protein	Number of *N*-glycosylation Sites	Theoretical pI	Calculated Mass (kDa) of Mature Protein
IML-1	291	1 (Asn^{237})	4.97	33.37
IML-2	308	1 (Asn^{253})	6.07	34.83
IML-3	292	1 (Asn^{134})	6.27	31.98
IML-4	300	2 (Asn^{58}, Asn^{228})	6.23	33.60

determined by Edman degradation), and both isoforms are glycoproteins confirmed by deglycosylation of the purified proteins with *N*-glycosidase F. IML-2 is present mainly as a monomer in solution with lower amounts of dimers and higher order oligomers. Polyclonal rabbit antiserum to recombinant IML-3 cross-reacts with at least three isoforms in hemolymph, which are converted to one smaller protein after treatment with *N*-glycosidase F (Yu et al., 2005), consistent with a hypothesis that the IML isoforms arise from difference in N-linked glycosylation. IML-4 is a 300-residue protein after removal of a signal peptide and contains two potential N-linked glycosylation sites (Yu et al., 2006).

Among the four *M. sexta* immulectins, IML-1 is 26%–27% identical to IML-2, IML-3, and IML-4, while IML-2 is 38% identical to IML-3 and 35% to IML-4. IML-3 and IML-4 are most similar with 56% identity in amino acid sequence, and rabbit polyclonal antiserum to recombinant IML-3 can cross-react with IML-4 (Yu et al., 2006). In mammalian C-type lectins, a tripeptide sequence within the CRD is critical for carbohydrate specificity (Weis et al., 1992; Zelensky and Gready, 2005). CRDs with a tripeptide of Glu-Pro-Asn are predicted to bind glucose/mannose, while a tripeptide of Gln-Pro-Asp is for galactose. In the carboxyl-terminal CRD2 of immulectins, a tripeptide of Glu-Pro-Asn with specificity for glucose/mannose is present in IML-1, IML-2, and IML-4, while IML-3 has a tripeptide of Gln-Pro-Asp with predicted specificity for galactose. However, neither conserved tripeptide is found in the amino-terminal CRD1 of the four immulectins. The amino-terminal and carboxyl-terminal CRDs within each *M. sexta* IML share only ~20% identity. Phylogenetic analysis of individual CRDs of *M. sexta* immulectins and other members of the IML family shows that the CRD1s are clustered in a group, which is distant from the group that contains CRD2s (Figure 26.2). This result suggests that the tandem CRDs diverged prior to an ancient gene duplication in the Lepidopteran lineage that gave rise to the tandem CRD structure of immulectins. CRD of rat MBP-A and CRD4 of the macrophage mannose receptor are placed in the group with the CRD2s, suggesting that mammalian CRDs and CRD2s of immulectins may have evolved from the same ancestor. However, *B. mori* multibinding protein (BmMBP) (Watanabe et al., 2006) is an exception, as both of its CRDs cluster with the CRD2 group, suggesting that its evolutionary history differs from most immulectins.

26.2.2 Agglutinating Activity of Immulectins

M. sexta immulectins agglutinate microorganisms and some animal erythrocytes. Recombinant IML-1, IML-3, and IML-4 agglutinate Gram-negative *Escherichia coli*, Gram-positive *Staphylococcus aureus*, and yeast (*Saccharomyces cerevisiae*) in a concentration- and calcium-dependent manner

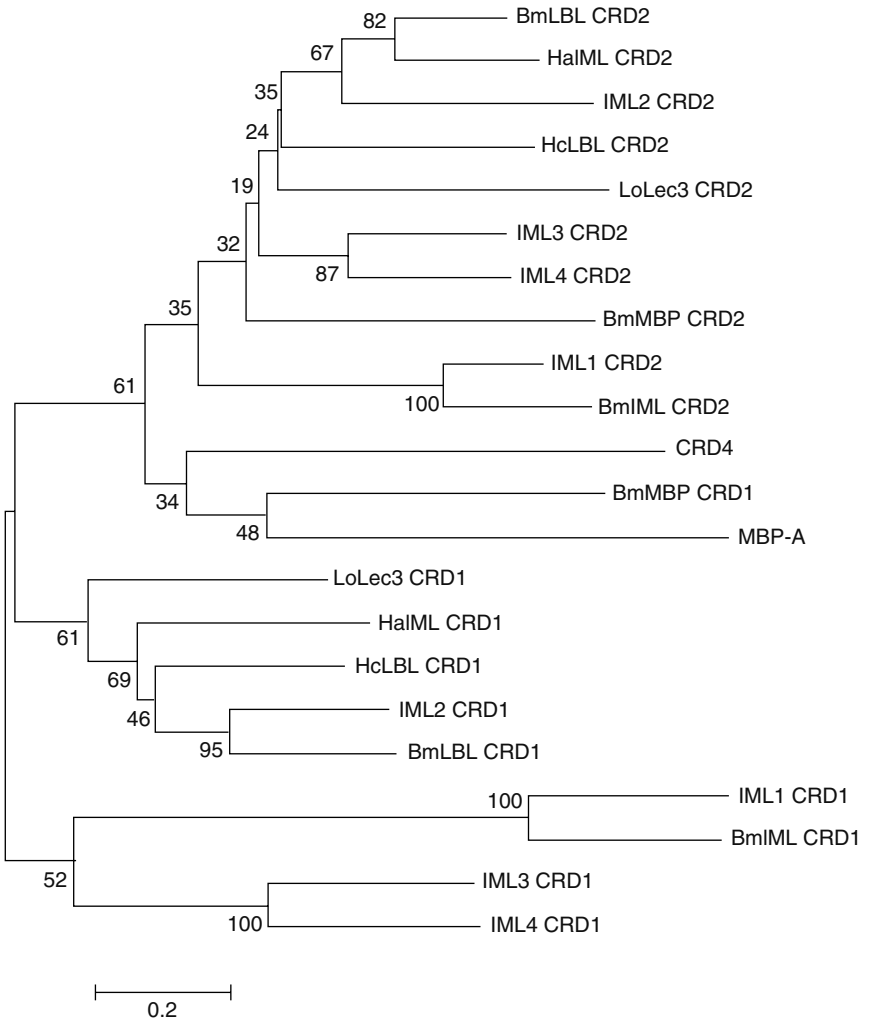


FIGURE 26.2 Phylogenetic analysis of individual CRDs of the IML family. Phylogenetic analysis of individual amino-terminal CRD1s (107–113 residues) and carboxyl-terminal CRD2s (147–154 residues) of insect IML members was conducted using MEGA version 3.1 by the neighbor-joining method. (From Kumar, S., Tamura, K., and Nei, M., *Brief. Bioinformatics*, 5, 150, 2004.) IML-1, -2, -3, and -4: *M. sexta* IML-1, -2, -3, and -4 (accession numbers: AF053131, AF242202, AY768811, and AY768812); BmIML: *B. mori* IML (accession number: AY297159); BmLBL: *B. mori* LPS-binding lectin (accession number: AJ011573); BmMBP: *B. mori* multibinding protein (accession number: AB219148); HcLBL: *H. cunea* LPS-binding lectin (accession number: AF023275); HaIML: *H. armigera* lectin (accession number: ABF83203); LoLec3: *L. oblique* lectin 3 (accession number: AAV91450); MBP-A: CRD (131–238 residues) of rat mannose binding protein A (accession number: P19999); CRD4: the CRD4 (621–764 residues) of the macrophage mannose receptor (accession number: P22897).

(Yu et al., 1999, 2005, 2006). Native IML-2 purified from hemolymph agglutinates *E. coli* and the agglutinating activity is calcium-dependent (Yu and Kanost, 2000). In addition, IML-2 agglutinates erythrocytes of horse, human group A, B and O, sheep, and rabbit with a minimum agglutinating concentration of 0.6 μg/mL for horse erythrocytes. This agglutinating activity is inhibited by D-xylose and D-glucose and by lipopolysaccharides (LPS) from *E. coli* (Yu and Kanost, 2000).

26.2.3 Induced Expression of Immulectins

mRNAs of all four immulectins are present in the fat body but not in the hemocytes of larvae (Yu et al., 1999, 2005, 2006; Yu and Kanost, 2000). IML-1 mRNA is not detected in the fat body of control larvae (injected with saline), but its synthesis is induced in the fat body after larvae are injected with *E. coli* (Yu et al., 1999). IML-2 mRNA is expressed at a constitutively low level in the fat body of control larvae and is induced to a much higher level in the fat body after injection of *E. coli*, *Micrococcus lysodeikticus*, or *S. cerevisiae* (Yu and Kanost, 2000). Two transcripts of IML-3 are detected in the fat body. A 5 kb transcript is present at a low level in the fat body of control larvae, but disappears after larvae are injected with bacteria. However, a much more abundant 3.1 kb transcript, which is absent in the fat body of control larvae, is induced in the fat body of larvae injected with bacteria (Yu et al., 2005). This 3.1 kb IML-3 transcript may be generated from a change in pre-mRNA splicing after larvae are injected with bacteria. The full-length cDNA clone we obtained for IML-3 is 3.8 kb, including an unusually long (2.8 kb) noncoding region at the 3′ end, which contains eight potential polyadenylation signal sites (Yu et al., 2005). A single IML-4 transcript is present in the fat body of control larvae and increases in amount after larvae are injected with bacteria (Yu et al., 2006).

M. sexta immulectins are secreted into hemolymph. IML-1 is not detectable in hemolymph of control larvae, and its concentration, though elevated after bacterial injection, remains low (less than 1 μg/mL) (Yu et al., 1999). IML-2 is present constitutively at a low concentration in hemolymph of control larvae, with an average of ~18 μg/mL (a range of 3.8–36.5 μg/mL in 36 larvae measured). After injection of *E. coli*, the concentration of IML-2 in hemolymph consistently declines within 2 h postinjection but then increases to the original level by 6 h and continues to increase up to 48 h postinjection. The concentration of IML-2 increases three- to four-fold in hemolymph by 24 h after larvae are injected with Gram-negative bacteria or LPS (Yu and Kanost, 2000). In hemolymph of naive larvae, IML-3 is present at a low concentration. After injection with bacteria, IML-3 concentration in hemolymph drops within 2 h but recovers to the original level by 12–24 h postinjection and slightly increases up to 48 h postinjection (Yu et al., 2005). The concentration of IML-3 in the control larvae (injected with saline) changes in a similar pattern, suggesting that wounding and microbial infections have similar effects on IML-3 concentrations in hemolymph. IML-4 is undetectable in hemolymph of naive larvae, but its concentration in hemolymph increases when larvae are injured (injected with saline) or infected with microorganisms (Yu et al., 2006). IML-2 is also detected on the surface of hemocytes (Ling and Yu, 2006a), while IML-3 is translocated from plasma to the cytoplasm and nucleus of hemocytes upon immune challenge (Yu et al., 2005; E. Ling and X.-Q. Yu, unpublished results).

26.2.4 Binding Specificity of Immulectins

M. sexta immulectins have a broad carbohydrate-binding specificity. IML-1 binds to bacterial LPS, but affinity of IML-1 for monosaccharides, including mannose, glucose, *N*-acetylglucosamine and *N*-acetylgalactosamine, is not detected (Yu et al., 1999). Enzyme-linked immunosorbent assay (ELISA) shows that IML-2 binds to immobilized LPS, lipoteichoic acid (LTA), mannan and laminarin, and protein pull-down assays show that IML-2 also binds to peptidoglycan and curdlan (Yu and Kanost, 2000; Yu and Ma, 2006). IML-2 binds to smooth LPS from five different Gram-negative bacteria (*E. coli* strains 026:B6 and 0111:B4, *Pseudomonas aeruginosa*, *Serratia marcescens*, *Salmonella minnesota*), and to four rough mutants of LPS (R-LPS), which lack the O-specific antigen and different regions of the core carbohydrate, and a smaller amount of IML-2 also binds to lipid A (Yu and Ma, 2006). These results suggest that IML-2 binds to the core carbohydrate and O-specific antigen of LPS, with perhaps lower affinity to lipid A. IML-2 also has affinity for D-xylose and D-glucose, as these monosaccharides inhibit its hemagglutination activity (Yu and Kanost, 2000). IML-3 and IML-4 bind to LPS, LTA and laminarin (Yu et al., 2005, 2006), as well as to peptidoglycan and curdlan (X.-Q. Yu, unpublished results).

Surprisingly, calcium is not required for binding of *M. sexta* immulectins to microbial components. At low IML-2 concentrations (0–50 nM), very little IML-2 binds to LPS from *E. coli* 026:B6 and 0111:B4 in the presence of calcium. Binding of IML-2 to LPS substantially increases when EGTA is present. However, the addition of spermine (a polyamine) significantly increases IML-2 binding to LPS. Binding of IML-2 to LTA and laminarin also increases when EGTA or spermine is added to the binding buffer, but the increase is not as great as that for LPS (Yu and Ma, 2006). Nearly the same amount of LPS or LTA bind to immobilized IML-4 in the presence of calcium or EGTA (Yu et al., 2006). These results suggest that calcium is not required for IML-2 and IML-4 binding to microbial components, and polyamine may increase IML-2's binding activity through water-mediated hydrogen-bonding. This calcium-independent binding seems to be shared by other immulectins, because *H. cunea* LPS-binding lectin, a member of the IML family, and *M. sexta* IML-1 and IML-3 all bind to microbial components in the absence of calcium (Shin et al., 2000; X.-Q. Yu, unpublished results). The unique property of calcium-independent binding of immulectins has not been observed in mammalian C-type lectins, because calcium is a direct ligand for mammalian C-type lectins in binding to carbohydrates (Weis et al., 1992). However, calcium is required for agglutination of microorganisms by immulectins (Yu et al., 1999, 2005, 2006; Yu and Kanost, 2000). This may be because immulectins still bind to calcium, but at a site different from the one for carbohydrate binding. Calcium binding may promote interactions between immulectins to form dimers or oligomers, which then can agglutinate microorganisms (Figure 26.3).

In mammalian C-type lectins such as MBL and the macrophage mannose receptor, calcium protects the CRD from proteinase digestion (Ng et al., 1996; Napper et al., 2001). This is also the case for IML-2 (Yu and Ma, 2006). In the presence of calcium, little IML-2 is digested by trypsin at 200 μg/mL; however, in the presence of EDTA, IML-2 is nearly completely digested by trypsin at 5 μg/mL. These results suggest that IML-2 does bind to calcium. We used isothermal titration calorimetry (ITC) to study direct binding of calcium to IML-2. We determined that recombinant CRD2-II of IML-2 (Yu and Kanost, 2003) binds relatively weakly to calcium, with a binding constant of

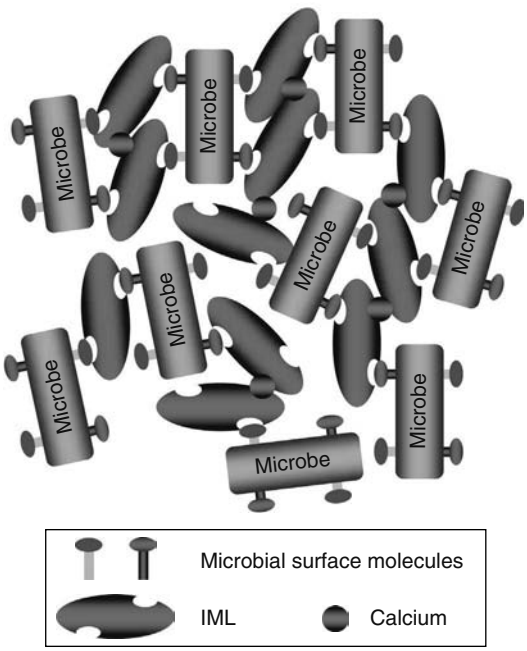


FIGURE 26.3 (See CD for color figure.) Calcium binding enhances agglutinating activity of immulectins. Binding of calcium to immulectins may promote interactions between immulectins to form dimers and oligomers, which then can agglutinate microorganisms by cross-binding to microbial surface molecules.

~7×10^3 M^{-1} (X.-Q. Yu, unpublished results). Calcium binding apparently causes conformational changes in IML-2 such that it resists trypsin digestion. To test whether calcium binding to CRD2-II induces a conformational change, we used circular dichroism (CD) to measure the secondary structures of CRD2-II in the presence of increasing concentrations of calcium. We found that the CD spectra of CRD2-II changed significantly when calcium concentrations increased (X.-Q. Yu, unpublished results), indicating that calcium binding to CRD2-II indeed induces a conformational change. A conformational change in IML-2 (after binding to calcium) may form a more compact structure, which is then resistant to proteinase digestion. This property may be particularly important for IML-2 to function as a PRR because microbial infection activates many proteinases in hemolymph (Kanost et al., 2004), which might otherwise digest IML-2.

26.2.5 IML-1 and IML-2 in Prophenoloxidase Activation

Activation of prophenoloxidase (proPO) in the hemolymph is mediated by a serine proteinase cascade, which has some similarities to the vertebrate complement system (Ashida and Brey, 1998; Cerenius and Soderhall, 2004). Recognition of pathogen-associated molecular pattern (PAMP) by different PRRs triggers activation of the serine proteinase cascade, leading to the conversion of the proPO zymogen to active phenoloxidase (PO). ProPO is converted to PO by proPO-activating proteinases (PAPs) (Jiang et al., 1998, 2003a,b). However, active PAP alone cannot convert proPO to functional PO, but requires serine proteinase homologs (SPHs) as essential cofactors (Jiang et al., 2003a,b; Yu et al., 2003). SPHs contain an amino-terminal clip domain and a carboxyl-terminal serine proteinase-like domain (Yu et al., 2003). The serine residue in the catalytic triad of the proteinase-like domain is mutated to a glycine. Thus, SPHs lack enzymatic activity, but may still retain ability to bind protein substrates. Binding of SPHs to proPO may facilitate its cleavage by PAP at the correct site to generate active PO.

IML-1 and IML-2 bind LPS to stimulate proPO activation in plasma (Yu et al., 1999; Yu and Kanost, 2000). For example, the addition of purified IML-2 alone or IML-2/mannan complex to plasma does not lead to activation of proPO. However, addition of IML-2/LPS complex results in significant activation of proPO within 10 min, and PO activity continues to increase up to 40 min. PO activity is barely observed when LPS alone is incubated with the plasma for 40 min (Yu and Kanost, 2000). Experimental results indicate that IML-1 and IML-2 bind to PAMPs like LPS and trigger activation of the serine proteinase cascade, leading to proPO activation. Although IML-2 has affinity for mannan, an IML-2/mannan complex does not activate proPO, suggesting that the selective interaction of IML-2 with LPS stimulates activation of the cascade.

Formation of a protein complex that includes immulectins and SPHs may target proPO activation to microbial surfaces. SPHs are copurified with IML-2, and the recombinant serine proteinase-like domain of SPH1 (Yu et al., 2003) can pull down IML-2, proPO, and PAP1 in hemolymph. Thus, a protein complex composed of IML-2, SPHs, proPO and PAP may form in hemolymph. Such a complex may localize proPO activation to the surface of pathogens because of IML-2's activity as a PRR. Localization of active PO at the surface of a pathogen would result in high concentration of PO products near the target of the immune response and limit systematic PO activation, which would harm the insect host. In addition, elevated local concentrations of IML-2 on a pathogen surface may enhance recruitment of other hemolymph proteins to stimulate additional humoral or cellular responses.

26.2.6 Immulectins Promote Phagocytosis and Encapsulation

Insect cellular immune responses are mediated by hemocytes (Lavine and Strand, 2002). As PRRs, immulectins can enhance hemocyte-mediated phagocytosis and encapsulation. We have demonstrated that different *M. sexta* hemocyte types have distinct functions in phagocytosis of foreign particles and apoptotic cells. Microsphere beads are phagocytosed primarily by plasmatocytes, whereas only granulocytes phagocytose apoptotic cells (Ling and Yu, 2006a). IML-2 is present on

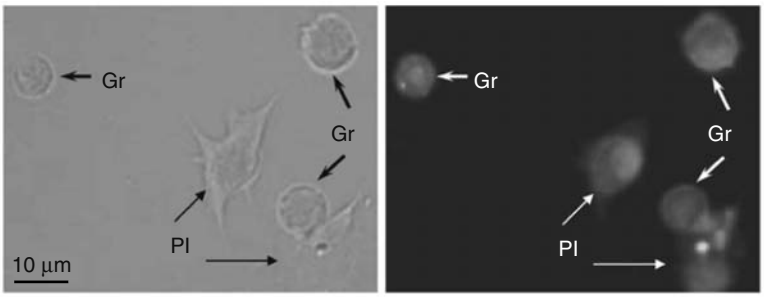


FIGURE 26.4 (See color insert following blank page 170. Also see CD for color figure.) IML-2 is detected on the surface of hemocytes. Hemocytes were collected from the fifth instar *M. sexta* larvae and fixed with paraformaldehyde on a slide. IML-2 on the surface of hemocytes was detected by rabbit polyclonal antibody to IML-2 followed by fluorescein isothiocyanate (FITC)-labeled goat-anti-rabbit IgG, and nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI). (A): A bright phase view of hemocytes. (B): The same hemocytes in (A) viewed by fluorescence microscopy. Green = IML-2; blue = nuclei; Pl = plasmatocytes; Gr = granulocytes.

the surface of both plasmatocytes and granulocytes (Figure 26.4), and it selectively distributes to plasmatocytes or granulocytes depending upon the immune status (Ling and Yu, 2006a). In naive larvae, IML-2 binds to the surface of granulocytes, but after immune challenge, more IML-2 is bound to plasmatocytes. Coupling of IML-2 to microsphere beads enhances their phagocytosis by hemocytes (Ling and Yu, 2006a). Thus, IML-2 may act as an opsonin to enhance phagocytic activity of hemocytes due to its interactions with both pathogens and hemocytes (Figure 26.5).

Cellular encapsulation is a complex process that requires participation of plasma proteins and hemocytes to form a multilayered capsule of hemocytes surrounding a parasite or other nonself object too large to be phagocytosed. In a Lepidopteran species, the soybean looper *Pseudoplusia incudens*, capsule formation begins with adhesion of granulocytes to the foreign surface, followed

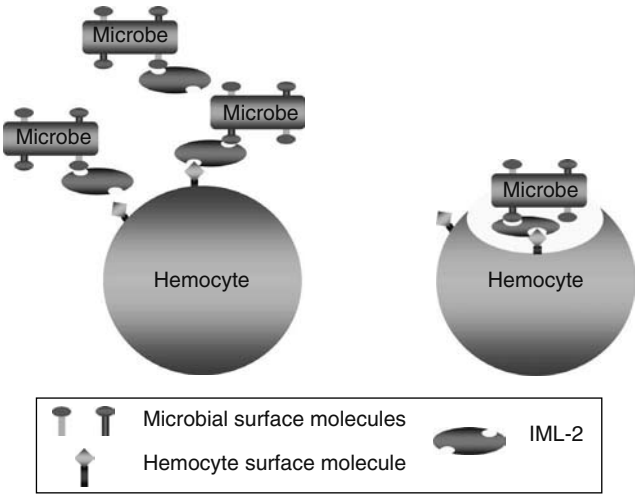


FIGURE 26.5 (See color insert following blank page 170. Also see CD for color figure.) IML-2 acts as an opsonin to enhance phagocytic activity of hemocytes. IML-2 is a PRR that can bind to microbes. Since IML-2 also binds to hemocytes, it can enhance phagocytosis of the IML-2-bound microbes by attracting the microbes to hemocyte surface.

by multiple layers of plasmatocytes, and terminates with a single layer of granulocytes (Pech and Strand, 1996). In *M. sexta*, immulectins can enhance cellular encapsulation (Yu et al., 2005, 2006; Ling and Yu, 2006b). In *in vitro* encapsulation assays, almost 100% of agarose beads coated with any one of the four *M. sexta* immulectins are encapsulated by hemocytes. In contrast, less than 10% of agarose beads coated with a control protein (CP36, an *M. sexta* cuticle protein) (Suderman et al., 2003) are encapsulated (Ling and Yu, 2006b). Preincubation of the IML-2-coated beads with IgG to IML-2 inhibits their encapsulation, further indicating that hemocytes are attracted to IML-2-coated surface. It appears that immulectins serve as PRRs to recognize and bind to the surface of pathogens and then hemocytes are recruited to the IML-bound surface to initiate the encapsulation process (Figure 26.6). Immulectins may also participate in hemocyte–hemocyte interactions through binding to hemocyte surface glycoproteins to enhance hemocyte capsule formation.

Cellular encapsulation is often accompanied by melanization, a process that requires PO activity or activation of proPO associated with hemocytes or the parasite surface. We have observed that IML-2 and IML-4 enhance both encapsulation and melanization of the beads, while IML-1 and IML-3 enhance encapsulation but not melanization (Yu et al., 2005, 2006; Ling and Yu, 2006b). Insect proPO is synthesized in oenocytoids in *M. sexta* larvae (Jiang et al., 1997), and released into hemolymph either by rupture of hemocytes or by exocytosis, as proPO does not contain a secretion signal. ProPO binds to the surface of granulocytes and spherule cells from naive *M. sexta* larvae (Ling and Yu, 2005). When larvae are injured or immune challenged, hemocyte surface proPO is activated to PO, and the number of hemocytes with surface proPO/PO increases, as does the number of melanized hemocytes (Ling and Yu, 2005).

Activation of hemocyte surface proPO may initiate melanization of these hemocytes, which then leads to the systemic melanization of hemocyte capsules. ProPO may bind to hemocytes through IML-2, because IML-2 is detected on the surface of granulocytes (Ling and Yu, 2006a), and it can directly interact with proPO/PO (Yu and Kanost, 2004). Active PO on the surface of hemocytes may recruit more hemolymph PO to hemocytes since PO tends to aggregate (Ashida and Brey, 1998).

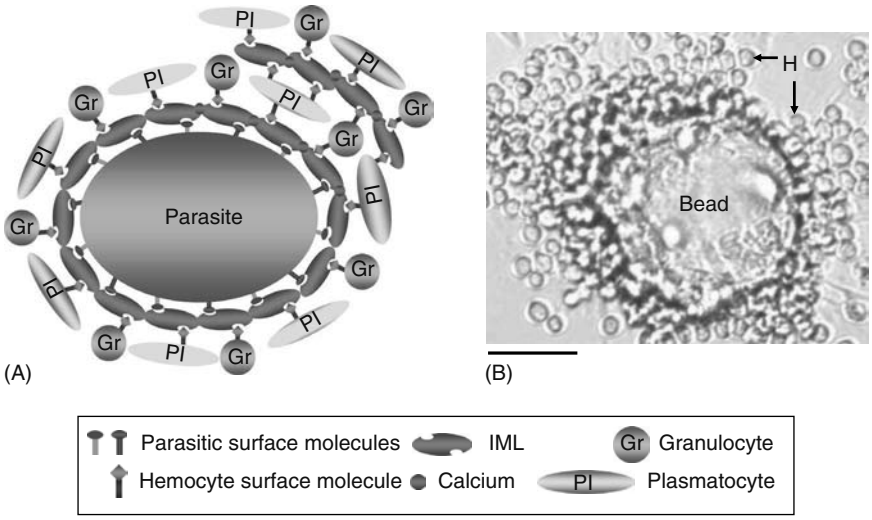


FIGURE 26.6 (See CD for color figure.) Immulectins promote cellular encapsulation. (A) Cartoon presentation of immulectins in cellular encapsulation. Immulectins bind to a parasite through its surface molecules and recruit hemocytes to the IML-coated parasite. Then hemocyte surface immulectins recruit more hemocytes to the parasite to promote formation of hemocyte capsules. (B) An agarose bead coated with recombinant IML-4 is encapsulated by *M. sexta* hemocytes. H = hemocytes. Bar is 50 μm.

26.2.7 IML-2 in Protection of Larvae from Bacterial Infection

IML-2 is a key PRR that participates in both cellular and humoral immune responses in *M. sexta*. Immunodepletion of endogenous IML-2 in hemolymph inhibits clearance of a Gram-negative bacterium, *S. marcescens*, and decreases larvae survival after *S. marcescens* infection (Yu and Kanost, 2003). Furthermore, RNAi silencing of IML-2 gene expression markedly decreases the ability of *M. sexta* larvae to withstand infection by the insect pathogen *Photorhabdus*. IML-2 knockdown causes the greatest reduction in host resistance to infection among the three PRR genes tested (IML-2, hemolin, and PGRP) (Eleftherianos et al., 2006a). The decreased resistance to infection in the IML-2 knockdown larvae is associated with reduced hemolymph PO activity, a result consistent with the idea that IML-2 mediates formation of a protein complex required for proPO activation (Yu et al., 2003). IML-2 can also bind to the surface of a nematode, *C. elegans*, and coating of nematodes by IML-2 enhances melanization of the worms in the hemocoel of *M. sexta* larvae (Yu and Kanost, 2004). It is apparent that IML-2 recognizes a wide variety of pathogens, including Gram-negative bacteria, and even some eukaryotic parasites such as nematodes. Prior infection of *M. sexta* with the nonpathogenic bacterium *E. coli* elicits effective immunity against subsequent infection by *Photorhabdus luminescens* TT01. This protective "immunization" is less effective when IML-2 is knocked down by RNAi (Eleftherianos et al., 2006b).

26.3 CONCLUSION AND FUTURE DIRECTIONS

M. sexta immulectins serve as PRRs in innate immunity. They have a broad binding ability, and thus can recognize a variety of microorganisms. To enhance cellular immune responses such as phagocytosis or encapsulation, immulectins must also interact with hemocyte surface molecules or receptors. So far, we have little information about hemocyte surface molecules or receptors with which immulectins can interact. One candidate may be integrins, since integrins are surface molecules involved in cell adhesion. A hemocyte specific β-integrin in *M. sexta* has been shown to participate in the encapsulation process (Levin et al., 2005). Among the four *M. sexta* immulectins, only IML-2 contains an RGD (Arg-Gly-Asp) motif, which can be recognized by some integrins. IML-2 might also interact with integrins or other cell adhesion proteins through their carbohydrate chains, as many of them are glycoproteins, and thus increase hemocyte–hemocyte interaction to promote encapsulation.

In proPO activation, immulectins interact with plasma proteins including proteinases (PAPs) and cofactors (SPHs) as shown by formation of a protein complex composed of IML-2, SPHs, proPO, and PAP1. This brings up two questions: (1) Which domains of IML-2 participate in these interactions? (2) Does the same or a different domain interact with the different plasma proteins? At this moment, detailed structural information is lacking for immulectins. In future work, identification of regions in immulectins that interact with different proteins and determination of three-dimensional structures are needed, which will provide knowledge for the design of mutagenesis experiments.

ACKNOWLEDGMENT

This work was supported by the National Institutes of Health Grant GM66356.

REFERENCES

Ashida, M. and Brey, P.T., 1998. Recent advances on the research of the insect prophenolxidase cascade. In: *Molecular Mechanisms of Immune Responses in Insects*. Brey P. T. and Hultmark D. eds., Chapman and Hall, London, United Kingdom, pp. 135–172.

Blandin, S. and Levashina, E.A., 2004. Thioester-containing proteins and insect immunity. *Mol. Immunol*. 40, 903–908.

Cerenius, L. and Soderhall, K., 2004. The prophenoloxidase-activating system in invertebrates. *Immunol. Rev.* 198, 116–126.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Muller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von, M.C., Xu, J., Zheng, L., Bork, P., and Kafatos, F.C., 2002. Immunity-related genes and gene families in *Anopheles gambiae. Science* 298, 159–165.

Dodd, R.B. and Drickamer, K., 2001. Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity. *Glycobiology* 11, 71R–79R.

Drickamer, K., 1993. Evolution of Ca^{2+}-dependent animal lectins. *Prog. Nucleic Acid Res. Mol. Biol.* 45, 207–232.

Drickamer, K. and Dodd, R.B., 1999. C-type lectin-like domains in *Caenorhabditis elegans*: Predictions from the complete genome sequence. *Glycobiology* 9, 1357–1369.

Drickamer, K. and Fadden, A.J., 2002. Genomic analysis of C-type lectins. *Biochem. Soc. Symp.* 69, 59–72.

Eleftherianos, I., Millichap, P.J., Ffrench-Constant, R.H., and Reynolds, S.E., 2006a. RNAi suppression of recognition protein mediated immune responses in the tobacco hornworm *Manduca sexta* causes increased susceptibility to the insect pathogen *Photorhabdus. Dev. Comp. Immunol.* 30, 1099–1107.

Eleftherianos, I., Marokhazi, J., Millichap, P.J., Hodgkinson, A.J., Sriboonlert, A., Ffrench-Constant, R.H., and Reynolds, S.E., 2006b. Prior infection of *Manduca sexta* with non-pathogenic *Escherichia coli* elicits immunity to pathogenic *Photorhabdus luminescens*: Roles of immune-related proteins shown by RNA interference. *Insect Biochem. Mol. Biol.* 36, 517–525.

East, L. and Isacke, C.M., 2002. The mannose receptor family. *Biochim. Biophys. Acta* 1572, 364–386.

Holmskov, U., Thiel, S., and Jensenius, J.C., 2003. Collections and ficolins: Humoral lectins of the innate immune defense. *Annu. Rev. Immunol.* 21, 547–578.

Jack, D.L. and Turner, M.W., 2003. Anti-microbial activities of mannose-binding lectin. *Biochem. Soc. Trans.* 31, 753–757.

Jack, D.L., Klein, N.J., and Turner, M.W., 2001. Mannose-binding lectin: Targeting the microbial world for complement attack and opsonophagocytosis. *Immunol. Rev.* 180, 86–99.

Janeway, C.A. and Medzhitov, R., 2002. Innate immune recognition. *Annu. Rev. Immunol.* 20, 197–216.

Jiang, H., Wang, Y., Ma, C., and Kanost, M.R., 1997. Subunit composition of pro-phenol oxidase from *Manduca sexta*: Molecular cloning of subunit ProPO-P1. *Insect Biochem. Mol. Biol.* 27, 835–850.

Jiang, H., Wang, Y., and Kanost, M.R., 1998. Pro-phenol oxidase activating proteinase from an insect, *Manduca sexta*: A bacteria-inducible protein similar to *Drosophila* easter. *Proc. Natl Acad. Sci. USA* 95, 12220–12225.

Jiang, H., Wang, Y., Yu, X.-Q., and Kanost, M.R., 2003a. Prophenoloxidase-activating proteinase-2 from hemolymph of *Manduca sexta*. A bacteria-inducible serine proteinase containing two clip domains. *J. Biol. Chem.* 278, 3552–3561.

Jiang, H., Wang, Y., Yu, X.-Q., Zhu, Y., and Kanost, M.R., 2003b. Prophenoloxidase-activating proteinase-3 (PAP-3) from *Manduca sexta* hemolymph: A clip-domain serine proteinase regulated by serpin-1J and serine proteinase homologs. *Insect Biochem. Mol. Biol.* 33, 1049–1060.

Jiang, H., Ma, C., Lu, Z.Q., and Kanost, M.R., 2004. Beta-1,3-glucan recognition protein-2 (betaGRP-2) from *Manduca sexta*; an acute-phase protein that binds beta-1,3-glucan and lipoteichoic acid to aggregate fungi and bacteria and stimulate prophenoloxidase activation. *Insect Biochem. Mol. Biol.* 34, 89–100.

Kanost, M.R., Jiang, H., and Yu, X.-Q., 2004. Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunol. Rev.* 198, 97–105.

Kim, Y.S., Ryu, J.H., Han, S.J., Choi, K.H., Nam, K.B., Jang, I.H., Lemaitre, B., Brey, P.T., and Lee, W.J., 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. *J. Biol. Chem.* 275, 32721–32727.

Koizumi, N., Imamura, M., Kadotani, T., Yaoi, K., Iwahana, H., and Sato, R., 1999. The lipopolysaccharide-binding protein participating in hemocyte nodule formation in the silkworm *Bombyx mori* is a novel member of the C-type lectin superfamily with two different tandem carbohydrate-recognition domains. *FEBS Lett.* 443, 139–143.

Kumar, S., Tamura, K., and Nei, M., 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinformatics* 5, 150–163.

Lavine, M.D. and Strand, M.R., 2002. Insect hemocytes and their role in immunity. *Insect Biochem. Mol. Biol.* 32, 1295–1309.

Levin, D.M., Breuer, L.N., Zhuang, S., Anderson, S.A., Nardi, J.B., and Kanost, M.R., 2005. A hemocyte-specific integrin required for hemocytic encapsulation in the tobacco hornworm, *Manduca sexta*. *Insect Biochem. Mol. Biol.* 35, 369–380.

Ling, E. and Yu, X.-Q., 2005. Prophenoloxidase binds to the surface of hemocytes and is involved in hemocyte melanization in *Manduca sexta*. *Insect Biochem. Mol. Biol.* 35, 1356–1366.

Ling, E. and Yu, X.-Q., 2006a. Hemocytes from the tobacco hornworm *Manduca sexta* have distinct functions in phagocytosis of foreign particles and self dead cells. *Dev. Comp. Immunol.* 30, 301–309.

Ling, E. and Yu, X.-Q., 2006b. Cellular encapsulation and melanization are enhanced by immulectins, pattern recognition receptors from the tobacco hornworm *Manduca sexta*. *Dev. Comp. Immunol.* 30, 289–299.

Liu, Y.C., Li, F.H., Dong, B., Wang, B., Luan, W., Zhang, X.J., Zhang, L.S., and Xiang, J.H., 2007. Molecular cloning, characterization and expression analysis of a putative C-type lectin (Fclectin) gene in Chinese shrimp *Fenneropenaeus chinensis*. *Mol. Immunol.* 44, 598–607.

Ma, C. and Kanost, M.R., 2000. A β-1,3-glucan-recognition protein from an insect, *Manduca sexta*, agglutinates microorganisms and activates the phenoloxidase cascade. *J. Biol. Chem.* 275, 7505–7514.

Napper, C.E., Dyson, M.H., and Taylor, M.E., 2001. An extended conformation of the macrophage mannose receptor. *J. Biol. Chem.* 276, 14759–14766.

Ng, K.K., Drickamer, K., and Weis, W.I., 1996. Structural analysis of monosaccharide recognition by rat liver mannose-binding protein. *J. Biol. Chem.* 271, 663–674.

Ochiai, M. and Ashida, M., 2000. A pattern recognition protein for β-1,3-glucan. *J. Biol. Chem.* 275, 4995–5002.

Pech, L.L. and Strand, M.R., 1996. Granular cells are required for encapsulation of foreign targets by insect haemocytes. *J. Cell Sci.* 109, 2053–2060.

Shin, S.W., Park, S.S., Park, D.S., Kim, M.G., Kim, S.C., Brey, P.T., and Park, H.Y., 1998. Isolation and characterization of immune-related genes from the fall webworm, *Hyphantria cunea*, using PCR-based differential display and subtractive cloning. *Insect Biochem. Mol. Biol.* 28, 827–837.

Shin, S.W., Park, D.S., Kim, S.C., and Park, H.Y., 2000. Two carbohydrate recognition domains of *Hyphantria cunea* lectin bind to bacterial lipopolysaccharides through O-specific chain. *FEBS Lett.* 467, 70–74.

Steiner, H., 2004. Peptidoglycan recognition proteins: On and off switches for innate immunity. *Immunol. Rev.* 198, 83–96.

Suderman, R.J., Andersen, S.O., Hopkins, T.L., Kanost, M.R., and Kramer, K.J., 2003. Characterization and cDNA cloning of three major proteins from pharate pupal cuticle of *Manduca sexta*. *Insect Biochem. Mol. Biol.* 33, 331–343.

Vasta, G.R., Quesenberry, M., Ahmed, H., and O'Leary, N., 1999. C-type lectins and galectins mediate innate and adaptive immune functions: Their roles in the complement activation pathway. *Dev. Comp. Immunol.* 23, 401–420.

Watanabe, A., Miyazawa, S., Kitami, M., Tabunoki, H., Ueda, K., and Sato, R., 2006. Characterization of a novel C-type lectin, *Bombyx mori* multibinding protein, from the *B. mori* hemolymph: Mechanism of wide-range microorganism recognition and role in immunity. *J. Immunol.* 177, 4594–4604.

Weis, W.I., Drickamer, K., and Hendrickson, W.A., 1992. Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* 360, 127–134.

Weis, W.I., Taylor, M.E., and Drickamer, K., 1998. The C-type lectin superfamily in the immune system. *Immunol. Rev.* 163, 19–34.

Yu, X.-Q. and Kanost, M.R., 2000. Immulectin-2, a lipopolysaccharide-specific lectin from an insect, *Manduca sexta,* is induced in response to Gram-negative bacteria. *J. Biol. Chem.* 275, 37373–37381.

Yu, X.-Q. and Kanost, M.R., 2003. *Manduca sexta* lipopolysaccharide-specific immulectin-2 protects larvae from bacterial infection. *Dev. Com. Immunol.* 27, 189–196.

Yu, X.-Q. and Kanost, M.R., 2004. Immulectin-2, a pattern recognition receptor that stimulates hemocyte encapsulation and melanization in the tobacco hornworm, *Manduca sexta*. *Dev. Com. Immunol.* 28, 891–900.

Yu, X.-Q. and Ma, Y., 2006. Calcium is not required for Immulectin-2 binding, but protects the protein from proteinase digestion. *Insect Biochem. Mol. Biol.* 36, 505–516.

Yu, X.-Q., Gan, H., and Kanost, M.R., 1999. Immulectin, an inducible C-type lectin from an insect, *Manduca sexta*, stimulates activation of plasma prophenol oxidase. *Insect Biochem. Mol. Biol.* 29, 585–597.

Yu, X.-Q., Zhu, Y.F., Ma, C., Fabrick, J.A., and Kanost, M.R., 2002. Pattern recognition proteins in *Manduca sexta* plasma. *Insect Biochem. Mol. Biol.* 32, 1287–1293.

Yu, X.-Q., Jiang, H., Wang, Y., and Kanost, M.R., 2003. Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. *Insect Biochem. Mol. Biol.* 33, 197–208.

Yu, X.-Q., Tracy, M.E., Ling, E., Scholz, F.R., and Trenczek, T., 2005. A novel C-type immulectin-3 from *Manduca sexta* is translocated from hemolymph into the cytoplasm of hemocytes. *Insect Biochem. Mol. Biol.* 35, 285–295.

Yu, X.-Q., Ling, E., Tracy, M.E., and Zhu, Y., 2006. Immulectin-4 from the tobacco hornworm Manduca sexta binds to lipopolysaccharide and lipoteichoic acid. *Insect Mol. Biol.* 15, 119–128.

Zelensky, A.N. and Gready, J.E., 2004. C-type lectin-like domains in *Fugu rubripes. BMC Genomics* 5, 51.

Zelensky, A.N. and Gready, J.E., 2005. The C-type lectin-like domain superfamily. *FEBS J.* 272, 6179–6217.

27 Galectins as Novel Regulators of Immune Cell Homeostasis and Inflammation

Gabriel A. Rabinovich, Marta A. Toscano, Juan M. Ilarregui, and Linda G. Baum

CONTENTS

27.1 INTRODUCTION

Galectins are a family of highly conserved glycan-binding proteins with affinity for β-galactoside-containing oligosaccharides [1,2]. Since the discovery of discoidin-1 in the cellular slime mold *Dictyostelium discoideum* [3] and electrolectin in the electric organ tissue of the electric eel [4], the family of galectins has received increasing attention. However, it was only in the 1990s that experimental evidence emerged, illuminating a role for galectins in the regulation of physiological and pathological processes, particularly in the control of immune cell homeostasis and inflammation [5]. Members of the galectin family are defined by a conserved carbohydrate recognition domain (CRD) with a canonical amino acid sequence and affinity for β-galactosides [2,6]. To date, 15 mammalian galectins have

been identified, which can be subdivided into those that have one CRD (prototype) and those that have two CRDs in tandem (tandem-repeat type). In addition, galectin-3, a one-CRD galectin, is unique in that it contains unusual tandem repeats of short amino acid stretches fused onto the CRD (chimera-type) [2,7]. Many galectins bind carbohydrate moieties in a bivalent or multivalent manner. Some one-CRD galectins exist as dimers; two-CRD galectins have two carbohydrate-binding sites, and galectin-3 forms oligomers when it binds to multivalent carbohydrates [8]. Galectin-mediated cross-linking of cell surface glycoconjugates can trigger a cascade of transmembrane signaling and modulate processes that include apoptosis, cytokine secretion, cell adhesion, and migration [8].

Galectins can be found inside and outside cells, and have distinct functions in each location [9]. Whether endogenously expressed, or rapidly internalized from the cell surface, galectins have been implicated in important intracellular functions such as pre-mRNA splicing, regulation of cell growth, cell cycle progression, and protein sorting [9]. Although galectins do not contain signal peptides to direct them through the classical endoplasmic reticulum (ER)-Golgi apparatus secretory system, they can be secreted by other unorthodox secretory pathways [6]. Once outside the cell, galectins bind to and cross-link multiple glycoconjugates found on the cell surface or in the extracellular matrix (ECM) [8].

Although most mammalian galectins bind preferentially to glycoconjugates containing the ubiquitous disaccharide *N*-acetyllactosamine [Galβ1-3GlcNAc or Galβ1-4GlcNAc], binding to individual lactosamine units is of relatively low affinity ($K_d \sim 1$ mM), and arrangement of lactosamine disaccharides in repeating chains (polylactosamine) increases binding avidity. Moreover, detailed structural analysis of the CRDs suggests subtle differences in carbohydrate-binding specificities of individual members of this family [10]. Whether differences in saccharide specificity might be responsible for distinct biological effects in response to individual galectin binding still remains to be elucidated.

Some galectins are distributed in a wide variety of tissues, whereas others have a more restricted localization [11]. Within the immune system, galectins are found in activated macrophages, activated B cells, and activated T cells [12–14]. Recent studies using gene expression arrays have indicated elevated expression of galectin-1 in $CD4^+$ $CD25^+$ regulatory T cells [15,16]. The expression of galectins is modulated during the activation and differentiation of immune cells and may be significantly altered under several pathological conditions [17].

Accumulating evidence has shown that galectins play a role in the initiation, activation, and resolution phases of innate and adaptive immune responses by promoting anti-inflammatory or proinflammatory effects (Figures 27.1 and 27.2). In this regard, it has been recently hypothesized that the same galectin may exert pro- or anti-inflammatory effects depending on multiple factors, such as the concentration reached in inflammatory foci, extracellular microenvironment, and the particular target cells impacted [5]. It has been suggested that multivalency of individual members of the galectin family and their cross-linking properties might determine different biological responses by inducing aggregation of specific cell surface glycoreceptors, which in many cases, are associated with different signal transduction events [18]. We will discuss here the anti-inflammatory (Figure 27.1) and proinflammatory (Figure 27.2) activities of different members of the galectin family on the physiology of different immune cell types. Finally, we will illustrate the impact of these immunoregulatory activities in the development and progression of chronic inflammatory disorders, autoimmunity, and cancer.

27.2 IMPACT OF GALECTINS IN T-CELL PHYSIOLOGY

T-cell homeostasis is attributed to multiple distinct checkpoints that are interconnected and intervene at defined stages of the life cycle of immune cells to guarantee clonal expansion, differentiation, and functional inactivation of antigen-experienced or potentially autoaggressive T lymphocytes [19]. During the past decade, several laboratories have made important contributions to the elucidation of the role galectins

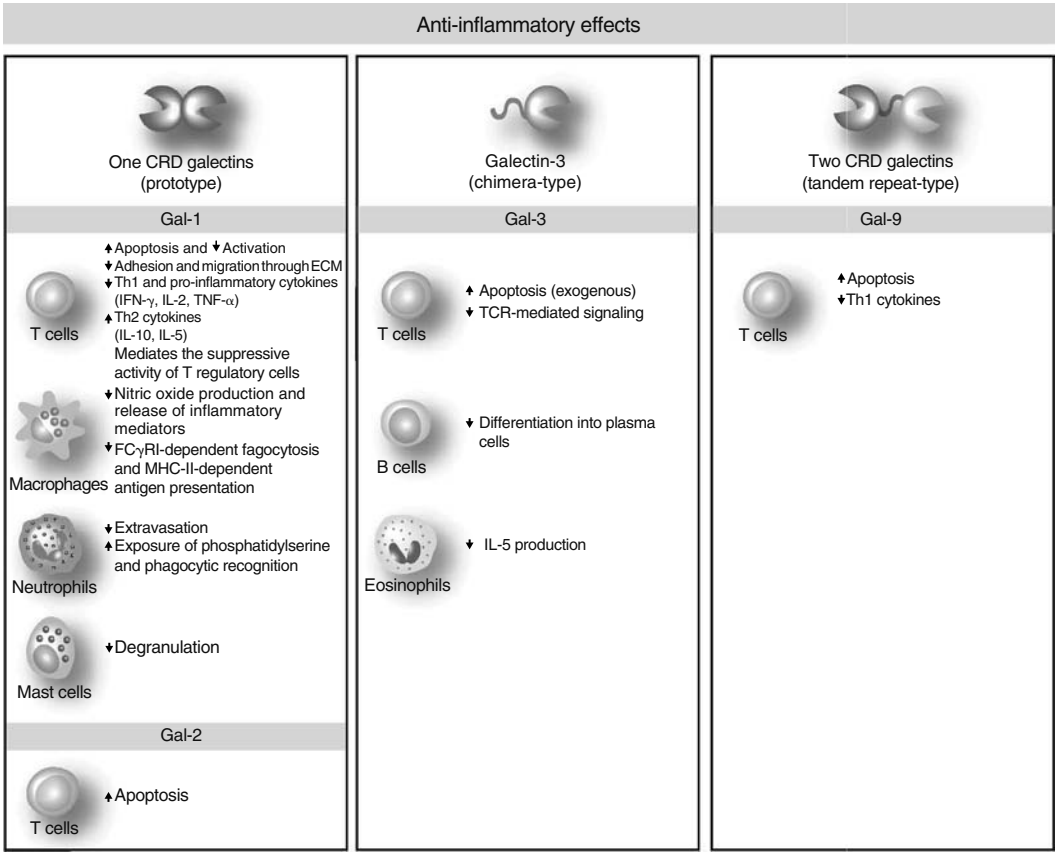


FIGURE 27.1 (See CD for color figure.) Anti-inflammatory effects of galectins on different cell types including T cells, B cells, macrophages, neutrophils, eosinophils and mast cells. FcgRI: Fc receptor type I for immunoglobulin G; MHC: Major histocompatibility complex; TCR: T cell receptor; CRD: Carbohydrate recognition domain.

in T-cell physiology and homeostasis. There is now consistent evidence that cross-linkage of T-cell surface receptors by galectins can trigger different transmembrane signaling events through which diverse processes such as survival, activation, cytokine secretion, and cell migration are modulated [7,8].

27.2.1 Galectins in the Modulation of T-Cell Survival at Central and Peripheral Immune Cell Compartments

A critical process for dampening potential harmful T-cell responses is the finetuning and regulation of T-cell survival. Mounting evidence indicates that galectin-1 can control T-cell proliferation and apoptosis of human and murine T cells during development in the thymus and after stimulation in the periphery [12,13,20–26]. Different cell surface glycoconjugates appear to be primary receptors for galectin-1, such as CD45, CD43, CD2, CD3, and CD7 [22,27–29]. Interestingly, galectin-1 binding to T cells results in a marked redistribution of these glycoproteins into segregated membrane microdomains [28]. Furthermore, it has been demonstrated that CD45 can positively or negatively regulate galectin-1-induced T-cell death depending on its glycosylation status [27]. Hence, $CD45^+$ T cells lacking the core 2 β-1,6-*N*-acetylglucosaminyltransferase (C2GnT), an enzyme responsible for creating branched structures on *O*-glycans of T-cell surface glycoproteins, are resistant to galectin-1-induced death [27,30]. In addition, it has been demonstrated that T-cell susceptibility to galectin-

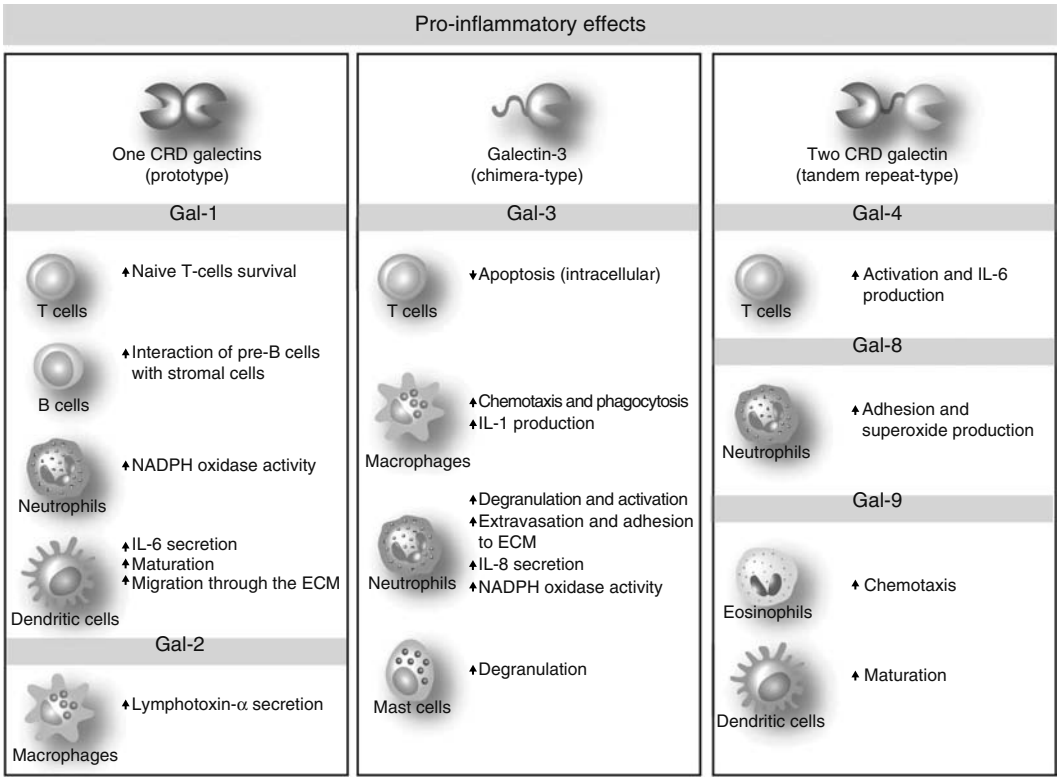


FIGURE 27.2 (See CD for color figure.) Pro-inflammatory effects of galectins on different cell types including T cells, B cells, dendritic cells, macrophages, neutrophils, eosinophils and mast cells. *ECM: extracellular matrix.

1-induced cell death can also be regulated by the ST6Gal I sialyltransferase, which selectively modifies *N*-glycans on CD45 to negatively regulate T-cell death [31]. Therefore, highly significant factors that determine the responsiveness of cells to galectin-mediated signals include the repertoire of potentially glycosylated molecules expressed on the cell surface and the activities of specific glycosyltransferases, which are responsible for generating galectin ligands. These variables can dramatically change according to the developmental stage, state of activation, and differentiation, and pathophysiologic status of lymphoid cells (reviewed in Ref. [32]).

In this regard, CD7 has been identified as a critical receptor for galectin-1-induced apoptosis [22] and it has been shown that $CD4^+$ $CD7^-$ leukemic T cells are protected from galectin-1-triggered T-cell death [33,34]. Interestingly, recent evidence indicates that haploinsufficiency of C2GnT in T-cell lymphoma results in altered cellular glycosylation and resistance to galectin-1-induced apoptosis [35]. These findings may represent a novel escape mechanism displayed by distinct types of lymphoma and leukemia cells in order to survive in galectin-1-enriched microenvironments.

The signal transduction events that lead to apoptosis induced by galectin-1 involve several intracellular mediators of apoptosis, including the induction of specific transcription factors (i.e., AP-1) and modulation of BCL2 protein production [36], activation of caspases and cytochrome *c* release [21], and involvement of proximal signals such as $p56^{lck}$ and ZAP70 [37]. Interestingly, recent evidence indicates that acid sphingomyelinase-mediated release of ceramide is essential to trigger the mitochondrial pathway of apoptosis induced by galectin-1 [38]. In addition, Allione et al. have shown that galectin-1 can influence apoptosis and cell cycle progression in an indirect manner; exposure to this protein upregulates the expression of both the α- and β-chains of interferon-γ

(IFN-γ) receptor on activated T lymphocytes and the increased expression of both chains renders the cells more susceptible to IFN-γ -induced apoptosis [39].

Interestingly, Endharti et al. demonstrated that, in contrast to the proapoptotic role of galectin-1 on activated T cells, secretion of this protein by stromal cells is capable of supporting the survival of naive T cells without promoting proliferation [40]. In addition, recent evidence showed that galectin-1 treatment can result in exposure of phosphatidylserine (an early apoptotic marker involved in the phagocytosis of apoptotic cells) on the plasma membrane of human neutrophils; however, the authors did not find phosphatidylserine exposure or apoptosis in peripheral T cells [41]. Thus, it seems apparent that galectin-1 might trigger different signals (i.e., apoptosis or survival) and even different apoptosis end points (full apoptosis, phosphatidylserine exposure, or none) depending on a number of factors including the activation state of the cells, spatiotemporal expression of specific glycosyltransferases, general context of the cell culture, biochemical properties of purified galectin-1 (monomeric vs. dimeric forms), or the nature of the target cell (activated, differentiated, or resting peripheral T cells). These apparent discrepancies remain to be elucidated in future work demonstrating the potential role of endogenous galectin-1 in the regulation of apoptosis *in vivo* and the effects of different microenvironments in the apoptotic or immunoregulatory activity of this carbohydrate-binding protein.

Galectin-2, another one-CRD member of the galectin family structurally related to galectin-1, also promotes T-cell apoptosis probably via binding to cell surface β-integrins on the surface of T cells [42]. Analysis of the mechanisms involved in the proapoptotic properties of this protein revealed involvement of caspases-3 and -9, enhanced cytochrome *c* release, disruption of the mitochondrial membrane potential, and an increase of the Bax/Bcl-2 ratio [42]. Thus, although through activation of different receptors and intracellular pathways, galectin-2 can also modulate T-cell survival similarly to galectin-1, suggesting that the existence of a conserved galectin-mediated mechanism responsible of regulating T-cell fate.

Galectin-3, the only chimera-type galectin, has been shown to act in a dual manner, either protecting cells from apoptosis or promoting apoptosis depending on whether the protein acts intracellularly [43,44] or whether it is added exogenously [45,46]. Yang et al. demonstrated that T-cell transfectants overexpressing galectin-3 display higher growth rates than control transfectants and are protected from apoptosis induced by a variety of agents including anti-Fas antibodies and staurosporine [44]. Furthermore, Akahani et al. showed that intracellular galectin-3 inhibits apoptosis through a cysteine protease pathway, highlighting the presence in the *gal-3* gene of a NWGR motif, which is highly conserved in the BH1 domain of the *bcl-2* gene family [43]. Interestingly, recent findings indicated that extracellular galectin-3 can signal apoptosis of human T cells through binding to cell surface glycoconjugates such as CD45, CD71, CD7, and CD29 [45,46]. Furthermore, Hahn et al. suggested a functional cross-talk between intracellular and extracellular galectins in the regulation of T-cell death; the authors found that galectin-1-induced cell death is inhibited by intracellular expression of galectin-3 [20]. Interestingly, recent evidence indicated, using galectin-3-deficient mice, that galectin-3 was at least in part responsible of accelerating thymocyte depletion during intracellular infection with the protozoan parasite *Trypanosoma cruzi* [47].

Galectin-9, a two-CRD member of the galectin family, has been shown to induce apoptosis of immature cortical thymocytes [48]. Furthermore, this lectin has also been shown to induce death of peripheral $CD4^+$ and $CD8^+$ T cells through a Ca^{2+}-calpain-caspase-1-mediated pathway [49]. Interestingly, recent evidence indicates that galectin-9-induced apoptosis requires the dimeric form of this protein and that galectin-9 treatment of the Jurkat T-cell line promotes loss of mitochondrial membrane potential and the release of apoptosis-inducing factor (AIF) [50].

Interestingly, recent findings highlight a novel role for galectin-9 as a binding partner of the Th1-specific receptor Tim-3 [51]. In elegant studies, Zhu et al. demonstrated that galectin-9 can trigger apoptosis and immunosuppression through binding to Tim-3 [51]. The authors found that galectin-9 induces intracellular calcium flux, T-cell aggregation, and T-cell death through a Tim-3-dependent mechanism [51].

Although there is still scarce evidence for the participation of other members of the galectin family in immunoregulation, galectins-7, -8, and -12 have been shown to regulate cell survival in other cell types that, under special circumstances, might be involved in immunoregulation. For example, galectin-7 has been shown to increase the susceptibility of keratinocytes to ultraviolet-B-induced apoptosis [52]. Gene profile experiments revealed that galectin-7 functions intracellularly to induce apoptosis upstream of JNK activation and cytochrome *c* release [53]. Furthermore, galectin-8 has been shown to modulate tumor cell survival through binding to β-integrins [54] and galectin-12 has been shown to modulate cell-cycle progression and survival of adipocytes [55]. Whether modulation of cell survival by these galectins may compromise other immunological processes (e.g., skin immunity, adipocyte-immune connections, or tumor immunity) still remains to be elucidated. Targeted disruption of galectin genes in single or double knockout mice will be critical to dissect the impact of endogenous galectins in the regulation of cell fate and survival during the development of pathophysiologic processes.

27.2.2 Galectin-Mediated Control of TCR Signaling and Immunological Synapse

Another critical immunologic checkpoint that controls the decision between immune cell activation or immune cell tolerance is at the level of the T-cell receptor (TCR)-mediated signaling and the immunological synapse.

In this regard, galectin-1 has been shown to affect early TCR-dependent signals during T-cell activation. Vespa et al. found that galectin-1 inhibits TCR-induced interleukin-2 (IL-2) production and proliferation in a murine T-cell hybridoma clone and freshly isolated mouse thymocytes [56]. Interestingly, the same group further demonstrated that galectin-1 antagonizes TCR signals that require costimulation such as IL-2 production, while allowing TCR responses that only require partial TCR signals such as CD69 upregulation and apoptosis. The authors clearly showed that galectin-1 can modulate the immunological synapse by blocking TCR/costimulator-dependent lipid raft clustering at the TCR constant site [57]. The effects of galectin-1 at the level of the immunological synapse might also provide alternative explanations for the ability of galectin-1 to differentially regulate cell death in different cell types.

In addition, Demetriou et al. reported that galectin-3 may play a critical role in restricting TCR complex-initiated signal transduction. The authors hypothesized that galectin-3 might form multivalent complexes with *N*-glycans on the TCR, thereby restraining the lateral mobility of TCR complexes [58]. This effect was abrogated in mice deficient in β1,6-*N*-acetylglucosaminiltransferase (Mgat5), a crucial enzyme in the N-glycosylation pathway. These mice showed enhanced delayed-type hypersensitivity responses and increased susceptibility to autoimmunity. Thus, galectin-3 may influence T-cell interactions with antigen-presenting cells and may control full T-cell activation by negatively regulating the immunological synapse [58].

On the other hand, Hokama et al. demonstrated using *in vivo* and *in vitro* experiments that galectin-4, which is highly expressed by epithelial intestinal cells, favors $CD4^+$ T-cell activation and induces IL-6 production through a PKCθ-dependent pathway [59]. These findings were confirmed in an experimental model of intestinal inflammation, indicating that galectin-4 may function as a positive regulator of T-cell activation *in vivo* [59]. Thus, it seems evident that different members of the galectin family may differentially affect the immunological synapse to positively or negatively regulate T-cell activation.

27.2.3 Galectin Regulation of T-Cell-Derived Cytokines

Although the precise mechanisms still remain to be elucidated, different members of the galectin family have been shown to positively or negatively influence the production of a wide range of anti-inflammatory or proinflammatory cytokines (Figures 27.1 and 27.2). We have reported that recombinant galectin-1, at low concentrations (~0.01–0.1 μM), can inhibit the secretion of proinflammatory cytokines, such as tumor necrosis factor-α (TNF-α) and IFN-γ by activated T cells without inducing

T-cell apoptosis [60]. Interestingly, this concentration is significantly lower than that required for galectin-1 to dimerize (~7 μM). In addition, we have shown that galectin-1 inhibits the allogeneic T-cell response through apoptotic and nonapoptotic mechanisms [26]. Interestingly, in this study, we found selective inhibition of Th1 cytokine production in the viable nonapoptotic T-cell population [26], suggesting that several galectin-1-mediated mechanisms may operate to achieve immunosuppression *in vivo* (see below). Furthermore, van der Leij et al. reported a marked increase in IL-10 mRNA and protein levels in nonactivated and activated $CD4^+$ and $CD8^+$ T cells following exposure to recombinant galectin-1 [61]. In addition, the authors generated leucine-zipper based stable galectin-1 homodimers and recently found that this stable dimeric galectin-1 can efficiently induce apoptosis, increase IL-10, and decrease IL-2 secretion at markedly lower concentrations compared to nonmodified recombinant galectin-1 [61].

However, one of the most consistent findings among the literature is the ability of galectin-1 to skew the balance from a Th1- toward a Th2-polarized immune response in different experimental models of chronic inflammation, autoimmunity, and cancer [62–68]. Investigation of cytokine balance in draining lymph nodes and spleens from mice treated with recombinant galectin-1 revealed decreased levels of IFN-γ and IL-2 production and high levels of IL-5, IL-10, and TGF-β production (Table 27.1). Recent findings provide an explanatory mechanism to explain this Th2 bias, indicating that Th1 and Th17 cells have the repertoire of cell surface glycans that are critical for galectin-1 binding and subsequent cell death, while Th2 cells are protected from galectin-1 through differential sialylation of cell surface glycoproteins [66]. Similar to galectin-1, Sturm et al. found that galectin-2 can also shift the balance toward a Th2 cytokine profile in T-cell culture [42]. Conversely, galectin-3 has been shown to specifically reduce the levels of IL-5 secretion by T-cell lines, suggesting its potential ability to skew the cytokine balance toward a Th1 response [69]. On the other hand, galectin-4, a two-CRD member of the galectin family has been shown to favor IL-6 production through a PKCθ-dependent pathway [59]. Thus, it seems apparent that different galectin members can selectively regulate the pro- or anti-inflammatory cytokine pattern and imprint their individual signatures in the regulation of an ongoing immune response. Whether Th1/Th2 shifts are associated with the ability of different galectin members to specifically target the survival or signaling of selected immune cell types is currently under investigation.

27.2.4 Impact of Galectins in Regulatory T-Cell Responses

Several mechanisms that control the immune system to prevent or minimize damage caused by reactivity to self-antigens and overexuberant immune response to pathogens have been described. In addition to activation-induced cell death and induction of anergy, avoidance of collateral damage to the host is also achieved by active immune suppression mediated by T regulatory cell populations [70]. We found that treatment with recombinant galectin-1 in the efferent phase of autoimmune ocular inflammation results in increased IL-10 and TGF-β production and expansion of $CD4^+$ $CD25^+$ T regulatory cells *in vivo* [68]. Adoptive transfer of regulatory T cells obtained from galectin-1-treated mice prevented the development of autoimmune disease in naive recipient mice [68]. Interestingly, recent studies demonstrated using DNA microarray analysis that galectin-1 is overexpressed in naturally occurring regulatory T cells [16] and that blockade of galectin-1 significantly reduced the suppressive effects of human and mouse $CD4^+$ $CD25^+$ regulatory T cells [15]. These findings indicate that galectin-1 is a key effector of the suppressive activity of T regulatory cells, thus providing another potential mechanism to understand the immunoregulatory activities of this carbohydrate-binding protein.

27.2.5 Galectins in T-Lymphocyte Adhesion and Migration

Adhesion and migration of immune cells across blood-vessel walls and through ECM barriers is instrumental in maintaining homeostasis and competent performance in immunological reactions, such as in inflammation [71]. We found that exposure to galectin-1 inhibited T-cell adhesion to ECM glycoproteins such as fibronectin and laminin [60]. In addition, galectin-1 presented on the

TABLE 27.1
Impact of Galectins in Autoimmunity and Chronic Inflammation

Experimental Models	Strategies Used	Clinical Outcome	Potential Mechanisms Involved	References
Experimental autoimmune myasthenia gravis (EAMG)	• Injection of electrolectin to rabbits	• Complete clinical recovery and delayed onset	• No changes in circulating autoantibodies or modifications at the muscular level	[110]
Experimental autoimmune encephalomyelitis (EAE)	• Prophylactic administration of Gal-1 to MBP-immunized Lewis rats	• Prevention of clinical and histopathological signs of the disease	• N.D. (Blockade of activation of pathogenic T cells?)	[111]
	• Gal-9 injection to MOG-immunized C57BL/6 mice • siRNA *gal-9* to PLP-immunized SJL mice	• Reduced severity and mortality • Increased severity of the disease	• Selective loss of IFN-γ-producing cells. Apoptosis of Tim-3 + Th1 cells	[51]
Collagen-induced arthritis (CIA)	• Gal-1 gene therapy and protein administration to DBA/1 mice	• Suppression of clinical and histopathological manifestations	• Increased IL-5 and decreased IFN-γ production • Increased T-cell susceptibility to activation induced cell death	[64]
Concanavalin A-induced hepatitis	• Prophylactic administration of Gal-1 in BALB/c mice	• Prevention of liver injury and T-helper cell liver infiltration	• Suppressed tumor necrosis factor-α and IFN-γ production • Increased apoptosis of activated T cells	[113]
Inflammatory bowel disease (TNBS-induced colitis)	• Prophylactic and therapeutic administration of Gal-1 in BALB/c mice	• Suppression of clinical and histopathological manifestations	• Reduced ability of mucosal T cells to produce IFN-γ • Increased number of apoptotic T cells within mucosal tissue	[67]
	• Epithelial-derived Gal-4	• Exacerbates intestinal inflammation	• Stimulates IL-6 production by $CD4^+$ T cells	[59]
IRBP-induced experimental autoimmune uveitis (EAU)	• Administration of Gal-1 during the afferent or efferent phase of EAU in B10.RIII mice	• Suppression of ocular inflammatory disease	• Ability to counteract Th1-mediated responses • Promotion of a Th2 and T-regulatory-mediated anti-inflammatory responses	[68]
Nephrotoxic nephritis (induced by antiglomerular basement membrane serum)	• Gal-1, Gal-3, Gal-9 administration to Wistar Kyoto rats	• Clinical recovery	• Gal-9 induces apoptosis of activated $CD8^+$ cells	[112]

TABLE 27.1 (continued)
Impact of Galectins in Autoimmunity and Chronic Inflammation

Experimental Models	Strategies Used	Clinical Outcome	Potential Mechanisms Involved	References
			• Gal-1 and Gal-3 suppress the accumulation of macrophages	
Nephropatic diabetes (model of type 2 diabetes)	• Gal-9 i.p. administration to male *db/db* mice	• Reduced albuminuria, glomerular hypertrophy, and mesangial matrix expansion	• Inhibition of glomerular expression of TGF-β_1 • Reduced cell arrest in the G0/G1 phase	[115]
Graft versus host disease	• Gal-1 administration to mice	• Increased host survival following allogeneic hematopoietic stem cell transplant	• Reduced production of IFN-γ and IL-2 • Reduced alloreactivity	[62]
Type 1 diabetes (induced by adoptive transfer of diabetogenic splenocytes into NOD*rag1*$^{-/-}$ mice)	• Prophylactic administration of DC genetically engineered to synthesize transgenic gal-1	• Delayed onset	• Deletion of β cell-reactive T cells	[63]
OVA-induced asthma	• Therapeutic gal-3 gene therapy to A/J mice	• Improvement in airway hyperresponsiveness and remodeling	• Eosinophil depletion by IL-5 inhibition	[116]
	• Challenge of Gal-3-deficient C57BL/6 mice	• Gal-3$^{-/-}$ mice developed less airway hyper-responsiveness and lower goblet cell metaplasia	• Fewer eosinophils and IgE levels • Lower Th2 response and higher Th1 response	[117]

Note: Gal, galectin; MBP, myelin-basic protein; MOG, myelin oligodendrocyte glycoprotein; DC, dendritic cell; siRNA, small interfering RNA; PLP, myelin proteolipid protein; TNBS, 2,4,6-trinitrobenzene sulfonic acid; IRBP, interphotoreceptor retinoid-binding protein; OVA, ovalbumin; AHR, airway hyperresponsiveness.

surface of the ECM reduced the ability of T cells to migrate through the matrix; this effect required CD43 clustering but was independent of the presence core 2 *O*-glycans [72]. Hence, it is becoming increasingly apparent that apoptosis may only partially explain the immunosuppressive properties of galectin-1; T cells that are refractory to apoptosis may be subject to suppression of proinflammatory cytokine secretion, inhibition of their migratory capacity, and targeting for phagocytic removal.

27.3 IMPACT OF GALECTINS IN B-CELL PHYSIOLOGY

Whereas compelling evidence has been accumulated regarding the effects of galectin-1 on T-cell fate, limited information is available on how galectin-1 may impact on B lymphocytes. A pioneer study by Gauthier et al. demonstrated that galectin-1 expressed by stromal cells acts as a ligand of the pre-B-cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells [73].

The authors found that pre-BCR binding to stromal cells depends upon Gal-1 anchoring to glycosylated counterreceptors and these complexes relocalize at the contact zone to form the immunological synapse [73]. The authors recently extended their findings showing that $\alpha_4\beta_1$ (VLA-4), $\alpha_5\beta_1$ (VLA-5), and $\alpha_4\beta_7$ integrins are the major receptors for galectin-1 during pre-BCR relocalization, activation, and signalling [74]. Thus, galectin-1 is critical during B-cell progenitor development in the bone marrow compartment. At the peripheral level, a significant increase in galectin-1 expression was found in stimulated B cells receiving signals via cross-linking of the BCR and CD40 [14]. Interestingly, no changes in B-cell survival were found following exposure of galectin-1 *in vitro* independently of the activation state of these cells [14]. In search for genes differentially transcribed in anergic B cells, Clark et al. recently found, using representational difference analysis (RDA), that galectins-1 and -3 are significantly overexpressed in anergic B cells tolerized by self-antigens. These findings prompt further exploration of the role for galectins as regulator of B-cell tolerance and homeostasis [75].

Regarding other members of the galectin family, we have provided evidence that galectin-3 is a critical mediator of B-cell differentiation and survival [76]. Antisense-mediated blockade of intracellular galectin-3 abrogated IL-4-induced survival of activated B cells, favoring the differentiation toward antibody-secreting plasma cells. Furthermore, B cells with restrained galectin-3 expression failed to downregulate the *Blimp-1* transcription factor (which is critical for the plasma cell pathway) after IL-4 stimulation [76]. Finally, inhibition of galectin-3 *in vivo* skewed the balance toward plasma cell differentiation, antibody secretion, and parasite clearance during intracellular parasite infection [76]. Taken together, these data indicate that galectins can also modulate B-cell maturation and differentiation both at the central and peripheral immune compartments.

27.4 IMPACT OF GALECTINS IN THE PHYSIOLOGY OF MONOCYTES, MACROPHAGES, AND DENDRITIC CELLS

The decision between activation or tolerance within the monocyte and macrophage compartment is determined by different stimuli which may act in concert to induce five different states of macrophage activation: (1) "innate activation," which is induced by microbial products that are recognized by pattern recognition receptors and is responsible for the production of proinflammatory cytokines; (2) "humoral activation," which is triggered by cross-linking of Fc and complement receptors and is critical for functions such as phagocytosis and secretion of pro- and anti-inflammatory cytokines; (3) "classical activation," which is typically mediated by the priming stimulus IFN-γ and is critical for the microbicidal activity of macrophages and DTH responses; (4) "alternative activation," which is triggered by IL-4 and IL-13, induces arginase expression, and determines the generation of Th2 responses; and (5) "deactivation," which is triggered by anti-inflammatory cytokines and uptake of apoptotic cells and results in MHC-II downregulation and secretion of high levels of IL-10, TGF-β, and PGE_2 [77].

Previous results from our laboratory showed that galectin-1, similarly to Th2 and Th3 cytokines, inhibits nitric oxide synthesis, favoring instead the expression of arginase (the alternative metabolic pathway of L-arginine) in activated peritoneal macrophages [78]. Recently, we have shown that galectin-1 can differentially regulate the expression and function of critical regulatory molecules (i.e., FcγRI and MHC-II) on human monocytes and mouse macrophages through a nonapoptotic ERK1/2-mediated pathway [79]. This effect was clearly observed in macrophages recruited in response to inflammatory stimuli following treatment with recombinant galectin-1 and further confirmed in galectin-1-deficient (*gal-1*$^{-/-}$) mice [79]. This result, together with our previous observations that galectin-1 favors arginase activity [78], suggests that this endogenous lectin might promote a state of "alternative activation" or "deactivation" in elicited macrophages. Consistently, we found in a model of *T. cruzi* infection that galectin-1 induces a biphasic modulation of parasite replication in peritoneal macrophages isolated from infected mice [80]. While low

doses of galectin-1 increased parasite replication and favors IL-10 production (favoring a state of "deactivation" or "alternative activation"), high doses of galectin-1 decreased the number of intracellular amastigotes [65]. On the other hand, other studies suggested that galectin-2 can regulate lymphotoxin-α secretion, which positively affects the degree of macrophage-induced inflammation during myocardial infarction [81].

Regarding the impact of galectin-3 in monocyte and macrophage functions, Liu's group made a major contribution. In a pioneer work, they showed that targeted disruption of the galectin-3 gene results in attenuated peritoneal inflammatory infiltrates [82]. In addition, the authors showed that peritoneal macrophages from galectin-3-deficient mice were more prone to undergo apoptosis than those from galectin-3-sufficient mice [82]. More recently, the same group reported a critical role for galectin-3 in phagocytosis by macrophages, as galectin-3-deficient cells exhibited reduced phagocytic capacity [83]. In addition, they demonstrated that galectin-3 promotes chemotaxis of human monocytes through a Pertussis toxin-sensitive G-protein-mediated pathway [84]. Furthermore, Bernardes et al. reported the presence of a significantly reduced number of infiltrating macrophages following infection of galectin-3-deficient mice with the *Toxoplasma gondii* protozoan parasite [85]. Therefore, it seems apparent that galectin-3 plays a critical proinflammatory role within the monocyte/macrophage compartment.

Recent evidence also indicates that macrophages differently sense the fungi *Candida albicans* and *Saccharomyces cerevisiae* through a mechanism involving Toll-like receptor (TLR)-2 and galectin-3 [86]. Thus, it seems that galectin-3 serves as a receptor for different bacteria, fungi and parasites as has also been demonstrated for *Leishmania major* [87] and *Neisseria gonorrhoeae* [88].

Concerning the impact of galectins in dendritic cell function, Lee's group reported the ability of galectin-1 to augment the secretion of proinflammatory cytokines and to influence dendritic cell migration through the ECM [89,90]. On the other hand, it has been demonstrated that dendritic cells engineered to overexpress galectin-1 are highly activated; these transgenic cells can stimulate naive T cells and induce apoptosis of activated T cells [91], consistent with findings described in previous sections. In addition, Hirashima's group has been demonstrated that galectin-9 promotes the maturation of monocyte-derived dendritic cells at similar levels as lipopolysaccharides [92], and Kiss' group showed that galectin-3 is upregulated in dendritic cells following *T. cruzi* infection, with critical implications for adhesion and migration of these cells [93].

27.5 ROLE OF GALECTINS IN NEUTROPHIL-MEDIATED INFLAMMATORY RESPONSES

The acute inflammatory response involves the release of soluble proinflammatory mediators, but is self-limiting and is resolved through the release of endogenous anti-inflammatory products and the clearance of inflammatory cells. Galectin-1 has been shown to attenuate the acute inflammatory response [94,95]. We have shown that galectin-1 ameliorates phospholipase A_2-induced edema and blocks neutrophil extravasation *in vivo* [95]. In addition, La et al. found that galectin-1 inhibits transendothelial migration and chemotaxis of neutrophils [94]. Furthermore, Stowell et al. reported that galectins-1, -2, and -4 can induce phosphatidylserine exposure in activated human neutrophils without affecting their survival [41].

In contrast, galectin-3 has been shown to act in most cases as a "proinflammatory cytokine." Galectin-3 null mutant mice exhibit an attenuated inflammatory response after intraperitoneal injection of thioglycollate, and have significantly reduced recovery of granulocytes compared to wild-type animals [82,96]. Interestingly, Karlsson and colleagues showed that both galectin-1 and galectin-3 are able to activate NADPH oxidase in primed neutrophils [97,98]. Furthermore, galectin-3 induces neutrophil adhesion to laminin [99] and to endothelial cells [100] and promotes IL-8 secretion [101], suggesting a critical role for this protein in the initiation of innate immune

responses. In this context, it has been recently shown that galectin-3 acts in concert with soluble fibrinogen to regulate neutrophil activation, degranulation, and survival through alternative activation of mitogen-activated protein kinases (ERK1/2 or p38)-mediated pathways [102]. In addition, a recent study reports a fine regulation of galectin-1 and galectin-3 expression during the early and late phases of inflammation in a model of rat peritonitis [103]. Furthermore, recent studies shed light to the only immunoregulatory role for galectin-8 reported to date, which involves the positive regulation of neutrophil functions associated to microbial killing including chemotaxis and transendothelial migration [104]. Further studies are warranted to dissect the precise role of individual members of the galectin family in neutrophil-mediated immune functions and the regulation of the inflammatory response.

27.6 ROLE OF GALECTINS IN EOSINOPHIL AND MAST CELL FUNCTIONS

Eosinophils, basophils, and mast cells are key effectors during Th2-mediated allergic reactions. Scarce information is available on how galectins regulate cells involved in allergic reactions. In this regard, galectin-3 has been shown to downregulate IL-5 (an archetypal Th2 cytokine) on eosinophilic cell lines [69] and galectin-1 has been shown to inhibit eosinophil migration *in vitro* [105].

Interestingly, galectin-9 has been identified as a potent eosinophil-specific chemoattractant (so-called ecalectin) and a potent eosinophil-activating factor [106]. Later, Hirashima's group associated galectin-9 with the control of eosinophil apoptosis [107], suggesting that galectin-9 may serve as a differential regulator of the survival, chemotaxis, and maturation of eosinophils and other immune cell types.

Morevoer, Zuberi et al. studied the effects of galectin-3 in basophils and mast-cell activation *in vitro*. The authors concluded that this lectin has the potential to activate and degranulate these cells, culminating in augmentation of an inflammatory response [108,109]. Conversely, in a model of acute inflammation, galectin-1 treatment inhibited degranulation of tissue mast cells [96].

27.7 EXPERIMENTAL MODELS USED TO EXPLORE THE ROLE OF GALECTINS *IN VIVO*

27.7.1 Galectins in the Regulation of Th1-Mediated Autoimmunity and Chronic Inflammation

Galectin-1, a prototype member of the galectin family, has been proposed to be, in general, a negative regulator of the immune response [5]. Early in the 1980s, Levi et al. reported the preventive and therapeutic effects of electrolectin, a galectin-1 homolog purified from the fish *Electrophorus electricus*, in an experimental model of autoimmune myasthenia gravis in rabbits [110]. Since then, the anti-inflammatory properties of galectin-1 have been evaluated in several models of chronic inflammation and autoimmunity including experimental autoimmune encephalomyelitis (EAE) [111], collagen-induced arthritis [64], hapten-induced colitis [67], interphotoreceptor-binding protein-induced uveitis [68], concanavalin A-induced hepatitis [112], nephrotoxic serum nephritis [113], and autoimmune diabetes [63].

In 1990, Offner et al. demonstrated that galectin-1 prevented the development of clinical and histopathological signs of experimental autoimmune EAE in Lewis rats [111]. Although the mechanisms of action of galectin-1 were not investigated in this study, the authors proposed that galectin-1 might block the sensitization and activation of encephalitogenic T cells [111].

In 1999, our group demonstrated that a single cell injection of syngeneic fibroblasts engineered to secrete galectin-1 at the day of the disease onset abrogated clinical and histopathological manifestations of collagen type-II induced arthritis (CIA), an experimental model of rheumatoid arthritis in DBA/1 mice [64]. This effect was also observed in response to daily injection of recombinant

galectin-1. Insights into the mechanisms involved in this process revealed a critical role for galectin-1 in shifting the balance from a Th1- toward a Th2-polarized immune response as reflected by reduced levels of IFN-γ and increased levels of IL-5 in draining lymph nodes from galectin-1-treated mice. In addition, sera from galectin-1-treated mice showed reduced levels of anticollagen type II IgG2a and increased levels of anticollagen type II IgG1 antibodies. In addition, lymph node cells from mice engaged in the galectin-1 gene therapy protocol had increased susceptibility to antigen-induced apoptosis [64].

Similarly, Santucci et al. found that galectin-1 treatment prevented liver injury and T-helper cell liver infiltration in a model of concanavalin A-induced hepatitis [112]. The authors demonstrated the protective effects of galectin-1 in this model and confirmed that galectin-1 acts *in vivo* by promoting selective elimination of antigen-activated T cells [66].

The preventive and therapeutic effects of galectin-1 have also been demonstrated in the hapten model of inflammatory bowel disease induced by intrarectal delivery of 2,4,6-trinitrobenzene sulfonic acid (TNBS). Here, galectin-1 treatment induced a reduction in the number of antigen-activated mucosal T cells and a decreased secretion of proinflammatory and Th1 cytokines [67].

Given the potential role of galectin-1 in the maintenance of immune privilege in organs such as the eye (similar to TGF-β, IL-10, or Fas ligand), we have also investigated the immunoregulatory effects of this protein in experimental autoimmune uveitis (EAU), a Th1-mediated model of retinal disease [68]. Treatment with galectin-1 either early or late during the course of EAU was sufficient to suppress clinical ocular pathology and to counteract pathogenic Th1 cells. Administration of galectin-1 ameliorated retinal inflammation by skewing the uveitogenic response toward nonpathogenic Th2- or T-regulatory-mediated anti-inflammatory responses [68]. These results highlight the ability of this endogenous lectin to counteract Th1-mediated responses through different, but potentially overlapping anti-inflammatory mechanisms. In addition, we found a striking correlation between the levels of antiretinal galectin-1 autoantibodies in sera from uveitic patients and the severity of autoimmune retinal inflammation [114].

Regarding the immunosuppressive activity of galectin-1 in a transplantation setting, Baum et al. investigated the efficacy of galectin-1 treatment in a murine model of graft versus host disease (GVHD) and found that 68% of galectin-1-treated mice survived compared to 3% of vehicle-treated mice [62]. Similar to findings in autoimmune models, Th1 cytokines were markedly reduced, while production of Th2 cytokines was similar between galectin-1-treated and control animals [62].

In addition, recent evidence indicates that dendritic cells engineered to overexpress galectin-1 can delay the onset of autoimmune diabetes and insulitis when targeted to inflammatory sites [63]. Interestingly, this therapeutic effect was accompanied by increased percentage of apoptotic T cells and reduced number of IFN-γ-secreting $CD4^+$ T cells in pancreatic lymph nodes [63]. Thus, galectin-1 can restore immune cell tolerance in several autoimmune settings by acting as an anti-inflammatory and immunoregulatory cytokine. From a therapeutic standpoint, these findings suggest the potential use of galectin-1 for the selective treatment of Th1-mediated inflammatory disorders.

While these findings suggest a pivotal role for galectin-1 in the maintenance and reestablishment of T-cell tolerance and homeostasis, targeted disruption of the galectin-1 gene in null mutant mice resulted in the absence of major spontaneous abnormalities, suggesting potential redundancy between different members of the galectin family [2,7]. However, in contrast to these previous assumptions, recent work clearly indicates that galectin family members are not redundant and that there are subtle, but functionally relevant differences in the specificity and function of individual members of the galectin family in the regulation of inflammatory responses [2,7,79,115]. In this regard, a study published by Tsuchiyama et al. showed that administration of different members of the galectin family (galectins-1, -3, and -9) have different effects in the regulation of the inflammatory response in a model of nephrotoxic serum nephritis in Wistar Kyoto rats [113]. Additionally, recent evidence indicates that galectin-9 may inhibit glomerular hypertrophy in *db/db* diabetic mice via inhibition of cyclin-dependent kinase inhibitors [116].

27.7.2 Galectins in the Regulation of Th2-Mediated Allergic Disorders

The observation that galectin-3 inhibits IL-5 production in eosinophils and T cells prompted Lopez et al. to investigate the potential therapeutic role of galectin-3 gene therapy in a model of chronic airway inflammation [117]. The authors found that intranasal instillation of a plasmid encoding galectin-3 resulted in an improvement of clinical and immunological manifestations of chronic airway inflammation, including normalization of the eosinophil count and attenuation of hyper-responsiveness to methacholine [117]. This result suggested a role for galectin-3 in counteracting allergic Th2 responses *in vivo*. However, another study indicated that galectin-3-deficient mice develop higher Th1 responses in a model of airway inflammation [118]. In fact, Liu and colleagues showed that galectin-3-deficient mice developed significantly less airway hyper-responsiveness and increased levels of Th1-mediated reactions following antigenic challenge in a murine model of asthma [118]. The discrepancies between these two studies might be possibly explained by the different strategies used by the authors to evaluate the role of galectin-3 *in vivo* (i.e., gal-3 gene therapy vs. induction of allergic inflammation in gal-3 deficient mice).

27.7.3 Galectins in Tumor Immunity

A growing body of experimental evidence suggests that tumors have acquired several strategies to evade immune response and thus behave as immune privileged sites [119]. Interestingly, expression of galectin-1 (as well as other galectins) in cancer cells and cancer-associated stroma positively correlates with the aggressiveness of different tumor types [120]. This suggests that secretion of galectin-1 by tumor cells may be a mechanism by which immunosuppressive microenvironments can be created at tumor sites. This hypothesis was investigated using combined *in vitro* and *in vivo* strategies, demonstrating the role of galectin-1 in tumor-induced immunosuppression and tumor-immune escape [65]. Blockade of the immunosuppressive and proapoptotic activity of galectin-1 within tumor tissue resulted in heightened T-cell-mediated tumor rejection with increased survival of IFN-γ-producing Th1 cells [65]. Supporting these findings, Le and colleagues [121] have recently identified galectin-1 as a molecular link between tumour hypoxia and tumor-immune privilege. The authors found a strong inverse correlation between galectin-1 expression and the presence of T cells in human tumour sections corresponding to head and neck squamous cell carcinoma patients [121]. In addition, it has been recently shown that endothelial cell expression of galectin-1 induced by prostate cancer cells inhibits T-cell transendothelial migration [72]. Finally, it has been recently found that immunosuppressive cytokines such as TGF-β1 and cytostatic drugs such as cyclophosphamide at low doses can modulate galectin-1 expression in tumor and immune cells [122,123]. Taken together, these results support the concept that galectin-1 contributes to immune privilege of tumors by negatively regulating the survival or migration of effector T cells or by skewing the balance toward a Th2-predominant cytokine milieu (Figure 27.3). Given its potent immunosuppressive effects, galectin-1 may be a useful target for therapeutic intervention in cancer. Since galectins-2, -3, -4, and -9 also regulate T-cell responses, future studies are warranted to investigate the potential role of these proteins in tumor immunity.

27.8 CONCLUSION AND FUTURE DIRECTIONS

As illustrated in this chapter, galectins can modulate a plethora of immune responses by acting intracellularly and extracellularly, as chemokines, cytokines, growth factors, death triggers, and survival inducers. Under distinct physiological or pathological conditions, different members of the galectin family may provide inhibitory or stimulatory signals to control immune cell homeostasis and regulate inflammation following an antigenic challenge. However, like many other cytokines and growth

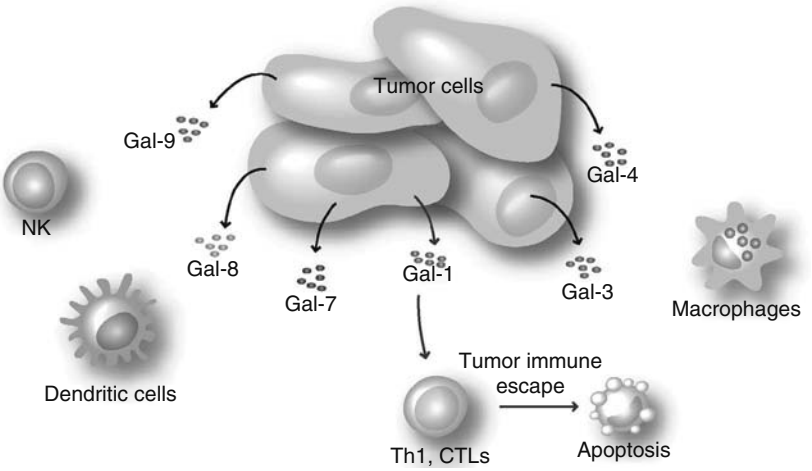


FIGURE 27.3 (See color insert following blank page 170. Also see CD for color figure.) Synthesis of galectins by tumor cells and modulation of tumor immunity.

factors (e.g., TGF-β), galectins may exhibit a "double-edged sword" effect depending on many different intrinsic factors such as the physicochemical properties of the isolated protein (monomer/dimer equilibrium), its stability in culture, its concentration and oxidation state, and other extrinsic factors such as the target cell type and the general context or microenvironment.

The mechanisms underlying galectin-mediated regulation of cytokine production and the association of this effect with the regulation of T-cell survival, migration, and differentiation still remains to be elucidated. Further studies are warranted to investigate the kinetics of galectin expression in different immune cell types and the potential overlapping in galectin functions throughout the development of an ongoing immune response.

Given the broad spectrum of immunoregulatory effects in autoimmune diseases, inflammatory processes, and cancer, galectins have been postulated as candidates for the design of novel anti-inflammatory drugs [124] and as targets for anticancer therapies [125]. However, before galectin-based therapeutic agents can be fully realized, a more thorough understanding of the lesser studied galectins and the mechanisms involved in their different immunoregulatory functions is required. To what extent is there functional redundancy and specificity of action within the galectin family? What is the precise explanation of the different functions exerted by the similar galectins in different environmental contexts? What are the precise mechanisms involved in the anti-inflammatory and immunoregulatory effects of different members of the galectin family? What are the levels of galectins attained *in vivo* during an inflammatory reaction, infectious process or tumor dissemination? Increased understanding of the role of galectins in immunoregulation, inflammation, and cancer should provide more insights into how the regulation of galectin expression and activity can be exploited for therapeutic purposes.

ACKNOWLEDGMENTS

We apologize to many authors whose excellent papers could not be cited in this chapter. Work in G.A.R.'s laboratory is supported by The Cancer Research Institute "Elaine R. Shepard Memorial Investigator Award" (United States), Mizutani Foundation for Glycoscience (Japan), Agencia Nacional de Promoción de Ciencia y Tecnológica (PICT 2003-05-13787), University of Buenos Aires (M091), Fundación Sales and Fundación Florencio Fiorini (Argentina). G.A.R. is a member

of the research Career of CONICET and a Fellow of the "John Simon Guggenheim Memorial Foundation" (United States).

REFERENCES

1. Barondes, S.H., et al., Galectins: A family of animal beta-galactoside-binding lectins, *Cell*, 76, 597, 1994.
2. Leffler, H., et al., Introduction to galectins, *Glycoconj J*, 19, 433, 2004.
3. Barondes, S.H., Vertebrate lectins: Properties and functions, in *The Lectins: Properties, Functions and Applications in Biology and Medicine*, Liener, I.E., Sharon, N. and Goldstein, I.J. Eds., Academic Press, Orlando, FL, 1986, p. 467.
4. Teichberg, V.I., et al., A beta-D-galactoside binding protein from electric organ tissue of *Electrophorus electricus*, *Proc Natl Acad Sci USA*, 72, 1383, 1975.
5. Rabinovich, G.A., et al., Galectins and their ligands: Amplifiers, silencers or tuners of the inflammatory response?, *Trends Immunol*, 23, 313, 2002.
6. Cooper, D.N., Galectinomics: Finding themes in complexity, *Biochim Biophys Acta*, 1572, 209, 2002.
7. Liu, F.T. and Rabinovich, G.A., Galectins as modulators of tumour progression, *Nat Rev Cancer*, 5, 29, 2005.
8. Brewer, C., Binding and cross-linking properties of galectins, *Biochim Biophys Acta*, 1572, 255, 2002.
9. Liu, F.T., Patterson, R.J., and Wang, J.L., Intracellular functions of galectins, *Biochim Biophys Acta*, 1572, 263, 2002.
10. Hirabayashi, J., et al., Oligosaccharide specificity of galectins: A search by frontal affinity chromatography, *Biochim Biophys Acta*, 1572, 232, 2002.
11. Ilarregui, J.M., et al., The coming of age of galectins as immunomodulatory agents: Impact of these carbohydrate binding proteins in T cell physiology and chronic inflammatory disorders, *Ann Rheum Dis*, 64 Suppl 4, iv96, 2005.
12. Blaser, C., et al., Beta-galactoside-binding protein secreted by activated T cells inhibits antigen-induced proliferation of T cells, *Eur J Immunol*, 28, 2311, 1998.
13. Rabinovich, G.A., et al., Activated rat macrophages produce a galectin-1-like protein that induces apoptosis of T cells: Biochemical and functional characterization, *J Immunol*, 160, 4831, 1998.
14. Zuniga, E., et al., Regulated expression of galectin-1 during B-cell activation and implications for T-cell apoptosis, *J Leukoc Biol*, 70, 73, 2001.
15. Garin, M.I., et al., Galectin-1: A key effector of regulation mediated by CD4+ CD25+ T cells, *Blood*, 109, 2058, 2007.
16. Sugimoto, N., et al., Foxp3-dependent and -independent molecules specific for CD25+ CD4+ natural regulatory T cells revealed by DNA microarray analysis, *Int Immunol*, 18, 1197, 2006.
17. Rabinovich, G.A., Rubinstein, N., and Toscano, M.A., Role of galectins in inflammatory and immunomodulatory processes, *Biochim Biophys Acta*, 1572, 274, 2002.
18. Sacchettini, J.C., Baum, L.G., and Brewer, C.F., Multivalent protein–carbohydrate interactions. A new paradigm for supermolecular assembly and signal transduction, *Biochemistry*, 40, 3009, 2001.
19. Van Parijs, L. and Abbas, A.K., Homeostasis and self-tolerance in the immune system: Turning lymphocytes off, *Science*, 280, 243, 1998.
20. Hahn, H.P., et al., Galectin-1 induces nuclear translocation of endonuclease G in caspase- and cytochrome c-independent T cell death, *Cell Death Differ*, 11, 1277, 2004.
21. Matarrese, P., et al., Galectin-1 sensitizes resting human T lymphocytes to Fas (CD95)-mediated cell death via mitochondrial hyperpolarization, budding, and fission, *J Biol Chem*, 280, 6969, 2005.
22. Pace, K.E., et al., CD7 delivers a pro-apoptotic signal during galectin-1-induced T cell death, *J Immunol*, 165, 2331, 2000.
23. Perillo, N.L., et al., Apoptosis of T cells mediated by galectin-1, *Nature*, 378, 736, 1995.
24. Perillo, N.L., et al., Galectin-1, an endogenous lectin produced by thymic epithelial cells, induces apoptosis of human thymocytes, *J Exp Med*, 185, 1851, 1997.
25. Rabinovich, G.A., et al., Specific inhibition of lymphocyte proliferation and induction of apoptosis by CLL-I, a beta-galactoside-binding lectin, *J Biochem*, 122, 365, 1997.
26. Rabinovich, G.A., et al., Induction of allogenic T-cell hyporesponsiveness by galectin-1-mediated apoptotic and non-apoptotic mechanisms, *Cell Death Differ*, 9, 661, 2002.
27. Nguyen, J.T., et al., CD45 modulates galectin-1-induced T cell death: Regulation by expression of core 2 O-glycans, *J Immunol*, 167, 5697, 2001.

28. Pace, K.E., et al., Restricted receptor segregation into membrane microdomains occurs on human T cells during apoptosis induced by galectin-1, *J Immunol*, 163, 3801, 1999.
29. Walzel, H., et al., Involvement of CD2 and CD3 in galectin-1 induced signaling in human Jurkat T-cells, *Glycobiology*, 10, 131, 2000.
30. Galvan, M., et al., Expression of a specific glycosyltransferase enzyme regulates T cell death mediated by galectin-1, *J Biol Chem*, 275, 16730, 2000.
31. Amano, M., et al., The ST6Gal I sialyltransferase selectively modifies N-glycans on CD45 to negatively regulate galectin-1-induced CD45 clustering, phosphatase modulation, and T cell death, *J Biol Chem*, 278, 7469, 2003.
32. Rabinovich, et al., Functions of cell surface galectin-glycoprotein lattices, *Curr Opin Struct Biol*, 17, 513, 2007.
33. Rappl, G., et al., CD4+ CD7-leukemic T cells from patients with Sezary syndrome are protected from galectin-1-triggered T cell death, *Leukemia*, 16, 840, 2002.
34. Roberts, A.A., et al., Galectin-1-mediated apoptosis in mycosis fungoides: The roles of CD7 and cell surface glycosylation, *Mod Pathol*, 16, 543, 2003.
35. Cabrera, P.V., et al., Haploinsufficiency of C2GnT-I glycosyltransferase renders T lymphoma cells resistant to cell death, *Blood*, 108, 2399, 2006.
36. Rabinovich, G.A., et al., Molecular mechanisms implicated in galectin-1-induced apoptosis: Activation of the AP-1 transcription factor and downregulation of Bcl-2, *Cell Death Differ*, 7, 747, 2000.
37. Ion, G., et al., Role of p56lck and ZAP70-mediated tyrosine phosphorylation in galectin-1-induced cell death, *Cell Death Differ*, 12, 1145, 2005.
38. Ion, G., et al., Acid sphingomyelinase mediated release of ceramide is essential to trigger the mitochondrial pathway of apoptosis by galectin-1, *Cell Signal*, 18, 1887, 2006.
39. Allione, A., et al., Beta-galactoside-binding protein (beta GBP) alters the cell cycle, up-regulates expression of the alpha- and beta-chains of the IFN-γ receptor, and triggers IFN-γ-mediated apoptosis of activated human T lymphocytes, *J Immunol*, 161, 2114, 1998.
40. Endharti, A.T., et al., Galectin-1 supports survival of naive T cells without promoting cell proliferation, *Eur J Immunol*, 35, 86, 2005.
41. Stowell, S.R., et al., Human galectin-1, -2, and -4 induce surface exposure of phosphatidylserine in activated human neutrophils but not in activated T cells, *Blood*, 109, 219, 2007.
42. Sturm, A., et al., Human galectin-2: Novel inducer of T cell apoptosis with distinct profile of caspase activation, *J Immunol*, 173, 3825, 2004.
43. Akahani, S., et al., Galectin-3: A novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family, *Cancer Res*, 57, 5272, 1997.
44. Yang, R.Y., Hsu, D.K., and Liu, F.T., Expression of galectin-3 modulates T-cell growth and apoptosis, *Proc Natl Acad Sci USA*, 93, 6737, 1996.
45. Fukumori, T., et al., CD29 and CD7 mediate galectin-3-induced type II T-cell apoptosis, *Cancer Res*, 63, 8302, 2003.
46. Stillman, B.N., et al., Galectin-3 and galectin-1 bind distinct cell surface glycoprotein receptors to induce T cell death, *J Immunol*, 176, 778, 2006.
47. Silva-Monteiro, E., et al., Altered expression of galectin-3 induces cortical thymocyte depletion and premature exit of immature thymocytes. *Am J Pathol*, 170, 546, 2007.
48. Wada, J., et al., Developmental regulation, expression, and apoptotic potential of galectin-9, a beta-galactoside binding lectin, *J Clin Invest*, 99, 2452, 1997.
49. Kashio, Y., et al., Galectin-9 induces apoptosis through the calcium-calpain-caspase-1 pathway, *J Immunol*, 170, 3631, 2003.
50. Lu, L.H., et al., Characterization of galectin-9-induced death of Jurkat T-cells, *J Biochem*, 141, 157, 2007.
51. Zhu, C., et al., The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity, *Nat Immunol*, 6, 1245, 2005.
52. Bernerd, F., Sarasin, A., and Magnaldo, T., Galectin-7 overexpression is associated with the apoptotic process in UVB-induced sunburn keratinocytes, *Proc Natl Acad Sci USA*, 96, 11329, 1999.
53. Kuwabara, I., et al., Galectin-7 (PIG1) exhibits pro-apoptotic function through JNK activation and mitochondrial cytochrome *c* release, *J Biol Chem*, 277, 3487, 2002.
54. Hadari, Y.R., et al., Galectin-8 binding to integrins inhibits cell adhesion and induces apoptosis, *J Cell Sci*, 113, 2385, 2000.
55. Yang, R.Y., et al., Cell cycle regulation by galectin-12, a new member of the galectin superfamily, *J Biol Chem*, 276, 20252, 2001.

56. Vespa, G.N., et al., Galectin-1 specifically modulates TCR signals to enhance TCR apoptosis but inhibit IL-2 production and proliferation, *J Immunol*, 162, 799, 1999.
57. Chung, C.D., et al., Galectin-1 induces partial TCR zeta-chain phosphorylation and antagonizes processive TCR signal transduction, *J Immunol*, 165, 3722, 2000.
58. Demetriou, M., et al., Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation, *Nature*, 409, 733, 2001.
59. Hokama, A., et al., Induced reactivity of intestinal CD4(+) T cells with an epithelial cell lectin, galectin-4, contributes to exacerbation of intestinal inflammation, *Immunity*, 20, 681, 2004.
60. Rabinovich, G.A., et al., Specific inhibition of T-cell adhesion to extracellular matrix and proinflammatory cytokine secretion by human recombinant galectin-1, *Immunology*, 97, 100, 1999.
61. van der Leij, J., et al., Dimeric galectin-1 induces IL-10 production in T-lymphocytes: An important tool in the regulation of the immune response, *J Pathol*, 204, 511, 2004.
62. Baum, L.G., et al., Amelioration of graft versus host disease by galectin-1, *Clin Immunol*, 109, 295, 2003.
63. Perone, M.J., et al., Dendritic cells expressing transgenic galectin-1 delay onset of autoimmune diabetes in mice, *J Immunol*, 177, 5278, 2006.
64. Rabinovich, G.A., et al., Recombinant galectin-1 and its genetic delivery suppress collagen-induced arthritis via T cell apoptosis, *J Exp Med*, 190, 385, 1999.
65. Rubinstein, N., et al., Targeted inhibition of galectin-1 gene expression in tumor cells results in heightened T cell-mediated rejection; A potential mechanism of tumor-immune privilege, *Cancer Cell*, 5, 241, 2004.
66. Toscano, M.A., et al., Differential glycosylation of Th1, Th2 and Th-17 effector cells selectively regulates susceptibility to cell death, *Nat Immunol*, 8, 825, 2007.
67. Santucci, L., et al., Galectin-1 suppresses experimental colitis in mice, *Gastroenterology*, 124, 1381, 2003.
68. Toscano, M.A., et al., Galectin-1 suppresses autoimmune retinal disease by promoting concomitant Th2- and T regulatory-mediated anti-inflammatory responses, *J Immunol*, 176, 6323, 2006.
69. Cortegano, I., et al., Galectin-3 down-regulates IL-5 gene expression on different cell types, *J Immunol*, 161, 385, 1998.
70. Shevach, E.M., CD4+ CD25+ suppressor T cells: More questions than answers, *Nat Rev Immunol*, 2, 389, 2002.
71. Butcher, E.C. and Picker, L.J., Lymphocyte homing and homeostasis, *Science*, 272, 60, 1996.
72. He, J. and Baum, L.G., Endothelial cell expression of galectin-1 induced by prostate cancer cells inhibits T-cell transendothelial migration, *Lab Iinvest*, 86, 578, 2006.
73. Gauthier, L., et al., Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering, *Proc Natl Acad Sci USA*, 99, 13014, 2002.
74. Rossi, B., et al., Clustering of pre-B cell integrins induces galectin-1-dependent pre-B cell receptor relocalization and activation, *J Immunol*, 177, 796, 2006.
75. Clark, A.G., et al., Multifunctional regulators of cell growth are differentially expressed in anergic murine B cells, *Mol Immunol*, 44, 1274, 2007.
76. Acosta-Rodriguez, E.V., et al., Galectin-3 mediates IL-4-induced survival and differentiation of B cells: Functional cross-talk and implications during *Trypanosoma cruzi* infection, *J Immunol*, 172, 493, 2004.
77. Gordon, S., Alternative activation of macrophages, *Nat Rev Immunol*, 3, 23, 2003.
78. Correa, S.G., et al., Opposite effects of galectin-1 on alternative metabolic pathways of L-arginine in resident, inflammatory, and activated macrophages, *Glycobiology*, 13, 119, 2003.
79. Barrionuevo, P., et al., A novel function for galectin-1 at the crossroad of innate and adaptive immunity: Galectin-1 regulates monocyte/macrophage physiology through a nonapoptotic ERK-dependent pathway, *J Immunol*, 178, 436, 2007.
80. Zuniga, E., et al., Regulated expression and effect of galectin-1 on *Trypanosoma cruzi*-infected macrophages: Modulation of microbicidal activity and survival, *Infect Immun*, 69, 6804, 2001.
81. Ozaki, K., et al., Functional variation in LGALS2 confers risk of myocardial infarction and regulates lymphotoxin-alpha secretion in vitro, *Nature*, 429, 72, 2004.
82. Hsu, D.K., et al., Targeted disruption of the galectin-3 gene results in attenuated peritoneal inflammatory responses, *Am J Pathol*, 156, 1073, 2000.
83. Sano, H., et al., Critical role of galectin-3 in phagocytosis by macrophages, *J Clin Invest*, 112, 389, 2003.

84. Sano, H., et al., Human galectin-3 is a novel chemoattractant for monocytes and macrophages, *J Immunol*, 165, 2156, 2000.
85. Bernardes, E.S., et al., *Toxoplasma gondii* infection reveals a novel regulatory role for galectin-3 in the interface of innate and adaptive immunity, *Am J Pathol*, 168, 1910, 2006.
86. Jouault, T., et al., Specific recognition of *Candida albicans* by macrophages requires galectin-3 to discriminate *Saccharomyces cerevisiae* and needs association with TLR2 for signaling, *J Immunol*, 177, 4679, 2006.
87. Pelletier, I. and Sato, S., Specific recognition and cleavage of galectin-3 by *Leishmania major* through species-specific polygalactose epitope, *J Biol Chem*, 277, 17663, 2002.
88. John, C.M., et al., Galectin-3 binds lactosaminylated lipooligosaccharides from *Neisseria gonorrhoeae* and is selectively expressed by mucosal epithelial cells that are infected, *Cell Microbiol*, 4, 649, 2002.
89. Fulcher, J.A., et al., Galectin-1-matured human monocyte-derived dendritic cells have enhanced migration through extracellular matrix, *J Immunol*, 177, 216, 2006.
90. Levroney, E.L., et al., Novel innate immune functions for galectin-1: Galectin-1 inhibits cell fusion by Nipah virus envelope glycoproteins and augments dendritic cell secretion of proinflammatory cytokines, *J Immunol*, 175, 413, 2005.
91. Perone, M.J., et al., Transgenic galectin-1 induces maturation of dendritic cells that elicit contrasting responses in naive and activated T cells, *J Immunol*, 176, 7207, 2006.
92. Dai, S.Y., et al., Galectin-9 induces maturation of human monocyte-derived dendritic cells, *J Immunol*, 175, 2974, 2005.
93. Vray, B., et al., Up-regulation of galectin-3 and its ligands by *Trypanosoma cruzi* infection with modulation of adhesion and migration of murine dendritic cells, *Glycobiology*, 14, 647, 2004.
94. La, M., et al., A novel biological activity for galectin-1: Inhibition of leukocyte–endothelial cell interactions in experimental inflammation, *Am J Pathol*, 163, 1505, 2003.
95. Rabinovich, G.A., et al., Evidence of a role for galectin-1 in acute inflammation, *Eur J Immunol*, 30, 1331, 2000.
96. Colnot, C., et al., Maintenance of granulocyte numbers during acute peritonitis is defective in galectin-3-null mutant mice, *Immunology*, 94, 290, 1998.
97. Almkvist, J., et al., Activation of the neutrophil nicotinamide adenine dinucleotide phosphate oxidase by galectin-1, *J Immunol*, 168, 4034, 2002.
98. Karlsson, A., et al., Galectin-3 activates the NADPH-oxidase in exudated but not peripheral blood neutrophils, *Blood*, 91, 3430, 1998.
99. Kuwabara, I. and Liu, F.T., Galectin-3 promotes adhesion of human neutrophils to laminin, *J Immunol*, 156, 3939, 1996.
100. Sato, S., et al., Role of galectin-3 as an adhesion molecule for neutrophil extravasation during streptococcal pneumonia, *J Immunol*, 168, 1813, 2002.
101. Nieminen, J., St-Pierre, C., and Sato, S., Galectin-3 interacts with naive and primed neutrophils, inducing innate immune responses, *J Leukoc Biol*, 78, 1127, 2005.
102. Fernandez, G.C., et al., Galectin-3 and soluble fibrinogen act in concert to modulate neutrophil activation and survival: Involvement of alternative MAPK pathways, *Glycobiology*, 15, 519, 2005.
103. Gil, C.D., et al., Inflammation-induced modulation of cellular galectin-1 and -3 expression in a model of rat peritonitis, *Inflamm Res*, 55, 99, 2006.
104. Nishi, N., et al., Galectin-8 modulates neutrophil function via interaction with integrin alphaM, *Glycobiology*, 13, 755, 2003.
105. Delbrouck, C., et al., Galectin-1 is overexpressed in nasal polyps under budesonide and inhibits eosinophil migration, *Lab Invest*, 82, 147, 2002.
106. Matsumoto, R., et al., Biological activities of ecalectin: A novel eosinophil-activating factor, *J Immunol*, 168, 1961, 2002.
107. Saita, N., et al., Association of galectin-9 with eosinophil apoptosis, *Int Arch Allergy Immunol*, 128, 42, 2002.
108. Zuberi, R.I., Frigeri, L.G., and Liu, F.T., Activation of rat basophilic leukemia cells by epsilon BP, an IgE-binding endogenous lectin, *Cell Immunol*, 156, 1, 1994.
109. Chen, H.Y., et al., Role of galectin-3 in mast cell functions: Galectin-3-deficient mast cells exhibit impaired mediator release and defective JNK expression, *J Immunol*, 177, 4991, 2006.
110. Levi, G., Tarrab-Hazdai, R., and Teichberg, V.I., Prevention and therapy with electrolectin of experimental autoimmune myasthenia gravis in rabbits, *Eur J Immunol*, 13, 500, 1983.
111. Offner, H., et al., Recombinant human beta-galactoside binding lectin suppresses clinical and histological signs of experimental autoimmune encephalomyelitis, *J Neuroimmunol*, 28, 177, 1990.

112. Santucci, L., et al., Galectin-1 exerts immunomodulatory and protective effects on concanavalin-A-induced hepatitis in mice, *Hepatology*, 31, 399, 2000.
113. Tsuchiyama, Y., et al., Efficacy of galectins in the amelioration of nephrotoxic serum nephritis in Wistar Kyoto rats, *Kidney Int*, 58, 1941, 2000.
114. Romero, M.D., et al., Circulating anti-galectin-1 antibodies are associated with the severity of ocular disease in autoimmune and infectious uveitis, *Invest Ophthalmol Vis Sci*, 47, 1550, 2006.
115. Thijssen, V.L., et al., Galectin-1 is essential in tumor angiogenesis and is a target for antiangiogenesis therapy, *Proc Natl Acad Sci USA*, 103, 15975, 2006.
116. Baba, M., et al., Galectin-9 inhibits glomerular hypertrophy in db/db diabetic mice via cell-cycle-dependent mechanisms, *J Am Soc Nephrol*, 16, 3222, 2005.
117. Lopez, E., et al., Inhibition of chronic airway inflammation and remodeling by galectin-3 gene therapy in a murine model, *J Immunol*, 176, 1943, 2006.
118. Zuberi, R.I., et al., Critical role for galectin-3 in airway inflammation and bronchial hyperresponsiveness in a murine model of asthma, *Am J Pathol*, 165, 2045, 2004.
119. Rabinovich, G.A., Gabrilovich, D., and Sotomayor, E.M., Immunosuppressive strategies that are mediated by tumor cells, *Annu Rev Immunol*, 25, 267, 2007.
120. Camby, I., et al., Galectin-1: A small protein with major functions, *Glycobiology*, 16, 137R, 2006.
121. Le, Q.T., et al., Galectin-1: A link between tumor hypoxia and tumor immune privilege, *J Clin Oncol*, 23, 8932, 2005.
122. Daroqui, C.M., et al., Regulation of galectin-1 expression by transforming growth factor beta1 in metastatic mammary adenocarcinoma cells: Implications for tumor-immune escape, *Cancer Immunol Immunother*, 56, 491, 2007.
123. Zacarias Fluck, M.F., et al., Low-dose cyclophosphamide modulates galectin-1 expression and function in an experimental rat lymphoma model, *Cancer Immunol Immunother*, 56, 237, 2007.
124. Parsonage, G., et al., Roles of galectins in chronic inflammatory microenvironments, *Future Rheumatol*, 1, 441, 2006.
125. Ingrassia, L., et al., Anti-galectin compounds as potential anti-cancer drugs, *Curr Med Chem*, 13, 3513, 2006.

28 Interactions of Galectins with Leukocytes

Sean R. Stowell and Richard D. Cummings

CONTENTS

28.1 INTRODUCTION

It seems fitting that the scientists credited for the discovery of the first vertebrate "galectin" also described its first and most impressive effects on leukocytes. Levi and Tiechburg demonstrated that prophylactic administration of "electrolectin" (equivalent to galectin-1 of the electric organ tissue of *Electrophorus electricus* [1]) prevented the formation of experimentally induced myasthenia gravis [2]. Several years later, Offner et al. demonstrated that human galectin-1 (Gal-1) also suppressed the clinical signs and pathology associated with experimental autoimmune encephalomyelitis (EAE) in rats [3], suggesting a broad immunosuppressive spectrum for Gal-1. Interestingly, both of these authors attributed the immunomodulatory effects of Gal-1 to suppressor T cells [2–4]. However, suppressor T cells quickly fell out of vogue [5], requiring an alternative explanation for the potent immunosuppressive activities of Gal-1. In 1995, Baum and colleagues suggested that Gal-1 reduced immune responses by directly inducing apoptotic cell death in activated, but not resting, T cells [6].

Following this report by Baum and colleagues [6], significant work on galectins focused on the proapoptotic potential of galectin family members toward different leukocytes. However, other studies also described apoptosis-independent activities of various galectin family members in the regulation of both innate and adaptive immune responses. Taken together, these studies suggest that the galectin family of β-galactoside proteins, which includes over a dozen vertebrate members, possess a wide variety of activities and likely regulates many aspects of immunity. We will begin our evaluation of the effects of galectins on leukocytes by first describing known functions of galectin family members in aspects of innate immunity.

28.2 INNATE IMMUNITY

Innate immunity is the constitutive immunity that is not inducible to specific antigens and does not require prior sensitization, via an antigen from, for example, an infection or vaccination and is in contrast to acquired or adaptive immunity. Since specific antigens do not stimulate it, innate immunity is generally nonspecific and immediate in action. A key feature of the innate immune response is the involvement of antimicrobial proteins and peptides, chemokines, Toll-like receptors, and glycan-binding proteins, including Ca^{2+}-dependent (C-type) lectins and galectins that recognize specific glycans on invading pathogens and on host glycoconjugates. The major cells important in innate immune responses include myeloid and lymphoid dendritic cells (DC), macrophages, mast cells, natural killer cells, gamma/delta T cells, and B-1 cells. Many if not all of these cells express various members of the galectin family and some of these galectins have been shown to have activities related to innate immune responses. This aspect of the review focuses on Gal-1, -2, -3, -4, -8, -9, -10, and -14, since little is known about the roles of the other galectins in innate immunity. A summary of many of the activities for galectins in both innate and adaptive immune responses are shown in Table 28.1.

28.2.1 Galectin-1

Gal-1 suppresses several aspects of innate immune function. For example, administration of human Gal-1 attenuates leukocyte extravasation, blocks IL-1β induced neutrophil migration *in vivo* in the mouse peritoneal cavity, and IL-8 induced chemotaxis of neutrophils *in vitro* [7]. Such results suggest a role for Gal-1 in blocking leukocyte, and in particular neutrophil, recruitment into inflamed tissue. Similarly, Gal-1 treatment also reduces neutrophil and mast-cell infiltration following carrageenin-induced rat peritonitis [8]. However, chemotaxis inhibition appears not only to apply to neutrophils, because Gal-1 also reduces eosinophil migration in nasal polyps [9]. Inhibition of leukocyte migration may protect viable cells from the potential harmful effects of innate immune cells once activated.

In addition to inhibiting chemotaxis of leukocytes, Gal-1 also blocks degranulation of mast cells and reduces PGE_2 production and arachidonic acid release from macrophages [10]. These results suggest that Gal-1 not only inhibits leukocyte recruitment once initiated, but also may be important in attenuating the signaling events prior to leukocyte infiltration, which initiate inflammatory events. Consistent with this hypothesis, Gal-1 reduces foot pad edema following phospholipase A2 injection with an accompanying reduction in neutrophil and mast-cell recruitment *in vivo* [10]. Taken together, these studies indicate that in certain conditions galectin-1 may exhibit anti-inflammatory activities.

Interestingly, Gal-1 seems to also exhibit activities consistent with proinflammatory functions. For example, Gal-1 activates the NADPH-dependent respiratory burst in neutrophils in exudated or (formyl-Methionyl-Leucyl-Phenylalanine) fMLP-activated, but not resting, human neutrophils [11], implicating a potential role for Gal-1 in the initiation of neutrophil activity. Furthermore, transgenic overexpression of Gal-1 in murine DC increases DC activation [12], suggesting that Gal-1 may serve as a danger signal, activating DC following release from damaged tissue. By contrast, adoptive transfer of transgenic DCs delayed the onset of type I diabetes [12], suggesting that Gal-1 may actually inhibit DC presentation and activation of naive T cells.

TABLE 28.1
Partial Summary of Galectin Activities and Interactions with Leukocytes

General immune functions of Galectins

Gal-1 reduces pathology associated with graft-versus-host disease, Con A-induced hepatitis, experimental allergic encephalomyelitis, myasthenia gravis, and rheumatoid arthritis
Gal-1 reduces acute inflammatory responses
Gal-2 is involved in the pathogenesis of atheroma formation
Gal-3 exacerbates Th2 immune responses (asthma)

Monocytes/macrophages

Gal-3 (also called Mac-2 antigen)
- enhances phagocytosis
- enhances respiratory burst
- enhances LPS-induced IL-1β secretion
- inhibits apoptosis (intracellularly)

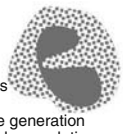
Mast cells

Gal-1 inhibits degranulation
Gal-3 induces degranulation independent of antigen-mediated IgE stimulation

B cells

Gal-3 blocks IL-4 induced survival of activated B cells
- favors plasma cell differentiation
- exhibits an antiapoptotic role in B-cell lymphomas

Eosinophils

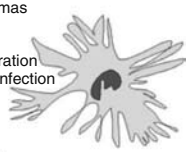


Gal-3 inhibits IL-5 production
Gal-9 induces
- chemotaxis
- activation
- superoxide generation
- moderate degranulation

Gal-10 (Charcot-Leyden crystal protein) is highly expressed; function unknown
Gal-14 expressed but function unknown

Dendritic cells

Gal-1 and Gal-9 induce maturation
Gal-3 expression by *T. cruzi* infection



Thymocytes/T cells

Gal-1 involved in Treg effector function
- conflicting results on effects on viability
- mediates adhesion of thymocytes to thymic epithelium
- induces apoptosis in CD4+CD8+double-positive thymocytes (DTT present)
- induces shift in TH1 response to TH2 (decreases IFNγ: increases IL-5)
- reduces TNFα, IL-1β, IL-12, IL-2 and IFNγ
- enhances Treg formation
- increases IL-10 production in both naive and activated T cells

Gal-2 induces T cell apoptosis (DTT present)
- decreases IFNγ and TNFα while increasing IL-10 and IL-5

Gal-3 blocks apoptosis when overexpressed intracellularly
- conflicting results on the effects of endogenous intracellular involvement in viability
- conflicting reports on the effects of extracellular effects on viability
- promotes adhesion of thymocytes to thymic epithelium
- enhances Th2 immune responses
- enhances adhesion of niave T cells to DCs
- binds TCR, reducing TCR-mediated T cell activation

Gal-4 induces IL-6 production
Gal-8 activates Rac-1
Gal-9 induces apoptosis in thymocytes and T cells
- induces selective loss of CD4+Th1 cells
- induces selective loss of CD8+T cells

Gal-10 involved in Treg effector function
Gal-11 involved in Treg effector function

Endothelium

Gal-3 enhances pro-inflammatory cytokine release
Gal-1 expression upregulated by activation
Gal-3 expression upregulated in tumor endothelial cells
Gal-9 expression induced by virus infection

Neutrophils

Gal-1 induces apoptosis-independent PS exposure (Ca^{2+}-dependent) (preaparesis)
- inhibits chemotaxis
- inhibits extravasation
- activates NADPH-dependent respiratory burst

Gal-3 induces chemotaxis
- enhances extravasation
- activates NADPH-dependent respiratory burst
- induces neutrophil activation
- induces release of IL-8
- mediates interaction with laminin and fibronectin (both directly and indirectly)
- enhances leukocyte adhesion to endothelium

Gal-2 induces apoptosis-independent PS exposure (Ca^{2+}-dependent)
Gal-4 induces apoptosis- independent PS exposure (Ca^{2+}-independent)
Gal-8 activates NADPH dependent respiratory burst
- modulates integrin-mediated neutrophil adhesion

Although apoptosis-independent mechanisms of neutrophil removal exist [13,14], the mechanism underlying these events remains enigmatic. Recently, Dias-Baruffi and colleagues demonstrated that Gal-1 induces cell surface exposure of phosphatidylserine (PS) in activated, but not resting, neutrophils; this is a membrane scrambling event normally associated with apoptotic cell death [15–17]. PS serves as the primary receptor for recognition and phagocytic clearance of apoptotic cells [18]. Importantly, Gal-1-induced PS exposure occurs in the absence of cell death, while rendering cells sensitive to phagocytic removal [15–17], providing the first possible mechanism whereby neutrophils may be signaled for removal independent of cell death [13,14]. Although the entire mechanism of Gal-1-induced nonapoptotic PS exposure remains to be elucidated, Gal-1-induced transient mobilization of intracellular Ca^{2+} and exposure of PS required activation of Src kinases [16] (Figure 28.1). The proteins responsible for surface expression are not clear, but at least two types of membrane proteins may be involved—phospholipids scramblases [19] and aminophospholipid translocases [20], both of which are independently and oppositely regulated by intracellular Ca^{2+} concentrations. This signaling

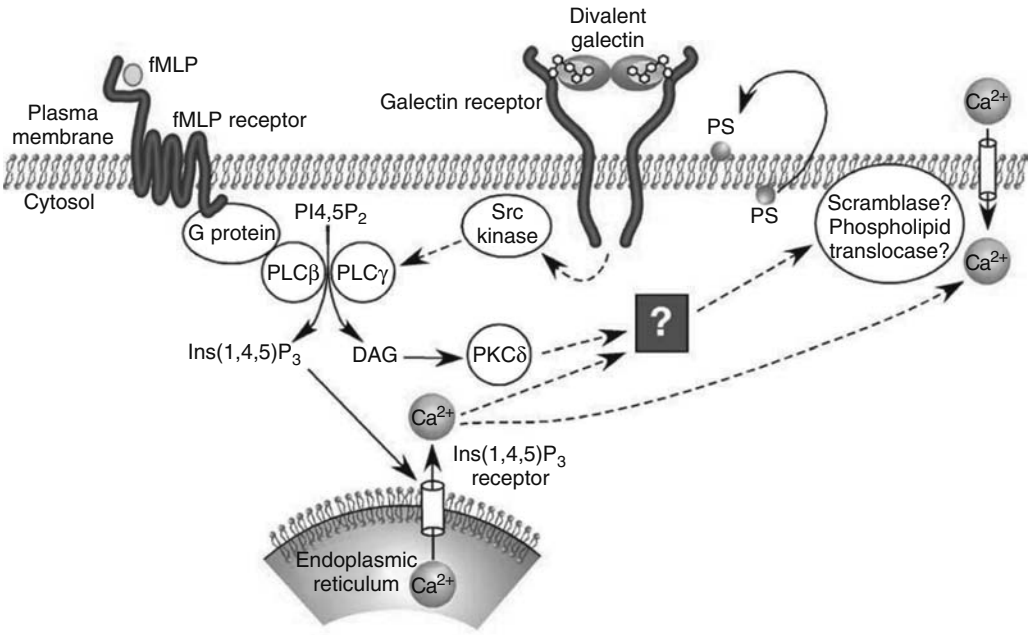


FIGURE 28.1 (See color insert following blank page 170. Also see CD for color figure.) Model of dGal-1 signaling in fMLP-activated neutrophils for cell surface exposure of PS. Following activation of neutrophils through the fMLP receptor, Ca^{2+} is released from intracellular stores through G-protein-coupled, PLCβ-induced cleavage of $PI4,5P_2$ to $Ins(1,4,5)P_3$ and DAG. Independently, the binding of dimeric dGal-1 (or possibly other divalent galectins) to a cell-surface glycoprotein receptor activates Src kinases, with subsequent activation of PLCγ, release of additional $Ins(1,4,5)P_3$ and DAG, and a second transient rise in cytosolic Ca^{2+}. The dual transient rises in cytosolic Ca^{2+} are necessary but not sufficient to redistribute PS to the cell surface. The IP_3-mediated release of Ca^{2+} opens plasma membrane ion channels, but extracellular Ca^{2+} is not required to mobilize PS to the cell surface. The DAG may activate PKCδ, which may contribute to PS exposure by activating a phospholipid scramblase and aminophospholipid translocase. The dashed arrows represent pathways for which signaling intermediates and/or experimental evidence remain incomplete. (From Karmakar, S., Cummings, R. D., and McEver, R. P., *J. Biol. Chem.,* 280, 28623, 2005.)

pathway appears to be unique to galectins and represents one of the best defined signaling cascades known for galectins in leukocytes independently of cell death.

The unique ability of galectins to signal PS exposure in living cells challenges current models of cell turnover and removal, which invoke only cellular apoptosis and necrosis. Galectin-induced cell removal could occur independently of cell death through the phagocytosis of living cells. We propose that this process be termed "preaparesis" (from the Latin term *preaparare*), signifying that it prepares cells for phagocytic removal by causing PS exposure without accompanying apoptosis [20b].

28.2.2 Galectin-2

Gal-2 shares the greatest degree of homology to Gal-1 [21]. Similar to Gal-1, Gal-2 induces apoptosis-independent PS exposure in activated neutrophils [17]. Gal-2 also induces Ca^{2+} flux in both resting and activated neutrophils, similar to that previously reported for Gal-1 [16,17], suggesting they may signal through identical counter receptors or at least convergent pathways.

An *Escherichia coli* two-hybrid system coupled with phage display identified Gal-2 as a putative ligand for lymphotoxin α (LT α) [22]. LT α is a member of the TNFα cytokine family and is potentially involved in leukocyte activation and the pathogenesis of atherosclerosis [23]. Decreased Gal-2

expression correlates with concomitant decreases in LT α secretion, suggesting a possible function for this interaction in LT α secretory regulation [22]. A 3279 C-T polymorphism in the first intron of Gal-2 reduces expression of Gal-2 and increases susceptibility to myocardial infarction (MI) [22] in a Japanese population. However, it is unclear why reduced Gal-2 expression and therefore LT α secretion would put individuals at greater risk for MI. Furthermore, an analogous study examining a British population found no correlation between this polymorphism and risk for MI [24]. By contrast, another study suggests that the 3279 polymorphism actually conveys a decreased risk for MI [25]. Although additional genetic studies seeking to further establish this linkage have been conducted [26], studies to date have not defined the biochemical or cellular association between LT α and Gal-2.

28.2.3 Galectin-3

In contrast to the putative anti-inflammatory effects of Gal-1 on innate immunity, many studies suggest that Gal-3 activates innate immune function at several levels. For example, Gal-3 is also known as the Mac-2 antigen [27] and is present on thioglycollate-elicited and activated macrophages [28]. Gal-3 recognizes both IgE [29,30] and the high-affinity IgE receptor (FcεRI) [31] on the surface of mast cells. Binding and presumably cross-linking of FcεRI by Gal-3 induces degranulation [31], releasing potent proinflammatory constituents such as histamine [32]. In addition, early studies also demonstrated that blocking antibodies against Gal-3 attenuated IgE-mediated killing of parasites, suggesting a possible role in IgE effector functions [33]. Gal-3 null mast cells also exhibit impaired degranulation [34]. Because Gal-3 is expressed in nearly every tissue examined [35], release of Gal-3 from damaged tissue may trigger mast-cell degranulation independently of antigen, allowing immune system activation. In this way, Gal-3 may serve as a danger signal, signifying tissue damage and destruction, as suggested by others [36].

Besides activating mast cells, Gal-3 also exhibits activating effects on other key players involved in early immunity. Gal-3 induces respiratory burst in neutrophils [37,38] and monocytes [39], increasing the microbialcidal activity of these cells [40]. Gal-3 also accentuates lipopolysaccharide (LPS)-induced IL-1 secretion [41], a key cytokine responsible for activating endothelial cells, fever response, and inducing the expression of other proinflammatory mediators [42,43], further suggesting that Gal-3 also induces early steps in innate immune activation.

Gal-3 also induces neutrophil activation and preformed IL-8 secretion [41] and enhances phagocytic activity in both neutrophils [44] and macrophages [45], and exhibits chemotactic activity toward monocytes [46]. Interestingly, IL-8 itself induces chemotaxis of neutrophils, allowing a "trail" to be left for others to follow [47]. Gal-3 also has adhesion-promoting activity toward leukocyte in their interactions with extracellular matrix (ECM) components, such as laminin and fibronectin, through direct and indirect mechanisms [48]. This provides at least one possible mechanism whereby enhancement of chemotaxis may occur [46]. Gal-3 null mice exhibit increased macrophage apoptosis [49], and impaired macrophage phagocytosis [45], further suggesting a role for Gal-3 in the regulation of macrophage function, survival, and recruitment. Release of Gal-3 from damaged tissue or through some stimulated export system may, therefore, not only serve to activate innate immune responses to the presence of danger, but also direct phagocytes, in a concentration-dependent manner, toward areas where the greatest danger exists.

Gal-3 may not only indirectly activate the endothelium, increasing the extravasation of leukocytes through inducing cytokine release [31,41], but may also directly interact with leukocytes by serving as adhesion molecule in extravasation of neutrophils, as observed during streptococcal pneumonia [50] and monocyte movement into xenograph-transplanted tissue [51]. This latter activity primarily occurs through recognition of α-Gal-containing glycoconjugates [51]. However, although addition of recombinant Gal-3 enhances neutrophil adhesion to endothelial cells [50], it remains to be determined whether endogenous Gal-3, possibly upregulated during endothelial activation [35], may mediate these interactions. Furthermore, although haptens completely inhibit Gal-3 interaction with endothelial cells, no studies to date demonstrate that similar haptens inhibit extravasation

in vivo or *in vivo*. Consistent with the possibility that Gal-3 may be directly involved in extravasation, Gal-3 null mice exhibit attenuated leukocyte infiltration [49,52]. However, whether reduced leukocyte influx in these mice reflects lack of direct interaction with leukocytes or reduced release of endothelial activating cytokines induced by Gal-3 [31,41] remains to be determined.

Most proinflammatory roles for Gal-3 require the full-length protein [38,46,48,50,53,54]. Removal of the N-terminal domain, responsible for oligomerization of the protein [55], prevents many of these biological activities of Gal-3 [38,46,48,50,53,54]. Interestingly, activated neutrophils cleave the N-terminal domain of Gal-3 [53]. Furthermore, *Leishmania major* also cleaves the N-terminal domain of Gal-3, suggesting a possible mechanism of immune escape [56]. Cleavage and therefore generation of either the nonfunctional C terminal domain or even a product with the potential to provide direct negative feedback provides one plausible homeostatic mechanism for reducing proinflammatory effects of Gal-3. Consistent with this, the C-terminal domain of Gal-3, which has no carbohydrate-binding activity, actually inhibits monocyte chemotaxis [46]. Furthermore, differential activities of full-length Gal-3 toward cells may also result from unique and functional distinct oligomeric complexes formed on the surface of leukocytes [57,58]. This could allow Gal-3 to exhibit differential effects depending on the activation state and microenvironment of the target cell.

Gal-3 expression increases during differentiation of promyelocytic HL60 toward monocytes, suggesting a role in myeloid differentiation [59]. Gal-3 expression and surface association also increases upon monocyte activation *in vitro* [39], suggesting that Gal-3 may not only regulate immunity but also immune cell development and maturation. However, as these expression levels reflect correlative association, the requirements for Gal-3 during leukocyte development and maturation remain to be determined. Furthermore, it should be noted that in Gal-3 null mice there does not appear to be any alteration in leukocyte maturation or cell numbers in homeostasis.

28.2.4 Galectin-4

Only one study has examined potential involvement of Gal-4 in innate immunity. Similar to Gal-1 and Gal-2, Gal-4 also induces nonapoptotic PS exposure in activated neutrophils. However, unlike Gal-1 and Gal-2 [15–17], Gal-4-induced PS exposure occurs independent of intracellular Ca^{2+} mobilization [17], indicating that at least two pathways exist whereby galectin family members may modulate the turnover of neutrophils independent of apoptosis. Furthermore, cells exhibited increased sensitivity to Gal-4 when compared to Gal-1 and Gal-2, possibly due to the tandem repeat nature of this galectin [17].

28.2.5 Galectin-8

Several studies suggest that Gal-8, like Gal-3, may accentuate innate immune functions. For example, Gal-8 induces superoxide production in activated neutrophils [60], possibly through interactions with α M integrin. Gal-8 also accentuates the ability of matrix metalloproteinase 3 (MMP-3) to cleave MMP-9, a protease involved in the cleavage of ECM components, suggesting that Gal-8 may facilitate neutrophil migration through ECM [60].

28.2.6 Galectin-9

Many studies implicate Gal-9 as a key factor in eosinophil biology. Gal-9 induces chemotaxis of eosinophils following release from activated T cells [61,62]. In addition to inducing chemotaxis, Gal-9 also induces eosinophil activation, superoxide generation, and moderate degranulation [61,63]. HFL-1 fibroblast cells, derived from human lung tissue, exhibit increased Gal-9 expression and Gal-9-mediated eosinophil adhesion following IFNγ treatment, which may be important in retaining these cells within an inflammatory setting [64]. While it is clear that Gal-9 induces

eosinophil migration, conflicting results exist concerning the effects of Gal-9 on eosinophil viability. While one study suggests that Gal-9 induces apoptosis in eosinophils isolated from normal donors while protecting eosinophils from apoptosis isolated from patients with eosinophillia [65], another study indicates that Gal-9 decreases eosinophil apoptosis isolated from normal volunteers [63]. Gal-9 also induces maturation of DCs, implicating that it, along with Gal-3, may initiate adaptive immune function [66]. Finally, Gal-9 expression decreases during HL60 cell differentiation into eosinophils and monocytes, suggesting a possible role in myeloid linage development [59].

28.2.7 Galectin-10

Gal-10, also known as Charcot-Leyden crystal protein, was first identified in the blood and spleen of postmortem specimens from a leukemia patient [67] and also later in the sputum of an asthmatic patient [67]. Thus, it is fair to conclude that Gal-10 was actually the first galectin discovered, although scientists did not know it was a carbohydrate-binding protein until 1993 when the sequence was obtained and it was noted it had similarity to members of the S-type lectin family (galectins) such as Mac-2 (Gal-3) and Gal-1 [68]. Interestingly, the Charcot-Leyden protein was originally thought to have lysophospholipase activity, but it was subsequently shown that the protein is not a lysophospholipase but it binds to lysophospholipase inhibitor [69]. Gal-10 often crystallizes in patients demonstrating pronounced eosinophillia [70]. The functions of Gal-10 remain enigmatic. In a similar fashion to Gal-3 and Gal-9, Gal-10 expression levels change during myeloid development, increasing during eosinophilic differentiation of HL60 cells [59], although the functional relevance of this remains unclear. In addition, high expression of Gal-10 has been observed in a duodenal biopsy specimen of the aberrant clonal T cells in a patient with celiac disease [71].

28.2.8 Galectin-14

Gal-14 was recently isolated from eosinophils [72]. Among leukocytes, Gal-14 seems to also exhibit the highest expression in eosinophils, suggesting a specific role for this protein in eosinophil function [72]. Furthermore, lung tissue expression and release of Gal-14 into alveoli significantly increases following allergic challenge [72], also implicating the protein as having potential roles in immunity.

28.3 ADAPTIVE IMMUNITY

Adaptive immunity responds to a stimulating agent, such as proteins, carbohydrates, or pathogens, to improve recognition of the agent and after elimination of neutralization of the agent, the system retains immunological memory. In this way, the adaptive immune response can be faster and more robust to a second challenge by the agent. Thus, the adaptive immune response is distinguishable from the innate immune response, although they can cooperate to neutralize or eliminate challenges. A key feature of the adaptive immune response is the involvement of specialized cells, such as T and B cells, which have rearranged gene segments to produce antigen receptors (T-cell receptors [TCR] and antibodies) that allow a response to almost any type of antigen. Effector cells in adaptive immunity include B and T cells, and antigen-presenting cells (APC), such as DC, macrophages, and B cells. APCs process antigen and present them on the surface to immature or naive T cells, leading to mature helper T cells and mature cytotoxic or NK cells. Evidence has accumulated over the past 30 years that galectin interactions with cells of the adaptive immune system can lead to altered chemotaxis, turnover or apoptosis, and activation. This aspect of the chapter focuses on Gal-1, -2, -3, -4, -8, and -9, since little is known about the roles of the other galectin in adaptive immunity.

28.3.1 Galectin-1

Following the observation by Levi and Tiechburg that prophylactic administration of Gal-1 prevented the onset of experimental myasthenia gravis [2], it was some years later that Offner et al. demonstrated that Gal-1 also reduced the pathology associated with EAE [3]. Taken together, these results strongly implicated Gal-1 in adaptive immune regulation. However, the mechanism underlying these observations remained enigmatic. Baum et al. found high expression of Gal-1 in thymic epithelial cells [73], suggesting a potential role for Gal-1 in central tolerance. Gal-1 expressed in thymic epithelium mediates interactions with T leukemic MOLT-4 cells and preferentially recognizes cortical, immature thymocytes expressing Core 2 *O*-glycans [73]. Shortly thereafter, Perillo et al. demonstrated that Gal-1 induced apoptosis in activated T cells, several T leukemic cell lines [6], and immature thymocytes [74]. These results suggested that Gal-1-suppression of adaptive immune responses stemmed from the induction of apoptosis at central and peripheral sites.

Following the original description of the proapoptotic activity of Gal-1 toward activated T cells, scores of papers followed addressing this important potential activity. Subsequently, studies demonstrated that Gal-1 preferentially induces apoptosis in $CD4^+CD8^+$ double positive thymocytes, although all populations exhibited a certain degree of susceptibility [75]. Gal-1-induced apoptosis in thymocytes likely occurred through modulation of the TCR [75]. Further studies suggested that Gal-1 interacts with CD45, CD43, and CD7 to induce apoptosis in T cells, possibly due to differential segregation of these ligands into macrodomains on the cell surface [76,77]. For example, segregation of CD45 into separate membrane domains by Gal-1 may allow signaling events to occur [77,78], either through physically eliminating CD45 phosphatase activity from the active Gal-1 receptor [77,78] or intrinsically reducing CD45 phosphatase activity [79]. However, the absolute requirements for CD45 were not clear [80], and other studies have shown that CD45 expression is not required for Gal-1 to induce apoptosis in T cells [78]. Subsequent studies demonstrated that CD7 actually delivers Gal-1 apoptotic signals in T cells [81–83]. Ultimately, Gal-1 increases transcriptional factor AP-1 and prevents activation-induced increases in the antiapoptotic factor bcl-2 [84], providing some insight into the intracellular consequences of Gal-1 receptor binding. Furthermore, Gal-1 induction of cell death appears to occur independently of caspase activation [85], although additional findings suggest that Gal-1 does signal caspase activation [86].

Efforts to identify carbohydrate classes responsible for Gal-1-induced cell death suggested a requirement for Core 2 *O*-glycans, since it was found that susceptibility toward Gal-1-induced apoptosis in murine T lymphoma cell lines correlated with expression of the Core 2-β-1,6-*N*-acetylglucosaminyl transferase (Core 2 GnT) [78,87]. Importantly, expression of Core 2 GnT in BW5147 cells, which lack expression this enzyme, made the cells susceptible to Gal-1-induced apoptosis [78,87]. HIV infection of T cells increases Core 2 GnT expression and susceptibility to Gal-1 [88]. Furthermore, haploinsufficiency of Core 2 GnT-I also decreases the ability of Gal-1 to induce apoptosis [89], corroborating these earlier findings. However, in apparent contradiction to the conclusion that Core 2 *O*-glycans are required for Gal-1-induced apoptosis, treatment of activated human T cells with benzyl-GalNAc, an inhibitor of *O*-glycan synthesis, increased susceptibility to Gal-1 [6]. Additional studies on murine $CD8^+$ T cells suggested that Gal-1-induced apoptosis occurred independently of Core 2 GnT expression [90]. In addition, T leukemic Jurkat cells, which lack *O*-glycan elongation beyond the Tn antigen GalNAcα-Ser/Thr [91], not only retain susceptibility to Gal-1, but require *N*-glycans for Gal-1-induced apoptosis [92]. The utilization of T-cell lines, which often accumulate unidentified mutations, may in part account for these differences. For example, although some studies suggest that Jurkat and CEM cells fail to respond to Gal-1 [6,74], other studies suggest that Gal-1 induces apoptotic cell death in these cells [85,92], making it difficult to fully interpret the activity of Gal-1 toward T-cell viability in a physiological setting.

As previously described in Section 28.2 on innate immunity, Dias-Baruffi et al. demonstrated that Gal-1 could induce PS exposure in activated, but not resting neutrophils [15]. That study also demonstrated that Gal-1 induced PS exposure in T leukemic MOLT-4 cells without inducing DNA

fragmentation, altering cell size, or inhibiting the rate of cell division [15]. However, these studies contrasted to other studies demonstrating that Gal-1 induced apoptosis in MOLT-4 cells [6,74]. Recent evidence suggests that differential responses to Gal-1 may be in part due to the inclusion of the reducing agent dithiothreitol (DTT) during treatment conditions [17]. Gal-1 loses it activity within a day at 37°C in the absence of reducing agents [93], so some investigators had included DTT during treatments of cells with galectins. In an effort to control for activity loss, the effects of Gal-1 on T-cell viability were conducted at various concentrations of DTT, ranging from 1.1 to 3 mM [6,12,90,94–98]. Indeed, Gal-1 fails to alter T-cell viability in the absence of DTT [15,90]. However, a mutated form of Gal-1, C2S-Gal-1, which exhibits significant resistance to oxidative inactivation [93], does not induce apoptosis in T cells [15]. Furthermore, cells incubated with Gal-1 in the presence of reduced glutathione, an alternative cell membrane impermeable reducing agent [99], also fails to show signs of altered viability [15]. These results show that Gal-1 signals T cells to undergo apoptosis under cell stress induced by the toxic actions of DTT. DTT induces the unfolded protein response and itself can induce apoptosis in cells [100–103]. Interestingly, a recent report demonstrated that other transformed cell lines only become sensitive to Gal-1-induced apoptosis following environmental stress, such as hyperthermia, reduced CO_2, serum starvation, or infection with mycoplasm [104]. These results indicate that cells may have a Gal-1-sensitive pathway that may only be active during extreme cell stress. Interestingly, although Gal-1 fails to alter either PS exposure or viability in primary activated T cells without added DTT, it induces PS exposure in the absence of DTT in T leukemic MOLT-4 and CEM cell lines [15,17]. These results demonstrate the importance of looking at primary cells when evaluating the impact of Gal-1 on leukocyte viability. Furthermore, more definitive markers of apoptotic cell death need to be utilized when assessing whether Gal-1 induces apoptotic cell death.

Many studies also exhibit conflicting data concerning the concentration of Gal-1 required to induce apoptotic cell death. For example, early studies describing a proapoptotic function for Gal-1 demonstrated a minimal operating concentration of 10 μM or about 150 μg/mL [6], with most studies conducted at 20 μM [74] and no significant effects observed below 5 μM [6]. Subsequent studies examining Gal-1-induced PS exposure in the absence of cell death demonstrated similar operating concentrations [15]. However, other studies have utilized more than 10-fold lower concentrations of Gal-1 (4 μg/mL or 0.3 μM) [84,105] and found robust responses to Gal-1.

In spite of the use of DTT during *in vitro* assays, many studies continue to demonstrate a potent immunosuppressive activity for Gal-1 *in vivo*. These results demonstrate that Gal-1 possesses potent immunosuppressive activity, although recent studies [15] suggest that apoptosis-independent mechanisms responsible for this immunomodulation probably exist. New studies suggest that Gal-1 may actually modulate adaptive immune responses by altering cytokine production, shifting Th1/Th2 balance in adaptive immune responses. For example, injection of genetically engineered fibroblast that secrete Gal-1 significantly reduces the clinical and histopathological features associated with arthritis [105]. Interestingly, these effects were also accompanied by a shift in the adaptive immune response from a Th1 to a Th2 immune profile, indicated by a reduction in IFNγ and increases in IL-5 [105]. Gal-1 administration, while preventing Con A-induced hepatitis, also significantly reduced Con A-induced levels of IFNγ and TNFα [94]. Similarly, Gal-1 reduces potent Th1 cytokines TNFα, IL-1β, IL-12, and IFNγ normally observed following the induction of experimental colitis by 2,4,6-trinitrobenzene sulfonic acid (TNBS) [95]. Gal-1 also reduced levels of IL-2 and IFNγ in experimental models of graft-versus-host disease (GvHD). Interestingly, increased levels of CD4+CD25+ regulatory T cells (Tregs) also occurred following Gal-1 treatment in GvHD, suggesting a possible role for regulatory T cells in the immunosuppressive activities of Gal-1. Consistent with this, Gal-1 induces IL-10 [98], a key cytokine involved in Treg function [106], in both naive and activated T cells and increases IL-10 levels *in vivo* in animals with experimental autoimmune uveitis (EAU) [107]. Adoptive transfer of Tregs from EAU mice treated with Gal-1 prevented the development of EAU in nontreated mice, further suggesting that Gal-1 may actually mediate its immunosuppressive effect through Tregs [107]. Tregs isolated from Gal-1 null mice exhibit significantly compromised

ability to suppress [108], providing the first genetic data implicating Gal-1 in T-cell function. Interestingly, the possibility that Tregs may be the key players in mediating the immunosuppressive effects of Gal-1 suggests that Tiechburg's original hypothesis [2] may be correct.

Finally, it is important to note that in contrast to the proapoptotic effect of T-cell viability, several papers suggest that Gal-1 may actually protect T cells from death or even exhibit a certain degree of mitogenicity. One of Tiechburg's early papers demonstrated that Gal-1 actually increased the incorporation of radiolabeled thymidine in inguinal rabbit lymphocytes at lower concentrations (<3 μg/mL) [2]. Recently, Endharti et al. also demonstrated that Gal-1 may also be important in sustaining naive T-cell survival [109], suggesting that Gal-1 may serve as key stromal derived trophic factor for lymphocytes.

28.3.2 Galectin-2

Few studies have examined potential regulatory roles of Gal-2 in adaptive immunity. One study suggested that Gal-2 may induce apoptosis in T cells through activation of caspases [110]. However, this study used experimental conditions in which Gal-1 also induced apoptosis, raising the possibility that the responses may be due to the inclusion of DTT. Interestingly, Gal-2-induced apoptosis required significantly more time than previously reported for Gal-1, requiring 24 h for robust PS exposure to be realized [110], in contrast to the 30 min required for maximal cell death induced by Gal-1 [6]. This was not likely a reflection of relative affinity, as Gal-2 was more potent than Gal-1 in this study [110]. In contrast to its reported effects on viability, Gal-2 also decreased IFNγ and TNFα levels while inducing both IL-10 and IL-5 [110], suggesting that Gal-2 may regulate adaptive immunity by shifting TH1 responses to more tolerable TH2 outcomes. By contrast, Gal-2 was recently reported to induce PS exposure in T leukemic MOLT-4 cells independent of apoptosis, while exhibiting no effect on PS distribution or viability on primary activated T cells unless coincubated with DTT [17]. No *in vivo* studies have been conducted to date to determine whether Gal-2 might exhibit similar immunosuppressive features as Gal-1. This may be important, since Gal-1-null mice exhibit a moderate phenotype following immunological challenge, suggesting that other members may compensate *in vivo*.

28.3.3 Galectin-3

In contrast to the putative proapoptotic activities of extracellular Gal-1, several studies suggest that Gal-3 actually inhibits apoptosis intracellularly when it is overexpressed [111]. For example, Gal-3 increases the growth rate and decreases sensitivity to proapoptotic agents such as anti-Fas and staurosporine [111] when overexpressed in T leukemic Jurkat cells, presumably through interactions with bcl-2 [111,112]. However, it should be noted that no alterations in T-cell viability have been reported in Gal-3 null mice. Attempts to resolve the potential involvement of Gal-3 in leukocyte viability utilizing siRNA knockdown approaches also give conflicting results. For example, knockdown of Gal-3 increased proliferation and reduced activation induced cell death in primary peripheral blood mononuclear cell (PBMCs) [113] in one study, while Gal-3 knockdown expression significantly attenuates T-cell proliferation [114] in another study. Knockdown of Gal-3 in B cells blocked IL-4-induced survival of activated B cells, favoring plasma cell differentiation, implicating roles for Gal-3 in memory B-cell development [115]. Similarly, Gal-3 exhibits an antiapoptotic role in B-cell lymphomas [116]. In contrast, intracellular Gal-3 may simply modulate the type of signaling response initiated by death receptor ligation [117].

Early reports suggested that extracellular Gal-3 exhibited no direct effects on T-cell viability [6,118] and may even antagonize the effects of Gal-1-induced apoptosis [118]. Similarly, Gal-3 also exhibited no effect on thymocyte viability [119], although Gal-3 seems to regulate other aspects of central tolerance, such as adhesion of thymocytes to the thymic epithelium [120]. By contrast, recent reports suggest that Gal-3 induces apoptosis in T cells and thymocytes [121,122]. Consistent with

this, a positive correlation between Gal-3, but not Gal-1, and lymphocyte apoptosis was found within melanoma neoplastic lesions, suggesting that Gal-3 may actually exhibit proapoptotic effects [123]. Furthermore, patients with inflammatory bowel diseases exhibit decreased expression of Gal-3, which suggests that chronic inflammation may in part result from decreased regulatory capacity of adaptive immunity by these tissues [113].

In contrast to the effects of Gal-3 on leukocyte viability, Gal-3 may also modulate other aspects of adaptive immunity. For example, Gal-3 inhibits IL-5 production, a key cytokine in Th2-adaptive immune responses, in eosinophil cell line EoL-3, PBMCs, and $CD4^+$ T cell lines. By contrast, Gal-3-null mice exhibit reduced eosinophil infiltration and airway hyperresponsiveness following allergic challenge [124], suggesting that Gal-3 actually stimulates Th2 responses when present. Lactose or anti-Gal-3 blocking antibodies also inhibit polymorphonuclear leukocyte-mediated increases in spontaneous IgE production in patients with IgE-associated atopic eczema/dermatitis syndrome (AEDS), further implicating Gal-3 in the development of TH2 adaptive immune responses. Gal-3 may also regulate T-cell activation. For example, Gal-3 enhances adhesion of naive T cells to DCs, possibly facilitating activation and proliferation [125]. By contrast, Gal-3 may also bind TCR, antagonizing immunological synapse formation, a process likely regulated by β-1,6-*N*-acetylglucosaminyltransferase V (Mgat5) [126].

28.3.4 Galectin-4 and Galectin-8

Very few studies have examined potential roles of either Gal-4 or Gal-8 in adaptive immunity. Gal-4 induces IL-6 production and is implicated in the pathogenesis of inflammatory bowel disease [127]. Furthermore, Gal-4 exhibits no effect on T-cell viability [17,127].

Gal-8 induces Rac-1 activation and profound morphological changes in T leukemic Jurkat cells [128,129]. In addition, anti-Gal-8 blocking antibodies have been demonstrated in patients with systemic lupus erythematosus (SLE). However, the physiological consequences of both these observations are unknown [128,129].

28.3.5 Galectin-9

The first studies implicating a potential role for Gal-9 in adaptive immunity demonstrated that Gal-9 induced apoptosis in thymocytes [119]. However, these studies, like Gal-1, utilized 3 mM DTT in all treatment conditions [119]. However, consistent with a potential role for Gal-9 in T-cell turnover, administration of Gal-9 not only reduced pathology associated with experimental autoimmune nephritis, but also increased apoptosis in $CD8^+$, but not $CD4^+$ T cells and restored kidney function [130]. Subsequent studies demonstrated that Gal-9 induces apoptosis in several T leukemic cell lines through a calpain- and caspase-1-dependent pathway [131]. Gal-9 has also been reported to selectively induce apoptosis in $CD4^+$ TH1 cells, possibly through interactions with Tim3, a TH1-specific receptor [132] in the absence of DTT, providing a potential feedback mechanism in the regulation of TH1-mediated adaptive immune responses. Consistent with this, administration of Gal-9 induced selective loss of TH1 IFNγ producing T cells [132].

28.3.6 Galectin-10

In addition to being expressed in eosinophils, recent evidence utilizing a proteomics approach identified Gal-10 as a specific intracellular marker of $CD4^+$ $CD25^+$ Tregs [133]. Knockdown of Gal-10 in $CD4^+$ $CD25^+$ Tregs rendered them sensitive to proliferation-inducing stimulation, changed cytokine production, and significantly compromised their suppressive behavior [133]. Ectopic expression of Gal-10 also induced caspase-dependent apoptosis in $CD4^+$ T cells. Although these alterations were attributed to intracellular Gal-10, it is not clear whether the detection methods utilized enabled Gal-10 recognition if bound to cell surface ligands [133].

28.4 CONCLUSION AND FUTURE DIRECTIONS

Galectins may serve as key immune regulatory elements in a manner analogous to Matzinger's danger theory of host–parasite interactions [134,135], signaling danger not only when pathogens induce host-cell death, but perhaps also when the immune system itself becomes a danger. Galectins have nearly ubiquitous expression [35,136], and are able to induce key changes in leukocytes, including apoptotic death, removal, and turnover, or immunosuppressive alterations in cytokine production. A partial summary of these activities is shown in Table 28.1. These activities place galectins in a key position to signal leukocyte elimination or shift toward a more tolerable Th2 outcome following release during excessive immune mediated tissue damage.

ACKNOWLEDGMENT

The work of the authors was supported by NIH Grant P01 HL085607 to R.D.C.

REFERENCES

1. Teichberg, V. I., Silman, I., Beitsch, D. D., and Resheff, G. A beta-D-galactoside binding protein from electric organ tissue of *Electrophorus electricus. Proc Natl Acad Sci USA* 72, 1383–1387, 1975.
2. Levi, G., Tarrab-Hazdai, R., and Teichberg, V. I. Prevention and therapy with electrolectin of experimental autoimmune myasthenia gravis in rabbits. *Eur J Immunol* 13, 500–507, 1983.
3. Offner, H. et al. Recombinant human beta-galactoside binding lectin suppresses clinical and histological signs of experimental autoimmune encephalomyelitis. *J Neuroimmunol* 28, 177–184, 1990.
4. Levi, G. and Teichberg, V. I. Selective interactions of electrolectins from eel electric organ and mouse thymus with mouse immature thymocytes. *Immunol Lett* 7, 35–39, 1983.
5. Moller, G. Do suppressor T cells exist? *Scand J Immunol* 27, 247–250, 1988.
6. Perillo, N. L., Pace, K. E., Seilhamer, J. J., and Baum, L. G. Apoptosis of T cells mediated by galectin-1. *Nature* 378, 736–739, 1995.
7. La, M. et al. A novel biological activity for galectin-1: Inhibition of leukocyte-endothelial cell interactions in experimental inflammation. *Am J Pathol* 163, 1505–1515, 2003.
8. Gil, C. D., Cooper, D., Rosignoli, G., Perretti, M., and Oliani, S. M. Inflammation-induced modulation of cellular galectin-1 and -3 expression in a model of rat peritonitis. *Inflamm Res* 55, 99–107, 2006.
9. Delbrouck, C. et al. Galectin-1 is overexpressed in nasal polyps under budesonide and inhibits eosinophil migration. *Lab Invest* 82, 147–158, 2002.
10. Rubinstein, N., Ilarregui, J. M., Toscano, M. A., and Rabinovich, G. A. The role of galectins in the initiation, amplification and resolution of the inflammatory response. *Tissue Antigens* 64, 1–12, 2004.
11. Almkvist, J., Dahlgren, C., Leffler, H., and Karlsson, A. Activation of the neutrophil nicotinamide adenine dinucleotide phosphate oxidase by galectin-1. *J Immunol* 168, 4034–4041, 2002.
12. Perone, M. J. et al. Dendritic cells expressing transgenic galectin-1 delay onset of autoimmune diabetes in mice. *J Immunol* 177, 5278–5289, 2006.
13. Lagasse, E. and Weissman, I. L. bcl-2 inhibits apoptosis of neutrophils but not their engulfment by macrophages. *J Exp Med* 179, 1047–1052, 1994.
14. Shi, J., Gilbert, G. E., Kokubo, Y., and Ohashi, T. Role of the liver in regulating numbers of circulating neutrophils. *Blood* 98, 1226–1230, 2001.
15. Dias-Baruffi, M. et al. Dimeric galectin-1 induces surface exposure of phosphatidylserine and phagocytic recognition of leukocytes without inducing apoptosis. *J Biol Chem* 278, 41282–41293, 2003.
16. Karmakar, S., Cummings, R. D., and McEver, R. P. Contributions of Ca^{2+} to Galectin-1-induced exposure of phosphatidylserine on activated neutrophils. *J Biol Chem*, 280, 28623–28631, 2005.
17. Stowell, S. et al. Human Galectins-1, -2, and -4 induce surface exposure of phosphatidylserine in activated human neutrophils but not activated T cells. *Blood*, 109, 219–227, 2006.
18. Williamson, P. and Schlegel, R. A. Hide and seek: The secret identity of the phosphatidylserine receptor. *J Biol* 3, 14, 2004.
19. Sahu, S. K., Gummadi, S. N., Manoj, N., and Aradhyam, G. K. Phospholipid scramblases: An overview. *Arch Biochem Biophys* 462, 103–114, 2007.
20. Williamson, P. and Schlegel, R. A. Transbilayer phospholipid movement and the clearance of apoptotic cells. *Biochim Biophys Acta* 1585, 53–63, 2002.

20b. Stowell, S. R., Qian, Y., et al. Differential roles of galectin-1 and galectin-3 in regulating leukocyte viability and cytokine secretion. *J Immunol* 180, 3091–3102, 2008.
21. Gitt, M. A., Massa, S. M., Leffler, H., and Barondes, S. H. Isolation and expression of a gene encoding L-14-II, a new human soluble lactose-binding lectin. *J Biol Chem* 267, 10601–10606, 1992.
22. Ozaki, K. et al. Functional variation in LGALS2 confers risk of myocardial infarction and regulates lymphotoxin-alpha secretion in vitro. *Nature* 429, 72–75, 2004.
23. Naoum, J. J. et al. Lymphotoxin-alpha and cardiovascular disease: Clinical association and pathogenic mechanisms. *Med Sci Monit* 12, RA121–RA124, 2006.
24. Mangino, M., Braund, P., et al. LGALS2 functional variant rs7291467 is not associated with susceptibility to myocardial infarction in Caucasians. *Atherosclerosis*, 194, 112–115, 2007.
25. Asselbergs, F. W., Pai, J. K., Rexrode, K. M., Hunter, D. J., and Rimm, E. B. Effects of lymphotoxin-alpha gene and galectin-2 gene polymorphisms on inflammatory biomarkers, cellular adhesion molecules and risk of coronary heart disease. *Clin Sci (Lond)* 112, 291–298, 2007.
26. Topol, E. J., Smith, J., Plow, E. F., and Wang, Q. K. Genetic susceptibility to myocardial infarction and coronary artery disease. *Hum Mol Genet 15 Spec No* 2, R117–R123, 2006.
27. Cherayil, B. J., Weiner, S. J., and Pillai, S. The Mac-2 antigen is a galactose-specific lectin that binds IgE. *J Exp Med* 170, 1959–1972, 1989.
28. Frigeri, L. G. and Liu, F. T. Surface expression of functional IgE binding protein, an endogenous lectin, on mast cells and macrophages. *J Immunol* 148, 861–867, 1992.
29. Liu, F. T., Albrandt, K., Mendel, E., Kulczycki, A., Jr., and Orida, N. K. Identification of an IgE-binding protein by molecular cloning. *Proc Natl Acad Sci USA* 82, 4100–4104, 1985.
30. Robertson, M. W., Albrandt, K., Keller, D., and Liu, F. T. Human IgE-binding protein: A soluble lectin exhibiting a highly conserved interspecies sequence and differential recognition of IgE glycoforms. *Biochemistry* 29, 8093–8100, 1990.
31. Frigeri, L. G., Zuberi, R. I., and Liu, F. T. Epsilon BP, a beta-galactoside-binding animal lectin, recognizes IgE receptor (Fc epsilon RI) and activates mast cells. *Biochemistry* 32, 7644–7649, 1993.
32. Crone, C. Modulation of solute permeability in microvascular endothelium. *Fed Proc* 45, 77–83, 1986.
33. Truong, M. J. et al. IgE-binding molecules (Mac-2/epsilon BP) expressed by human eosinophils. Implication in IgE-dependent eosinophil cytotoxicity. *Eur J Immunol* 23, 3230–3235, 1993.
34. Chen, H. Y. et al. Role of galectin-3 in mast cell functions: Galectin-3-deficient mast cells exhibit impaired mediator release and defective JNK expression. *J Immunol* 177, 4991–4997, 2006.
35. Dumic, J., Dabelic, S., and Flogel, M. Galectin-3: An open-ended story. *Biochim Biophys Acta* 1760, 616–635, 2006.
36. Sato, S. and Nieminen, J. Seeing strangers or announcing "danger": Galectin-3 in two models of innate immunity. *Glycoconj J* 19, 583–591, 2004.
37. Almkvist, J., Faldt, J., Dahlgren, C., Leffler, H., and Karlsson, A. Lipopolysaccharide-induced gelatinase granule mobilization primes neutrophils for activation by galectin-3 and formylmethionyl-Leu-Phe. *Infect Immun* 69, 832–837, 2001.
38. Yamaoka, A., Kuwabara, I., Frigeri, L. G., and Liu, F. T. A human lectin, galectin-3 (epsilon bp/Mac-2), stimulates superoxide production by neutrophils. *J Immunol* 154, 3479–3487, 1995.
39. Liu, F. T. et al. Expression and function of galectin-3, a beta-galactoside-binding lectin, in human monocytes and macrophages. *Am J Pathol* 147, 1016–1028, 1995.
40. Djaldetti, M., Salman, H., Bergman, M., Djaldetti, R., and Bessler, H. Phagocytosis—the mighty weapon of the silent warriors. *Microsc Res Tech* 57, 421–431, 2002.
41. Jeng, K. C., Frigeri, L. G., and Liu, F. T. An endogenous lectin, galectin-3 (epsilon BP/Mac-2), potentiates IL-1 production by human monocytes. *Immunol Lett* 42, 113–116, 1994.
42. Ward, P. A., Warren, J. S., and Johnson, K. J. Oxygen radicals, inflammation, and tissue injury. *Free Radic Biol Med* 5, 403–408, 1988.
43. Pober, J. S. Effects of tumour necrosis factor and related cytokines on vascular endothelial cells. *Ciba Found Symp* 131, 170–184, 1987.
44. Fernandez, G. C., Ilarregui, J. M., et al. Galectin-3 and soluble fibrinogen act in concert to modulate neutrophil activation and survival. Involvement of alternative MAPK-pathways. *Glycobiology*, 15, 519–527, 2005.
45. Sano, H. et al. Critical role of galectin-3 in phagocytosis by macrophages. *J Clin Invest* 112, 389–397, 2003.
46. Sano, H. et al. Human galectin-3 is a novel chemoattractant for monocytes and macrophages. *J Immunol* 165, 2156–2164, 2000.
47. Baggiolini, M. and Clark-Lewis, I. Interleukin-8, a chemotactic and inflammatory cytokine. *FEBS Lett* 307, 97–101, 1992.

48. Kuwabara, I. and Liu, F. T. Galectin-3 promotes adhesion of human neutrophils to laminin. *J Immunol* 156, 3939–3944, 1996.
49. Hsu, D. K. et al. Targeted disruption of the galectin-3 gene results in attenuated peritoneal inflammatory responses. *Am J Pathol* 156, 1073–1083, 2000.
50. Sato, S. et al. Role of galectin-3 as an adhesion molecule for neutrophil extravasation during streptococcal pneumonia. *J Immunol* 168, 1813–1822, 2002.
51. Jin, R., Greenwald, A., Peterson, M. D., and Waddell, T. K. Human monocytes recognize porcine endothelium via the interaction of galectin 3 and alpha-GAL. *J Immunol* 177, 1289–1295, 2006.
52. Colnot, C. et al. Maintenance of granulocyte numbers during acute peritonitis is defective in galectin-3-null mutant mice. *Immunology* 94, 290–296, 1998.
53. Nieminen, J., St-Pierre, C., and Sato, S. Galectin-3 interacts with naive and primed neutrophils, inducing innate immune responses. *J Leukoc Biol* 78, 1127–1135, 2005.
54. Karlsson, A., Follin, P., Leffler, H., and Dahlgren, C. Galectin-3 activates the NADPH-oxidase in exudated but not peripheral blood neutrophils. *Blood* 91, 3430–3438, 1998.
55. Hsu, D. K., Zuberi, R. I., and Liu, F. T. Biochemical and biophysical characterization of human recombinant IgE-binding protein, an S-type animal lectin. *J Biol Chem* 267, 14167–14174, 1992.
56. Pelletier, I. and Sato, S. Specific recognition and cleavage of galectin-3 by *Leishmania major* through species-specific polygalactose epitope. *J Biol Chem* 277, 17663–17670, 2002.
57. Massa, S. M., Cooper, D. N., Leffler, H., and Barondes, S. H. L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity. *Biochemistry* 32, 260–267, 1993.
58. Nieminen, J., Kuno, A., Hirabayashi, J., and Sato, S. Visualization of galectin-3 oligomerization on the surface of neutrophils and endothelial cells using fluorescence resonance energy transfer. *J Biol Chem* 282, 1374–1383, 2007.
59. Abedin, M. J., Kashio, Y., Seki, M., Nakamura, K., and Hirashima, M. Potential roles of galectins in myeloid differentiation into three different lineages. *J Leukoc Biol* 73, 650–656, 2003.
60. Nishi, N. et al. Galectin-8 modulates neutrophil function via interaction with integrin alphaM. *Glycobiology* 13, 755–763, 2003.
61. Matsumoto, R. et al. Human ecalectin, a variant of human galectin-9, is a novel eosinophil chemoattractant produced by T lymphocytes. *J Biol Chem* 273, 16976–16984, 1998.
62. Matsushita, N. et al. Requirement of divalent galactoside-binding activity of ecalectin/galectin-9 for eosinophil chemoattraction. *J Biol Chem* 275, 8355–8360, 2000.
63. Matsumoto, R., Hirashima, M., Kita, H., and Gleich, G. J. Biological activities of ecalectin: A novel eosinophil-activating factor. *J Immunol* 168, 1961–1967, 2002.
64. Asakura, H. et al. Selective eosinophil adhesion to fibroblast via IFN-gamma-induced galectin-9. *J Immunol* 169, 5912–5918, 2002.
65. Saita, N. et al. Association of galectin-9 with eosinophil apoptosis. *Int Arch Allergy Immunol* 128, 42–50, 2002.
66. Dai, S. Y. et al. Galectin-9 induces maturation of human monocyte-derived dendritic cells. *J Immunol* 175, 2974–2981, 2005.
67. Charcot, J. M. and Robin, C. Observation de leucocythemie. *C R Seances Mem Soc Biol Paris* 5, 44–52, 1854.
68. Ackerman, S. J. et al. Molecular cloning and characterization of human eosinophil Charcot-Leyden crystal protein (lysophospholipase). Similarities to IgE binding proteins and the S-type animal lectin superfamily. *J Immunol* 150, 456–468, 1993.
69. Ackerman, S. J. et al. Charcot-Leyden crystal protein (galectin-10) is not a dual function galectin with lysophospholipase activity but binds a lysophospholipase inhibitor in a novel structural fashion. *J Biol Chem* 277, 14859–14868, 2002.
70. Gleich, G. J., Loegering, D. A., Mann, K. G., and Maldonado, J. E. Comparative properties of the Charcot-Leyden crystal protein and the major basic protein from human eosinophils. *J Clin Invest* 57, 633–640, 1976.
71. De Re, V. et al. Proteins specifically hyperexpressed in a coeliac disease patient with aberrant T cells. *Clin Exp Immunol* 148, 402–409, 2007.
72. Dunphy, J. L. et al. Isolation and characterization of a novel eosinophil-specific galectin released into the lungs in response to allergen challenge. *J Biol Chem* 277, 14916–14924, 2002.
73. Baum, L. G. et al. Human thymic epithelial cells express an endogenous lectin, galectin-1, which binds to core 2 O-glycans on thymocytes and T lymphoblastoid cells. *J Exp Med* 181, 877–887, 1995.
74. Perillo, N. L., Uittenbogaart, C. H., Nguyen, J. T., and Baum, L. G. Galectin-1, an endogenous lectin produced by thymic epithelial cells, induces apoptosis of human thymocytes. *J Exp Med* 185, 1851–1858, 1997.

75. Vespa, G. N. et al. Galectin-1 specifically modulates TCR signals to enhance TCR apoptosis but inhibit IL-2 production and proliferation. *J Immunol* 162, 799–806, 1999.
76. Symons, A., Cooper, D. N., and Barclay, A. N. Characterization of the interaction between galectin-1 and lymphocyte glycoproteins CD45 and Thy-1. *Glycobiology* 10, 559–563, 2000.
77. Pace, K. E., Lee, C., Stewart, P. L., and Baum, L. G. Restricted receptor segregation into membrane microdomains occurs on human T cells during apoptosis induced by galectin-1. *J Immunol* 163, 3801–3811, 1999.
78. Nguyen, J. T. et al. CD45 modulates galectin-1-induced T cell death: Regulation by expression of core 2 O-glycans. *J Immunol* 167, 5697–5707, 2001.
79. Walzel, H., Schulz, U., Neels, P., and Brock, J. Galectin-1, a natural ligand for the receptor-type protein tyrosine phosphatase CD45. *Immunol Lett* 67, 193–202, 1999.
80. Fajka-Boja, R. et al. Receptor tyrosine phosphatase, CD45 binds galectin-1 but does not mediate its apoptotic signal in T cell lines. *Immunol Lett* 82, 149–154, 2002.
81. Pace, K. E., Hahn, H. P., Pang, M., Nguyen, J. T., and Baum, L. G. CD7 delivers a pro-apoptotic signal during galectin-1-induced T cell death. *J Immunol* 165, 2331–2334, 2000.
82. Rappl, G. et al. CD4+CD7- leukemic T cells from patients with Sezary syndrome are protected from galectin-1-triggered T cell death. *Leukemia* 16, 840–845, 2002.
83. Roberts, A. A. et al. Galectin-1-mediated apoptosis in mycosis fungoides: The roles of CD7 and cell surface glycosylation. *Mod Pathol* 16, 543–551, 2003.
84. Rabinovich, G. A. et al. Molecular mechanisms implicated in galectin-1-induced apoptosis: Activation of the AP-1 transcription factor and downregulation of Bcl-2. *Cell Death Differ* 7, 747–753, 2000.
85. Hahn, H. P. et al. Galectin-1 induces nuclear translocation of endonuclease G in caspase- and cytochrome c-independent T cell death. *Cell Death Differ* 11, 1277–1286, 2004.
86. Matarrese, P. et al. Galectin-1 sensitizes resting human T lymphocytes to Fas (CD95)-mediated cell death via mitochondrial hyperpolarization, budding, and fission. *J Biol Chem* 280, 6969–6985, 2005.
87. Galvan, M., Tsuboi, S., Fukuda, M., and Baum, L. G. Expression of a specific glycosyltransferase enzyme regulates T cell death mediated by galectin-1. *J Biol Chem* 275, 16730–16737, 2000.
88. Lanteri, M. et al. Altered T cell surface glycosylation in HIV-1 infection results in increased susceptibility to galectin-1-induced cell death. *Glycobiology* 13, 909–918, 2003.
89. Cabrera, P. V. et al. Haploinsufficiency of C2GnT-I glycosyltransferase renders T lymphoma cells resistant to cell death. *Blood* 108, 2399–2406, 2006.
90. Carlow, D. A., Williams, M. J., and Ziltener, H. J. Modulation of O-glycans and N-glycans on murine CD8 T cells fails to alter annexin V ligand induction by galectin 1. *J Immunol* 171, 5100–5106, 2003.
91. Ju, T. and Cummings, R. D. A unique molecular chaperone Cosmc required for activity of the mammalian core 1 beta 3-galactosyltransferase. *Proc Natl Acad Sci USA* 99, 16613–16618, 2002.
92. Walzel, H. et al. Effects of N-glycan processing inhibitors on signaling events and induction of apoptosis in galectin-1-stimulated Jurkat T lymphocytes. *Glycobiology* 16, 1262–1271, 2006.
93. Hirabayashi, J. and Kasai, K. Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. *J Biol Chem* 266, 23648–23653, 1991.
94. Santucci, L. et al. Galectin-1 exerts immunomodulatory and protective effects on concanavalin A-induced hepatitis in mice. *Hepatology* 31, 399–406, 2000.
95. Santucci, L. et al. Galectin-1 suppresses experimental colitis in mice. *Gastroenterology* 124, 1381–1394, 2003.
96. Rabinovich, G. A. et al. Induction of allogenic T-cell hyporesponsiveness by galectin-1-mediated apoptotic and non-apoptotic mechanisms. *Cell Death Differ* 9, 661–670, 2002.
97. Pace, K. E., Hahn, H. P., and Baum, L. G. Preparation of recombinant human galectin-1 and use in T-cell death assays. *Methods Enzymol* 363, 499–518, 2003.
98. van der Leij, J. et al. Dimeric galectin-1 induces IL-10 production in T-lymphocytes: An important tool in the regulation of the immune response. *J Pathol* 204, 511–518, 2004.
99. Soltysiak-Pawluczuk, D., Naciazek-Wieniawska, A., Danysz, A., and Czarnomska, A. Effects of thiol compounds on methotrexate uptake by murine lymphocytes from thymus and thymic lymphosarcoma. *Cancer Lett* 65, 251–257, 1992.
100. Murray, J. I. et al. Diverse and specific gene expression responses to stresses in cultured human cells. *Mol Biol Cell* 15, 2361–2374, 2004.
101. Frand, A. R. and Kaiser, C. A. The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Mol Cell* 1, 161–170, 1998.

102. Tartier, L., McCarey, Y. L., Biaglow, J. E., Kochevar, I. E., and Held, K. D. Apoptosis induced by dithiothreitol in HL-60 cells shows early activation of caspase 3 and is independent of mitochondria. *Cell Death Differ* 7, 1002–1010, 2000.
103. Braakman, I., Helenius, J., and Helenius, A. Manipulating disulfide bond formation and protein folding in the endoplasmic reticulum. *Embo J* 11, 1717–1722, 1992.
104. Wiest, I., Seliger, C., Walzel, H., Friese, K., and Jeschke, U. Induction of apoptosis in human breast cancer and trophoblast tumor cells by galectin-1. *Anticancer Res* 25, 1575–1580, 2005.
105. Rabinovich, G. A. et al. Recombinant galectin-1 and its genetic delivery suppress collagen-induced arthritis via T cell apoptosis. *J Exp Med* 190, 385–398, 1999.
106. Asseman, C., Mauze, S., Leach, M. W., Coffman, R. L., and Powrie, F. An essential role for interleukin 10 in the function of regulatory T cells that inhibit intestinal inflammation. *J Exp Med* 190, 995–1004, 1999.
107. Toscano, M. A. et al. Galectin-1 suppresses autoimmune retinal disease by promoting concomitant Th2- and T regulatory-mediated anti-inflammatory responses. *J Immunol* 176, 6323–6332, 2006.
108. Garin, M. I. et al. Galectin-1: A key effector of regulation mediated by CD4+CD25+ T cells. *Blood* 109, 2058–2065, 2007.
109. Endharti, A. T., Zhou, Y. W., Nakashima, I., and Suzuki, H. Galectin-1 supports survival of naive T cells without promoting cell proliferation. *Eur J Immunol* 35, 86–97, 2005.
110. Sturm, A. et al. Human galectin-2: Novel inducer of T cell apoptosis with distinct profile of caspase activation. *J Immunol* 173, 3825–3837, 2004.
111. Yang, R. Y., Hsu, D. K., and Liu, F. T. Expression of galectin-3 modulates T-cell growth and apoptosis. *Proc Natl Acad Sci USA* 93, 6737–6742, 1996.
112. Akahani, S., Nangia-Makker, P., Inohara, H., Kim, H. R., and Raz, A. Galectin-3: A novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family. *Cancer Res* 57, 5272–5276, 1997.
113. Muller, S. et al. Galectin-3 modulates T cell activity and is reduced in the inflamed intestinal epithelium in IBD. *Inflamm Bowel Dis* 12, 588–597, 2006.
114. Joo, H. G. et al. Expression and function of galectin-3, a beta-galactoside-binding protein in activated T lymphocytes. *J Leukoc Biol* 69, 555–564, 2001.
115. Acosta-Rodriguez, E. V. et al. Galectin-3 mediates IL-4-induced survival and differentiation of B cells: Functional cross-talk and implications during *Trypanosoma cruzi* infection. *J Immunol* 172, 493–502, 2004.
116. Hoyer, K. K. et al. An anti-apoptotic role for galectin-3 in diffuse large B-cell lymphomas. *Am J Pathol* 164, 893–902, 2004.
117. Fukumori, T. et al. Endogenous galectin-3 determines the routing of CD95 apoptotic signaling pathways. *Cancer Res* 64, 3376–3379, 2004.
118. Ahmad, N. et al. Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. *J Biol Chem* 279, 10841–10847, 2004.
119. Wada, J., Ota, K., Kumar, A., Wallner, E. I., and Kanwar, Y. S. Developmental regulation, expression, and apoptotic potential of galectin-9, a beta-galactoside binding lectin. *J Clin Invest* 99, 2452–2461, 1997.
120. Villa-Verde, D. M. et al. Galectin-3 modulates carbohydrate-dependent thymocyte interactions with the thymic microenvironment. *Eur J Immunol* 32, 1434–1444, 2002.
121. Fukumori, T. et al. CD29 and CD7 mediate galectin-3-induced type II T-cell apoptosis. *Cancer Res* 63, 8302–8311, 2003.
122. Stillman, B. N. et al. Galectin-3 and galectin-1 bind distinct cell surface glycoprotein receptors to induce T cell death. *J Immunol* 176, 778–789, 2006.
123. Zubieta, M. R. et al. Galectin-3 expression correlates with apoptosis of tumor-associated lymphocytes in human melanoma biopsies. *Am J Pathol* 168, 1666–1675, 2006.
124. Zuberi, R. I. et al. Critical role for galectin-3 in airway inflammation and bronchial hyperresponsiveness in a murine model of asthma. *Am J Pathol* 165, 2045–2053, 2004.
125. Swarte, V. V., Mebius, R. E., Joziasse, D. H., Van den Eijnden, D. H., and Kraal, G. Lymphocyte triggering via L-selectin leads to enhanced galectin-3-mediated binding to dendritic cells. *Eur J Immunol* 28, 2864–2871, 1998.
126. Demetriou, M., Granovsky, M., Quaggin, S., and Dennis, J. W. Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation. *Nature* 409, 733–739, 2001.
127. Hokama, A. et al. Induced reactivity of intestinal CD4(+) T cells with an epithelial cell lectin, galectin-4, contributes to exacerbation of intestinal inflammation. *Immunity* 20, 681–693, 2004.

128. Carcamo, C. et al. Galectin-8 binds specific beta1 integrins and induces polarized spreading highlighted by asymmetric lamellipodia in Jurkat T cells. *Exp Cell Res* 312, 374–386, 2006.
129. Pardo, E. et al. [Antibodies against galectin-8 in patients with systemic lupus erythematosus]. *Rev Med Chil* 134, 159–166, 2006.
130. Tsuchiyama, Y. et al. Efficacy of galectins in the amelioration of nephrotoxic serum nephritis in Wistar Kyoto rats. *Kidney Int* 58, 1941–1952, 2000.
131. Kashio, Y. et al. Galectin-9 induces apoptosis through the calcium-calpain-caspase-1 pathway. *J Immunol* 170, 3631–3636, 2003.
132. Zhu, C. et al. The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nat Immunol* 6, 1245–1252, 2005.
133. Kubach, J., Lutter, P., et al. Human CD4+CD25+ regulatory T cells: Proteome analysis identifies galectin-10 as a novel marker essential for their anergy and suppressive function. *Blood*, 110, 1550–1558, 2007.
134. Matzinger, P. Tolerance, danger, and the extended family. *Annu Rev Immunol* 12, 991–1045, 1994.
135. Matzinger, P. The danger model: A renewed sense of self. *Science* 296, 301–305, 2002.
136. Camby, I., Le Mercier, M., Lefranc, F., and Kiss, R. Galectin-1: A small protein with major functions. *Glycobiology* 16, 137R–157R 2006.

29 Regulation of Immune Responses by Galectin-3

Daniel K. Hsu and Fu-Tong Liu

CONTENTS

29.1 INTRODUCTION

Galectins are an ancient family of animal lectins present in phylogenically distant organisms defined by selectivity for β-galactose and the presence of consensus amino-acid sequences [1]. Family members are observed in vertebrates, insects, nematodes, fungi and marine sponges, and homologs are present in plants [2] and viruses [3]. Fifteen members are described in mammals and others are present in expression databases [2]. The conserved galectin carbohydrate recognition domain (CRD) consists of approximately 130 amino acids that provide lectin activity. Family members contain a single CRD domain (galectin-1, -2, -5, -7, -10, -11, -13, -14, and -15) or two domains linked by a nonconserved peptide (galectin-4, -6, -8, -9, and -12). Galectin-3 is the only member of the third form with a single CRD domain and N-terminal region of approximately 120 amino acids composed of tandem repeats of short amino-acid segments.

Galectin-3 oligomerizes in the presence of multivalent carbohydrate ligands at approximately 1 μM [4–6]. A quantitative study revealed formation of galectin-3 pentamers on binding to multivalent carbohydrates [7]. Oligomerization of galectin-3 also occurs by covalent transglutaminase activity [8], resulting in the formation of stable multivalent lectin [9,10]. The affinity of galectin-3 for galactose is low, but that for certain galactose-containing oligosaccharides can be significantly higher [11].

Like other galectins, galectin-3 lacks a classical signal peptide and transmembrane domain required for secretion through the classical secretory pathway and for display on the cell surface, respectively (reviewed in Refs. [12,13]). Nevertheless, the protein is released to extracellular spaces and possible mechanisms have been investigated (reviewed in Ref. [12]).

Secreted galectin-3 may affect cells by an autocrine or paracrine mechanism by binding to and cross-linking glycoproteins and glycolipids present on the cell surfaces and extracellular matrices that contain appropriate oligosaccharides. Galectin-3's effectiveness in this process is likely to be associated with its ability to form oligomers. From earlier studies, it appeared that oligomerization of galectin-3 *in vitro* required the protein being present at micromolar concentrations. Recent findings with galectin-1 suggest these proteins may be more effective when presented by solid-phase (bound to extracellular matrix) repositories [14]. Another recent study demonstrates that oligomerization occurring on cell surfaces mediates cell activation and adherence [15], but local physiological concentrations required to effect these functions on cells *in vivo* are yet to be determined.

Although all galectins bind to galactose, they exhibit individual selectivity for glycans. For example, galectin-3 (as well as galectin-1), but not galectin-7, causes aggregation of neutrophils, thymocytes, and platelets [16]. Similarly, galectin-3 (as well as galectin-7), but not galectin-1, renders human erythrocytes sensitive to detergent-induced hemolysis of human erythrocytes.

Galectin-3 has been observed to shuttle between the cytoplasm and the nucleus [17] and has been described to play a role in pre-mRNA splicing [13,18]. A number of other intracellular functions have been reported and, for some of these, intracellular proteins with which galectin-3 interacts and possibly mediate these functions have been identified. Notably, galectin-3 binds to them through protein–protein interactions and not lectin–carbohydrate interactions [13,19]. A role for galectin-3 in glycoprotein trafficking was recently described by Delacour et al. [20]. They demonstrated the presence of galectin-3 in vesicles containing glycoproteins destined for membrane nonlipid raft microdomains at the apical side of the cell. Galectin-3 knockdown with small interfering RNA (siRNA) resulted in misdirection of apical proteins to basolateral cell surfaces. This study implicates galectin-3 with a role in controlling the composition of cell surface glycoproteins.

29.2 EXPRESSION OF GALECTIN-3 IN IMMUNE CELLS

Galectin-3 is widely distributed in tissues [21] and the expression profile in immune cells is summarized in Table 29.1.

29.2.1 Lymphoid Cells

Galectin-3 is absent in resting B cells but its expression is inducible by B cell receptor (BCR) cross-linkage or exposure to lipopolysaccharide (LPS), and by activation/differentiation stimuli with IL-4 or CD40 cross-linking [22]. In addition, galectin-3 was detected in B cells of mice infected with *Trypanosoma cruzi* and the lectin levels increased significantly when the cells were treated with IL-4 [22]. Similarly, galectin-3 is absent in resting mouse T cells, but expressed in CD4+ and CD8+ after activation with anti-CD3 antibody or the mitogen Concanavalin A (ConA) followed by culture with IL-2, IL-4, and IL-7 [23]. Galectin-3 has sometimes been detected in T cell subsets, as it appears to be present in subpopulations of mouse CD8+ T cells and this

TABLE 29.1
Expression of Galectin-3 in Immune Cells

Cell Type	Subset/Treatment	Animal Species
B cell	Induced by LPS, IL-4, or cross-linkage of BCR or CD40	Mouse
	Induced by parasite infection in the host	Mouse
	Induced by BCR ligation in anergic B cells	Mouse
T cell	CD4+ and CD8+ activated by anti-CD3 or Con A	Mouse
	Induced by HTLV-1 and HIV infection	Human
	CD8+ infection with *Yersinia*	Mouse
Langerhans cell	Cell surface and Birbeck granules (acquired from other cells?)	Human
Dendritic cell	Induced by parasite infection	Mouse
Myeloid	Increased upon maturation	Human
Mast cells		Human and mouse
Neutrophils		Human
Eosinophils		Human
Monocytes/ macrophages	Primary cells and cell lines/induced upon cell differentiation	Human and mouse

expression is upregulated in animals infected with *Yersinia* [24]. Viral induction of galectin-3 expression was observed in T cells infected with human T lymphotropic virus-1 (HTLV-1) [25] or human immunodeficiency virus (HIV) type 1 [26,27]. This is related to action on the galectin-3 promoter by HTLV-1 Tax transactivator protein [25] and HIV-1 Tat protein [26,27], respectively.

29.2.2 Dendritic Cells

Galectin-3 is present in mouse Langerhans cells [28] and on surfaces of human skin Langerhans cells [29]. As galectin-3 mRNA is undetectable in the latter, the lectin may be acquired from neighboring epidermal keratinocytes. Galectin-3 was also detected in Birbeck granules of Langerhans cells [30]. Expression of galectin-3 in splenic dendritic cells and a dendritic cell line was highly induced when mice were infected with *T. cruzi* infection [31].

29.2.3 Myeloid Cells

Expression of galectin-3 was observed in myelopoietic but not lymphopoietic cells of rat bone marrow [32]. In human bone marrow myeloid cells, galectin-3 expression is influenced by cell differentiation, being expressed at low levels in CD34+ early myeloid cells, but at higher levels in more mature cells [33]. Galectin-3 levels in human monocytes increase dramatically during differentiation into macrophages and in mouse peritoneal macrophages during *in vitro* culture [34,35]. Both galectin-3 protein and mRNA levels were induced by treatment of the human promyelocytic line HL-60 and the monocytic cell line THP-1 with phorbol 12-myristate 13-acetate [36,37]. Galectin-3 expression in resident mouse peritoneal macrophages increased when the cells were loaded with acetylated low-density lipoprotein (LDL) or oxidized LDL, and the PKC-Ras-Raf-MKK1-MKK3-p38 pathway appears to be involved in the regulation of lectin expression [38]. In addition, galectin-3 expression was upregulated in macrophages treated with 1,25-dihydroxyvitamin D3 [38,39], in macrophage and microglia activated during phagocytosis of myelin [40], and in microglia exposed to granulocyte macrophage colony stimulating factor [41]. Finally, galectin-3 is also present in mast cells [42], neutrophils [43], and eosinophils [44].

29.3 HOMEOSTATIC REGULATION OF IMMUNE CELLS

Galectin-3 has been shown to function in the regulation of cell differentiation, growth, and survival, and in modulation of cell adhesion, activation, and chemoattraction (Table 29.2). Several functions were demonstrated by exogenous delivery of galectin-3 to cells, while others involved functional inhibitors, gene suppression, or ectopic protein expression. The roles of galectin-3 in homeostatic regulation of immune cells are described in this section, and regulation of immune cell functions are covered in the next section.

29.3.1 Regulation of Cell Differentiation

Acosta-Rodriguez et al. [22] reported that galectin-3 expression favors differentiation of B cells into plasma cells over memory cells. They showed that galectin-3 contributed to IL-4-induced downregulation of Blimp-1, which is known to be able to drive the development of memory cells [22]. In another study comparing B cells from wild-type C57BL/6 mice exhibiting no autoimmune reactivity and autoimmune MRL/Mp mice, galectin-3 was found to be significantly elevated in anergic wild-type B cells after activation by BCR ligation, as opposed to low unchanged levels in activated self-reactive B cells [45]. The results suggest an association between galectin-3 expression and tolerance to self-antigens.

TABLE 29.2
Functions of Galectin-3 in Immune Cells Demonstrated *in Vitro*

Function	Comment
Induces B cell differentiation	Endogenous (antisense)
Mediates Con A/cytokine-induced T cell growth	Endogenous (antisense)
Promotes growth of T cell line	Endogenous (gene transfection)
Promotes G-CSF-driven growth of myeloid cells	Exogenous
Inhibits GM-CSF-driven growth of bone marrow cells	Exogenous
Inhibits apoptosis in B cell line	Endogenous (gene transfection)
Inhibits apoptosis in T cell line (intracellular)	Endogenous (gene transfection)
Induces apoptosis in T cell line (extracellular)	Exogenous
Inhibits apoptosis in macrophages	Endogenous (gene knockout)
Inhibits adhesion of thymocytes to thymic nurse cells	Exogenous
Promotes T cell adhesion to DC and macrophage	Endogenous
Induces IL-2 production by T cell line	Exogenous
Promotes IgE production in B cells	Exogenous
Regulates TCR signaling	Endogenous
Attenuates Th1 response to parasite antigen in DC	Endogenous (gene knockout)
Induces mast cell degranulation	Exogenous
Regulates JNK1 expression and cell function in mast cells	Endogenous (gene knockout)
Induces oxidative burst and IL-1 production in monocytes	Exogenous
Induces migration of monocytes/macrophages	Exogenous
Suppresses IL-5 production in eosinophils, PBMC, and T cell line	Exogenous
Potentiates macrophage phagocytosis	Endogenous (gene knockout)
Induces oxidative burst in neutrophils	Exogenous

Note: Exogenous: Demonstrated with recombinant galectin-3 added exogenously to cells. Endogenous: Demonstrated by inhibiting galectin-3 functions (using appropriate saccharide inhibitors or neutralizing antibodies), by inhibiting galectin-3 expression (using the antisense or siRNA strategy, or using cells from knockout mice) or by ectopically expressing galectin-3 (using the gene transfection approach).

29.3.2 Regulation of Cell Growth

The ability of galectin-3 to form multimers upon binding to glycans on the cell surface [46,15] suggests a means by which transmembrane signaling events can be initiated. It is possibly through this mechanism that galectin-3 influences the growth of various cell types (reviewed in Refs. [47,48]), including immune cells. While recombinant human galectin-3 promotes granulocyte-colony stimulating factor (G-CSF)-driven proliferation of CD34+ early myeloid cells *in vitro* [33], recombinant rat galectin-3 suppresses granulocyte macrophage colony stimulating factor (GM-CSF)-induced proliferation of rat bone marrow cells [32]. These findings suggest selective function of galectin-3 delivered extracellularly depending on the target cells.

By using galectin-3-specific antisense oligonucleotides, it was shown that decrease in galectin-3 expression in T cells resulted in attenuated cell growth induced by Con A and IL-2 [23]. It is unknown if galectin-3 is secreted by the T cells and influences cell growth by autocrine action, or functions intracellularly by affecting signaling pathways.

29.3.3 Regulation of Apoptosis (Induced by Exogenous Delivery)

Galectin-3 has been shown to induce apoptosis in T cells, which is also a demonstrated function of galectin-1 and -9. Galectin-3 is capable of inducing apoptosis in human T leukemic cell lines, human peripheral blood mononuclear cells (PBMC), and activated mouse T cells [49]. Interestingly, sensitivity of cells to apoptosis induced by galectin-3 is dependent on levels of the galectin-3 itself expressed in these T cell lines (e.g., galectin-3-negative Jurkat, CEM, and MOLT-4 were significantly more sensitive to apoptosis than galectin-3-positive SKW6.4 and H9 [49]). This is consistent with documented antiapoptotic activity of intracellular galectin-3 (described below). Neutralizing antibodies to CD7 and CD29 (β1 integrin) inhibited galectin-3-induced apoptosis [49], suggesting involvement of these antigens in galectin-3-mediated apoptosis. A recent study confirmed the capacity of galectin-3 to induce cell death in T cells and described differential sensitivities of cells to galectin-1 and galectin-3 [50]. Galectin-3 acts preferentially on double-negative thymocytes, while galectin-1 kills both double-negative and double-positive thymocytes and acts through CD45 and CD71 [50]. Galectin-3-induced apoptosis appears to involve activation of mitochondrial events, including cytochrome *c* release, and caspase-3 activation, but not caspase-8 activation [49]. As reported for galectin-1 [51–53], T cell killing by galectin-3 may be relevant to immune escape by tumor cells expressing the lectin. A potential role of galectin-3 in this regard was suggested by the observed correlation between galectin-3 expression in tumors and the degree of apoptosis in tumor-associated T cell [54].

29.3.4 Regulation of Apoptosis (Determined by Gene Transfection)

Galectin-3 was found to confer antiapoptotic activity in a number of cell types and diverse apoptotic stimuli by using gene transfection or antisense nucleotide methods in many cell types (reviewed in Ref. [13]), and some data exist for immune cells. Galectin-3 expression in human B lymphoma cells transfected with cDNA conferred resistance to apoptosis induced by anti-Fas antibody [55]. Conversely, a galectin-3-positive B cell line transfected to express the amino-terminal domain of galectin-3 (functioning in dominant negative fashion) exhibited elevated sensitivity to anti-Fas-induced apoptosis [55]. Similarly, galectin-3 expression in the human T cell line Jurkat resulted in resistance to apoptosis induced by anti-Fas antibody and staurosporine [56], and its expression in another human T cell line CEM resulted in resistance to apoptosis induced by galectin-1 [57] and C(2)-ceramide [49].

Galectin-3 also regulates apoptosis in other immune cells. Galectin-3-deficient macrophages undergo apoptosis more readily than wild-type mice when treated with a combination of LPS and IFNγ [58]. The results confirmed the antiapoptotic activity of galectin-3 demonstrated by gene transfection approaches as described above.

The mechanisms by which intracellular galectin-3 confers resistance to apoptosis appear to involve engagement of apoptosis regulation pathways. Translocation of galectin-3 from the cytosol or the nucleus to mitochondria following exposure to apoptotic stimuli was described [59,60]. Loss of mitochondrial membrane potential is a feature associated with apoptosis, and expression of galectin-3 was observed to suppress the loss in the potential, suggesting that antiapoptotic action of galectin-3 may involve interaction with other regulators of apoptosis operating in mitochondria [60]. The antiapoptotic activity of galectin-3 appears to rely on phosphorylation of serine at position 6 [61], which is also required for export of the lectin from the nucleus upon exposure to apoptotic stimuli [62]. The antiapoptotic activity is also dependent on the presence of Asn-Trp-Gly-Arg (NWGR) motif in the C-terminal domain of galectin-3 [56,63]. This motif is contained in well-characterized members of the Bcl-2 family of apoptosis regulators and is indispensable in their functions [64]. It is present in galectin-3 from various animal species [56] and is located at the carbohydrate-binding site of human galectin-3, as evidenced in the x-ray crystal structure [65].

Galectin-3 has been found to bind to proteins implicated in regulation of apoptosis, including Bcl-2 [56] and synexin [59]. Synexin was also shown to be required for translocation of galectin-3 to perinuclear membranes [59]. Current knowledge indicates that mechanisms by which galectin-3 inhibit apoptosis are complex and rely on compartmentalization of galectin-3. In particular, galectin-3 is antiapoptotic when present in the cytosol, but proapoptotic when present in the nucleus [66].

29.4 FUNCTIONAL REGULATION OF IMMUNE CELLS

29.4.1 Regulation of Cell Adhesion

In a carbohydrate-dependent manner, galectin-3 binds to the extracellular matrix proteins laminin [5,67–69], fibronectin [69], elastin [70], and hensin [71] and can influence cell adhesion to extracellular matrices. It also interacts with integrins $\alpha 1\beta 1$ [72,73] and $\alpha M\beta 1$ (CD11b/18) [74] and potentiates cell adhesion through these integrins. As immune and inflammatory responses notably rely on cell adhesion, it is conceivable that galectin-3 can affect immune and inflammatory processes by modulating cell adhesion in various cell types (reviewed in Ref. [67]).

29.4.2 Lymphoid Cells

In an *in vitro* system representing reconstituted thymic microenvironment, Villa-Verde et al. [75] found galectin-3 to attenuate interaction of thymocytes with thymic nurse cells. The authors proposed that galectin-3 may influence development of thymocytes by modulating their adhesion to surrounding cells in thymus in a capacity as a deadhesion molecule. This function appears to operate via the lectin activity as it is inhibited by lactose, a natural galectin-3 ligand.

The function described above was demonstrated by delivering galectin-3 exogenously. Involvement of intrinsic galectin-3 in adhesion of mouse T cells was described by Swarte et al. [76] in cells triggered by antibody ligation of L-selectin that results in adhesion of T cells to dendritic cells or macrophages. This intercellular adhesion was inhibited by a known galectin-3 sugar ligand or antigalectin-3 antibody [76].

29.4.3 Myeloid Cells

Recombinant galectin-3 was found to promote adhesion of human neutrophils to laminin [68] and to an endothelial cell line [77]. Galectin-3 was found to bind to alpha-galactosylated glycan xenoantigens [78]. Additional results suggest that galectin-3 serves as a receptor on human monocytes for these xenoantigens and is responsible for infiltration of these cells to the xenograft.

Influence of galectin-3 on cell adhesion appears to rely on the cell type and the nature of glycoconjugates on immune cells, as well as relevance of these surface glycans to cell adhesion [67]. Interestingly, galectin-3 was observed to regulate the expression of integrins [79] and induce endocytosis of β1 integrins in certain cells [80]. This suggests that another mechanism by which this protein regulates cell adhesion may occur by controlling cell surface levels of adhesion molecules.

29.4.4 Cell Activation

As receptor cross-linking is a method by which transmembrane signal transduction initiates, galectin-3 may be expected to induce cell activation through its multivalent lectin activity.

29.4.4.1 Lymphoid Cells

Galectin-3 can mediate IgE production by B cells induced by polymorphonuclear cells from patients with IgE-associated atopic eczema/dermatitis syndrome [81], and activate human Jurkat T cells to produce IL-2 [25,82]. By these actions, it functions as a positive regulator. Galectin-3 was described to associate with the T-cell receptor (TCR) complex, in a fashion that is dependent on the glycosylation of TCR controlled by β1,6-*N*-acetylglucosaminyltransferase V (Mgat5), an enzyme in the N-glycosylation pathway [83]. Inhibition of galectin-3 on T cell surfaces by lactose pretreatment increased TCR clustering as well as signaling through the receptor. This suggests that galectin-3 can serve as a negative regulator of TCR-initiated signal transduction [83].

29.4.4.2 Myeloid Cells

Galectin-3 was previously shown to bind to IgE (hence the name IgE-binding protein) but also binds to IgE receptor (FcεRI) [84]. Galectin-3 induces mediator release in both IgE-sensitized and nonsensitized cells [84,85], possibly by cross-linking FcεRI-bound IgE, FcεRI, or both. Galectin-3 induces superoxide anion production [34] and potentiates LPS-induced IL-1 production [86] in human peripheral blood monocytes. It also induces superoxide anion production in human peripheral blood neutrophils primed with cytochalasin-D [87]. This was similarly demonstrated with exudated human neutrophils obtained from skin blister fluid and with peripheral blood neutrophils treated with stimuli such as fMLP [88]. Additional study revealed that neutrophils responded to galectin-3 after prior exposure to LPS, and it was shown that responsiveness was due to mobilization of granular proteins recognizable by galectin-3 to cell surfaces [89]. Interaction of galectin-3 with CD66a and CD66b on neutrophil surfaces has also been shown [90]. Comparative analyses of neutrophil activation by galectin-1 and -3 revealed differential activation by engagement of different receptor sets [91].

Galectin-3 can also exert a suppressive effect on myeloid cells, as exemplified by inhibition of IL-5 production in human eosinophils, an eosinophilic cell line, human PBMC, and an antigen-specific T cell line, due to downregulation of IL-5 mRNA levels [92]. Galectin-3 appears to downregulate IL-5 production by engaging FcγRII, as revealed by studying PBMC from FcγRII-deficient mice [93].

A function in phagocytosis has been demonstrated for galectin-3 by comparing macrophages from galectin-3-deficient (gal3$^{-/-}$) mice and wild-type mice [94]. Gal3$^{-/-}$ macrophages are defective in phagocytosis of opsonized erythrocytes and apoptotic thymocytes both *in vitro* and *in vivo*. In addition, in wild-type macrophages engulfing opsonized erythrocytes, galectin-3 was detected in phagocytic cups and concentrated in phagosomes. The mechanism by which intracellular galectin-3 participates in phagocytosis is yet to be clarified.

Studies of mast cell in gal3$^{-/-}$ mice have revealed a role for galectin-3 in regulation of mast cell functions. Gal3$^{-/-}$ mast cells produced lower levels of granular mediators and cytokines, when activated by cross-linkage of cell surface IgE receptor, compared to wild-type cells [95]. Reduced

IgE-mediated responses of mast cells *in vivo* were also observed in gal3$^{-/-}$ mice, as evidenced by diminished passive cutaneous anaphylactic reactions in these mice. The results indicate that galectin-3 is intimately involved in mast cell functions. Gal3$^{-/-}$ mast cells contained lower levels of c-jun-N-terminal kinase-1 (JNK1) mRNA and protein, suggesting that the lectin regulates JNK1 transcription [95]. Because JNK1 plays a central role in the signaling pathways leading to production of selected cytokines, this could explain how galectin-3 controls cytokine production.

29.4.5 Chemoattraction

Development and maturation of immune and inflammatory responses rely on migration of various immune cells induced by chemoattractants, such as chemokines. In this regard, galectin-3 was found to induce migration of human monocytes and macrophages, and alveolar macrophages, both *in vitro* and *in vivo* [96]. N-terminal and C-terminal domains are required for this activity, and the activity is at least partially mediated through a pertussis toxin (PTX)-sensitive (G protein-coupled) pathway also engaged by various chemokine receptors [96]. The identity of specific chemokine receptors in the chemotactic activity of galectin-3 remains to be established.

29.5 BIOLOGICAL EFFECTS DEMONSTRATED *IN VIVO*

In vitro studies suggest that galectin-3 can promote inflammatory responses through its functions on cell activation, cell migration, and inhibition of apoptosis (thus prolonging the survival of inflammatory cells). This has been confirmed by *in vivo* studies using gal3$^{-/-}$ mice, as demonstrated by attenuated infiltration relative to wild-type mice in a peritoneal model of inflammation by Colnot et al. [97]. We also observed attenuated inflammation in gal3$^{-/-}$ mice, as evident in consistently lower macrophage infiltrations in the peritoneal cavity, compared to wild-type mice, when mice were injected intraperitoneally with thioglycollate broth [58].

Our laboratory also described a proinflammatory role for galectin-3 in an antigen-specific mouse model of asthma. Ovalbumin-sensitized gal3$^{-/-}$ mice developed lower lung eosinophilia and significantly reduced airway hyperresponsiveness relative to similarly treated wild-type mice following airway antigen challenge [98]. Other investigators observed reduced eosinophil infiltration and airway hyperresponsiveness in response to airway antigen challenge in rats and mice after intranasal delivery of cDNA encoding galectin-3 [99,100]. These contrasting results may be explained by a potentiating role for endogenous galectin-3 in the airway inflammatory response, but a suppressive effect of pharmacological concentrations of galectin-3 applied to the airways. The study of airway inflammation using gal3$^{-/-}$ mice also revealed the regulation of Th1/Th2 polarization by galectin-3: Compared to wild type mice, gal3$^{-/-}$ mice showed (i) decreased IL-4 and IgE levels in bronchoalveolar fluid and serum and (ii) elevated IFNγ levels and IgG2a/IgG1 ratios in bronchoalveolar fluid and serum [98].

The role of galectin-3 in innate immunity against microorganisms has also been revealed by studying gal3$^{-/-}$ mice. Consistent with the role of galectin-3 in promoting the inflammatory response, gal3$^{-/-}$ mice infected with *Toxoplasma gondii* exhibited lower inflammatory scores in gut, liver, and brain compared to similarly infected wild-type mice [101]. However, cellular inflammation in lungs of gal3$^{-/-}$ mice was more pronounced than wild-type mice in late infection, suggesting complex cellular regulation by galectin-3. A polarized T helper response was also observed in gal3$^{-/-}$ mice infected with *T. gondii* relative to wild type in that serum levels of IL-12 and IFNγ were highly elevated in the former, and appear to result from increased secretion of IL-12 by dendritic cells [101]. Finally, in *T. gondii* infection by the oral route, gal3$^{-/-}$ mice displayed survival periods that are similar to wild-type mice, but infections by the peritoneal route revealed increased susceptibility of gal3$^{-/-}$ mice to parasite.

In a mouse infection model with *Mycobaterium tuberculosis*, Beatty et al. [102] observed that gal3$^{-/-}$ mice demonstrated impaired capacities in clearing late infections compared to wild-type

mice, suggesting involvement of galectin-3 in innate defense against *Mycobacterial* infection. In a mouse infection model with *Streptococcal pneumonia,* Sato et al. [77] observed the accumulation of galectin-3 in alveolar spaces of infected lungs, correlating with the onset of neutrophil extravasation. These results suggest involvement of galectin-3 in neutrophil adhesion during processes of extravasation in *S. pneumonia* lung infection.

29.6 CONCLUSION AND FUTURE DIRECTIONS

The functions demonstrated for galectin-3 can be described as either extracellular or intracellular. Extracellularly, galectin-3 can function by binding to cell surface glycans in a fashion that is dependent on lectin–carbohydrate interactions. Through this mode of action, it can affect cell growth and survival, activate, or inhibit cellular responses, modulate cell adhesion, and induce cell migration. These are mostly demonstrated with recombinant galectin-3 *in vitro*. Whether endogenous galectin-3 exerts these activities remains to be clarified. First, there is a question of whether the lectin is present extracellularly. In this regard, galectin-3 is secreted by cultured cells (reviewed in Ref. [12]) and is readily detectable in the extracellular environment and present under inflammatory conditions. For example, galectin-3 is detectable in tears from patients with ocular diseases [103]. In patients with rheumatoid arthritis, galectin-3 expression is upregulated in synovial tissues and present in synovial fluid from these patients [104]. Bronchoalveolar lavage fluid from mice with inflamed airways [98] and mice infected with *S. pneumonia* [77] also contain galectin-3.

The second question is whether the concentrations of extracellular galectin-3 would be sufficiently high for those responses demonstrated *in vitro* to occur *in vivo*. The extracellular functions described were mostly demonstrated by addition of relatively high concentrations of recombinant galectin-3. However, galectin-3 may be more effective in induction of extracellular effects when presented by the extracellular matrices, as described for galectin-1 [14]. In addition, stable dimers of galectin-3 may be formed through covalent transglutamination [8–10], which would render the protein more effective in cross-linking cell surface glycoproteins. Nevertheless, *in vivo* data are needed to definitively establish the role of extracellular functions of galectin-3 in physiological and pathological processes.

The cell surface receptors responsible for galectin-3's action also remain to be clarified. Galectin-3 has been shown to bind to a large number of different glycoproteins. It appears the lectin invariably binds to a number of different glycoproteins on the cell surfaces and thus does not have a specific receptor. It also appears that the lectin binds to different cell surface glycoproteins in different cell types. It would continue to be a challenge to establish which glycoproteins recognized by the lectin is responsible for its function exerted on a specific cell type, and the conditions under which galectin-3 can effect the function.

The intracellular functions identified for galectin-3 are not expected for a protein with lectin properties, but are consistent with the protein's intracellular localization. These involve some fundamental intracellular process, such as pre-mRNA splicing and apoptosis. Many intracellular functions of galectin-3 were revealed from studies involving gene transfection and antisense nucleotides to influence expression, and others from correlation with intracellular localization of the proteins. More recently, some of the intracellular functions have been confirmed by the use of cells from mice deficient in galectin-3. The possibility exists that some activities detected *in vivo* are due to the protein being released from cells and functioning extracellularly. In a number of cases, however, this is excluded by the fact that the activities in question are not affected by lactose added to the culture medium, which would inhibit galectin-3's carbohydrate-dependent extracellular action. In addition, the intracellular functions are supported by the identification of intracellular proteins with which galectin-3 interact.

It is to be noted that most of the described intracellular functions of galectin-3 are not dependent on its interactions with intracellular carbohydrates. However, as described earlier in this review, galectin-3 has been identified in post-Golgi intracellular vehicles transporting glycoproteins destined

for apical export [20]. The authors concluded that galectin-3 may be involved in sorting selected glycoproteins into these specific vehicles. This suggests an intriguing possibility in intracellular utilization of galectin-3's lectin function.

A large body of *in vitro* studies using a variety of different methodologies have demonstrated a number of functions of galectin-3 in various immune cells. *In vivo* studies, including those using gal3$^{-/-}$ mice also have begun to provide information on the functions of galectin-3 in adaptive and innate immunity. It may be anticipated that additional studies of these mice, as well as investigations employing small interference RNA (siRNA) and specific inhibitors will continue to shed light on the role of galectin-3 in regulation of the immune response.

ACKNOWLEDGMENTS

The authors thank members of their research group who contributed to the work cited in this chapter. The work has been supported by grants from the National Institutes of Health (RO1AI20958, RO1AI39620, and PO1 AI50498).

REFERENCES

1. Barondes SH, Castronovo V, Cooper DNW, Cummings RD, Drickamer K, Feizi T, Gitt MA, Hirabayashi J, Hughes C, Kasai K, Leffler H, Liu FT, Lotan R, Mercurio AM, Monsigny M, Pillai S, Poirer F, Raz A, Rigby PWJ, Rini JM, and Wang JL: Galectins: A family of animal b-galactoside-binding lectins [Letter to the Editor], *Cell* 1994, 76:597–598.
2. Cooper DN: Galectinomics: Finding themes in complexity, *Biochim Biophys Acta* 2002, 1572:209–231.
3. Kleiboeker SB: Sequence analysis of the fiber genomic region of a porcine adenovirus predicts a novel fiber protein, *Virus Res* 1995, 39:299–309.
4. Hsu DK, Zuberi R, and Liu FT: Biochemical and biophysical characterization of human recombinant IgE-binding protein, an S-type animal lectin, *J Biol Chem* 1992, 267:14167–14174.
5. Massa SM, Cooper DNW, Leffler H, and Barondes SH: L-29, an endogenous lectin, binds to glycoconjugate ligands with positive cooperativity, *Biochemistry* 1993, 32:260–267.
6. Ochieng J, Platt D, Tait L, Hogan V, Raz T, Carmi P, and Raz A: Structure–function relationship of a recombinant human galactoside-binding protein, *Biochemistry* 1993, 32:4455–4460.
7. Ahmad N, Gabius HJ, Andre S, Kaltner H, Sabesan S, Roy R, Liu B, Macaluso F, and Brewer CF: Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes, *J Biol Chem* 2004, 279:10841–10847.
8. Mehul B, Bawumia S, and Hughes RC: Cross-linking of galectin 3, a galactose-binding protein of mammalian cells, by tissue-type transglutaminase, *FEBS Lett* 1995, 360:160–164.
9. Mahoney S, Wilkinson M, Smith S, and Haynes LW: Stabilization of neurites in cerebellar granule cells by transglutaminase activity: Identification of midkine and galectin-3 as substrates, *Neurosci* 2000, 101:141–155.
10. van den Brûle FA, Liu FT, and Castronovo V: Transglutaminase-mediated oligomerization of galectin-3 modulates human melanoma cell interactions with laminin, *Cell Adhes Commun* 1998, 5:425–435.
11. Rini JM and Lobsanov YD: New animal lectin structures, *Curr Opin Struct Biol* 1999, 9:578–584.
12. Hughes RC: Secretion of the galectin family of mammalian carbohydrate-binding proteins, *Biochim Biophys Acta* 1999, 1473:172–185.
13. Liu FT, Patterson RJ, and Wang JL: Intracellular functions of galectins, *Biochim Biophys Acta* 2002, 1572:263–273.
14. He J and Baum LG: Presentation of galectin-1 by extracellular matrix triggers T cell death, *J Biol Chem* 2004, 279:4705–4712.
15. Nieminen J, Kuno A, Hirabayashi J, and Sato S: Visualization of galectin-3 oligomerization on the surface of neutrophils and endothelial cells using fluorescence resonance energy transfer, *J Biol Chem* 2007, 282: 1374–1383.
16. Timoshenko AV, Gorudko IV, Maslakova OV, Andre S, Kuwabara I, Liu FT, Kaltner H, and Gabius HJ: Analysis of selected blood and immune cell responses to carbohydrate-dependent surface binding of proto- and chimera-type galectins, *Mol Cell Biochem* 2003, 250:139–149.
17. Davidson PJ, Davis MJ, Patterson RJ, Ripoche MA, Poirier F, and Wang JL: Shuttling of galectin-3 between the nucleus and cytoplasm, *Glycobiology* 2002, 12:329–337.

18. Dagher SF, Wang JL, and Patterson RJ: Identification of galectin-3 as a factor in pre-mRNA splicing, *Proc Natl Acad Sci USA* 1995, 92:1213–1217.
19. Wang JL, Gray RM, Haudek KC, and Patterson RJ: Nucleocytoplasmic lectins, *Biochim Biophys Acta* 2004, 1673:75–93.
20. Delacour D, Cramm-Behrens CI, Drobecq H, Le Bivic A, Naim HY, and Jacob R: Requirement for galectin-3 in apical protein sorting, *Curr Biol* 2006, 16:408–414.
21. Chiariotti L, Salvatore P, Frunzio R, and Bruni CB: Galectin genes: Regulation of expression, *Glycoconj J* 2002, 19:441–449.
22. Acosta-Rodriguez EV, Montes CL, Motran CC, Zuniga EI, Liu FT, Rabinovich GA, and Gruppi A: Galectin-3 mediates IL-4-induced survival and differentiation of B cells: Functional cross-talk and implications during *Trypanosoma cruzi* infection, *J Immunol* 2004, 172:493–502.
23. Joo HG, Goedegebuure PS, Sadanaga N, Nagoshi M, von Bernstorff W, and Eberlein TJ: Expression and function of galectin-3, a beta-galactoside-binding protein in activated T lymphocytes, *J Leukoc Biol* 2001, 69:555–564.
24. Fahrer AM, Konigshofer Y, Kerr EM, Ghandour G, Mack DH, Davis MM, and Chien YH: Attributes of gammadelta intraepithelial lymphocytes as suggested by their transcriptional profile, *Proc Natl Acad Sci USA* 2001, 98:10261–10266.
25. Hsu DK, Hammes SR, Kuwabara I, Greene WC, and Liu FT: Human T lymphotropic virus-1 infection of human T lymphocytes induces expression of the b-galactose-binding lectin, galectin-3, *Am J Pathol* 1996, 148:1661–1670.
26. Schröder HC, Ushijima H, Theis C, Sève AP, Hubert J, and Müller WEG: Expression of nuclear lectin carbohydrate-binding protein 35 in human immunodeficiency virus type 1-infected Molt-3 cells, *J Acquir Immune Defic Syndr Hum Retrovirol* 1995, 9:340–348.
27. Fogel S, Guittaut M, Legrand A, Monsigny M, and Hébert E: The Tat protein of HIV-1 induces galectin-3 expression, *Glycobiology* 1999, 9:383–387.
28. Haines KA, Flotte TJ, Springer TA, Gigli I, and Thorbecke GJ: Staining of Langerhans cells with monoclonal antibodies to macrophages and lymphoid cells, *Proc Natl Acad Sci USA* 1983, 80:3448–3451.
29. Wollenberg A, De la Salle H, Hanau D, Liu FT, and Bieber T: Human keratinocytes release the endogenous b-galactoside-binding soluble lectin eBP (IgE-binding protein) which binds to Langerhans cells where it modulates their binding capacity for IgE glycoform, *J Exp Med* 1993, 178:777–785.
30. Smetana K, Holikova Z, Klubal R, Bovin NV, Dvoankova B, Bartkova J, Liu FT, and Gabius HJ: Coexpression of binding sites for A(B) histo-blood group trisaccharides with galectin-3 and lag antigen in human Langerhans cells, *J Leukocyte Biol* 1999, 6:644–649.
31. Vray B, Camby I, Vercruysse V, Mijatovic T, Bovin NV, Ricciardi-Castagnoli P, Kaltner H, Salmon I, Gabius HJ, and Kiss R: Up-regulation of galectin-3 and its ligands by *Trypanosoma cruzi* infection with modulation of adhesion and migration of murine dendritic cells, *Glycobiology* 2004, 14:647–657.
32. Krugluger W, Frigeri LG, Lucas T, Schmer M, Forster O, Liu FT, and Boltz-Nitulescu G: Galectin-3 inhibits granulocyte-macrophage colony-stimulating factor (GM-CSF)-driven rat bone marrow cell proliferation and GM-CSF-induced gene transcription, *Immunobiology* 1997, 197:97–109.
33. Le Marer N: Galectin-3 expression in differentiating human myeloid cells, *Cell Biol Int* 2000, 24:245–251.
34. Liu FT, Hsu DK, Zuberi RI, Kuwabara I, Chi EY, and Henderson WR, Jr.: Expression and function of galectin-3, a b-galactoside-binding lectin, in human monocytes and macrophages, *Am J Pathol* 1995, 147:1016–1029.
35. Dumic J, Lauc G, Hadzija M, and Flögel M: Transfer to in vitro conditions influences expression and intracellular distribution of galectin-3 in murine peritoneal macrophages, *Z Naturforsch (C)* 2000, 55:261–266.
36. Cherayil BJ, Chaitovitz S, Wong C, and Pillai S: Molecular cloning of a human macrophage lectin specific for galactose, *Proc Natl Acad Sci USA* 1990, 87:7324–7329.
37. Nangia-Makker P, Ochieng J, Christman JK, and Raz A: Regulation of the expression of galactoside-binding lectin during human monocytic differentiation, *Cancer Res* 1993, 53:1–5.
38. Kim K, Mayer EP, and Nachtigal M: Galectin-3 expression in macrophages is signaled by Ras/MAP kinase pathway and up-regulated by modified lipoproteins, *Biochim Biophys Acta* 2003, 1641:13–23.
39. Abu-Amer Y and Bar-Shavit Z: Modulation of vitamin D increased H2O2 production and MAC-2 expression in the bone marrow-derived macrophages by estrogen, *Calcif Tissue Int* 1994, 55:29–32.
40. Reichert F and Rotshenker S: Galectin-3/MAC-2 in experimental allergic encephalomyelitis, *Exp Neurol* 1999, 160:508–514.

41. Pesheva P, Urschel S, Frei K, and Probstmeier R: Murine microglial cells express functionally active galectin-3 in vitro, *J Neurosci Res* 1998, 51:49–57.
42. Frigeri LG and Liu FT: Surface expresson of functional IgE binding protein, an endogenous lectin, on mast cells and macrophages, *J Immunol* 1992, 148:861–869.
43. Truong MT, Gruart V, Kusnierz JP, Papin JP, Loiseau S, Capron A, and Capron M: Human neutrophils express immunoglobulin E (IgE)-binding proteins (Mac-2/eBP) of the S-type lectin family: Role in IgE-dependent activation, *J Exp Med* 1993, 177:243–248.
44. Truong MJ, Gruart V, Liu FT, Prin L, Capron A, and Capron M: IgE-binding molecules (Mac-2/eBP) expressed by human eosinophils Implication in IgE-dependent eosinophil cytotoxicity, *Eur J Immunol* 1993, 23:3230–3235.
45. Clark AG, Chen S, Zhang H, Brady GF, Ungewitter EK, Bradley JK, Sackey FN, and Foster MH: Multifunctional regulators of cell growth are differentially expressed in anergic murine B cells, *Mol Immunol* 2007, 44:1274–1285.
46. Brewer CF, Miceli MC, and Baum LG: Clusters, bundles, arrays and lattices: Novel mechanisms for lectin-saccharide-mediated cellular interactions, *Curr Opin Struct Biol* 2002, 12:616–623.
47. Hsu DK and Liu FT: Regulation of cellular homeostasis by galectins, *Glycoconj J* 2004, 19:507–515.
48. Yang RY and Liu FT: Galectins in cell growth and apoptosis, *Cell Mol Life Sci* 2003, 60:267–276.
49. Fukumori T, Takenaka Y, Yoshii T, Kim HR, Hogan V, Inohara H, Kagawa S, and Raz A: CD29 and CD7 mediate galectin-3-induced type II T-cell apoptosis, *Cancer Res* 2003, 63:8302–8311.
50. Stillman BN, Hsu DK, Pang M, Brewer CF, Johnson P, Liu FT, and Baum LG: Galectin-3 and galectin-1 bind distinct cell surface glycoprotein receptors to induce T cell death, *J Immunol* 2006, 176:778–789.
51. Perillo NL, Pace KE, Seilhamer JJ, and Baum LG: Apoptosis of T cells mediated by galectin-1, *Nature* 1995, 378:736–739.
52. Perillo NL, Uittenbogaart CH, Nguyen JT, and Baum LG: Galectin-1, an endogenous lectin produced by thymic epithelial cells, induces apoptosis of human thymocytes, *J Exp Med* 1997, 185:1851–1858.
53. Rabinovich G, Castagna L, Landa C, Riera CM, and Sotomayor C: Regulated expression of a 16-kD galectin-like protein in activated rat macrophages, *J Leukocyte Biol* 1996, 59:363–370.
54. Zubieta MR, Furman D, Barrio M, Bravo AI, Domenichini E, and Mordoh J: Galectin-3 expression correlates with apoptosis of tumor-associated lymphocytes in human melanoma biopsies, *Am J Pathol* 2006, 168:1666–1675.
55. Hoyer KK, Pang M, Gui D, Shintaku IP, Kuwabara I, Liu FT, Said JW, Baum LG, and Teitell MA: An anti-apoptotic role for galectin-3 in diffuse large B-cell lymphomas, *Am J Pathol* 2004, 164:893–902.
56. Yang RY, Hsu DK, and Liu FT: Expression of galectin-3 modulates T cell growth and apoptosis, *Proc Natl Acad Sci USA* 1996, 93:6737–6742.
57. Hahn HP, Pang M, He J, Hernandez JD, Yang RY, Li LY, Wang X, Liu FT, and Baum LG: Galectin-1 induces nuclear translocation of endonuclease G in caspase- and cytochrome c-independent T cell death, *Cell Death Differ* 2004, 11:1277–1286.
58. Hsu DK, Yang RY, Yu L, Pan Z, Salomon DR, Fung-Leung WP, and Liu FT: Targeted disruption of the galectin-3 gene results in attenuated peritoneal inflammatory responses, *Amer J Pathol* 2000, 156:1073–1083.
59. Yu F, Finley RL, Jr., Raz A, and Kim HR: Galectin-3 translocates to the perinuclear membranes and inhibits cytochrome c release from the mitochondria. A role for synexin in galectin-3 translocation, *J Biol Chem* 2002, 277:15819–15827.
60. Matarrese P, Tinari N, Semeraro ML, Natoli C, Iacobelli S, and Malorni W: Galectin-3 overexpression protects from cell damage and death by influencing mitochondrial homeostasis, *FEBS Lett* 2000, 473:311–315.
61. Yoshii T, Fukumori T, Honjo Y, Inohara H, Kim HR, and Raz A: Galectin-3 phosphorylation is required for its anti-apoptotic function and cell cycle arrest, *J Biol Chem* 2002, 277:6852–6857.
62. Takenaka Y, Fukumori T, Yoshii T, Oka N, Inohara H, Kim HR, Bresalier RS, and Raz A: Nuclear export of phosphorylated galectin-3 regulates its antiapoptotic activity in response to chemotherapeutic drugs, *Mol Cell Biol* 2004, 24:4395–4406.
63. Akahani S, Nangia-Makker P, Inohara H, Kim HRC, and Raz A: Galectin-3: A novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family, *Cancer Res* 1997, 57:5272–5276.
64. Muchmore SW, Sattler M, Liang H, Meadows RP, Harlan JE, Yoon HS, Nettesheim D, Chang BS, Thompson CB, Wong SL, Ng SL, and Fesik SW: X-ray and NMR structure of human Bcl-xL, an inhibitor of programmed cell death, *Nature* 1996, 381:335–341.

65. Seetharaman J, Kanigsberg A, Slaaby R, Leffler H, Barondes SH, and Rini JM: X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1 Å resolution, *J Biol Chem* 1998, 273:13047–13052.
66. Califice S, Castronovo V, Bracke M, and Van Den Brule F: Dual activities of galectin-3 in human prostate cancer: Tumor suppression of nuclear galectin-3 vs tumor promotion of cytoplasmic galectin-3, *Oncogene* 2004, 23:7527–7536.
67. Hughes RC: Galectins as modulators of cell adhesion, *Biochimie* 2001, 83:667–676.
68. Kuwabara I and Liu FT: Galectin-3 promotes adhesion of human neutrophils to laminin, *J Immunol* 1996, 156:3939–3944.
69. Sato S and Hughes RC: Binding specificity of a baby hamster kidney lectin for H type I and II chains, polylactosamine glycans, and appropriately glycosylated forms of laminin and fibronectin, *J Biol Chem* 1992, 267:6983–6990.
70. Ochieng J, Warfield P, Green-Jarvis B, and Fentie I: Galectin-3 regulates the adhesive interaction between breast carcinoma cells and elastin, *J Cell Biochem* 1999, 75:505–514.
71. Hikita C, Vijayakumar S, Takito J, Erdjument-Bromage H, Tempst P, and Al-Awqati Q: Induction of terminal differentiation in epithelial cells requires polymerization of hensin by galectin 3, *J Cell Biol* 2001, 151:1235–1146.
72. Ochieng J, Leite-Browning ML, and Warfield P: Regulation of cellular adhesion to extracellular matrix proteins by galectin-3, *Biochem Biophys Res Commun* 1998, 246:788–791.
73. Andre S, Kojima S, Yamazaki N, Fink C, Kaltner H, Kayser K, and Gabius HJ: Galectins-1 and -3 and their ligands in tumor biology Non-uniform properties in cell surface presentation and modulation of adhesion to matrix glycoproteins for various tumor cell lines, in biodistribution of free and liposome-bound galectins and in their expression by breast and colorectal carcinomas with/without metastatic propensity, *J Cancer Res Clin Oncol* 1999, 125:461–474.
74. Dong S and Hughes RC: Macrophage surface glycoproteins binding to galectin-3 (Mac-2-antigen), *Glycoconj J* 1997, 14:267–274.
75. Villa-Verde DM, Silva-Monteiro E, Jasiulionis MG, Farias-De-Oliveira DA, Brentani RR, Savino W, and Chammas R: Galectin-3 modulates carbohydrate-dependent thymocyte interactions with the thymic microenvironment, *Eur J Immunol* 2002, 32:1434–1444.
76. Swarte VV, Mebius RE, Joziasse DH, Van den Eijnden DH, and Kraal G: Lymphocyte triggering via L-selectin leads to enhanced galectin-3-mediated binding to dendritic cells, *Eur J Immunol* 1998, 28:2864–2871.
77. Sato S, Ouellet N, Pelletier I, Simard M, Rancourt A, and Bergeron MG: Role of galectin-3 as an adhesion molecule for neutrophil extravasation during streptococcal pneumonia, *J Immunol* 2002, 168:1813–1822.
78. Jin R, Greenwald A, Peterson MD, and Waddell TK: Human monocytes recognize porcine endothelium via the interaction of galectin 3 and alpha-GAL, *J Immunol* 2006, 177:1289–1295.
79. Mataresse P, Fusco O, Tinari N, Natoli C, Liu FT, Semeraro ML, Malorni W, and Iacobelli S: Galectin-3 overexpression protects from apoptosis by improving cell adhesion properties, *Int J Cancer* 2000, 85:545–554.
80. Furtak V, Hatcher F, and Ochieng J: Galectin-3 mediates the endocytosis of beta-1 integrins by breast carcinoma cells, *Biochem Biophys Res Commun* 2001, 289:845–850.
81. Kimata H: Enhancement of IgE production in B cells by neutrophils via galectin-3 in IgE-associated atopic eczema/dermatitis syndrome, *Int Arch Allergy Immunol* 2002, 128:168–170.
82. Dong S and Hughes RC: Galectin-3 stimulates uptake of extracellular Ca^{2+} in human Jurkat T-cells, *FEBS Lett* 1996, 395:165–169.
83. Demetriou M, Granovsky M, Quaggin S, and Dennis JW: Negative regulation of T-cell activation and autoimmunity by Mgat5 N-glycosylation, *Nature* 2001, 409:733–779.
84. Frigeri LG, Zuberi RI, and Liu FT: eBP, a b-galactoside-binding animal lectin, recognizes IgE receptor (FceRI) and activates mast cells, *Biochemistry* 1993, 32:7644–7649.
85. Zuberi RI, Frigeri LG, and Liu FT: Activation of rat basophilic leukemia cells by eBP, an IgE-binding endogenous lectin, *Cell Immunol* 1994, 156:1–12.
86. Jeng KCG, Frigeri LG, and Liu FT: An endogenous lectin, galectin-3 (eBP/Mac-2), potentiates IL-1 production by human monocytes, *Immunol Lett* 1994, 42:113–116.
87. Yamaoka A, Kuwabara I, Frigeri LG, and Liu FT: A human lectin, galectin-3 (eBP/Mac-2), stimulates superoxide production by neutrophils, *J Immunol* 1995, 154:3479–3487.
88. Karlsson A, Follin P, Leffler H, and Dahlgren C: Galectin-3 activates the NADPH-oxidase in exudated but not peripheral blood neutrophils, *Blood* 1998, 91:3430–3438.

89. Almkvist J, Faldt J, Dahlgren C, Leffler H, and Karlsson A: Lipopolysaccharide-induced gelatinase granule mobilization primes neutrophils for activation by galectin-3 and formylmethionyl-Leu-Phe, *Infect Immun* 2001, 69:832–837.
90. Feuk-Lagerstedt E, Jordan ET, Leffler H, Dahlgren C, and Karlsson A: Identification of CD66a and CD66b as the major galectin-3 receptor candidates in human neutrophils, *J Immunol* 1999, 163:5592–5598.
91. Almkvist J, Dahlgren C, Leffler H, and Karlsson A: Activation of the neutrophil nicotinamide adenine dinucleotide phosphate oxidase by galectin-1, *J Immunol* 2002, 168:4034–4041.
92. Cortegano I, delPozo V, Cardaba B, deAndres B, Gallardo S, delAmo A, Arrieta I, Jurado A, Palomino P, Liu FT, and Lahoz C: Galectin-3 down-regulates IL-5 gene expression on different cell types, *J Immunol* 1998, 161:385–389.
93. Cortegano I, Del Pozo V, Cárdaba B, Arrieta I, Gallardo S, Rojo M, Aceituno E, Takai T, Verbeek S, Palomino P, Liu FT, and Lahoz C: Interaction between galectin-3 and FcgammaRII induces down-regulation of IL-5 gene: implication of the promoter sequence IL-5REIII, *Glycobiology* 2000, 10:237–242.
94. Sano H, Hsu DK, Apgar JR, Yu L, Sharma BB, Kuwabara I, Izui S, and Liu FT: Critical role of galectin-3 in phagocytosis by macrophages, *J Clin Invest* 2003, 112:389–397.
95. Chen HY, Sharma BB, Yu L, Zuberi R, Weng IC, Kawakami Y, Kawakami T, Hsu DK, and Liu FT: Role of galectin-3 in mast cell functions: Galectin-3-deficient mast cells exhibit impaired mediator release and defective JNK expression, *J Immunol* 2006, 177:4991–4997.
96. Sano H, Hsu DK, Yu L, Apgar JR, Kuwabara I, Yamanaka T, Hirashima M, and Liu FT: Human galectin-3 is a novel chemoattractant for monocytes and macrophages, *J Immunol* 2000, 165:2156–2164.
97. Colnot C, Ripoche MA, Milon G, Montagutelli X, Crocker PR, and Poirier F: Maintenance of granulocyte numbers during acute peritonitis is defective in galectin-3-null mutant mice, *Immunology* 1998, 94:290–296.
98. Zuberi RI, Hsu DK, Kalayci O, Chen HY, Sheldon HK, Yu L, Apgar JR, Kawakami T, Lilly CM, and Liu FT: Critical role for galectin-3 in airway inflammation and bronchial hyperresponsiveness in a murine model of asthma, *Am J Pathol* 2004, 165:2045–2053.
99. del Pozo V, Rojo M, Rubio ML, Cortegano I, Cardaba B, Gallardo S, Ortega M, Civantos E, Lopez E, Martin-Mosquero C, Peces-Barba G, Palomino P, Gonzalez-Mangado N, and Lahoz C: Gene therapy with galectin-3 inhibits bronchial obstruction and inflammation in antigen-challenged rats through interleukin-5 gene downregulation, *Am J Respir Crit Care Med* 2002, 166:732–737.
100. Lopez E, Del Pozo V, Miguel T, Sastre B, Seoane C, Civantos E, Llanes E, Baeza ML, Palomino P, Cardaba B, Gallardo S, Manzarbeitia F, Zubeldia JM, and Lahoz C: Inhibition of chronic airway inflammation and remodeling by galectin-3 gene therapy in a murine model, *J Immunol* 2006, 176:1943–1950.
101. Bernardes ES, Silva NM, Ruas LP, Mineo JR, Loyola AM, Hsu DK, Liu FT, Chammas R, and Roque-Barreira MC: *Toxoplasma gondii* infection reveals a novel regulatory role for galectin-3 in the interface of innate and adaptive immunity, *Am J Pathol* 2006, 168:1910–1920.
102. Beatty WL, Rhoades ER, Hsu DK, Liu FT, and Russell DG: Association of a macrophage galactoside-binding protein with Mycobacterium-containing phagosomes, *Cell Microbiol* 2002, 4:167–176.
103. Hrdlickova-Cela E, Plzak J, Smetana K, Jr., Melkova Z, Kaltner H, Filipec M, Liu FT, and Gabius HJ: Detection of galectin-3 in tear fluid at disease states and immunohistochemical and lectin histochemical analysis in human corneal and conjunctival epithelium, *Br J Ophthalmol* 2001, 85:1336–1340.
104. Ohshima S, Kuchen S, Seemayer CA, Kyburz D, Hirt A, Klinzing S, Michel BA, Gay RE, Liu FT, Gay S, and Neidhart M: Galectin 3 and its binding protein in rheumatoid arthritis, *Arthritis Rheum* 2003, 48:2788–2795.

30 X-Lectins: A New Family with Homology to the *Xenopus laevis* Oocyte Lectin XL35

Jin Kyu Lee and Michael Pierce

CONTENTS

30.1 INTRODUCTION

In 1988, Herman Stillmark discovered a hemagglutinin (that he originally called ricin) in extracts of castor beans. Although lectins were first described in plants in relation to their cell agglutinating properties [1,2], molecules that bind carbohydrate residues have now been discovered in viruses, microorganisms, plants, and animals [1,3,4]. It has been shown that lectins serve to mediate diverse carbohydrate recognition events in plants and animal tissues of both embryonic and adult origins.

Animal lectins have been found to be associated with the cell surface, the cytoplasm, and the nucleus. At the cell surface, lectins can act as receptors involved in selective intercellular adhesions and cell migration [5,6], as well as in the recognition of circulating glycoproteins [7]. Lectins have also been shown to function as receptors for the extracellular matrix proteins, elastin, and laminin [8,9] for glycosaminoglycans that mediate the binding of the proteoglycans to the sugars of other matrix glycoproteins [10], and mediate the first steps in immune cell adhesion to endothelia [11]. Proteins with carbohydrate residue affinities also function in the innate immune system of vertebrates [12].

Although the number of animal lectins discovered continues to increase, these lectins can be classified into distinct families based on protein sequence homologies [13,14]. A recent classification indicates that most of fall into one of five major groups: Ca^{2+}-dependent lectin (C-type lectins), galectins (S-type lectins), mannose-6-phosphate receptors (P-type lectins), siglec and other immunoglobulin-like sugar binding lectins (I-type lectins), and lectins related in sequence to the

leguminous plant lectins (L-type lectins) [15]. In this chapter, we review a relatively new family of animal lectins, the "X-lectins", that have no C-type lectin domains (CRD) in their primary sequences, but have been shown to display binding activity to carbohydrate residues only in the presence of Ca^{2+}. These lectins appear to function in innate immunity, but members have been shown to have specialized functions as well.

The X-lectin family was first studied as the *Xenopus laevis* oocyte cortical granule lectin (XL35) in oocyte and embryos. The cDNA encoding XL35 was isolated from a *Xenopus* oocyte cDNA library and its predicted amino acid sequence does not display the C-type lectin motif, although it does require calcium for binding [16]. The lack of sequence similarity of the XL35 lectin with other known calcium-dependent lectin families suggested that XL35 represents the first member of a new family of lectins. XL35 was shown to consist of 313 amino acids with three potential N-linked oligosaccharide sites and includes predicted signal sequence.

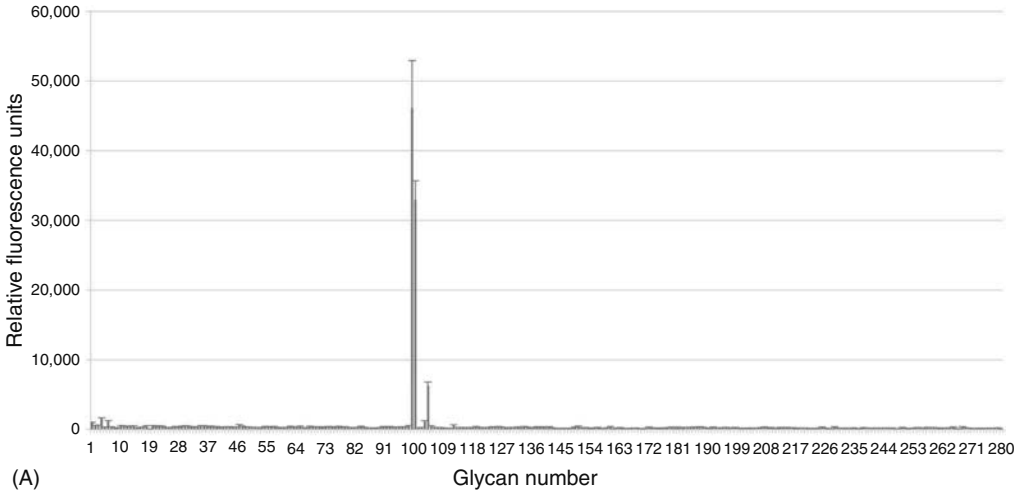
Several nucleic acid sequences that predict proteins homologous to XL35 have since been reported in frog, human, mouse, lamprey, trout, and ascidian. These predicted proteins also show high degrees of amino acid sequence homology to a common fibrinogen-like motif that may involve carbohydrate binding. Although their biological functions and carbohydrate binding specificities of all of these members have not been studied in detail, they appear to have common characteristics. Several independent studies on this new family of lectins strongly suggest that some members are expressed and stored in specialized vesicles that may be released upon infection by pathogens. In addition, some family members have been shown to bind to oligosaccharides of bacterial pathogens. Therefore, this family of lectins likely participates in pathogen surveillance as part of the innate immune system in humans, murine, and frogs. We proposed the name "X-lectin" family for these homologs, since these lectins are all clearly homologs of XL35 [17].

30.2 *XENOPUS LAEVIS* OOCYTE CORTICAL GRANULE LECTIN (XL35) WAS THE FIRST IDENTIFIED LECTIN AMONG THE X-LECTIN FAMILY

Studies by two groups have demonstrated that *X. laevis* oocytes and embryos contain soluble, calcium-dependent lectins that form an oligomeric structure with an apparent molecular weight of 500 kDa under nonreducing conditions [18,19]. Reducing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) reveals a monomer of about 43–45 kDa with size heterogeneity due to N-linked oligosaccharides, suggesting that the native structure of the lectin is a 12-mer [18,20]. The hydrophobicity of the XL35 C-terminus and the lack of oligomerization after limited acid digestion of XL35 suggest that the region responsible for forming oligomers lies in the C-terminus selection of the polypeptide [21]. Interestingly, mass spectrometric studies demonstrated that the majority of the purified XL35 polypeptide from the oocyte does not show cleavage of its secretion signal sequence, although removal of the signal peptide was detected on a small fraction, less than 10% [21]. This result suggests a unique biosynthetic pathway for XL35 in the oocyte that allows glycosylation and packaging into the cortical granules without cleavage of the signal sequence. XL35 is contained within the cortical granules of the oocytes, as well as in several other intracellular locations, and it is released from the cortical granules at fertilization [22]. Several lines of evidence suggest that the released multimeric lectins bind to oligosaccharide targets on glycoproteins in the egg jelly coat where they participate in the formation of the fertilization envelope that blocks sperm entry [23]. These ligands are expressed on ~500 kDa mucin-like glycoproteins cross-linked by disulfide bonds, each containing hundreds of O-linked saccharides [24]. XL35 has a remarkable ability to bind a wide variety of both monovalently and polyvalently presented D-galactopyranosides, and binding is calcium-dependent [25]. Hydrolysis of the purified XL35 ligand with a series of exoglycosidases showed that a terminal α-galactose was the ligand structure required for recognition by XL35 [24]. The ligand was rich in the potentially glycosylated β-hydroxy amino

acids, Ser, Thr, and Gly, which are typical of glycoproteins containing O-linked glycans such as mucins. The structure of several neutral oligosaccharides released from glycoproteins of *X. laevis* jelly coat by β-elimination has been reported [26]. Many of these oligosaccharide structures were found to contain a terminal α-galactose residue: Galα1–4(Fucα1–2)Galβ1–3GalNac-, Galα1–4(Fuc1–2)Galβ1–3(GlcNAcβ1–6)GalNac-, Galα1–4(Fucα1–2)Galβ1–3(Fucα1–2)Gal- and Galα1–4(Fuc1–2)Galβ1–3(Fucα1–2)Galβ1–3(GlcNAcβ1–6)GalNAc-.

Recent results from a glycan array screening analysis with fluorescent labeled XL35 showed the very clear carbohydrate binding specificity of XL35 (Figure 30.1) (http://www.functionalglycomics.org/glycomics/publicdata/selectedScreens.jsp). The glycan array production and assay was performed by Core H of the Consortium for Functional Glycomics (CFG) (http://www.functionalglycomics.org/fg/index.shtml). The array consists of 285 glycans in replicates of six. The results show that in the presence of calcium, specificity was restricted to glycans with a terminal alpha-linked galactose. XL35 prefers to be attached to a GalNAc residue, more than any other residue assayed (Figure 30.1). These results are consistent with earlier studies and graphically demonstrate the specificity of XL35.



(A)

Glycan number	Glycan name	Relative bound fluorescence units
101	Galα1–3GalNAcα–Sp8	46,030
102	Galα1–3GalNAcβ–Sp8	32,832
106	Galα1–3Galβ1–4Glcβ–Sp0	6,182
4	Ceruloplasmine	1,490
105	Galα1–3Galβ1–4GlcNAcβ–Sp8	1,080
6	Transferrin	1,049
1	AGP	922
47	[6OSO3]GlcNAcβ–Sp8	605

(B)

FIGURE 30.1 Glycan binding arrary analysis of XL35 shows it binds to alpha-linked galactose. The fluorescent labeled XL35 has been assayed for its binding specificities using a glycan array consisting of 285 different glycans structures. Its binding specificity was restricted to glycans with a terminal alpha-linked galactose. (A): Relative binding intensity of XL35 to 285 different glycan structures. (B): Glycan names bound to XL35 among 285 different glycans. The glycan array was produced and assayed by CFG. The full data are available at http://www.functionalglycomics.org/fg/index.shtml. (From Lee, J.K., et al., *Glycoconj. J.*, 21(8–9), 443, 2004. With permission.)

30.2.1 Homologs of XL35 in Frog Embryos

To examine the expression patterns of XL35 mRNA at fertilization and during embryo development, Northern analysis was performed on total RNA purified from Stage VI oocytes and from embryos at various stages of development. These results showed that relatively high levels of XL35 mRNA were present in Stage VI oocytes and persisted through gastrulation, after which it declined [16]. Compared to the levels of expression in gastrulae, low levels of XL35 mRNA was present in hatching tadpoles. Since it is highly unlikely that maternal mRNA persisted until tadpole stages, together with the observation of an increase of RNA levels at gastrulation, it appeared that XL35 mRNA is newly transcribed at the midblastula transition along with many other zygotic RNAs. The fact that these RNA are transcribed zygotically, as well as maternally, strongly supports the hypothesis that XL35 displays other functions in addition to its role in fertilization [20,27].

Nomura et al. [28] showed that monoclonal antibodies against human blood group-B-type trisaccharides (B-substance, Galα1–4(Fucα1–2)Galβ1–3-) completely block the Ca^{2+}-dependent cell–cell adhesion system in *X. laevis* embryonic (blastula stage) cells. Synthetic B-substance glycopeptides also disrupt this Ca^{2+} -dependent cell–cell adhesion. These authors purified membrane glycoproteins that reacted with B-substance saccharides and showed they are glycosylphosphatidyl inositol (GPI)-anchored proteins. Amino acid sequence analysis of the purified protein showed that these proteins are homologs of XL35 [28]. These results indicate that the GPI-anchored XL35 homologs that recognize the B-substance trisaccharide are directly involved in Ca^{2+} -dependent cell–cell adhesion of *Xenopus* embryonic blastula cells.

The oocyte cortical granule lectin from *Silurana tropicalis* (western clawed frog) has also been reported in GenBank (accession no: AY079196). The amino acid identity between the cortical granule lectin from *S. tropicalis* and XL35 was 85% (similarity, 95%). The total number of amino acids in the open reading frame was 320 compared to 313 in XL35. This protein is clearly a homolog of XL35 (Figure 30.2) and it will be interesting to determine if its binding specificity and functions are similar.

Recently, Ishino et al. have identified two amino acid sequences and cloned from *X. laevis* serum, termed the 35 kDa serum lectin (accession no: AB061238) and lectin type 2 (accession no: AB061239). These lectins showed a high degree of amino acid sequence homology with XL35 (Figure 30.2). The overall amino acid identity between the 35 kDa serum protein and XL35 was 59% (similarity, 81%), while there was 59% amino acid identity (similarity, 84%) between lectin type 2 and XL35 (Figure 30.2). The open reading frame for the 35 kDa serum lectin was predicted to encode 338 amino acids, while lectin type 2 was 315 amino acids. The predicted N-terminal region of these two lectins is composed of hydrophobic amino acids, suggesting the presence of a signal peptide sequence that causes proteins to enter the secretory pathway, similar to XL35. Earlier, Barondes and coworkers had reported that serum from estrogen-induced *X. laevis* contained a 69 kDa protein that was weakly reactive against antibody against the cortical granule lectin. The serum protein also bound to immobilized melibiose in a Ca^{2+} -dependent manner and a peptide mapping analysis suggested some similarity with the cortical granule lectin (XL35) [29]. It will be interesting to determine the relationship of the 35 kDa serum lectin and the lectin type 2 may be related to this 69 kDa protein.

30.3 MOUSE HOMOLOGS OF XL35: INTELECTIN-1 AND INTELECTIN-2

The first report on a mouse homolog of XL35 was the isolation of its cDNA using a large-scale in situ hybridization screening method. They named the protein encoded by this cDNA "intelectin," since it was shown to be expressed in small intestine. The term, intelectin, is now sometimes used for homologs of XL35. The intelectin amino acid sequence revealed a 61% homology with XL35 [30] (Figure 30.2). Northern blot analysis revealed the mRNA corresponding to the cDNA was 1.2kb in length, and expression was specific to small intestine. The mRNA of intelectin appeared to be expressed in small intestine Paneth cells in situ hybridization studies. Interestingly, an intelectin

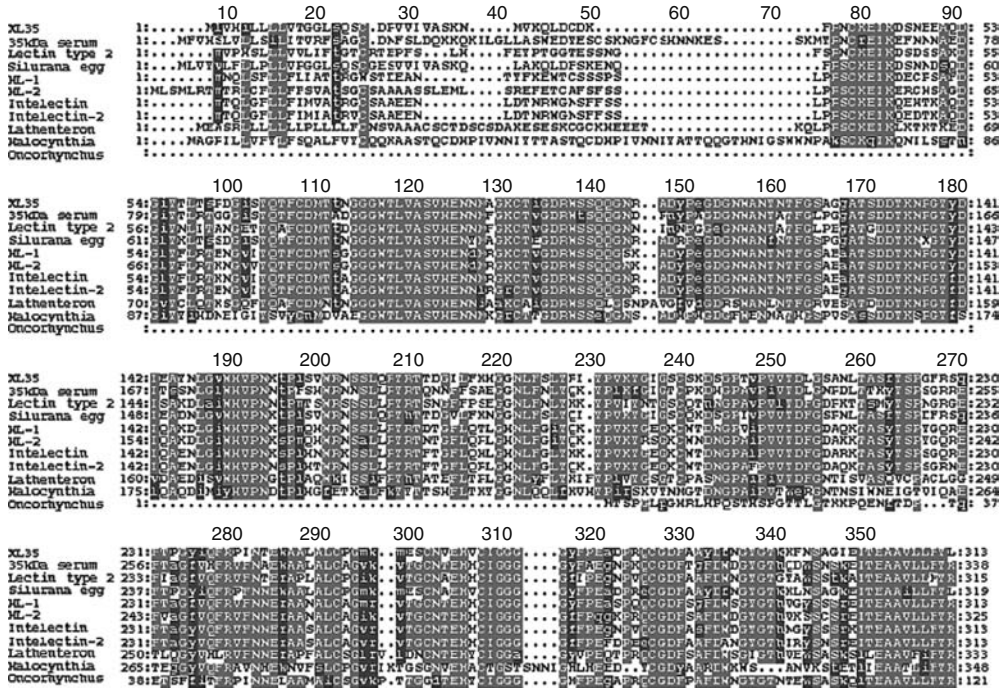


FIGURE 30.2 (See CD for color figure.) Amino acid sequences comparisons of X-lectin family. Amino acid sequences of all putative homologues from various eukaryotes are shown. Sequence information was obtained from the published paper or GenBank database. XL35, *Xenopus laevis* oocyte cortical granule lectin: 35 kDa serum, *X. laevis* 35 kDa serum lectin (AB061238): Lectin type 2, *X. laevis* lectin type 2 (AB061239): Silurana egg, *Silurana tropicalis* egg cortical granule lectin (AY079196): HL-1, human HL-1 (AY065972): HL-2, human HL-2 (AY065973): Intelectin, mouse intelectin (AB016496): Intelectin-2, mouse intelectin-2 (AY217760): Lathenteron, *Lathenteron japonicum* lamprey serum lectin (AB055981): Halocynthia, *Halocynthia roretzi* ascidian galactose lectin: Oncorhynchus, *Oncorhynchus mykiss* (rainbow trout) (AF281350): The identical amino acids are shown in red background characters, and similar amino acids are shown in blue background. (From Lee, J.K., et al., *Glycoconj. J.*, 21(8–9), 443, 2004. With permission.)

sequence has been also identified in a 10-day-old mouse pancreas cDNA library [31] (accession no: AK065973), and no further characterization of the protein encoded by this cDNA can be found.

Another mouse homolog, intelectin-2, has been detected by the proteomic analysis of mouse jejunal epithelium and its response to infection with the intestinal nematode, *Trichinella spiralis* [32]. Intelectin-2 was cloned from BALB/c mRNA extracted on day 14 of infection of *T. spiralis* and was found to have 91% amino acid identity with mouse intelectin (now termed intelectin-1) (Figure 30.2). Intelectin-2 transcripts were upregulated early (day 3) during infection with *T. spiralis* in BALB/c mice and levels remained high through to day 14 (time of parasite rejection), but returned to undetectable levels at day 56, following resolution of infection [33]. Immunohistochemistry of jejunal sections followed a similar pattern, with intense labeling of goblet and Paneth cells at day 14. However, intelectin-2 transcripts and protein were absent when C57BL/10 mice were infected with *T. spiralis*. Genomic polymerase chain reaction (PCR) and Southern blotting confirmed that the intelectin-2 gene is absent from the C57BL/10 genome. In uninfected BLAB/c mice, mRNA for intelectin-1 was expressed at high levels throughout the gut, but was not detected in significant amount in other tissues. Nematode infection did not cause any apparent change in intelectin-1 expression in the gut, however. These observations suggest that intelectin-2 may serve a protective role in the innate immune response to parasite infection and may be more similar to HL-2 function (see Section 30.4).

Mouse intelectin-1 is also expressed in the airway epithelial cells of IL-13-overexpressing transgenic mice. IL-13 overexpression in a mouse model is necessary and sufficient for the development for experimental asthma. These mice develop airway hyperreactivity and mucus overproduction. Intelectin was highly expressed (20 times in microarray and 146 times in PCR analysis) in airway epithelial cells from mice with asthma compared to controls [34]. Intelctin-1 in the airway might alter the response of subjects with asthma to infection or colonization with bacterial or fungal pathogens.

30.4 XL35 HOMOLOGS IN FISH HAVE IMPORTANT FUNCTIONS IN THEIR INNATE IMMUNE RESPONSE

In addition to human, mouse, and frog, sequence homologs of XL35 have been reported in several other eukaryotes, especially in fish (Table 30.1 and Figure 30.2), in which innate immune responses are the dominant defense system against pathogens. The cDNA encoding grass carp intelectin was isolated from a kidney cDNA library and termed gcIntL [35]. The deduced amino acid sequence of gcIntL consists of 318 amino acids and was about 55% identical and 74% similar to a human

TABLE 30.1
Summary of the X-Lectin Family from Various Sources

Organisms	Name	Tissues Found Proteins or mRNA	Proposed Biological Functions (Binding Specificity)	References
Xenopus laevis	XL35	Oocyte (cortical granule)	Transformation of fertilized envelop, embryogenesis, bind to *Cryptococcus neoformans* and candida (galactose)	[16,17]
	35 kDa serum lectin	Serum	?	Ishino et al. 2002
	Lectin type 2	Serum	?	Ishino et al. 2002
Silurana tropicalis	Egg cortical granule Lectin	Egg	?	Lindsay et al. 2002
Mouse	Intelectin	Small intestine (Paneth cell), epithelial	Induced by IL-13 transgenic mice in lung	Komiya et al. 1999
		Cells in lung	Involved in asthma pathway	[34]
	Intelectin-2	Paneth cells and goblet cells in intestinal epithelium	Induced by infection of *Trichinella spiralis*	[32]
Human	HL-1, intelectin, omentin, lactoferrin receptor	Heart, colon, thymus, small intestine, visceral adipose tissue, Paneth cells, goblet cells along with the entire crypt-villus axis	Recognize bacterial arabinogalactan of *Nocardia*, affinity to D-galactose and D-galactofuranosyl residue, lactoferrin receptor, adipocytokine, increase glucose transport	[16,44,45,53]
		-	Lactoferrin receptor	[50]

TABLE 30.1 (continued)
Summary of the X-Lectin Family from Various Sources

Organisms	Name	Tissues Found Proteins or mRNA	Proposed Biological Functions (Binding Specificity)	References
	HL-2	Small intestine (Paneath cell)	Bind to *Cryptococcus neoformans* and *Candida albicans*	[41]
Lathenteron japonicum	Lamprey serum lectin	Serum	?	Yoshimura et al. 2001
Ciona intestinalis		Serum	?	Satou et al. 2002
Oncorhynchus mykiss		Liver	Induced by the injection of *Vibrio anguillarum*	[39]
Halocynthia roretzi	Asidian galactose Lectin	Serum	Enhance phagocytosis by *H. roretzi*	[38]
	Lectin		Hemocytoes	

Source: From Lee, J.K., et al., *Glycoconj. J.*, 21(8–9), 443, 2004. With permission.

homolog, HL-1 (Figure 30.2). Using real-time quantitative reverse transcriptase polymerase chain reaction (RT-PCR) analysis, gcIntL transcripts were significantly induced in head kidney, trunk kidney, spleen, and intestine from lipopolysaccharide (LPS)-stimulated fish. RT-PCR and Western blotting analysis demonstrated that gcIntL mRNA and protein were detected in brain, gill, intestine, head kidney, trunk kidney, spleen, and heart. Furthermore, gcIntL protein could also be detected in gill, intestine, trunk kidney, head kidney, spleen, heart, and brain.

Yokosawa et al. investigated the defense mechanism in the ascidian *Halocynthia roretzi*, which occupies a phylogenic position between the vertebrate and invertebrates [36] (Figures 30.2 and 30.3). They isolated several candidate defense molecules from plasma and hemocytes. Among them, a galactose-specific lectin was purified from plasma and demonstrated to stimulate the production of superoxide anions by mammalian polymorphonuclear leukocytes [37]. The complete amino acid sequence of the galactose-specific lectin from the plasma of the ascidian *H. roretzi* was determined by sequential Edman degradation, analysis of peptide fragments derived by proteolytic fragmentation, and chemical cleavage of the reduced S-pyridylethylated lectin. The amino acid sequence was verified by cDNAs isolated from a *H. roretzi* hepatopancreas cDNA library. The protein consisted of a total of 348 amino acids, including a putative signal sequence [38]. The putative amino acid sequence showed ~40% identity (~70% similarity) to XL35 (Figure 30.2). The authors reported that this lectin functions as a phagocytosis-stimulating molecule [38]. Although its effect appeared to be weak, they

FIGURE 30.3 (See CD for color figure.) Dendogram of the X-lectin family. Sequence were compared by multiple sequence alignment using the CLUSTAL W algorithm. (From Thompson, J.D., Higgins, D.G., and Gibson, T.J., *Nucleic Acids Res.*, 22(22), 4673, 1994; Lee, J.K., et al., *Glycoconj. J.*, 21(8–9), 443, 2004. With permission.)

suggested that the lectin in the plasma may bind to and agglutinate invading foreign materials via galactose residues, which would enhance phagocytosis by the hemocytes.

Bayne et al. analyzed differentially expressed genes in the livers of *Oncorhynchus mykiss* (rainbow trout) in the course of an acute phase response using suppression subtractive hybridization of cDNAs from the livers of unstimulated trout and of trout given a potent inflammatory stimulus (after intraperitoneal injection with a *Vibrio* bacterin). They isolated a cDNA of putative homolog of XL35 along with 25 other genes thought to be potentially immune-related [39,40]. One of the isolated cDNAs showed a partial open reading frame of 121 amino acids that showed 46% identity (~73% similarity) to the C-terminal region of XL35 (Figure 30.2). These findings strongly suggest that adult fish also express at least one type of XL35 homolog in response to infection by a pathogenic microorganism.

Yoshimura et al. reported the cDNA sequence of a putative homolog of XL35 from lamprey, *Lethenteron japonicum* (accession no. AB055981), termed the *Lampetra japonica* serum lectin. This cDNA encoded 333 amino acids and showed 47% identity (~75% similarity) with XL35 (Figure 30.2). No functional studies on this lectin have been reported as yet.

Because invertebrates lack an adaptive immune system, they had to evolve effective intrinsic defense strategies against a variety of microbial pathogens. The innate immune system of invertebrates includes a hemolymph coagulation system, which participates both in defense against microbes and in hemostasis. In invertebrates, it has been shown that lectins present in the hemolymph play key roles in biological defense against invading foreign materials as nonspecific defense molecules [38]. Although X-lectin homolog function has just begun to be studied in fish, lamprey, and ascidian defense mechanism, they likely play an important role in innate immune responses.

30.5 HUMAN HOMOLOGS OF XL35 HAVE BEEN REPORTED WITH VARIOUS BIOLOGICAL FUNCTIONS AND NAMES: HL-1, HL-2, INTELECTIN, LACTOFERRIN RECEPTOR, AND OMENTIN

Since the molecular cloning of XL35, several mammalian homologs of XL35 have been identified. Two human XL35 homologs, termed HL-1 and HL-2 [41], were identified and cloned rapidly after XL35 cloning using the human expressed sequence tag (EST) data bank (GenBank, Accession number Z36760) [16,41]. These two cDNA sequences showed 85% identity to one another at the deduced amino acid level. The overall amino acid identity between HL-1 and XL35 was 60% (similarity, 74%) with a 56% amino acid identity (similarity, 74%) between HL-2 and XL35 (Figure 30.2). The open reading frame for HL-1 was the same size as that of XL35, 313 amino acids, while HL-2 was predicted to have 325 amino acids. Both HL-1 and HL-2 are predicted to have signal peptide sequences and are encoded at chromosome 1q21.3 and 1q22–23.5, respectively, the same locus that encodes the selectins [42]. BLAST analysis against a human genomic DNA database showed the distance between the two genes encoding HL-1 and HL-2 is only 7000 base pairs, implying a gene duplication event was likely involved.

Northern blot analysis showed selective expression of HL-1, in heart, small intestine, colon, and thymus, with lower levels in ovary, testis, and spleen. Other tissues showed detectable levels of expression: skeletal muscle, placenta, and spleen. Colon, liver, and thymus expressed HL-1 in their vascular endothelial cells using immunohistochemistry on tissue sections [41]. Human aortic endothelial cells express HL-1 but the expression was apparently unaltered in cultured cells by the addition of an endothelial cell activating agent [41]. HL-2, however, was expressed only in small intestine [41] by Northern and PCR analysis. Using a peptide-specific antibody, HL-2 was localized by immunostaining to the Paneth cells, specialized secretory cells whose main function is in pathogen surveillance (unpublished data). The Paneth cells in the small intestine of most mammals produce α-defensins and other antimicrobial proteins including lysozyme, secretory phospholipase A2, HIP/PAP, and RegIIIγ [12,43]. Microbial colonization of germ-free mice triggers epithelial expression

of RegIIIγ, a secreted C-type lectin. RegIIIγ binds intestinal bacteria but lacks the complement recruitment domains present in other microbe-binding mammalian C-type lectins. Recently, Cash et al. showed that RegIIIg is by itself an antimicrobial protein that binds it bacterial targets via interactions with peptidoglycan carbohydrate [12]. These findings strongly implicated the involvement of the human homologs of XL35 in host defense against pathogens in the small intestine, thus making them a part of the innate immune system.

Tsuji et al. purified and cloned a protein from human placenta and named this protein human intelectin (hIntL) [44] because of its homology to the mouse intelectin. Surprisingly, the deduced amino acid sequence of hIntL was the same as that of HL-1. The protein was reported to be expressed at very low levels as a secreted form from the rabbit kidney cell line RK-13. The hIntL/HL-1 was reported to be absorbed to galactose-Sepharose and was completely eluted with EDTA. About 50% of the absorbed HL-1 was eluted by buffers containing 100 mM galactose, *N*-acetylgalactosamine, or fructose, however. The same concentrations of mannose, glucose, *N*-acetylmannosamine, *N*-acetylglucosamine, sorbose, D-fucose, L-fucose, L-rhamnose, and 2-deoxy-D-glucose did not appear to elute hIntL from galactose-Sepharose. The protein was also reported to be effectively eluted by D-pentose, D-xylose, D-ribose, and 2-deoxy-D-ribose. Interestingly, however, hIntL was not eluted from galactose-Sepharose by melibiose or lactose. These results demonstrated that the carbohydrate-binding specificities of hIntL (HL-1) and XL35 are distinct, but suggested that the specificity was not clear. The hIntL also appeared to bind to the bacterial arabinogalactan from the cell wall of *Nocardia ruba* containing D-galactofuranosyl residues. This binding was completely inhibited by EDTA, D-ribose, D-galactose, and D-arabinose, but not D-glucose. Pentoses (D-ribose, D-xylose, D -lyxose, and D-arabinose) and D-galactose inhibited the binding of HL-1 to arabinogalactan more effectively than D-mannose or D-glucose. In conclusion, hIntL appears to be is a lectin that can at least partially recognize galactofuranose and likely plays a role in the recognition of bacteria-specific component in the host.

Another report suggested a possible, additional function for HL-1. The entire coding region of a lactoferrin receptor cDNA was reported to be cloned by PCR, based on amino acid sequences of a protein from fetal intestine that bound lactoferrin [45]. Surprisingly, the amino acid sequence of this lactoferrin receptor was reported to be a 100% match with that of HL-1. The apparent molecular mass of recombinantly expressed protein in a baculovirus-insect cell system was 136 kDa under nonreducing conditions and 34 kDa under reducing conditions, suggesting a tetramer under native conditions. Phosphoinositol phospholipase C treatment indicated that the lactoferrin receptor is GPI anchored. The C-terminal region of the GPI-anchored proteins should consist of a predominantly hydrophobic region of 8–20 amino acids, which directs the addition of preformed GPI anchor. There is a putative cleavage site in the sequence of lactoferrin receptor/HL-1 at residue 298 that could then be attached to the GPI anchor. This is a very interesting observation, in light of the finding of *Xenopus* homologs of XL35 that were GPI-linked and involved in cell adhesion, as mentioned above [28]. In addition, lactoferrin is known to have a variety of antimicrobial activities [46–48]. It is possible, therefore, that HL-1, in its role as the lactoferrin receptor, can modulate the antimicrobial effects of lactoferrin [45]. Since lactoferrin also functions in the mammalian embryo [49], it is also possible that HL-1 expressed in the mammalian oocyte and blastula, functions as a lactoferrin receptor, and mediates the function of lactoferrin during fertilization and early embryogenesis. Lactoferrin bound to this lactoferrin receptor in a Ca^{2+}-dependent manner during affinity chromatography. This binding may involve the glycans on lactoferrin, since it is a glycoprotein, although the role of oligosaccharides on lactoferrin binding to lactoferrin receptor (HL-1) was not examined. A recent publication suggested that lactoferrin receptors are present in both the secretory granules of lysozyme-positive Paneth cells in the bottom of the intestinal crypts, as well as in goblet cells along the crypt-villus axis [50]. But quantitatively, the major site of lactoferrin receptor localization was the enterocyte brush border in small intestine. This membrane is organized in stable glycolipid-based lipid raft microdomains and has long been known to harbor receptor for lactoferrin. Enterocytes synthesize a number

of brush border GPI-anchored proteins, showing this cell type harbors the enzymes required for anchorage of GPI in the endoplasmic reticulum. A major part of lactoferrin receptors in enterocytes was released by phosphatidylinositol-specific phospholipase C, indicating a membrane insertion by a GPI anchor. Interestingly, lactoferrin receptor (HL-1) was suggested to be a major component of microvillar lipid rafts and superrafts. This strategic localization suggests that the lactoferrin receptor (HL-1) serves as an organizer and stabilizer of the brush border membrane, preventing loss of digestive enzymes to the gut lumen and protecting the glycolipid microdomains from pathogens [50].

Another interesting study of HL-1 (Intelectin-1) has been done on malignant pleural mesothelioma (MPM), which is a fatal neoplasm with no acceptable curative approaches [51]. Serial analysis of gene expression (SAGE) was used to compare the gene expression pattern of a surgically resected MPM to the autologous normal mesothelium. Intelectin gene overexpression (>139-fold) was found in the tumor. Online SAGE datasets revealed intelectin to be consistently present in mesotheliomas, ovarian cancer, and colon cancer. Intelectin mRNA expression was found by RT-PCR in resected MPM tumors, and intelectin protein expression was confirmed by immunohistochemistry in MPM tumors. These observations suggest that intelectin overexpression in mesothelioma could have potential screening, and therapeutic implications.

30.5.1 HL-1 Is a Novel Adipocytokine and Enhances Glucose Transport: Omentin

Recently, a new adipocytokine expressed in human omental adipose tissue, termed Omentin, was reported and revealed to be 100% identical in amino acid sequence to that of HL-1 [52,53]. Omentin mRNA was predominantly expressed in visceral adipose tissue and was barely detectable in subcutaneous fat depots in human and rhesus monkeys. Whether there exists a unique type of cell in the omental adipose depot or whether a unique local environment induces omentin expression from a specific cell type is not clear, since an immunohistological study has not been performed. Omentin is among the first molecules known to exhibit dramatic difference in gene expression between the two major types of fat deposits. Furthermore, omentin enhanced only insulin-mediated glucose transport and did not stimulate basal glucose transport [53]. Very recently, Batista et al. showed that plasma omentin and mRNA expression levels are inversely related to obesity [54]. This result suggests that higher omentin levels in serum may be seen as a positive factor that opposes the obese state.

These observations strongly suggest that omentin is a novel adipokine that is expressed in omental adipose tissue in human and may regulate insulin action. It is certainly possible that omentin (HL-1) may be present in endothelial cells in the vessels that vascularize the fat tissue, which would be consistent with the location identified in other tissues. Various adipocytokines (adipocyte-secreted hormones) have been described, which profoundly affect insulin sensitivity and might potentially link obesity, insulin resistance, and cardiovascular disease [55,56]. Among these, adiponectin and visfatin appear as insulin-sensitizing adipocytokines, whereas TNF-a, IL-6, and resistin induce insulin resistance. One of the adipocytokines, leptin, is a fat-derived key regulator of appetite and energy expenditure in mice [56]. Due to their profound effects on whole-body glucose and energy metabolism, omentin is attracting interest as potential new therapeutics for diabetes mellitus and obesity related to pathogenesis. It will be interesting to define the omentin receptor on adipocytes and analyze the possible carbohydrate-binding specificity of omentin to its receptor, since it is highly homologous to a known lectin, XL35.

30.6 CONCLUSION AND FUTURE DIRECTIONS

A new family of lectins showing significant homology to the amino acid sequence of XL35 has been reported from several eukaryotes since the original cloning of XL35 (Figure 30.3). These findings demonstrate that a family of lectins, homologs of XL35, is present in broad range of species. We have proposed the name “X-lectin” family for these homologs of XL35. In addition to its function

in formation of the fertilization envelope, which blocks sperm entry [22,23,27], XL35 likely functions as well in cell–cell or cell–matrix adhesion events in the embryo [16,20,27]. Although several homologs of XL35 have been discovered in lower eukaryotes, as well as human or mouse, studies on their biological functions and carbohydrate-binding specificities have not been done in detail (Table 30.1).

None of these proteins has the C-type lectin domain (CRD) [57], even though several have been shown to display binding activity to carbohydrate residues only in the presence of Ca^{2+}. Instead of the CRD domain, all X-lectins, including the fish homologs, have a fibrinogen-like motif that is in the region of sequence that shows the highest degree of homology based on amino acid sequence alignment (Figures 30.2 and 30.4). Fibrinogen is a principal protein of vertebrate blood clotting and is found universally in blood from fish and echinoderms to mammals [58]. Early work on the evolution of vertebrate fibrinogen suggested a common origin of the arthropod hemolymph coagulation and the vertebrate blood coagulation systems. A thorough analysis of the evolution of vertebrate fibrinogen suggested most of the proteins bearing fibrinogen-related domains are lectins [59]. Members of the ficolin/opsonin/p35 lectin family also contain significant homology to this fibrinogen-like motif, although members of this family do not share any other similarity with the X-lectins [60,61]. The ficolin/opsonin/p35 family of proteins are found in serum where they are thought to bind to oligosaccharide structures on the surfaces of microorganisms, leading to the killing of bound microbes through complement activation and phagocytosis [62]. Collectins also bind to a wide range of

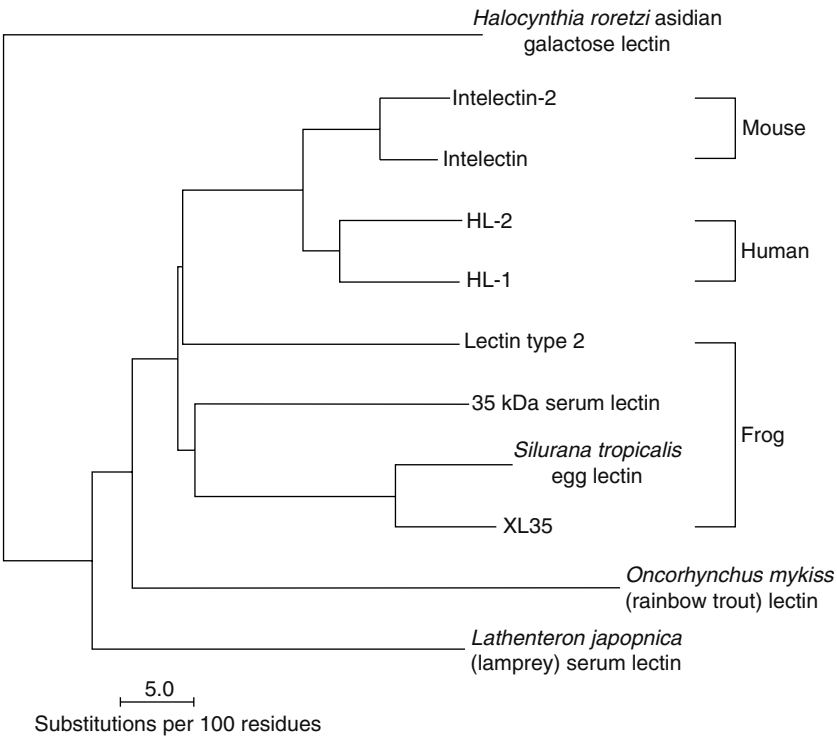


FIGURE 30.4 Comparison of fibrinogen-like motif between homologues of XL35 and ficolin/opsonin/p35 lectin family. The conserved fibrinogen-like motif in X-lectins may function in carbohydrate recognition. Only fibrinogen-like motif regions were compared: XL35, amino acids 66–134; HL-1, 66–134; HL-2, 77–165; mouse inteletin-1, 66–134; Fibrinogen-beta chain, 267–134; ficolin-1, 142–200; Hakata antigen, 117–175; p35, 129–187. Identical amino acids are shown in gray background characters and similar amino acids are shown in black. (From Lee, J.K., et al., *Glycoconj. J.*, 21(8–9), 443, 2004. With permission.)

sugar residues in a Ca^{2+}-dependent manner. In addition, ficolins bind to sugar residues that are rich on microbial surfaces, for example, *N*-acetyl-D-glucosamine, in a calcium-dependent manner [63]. Therefore, the conserved fibrinogen-like motif in X-lectins may function in carbohydrate recognition, as well.

Intestinal Paneth cells are the major producers of multiple peptides and proteins with antimicrobial activity in the intestine [64]. The most abundant and diverse of these are the defensins. They are highly microbicidal *in vitro* and probably important *in vivo*, yet their physiologic functions and mechanisms remain incompletely understood. Relative defensin deficiency may be a risk factor for Crohn's disease and infectious diarrhea. Antimicrobial lectins, particularly the hepatocarcinoma-intestine-pancreas/pancreatic-associated protein, RegIIIγ can lyse bacteria or interfere with their attachment to epithelial cells [12]. As discussed, several groups have shown that X-lectins members are expressed in mammalian Paneth cells. This location strongly suggests that mammal X-lectins function as a type of antimicrobial or microbicidal proteins secreted in Paneth cells in response to pathogens. Notably, the interactions of XL35 itself, after release from the cortical granules in complex and cross-linking with its high molecular weight glycoprotein ligand, participates in the formation of the fertilization membrane. This structure functions in the block of polyspermy, but may function as well as a barrier to microbial infection of the embryo.

It is likely that pathogenic infection can cause the induction of their transcription and release of members of the XL35 family from specialized vesicles. Some members are involved in the surveillance of the pathogens in the innate immune reaction, and their activities may involve opsonization, immobilization, or agglutination of the pathogens. The reg family is an example among proteins that have those biological functions. The *reg* gene family encodes a diverse group of secreted proteins that contain the conserved sequence motif found in C-type lectin domains. They bind to bacteria, but lack the complement recruitment domains present in other microbe-binding mammalian C-type lectins. Definition of the mechanisms of microbial cytotoxicity and carbohydrate-binding specificities will facilitate development of possible novel antimicrobial therapeutics based on this new lectin family. Importantly, as the X-lectins may recognize invariant saccharide components of pathogens, these molecules may be involved in the rapid recognition and control of microbial pathogens at the "front lines" of the innate immune response.

Very recently, HL-1 has been shown to function as a new member of the adipocytokine family in humans. Since there is now a link between several adipocytokines and the immune response [65], it is intriguing to speculate that through a gene duplication event, a cytotoxic lectin involved in the innate immune system (HL-2) gave rise to a homolog (HL-1), which is expressed in a subset of endothelial cells, including omental endothelia, which evolved to function as an adipocytokine. It will be very interesting to determine if HL-1 has retained a true carbohydrate-binding specificity for its receptor on adipocytes, or whether its ligand is completely peptide in nature, or a hybrid of peptide and glycan, such as observed with P-selectin, which binds portions of both the sialylated Lewis X glycan and sulfated tyrosine residues on its ligand, PSGL-1.

REFERENCES

1. Sharon, N., Lectins: Carbohydrate-specific reagents and biological recognition molecules. *J Biol Chem*, 2007. 282(5): 2753–2764.
2. Goldstein, I.J. and C.E. Hayes, The lectins: Carbohydrate-binding proteins of plants and animals. *Adv Car Chem Biochem*, 1978. 35: 127–340.
3. Sharon, N., Baterial lectins, cell–cell recognition and infectious disease. *FEBS Lett*, 1987. 217: 145–157.
4. Barondes, S., Soluble lectins: A new class of extracellular proteins. *Science*, 1984. 223: 1259–1264.
5. Lehmann, S., et al., An endogenous lectin and one of its neuronal glycoprotein ligands are involved in contact guidance of neuron migration. *Proc Natl Acad Sci USA*, 1990. 87(16): 6455–6459.
6. Rosen, S.D., Lymphocyte homing: Progress and prospects. *Curr Opin Cell Biol*, 1989. 1: 913–919.

7. Ashwell, G. and J. Harford, Carbohydrate specific receptors of the liver. *Ann Rev Biochem*, 1982. 51: 531–554.
8. Cooper, D., et al., P-selectin interacts with a beta 2-integrin to enhance phagocytosis. *J Immunol*, 1994. 153: 3199–3209.
9. Hinek, A., et al., The elastin receptor a galactoside-binding protein. *Science*, 1988. 239: 1539–1540.
10. Doege, K., et al., Complete primary structure of the rat cartilage proteoglycan core protein deduced from cDNA clones. *J Biol Chem*, 1987. 262(36): 17757–17767.
11. Rosen, S.D., Cell surface lectins in the immune system. *Sem Immunol*, 1993. 5: 237–247.
12. Cash, H.L., et al., Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science*, 2006. 313(5790): 1126–1130.
13. Drickamer, K. and M.E. Taylor, Biology of animal lectins. *Annu Rev Cell Biol*, 1993. 9: 237–264.
14. Powell, L.D. and A. Varki, I-type lectins. *J Biol Chem*, 1995. 270(24): 14243–14246.
15. Drickamer, K., Increasing diversity of animal lectin structures. *Curr Opin Struct Biol*, 1995. 5(5): 612–616.
16. Lee, J.K., et al., Cloning and expression of a *Xenopus laevis* oocyte lectin and characterization of its mRNA levels during early development. *Glycobiology,* 1997. 7: 367–372.
17. Lee, J.K., et al., The X-lectins: A new family with homology to the *Xenopus laevis* oocyte lectin XL-35. *Glycoconj J*, 2004. 21(8–9): 443–450.
18. Roberson, M.M. and S.H. Barondes, Lectin from embryos and oocytes of *Xenopus laevis*: Purification and properties. *J Biol Chem*, 1982. 257: 7520–7526.
19. Nishihara, T., et al., Isolation and characterization of a lectin from the cortical granules of *Xenopus laevis* eggs. *Biochemistry*, 1986. 25: 6013–6020.
20. Outenreath, R.L., M.M. Roberson, and S.H. Barondes, Endogenous lectin secretion into the extracellular matrix of early embryos of *Xenopus laevis. Dev Biol*, 1988. 125(1): 187–194.
21. Arrenz-Plaza, E., A.S. Tracy, A. Siriwardena, J.M. Pierce, and G.J. Boons, High-avidity, low-affinity multivalent interactions and the block to polyspermy in *Xenopus laevis. J Am Chem Soc*, 2002. 124: 13035–13046.
22. Wyrick, R.E., T. Nishihara, and J.L. Hedrick, Agglutination of jelly coat and cortical granule components and the block to polyspermy in the amphibian *Xenopus laevis. Proc Natl Acad Sci USA*, 1974. 71(5): 2067–2071.
23. Monk, B.C. and J.L. Hedrick, The cortical reaction in *Xenopus laevis* eggs: Cortical granule lectin release as determined by radioimmunoassay. *Zool Sci*, 1986. 3: 459–466.
24. Quill, T.A. and J.L. Hedrick, The fertilization layer mediated block to polyspermy in *Xenopus laevis*: Isolation of the cortical granule lectin ligand. *Arch Biochem Biophys*, 1996. 333: 326–332.
25. Arranz-Plaza, E., et al., High-avidity, low-affinity multivalent interactions and the block to ploypermy in *Xenopus laevis. J Am Chem Soc*, 2002.
26. Strecker, G., et al., Primary structure of 12 neutral oligosaccharide-alditols released from the jelly coats of the anuran *Xenopus laevis* by reductive beta-elimination. *Glycobiology*, 1995. 5(1): 137–146.
27. Roberson, M.M. and S.H. Barondes, *Xenopus laevis* lectin is localized at several sites in Xenopus oocytes, eggs, and embryos. *J Cell Biol*, 1983. 97(December): 1875–1881.
28. Nomura, K.H., et al., Involvement of blood-group-B-active trisaccharides in Ca^{2+}-dependent cell–cell adhesion in the Xenopus blastula. *Dev Genes Evol*, 1998. 208(1): 9–18.
29. Roberson, M., et al., Galactoside-binding serum lectin of *Xenopus laevis*. Estrogen-dependent hepatocyte synthesis and relationship to oocyte lectin. *JBC*, 1985. 260: 11027–11032.
30. Komiya, T., Y. Tanigawa, and S. Hirohashi, Cloning of the novel gene intelectin, which is expressed in intestinal Paneth cells in mice. *Biochem Biophys Res Commun*, 1998. 251(3): 759–762.
31. Carninci, P., et al., Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res*, 2000. 10(10): 1617–1630.
32. Pemberton, A.D., et al., Proteomic analysis of mouse jejunal epithelium and its response to infection with the intestinal nematode, *Trichinella spiralis. Proteomics*, 2004. 4(4): 1101–1108.
33. Pemberton, A.D., et al., Innate BALB/c enteric epithelial responses to *Trichinella spiralis*: Inducible expression of a novel goblet cell lectin, intelectin-2, and its natural deletion in C57BL/10 mice. *J Immunol*, 2004. 173(3): 1894–1901.
34. Kuperman, D.A., et al., Dissecting asthma using focused transgenic modeling and functional genomics. *J Allergy Clin Immunol*, 2005. 116(2): 305–311.
35. Chang, M.X. and P. Nie, Intelectin gene from the grass carp *Ctenopharyngodon idella*: cDNA cloning, tissue expression, and immunohistochemical localization. *Fish Shellfish Immunol*, 2006.

36. Yokosawa, H., et al., Galactose-specific lectin in the hemolymph of solitary ascidian, *Halocynthia roretzi*: Isolation and characterization. *Biochem Biophys Res Commun*, 1982. 107(2): 451–457.
37. Yokosawa, H., et al., Galactose-specific lectin in the hemolymph of solitary ascidian, *Halocynthia roretzi*. Molecular, binding and functional properties. *Biochim Biophys Acta*, 1986. 870(2): 242–247.
38. Abe, Y., et al., A unique primary structure, cDNA cloning and function of a galactose-specific lectin from ascidian plasma. *Eur J Biochem*, 1999. 261(1): 33–39.
39. Bayne, C.J., et al., Immune-relevant (including acute phase) genes identified in the livers of rainbow trout, *Oncorhynchus mykiss*, by means of suppression subtractive hybridization. *Dev Comp Immunol*, 2001. 25(3): 205–217.
40. Gerwick, L., G. Corley-Smith, and C.J. Bayne, Gene transcript changes in individual rainbow trout livers following an inflammatory stimulus. *Fish Shellfish Immunol*, 2007. 22(3): 157–171.
41. Lee, J.K., et al., Human homologs of the *Xenopus oocyte* cortical granule lectin XL35. *Glycobiology*, 2001. 11(1): 65–73.
42. Watson, M.L., et al., Genomic organization of the selectin family of leukocyte adhesion molecules on human and mouse chromosome 1. *J Exp Med*, 1990. 172(1): 263–272.
43. Ganz, T., Defensins and host defense. *Science*, 1999. 286(5439): 420–421.
44. Tsuji, S., et al., Human intelectin is a novel soluble lectin that recognizes galactofuranose in carbohydrate chains of bacterial cell wall. *J Biol Chem*, 2001. 276(26): 23456–23463.
45. Suzuki, Y.A., K. Shin, and B. Lonnerdal, Molecular cloning and functional expression of a human intestinal lactoferrin receptor. *Biochemistry*, 2001. 40(51): 15771–15779.
46. Teraguchi, S., et al., Glycans of bovine lactoferrin function as receptors for the type 1 fimbrial lectin of *Escherichia coli*. *Infect Immun*, 1996. 64(3): 1075–1077.
47. Schaible, U.E., et al., Correction of the iron overload defect in beta-2-microglobulin knockout mice by lactoferrin abolishes their increased susceptibility to tuberculosis. *J Exp Med*, 2002. 196(11): 1507–1513.
48. Otto, B.R., A.M. Verweij-van Vught, and D.M. MacLaren, Blood substitutes and infection. *Nature*, 1992. 358(6381): 23–24.
49. Ward, P.P., et al., Restricted spatiotemporal expression of lactoferrin during murine embryonic development. *Endocrinology*, 1999. 140(4): 1852–1860.
50. Wrackmeyer, U., et al., Intelectin: A novel lipid raft-associated protein in the enterocyte brush border. *Biochemistry*, 2006. 45(30): 9188–9197.
51. Wali, A., et al., Identification of intelectin overexpression in malignant pleural mesothelioma by serial analysis of gene expression (SAGE). *Lung Cancer*, 2005. 48(1): 19–29.
52. Schaffler, A., et al., Genomic structure of human omentin, a new adipocytokine expressed in omental adipose tissue. *Biochim Biophys Acta*, 2005. 1732(1–3): 96–102.
53. Yang, R.Z., et al., Identification of omentin as a novel depot-specific adipokine in human adipose tissue: Possible role in modulating insulin action. *Am J Physiol Endocrinol Metab*, 2006. 290(6): E1253–1261.
54. de Souza Batista, C.M., et al., Omentin plasma levels and gene expression are decreased in obesity. *Diabetes*, 2007. 56(6): 1655–1661.
55. Kralisch, S., et al., Therapeutic perspectives of adipocytokines. *Expert Opin Pharmacother*, 2005. 6(6): 863–872.
56. Rosen, E.D. and B.M. Spiegelman, Adipocytes as regulators of energy balance and glucose homeostasis. *Nature*, 2006. 444(7121): 847–853.
57. Drickamer, K., Ca^{2+}-dependent carbohydrate-recognition domains in animal proteins. *Curr Opin Structural Biol*, 1993. 3: 393–400.
58. Xu, X. and R.F. Doolittle, Presence of a vertebrate fibrinogen-like sequence in an echinoderm. *Proc Natl Acad Sci USA*, 1990. 87(6): 2097–2101.
59. Doolittle, R.F., The structure and evolution of vertebrate fibrinogen. *Ann N Y Acad Sci*, 1983. 408: 13–27.
60. Kilpatrick, D.C., T. Fujita, and M. Matsushita, P35, an opsonic lectin of the ficolin family, in human blood from neonates, normal adults, and recurrent miscarriage patients. *Immunol Lett*, 1997. 67: 109–112.
61. Ohashi, T. and H.P. Erickson, Two oligomeric forms of plasma ficolin have differential lectin activity. *J Biol Chem*, 1997. 272: 14220–14226.
62. Matsushita, M., et al., A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin. *J Biol Chem*, 1996. 271: 2448–2454.

63. Lu, J., et al., Collectins and ficolins: Sugar pattern recognition molecules of the mammalian innate immune system. *Biochim Biophys Acta*, 2002. 1572(2–3): 387–400.
64. Dann, S.M. and L. Eckmann, Innate immune defenses in the intestinal tract. *Curr Opin Gastroenterol*, 2007. 23(2): 115–120.
65. Lam, Q.L. and L.W. Lu, Role of leptin in immunity. *Cell Mol Immunol*, 2007. 4(1): 1–13.
66. Thompson, J.D., D.G. Higgins, and T.J. Gibson, CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 1994. 22(22): 4673–4680.

31 F-Type Lectins: A New Family of Recognition Factors

Gerardo R. Vasta, Eric W. Odom, Mario A. Bianchet, L. Mario Amzel, Keiko Saito, and Hafiz Ahmed

CONTENTS

31.1 INTRODUCTION

Interactions between humoral or cell-associated lectins with the exposed glycans on the cell surface of potential pathogens are considered key components of the innate immune response of vertebrates and invertebrates [1–4]. Lectins are generally organized as oligomers of noncovalently bound subunits, each displaying a carbohydrate recognition domain (CRD) that binds to the sugar ligand, usually a nonreducing terminal monosaccharide or oligosaccharide. Although lectin–ligand interactions are relatively weak as compared with other immune recognition molecules, high avidity for the target is achieved when multiple CRDs interact with ligand simultaneously [5]. In soluble collectins [2] such as the mannose-binding lectin, the subunits possessing a single CRD associate to form a "bouquet"-like oligomer with all CRDs facing the potential target surface. However, in the "cruciform" organization of the conglutinin subunits, the single CRDs are arranged in a manner that can also cross-link multiple targets. A less frequent organizational plan is the presence of tandem CRDs encoded within a polypeptide, such as observed in the macrophage mannose receptor [6], immulectins [7], and tandem repeat-type galectins [8].

Several lectin families representing distinct signature sequence motifs and structural folds, such as C-type [9], peptidoglycan recognition proteins [10], ficolins [11], pentraxins [12], and most recently galectins [13], have been implicated in immune surveillance. The F-type lectins, fucose-binding lectins found in ectothermic vertebrates, protochordates, and invertebrates are the most recent lectin family to be identified, characterized by unique fucose- and calcium-binding sequence

motifs [14]. Some of its members have been proposed to mediate immune recognition [15,16]. Resolution of the structure of the fucose-binding lectin from European eel (*Anguilla anguilla* agglutinin; AAA) also revealed a novel lectin fold (the "F-type" fold) ([17]; see also Chapter 2). This novel lectin family comprises a large number of proteins present from prokaryotes to vertebrates, exhibiting single, double, or greater multiples of the F-type motif, either tandemly arrayed or in mosaic combinations with other domains, yielding subunits of variable sizes even within a single species [14–16]. This widespread, heterogeneous distribution of the F-type domain suggests an extensive structural–functional diversification of this lectin family. Interestingly, the F-type fold is also shared with several proteins of distinct functional properties, including the C-terminal domain of human blood coagulation factors V [18] and VIII [19], the C-terminal domain of bacterial sialidases [20], the N-terminal domain of a fungal galactose oxidase [21], a human ubiquitin ligase subunit [22], a domain of the single-strand DNA break repair complex [23], the b1 domain of neuropilin [24], and the yeast allantoicase [25]. The structural analogy of these seemingly unrelated proteins is indicative of the archaic origin of this lectin fold.

Although preliminary searches in the sequence databases suggest that AAA-like proteins possessing multiple CRDs are phylogenetically widespread, so far only a few have been characterized biochemically or structurally. The fucose-binding lectin (MsaFBP32) isolated from serum of the striped bass (*Morone saxatilis*) is composed of two tandem domains that exhibit the F-type sequence motif [26]. Among the tandem lectins, MsaFBP32 and other tandem binary homologs appear unique in that although their N-terminal domain shows close similarity to the fucose recognition domain of the eel agglutinin, their C-terminal domain exhibits changes that potentially could confer a distinct specificity for fucosylated ligands. The F-type lectin (SauFBP32) from the serum of the gilt head bream *Sparus aurata* has recently been purified and characterized [27]. The SauFBP32 strongly cross-reacts with the fucose-binding lectin from the sea bass (*Dicentrarchus labrax*), indicating that the latter may be member of this emerging family of lectins [27,28].

31.2 STRUCTURAL ASPECTS OF F-TYPE LECTINS

The F-type lectin fold, described for the first time in the AAA-α-L-fucose (α-Fuc) complex, consists of a β-jelly roll sandwich composed of two β-sheets of three- and five antiparallel β-strands (Figure 31.1A) (17; see also Chapter 2). However, similar to collectins, the native structure of the AAA is a homotrimer (Figure 31.1B).

31.2.1 Carbohydrate-Binding Site of the F-Type Lectin: Structural Basis of Protein–Carbohydrate Interactions

As the lectin was cocrystallized with its ligand, α-Fuc, the structure revealed not only the nature of the amino acids that mediate recognition of the carbohydrate ligand, but also the type of interactions that are established. Binding to α-Fuc is mediated by hydrogen bonds from a trio of basic side chains that emerge from a shallow pocket and van der Waals contacts with a unique disulfide bridge formed by contiguous cysteines. The lectin establishes interactions with the ring O5 and the equatorial 3- and axial 4-OH groups of the α-Fuc using the nitrogen atoms of three conserved residues: Nε of His52 and the guanidinium groups of Arg79 and Arg86 (Figure 31.1C). van der Waals interactions are established between the disulfide bridge formed between consecutive cysteine residues (Cys82 and Cys83) and the bond between ring atoms C1 and C2 of the monosaccharide, and the C6, which docks loosely in a hydrophobic pocket, stacking against the aromatic rings of two residues His27 and Phe45. These residues, together with the residues Leu23 and Tyr46, form the binding pocket. Although AAA preferentially binds α-Fuc, it also binds 3-*O*-methyl-D-galactose (also 3-*O*-methyl-D-fucose), a galactose derivative that has topologically equivalent hydroxyls (i.e., axial hydroxyl on C4) and a hydrophobic moiety similar to α-Fuc (Figure 31.1D).

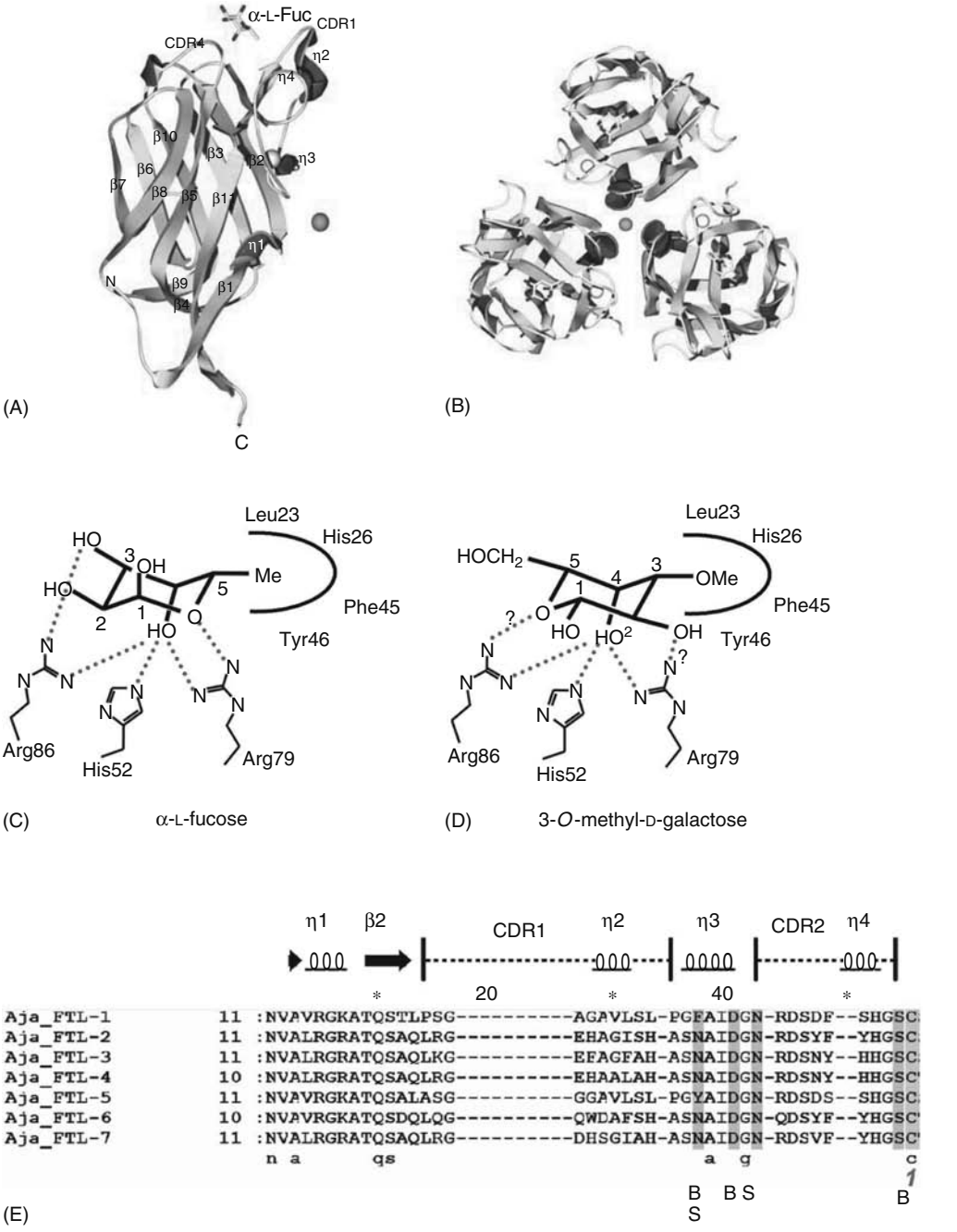


FIGURE 31.1 (See CD for color figure.) Crystallographic model of *Anguilla anguilla* agglutinin. (A) Ribbon diagram of AAA with helices colored magenta and β-sheets colored yellow. Bound α-Fuc is shown as a stick model above the lectin. Calcium is shown at right as a white sphere. Chlorine atom is illustrated by a green sphere. (B) Quaternary structure of AAA. The single chlorine ion marking the threefold axis of rotation is coordinated by a lysine (Lys16) from each subunit. (C) AAA binding to α-Fuc. (D) Docking of 3-*O*-methyl-D-galactose in AAA–α-Fuc complex. Additional interactions are possible with amino acid residues in the loops surrounding the binding pocket. (E) Amino acid sequence alignment of FTLD isoforms from the Japanese eel. The variability of critical residues in the binding pocket and surrounding loops in the multiple isoforms may contribute alternative interactions with terminal and subterminal sugar units and thus expand the range of diverse oligosaccharides recognition by the lectin isoform repertoire.

31.2.2 Extended Carbohydrate-Binding Site

Additional interactions of residues in the loops, designated CDR1–5 by analogy to immunoglobulin complementarity determining regions (CDRs), which surround the binding pocket of AAA with subterminal sugars of the blood group H type 1 (Fucα1–2 Galβ1–3GlcNAcβ1–3Galβ1–4Glc) and Le^a (Galβ1–3[Fucα1–4]GlcNAcβ1–3Galβ1–4Glc) trisaccharides allow fine discrimination of its cognate glycoconjugate ligands. Glu26 and His27 on CDR1 can recognize the equatorial hydroxyls (3-OH and 2-OH) in Gal and oxygen of the GlcNAc 2-N-Acetyl group in Le^a, or GlcNAc 6-OH and 4-OH groups in H. The OH group of Tyr46 in CDR2 can coordinate the glycosidic bond oxygen between Gal and GlcNAc moieties. Additionally, Asp81 and Arg79 in CDR4 may interact with the GlcNAc 6-OH group in Le^a and a water molecule observed in the structure of AAA–Fuc complex (W2) may bridge between Gal 4-OH group and Asp81 in H. In Le^x the 2-*N*-acetyl points toward the Fuc side of the trisaccharide impeding the adjustment of the Gal to remove clashes with the CDR1. It is the rigidity of the CDR1 loop that prevents binding of AAA to this trisaccharide. F-lectins other than the AAA contain a shorter CDR1, lacking five of the residues that form 3_{10}-helix h2. This erases the loop protruding feature observed in AAA, reducing or eliminating interactions between CDR1 and the putative oligosaccharide antennae and perhaps broadening the specificity of these fucolectins for Le^x oligosaccharides [17]. Organisms other than bacteria conserve the cysteines involved in the two interstrand disulfide bridges of AAA (50–146 and 104–128).

31.2.3 Diversity of Carbohydrate-Binding Specificity

The presence in single individuals of multiple F-type lectin isoforms, with amino acid substitutions at positions revealed by the structural analysis as key for ligand binding (Figure 31.1E) suggests potential diversity in recognition, a critical feature for proteins that mediate immune defense. Variability of critical residues in the binding pocket and surrounding loops in the multiple isoforms, as expressed in the Japanese eel [15], suggests that alternative interactions with terminal and subterminal sugar units may expand the range of diverse oligosaccharides recognized by the lectin isoform repertoire [17]. This is an intriguing observation for a protein proposed to mediate the recognition of potential microbial pathogens. The trimeric nature of the native structure of AAA (see Figure 31.1B) suggests that avidity is enhanced through cooperative ligand binding. Further, the threefold cyclic symmetry of the native AAA, very similar to that observed in collectins [29], would optimize the spacing and orientation of binding sites for recognition of glycoconjugates displayed on microbial surfaces. It is noteworthy that the distances between binding sites in the AAA trimer (26 Å) are about half those in mannose binding lectin (MBL) (54 Å), suggesting that they bind to differently arrayed surface glycans on the surface of microbes. Therefore, although F- and C-type lectins may recognize the same monosaccharide (MBL also binds fucose), their affinity for different microorganisms may be distinct, thereby expanding the immune recognition spectrum in those species that are endowed with both lectin types.

31.2.4 Role of Calcium Atom

In the 3D structure of F-type lectins, a calcium atom was identified as such based on coordination geometry and distance within a region rich in short 3_{10} helices. The calcium ion is heptacoordinated by oxygens contributed by both the peptide backbone and side chains from semicontinuous residues in a pentagonal bipyramidal geometry. In contrast to C-type lectins, in AAA calcium appears to play a role in structural stabilization, rather than participating in direct cation–saccharide interactions. F-lectins that require calcium for binding activity, such as Tachylectin-4 [16] and the Japanese eel lectin [15], form larger order assemblies, which may be susceptible to calcium chelation.

31.3 EVOLUTIONARY ASPECTS OF F-TYPE LECTINS

The F-type lectins were identified in a variety of taxa from prokaryotes to amphibians. However, the F-type lectin sequence motif appears to be absent from protozoa, fungi, nematodes, ascidians, and higher vertebrates such as reptiles, birds, and mammals [26].

31.3.1 F-Type Lectin Family: Domain Topology

Despite the fact that AAA and MsaFBP32 possess one and two F-type lectin domains (FTLDs), respectively, substantially diverse domain topologies, in some cases lineage-related, were identified based on sequence alignments of the F-type lectin motif [26] (Figure 31.2). For example, most teleost F-type lectins contain either duplicate or quadruplicate tandem domains, whereas in *Xenopus* spp., these lectins are composed of either triplicate or quintuple tandem F-type domains. Clearly, the F-type fold with its joined N- and C-terminals favors the formation of concatenated CRD topologies in numbers that appear lineage-related. These tandem arrays may yield mosaic proteins by including pentraxin (*Xenopus laevis*) or C-type domains (*Drosophila melanogaster* CG9095, malarial mosquito, and honey bee).

The F-type sequence motif is also present in both lophotrochozoan (i.e., molluscs and planaria) and ecdysozoan protostomes (i.e., horseshoe crabs and insects), invertebrate deuterostomes (i.e., echinoderm), elasmobranchs (i.e., skate), lobe- and ray-finned teleost fish, and amphibians (i.e., *X. laevis* and salamander) [3,14,26]. The F-type lectin sequence motif is not restricted to eukaryotes; homologs were identified in both Gram positive (i.e., *Streptococcus pneumoniae*) [30,31] and Gram negative (i.e., *Microbulbifer degradans*) [32] eubacteria. The absence of the F-type lectin sequence motif in protozoa, fungi, nematodes, ascidians, and higher vertebrates

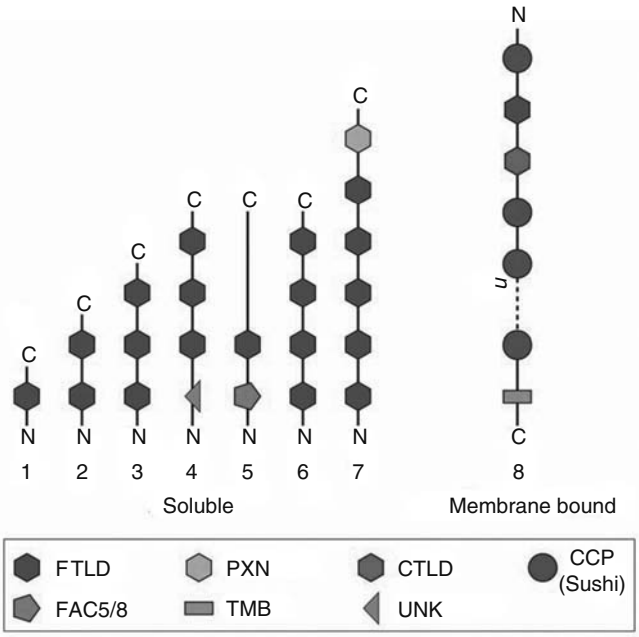


FIGURE 31.2 (See color insert following blank page 170. Also see CD for color figure.) Domain architecture of the F-type lectin family. FBPL, FBP-like lectin; UNK, unknown; FAC5/8, coagulation factors V and VIII; PXN, pentraxin; CCP, complement control protein; TMB, transmembrane. The subscript "*n*" indicates the extended number of CCP domains present in *furrowed*. Numbers indicate: 1, *A. anguilla* agglutinin (AAA); 2, MsaFBP32; 3, *X. tropicalis*; 4, *S. pneumoniae*; 5, *M. degradans*; 6, *O. mykiss*; 7, Xla-PXN-FBPL; 8, CG9095, *furrowed*.

suggests that it may have been selectively lost even in relatively closely related lineages. The paucity of bacteria possessing F-type CRDs suggests that it may have either been acquired through horizontal transfer from metazoans, or less likely, that most prokaryote lineages lost this CRD. Even the multiple duplicate tandem homologs present within modern teleost orders appear to be the product of independent duplications. The mottled phylogenetic distribution, diverse temporospatial expression, and varied domain architecture of the F-type family members point to a functionally plastic CRD, which has been specifically tailored in each lineage, and apparently lost its fitness value in some taxa. The absence of the F-type CRD in higher vertebrates is an evolutionary quandary that correlates with the appearance of cleidoic egg and the colonization of land by vertebrates [14,26].

31.3.2 Sequence Variants of the F-Type CRD

The 3D structure of AAA-α–Fuc complex enabled the thorough analysis of variations in the F-type sequence motif, focusing on positions relevant for ligand binding, metal binding, or tertiary structure. Sequence alignments revealed that most CRDs shared the α-Fuc-binding motif (HX24RXDX4(R/K) suggests that most F-type lectins bind this saccharide. Some CRDs, however, deviate from this motif, suggesting alternate specificities or even lack of saccharide-binding activity. An example is the *Drosophila* CG9095, where two residues of the sugar-binding triad are mutated to aliphatic residues unable to produce the H-bonds typical of the F-type CRD. This reduces the likelihood that the domain binds sugar, coinciding with the observation that CG9095's C-lectin domain is also unlikely to bind carbohydrate. The case of both domains of the binary DreV appears to be extreme since a large deletion is present in the CDR4 loop, which would most likely abrogate α-Fuc-binding. Another interesting pattern is that most duplicate tandem F-type lectins (i.e., striped bass, zebrafish, pufferfish, and stickleback) possess a unique combination of saccharide-binding motif and cystine bonds (N-terminal: CXHX24RGDCCXERXX16XX22C and C-terminal: CXHX24RDXXXERCX16CX22C). Specifically, one domain has lost the contiguous cystines that contact the saccharide ring while it gained the nested cystine. It is not clear what effect the loss of this cystine has on binding since the alkaline residue trio is conserved, but the predominance of this character only among duplicate tandem F-type lectins suggests that it has functional relevance. Future recombinant expression of individual domains should allow biochemical confirmation of these possible alternate specificities [14,26].

Unlike the carbohydrate-binding cleft of AAA, analysis of replacements of metal-binding residues is not clear because most interactions are provided by main-chain carboxyls and their conservation probably relates more to their configuration in the folded protein. However, differing degrees of conservation are evident for the three positions that interact through side-chains. Serine 49 is of special interest, since as a bidentate ligand, it is central to the pentagonal bipyramidal coordination geometry. In a few cases, this position is substituted with nonalcohol residues unable to produce the bidentate coordination (i.e., His, Asp, Pro, Gly, and Val). In most others, however, it is substituted by residues possessing oxygens (i.e., Asp, Gln, Glu, Thr, and Tyr) that likely are still able to form coordination bonds. In the former cases, it is possible that a water molecule may substitute in coordination, or there is a modification of the coordination geometry. A coordination number of six produces an octahedral geometry, which for Ca^{2+} is still permissible, as exemplified by several solved protein structures [33]. Nevertheless, there appears to be a strong selection for heptacoordination in F-type lectins to stabilize the large helix-rich loop. Similar support is given by the absolute conservation seen for Glu147, which also provides a side-chain ligand [14,26].

Like other proteins, sequence insertions or deletions (indels) are most permissible in loops where they have minimal effect on the core fold. Such is the case for the F-type lectin CDRs, where various indels reside. Interestingly, the CDR1 loop, which appears to interact with saccharides subterminal to α-Fuc, shows great divergence, suggesting that it might regulate binding to a wide diversity of glycoconjugates. Coincidentally, the intron that splits the exons coding for each domain of a binary CRD F-lectin from striped bass is localized close to the long irregular turn at the lower side of the barrel, which also presents length variability (in AAA: Glu123–Cys124), but it does not appear to be subject to junctional diversity during splicing [14,26].

31.3.3 Phylogeny of the F-Type Fold: Evolutionary Co-Option

Although the F-type lectin fold is unique among animal lectins, a structure-based search identified seven nonlectin proteins with negligible sequence similarity to F-type lectin motif, but shared the same jellyroll fold with AAA (Figure 31.3). These are the C domains of human blood coagulation factors V (FA58C) and VIII (PDB 1CZT) [34], the C-terminal domain of a bacterial sialidase (PDB 1EUT) [35], the NH_2-terminal domain of a fungal galactose oxidase (PDB 1GOF) [36,37], a subunit of the human APC10/DOC1 ubiquitin ligase (PDB 1XNA) [38], the N-terminal domain of the XRCC1 single-strand DNA repair complex (PDB 1JHJ) [39], and a yeast allantoicase (PDB 1SG3) [40]. Curiously, the FA58C domain of identical structure is present next to the F-type domain in the *M. bulbifer* protein previously described. In most of these proteins the AAA-like domain appears to mediate binding. However, allantoicase is the fist reported analog to exhibit intrinsic enzyme activity [14,17].

The F-lectin analogs in higher vertebrates may have originated from carbohydrate-binding domains, which later evolved new specificities. It also appears that all of these families share the placement of their binding site among the loops at the opening of the β-barrel. The FA58C domain, present as a tandem repeat at the C-terminal of part of multidomain circulating coagulation factors, is similar to the discoidin (DS) domain [41] named for a lectin from the slime mold [42,43] for which the binding site has not been determined. Interestingly, like the F-type domain, the FA58C domain is also frequently found as tandem replicates. However, its sequence is studded with hydrophobic residues that enable binding to cell surface phospholipids instead of saccharides [44]. In contrast, the b1 domain of neuropilin-1 receptor [45], a DS homolog, is devoid of the hydrophobic spikes and instead presents a polar cleft. Another set of DS homologs, the DS domain receptors [46],

FIGURE 31.3 (See color insert following blank page 170. Also see CD for color figure.) AAA shares structural similarity with diverse proteins. Green, AAA (PDB 1K12); Yellow, fungal galactose oxidase (PDB 1GOF); Gray, bacterial sialidase (PDB 1EUT); Blue, C2 domain of coagulation factor V (PDB 1CZT); Red, anaphase-promoting complex (PDB 1JHJ); Magenta, DNA single-strand break repair complex (PDB 1XNA). Inset: Overlay of yeast allontocaise (cyan) to the eel lectin.

which possess tyrosine kinase-signaling domains, are activated by binding to collagen [47]. It is evident from these examples that the animal DS domains have evolved to bind a diversity of ligands other than saccharides. Among the remaining structural analogs, the only jellyroll proteins of cytoplasmic localization [36] do not appear to bind glycoconjugates either. For XRCC1, the binding site for single-strand DNA breaks was mapped to a groove at the top of the barrel, but does not appear to involve interactions with the glycoside backbone. Although no binding activity has been attributed to APC10/DOC1 domain, the variable loops suggest it has a putative ligand. Curiously, the presence of a cation-binding loop is variable among these structural analogs since no cation has been detected in the structures FA58C, APC10/DOC1, or XRCC1 [14,17,26].

In summary, binary FBPLs have diversified through lineage-dependent gene duplications and speciation events producing a combination of paralogous relationships unique to teleosts. In contrast, in *Xenopus* spp. frogs not only are single domain F-type lectins expressed but also combinations of two, three, four, and chimeric protein-containing five tandem F-type domains adjacent to a pentraxin domain. Clearly, *Xenopus* spp. exhibits a greater diversity of F-type lectins than the pufferfish. Despite the diversity evident in this early tetrapod, however, no homologs are detectable in genomes of higher vertebrates, including reptile and avian representatives. This observation begs the question if this lectin family is uniquely restricted to invertebrates and cold-blooded vertebrates and was subsequently lost as such above the level of the amphibians. Specifically, the F-lectins may have either became truly extinct or have been co-opted into other biological roles, which in the course of evolution may have imposed structural constraints such as proposed for the C-1 and C-2 domains of the coagulation factors V and VIII [14,18,26].

31.4 CONCLUSIONS AND FUTURE DIRECTIONS

Experimental evidence and similarity search using available genomic databases have enabled greater insight into the diversity and complexity of lectin repertoires in invertebrates, protochordates, and ectothermic vertebrates. The identification in these taxa of members of the lectin families typical of mammals has resulted in the discovery of novel structural features, most likely revealing functional adaptations along the lineages, leading to the higher vertebrate taxa. Further, the identification of novel lectin families such as the F-type lectins underscores the fact that more research in nonmammalian model organisms will provide new information on both the structural, functional, and evolutionary aspects of lectin repertoires that may not be as obvious in mouse or humans. For example, elucidation of the detailed structural aspects of eel lectin–ligand binding and the potential of its multiple isoforms to achieve substantial diversity in oligosaccharide binding provides the structural basis for a tantalizing novel mechanism for generating diversity for nonself recognition in innate immunity, which resembles those operative through adaptive immunity in higher vertebrates. The ongoing genome, transcriptome, and proteome projects on additional model organisms representative of nonmammalian taxa will reveal not only the extent of their full lectin repertoires, but coupled to the structural analysis of selected components has the potential to uncover novel structural features on which a rigorous experimental assessment of their biological roles may be supported. In this regard, because of its external fertilization, rapidly developing transparent embryos, variety of established techniques for manipulation of gene expression, and a growing collection of mutations affecting early embryonic development that have been characterized and mapped, zebrafish may constitute an alternative model of choice for the elucidation of the biological roles of F-type lectins in embryogenesis and innate immunity.

ACKNOWLEDGMENTS

The work cited in this chapter was supported by grants MCB-00-77928 and IOB 0618409 from the National Science Foundation and R01 GM0705589-01 from the National Institutes of Health.

REFERENCES

1. McGreal, E. P., Martinez-Pomares, L., and Gordon, S. 2004. Divergent roles for C-type lectins expressed by cells of the innate immune system. *Mol Immunol* 41, 1109–1121.
2. Holmskov, U., Thiel, S., and Jensenius, J. C. 2003. Collections and ficolins: Humoral lectins of the innate immune defense. *Annu Rev Immunol* 21, 547–578.
3. Vasta, G. R., Ahmed, H., and Odom, E. W. 2004. Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. *Curr Opin Struct Biol* 14, 617–630.
4. Fujita, T., Matsushita, M., and Endo, Y. 2004. The lectin-complement pathway—its role in innate immunity and evolution. *Immunol Rev* 198, 185–202.
5. Weis, W. I. and Drickamer, K. 1996. Structural basis of lectin–carbohydrate recognition. *Annu Rev Biochem* 65, 441–473.
6. Taylor, M. E., Conary, J. T., Lennartz, M. R., Stahl, P. D., and Drickamer, K. 1990. Primary structure of the mannose receptor contains multiple motifs resembling carbohydrate-recognition domains. *J Biol Chem* 265, 12156–12162.
7. Yu, X. -Q., Zhu, Y. -F., Ma, C., Fabrick, J. A., and Kanost, M. R. 2002. Pattern recognition proteins in *Manduca sexta* plasma. *Insect Biochem Mol Biol* 32, 1287–1293.
8. Oda, Y., Herrmann, J., Gitt, M. A., Turck, C. W., Burlingame, A. L., Barondes, S. H., and Leffler, H. 1993. Soluble lactose-binding lectin from rat intestine with two different carbohydrate-binding domains in the same peptide chain. *J Biol Chem* 268, 5929–5939.
9. Weis, W. I., Taylor, M. E., and Drickamer, K. 1998. The C-type lectin superfamily in the immune system. *Immunol Rev* 163, 19–34.
10. Kim, M. S., Byun, M., and Oh, B. H. 2003. Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster. Nat Immun* 4, 787–793.
11. Endo, Y., Liu, Y., and Fujita, T. 2006. Structure and function of ficolins. *Adv Exp Med Biol* 586, 265–279.
12. Garlanda, C., Hirsch, E., Bozza, S., Salustri, A., De Acetis, M., Nota, R., Maccagno, A., Riva, F., Bottazzi, B., Peri, G., Doni, A., Vago, L., Botto, M., De Santis, R., Carminati, P., Siracusa, G., Altruda, F., Vecchi, A., Romani, L., and Mantovani, A. 2002. Non-redundant role of the long pentraxin PTX3 in anti-fungal innate immune response. *Nature* 420, 182–186.
13. Tasumi, S. and Vasta, G. R. 2007. A galectin of unique domain organization from hemocytes of the Eastern oyster (*Crassostrea virginica*) is a receptor for the protistan parasite *Perkinsus marinus. J Immunol* 179, 3086–3098.
14. Odom, E. W. 2004. F-Type Lectins: Biochemical, Genetic and Structural Characterization of a Novel Lectin Family in Lower Vertebrates. PhD thesis. University of Maryland College Park, MD.
15. Honda, S., Kashiwagi, M., Miyamoto, K., Takei, Y., and Hirose, S. 2000. Multiplicity, structures, and endocrine and exocrine natures of eel fucose-binding lectins. *J Biol Chem* 275, 33151–33157.
16. Saito, T., Hatada, M., Iwanaga, S., and Kawabata, S. 1997. A newly identified horseshoe crab lectin with binding specificity to O-antigen of bacterial lipopolysaccharides. *J Biol Chem* 272, 30703–30708.
17. Bianchet, M. A. and Odom, E. W., et al. 2002. A novel fucose recognition fold involved in innate immunity. *Nat Struct Biol* 9, 628–634.
18. Macedo-Ribeiro, S., Bode, W., Huber, R., Quinn-Allen, M. A., Kim, S. W., Ortel, T. L., Bourenkov, G. P., Bartunik, H. D., Stubbs, M. T., Kane, W. H., and Fuentes-Prior, P. 1999. Crystal structures of the membrane-binding C2 domain of human coagulation factor V. *Nature* 402, 434–439.
19. Pratt, K. P., Shen, B. W., Takeshima, K., Davie, E. W., Fujikawa, K., and Stoddard, B. L. 1999. Structure of the C2 domain of human factor VIII at 1.5 A resolution. *Nature* 402, 439–442.
20. Gaskell, A., Crennell, S., and Taylor, G. 1995. The three domains of a bacterial sialidase: A beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure (Camb.)* 3, 1197–1205.
21. Ito, N., Phillips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D., and Knowles, P. F. 1991. Novel thioether bond revealed by a 1.7 A crystal structure of galactose oxidase. *Nature* 350, 87–90.
22. Wendt, K. S., Vodermaier, H. C., Jacob, U., Gieffers, C., Gmachl, M., Peters, J. M., Huber, R., and Sondermann, P. 2001. Crystal structure of the APC10/DOC1 subunit of the human anaphase-promoting complex. *Nat Struct Biol* 8, 784–788.
23. Marintchev, A., Mullen, M. A., Maciejewski, M. W., Pan, B., Gryk, M. R., and Mullen, G. P. 1999. Solution structure of the single-strand break repair protein XRCC1 N-terminal domain. *Nat Struct Biol* 6, 884–893.
24. Lee, C. C., Kreusch, A., McMullan, D., Ng, K., and Spraggon, G. 2003. Crystal structure of the human neuropilin-1 b1 domain. *Structure (Camb.)* 11, 99–108.

25. Leulliot, N., Quevillon-Cheruel, S., Sorrel, I., Graille, M., Meyer, P., Liger, D., Blondeau, K., Janin, J., and van Tilbeurgh, H. 2004. Crystal structure of yeast allantoicase reveals a repeated jelly roll motif. *J Biol Chem* 279, 23447–23452.
26. Odom, E. W. and Vasta, G. R. 2006. Characterization of a binary tandem domain F-type lectin from striped bass (*Morone saxatilis*). *J Biol Chem* 281, 1698–1713.
27. Cammarata, M., Benenati, G., Odom, E. W., Salerno, G., Vizzini, A., Vasta, G. R., and Parrinello, N. 2007. Isolation and characterization of a fish F-type lectin from gilt head bream (*Sparus aurata*) serum. *Biochim Biophys Acta* 1770, 150–155.
28. Cammarata, M., Vazzana, M., Chinnici, C., and Parrinello, N. 2001. A serum fucolectin isolated and characterized from sea bass *Dicentrarchus labrax*. *Biochim Biophys Acta* 1528, 196–202.
29. Weis, W. I. and Drickamer, K. 1994. Trimeric structure of a c-type mannose-binding protein. *Structure* 2, 1227–1240.
30. Hoskins, J. and Alborn, W. E., Jr., et al. 2001. Genome of the bacterium *Streptococcus pneumoniae* strain r6. *J Bacteriol* 183(19), 5709–5717.
31. Tettelin, H. and Nelson, K. E., et al. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* 293(5529), 498–506.
32. Howard, M. B. and Ekborg, N. A., et al. 2003. Genomic analysis and initial characterization of the chitinolytic system of microbulbifer degradans strain 2–40. *J Bacteriol* 185(11), 3352–3360.
33. McPhalen, C. A. and Strynadka, N. C., et al. 1991. Calcium-binding sites in proteins: A structural perspective. *Adv Protein Chem* 42, 77–144.
34. Macedo-Ribeiro, S. and Bode, W., et al. 1999. Crystal structures of the membrane-binding c2 domain of human coagulation factor v. *Nature* 402, 434–439.
35. Gaskell, A. and Crennell, S., et al. 1995. The three domains of a bacterial sialidase: A beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure* 3, 1197–1205.
36. Ito, N. and Phillips, S. E., et al. 1991. Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. *Nature* 350, 87–90.
37. Firbank, S. J. and Rogers, M. S., et al. 2001. From the cover: Crystal structure of the precursor of galactose oxidase: An unusual self-processing enzyme. *Proc Natl Acad Sci USA* 98, 12932–12937.
38. Wendt, K. S. and Vodermaier, H. C., et al. 2001. Crystal structure of the apc10/doc1 subunit of the human anaphase-promoting complex. *Nat Struct Biol* 8, 784–788.
39. Marintchev, A. and Mullen, M. A., et al. 1999. Solution structure of the single-strand break repair protein xrcc1 n-terminal domain. *Nat Struct Biol* 6, 884–893.
40. Leulliot, N. and Quevillon-Cheruel1, S., et al. 2004. Crystal structure of yeast allantoicase reveals a repeated jelly-roll motif. *J Biol Chem* M401336200.
41. Baumgartner, S. and Hofmann, K., et al. 1998. The discoidin domain family revisited: New members from prokaryotes and a homology-based fold prediction. *Protein Sci* 7, 1626–1631.
42. Rosen, S. D. and Kafka, J. A., et al. 1973. Developmentally regulated, carbohydrate-binding protein in dictyostelium discoideum. *Proc Natl Acad Sci USA* 70, 2554–2557.
43. Poole, S. and Firtel, R. A., et al. 1981. Sequence and expression of the discoidin I gene family in *Dictyostelium discoideum*. *J Mol Biol* 153, 273–289.
44. Zwaal, R. F. and Comfurius, P., et al. 1998. Lipid–protein interactions in blood coagulation. *Biochim Biophys Acta* 1376, 433–453.
45. Lee, C. C. and Kreusch, A., et al. 2003. Crystal structure of the human neuropilin-1 b1 domain. *Structure* 11, 99–108.
46. Vogel, W. 1999. Discoidin domain receptors: Structural relations and functional implications. *FASEB J* 13, 77–82.
47. Vogel, W. and Gish, G. D., et al. 1997. The discoidin domain receptor tyrosine kinases are activated by collagen. *Mol Cell* 1, 13–23.

32 Biology of FREPs: Diversified Lectins with Fibrinogen-Related Domains from the Freshwater Snail *Biomphalaria glabrata*

Barbara A. Stout, Coen M. Adema, Si-Ming Zhang, and Eric S. Loker

CONTENTS

32.1 INTRODUCTION

The concept that invertebrates employ lectins to mediate nonself recognition in an immune context, proposed decades ago [1,2], holds true to date. Extensive experimental data in the literature and in sequence databases confirm that both invertebrate and vertebrate organisms use lectins as pattern recognition molecules (PRMs) to detect pathogen-associated molecular patterns (PAMPs) and activate defense responses [3–6]. In the absence of lymphocyte-derived defense components (T-cells, B-cells, antibodies, and T-cell receptors) [7], invertebrates may rely extensively on lectin-mediated pattern recognition of pathogens for their defense.

The internal defenses and the lectin-based, nonself recognition capabilities of gastropod molluscs have long been studied in the context of understanding why snails such as *Biomphalaria glabrata* support the larval development and transmission of digenetic trematodes such as the schistosome parasites that cause the widespread, debilitating human disease, schistosomiasis [8,9]. A multitude of studies have demonstrated the presence of various carbohydrate-specific agglutinating activities ascribed to lectins in the blood and tissues of snails [10–13]. Application of molecular techniques has made it feasible to characterize and identify many snail lectins at the level of nucleotide and amino acid sequence. Moreover, lectin sequences can be considered from the vantage point of comparative immunology and genomics. Targeted experiments and broad-scale gene discovery efforts have revealed lectins from snails including, but not limited to, galectin (see GenBank accession number EE722624), incilarin [14], C-type lectins including selectins [15–17], sialic acid binding lectin (SBL) [18], Group 1 and 2 molecules (G1M and G2M), and fibrinogen-related proteins (FREPs) [19–22].

Detailed studies aimed at characterization of FREPs as one particular group of lectins from *B. glabrata* were among the first to provide the insight that an ancestral function of fibrinogen-related domains (FReDs) may have been in nonself recognition of carbohydrates. Here we provide a brief overview of the role of FReDs in internal defense across the animal phyla, recount the discovery and basic structural features of FREPs, and discuss their role in defense responses against trematodes. We next consider the extensive diversity of these FReD-containing lectins from *B. glabrata*, explore the implications of this diversity for innate-type immunity across animal phylogeny, and conclude with a discussion of some remaining mysteries of FREP biology.

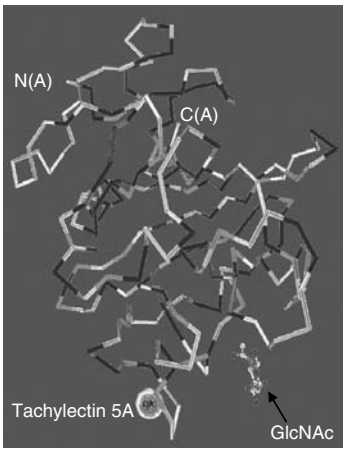
32.2 FIBRINOGEN-RELATED DOMAIN: A LECTIN DOMAIN BROADLY DISTRIBUTED AMONG ANIMALS, WITH AN INTERESTING HISTORY

Lectins have a wide variety of structural features that serve as carbohydrate-binding motifs [23]. For example, the carbohydrate-binding ability of some lectins, such as the ficolins, is ascribed to their FReD [24]. The crystal structure of the FReD-containing lectin tachylectin 5A, from the horseshoe crab *Tachypleus tridentatus* (Figure 32.1), confirms that binding to the carbohydrate *N*-acetyl-D-glucosamine (GlcNAc) can be accomplished via the FReD [25].

FReDs contain 24 canonical amino acids, mostly hydrophobic and including four cysteines [26]. Protein families containing FReDs include the fibrinogens, tenascins, microfibril-associated proteins, ficolins, tachylectins, and FREPs, the latter three being relevant to the topics of internal defense and a role in self/nonself recognition. As we and others have noted previously [21,25,27], the primordial function of the FReD domain in animals appears to have been in innate immunity and it continues to function in this capacity in both extant invertebrates and vertebrates (see Figure 32.2 for a summary).

In addition to innate immunity, FReD-containing proteins are also involved in development in invertebrates [29,30]. Based on the observation of a FReD-containing protein in the sponge *Suberites domuncula* that recognizes (1→3)-beta-D-glucan, which was interpreted to provide a molecular mechanism for recognizing fungi [31], it is likely that their role in innate immunity preceded their role in development.

The role of the best known of all FReD-containing proteins, the α, β, and γ fibrinogens involved in blood coagulation in vertebrates [25,32], would seem to represent a fascinating co-opting of this domain for quite a different function. As the vertebrate hemostasis system is complex, sufficiently so to have lead some to invoke divine intelligence in its design (see discussion by Aird [33]), this represents an example of the evolution of a highly regulated and multifactorial system.



(A)

```
TL-5A  PTDCADILLNGYRSSGGYRIWPKSWMTVGTLNVYCDMETDGGGWTV
       PT+C D                          L V CD  TDGGGWT+
FREP4  PTSCRDVISTEDRVVVTLA---------SGLEVMCDTTTDGGGWTI

TL-5A  IQRRGNYGNPSDYFYKPWKNYKLGFGNIEK-DFWLGNDRIFALTNQ
        QRR   G  S  FY+ WK Y+ GFG+    +F+LGN  IF  LT+
FREP4  FQRRFN-G--SIDFYRDWKEYRDGFGDYNIGEFYLGNENIFNLTSS

TL-5A  RNYMIRFDLKDKENDTRYAIYQDFWIENEDYLYCLHIGNYSGDAGN
       R Y +RFDL + EN   +A Y DF + +E+  Y L IG+YSG+AG+
FREP4  RKYELRFDL-EYENKKYFAHYSDFKLLDENNKYKLIIGSYSGNAGD

TL-5A  SFGRHNGHNFSTIDKDHD-THETHCAQTYKGGWWYDRCHESNLNGL
       S  RH    F+T DKD+D +   +CA   +G WWY  C + NLNG
FREP4  SMRRHVNKFFTTFDKDNDDSPNDNCAIIRRGAWWYQNCADVNLNGN

TL-5A  YLNGEHNSYADGIEW---RAWKGYHYSLPQVEMKIRPVEFNIIGN
       +  GE     DG+ W     W+   +S    E+KIR +
FREP4  WGRGE----PDGVFWDNITVWESVSFS----EIKIREI-------
```

(B)

FIGURE 32.1 (See CD for color figure.) (A) Lectin function of a FReD. This three-dimensional crystal structure shows the calcium-dependent binding of GlcNAc by the FReD of tachylectin 5A, a nonself recognition lectin from the invertebrate *Tachypleus tridentatus* (Japanese horseshoe crab) (17982, Molecular Modeling DataBase MMDB.) Superimposed on the white tube-style rendering of tachylectin 5A (TL-5A) are identical black and similar gray amino acid residues of the FReD of FREP4 of *B. glabrata* (AY012701) according to the alignment shown in B. N(A) and C(A) identify the terminal amino acids, GlcNAc: *N*-acetyl-D-glucosamine. The gray sphere labeled Ca denotes the calcium ion. (Adapted from Kairies, N., et al., *Proc. Natl Acad. Sci. USA*, 98(24), 13519, 2001.) (B) FREP4 is one member of an extensive gene families of FreD-containing calcium-dependent lectins from the hemolymph of *B. glabrata*. The sequence similarities shown concord with the notion that both FreDs function as lectins. Gaps in the alignment may reflect phylogenetic distance of the organisms from which the sequences were derived, and perhaps also differences in carbohydrate specificity. No crystal structure is available for FREP4. Brackets indicate cysteine residues that form intrachain disulfide bonds, the calcium-binding site is indicated by a dashed box and the carbohydrate-binding region is underlined.

32.3 FREPs FROM *B. GLABRATA* AS AN EXAMPLE OF FReDs: THEIR DISCOVERY AND BASIC STRUCTURAL FEATURES

Biomphalaria glabrata has attracted considerable attention because of its role as an intermediate host for the digenetic trematode *Schistosoma mansoni*, one of six species of this genus implicated in causing human schistosomiasis, a disease that still afflicts 200 million people, mostly in tropical Africa and South America [8,9]. As part of ongoing investigations to understand the internal defense system of *B. glabrata* and how it resists schistosome infection, another trematode, *Echinostoma paraensei*, which also infects *B. glabrata*, has frequently been used to probe the snail's defense responses as well. Studies employing *E. paraensei* have proven useful because this trematode species often provokes different and more conspicuous defense responses from *B. glabrata* than does *S. mansoni* [19,34].

It was noted that addition of secretory–excretory products (SEP) derived from cultured sporocysts of *E. paraensei* to the cell-free hemolymph (plasma) of snails previously exposed to this parasite provoked a conspicuous precipitation response in the plasma, a response not provoked in plasma from unexposed control snails. The precipitate was highly enriched in a 65 kDa band of snail origin (Figure 32.3). This precipitation reaction is inhibitable by the Ca^{2+} chelators EDTA and EGTA and the sugar L-fucose, prompting the conclusion that the polypeptides comprising this band are Ca^{2+}-dependent lectins [21], typical of other FReD lectins. The 65 kDa band was subsequently collected,

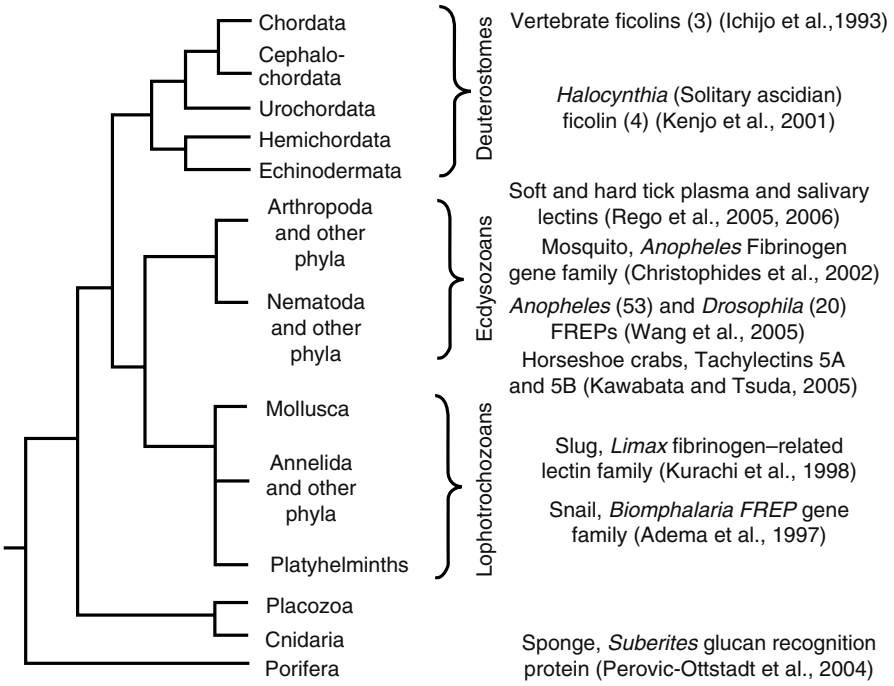


FIGURE 32.2 Studies implicating molecules with FReDs functioning in internal defense, represented next to a simplified overview of animal phylogeny, to make the general point that a defense function for FReDs has been reported from prebilaterians, and from both deuterostome and protostome lineages, in the latter including both ecdysozoans and lophotrochozoans. In parentheses is the number of gene variants discovered so far. (From Halanych, K. M., *Annu Rev Ecol Evol Syst*, 35, 229, 2004.)

subjected to Edman degradation, and shown to contain peptides with homology to fibrinogen [20,34]. These peptide sequences were used to make degenerate primers that were used to amplify more complete sequences from genomic DNA of *B. glabrata*. These amplicons were used to probe and isolate complete coding sequences from a *B. glabrata* cDNA library. Surprisingly, the sequences obtained revealed the presence of both a C-terminal FReD and an N-terminal IgSF domain. These molecules were christened fibrinogen-related proteins, or FREPs, at the time [21]. The original FREP comprising the 65 kDa band has since been identified as FREP4.

Since then, a series of studies [20,22,35–39] have revealed additional FREPs, now arrayed in 13 subfamilies [37]. Initially subfamilies were numbered in the order of recovery of different FREPs; more recently subfamily membership was defined by an 84% level of identity in nucleotide sequence of available upstream IgSF domains. Figure 32.4 provides an overview of structure of the better known FREP subfamilies. To our knowledge, FREPs are unique in nature in having one or two IgSF domains in conjunction with a FReD domain. Polymerase chain reaction (PCR) assays have yielded FreD sequences from several representatives of three families of gastropod molluscs: the Planorbidae (*Helisoma trivolvis*, *Bulinus truncatus*), Lymnaeidae (*Lymnaea stagnalis*), and Helicidae (*Helix aspersa*). Based on extensive sequence similarities, we assume that these sequences are parts of FREPs [21].

Interestingly, as is true of all other FReD-containing proteins that we know of, the FReD domain of FREPs is found at the C-terminus. This position is likely to be necessary for carbohydrates to sterically interact with the Ca^{2+} ion as well as the FReD. The FReD within FREPs displays regions of homology to areas in the human serum lectin P35 (ficolin 2) that contribute to Ca^{2+} -dependent

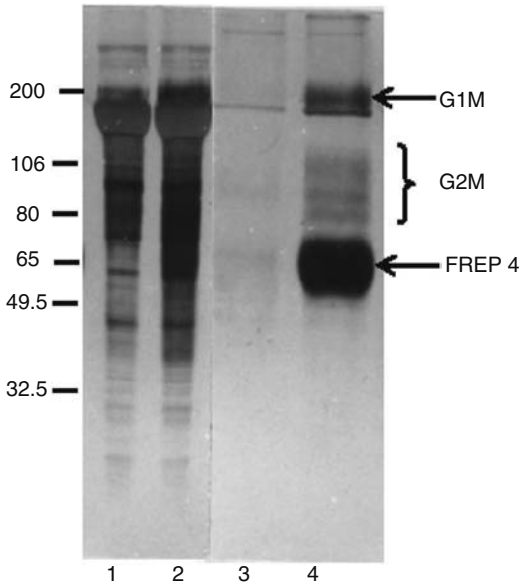


FIGURE 32.3 SDS-PAGE gel of cell-free hemolymph (plasma) polypeptides of *B. glabrata*. Lane 1 shows the typical plasma profile of a control snail not exposed to trematode infection. Note the prominent bulge comprised of hemoglobin, the dominant plasma polypeptide, shown at about 180 kDa. Lane 2 is plasma from a snail exposed 4 days prior to infection with *E. paraensei*. Note the overall denser appearance of the lane, with broad regions of heavier staining, as at ~200 kDa, and from just below the hemoglobin band to ~40 kDa. Lane 3 shows the small amount of material precipitated from the plasma of control snails following the addition of secretory–excretory products from *in vitro* cultured sporocysts of *E. paraensei*. Lane 4 shows what is precipitated from the plasma of echinostome-infected snails by a similar treatment. Note the appearance of material designated G1M and G2M, which has also been identified in other studies as trematode-responsive components (see text) and the broad band centered at 65 kDa, which was analyzed directly using both Edman degradation and mass spectroscopy and found to be comprised of FREP4. See Adema et al. for additional details. (From Adema, C. M., Hertel, L. A., Loker, E. S., Chapman and Hall, New York, 1997; Adema, C. M., et al., *Proc. Natl Acad. Sci. USA*, 94(16), 8691, 1997.)

binding of carbohydrate, consistent with the behavior of Ca^{2+} lectins [21,40,41]. Figure 32.1 shows evident similarity between the structure of horseshoe crab tachylectin and the FReD domain of FREP4. In general, the FReD domains in FREPs are more conserved than IgSF domains as supported by both sequence data and Southern blot analysis [35,36,39]. In addition, no obvious cleavage sites analogous to those of fibrinogen involved in hemostasis are observed in the FReDs in *B. glabrata* FREPs [21].

With respect to the IgSF domains, the two cysteine residues characteristic of such domains are spaced 80 amino acids apart [21] and the predicted structure of the IgSF domain of FREP2 and FREP4 has multiple β strands characteristic of IgSF members [35]. Eighteen of 20 residues conserved within IgSF domains of invertebrates are also present including two conserved cysteines that likely participate in the formation of a disulfide bond to form an intrachain loop [21]. The amino acid identity between IgSF regions of the different FREP subfamilies ranges from 42% to 81% [35–37]. For the subfamilies for which the IgSF structure is known, subfamilies 2 and 4 contain 1 IgSF domain, whereas FREP subfamilies 3, 7, 12, and 13 have two tandemly situated IgSF domains upstream of the FReD. For those with tandem IgSF domains, the N-terminal domain is designated IgSF1 and is followed by the IgSF2 domain, with a small connecting region between them. Neither the sequence of IgSF1 or IgSF2 easily yields to classification as C- or V-type IgSF domains [36,42].

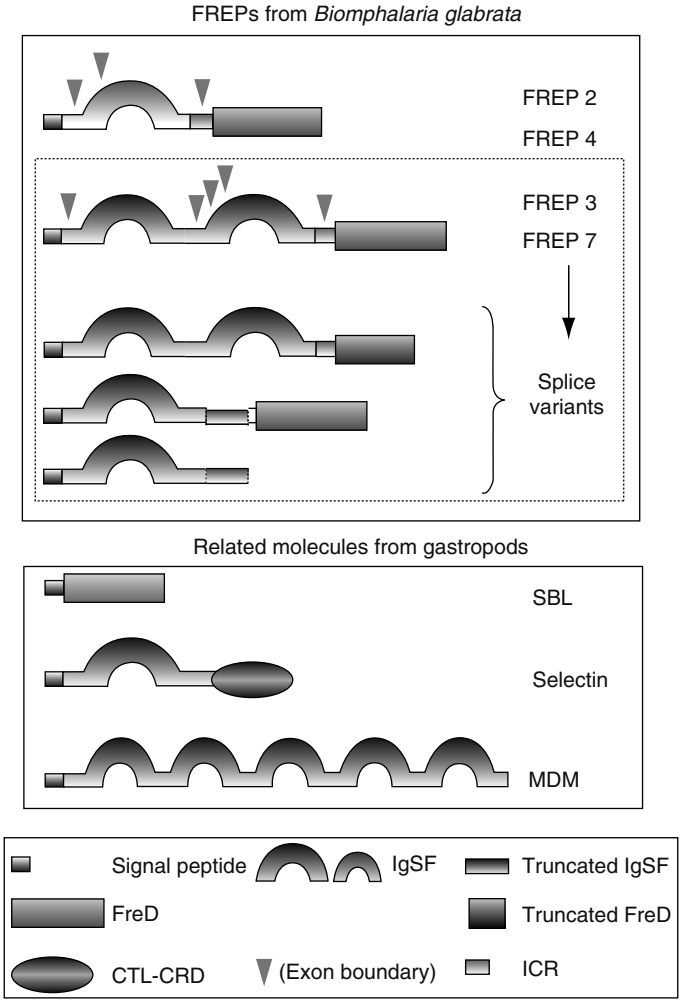


FIGURE 32.4 (See color insert following blank page 170. Also see CD for color figure.) Domain structures of FREPs and related molecules from *B. glabrata* and other gastropods. The two types of FREPs that have been identified each contain a downstream FreD in juxtaposition with upstream IgSF sequence, either one IgSF domain (as observed from FREP2 and FREP4) or two domains arranged in tandem (as in FREP3 and FREP7). An interceding region (ICR) separates the IgSF and FReD domains. The ICR shows no similarity to known sequences. Characterization of the full-length sequence of four different FREP genes showed that the FreD is encoded in a single exon. The three IgSF domains all differ in intron–exon organization. Arrowheads indicate the boundaries of sequences that are encoded by separate exons. Delineated by DNA sequence identity of <84%, gene subfamilies are recognized within both types of FREPs. Alternative splicing adds to diversity of FREPs with tandemly arranged IgSF domains. The resulting splice variants either have a truncated FreD, lack part of the second IgSF domain (resembling the one IgSF conformation), or consist of the leading IgSF with a truncated second IgSF domain only. SBL from the slug *L. flavus* consists of a single FreD. (From Kurachi, S., et al., *Eur. J. Biochem*., 254(2), 217, 1998.) The structure of selectin from *B. glabrata* is intriguingly similar to that of FREPs as it is comprised of an IgSF domain upstream of a C-type lectin carbohydrate recognition domain (CTL-CRD). (From Duclermortier, P., et al., *Parasitol. Res*., 85(6), 481, 1999; Guillou, F., et al., *Comp. Biochem. Physiol. B Biochem. Mol. Biol.,* 138(2), 175, 2004.) MDM from *Lymnaea stagnalis* consists of five IgSF domains and is thought to mediate nonself recognition. (From Hoek, R. M., et al., *Eur. J. Immunol.,* 26(4), 939, 1996.) Note that the IgSF domains of MDM are shorter than those of FREPs (~60 vs. ~80 amino acids enclosed in the intra-chain loop that characterizes IgSF sequences, respectively).

The region between the IgSF and FReD, called the interceding region (ICR), is variable in length and sequence amongst the various FREPs. Lysine (K), isoleucine (I), and glutamic acid (E) amino acid residues constitute a relatively high portion of the ICR from several FREPs, the significance of which is currently unknown. Furthermore, there is no detectable similarity between the amino acid sequence of the ICR of FREPs and other proteins [35,36].

In combining IgSF and FreD sequences, the structure of FREPs contains two types of domains, each of which could potentially bind (foreign) ligands. One idea is that one end of the FREP binds to a nonself entity and the other end binds to a receptor on a hemocyte surface, thus forming a bridge and activating the hemocyte, facilitating phagocytosis, release of toxic oxygen radicals, or an encapsulation reaction. If this model is correct, is nonself recognition achieved by the FReD domain or by IgSF domains? A consideration of the structure of several related IgSF and lectin sequences known from gastropods provides some insights into which domain may be the "business end" that imparts lectin activity to FREPs (also see Figure 32.3). Molluscan defense molecule (MDM) from the pond snail *L. stagnalis* [43] provides an indication that gastropod IgSF domains can possibly mediate nonself recognition. MDM consists of five tandemly arranged IgSF domains that have sequence similarity to hemolin (four IgSF domains in tandem), an antibacterial opsonin described from insects [44,45]. The repeated IgSF domains of soluble and membrane-bound hemolin can interact to mediate binding between ligands and cells [46]. However, the IgSF sequences of FREPs may not have similar functions because they differ considerably in both sequence and size from the IgSF domains of hemolin and MDM. Additional work is needed to characterize functional aspects of the IgSF domain of FREPs. By contrast, a clear indication for the use of FreD as lectins by gastropods is provided by SBLs from the slug *Limax flavus* [47]. Biochemical and sequence analysis revealed that these lectins consist only of a single FReD. Although some sequence differences are evident, the essential sequence and structural features of FReDs as identified by Doolittle [26] are conserved in both SBLs and the FReD domains of FREPs. This makes it highly likely that the FreD domain conveys lectin activity to FREPs, a conclusion supported by several recent reports that FreDs are functional domains of many lectins from both vertebrates and invertebrates (Figure 32.2) [21,31,47–52]. Finally, the structure of another lectin from *B. glabrata* termed selectin [15,53] is tantalizingly similar to that of FREPs. Both molecules are comprised of upstream IgSF domains (of different sequence but comparable size) and downstream lectin domains, either a FreD or a C-type lectin carbohydrate recognition domain (CTL-CRD). It is tempting to hypothesize that some snail lectins have a modular organization, where different lectin domains (FreD vs. CTL-CRD) provide diversity in recognition capability of particular polysaccharides, which can be processed by similar receptors (pathways) or cell types due to related N-terminal sequences (IgSF domains).

Such a model may help to explain why none of the FREPs we have isolated have membrane spanning motifs that are compatible with FREPs being displayed on hemocyte surfaces. Although our assumption is that FREPs must act in concert with hemocytes to be effective, they may have a direct toxic effect on trematode sporocysts, potentially offering an explanation for the degeneration of sporocysts without evident involvement of hemocytes in the *Echinostoma caproni–B. glabrata* model system [54].

32.4 ROLE OF FREPs AND RELATED MOLECULES IN DEFENSE AGAINST PATHOGENS

32.4.1 FREPs AND TREMATODES

The vast majority of the ~18,000 nominal species of digenetic trematodes undergo obligatory larval development in molluscs, often castrating and otherwise severely compromising their hosts in the process. Thus it seems reasonable to expect that snails would mount vigorous defense responses against trematodes as they attempt to establish infections. Are FREPs a part of the antitrematode defense response? There are several reasons to suggest this might be the case, as we outline below.

FREPs were first identified in the plasma of snails infected with *E. paraensei*, and FREPs were shown to be reactive with soluble antigens of this trematode [20]. They have also been shown to bind to the surfaces of *E. paraensei* sporocysts and miracidia [20,22]. Other studies have since shown that at least four FREPs are upregulated following exposure to *E. paraensei*; FREPs 2, 3, 4, and 7 with FREPs 2 and 4 first appearing in increased abundance at 1 or 2 days postexposure (dpe), respectively, and continuing until at least 16 dpe [55,56]. The kinetics of FREP production following exposure to *E. paraensei* are mirrored by an increase in circulating hemocyte numbers and it has been shown that FREPs are produced by hemocytes [21,55].

Even though a strong upregulation of FREPs is noted suggesting *B. glabrata* can recognize the presence of *E. paraensei*, and FREPs can bind to parasite surfaces, *B. glabrata* nonetheless typically succumbs to infection, suggesting that FREPs are inconsequential in defense against this parasite. However, *E. paraensei* is well known for its ability to interfere strongly with hemocyte spreading and formation of hemocyte capsules and can increase the susceptibility to infection by other trematodes [19,57–61]. Consequently, even though soluble recognition molecules like FREPs are present, they may be rendered irrelevant if hemocytes as the main effector cells in snails have been compromised. However, the ability of FREPs to precipitate parasite SEPs may have the effect of protecting at least some hemocytes such that the immune capacity of the snail is not completely compromised.

With respect to the response of *B. glabrata* to infection with the medically important trematode *S. mansoni*, quantitative PCR reveals that differential FREP expression is elicited by *B. glabrata* strains susceptible (M-line strain) or resistant (BS-90 strain) to infection by this parasite. When a susceptible strain is infected with *S. mansoni*, an upregulated FREP response does not occur. However, when snails from the BS-90 strain are exposed to *S. mansoni*, there is a 40-fold increase as early as 1 to 2 days post exposure in FREP2 expression which remains elevated for at least 8 days, suggesting that FREP2 may play a role in the resistance of the BS-90 strain to *S. mansoni*. A four-fold elevation of FREP4 expression also occurs [56]. The kinetics of FREP2 and FREP4 induction match the time course of *S. mansoni* sporocyst killing in resistant strains of *B. glabrata* [62,63].

It has been speculated that unlike *E. paraensei*, *S. mansoni* essentially avoids detection in compatible snails [19], possibly due to molecular mimicry of host tissues [64–68]. Mimicry could be due to either a parasite-encoded matching of host molecules or acquisition of host molecules. In either case, because of a lack of detection, a possible consequence is that a FREP response is not orchestrated. In resistant snails, recognition is somehow achieved and a FREP response is recorded. One of the biggest gaps in our understanding of FREP biology is if and how circulating FREPs may collaborate with hemocytes or other components of the snail's internal defense system to effect trematode killing.

In general, exposure to *E. paraensei* elicits a greater response in terms of both hemocyte and FREP production than does exposure to *S. mansoni*, and as noted above, these two responses are likely causally linked. Although this difference may be due to fundamentally different survival strategies of the two parasites [19], another possible explanation is that *E. paraensei* undergoes the initial stages of its larval development in the snail's ventricle, in close proximity to the amebocyte-producing organ (APO) where hemocytes are produced (see Figure 32.5). This by itself may trigger a stronger response from the APO, accounting for the differences between the two trematodes with respect to provoking FREP production.

32.4.2 FREPs and Other Potential Pathogens

Given the importance of digenetic trematodes to molluscan biology, it is possible that FREPs represent a class of molecules responsive only to these formidable pathogens. Conversely, given the role of FReD-containing molecules in defense in other organisms, it seems logical to expect that FREPs are also responsive to other pathogens as well. As part of an overall effort to better define FREP functions, this topic is currently under investigation in our laboratory and preliminary studies suggest

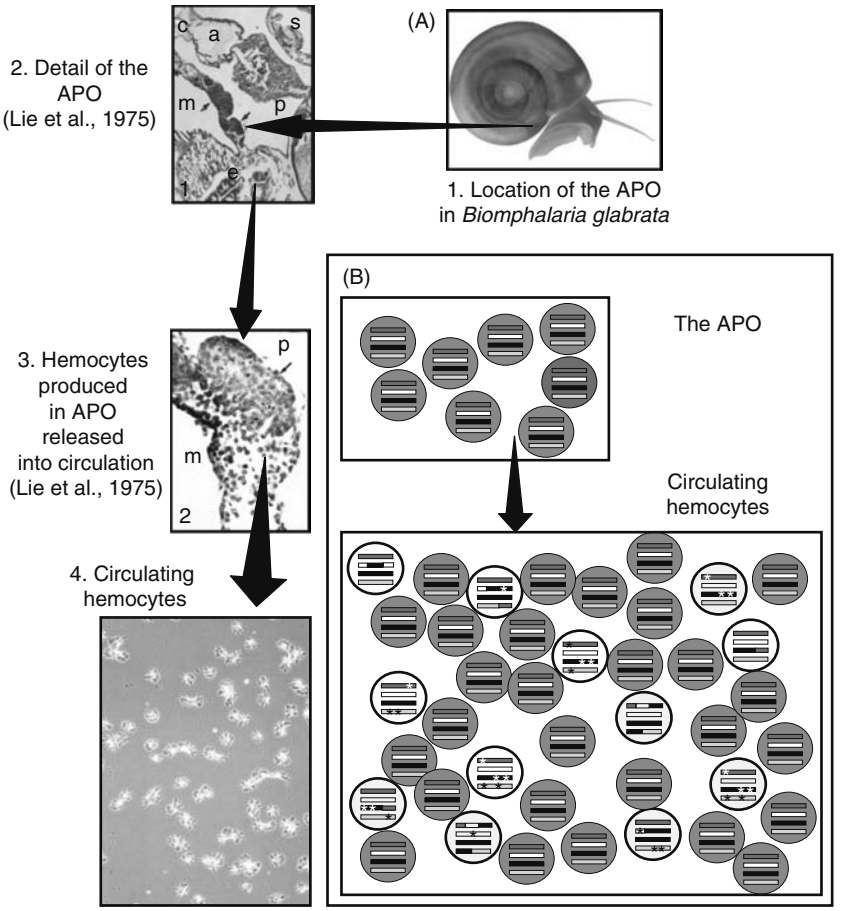


FIGURE 32.5 (See CD for color figure.) (A) This sequence of four figures shows (1) the general anatomical location of the hemopoietic organ of *B. glabrata*, between the pericardium and the posterior epithelium of the mantle cavity. This organ is called the APO, and is where amebocytes, now more commonly referred to as hemocytes, the circulating defense cells of snails, are produced; (2) a detail of the general region of the APO in a snail infected with echinostome trematodes, showing the APO (arrow) in a hypertrophied state; (3) further detail showing hemocytes being released from the APO into the snail's circulation; and (4) hemocytes taken from the general circulation of *B. glabrata*. (Photographs 2 and 3 reproduced from Lie, K. J., Heyneman, D., and Yau, P., *J. Parasitol*., 61(3), 574, 1975. With permission.) (B) A schematic view of one hypothesis for generation of diversified FREPs starting with an initially genetically uniform population of stem cells in the APO that contain the FREP source sequences. These cells undergo mitosis to produce hemocytes that in the process of differentiating may undergo mutational and/or recombinatorial events in their FREP-encoding genes. Most of the hemocytes produced express the original FREP source sequences but many express diversified FREPs. Point mutations are represented by asterices. Recombination events are represented by mosaic blocks.

that certain species of bacteria elicit upregulation from a subset of FREPs depending on the particular strain of *B. glabrata* exposed to the bacteria.

32.4.3 Other Functions for FREPs

Our results lead us to believe that the primary function of FREPs is to protect the snail host from pathogens, based primarily upon the upregulation of FREPs following infections. However, this

does not mean that FREPs may not function in other aspects of snail biology. For example, proteins such as Toll and DSCAM, originally identified as regulating development, were later discovered to also play a role the innate immune system of invertebrates [69–73]. It is possible that FREPs function in other capacities in physiological states that occur along with infections, such as stress or wound repair. It is important to retain an open mind regarding other possible functions for FREPs in *B. glabrata*.

32.4.4 Other Molecules with Lectin Activity That Accompany FREP Induction

Prior to the discovery of FREPs and the association of FREP4 with the 65 kDa plasma band, other conspicuous components in *B. glabrata* plasma designated as G1M and G2M (Figure 32.3), had for several years been under observation, primarily because they were upregulated following infection of *B. glabrata* with *E. paraensei* [34,74–76]. and to a lesser extent with *S. mansoni* [77]. G1M (group 1 molecules) range in molecular weight from 150 to 220 kDa and G2M (group 2 molecules) are 75–130 kDa [76]. Components of G1M vary between *E. paraensei*-infected snail strains susceptible or resistant to *S. mansoni* infection [75]. G1M/G2M also bind to trematode sporocyst SEPs *in vitro*, have lectin activity, can be harvested by affinity chromatography using monosaccharide-treated beads [75,77,78], and were implicated in having opsonic activity in phagocytosis assays [76]. They also share with the 65 kDa FREP4 band the property of migrating on SDS-PAGE gels as broad, diffuse bands.

It is possible that some proteins in the G1M and G2M fractions are FREPs because, for example, the expected molecular weights of FREPs 3 and 7 when glycosylated are between 90 and 100 kDa [36], putting them within the G2M molecular weight range.

Preliminary evidence from mass spectrometry of plasma proteins in the G1M/G2M molecular weight ranges reveal the presence of IgSF domains, as does the use of newly developed antibodies to the IgSF portions of recombinant FREPs (Zhang, unpublished results). However, FreD domains were not obvious from these assays; this has lead us to consider other possibilities, including that these G1M and G2M molecules may consist solely of IgSF domains, perhaps in analogy to the multi-IgSF domain containing molecule identified as MDM from the snail *L. stagnalis* [43].

32.5 SURPRISING EXTENT OF FREP DIVERSITY

Southern hybridization studies indicate that many (at least 24 bands identified) *B. glabrata* genes encode FReDs, and when probes for specific FREP subfamilies are used, multiple bands are again recovered (approximately between one and eight bands among the six FREP subfamilies probed). The Southern hybridization results are broadly consistent with our PCR-based approaches using both cDNA and genomic templates in revealing FREPs to be encoded by a complex gene family. In common with other invertebrates (Figure 32.2), *B. glabrata* has an inherently diverse set of FReD-encoding genes, suggestive of an important role in snail biology.

In addition to an abundance of FReD encoding genes, other surprising sources of diversity have been noted. The 65 kDa plasma band of *B. glabrata*, which is known to be comprised of FREP4, migrates as a broad, diffuse band, as do other trematode-responsive bands such as G2M, which may also be comprised of FREPs, as noted above. The breadth of the band centered at 65 kDa was suggestive of inherent diversity. Treatment of material comprising this band with N-glycosidase F to remove N-linked carbohydrate side chains surprisingly did not alter the band's breadth, though it did provide an expected shift to a lower molecular weight. This pattern was consistent with the idea that the FREP4 band is comprised of a variety of related polypeptides [21].

This finding prompted us to explore further the phenomenon of FREP diversification, particularly because expressed sequence tag (EST) studies of this and other FREPs revealed a surprising level of sequence diversity [35,37]. This was especially true of FREP3 [38]. It was subsequently

shown that in just one region of one FREP gene, the IgSF1 region of FREP3, using PCR amplification with high-fidelity polymerases and primers specific for the IgSF1 domain of the FREP3 gene subfamily, 45 DNA sequences were recovered from one snail, and 37 sequences were recovered from a second snail, with only one of the sequences held in common between the two individuals. A total of 314 unique DNA sequences have been recovered from 22 snails, which would encode 204 unique amino acid sequences. Similar levels of FREP sequence diversity were evident from EST sequence data obtained from *B. glabrata* derived independently by Mitta and coworkers in France [53]. Southern analysis indicated that only two to five FREP3 loci existed in the *B. glabrata* genome, a paradox that was subsequently resolved when a computational analysis revealed that the recovered diversity within IgSF1 region of FREP3 from a single snail could be explained as a consequence of recombinatorial events or point mutations (or both) in as few as two to five "source sequences," which can be conceptualized as alleles. Our hypothesis is that a limited number of source sequences (no more than 10) are carried among the two to five FREP3 loci in the genome, and these are diversified among differentiated cells, with hemocytes being the most likely source of the diversified sequences (Figure 32.5). We also note that diversified FREP3 sequences were also found among mRNAs recovered from a single individual snail as well, suggesting expression of diversified FREP3 molecules is a possibility (see Zhang et al. [38] for additional details, including the use of controls to check for PCR artifacts.)

Furthermore, yet additional means by which snails might present diversified FREPs as part of a complex defense response are possible. Three types of truncated cDNAs have been identified, which are identical to coding regions of previously characterized FREP genes. The truncated forms were missing partial or entire exons and thus are likely be products of alternative splicing [37] (see Figure 32.4). Additional possibilities for alternative splicing are indicated by the presence of a mini exon with appropriate type 1 splice sites in the FREP2 gene [35]. Finally, FREPs may associate with one another in complex homo- and hetero-oligomerization patterns [20], or may also interact with other host proteins to increase the diversity of response to pathogens.

It should be pointed out that also other lectins in gastropods have been found to be diverse at least at some level. Three different groups of selectin sequences were recorded from different strains of *B. glabrata* [17]. *L. flavus* (slug) SBLs were assigned to three main groups, with many minor sequence variants evident with each group. This prompted Kurachi et al. to speculate that the sequence differences could not result from alternative splicing alone [47]. Along with the diversity of FREPs, selectin and SBL may represent yet more examples of systems that concord with the recent realization that mechanisms exist that generate diversity of innate immunity factors in invertebrates [42].

32.5.1 Biological Significance of FREP Diversification

Invertebrate immune systems have been characterized as being "innate" and to have relatively limited number of recognition proteins that can bind relatively invariant PAMPs. In part because of the revelation of the extensive diversity found in FREPs in *B. glabrata* and in other categories of putative recognition molecules in other invertebrates and even jawless vertebrates such as the lamprey [42,80,81], the line between adaptive and innate immunity has become less distinct.

Furthermore, it should not be concluded that invertebrate defense systems are incapable of rapid mobilization of responses, or even of heightened responses upon secondary exposure to particular pathogens. A phenomena has been observed in *B. glabrata* referred to as acquired resistance in which snails that have been exposed to irradiated *E. paraensei* miracidia become resistant to infection when later challenged with normal miracidia [82,83]. The irradiated parasites migrate to the heart as normal but are then encapsulated and destroyed 2–9 days postinfection. When challenged with normal miracidia 5–10 days following the first exposure, miracidia of the challenge group are destroyed by hemocytes in the head/foot area and do not reach the heart [55,82,83]. A study of FREP

expression in snails sensitized by exposure to irradiated miracidia and later challenged with normal miracidia, and thus exhibiting acquired resistance, produced not only higher levels of FREPs than snails exposed only to irradiated parasites, but also produced them faster, and their response included some FREPs that normally were not induced following exposure to trematodes [55]. Thus the phenomenon of acquired resistance may be explained, at least in part, by the upregulation of FREPs induced by the prior exposure to irradiated miracidia.

There are several other examples of heightened secondary responses in invertebrates such as in the crustacean *Daphnia* in response to bacteria [84], a copepod in response to a tapeworm [85], *Anopheles gambiae* in response to bacteria and *Plasmodium* parasites [86], and the bumble bee *Bombus terrestris* in response to bacteria [87]. Although such heightened secondary responses have a very different underlying mechanism than the long-term memory responses of vertebrate adaptive immune systems, they nonetheless reinforce the concept that invertebrate defense capabilities are far from static and may be considerably more complex than often considered, thus further blurring the distinctions between immune capabilities in invertebrates and vertebrates.

32.6 SOME REMAINING PUZZLES AND SOME SPECULATIONS

As noted above, many issues remain to be fully clarified with respect to FREP biology. In addition, the following is a list of some of the puzzles that continue to intrigue us about FREPs.

What is the exact mechanism by which FREPs are diversified? Preliminary evidence from the sequencing of BAC clones suggests that FREP genes occur in tandem in the genome, and this arrangement may facilitate the process of gene conversion as a way to generate variants from source sequences. Introns of several FREP genes have regions of sequence similarity to a reverse transcriptase derived from non-long-terminal repeat retrotransposons present in the genome of *S. mansoni* [35]. Because trematodes are common parasites of *B. glabrata*, it is conceivable that a transfer of retrotransposons between parasite and host has taken place and perhaps retroelements in some way contribute to the diversification of host FREP genes. It would make sense that somatic diversification of FREP genes would be confined to tissues and cells of the internal defense system, particularly hemocytes, and that exposure to pathogens might trigger diversification, but definitive evidence on these points is lacking. Although hemocytes are likely the source of most of the FREP molecules and the site of somatic diversification, agglutinins with lectin properties are produced in the albumen gland of *B. glabrata* [10,88–90]. It is possible that some of these albumen gland-derived agglutinins are FREPs, but this awaits investigation.

Although the hemolymph of molluscs does not form obvious clots [91], do plasma polymerization reactions, potentially involving FREPs and other plasma components, nonetheless occur but usually avoid detection? Although no gel or clot is formed when *B. glabrata* plasma is mixed with activators of blood coagulation in vertebrates or of the crayfish *Pacifastacus leniusculus* [21], we have noted other suspicious precipitates in the hemolymph of echinostome-infected snails (Figure 32.6) and Matricon-Gondran et al. have observed the rapid formation of tubular double-helical filaments in *B. glabrata* injected with various foreign substances, particularly in snails infected with echinostomes [92]. The relationship between these phenomena remains unclear.

As a final issue to ponder, from a more theoretical point of view, if indeed *B. glabrata* and other invertebrates are able to produce diversified recognition molecules, on the one hand this makes sense given that some pathogens like bacteria or viruses in particular may be capable of rapid genetic change relative to their host, and diversified receptors offer a means of keeping pace with pathogens. On the other hand, if FREP diversification results in only small quantities of many different FREPs, of what good would such a response be, particularly against potentially overwhelming infections as resulting from colonization by trematodes? One speculation is that some diversification of FREP sequences may enable *B. glabrata* to prevent tracking by pathogens [38]. Mathematical modeling

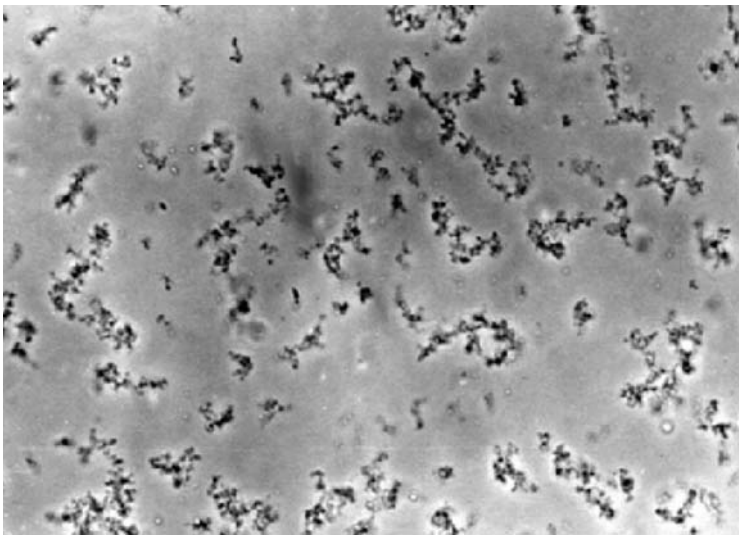


FIGURE 32.6 A micrograph of plasma of *B. glabrata* derived from a snail exposed to *E. paraensei* for 4 days. Note the precipitate that tends to form within minutes following bleeding of the snail. Similar precipitates typically do not form in unexposed control snails. (From Adema, C. M., Hertel, L. A., Loker, E. S. In *Parasite Effects on Host Physiology and Behavior*, (Ed) Beckage, N., Chapman Press, New York, 1997.)

indicated that random diversification of FREPs, leading to some variation in binding specificity, even without selection, clonal expansion or tolerance induction, can increase immunological competence of snails, under the assumptions that FREPs function as opsonins/recognition factors [80]. Another alternative speculation does invoke the radical idea of a secondary expansion of a population of hemocytes producing a given FREP specificity, i.e., clonal expansion. A related question is with all these different FREPs circulating throughout the snail's system, does the probability of generating autoreactive FREPs become a significant concern? Future studies need to consider these possibilities.

32.7 CONCLUSION AND FUTURE DIRECTIONS

In conclusion, many fascinating questions relating to FREPs remain, paramount among them a full reckoning of their functional relevance and details of the mechanisms for generating diversity. The ongoing *B. glabrata* genome project will greatly assist in interpreting all aspects of snail biology, including FREP structure and function. In addition, new tools such as RNAi are recently available for *B. glabrata* and have been shown to be capable of down regulating specific FREPs, thus providing an exciting new means of assessing the phenotypic effects of diminished FREP expression [93]. Also, the expression of recombinant FREP proteins and the production of corresponding antibodies offer the potential for availability of very specific reagents to dissect FREP functions. The available toolkit for molluscan biologists has never been better, so the prospects for new and exciting discoveries are currently very high.

ACKNOWLEDGMENTS

The work described was supported by Grant Number P20 RR018754 from the National Center for Research Resources (NCRR) and Grant Number R01 A1024340 (ESL), AI052363 (CMA), and

AI067686 (SMZ) from the National Institute of Allergy and Infectious Disease, components of the National Institutes of Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NCRR, NIAID, or NIH.

We also acknowledge our colleagues Dr. Thomas Kepler who made possible the analysis of FREP diversity and also Dr. Lynn A. Hertel for the many contributions in unraveling the complexities of FREP biology.

REFERENCES

1. Sminia, T., van der Knaap, W. P. W., and Edelenbosch, R., The role of serum factors in phagocytosis of foreign particles by blood cells of Lymnaea *stagnalis*, *Dev Comp Immunol* 3, 37–44, 1979.
2. Komano, H., Mizuno, D., and Natori, S., Purification of lectin induced in the hemolymph of *Sacrophaga peregrina* larvae on injury, *J Biol Chem* 255(7), 2919–2924, 1980.
3. Dommett, R. M., Klein, N., and Turner, M. W., Mannose-binding lectin in innate immunity: Past, present and future, *Tissue Antigens* 68(3), 193–209, 2006.
4. Nieminen, J., St-Pierre, C., and Sato, S., Galectin-3 interacts with naive and primed neutrophils, inducing innate immune responses, *J Leukoc Biol* 78(5), 1127–1135, 2005.
5. Kawabata, S. and Tsuda, R., Molecular basis of non-self recognition by the horseshoe crab tachylectins, *Bio chim Biophys Acta* 1572 (2–3), 414–421, 2002.
6. Vasta, G. R. et al., Animal lectins as self/non-self recognition molecules. Biochemical and genetic approaches to understanding their biological roles and evolution, *Ann N Y Acad Sci* 712, 55–73, 1994.
7. Cannon, J. P. et al., Individual protochordates have unique immune-type receptor repertoires, *Curr Biol* 14(12), R465–R466, 2004.
8. Morgan, J. A. et al., *Schistosoma mansoni* and *Biomphalaria*: Past history and future trends, *Parasitology* 123 Suppl, S211–28, 2001.
9. Lockyer, A. J., Jones, C. S., Noble, L. R., and Rollinson, D., Trematodes and snails: An intimate association, *Can J Zool-Revue Canadienne Zool* 82(2), 251–269, 2004.
10. Jeong, K. H. et al., Distribution and variation of hemagglutinating activity in the hemolymph of *Biomphalaria glabrata*, *J Invertebr Pathol* 38(2), 256–263, 1981.
11. Renwrantz, L. et al., Discriminative ability and function of the immunobiological recognition system of the snail *Helix pomatia*, *J Compar Physiol* 141, 477–488, 1981.
12. van der Knaap, W. P. W. et al., Some properties of an agglutinin in the haemolymph of the pond snail *Lymnaea stagnalis*, *Biol Bull* 162, 404–412, 1982.
13. Renwrantz, L. and Lackie, A., Lectins in molluscs and arthropods: Their occurrence, origin and roles in immunity. *Symposium of the Zoological Society of London*, Clarendon Press, Oxford, 1986, pp. 81–93.
14. Yuasa, H. J. et al., Cloning and sequencing of three C-type lectins from body surface mucus of the land slug, *Incilaria fruhstorferi*, *Comp Biochem Physiol B Biochem Mol Biol* 119(3), 479–484, 1998.
15. Duclermortier, P. et al., *Biomphalaria glabrata* embryonic cells express a protein with a domain homologous to the lectin domain of mammalian selectins, *Parasitol Res* 85(6), 481–486, 1999.
16. Fischer, E., Wagner, M., and Bertsch, T., *Cepaea hortensis* agglutinin-I, specific for O-glycosidically linked sialic acids, selectively labels endothelial cells of distinct vascular beds, *Histochem J* 32(2), 105–109, 2000.
17. Guillou, F. et al., Use of individual polymorphism to validate potential functional markers: Case of a candidate lectin (BgSel) differentially expressed in susceptible and resistant strains of *Biomphalaria glabrata*, *Comp Biochem Physiol B Biochem Mol Biol* 138(2), 175–181, 2004.
18. Gerlach, D., Schlott, B., and Schmidt, K. H., Cloning and expression of a sialic acid-binding lectin from the snail *Cepaea hortensis*, *FEMS Immunol Med Microbiol* 40(3), 215–221, 2004.
19. Loker, E. S. and Adema, C. M., Schistosomes, echinostomes and snails: Comparative immunobiology, *Parasitol Today* 11(3), 120–124, 1995.
20. Adema, C. M., Hertel, L. A., Loker, E. S., Infection with *Echinostoma paraensei* (Digenea) induces parasite-reactive polypeptides in the hemolymph of the gastropod host *Biomphalaria glabrata*, In Parasite Effects on Host Physiology and Behavior, (Ed) Beckage, N., Chapman Press, New York, 1997.
21. Adema, C. M. et al., A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection, *Proc Natl Acad Sci USA* 94(16), 8691–8696, 1997.
22. Adema, C. M., Hertel, L. A., and Loker, E. S., Evidence from two planorbid snails of a complex and dedicated response to digenean (echinostome) infection, *Parasitology* 119(Pt 4), 395–404, 1999.

23. Dodd, R. B. and Drickamer, K., Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity, *Glycobiology* 11(5), 71R–79R, 2001.
24. Lu, J. H. and Le, Y., Ficolins and the fibrinogen-like domain, *Immunobiology* 199(2), 190–199, 1998.
25. Kairies, N. et al., The 2.0-Å crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems, *Proc Natl Acad Sci USA* 98(24), 13519–13524, 2001.
26. Doolittle, R. F., A detailed consideration of a principal domain of vertebrate fibrinogen and its relatives, *Protein Sci* 1(12), 1563–1577, 1992.
27. Gokudan, S. et al., Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen, *Proc Natl Acad Sci USA* 96(18), 10086–10091, 1999.
28. Halanych, K. M., The new view of animal phylogeny, *Annu Rev Ecol Evol Syst* 35, 229–256, 2004.
29. Baker, N. E., Mlodzik, M., and Rubin, G. M., Spacing differentiation in the developing *Drosophila* eye: A fibrinogen-related lateral inhibitor encoded by scabrous, *Science* 250(4986), 1370–1377, 1990.
30. Xu, X. and Doolittle, R. F., Presence of a vertebrate fibrinogen-like sequence in an echinoderm, *Proc Natl Acad Sci USA* 87(6), 2097–2101, 1990.
31. Perovic-Ottstadt, S. et al., A (1→3)-beta-D-glucan recognition protein from the sponge *Suberites domuncula*. Mediated activation of fibrinogen-like protein and epidermal growth factor gene expression, *Eur J Biochem* 271(10), 1924–1937, 2004.
32. Davidson, C. J., Tuddenham, E. G., and McVey, J. H., 450 million years of hemostasis, *J Thromb Haemost* 1(7), 1487–1494, 2003.
33. Aird, W. C., Hemostasis and irreducible complexity, *J Thromb Haemost* 1(2), 227–230, 2003.
34. Loker, E. S. and Hertel, L. A., Alterations in *Biomphalaria glabrata* plasma induced by infection with the digenetic trematode *Echinostoma paraensei*, *J Parasitol* 73(3), 503–513, 1987.
35. Leonard, P. M. et al., Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail *Biomphalaria glabrata*, *Gene* 269(1–2), 155–165, 2001.
36. Zhang, S. M. et al., Parasite-responsive IgSF members in the snail *Biomphalaria glabrata*: Characterization of novel genes with tandemly arranged IgSF domains and a fibrinogen domain, *Immunogenetics* 53(8), 684–694, 2001.
37. Zhang, S. M. and Loker, E. S., The FREP gene family in the snail *Biomphalaria glabrata*: Additional members, and evidence consistent with alternative splicing and FREP retrosequences. Fibrinogen-related proteins, *Dev Comp Immunol* 27(3), 175–187, 2003.
38. Zhang, S. M. et al., Diversification of Ig superfamily genes in an invertebrate, *Science* 305(5681), 251–254, 2004.
39. Zhang, S. M. and Loker, E. S., Representation of an immune responsive gene family encoding fibrinogen-related proteins in the freshwater mollusc *Biomphalaria glabrata*, an intermediate host for *Schistosoma mansoni*, *Gene* 341, 255–266, 2004.
40. Drickamer, K., Multiplicity of lectin–carbohydrate interactions, *Nat Struct Biol* 2(6), 437–439, 1995.
41. Matsushita, M. et al., A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin, *J Biol Chem* 271(5), 2448–2454, 1996.
42. Litman, G. W., Cannon, J. P., and Rast, J. P., New insights into alternative mechanisms of immune receptor diversification, *Adv Immunol* 87, 209–236, 2005.
43. Hoek, R. M. et al., A new Ig-superfamily member, molluscan defence molecule (MDM) from *Lymnaea stagnalis*, is down-regulated during parasitosis, *Eur J Immunol* 26(4), 939–944, 1996.
44. Sun, S. C. et al., Hemolin: An insect-immune protein belonging to the immunoglobulin superfamily, *Science* 250(4988), 1729–1732, 1990.
45. Schmidt, O. et al., Specific immune recognition of insect hemolin, *Dev Comp Immunol* 17(3), 195–200, 1993.
46. Su, X. D. et al., Crystal structure of hemolin: A horseshoe shape with implications for homophilic adhesion, *Science* 281(5379), 991–995, 1998.
47. Kurachi, S. et al., Sialic-acid-binding lectin from the slug *Limax flavus*—cloning, expression of the polypeptide, and tissue localization, *Eur J Biochem* 254(2), 217–222, 1998.
48. Adams, M. D. et al., The genome sequence of *Drosophila melanogaster*, *Science* 287(5461), 2185–2195, 2000.
49. De Gregorio, E. et al., Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays, *Proc Natl Acad Sci USA* 98(22), 12590–12595, 2001.
50. Christophides, G. K. et al., Immunity-related genes and gene families in *Anopheles gambiae*, *Science* 298(5591), 159–165, 2002.

51. Zdobnov, E. M. et al., Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*, *Science* 298(5591), 149–159, 2002.
52. Wang, X., Zhao, Q., and Christensen, B. M., Identification and characterization of the fibrinogen-like domain of fibrinogen-related proteins in the mosquito, *Anopheles gambiae*, and the fruitfly, *Drosophila melanogaster*, genomes, *BMC Genomics* 6, 114, 2005.
53. Mitta, G. et al., Gene discovery and expression analysis of immune-relevant genes from *Biomphalaria glabrata* hemocytes, *Dev Comp Immunol* 29(5), 393–407, 2005.
54. Ataev, G. L. and Coustau, C., Cellular response to *Echinostoma caproni* infection in *Biomphalaria glabrata* strains selected for susceptibility/resistance, *Dev Comp Immunol* 23(3), 187–198, 1999.
55. Hertel, L. A., Susceptibility and resistance of snails to trematode infection: Investigations into the role of hemolymph components and symbionts in the snail, *Biomphalaria glabrata*, PhD Dissertation University of New Mexico, 2004.
56. Hertel, L. A., Adema, C. M., and Loker, E. S., Differential expression of FREP genes in two strains of *Biomphalaria glabrata* following exposure to the digenetic trematodes *Schistosoma mansoni* and *Echinostoma paraensei*, *Dev Comp Immunol* 29(4), 295–303, 2005.
57. Lie, K. J., Heyneman, D., and Jeong, K. H., Studies on resistance in snails. 7. Evidence of interference with the defense reaction in *Biomphalaria glabrata* by trematode larvae, *J Parasitol* 62(4), 608–615, 1976.
58. Lie, K. J. and Heyneman, D., Studies on resistance in snails: interference by nonirradiated echinostome larvae with natural resistance to *Schistosoma mansoni* in *Biomphalaria glabrata*, *J Invertebr Pathol* 29(2), 118–125, 1977.
59. Lie, K. J., Jeong, K. H., and Heyneman, D., Selective interference with granulocyte function induced by *Echinostoma paraensei* (Trematoda) larvae in *Biomphalaria glabrata* (Mollusca), *J Parasitol* 67(6), 790–796, 1981.
60. Loker, E. S., Cimino, D. F., and Hertel, L. A., Excretory-secretory products of *Echinostoma paraensei* sporocysts mediate interference with *Biomphalaria glabrata* hemocyte functions, *J Parasitol* 78(1), 104–115, 1992.
61. Humbert, E. and Coustau, C., Refractoriness of host haemocytes to parasite immunosuppressive factors as a putative resistance mechanism in the *Biomphalaria glabrata–Echinostoma caproni* system, *Parasitology* 122 (Pt 6), 651–660, 2001.
62. Sullivan, J. T. and Richards, C. S., *Schistosoma mansoni*, NIH-SM-PR-2 strain, in susceptible and nonsusceptible stocks of *Biomphalaria glabrata*: Comparative histology, *J Parasitol* 67(5), 702–708, 1981.
63. Loker, E. S. and Bayne, C. J., In vitro encounters between *Schistosoma mansoni* primary sporocysts and hemolymph components of susceptible and resistant strains of *Biomphalaria glabrata*, *Am J Trop Med Hyg* 31(5), 999–1005, 1982.
64. Yoshino, T. P. and Boswell, C. A., Antigen sharing between larval trematodes and their snail hosts: How real a phenomenon in immune evasion?, *Symp Zool Soc Lond* 56, 221–238, 1986.
65. Damian, R. T., Molecular mimicry revisited, *Parasitol Today* 3(9), 263–266, 1987.
66. Damian, R. T., Parasite immune evasion and exploitation: Reflections and projections, *Parasitology* 115 Suppl, S169–S175, 1997.
67. Thompson, R. C., Molecular mimicry in schistosomes, *Trends Parasitol* 17(4), 168, 2001.
68. Lehr, T. et al., Structural characterization of N-glycans from the freshwater snail *Biomphalaria glabrata* cross-reacting with *Schistosoma mansoni* glycoconjugates, *Glycobiology*, 17(1), 82–103, 2007. E pub 2006 Sep 13. 2006.
69. Anderson, K. V., Bokla, L., and Nusslein-Volhard, C., Establishment of dorsal-ventral polarity in the Drosophila embryo: The induction of polarity by the Toll gene product, *Cell* 42(3), 791–798, 1985.
70. Anderson, K. V., Jurgens, G., and Nusslein-Volhard, C., Establishment of dorsal–ventral polarity in the Drosophila embryo: Genetic studies on the role of the Toll gene product, *Cell* 42(3), 779–789, 1985.
71. Lemaitre, B. et al., The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults, *Cell* 86(6), 973–983, 1996.
72. Yamakawa, K. et al., DSCAM: A novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system, *Hum Mol Genet* 7(2), 227–237, 1998.
73. Watson, F. L. et al., Extensive diversity of Ig-superfamily proteins in the immune system of insects, *Science* 309(5742), 1874–1878, 2005.
74. Couch, L., Hertel, L. A., and Loker, E. S., Humoral response of the snail *Biomphalaria glabrata* to trematode infection: Observations on a circulating hemagglutinin, *J Exp Zool* 255(3), 340–349, 1990.

75. Monroy, F., Hertel, L. A., and Loker, E. S., Carbohydrate-binding plasma proteins from the gastropod *Biomphalaria glabrata*: Strain specificity and the effects of trematode infection, *Dev Comp Immunol* 16(5), 355–366, 1992.
76. Uchikawa, R. and Loker, E. S., *Echinostoma paraensei* and *Schistosoma mansoni*: Adherence of unaltered or modified latex beads to hemocytes of the host snail *Biomphalaria glabrata*, *Exp Parasitol* 75(2), 223–232, 1992.
77. Monroy, F. P. and Loker, E. S., Production of heterogeneous carbohydrate-binding proteins by the host snail *Biomphalaria glabrata* following exposure to *Echinostoma paraensei* and *Schistosoma mansoni*, *J Parasitol* 79(3), 416–423, 1993.
78. Uchikawa, R. and Loker, E. S., Lectin-binding properties of the surfaces of in vitro-transformed *Schistosoma mansoni* and *Echinostoma paraensei* sporocysts, *J Parasitol* 77(5), 742–748, 1991.
79. Lie, K. J., Heyneman, D., and Yau, P., The origin of amoebocytes in *Biomphalaria glabrata*, *J Parasitol* 61(3), 574–576, 1975.
80. Loker, E. S. et al., Invertebrate immune systems—not homogeneous, not simple, not well understood, *Immunol Rev* 198, 10–24, 2004.
81. Litman, G. W., Cannon, J. P., and Dishaw, L. J., Reconstructing immune phylogeny: New perspectives, *Nat Rev Immunol* 5(11), 866–879, 2005.
82. Lie, K. J. and Heyneman, D., Acquired resistance to echinostomes in four *Biomphalaria glabrata* strains, *Int J Parasitol* 9(6), 533–537, 1979.
83. Lie, K. J., Jeong, K. H., and Heyneman, D., Further characterization of acquired resistance in *Biomphalaria glabrata*, *J Parasitol* 68(4), 529–531, 1982.
84. Little, T. J. et al., Maternal transfer of strain-specific immunity in an invertebrate, *Curr Biol* 13(6), 489–492, 2003.
85. Kurtz, J. and Franz, K., Innate defence: Evidence for memory in invertebrate immunity, *Nature* 425(6953), 37–38, 2003.
86. Kurtz, J. and Armitage, S. A., Alternative adaptive immunity in invertebrates, *Trends Immunol*, 27(11), 493–627, 2006.
87. Sadd, B. M. and Schmid-Hempel, P., Insect immunity shows specificity in protection upon secondary pathogen exposure, *Curr Biol* 16(12), 1206–1210, 2006.
88. Michelson, E. H. and Dubois, L., Agglutinins and lysins in the molluscan family Planorbidae: A survey of hemolymph, egg-masses, and albumen-gland extracts, *Biol Bull* 153(1), 219–227, 1977.
89. Stein, P. C. and Basch, P. F., Purification and binding properties of hemagglutinin from *Biomphalaria glabrata*, *J Invertebr Pathol* 33(1), 10–8, 1979.
90. Vergote, D. et al., Characterisation of proteins differentially present in the plasma of *Biomphalaria glabrata* susceptible or resistant to *Echinostoma caproni*, *Int J Parasitol* 35(2), 215–224, 2005.
91. Millar, D. A. and Ratcliffe, N. A., *Immunology: A Comparative Approach*, Wiley, New York, 1994.
92. Matricon-Gondran, M. and Letocart, M., Internal defenses of the snail *Biomphalaria glabrata*, *J Invertebr Pathol* 74(3), 248–254, 1999.
93. Jiang, Y., Loker, E. S., and Zhang, S. M., In vivo and in vitro knockdown of FREP2 gene expression in the snail *Biomphalaria glabrata* using RNA interference, *Dev Comp Immunol* 30(10), 855–866, 2006.

33 Lectins in Sand Fly–*Leishmania* Interactions

Shaden Kamhawi and Jesus G. Valenzuela

CONTENTS

33.1 INTRODUCTION

From field and laboratory observations, it is evident that there is a close evolutionary fit between *Leishmania* species and their sand fly vectors (Killick-Kendrick, 1985). This translates into an observed restricted vector competence in certain sand flies such as *Phlebotomus papatasi*. Despite its wide distribution and prevalence in foci with more than one *Leishmania* species, *P. papatasi* has only been found infected with *Leishmania major* strains. This natural specificity was further characterized by the loss of experimental infections of *P. papatasi* with all tested species of *Leishmania* apart from *L. major* (Killick-Kendrick et al., 1994; Pimenta et al., 1994). *Phlebotomus sergenti*, a natural vector of *Leishmania tropica* also shares this specificity (Killick-Kendrick et al., 1995; Kamhawi et al., 2000). In contrast, natural vectors of visceral leishmaniasis such as *Lutzomyia longipalpis, Phlebotomus argentipes*, and *Phlebotomus halepensis* exhibit the opposite phenotype supporting maturity experimental infections with various *Leishmania* species including *L. major* and *L. tropica* (Pimenta et al., 1994; Sadlova et al., 2003). Studies aimed at elucidating this observed vector–parasite specificity focused on *Leishmania* glycoconjugates and their interaction with the sand fly midgut to which the parasite's life cycle is restricted.

Lectins are proteins or glycoproteins capable of binding specific carbohydrate moieties through one or more binding sites. Overall, multiple functions have been attributed to lectins of blood sucking insects. These include regulatory functions in morphogenesis; recognition of nonself in innate defense mechanisms and an involvement in pathogen–vector interactions (Pace and Baum, 2004). Sand fly lectin activity was first reported by Wallbanks et al. (1986). The authors showed that head and gut extracts of various species of sand flies were able to agglutinate several species of *Leishmania* parasites. The authors hypothesized that lectin–*Leishmania* interactions may be necessary for a successful infection in the fly. Further studies continued to focus on the role of sand fly lectins in vector competence for *Leishmania* parasites. Despite finding that lectins are major determinants of vector competence, few of the sand fly lectins have been fully characterized and a lot of work remains to be done to truly understand their significance in sand fly–*Leishmania* interactions.

33.2 BRIEF ACCOUNT OF THE LIFE CYCLE OF *LEISHMANIA* PARASITES IN A SAND FLY VECTOR

In order to understand the function of lectins in sand fly–*Leishmania* interactions, it is important to understand the life cycle of parasites in a natural vector. Amastigotes, the mammalian stage of the parasite, is taken up by the sand fly when it feeds on infected blood. Amastigotes differentiate into promastigotes, the flagellated stage that exists in the fly. Procyclic promastigotes are the first stage that appears in the blood bolus where they multiply within the confinement of a peritrophic membrane secreted around the blood meal. Procyclics differentiate into nondividing nectomonads, large parasites whose function is to escape the peritrophic membrane and anchor themselves to epithelial cells of the sand fly midgut. Nectomonads migrate forward into the anterior thoracic midgut and differentiate to smaller dividing leptomonads. These ultimately give rise to two stages, haptomonads that form a parasitic plug at the stomodeal valve and infectious metacyclics, free swimming and highly motile parasites adapted for transmission back to the mammalian host. The duration of the life cycle lasts 6–9 days. For a detailed account of the *Leishmania* life cycle in the sand fly, refer to reviews by Sacks and Kamhawi (2001) and Kamhawi (2006).

During their development and anterior migration within the digestive tract of the sand fly, the parasites face adverse conditions that they must overcome if they are to be transmitted to the mammalian host. Two main events include (1) the secretion of lectins and digestive enzymes that are harmful to the parasites and (2) expulsion of the degenerating peritrophic membrane with the undigested blood, making it necessary for the parasites to anchor themselves to the midgut to avoid coexpulsion. *Leishmania* parasites survive these challenges through sophisticated adaptations of their cell surface and secreted glycoconjugates.

33.3 SURFACE SUGARS OF *LEISHMANIA* PARASITES

Sugars cover the surface of *Leishmania* promastigotes and have been intimately associated with survival of the parasites in the sand fly. These include cell surface and secreted glycoconjugates that contain a common phosphoglycan structure composed of phosphodiester-linked PO_4-6Galactose(β1-4) Mannoseα1 disaccharide repeats (McConville and Ferguson, 1993; Ilg et al., 1994; Mengeling and Turco, 1998). Cell surface molecules such as lipophosphoglycan (LPG) and proteophosphoglycans (PPG) are glycosylphosphatidylinositol (GPI) anchored molecules. Secreted molecules include secreted PPG (sPPG) and secreted acid phosphatase (sAP) (Figure 33.1A). LPG is the largest and most abundant surface molecule in *Leishmania* promastigotes and forms a glycocalyx that covers the entire surface of the parasite. LPG of all *Leishmania* species and strains share the GPI anchor, glycan core, and phosphoglycan backbone of phosphorylated Gal-Man disaccharide repeats. However, they differ in the nature, number, and length of side chain oligosaccharides that branch off the backbone (Turco et al., 1987; McConville et al., 1992, 1995; Mahoney et al., 1999; Dobson et al., 2003a,b) (Figure 33.1B). These variations have been implicated in the specificity of certain sand fly species to a particular parasite, e.g., the restricted vector competence of *P. papatasi* to *L. major* and of *P. sergenti* to *L. tropica* (Pimenta et al., 1994; Kamhawi et al., 2000).

33.4 SAND FLY LECTINS AND EARLY SURVIVAL OF *LEISHMANIA* PARASITES

Early killing of *Leishmania* parasites has been observed in experimental infections of sand flies including natural vector–parasite pairs (Pimenta et al., 1997; Nieves and Pimenta, 2002; Rogers et al., 2002). *Leishmania* parasites have evolved species-specific glycoconjugates to combat the harmful environment of the early blood fed midgut. Released glyconjugates of *Leishmania* were shown to promote parasite survival at this early stage of the life cycle (Schlein et al., 1990; Sacks et al., 2000). Mutants deficient in the expression of phosphoglycan-containing molecules were completely killed during their first 24 h in the blood fed midgut of their natural vector (Sacks et al., 2000). The protective function of released glycoconjugates seems to be species specific.

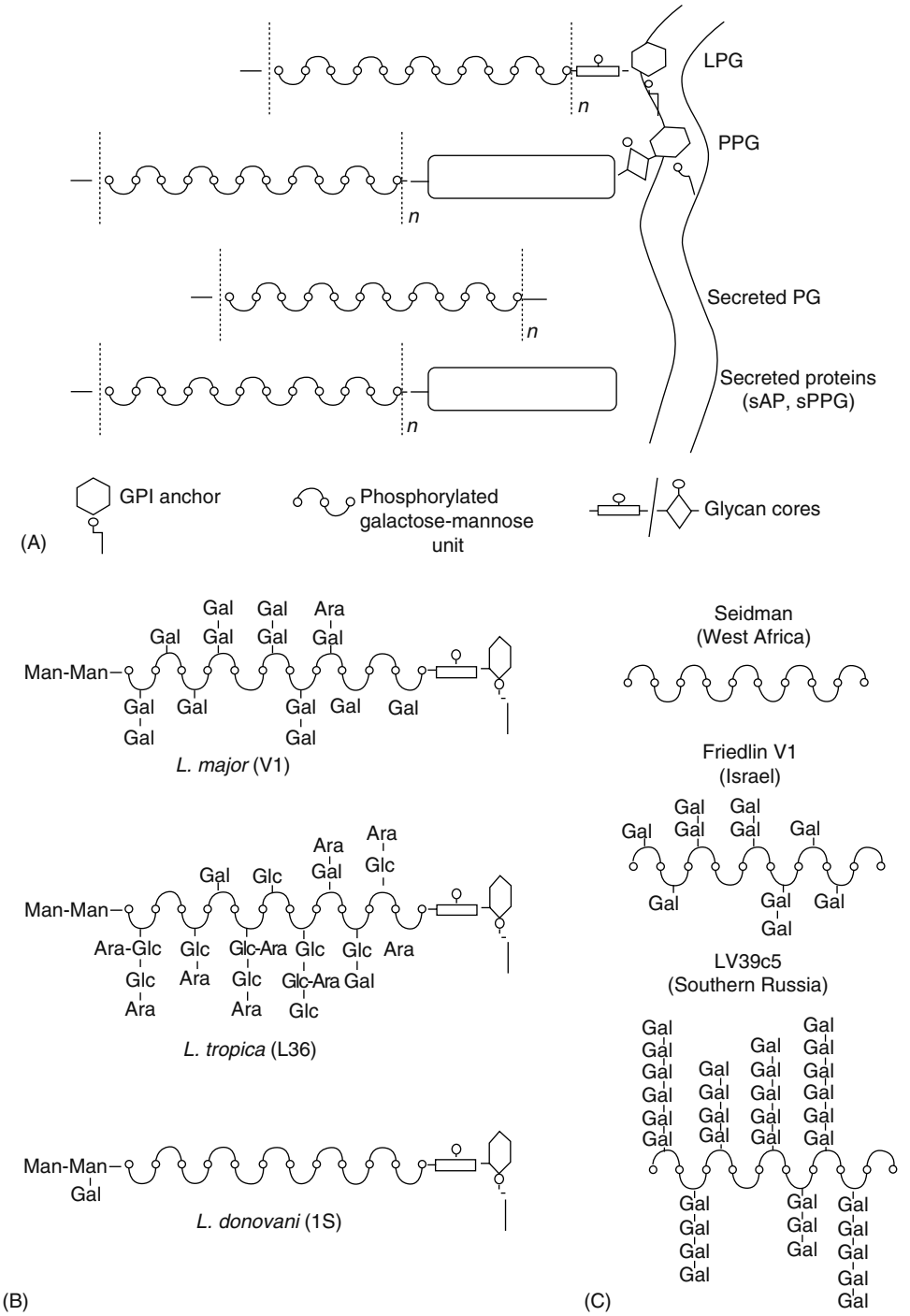


FIGURE 33.1 Phosphoglycans of *Leishmania* parasites. (A) Cell surface phosphoglycans include the lipophosphoglycan (LPG) and proteophosphoglycans (PPG), which are anchored by glycosylphosphatidylinositol (GPI). Secreted phosphoglycans include the secreted proteophosphoglycans (sPPG) and the secreted acid phosphatase (sAP). These glycoconjugates share a common phosphodiester-linked PO4–6Galactose(β1–4) Mannoseα1 disaccharide repeats. (B) Interspecific LPG polymorphisms in *L. major* (Friedlin V1), *L. tropica* (L36), and *L. donovani* (1S), depicting the variation in the nature and number of side chain oligosaccharides that branch off the LPG backbone. (C) Intraspecific LPG polymorphisms in *L. major* depicting differences in the number and length of polygalactose side chains in the Seidman, Friedlin V1, and LV39c5 strains.

Schlein et al. (1990) showed that released glycoconjugates from cultures of *L. major* promastigotes promoted the early survival of a glycoconjugate-deficient *L. major* strain in the midgut of *P. papatasi*, a natural vector of *L. major*, and delayed blood meal digestion. The released glycoconjugates from *Leishmania donovani,* a parasite that cannot survive in *P. papatasi*, was not able to promote survival. Most studies have attributed the early killing of *Leishmania* parasites and the delay of blood digestion to the corresponding secretion of digestive enzymes, shown to be harmful to the developing parasites (Sacks and Kamhawi, 2001). However, the ingestion of blood also induces the secretion of lectins, observed in several sand fly species (Wallbanks et al., 1986; Volf et al., 1994). Midgut lectin activity peaks around 2 days postblood meal, increasing two- to 16-fold depending on the sand fly species and protein content of the meal (Volf and Killick-Kendrick, 1996; Volf and Palanova, 1996). Moreover, lectin activity in unfed female sand flies was 50 times higher than in males, indicating that they are more relevant for the blood-sucking females (Volf and Killick-Kendrick, 1996). The agglutination is optimal at pH 7.0–7.5, the state of the midgut following the ingestion of blood (Palanova and Volf, 1997) and returns to baseline levels only after defecation (Volf and Killick-Kendrick, 1996). Moreover, gut extracts of female sand flies can agglutinate several species of *Leishmania* (Wallbanks et al., 1986; Svobodova et al., 1996). From the above observations, the contribution of lectin activity to the observed early killing of *Leishmania* cannot be excluded.

Lectins secreted by sand flies seem to be conserved in unfed and fed females. The carbohydrate specificity of the lectin activity in five sand fly species remained the same in unfed and fed flies, indicating that the same lectins are expressed at a higher level in the fed flies (Palanova and Volf, 1997). However, the magnitude of agglutination varied in several sand fly species and marked differences were observed in their agglutination of different *Leishmania* species and strains (Svobodova et al., 1996). This variability in parasite agglutination suggests a role for sand fly lectins in modulating the early survival of *Leishmania* parasites within the blood-fed midgut. This is reinforced by the enhancement of *L. major* infections following inhibition of midgut lectin activity in its natural vector *Phlebotomus duboscqi*, significantly increasing parasite loads in mature infections (Volf et al., 1998, 2001). In addition, the LPG of *L. major* promastigotes strongly inhibited the activity of lectins in the midgut of another natural vector, *P. papatasi* (Volf and Palanova, 1996). In another vector–parasite interaction, a molecule with lectin and trypsin activities (named *Gpl*) was identified in the midgut of the tsetse fly *Glossina morsitans* (Abubakar et al., 2006). This protein was able to stimulate the differentiation of trypanosomes *in vitro*. Additionally, the recombinant protein was able to bind D(+)-glucosamine, agglutinate trypanosomes, and rabbit red blood cells, and also has proteolytic activities (Abubakar et al., 2006). Due to the effects of this protein on parasite transformation, it was proposed that *Gpl* is important for the establishment of the parasites in these flies (Abubakar et al., 2006).

Taken together, these facts suggest that sand fly lectins together with digestive enzymes are early determinants of vector competence, influencing the survival of some parasite species/strains in certain sand fly vectors.

33.5 SAND FLY LECTINS AND MIDGUT ATTACHMENT OF *LEISHMANIA* PARASITES

The ability of parasites to attach to the midgut epithelium of sand flies is considered a major determinant of vector competence. Without attachment, the parasites are lost due to midgut peristalsis, aimed at expelling the degenerating peritrophic membrane and undigested blood meal remnants. Several studies have shown that LPG, the major glycoconjugate on the surface of *Leishmania* promastigotes, is the ligand-mediating attachment of the parasites to the sand fly midgut (Pimenta et al., 1992; Butcher et al., 1996; Sacks et al., 2000). Moreover, differences in the composition of the LPG side chain sugars controls to a large extent species-specific vector competence (Pimenta et al., 1994; Mahoney et al., 1999; Kamhawi et al., 2000).

A pattern for LPG structure diversity emerged in which the LPG structure of *Leishmania* species that persist in restricted vectors, such as *L. major* and *L. tropica* in *P. papatasi* and *P. sergenti*, respectively, is complex with numerous side chains of sugar polymers (McConville et al., 1990; McConville et al., 1995). Conversely, permissive vectors, such as *Lu. longipalpis* and *P. argentipes,* are naturally infected by *Leishmania* species (*L. chagasi* and *L. donovani*) that possess simple LPG molecules with few or no side chain branches (Turco et al., 1987; Mahoney et al., 1999; Soares et al., 2002) (Figure 33.1B). This indicates that receptors on sand fly midguts are variable and their complexity provides an evolutionary drive manifested in LPG polymorphisms. This theory is reinforced by the fact that intraspecific variation in the LPG structure of some *L. major* strains rendered them incapable of infecting *P. papatasi*. For example, *L. major* Seidman, a strain isolated from West Africa, whose LPG naturally lacks polygalactose side chains (Figure 33.1C) cannot grow in *P. papatasi* but grows in *P. duboscqi*, a sister species that belongs to the same subgenera *Phlebotomus*. Interestingly, the wide geographical distribution of *P. papatasi* does not extend to West Africa where *P. duboscqi* prevails.

Until recently, sand fly midgut receptors that bind *Leishmania* LPGs were unknown. Since several studies implicated the sugar moieties of LPG in midgut attachment (Pimenta et al., 1992), it was reasonable to suspect that the sand fly midgut receptor was a lectin. Indeed, a tandem repeat galectin, termed PpGalec, was identified as the midgut receptor for *L. major* procyclic LPG in the sand fly *P. papatasi* (Kamhawi et al., 2004). *PpGalec* was recognized from a cDNA library of dissected female *P. papatasi* midguts using massive cDNA sequencing and bioinformatic approaches (Kamhawi et al., 2004). The gene was strongly upregulated in adults compared to larval and pupal stages, constitutively expressed, and located on the surface of most midgut epithelial cells. PpGalec was strongly expressed by these cells, forming abundant clusters on the luminal surface of the midgut (Kamhawi et al., 2004) (Figure 33.2). These are all important functional prerequisites for a *Leishmania* midgut receptor. Furthermore, recombinant PpGalec showed a binding specificity to live parasites as well as to purified procyclic LPG of *L. major* (Friedlin V1). Procyclic LPG of this strain of *L. major* has side chains of—one to three galactose molecules (Figure 33.1B and C). This is compatible with the binding specificity reported for galectins that have a cation-independent affinity to β-galactosides and typically bind to type I Gal_1,3GlcNAc or type II Gal_1,4GlcNAc units (Pace and Baum, 2004; Rabinovich and Gruppi, 2005). The strongest evidence of the midgut receptor function of PpGalec was the impairment of *in vivo* development of *L. major* (Friedlin V1) in *P. papatasi* by antibodies against rPpGalec (Kamhawi et al., 2004). The anti-PpGalec antibodies occupy attachment sites used by the parasite for attachment resulting in loss of infection (Figure 33.2). This has implications for the possible use of midgut receptors as transmission-blocking vaccines. Additionally, this represents the first receptor for a parasite identified from an insect vector.

An important aspect of the observed binding specificity of PpGalec to *L. major* (Friedlin V1) was that it was only strain specific. *L. major* strains that possessed no galactose side chains (*L. major* Seidman), as well as those with very long polygalactose side chains (*L. major* LV39c5) (Figure 33.1C) showed no binding to PpGalec (Joshi et al., 1998; Dobson et al., 2003a; Kamhawi et al., 2004). PpGalec was identified from a colony of *P. papatasi* that originated from the Middle East where the *L. major* Friedlin V1 strain was isolated. The *L. major* LV39c5 and the *L. major* Seidman strains were isolated from the former USSR and West Africa, respectively (Joshi et al., 1998; Dobson et al., 2003b). Despite the fact that *P. papatasi* is the vector of *L. major* in the former USSR, geographical isolation may have resulted in intraspecific variations in PpGalec that explains the poor growth of *L. major* LV39c5 in flies originating from the Middle East (Kamhawi et al., 2004). The sand fly vector for *L. major* in West Africa is *P. duboscqi*, a sister species of *P. papatasi*. A close homolog of PpGalec was identified in *P. duboscqi* (Valenzuela and Kamhawi, unpublished material). However, this fly supports the growth of *L. major* Seidman (Joshi et al., 2002) and argues for the presence of a secondary receptor, which is absent or lost in *P. papatasi*, and which renders *P. duboscqi* more permissive to varied *L. major* strains. Another possibility is that the galectin-like

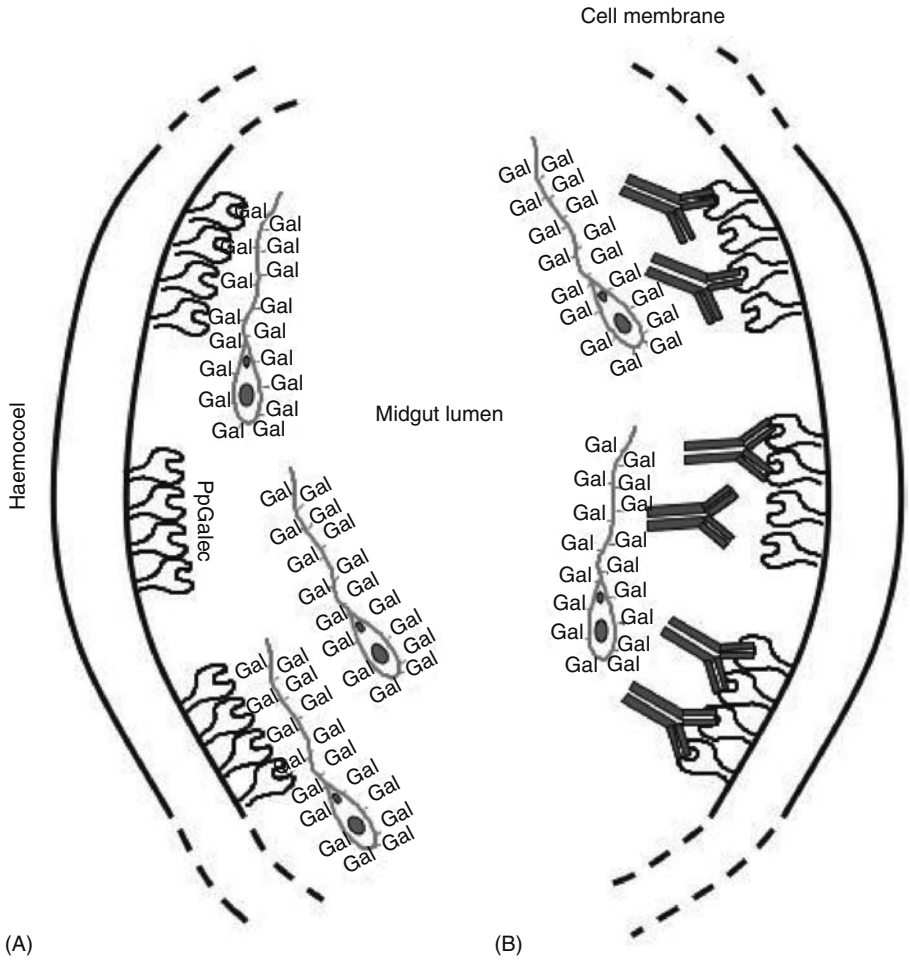


FIGURE 33.2 A cartoon depicting the molecular associations between the LPG of *Leishmania major* and the galectin receptor (PpGalec) on the surface of the sand fly midgut. (A) Poly-galactose residues of procyclic *L. major* LPG interacting with clusters of cells expressing the galectin receptor (PpGalec) at the luminal surface of the sand fly midgut. (B) *Leishmania* LPG and PpGalec interaction disrupted by anti-PpGalec antibodies. This results in loss of parasite attachment and termination of their development in the sand fly midgut.

molecule in *P. duboscqi* does not have the same sugar-binding properties as PpGalec. The role of intraspecific variations of sand fly vectors in vector competence needs to be promptly addressed.

The true function of PpGalec in the fly remains unknown. Galectins are multifunctional molecules and apart from their involvement in host–pathogen interactions (Kamhawi et al., 2004; Young and Meeusen, 2004; Rabinovich and Gruppi, 2005), they function in differentiation, cell–cell, and cell–matrix interactions, immunity, and cell signaling (Rabinovich et al., 2002; Pace and Baum, 2004; Rabinovich and Gruppi, 2005). Therefore, it is difficult to predict the function of PpGalec. Based on its midgut specificity and its weak expression in males (that only take sugar meals) and larval and pupal developmental stages, potential functions of PpGalec may involve blood digestion, cell signaling, and defense, and less likely development.

Other lectins, assigned by sequence homology as c-Type and P-Type lectins, were identified from a cDNA library of *P. sergenti* (Jochim and Valenzuela, unpublished material). However, their relevance as midgut receptors for *Leishmania* parasites remains to be investigated. Until recently,

the vector competence of *P. sergenti* was thought to be restricted to *L. tropica* (Killick-Kendrick, 1999; Kamhawi et al., 2000). However, Gebre-Michael et al. (2004) reported the isolation of *Leishmania aethiopica* from this species in Ethiopia. Moreover, *L. tropica* strains exhibiting variations in the sugar moieties forming the LPG side chains were isolated from other flies including *Phlebotomus saevus* in Ethiopia and *Phlebotomus arabicus* from Israel (Jacobson, 2003; Gebre-Michael et al., 2004; Soares et al., 2004). In addition, *P. arabicus,* subgenus *Adlerius*, was able to transmit *L. tropica* to hyraxes, a natural reservoir of the parasite, by bite (Svobodova et al., 2006). This confirms that *P. arabicus–L. tropica* represent a competent sand fly–*Leishmania* pair in nature. The restricted competence of certain sand fly vectors to *Leishmania* parasites at the species level has been confounded by the various associations uncovered between different *Leishmania* and sand fly strains of the same species and highlights the complexity of sand fly–*Leishmania* interactions.

To date, PpGalec remains the only fully characterized midgut receptor of a competent *Leishmania*–sand fly pair. Finding PpGalec was facilitated by the restricted vector competence of *P. papatasi* to *L. major* in the Middle East. Finding the receptor in permissive vectors that support the full development of several *Leishmania* species has been more elusive. Several species belonging to different subgenera are considered permissive as a result of experimental and natural infections with several *Leishmania* species. These include *Lu. longipalpis, P. argentipes, P. halepensis*, and *P. arabicus* (Walters, 1993; Pimenta et al., 1994; Sadlova et al., 2003; Svobodova et al., 2006; Volf and Myskova, 2007). Recently, permissive vectors but not restricted ones were shown to possess O-glycosylated midgut proteins bearing *N*-acetylgalactosamine (GalNAc) epitopes (Evangelista and Leite, 2002; Svobodova et al., 2006; Volf and Myskova, 2007). Moreover, *Lu. longipalpis* supported the full growth of LPG mutants of *L. major* and *L. mexicana* (Rogers et al., 2004; Volf and Myskova, 2007), indicating that midgut attachment in permissive vectors occurs via a LPG-independent manner (Volf and Myskova, 2007). This led to the proposal of a novel mechanism for midgut attachment in *Leishmania*–sand fly interactions in which lectin molecules expressed by the parasites act as a receptor for a midgut surface O-glycosylated proteins (Volf and Myskova, 2007). Indeed, a lectin with GalNAc-specific activity has been reported on the surface of *Leishmania* promastigotes (Schottelius, 1992; Kock et al., 1997). Identification and isolation of the sand fly protein bearing the O-glycosylated moiety that may function as a ligand for the parasite lectin will be necessary to demonstrate the specificity of this proposed association. Furthermore, other biological questions will need to be addressed such as the necessity for attached parasites to detach in order to be transmitted. In vectors where LPG is the ligand for midgut attachment, detachment of parasites is mediated by changes in the LPG structure as they differentiate from procyclics (attached) to infectious metacyclics (freely motile) (Sacks and Kamhawi, 2001). It remains that the hypothesis of a parasite lectin-binding to a sand fly sugar is supported by exciting findings and emphasizes the importance of lectins in the elaborate adaptations of *Leishmania* parasites to their sand fly vectors.

33.6 TANDEM CARBOHYDRATE-BINDING DOMAINS OF GALECTINS FROM OTHER INSECTS

The galectin (PpGalec), identified from the sand fly *P. papatasi,* represents the first molecular characterization of a lectin receptor for a pathogen in the midgut of a disease vector (Kamhawi et al., 2004). PpGalec has a molecular weight of 35 kDa; has two carbohydrate recognition domains (CRDs) similar to galectins 4, 6, 8, 9, and 12 from mammals; and does not have a Ca^{2+}-binding domain nor a transmembrane domain. Additionally, it does not have a secretory signal peptide and probably follows a nonclassical secretory pathway as reported for other galectins (Rabinovich et al., 2002; Pace and Baum, 2004; Rabinovich and Gruppi, 2005).

Galectins are widespread in different organisms including mammals and nonmammals (Pace and Baum, 2004). Due to recent advances in large-scale DNA sequencing of various organisms including insects (Dimopoulos et al., 2000; Valenzuela et al., 2002; Kriventseva et al., 2005), various

transcripts or genes coding for proteins with homologies to tandem repeat galectins (as well as other types of galectins) have been identified (Pace and Baum, 2004). Scanning of the GenBank database for insect galectins with molecular weights similar to PpGalec and with two potential CRDs, 8, 15, and 6 galectin homologs were identified from *Aedes aegypti*, *Anopheles gambiae*, and *Drosophila melanogaster*, respectively. Additionally, an expressed sequence tag (EST) database of *Lu. longipalpis* generated four galectin homologs (Dillon et al., 2006). Recently, a galectin from the tick *Ornithodoros moubata* was reported (Huang et al., 2006). Sequence comparison among one galectin from each insect group and galectins from mammals shows that these proteins share conserved CRDs (N-terminal and C-terminal) (Figure 33.3A). Phylogenetic analysis of insect galectins shows that these molecules form clusters apart from the mammalian galectins and the *O. moubata* galectin (Figure 33.3B). Moreover, insect galectins cluster in three distinct groups, one includes *D. melanogaster and G. morsitans*, the other includes the mosquitoes *A. aegypti* and *A. gambiae,* and the third cluster includes the sand flies *P. papatasi* and *Lu. longipalpis* (Figure 33.3B). The amino acids involved in carbohydrate binding were conserved in the majority of the sequences (Figure 33.3, gray-shaded amino acids). The variation noted in some of the amino acids responsible for carbohydrate binding may result in the different carbohydrate specificities of these galectins, leading ultimately to a potential difference in their function (Figure 33.3). The role of *G. morsitans*, *A. aegypti*, and *A. gambiae* galectins has not been elucidated. The *G. morsitans* galectin has an arginine for histidine and a tyrosine for arginine substitutions in the N-terminal CRD and a serine for aspargine substitution in the C-terminal CRD. In the *A. aegypti* and *A. gambiae* galectins, all the amino acids relevant to sugar binding were highly conserved (Figure 33.3A). Dmgal*,* a *D. melanogaster* galectin expressed in developing immune, neural, and muscle tissues, was proposed to function in cell–cell interactions or to cross-link glycoprotein receptors on the cell surface to trigger signal transduction pathways. Interestingly, Dmgal recombinant protein was able to bind to a β-lactose affinity column, suggesting Dmgal may be a functional galectin (Pace et al., 2002).

The *P. papatasi* galectin (PpGalec) was able to agglutinate parasites and to bind specifically to the β-galactose sugar moiety of *L. major* LPG. Interestingly, this galectin as well as the putative galectin from *Lu. longipalpis* had tyrosine for tryptophane substitution in the N-terminal CRD. In addition, PpGalec had another alanine for valine substitution in the C-terminal CRD (Figure 33.3A). PpGalec is constitutively and specifically expressed in the midgut of *P. papatasi* and functions as a receptor for *L. major* (Kamhawi et al., 2004). In *Lu. longipalpis,* an EST made from unfed, blood-fed, and infected (with several pathogens) whole sand flies resulted in the identification of four galectin homologs (Dillon et al., 2006). However, following the sequencing of approximately 8000 *Lu. longipalpis* midgut transcripts (Jochim and Valenzuela, unpublished data), only one partially full-length galectin transcript was identified. This suggests that these four putative galectins may not be very abundant in the midgut and are probably expressed in other tissues. The tick *O. moubata* galectin (Huang et al., 2006), which also bears similarities to PpGalec, was expressed in different tissues including hemocytes, midgut, and reproductive organs. Automated frontal affinity chromatography showed that the tick galectin was able to bind lactosamine type disaccharides, Galβ1–3(4) Glc(NAc), providing an insight into the nature of its ligands. This approach may be relevant for the determination of the function of other putative galectins and their binding affinities for different oligosaccharides.

33.7 CONCLUSION AND FUTURE DIRECTIONS

Lectins are multifunctional regulatory molecules involved in almost every aspect of biology. Therefore, it is not surprising to find that they play a significant part in *Leishmania*–sand fly interactions. In sand flies, some lectins are universally induced by blood and can aggregate *Leishmania* parasites, indicating that they function in the innate immunity of the insect. However, a tandem repeat galectin identified from *P. papatasi* is constitutively expressed in adult females and is present in some but not all sand fly species. This galectin has been exploited by *L. major* parasites that evolved sugar moieties on their surface with affinity for PpGalec. The attachment of the LPG polygalactose side chains of

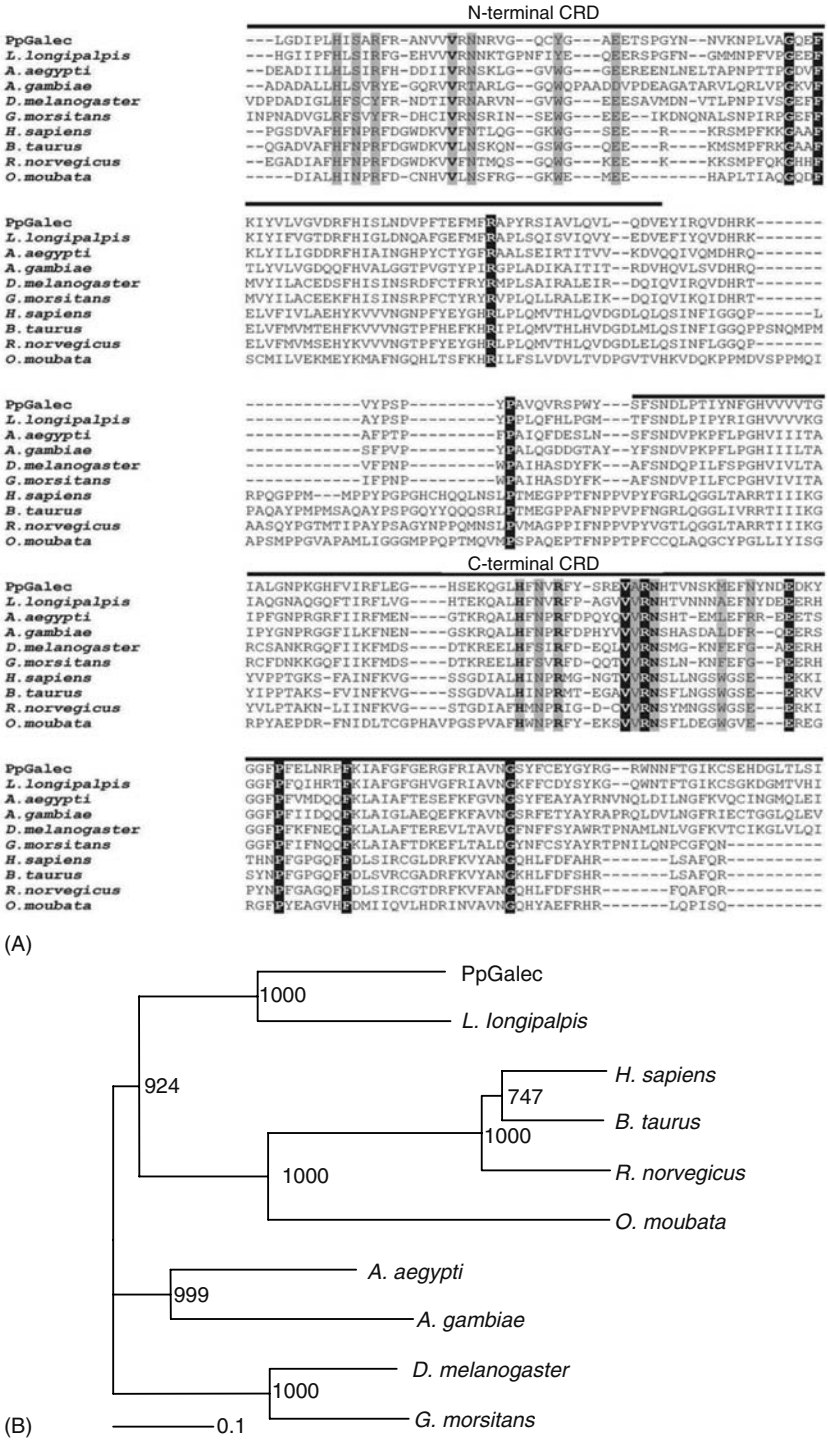


FIGURE 33.3 Sequence analysis of tandem repeat galectins. (A) Clustal alignments of galectin homologs from insects including the sand flies *P. papatasi* (PpGalec) and *Lu. longipalpis*, the mosquitoes *A. aegypti* and *A. gambiae*, the fruit fly *D. melanogaster*, and the tse tse fly *G. morsitans*, and from mammals including *H. sapiens*, *B. taurus*, and *R. norvegicus* and from the soft tick *O. moubata*. Black shading and or bold indicates amino acid sequence identity and gray regions indicate amino acids involved in β-galactoside binding. The two black bold lines on top of the sequences indicate the two carbohydrate recognition domains (CRD). (B) Phylogenetic tree analysis of the various insect galectins aligned in (A). Scale represents 0.1 nucleotide substitutions per site.

L. major to PpGalec is a key survival step necessary for successful transmission of the parasites to the mammalian host. However, the true function of PpGalec in the sand fly remains unknown.

Other galectins have been identified from several insects using high-throughput approaches. Similar to PpGalec, the function of these galectins is unknown. Further biological tests or functional genomic approaches are needed to understand the biological significance of these galectins. Moreover, because mammalian and insect galectins are highly conserved, the study of insect galectins may provide a more accessible model to uncover new functions in this intriguing family of proteins.

REFERENCES

Abubakar, L. U., Bulimo, W. D., Mulaa, F. J., and Osir, E. O. 2006. Molecular characterization of a tsetse fly midgut proteolytic lectin that mediates differentiation of African trypanosomes. *Insect Biochem Mol Biol*, **36**, 344–352.

Butcher, B. A., Turco, S. J., Hilty, B. A., Pimenta, P. F., Panunzio, M., and Sacks, D. L. 1996. Deficiency in beta1,3-galactosyltransferase of a Leishmania major lipophosphoglycan mutant adversely influences the *Leishmania*–sand fly interaction. *J Biol Chem*, **271**, 20573–20579.

Dillon, R. J., Ivens, A. C., Churcher, C., Holroyd, N., Quail, M. A., Rogers, M. E., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Lehane, M. J., and Bates, P. A. 2006. Analysis of ESTs from *Lutzomyia longipalpis* sand flies and their contribution toward understanding the insect–parasite relationship. *Genomics*, **88**, 831–840.

Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M. B., and Kafatos, F. C. 2000. *Anopheles gambiae* pilot gene discovery project: Identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines. *Proc Natl Acad Sci USA*, **97**, 6619–6624.

Dobson, D. E., Mengeling, B. J., Cilmi, S., Hickerson, S., Turco, S. J., and Beverley, S. M. 2003a. Identification of genes encoding arabinosyltransferases (SCA) mediating developmental modifications of lipophosphoglycan required for sand fly transmission of leishmania major. *J Biol Chem*, **278**, 28840–28848.

Dobson, D. E., Scholtes, L. D., Valdez, K. E., Sullivan, D. R., Mengeling, B. J., Cilmi, S., Turco, S. J., and Beverley, S. M. 2003b. Functional identification of galactosyltransferases (SCGs) required for species-specific modifications of the lipophosphoglycan adhesin controlling *Leishmania major*–sand fly interactions. *J Biol Chem*, **278**, 15523–15531.

Evangelista, L. G. and Leite, A. C. 2002. Histochemical localization of *N*-acetyl-galactosamine in the midgut *Lutzomyia longipalpis* (Diptera: Psychodidae). *J Med Entomol*, **39**, 432–439.

Gebre-Michael, T., Balkew, M., Ali, A., Ludovisi, A., and Gramiccia, M. 2004. The isolation of *Leishmania tropica* and *L. aethiopica* from *Phlebotomus* (Paraphlebotomus) species (Diptera: Psychodidae) in the Awash Valley, northeastern Ethiopia. *Trans R Soc Trop Med Hyg*, **98**, 64–70.

Huang, X., Tsuji, N., Miyoshi, T., Nakamura-Tsuruta, S., Hirabayashi, J., and Fujisaki, K. 2006. Molecular characterization and oligosaccharide binding properties of a galectin from the argasid tick Ornithodoros moubata. *Glycobiology*, **17**, 313–323.

Ilg, T., Stierhof, Y. D., Wiese, M., Mcconville, M. J., and Overath, P. 1994. Characterization of phosphoglycan-containing secretory products of Leishmania. *Parasitology*, **108 Suppl**, S63–S71.

Jacobson, R. L. 2003. *Leishmania tropica* (Kinetoplastida: Trypanosomatidae)—a perplexing parasite. *Folia Parasitol (Praha)*, **50**, 241–250.

Joshi, P. B., Kelly, B. L., Kamhawi, S., Sacks, D. L., and Mcmaster, W. R. 2002. Targeted gene deletion in *Leishmania major* identifies leishmanolysin (GP63) as a virulence factor. *Mol Biochem Parasitol*, **120**, 33–40.

Joshi, P. B., Sacks, D. L., Modi, G., and Mcmaster, W. R. 1998. Targeted gene deletion of *Leishmania major* genes encoding developmental stage-specific leishmanolysin (GP63). *Mol Microbiol*, **27**, 519–530.

Kamhawi, S. 2006. Phlebotomine sand flies and *Leishmania* parasites: Friends or foes? *Trends Parasitol*, **22**, 439–445.

Kamhawi, S., Modi, G. B., Pimenta, P. F., Rowton, E., and Sacks, D. L. 2000. The vectorial competence of *Phlebotomus sergenti* is specific for *Leishmania tropica* and is controlled by species-specific, lipophosphoglycan-mediated midgut attachment. *Parasitology*, **121** (Pt 1), 25–33.

Kamhawi, S., Ramalho-Ortigao, M., Pham, V. M., Kumar, S., Lawyer, P. G., Turco, S. J., Barillas-Mury, C., Sacks, D. L., and Valenzuela, J. G. 2004. A role for insect galectins in parasite survival. *Cell*, **119**, 329–341.

Killick-Kendrick, R. 1985. Some epidemiological consequences of the evolutionary fit between *Leishmaniae* and their phlebotomine vectors. *Bull Soc Pathol Exot Filiales*, **78**, 747–755.

Killick-Kendrick, R. 1999. The biology and control of phlebotomine sand flies. *Clin Dermatol*, **17**, 279–289.

Killick-Kendrick, R., Killick-Kendrick, M., and Tang, Y. 1994. Anthroponotic cutaneous leishmaniasis in Kabul, Afghanistan: The low susceptibility of *Phlebotomus papatasi* to *Leishmania tropica*. *Trans R Soc Trop Med Hyg*, **88**, 252–253.

Killick-Kendrick, R., Killick-Kendrick, M., and Tang, Y. 1995. Anthroponotic cutaneous leishmaniasis in Kabul, Afghanistan: The high susceptibility of *Phlebotomus sergenti* to *Leishmania tropica*. *Trans R Soc Trop Med Hyg*, **89**, 477.

Kock, N. P., Gabius, H. J., Schmitz, J., and Schottelius, J. 1997. Receptors for carbohydrate ligands including heparin on the cell surface of *Leishmania* and other trypanosomatids. *Trop Med Int Health*, **2**, 863–874.

Kriventseva, E. V., Koutsos, A. C., Blass, C., Kafatos, F. C., Christophides, G. K., and Zdobnov, E. M. 2005. AnoEST: Toward *A. gambiae* functional genomics. *Genome Res*, **15**, 893–899.

Mahoney, A. B., Sacks, D. L., Saraiva, E., Modi, G., and Turco, S. J. 1999. Intra-species and stage-specific polymorphisms in lipophosphoglycan structure control *Leishmania donovani*–sand fly interactions. *Biochemistry*, **38**, 9813–9823.

Mcconville, M. J. and Ferguson, M. A. 1993. The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem J*, **294** (Pt 2), 305–324.

Mcconville, M. J., Schnur, L. F., Jaffe, C., and Schneider, P. 1995. Structure of Leishmania lipophosphoglycan: Inter- and intra-specific polymorphism in Old World species. *Biochem J*, **310** (Pt 3), 807–818.

Mcconville, M. J., Thomas-Oates, J. E., Ferguson, M. A., and Homans, S. W. 1990. Structure of the lipophosphoglycan from *Leishmania major*. *J Biol Chem*, **265**, 19611–19623.

Mcconville, M. J., Turco, S. J., Ferguson, M. A., and Sacks, D. L. 1992. Developmental modification of lipophosphoglycan during the differentiation of *Leishmania major* promastigotes to an infectious stage. *Embo J*, **11**, 3593–3600.

Mengeling, B. J. and Turco, S. J. 1998. Microbial glycoconjugates. *Curr Opin Struct Biol*, **8**, 572–577.

Nieves, E. and Pimenta, P. F. 2002. Influence of vertebrate blood meals on the development of Leishmania (Viannia) braziliensis and Leishmania (Leishmania) amazonensis in the sand fly *Lutzomyia migonei* (Diptera: Psychodidae). *Am J Trop Med Hyg*, **67**, 640–647.

Pace, K. E. and Baum, L. G. 2004. Insect galectins: Roles in immunity and development. *Glycoconj J*, **19**, 607–614.

Pace, K. E., Lebestky, T., Hummel, T., Arnoux, P., Kwan, K., and Baum, L. G. 2002. Characterization of a novel *Drosophila melanogaster* galectin. Expression in developing immune, neural, and muscle tissues. *J Biol Chem*, **277**, 13091–13098.

Palanova, L. and Volf, P. 1997. Carbohydrate-binding specificities and physico-chemical properties of lectins in various tissue of phlebotominae sandflies. *Folia Parasitol (Praha)*, **44**, 71–76.

Pimenta, P. F., Modi, G. B., Pereira, S. T., Shahabuddin, M., and Sacks, D. L. 1997. A novel role for the peritrophic matrix in protecting Leishmania from the hydrolytic activities of the sand fly midgut. *Parasitology*, **115** (Pt 4), 359–369.

Pimenta, P. F., Saraiva, E. M., Rowton, E., Modi, G. B., Garraway, L. A., Beverley, S. M., Turco, S. J., and Sacks, D. L. 1994. Evidence that the vectorial competence of phlebotomine sand flies for different species of Leishmania is controlled by structural polymorphisms in the surface lipophosphoglycan. *Proc Natl Acad Sci USA*, **91**, 9155–9159.

Pimenta, P. F., Turco, S. J., Mcconville, M. J., Lawyer, P. G., Perkins, P. V., and Sacks, D. L. 1992. Stage-specific adhesion of Leishmania promastigotes to the sandfly midgut. *Science*, **256**, 1812–1815.

Rabinovich, G. A. and Gruppi, A. 2005. Galectins as immunoregulators during infectious processes: From microbial invasion to the resolution of the disease. *Parasite Immunol*, **27**, 103–114.

Rabinovich, G. A., Rubinstein, N., and Fainboim, L. 2002. Unlocking the secrets of galectins: A challenge at the frontier of glyco-immunology. *J Leukoc Biol*, **71**, 741–752.

Rogers, M. E., Chance, M. L., and Bates, P. A. 2002. The role of promastigote secretory gel in the origin and transmission of the infective stage of *Leishmania mexicana* by the sandfly *Lutzomyia longipalpis*. *Parasitology*, **124**, 495–507.

Rogers, M. E., Ilg, T., Nikolaev, A. V., Ferguson, M. A., and Bates, P. A. 2004. Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. *Nature*, **430**, 463–467.

Sacks, D. and Kamhawi, S. 2001. Molecular aspects of parasite–vector and vector–host interactions in leishmaniasis. *Annu Rev Microbiol*, **55**, 453–483.

Sacks, D. L., Modi, G., Rowton, E., Spath, G., Epstein, L., Turco, S. J., and Beverley, S. M. 2000. The role of phosphoglycans in Leishmania–sand fly interactions. *Proc Natl Acad Sci USA*, **97**, 406–411.

Sadlova, J., Hajmova, M., and Volf, P. 2003. Phlebotomus (Adlerius) halepensis vector competence for *Leishmania major* and *Le. tropica*. *Med Vet Entomol*, **17**, 244–250.

Schlein, Y., Schnur, L. F., and Jacobson, R. L. 1990. Released glycoconjugate of indigenous *Leishmania major* enhances survival of a foreign *L. major* in *Phlebotomus papatasi*. *Trans R Soc Trop Med Hyg*, **84**, 353–355.

Schottelius, J. 1992. Neoglycoproteins as tools for the detection of carbohydrate-specific receptors on the cell surface of *Leishmania*. *Parasitol Res*, **78**, 309–315.

Soares, R. P., Barron, T., Mccoy-Simandle, K., Svobodova, M., Warburg, A., and Turco, S. J. 2004. *Leishmania tropica*: Intraspecific polymorphisms in lipophosphoglycan correlate with transmission by different Phlebotomus species. *Exp Parasitol*, **107**, 105–114.

Soares, R. P., Macedo, M. E., Ropert, C., Gontijo, N. F., Almeida, I. C., Gazzinelli, R. T., Pimenta, P. F., and Turco, S. J. 2002. *Leishmania chagasi*: Lipophosphoglycan characterization and binding to the midgut of the sand fly vector *Lutzomyia longipalpis*. *Mol Biochem Parasitol*, **121**, 213–224.

Svobodova, M., Volf, P., and Killick-Kendrick, R. 1996. Agglutination of Leishmania promastigotes by midgut lectins from various species of phlebotomine sandflies. *Ann Trop Med Parasitol*, **90**, 329–336.

Svobodova, M., Volf, P., and Votypka, J. 2006. Experimental transmission of Leishmania tropica to hyraxes (Procavia capensis) by the bite of Phlebotomus arabicus. *Microbes Infect*, **8**, 1691–1694.

Turco, S. J., Hull, S. R., Orlandi, P. A., Jr., Shepherd, S. D., Homans, S. W., Dwek, R. A., and Rademacher, T. W. 1987. Structure of the major carbohydrate fragment of the *Leishmania donovani* lipophosphoglycan. *Biochemistry*, **26**, 6233–6238.

Valenzuela, J. G., Pham, V. M., Garfield, M. K., Francischetti, I. M., and Ribeiro, J. M. 2002. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. *Insect Biochem Mol Biol*, **32**, 1101–1122.

Volf, P., Kiewegova, A., and Svobodova, M. 1998. Sandfly midgut lectin: Effect of galactosamine on *Leishmania major* infections. *Med Vet Entomol*, **12**, 151–154.

Volf, P. and Killick-Kendrick, R. 1996. Post-engorgement dynamics of haemagglutination activity in the midgut of phlebotomine sandflies. *Med Vet Entomol*, **10**, 247–250.

Volf, P., Killick-Kendrick, R., Bates, P. A., and Molyneux, D. H. 1994. Comparison of the haemagglutination activities in gut and head extracts of various species and geographical populations of phlebotomine sandflies. *Ann Trop Med Parasitol*, **88**, 337–340.

Volf, P. and Myskova, J. 2007. Sand flies and Leishmania: Specific versus permissive vectors. *Trends Parasitol*, **23**, 91–92.

Volf, P. and Palanova, L. 1996. The relationship between protein content of the meal and lectin secretion in the midgut of *Phlebotomus duboscqi*. *Ann Trop Med Parasitol*, **90**, 567–570.

Volf, P., Svobodova, M., and Dvorakova, E. 2001. Bloodmeal digestion and *Leishmania major* infections in *Phlebotomus duboscqi*: Effect of carbohydrates inhibiting midgut lectin activity. *Med Vet Entomol*, **15**, 281–286.

Wallbanks, K. R., Ingram, G. A., and Molyneux, D. H. 1986. The agglutination of erythrocytes and Leishmania parasites by sandfly gut extracts: Evidence for lectin activity. *Trop Med Parasitol*, **37**, 409–413.

Walters, L. L. 1993. Leishmania differentiation in natural and unnatural sand fly hosts. *J Eukaryot Microbiol*, **40**, 196–206.

Young, A. R. and Meeusen, E. N. 2004. Galectins in parasite infection and allergic inflammation. *Glycoconj J*, **19**, 601–606.

34 Ficolins: The Structural Basis for Recognition Plasticity

Misao Matsushita, Yuichi Endo, and Teizo Fujita

CONTENTS

34.1 INTRODUCTION

In 1993, Ichijo et al. cloned and characterized new transforming growth factor-β1 (TGF-β1)-binding proteins from porcine uterus membranes [1]. The proteins that they termed "ficolin" had both collagen-like and fibrinogen-like domains. Since then, many proteins possessing these structural characteristics have been identified in both vertebrates and invertebrates at the protein and cDNA level (Table 34.1). Ficolins are now recognized as a family of proteins that have both collagen-like and fibrinogen-like domains, [2–6] and their mRNA has been identified in a variety of tissues. To date, the majority of ficolins characterized at the protein level are lectins that have a common binding specificity for *N*-acetyl group-containing sugar such as *N*-acetylglucosamine (GlcNAc). Ficolins that are present in serum might have a crucial role in innate immunity by activating the complement system via the lectin pathway in a manner similar to mannose-binding lectin (MBL) of the collectin family [7]. It is possible that the nonserum ficolins perform a similar role to the serum ficolins.

34.2 HUMAN FICOLINS

In humans, there are three kinds of ficolins: L-ficolin, H-ficolin, and M-ficolin, which will be discussed in the following sections.

TABLE 34.1
Characteristics of Ficolins

Species	Ficolins	mRNA Expression	Protein Identified	Binding Substance	Function	Gene Locus	Number of Exons
Human	L-ficolin	Liver	Serum/plasma	GlcNAc (acetyl group) elastin corticosteroid	Complement activation opsonin	9q34	8
	H-ficolin	Liver Lung Glioma cell	Serum/plasma Bile duct Bronchus, alveolus	GlcNAc GalNAc Fucose	Complement activation	1p35.3	8
	M-ficolin	Lung Monocyte Spleen	Neutrophil Monocyte Alveolar epithelial cell	GlcNAc	Complement activation Phagocytic receptor	9q34	9
Mouse	Ficolin-A	Liver Spleen	Serum/plasma	GlcNAc, GalNAc Elastin	Complement activation	2A3	10
	Ficolin-B	Bone marrow Spleen	Macrophage	GlcNAc, GalNAc Sialic acid	ND	2A3	9
Pig	Ficolin α	Liver Lung Bone marrow	Serum/plasma	GlcNAc Elastin TGF-β1	ND		
	Ficolin β	Neutrophil Bone marrow	Neutrophil	TGF-β1	ND		
Hedgehog	Erinacin	ND	Muscle	Metalloprotease	Antihemorrhagic		
Xenopus	XeFCN1	Liver, spleen heart	Serum	GlcNAc, GalNAc	ND		
	XeFCN2	Lung, spleen leukocyte	NI	ND	ND		
	XeFCN3	ND	NI	ND	ND		
	XeFCN4	Lung, spleen	NI	ND	ND		
Ascidian	p40	Hepatopancreas	hemolymph plasma	GlcNAc, GalNAc	ND		
	p50	Hepatopancreas	hemolymph plasma	GlcNAc	ND		

Note: ND, not determined; NI, not identified.

34.2.1 L-Ficolin

L-ficolin (synonymous with L-ficolin/P35, ficolin 2, ficolin L, EBP-37, or hucolin) is an oligomeric protein consisting of 35 kDa subunits [8,9]. Each subunit is composed of four domains: an NH_2-terminl region having two cysteine residues (Cys7 and Cys27); a collagen-like sequence; a short segment (neck domain); and the COOH-terminus of a fibrinogen-like domain (Figure 34.1A). The fibrinogen-like domain forms a globular structure and the overall structure of L-ficolin looks

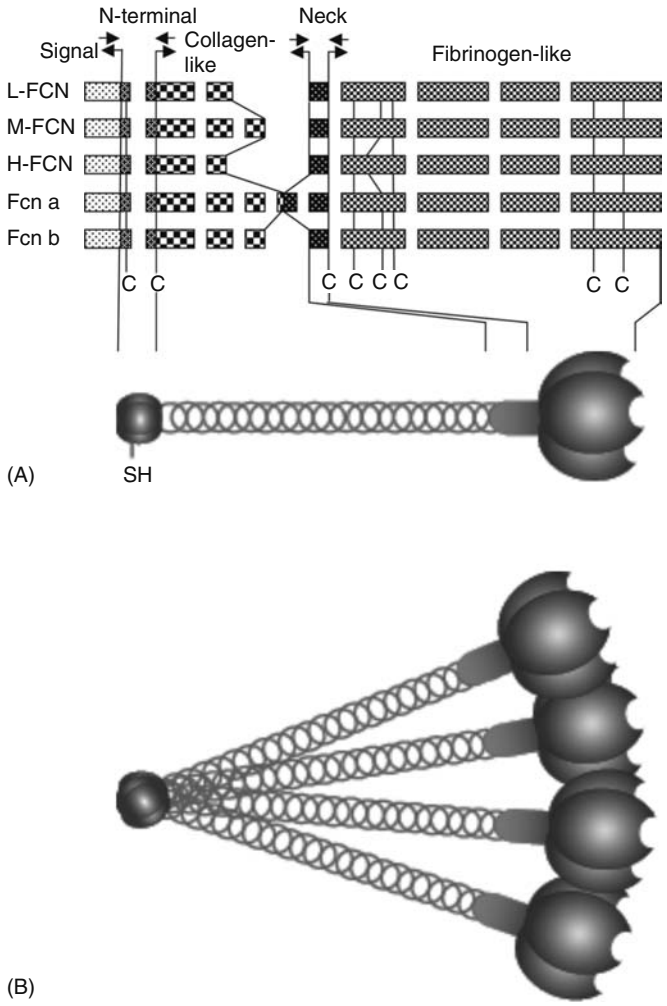


FIGURE 34.1 (See CD for color figure.) (A) Exon organization of the genes encoding human and mouse ficolin and the exon-encoding domain structure of ficolin. The boxes depict the exons of the L-ficolin (L-FCN), M-ficolin (M-FCN), H-ficolin (H-FCN), mouse ficolin A (Fcn a), and mouse ficolin B (Fcn b) genes. The letter C denotes the conserved cysteine residues in ficolin. (B) The tetrameric structure of ficolin.

like a "bouquet." Based on electron microscopy studies of plasma L-ficolin and analyses of recombinant L-ficolin, a tetramer structure that consists of four triple helices formed by 12 subunits has been proposed for L-ficolin (Figure 34.1B) [6,10,11]. The oligomeric structure of L-ficolin is formed by the cross-linking of subunits via disulfide bridges that involve Cys7 and Cys27 residues [11]. The L-ficolin gene (*FCN2*) is located on chromosome 9q34 and contains eight exons [12]: the first exon encodes the 5′ untranslated region, the signal peptide, and nine NH_2-terminal amino acids; the second and third exons encode the collagen-like domain; the fourth exon encodes the neck domain; exons five to seven encode the upstream portion of the fibrinogen-like domain; and the last exon encodes the remaining fibrinogen-like domain and 3′ untranslated region. L-ficolin mRNA is expressed mainly in liver and its protein product is secreted in the serum. The serum concentration of L-ficolin varies in normal adults, and the average value is approximately 3 μg/mL [13–19], whereas it is lower in umbilical cord blood (median 2.5 μg/mL).

L-ficolin exhibits a lectin activity toward GlcNAc. It has been shown that L-ficolin recognizes an acetyl group in GlcNAc [20], suggesting that L-ficolin is able to bind the GlcNAc residue next

to galactose at the nonreducing terminal of complex-type oligosaccharides [8]. L-ficolin also recognizes β-(1.3)-D-glucan [21]. The fibrinogen-like domain of L-ficolin is responsible for lectin activity [13]. Two single nucleotide polymorphisms (SNPs) have been found in exon 8 of *FCN2*, which lead to the substitution of threonine by methionine at codon 236 or to the substitution of alanine by serine at codon 258 [19,22]; these SNPs are associated with decreased and increased GlcNAc binding, respectively, when compared to wild-type L-ficolin. The polymorphisms at positions –986, –602, and –64 in the promoter region of *FCN2* are associated with marked changes in L-ficolin serum concentrations. L-ficolin binds to *Salmonella typhimurium* TV119 [8], *Escherichia coli*, several capsulated *Staphylococcus aureus* serotypes (serotypes T-1, T-8, T-9, T-11, and T-12), and capsulated *Streptococcus pneumoniae* serotypes (11A, 11D, and 11F) but L-ficolin does not bind to noncapsulated strains of bacteria [18]. In addition to the lectin activity, L-ficolin binds elastin [23]and corticosteroid [24].

The acetyl group-recognizing tachylectins 5A and 5B (TL5A and TL5B), which are found in the hemolymph plasma of the horseshoe crab *Tachypleus trimentatus*, consist of an NH_2-terminal Cys-containing segment and a COOH-terminal fibrinogen-like domain, which, to date, have the highest sequence identity to ficolins [25]. Based on the crystal structure of TL5A, the *N*-acetyl binding site is adjacent to and coordinated with the Ca^{2+}-binding site [25]. However, the x-ray structure of the fibrinogen-like domain of L-ficolin has revealed that L-ficolin has several carbohydrate-binding sites, including a site for GlcNAc-binding, suggesting a versatile recognition property of L-ficolin [26]. Interestingly, the GlcNAc-binding site on L-ficolin is different from that of TL5A despite the presence of a Ca^{2+} binding site on L-ficolin which is homologous to that of TL5A.

34.2.2 H-Ficolin

H-ficolin (synonymous with Hakata antigen, ficolin H, ficolin 3, or β2 thermolabile macroglycoprotein) was first identified as a serum antigen that is recognized by an autoantibody present in patients with systemic lupus erythematosus [27]. H-ficolin is an oligomer of 34 kDa subunits, which each have the same domain organization as L-ficolin. As with L-ficolin, the subunits are linked by disulfide bridges but under nonreducing conditions, H-ficolin resolves in more than 10 bands by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) [28]. However, based on electron microscopy studies, H-ficolin is thought to have a hexameric structure [29]. The H-ficolin gene (*FCN3*) is located on chromosome 1p35.3 and consists of eight exons whose organization is similar to that of L-ficolin (Figure 34.1A). H-ficolin mRNA is expressed in the liver and lung. In the liver, H-ficolin is produced by bile duct epithelial cells and hepatocytes and is secreted in bile and serum [30]. The concentration of H-ficolin in serum from normal individuals ranges from 7 to 23 μg/mL. In the lung, H-ficolin is produced by both ciliated bronchial epithelial cells and Type II alveolar epithelial cells and is secreted into the bronchus and alveolus. H-ficolin is also produced in glioma cell lines [31]. H-ficolin has been shown to bind GlcNAc, *N*-acetylgalactosamine (GalNAc), and fucose, but not to mannose and lactose [29]. It appears, however, that the binding affinity of H-ficolin for GlcNAc is very weak compared to that of L-ficolin [32]. H-ficolin agglutinates human erythrocytes coated with lipopolysaccharide derived from *S. typhimurium*, *Salmonella minnesota*, and *E. coli* (O111) [29]. H-ficolin binds to *Aerococcus viridans* and its polysaccharide preparations.

34.2.3 M-Ficolin

M-ficolin (synonymous with L-ficolin/P35-related protein, ficolin M, or ficolin 1) mRNA is expressed in monocytes, the lung, and the spleen [12,33–35]. The M-ficolin gene (*FCN1*) is located on chromosome 9q34, as with L-ficolin, and its exon organization resembles that of L-ficolin [12]. M-ficolin has an extra exon encoding an additional segment of four Gly-Xaa-Yaa repeats (Figure 34.1A). M-ficolin has been reported to be expressed on the surface of peripheral blood monocytes and promonocytic U937 cells [35]. The recombinant fibrinogen-like domain of M-ficolin has affinity

for GlcNAc and an antibody against the recombinant protein inhibits phagocytosis of *E. coli* by U937 cells, suggesting that M-ficolin might play a role in innate immunity by acting as a lectin-like phagocytic receptor for microbial pathogens. In contrast, M-ficolin has been shown to be located in secretory granules in the cytoplasm of peripheral neutrophils and monocytes, and in type II, alveolar epithelial cells in the lung, suggesting that M-ficolin is a secretory protein [36]. Recombinant M-ficolin exhibits lectin activity toward acetylated compounds including GlcNAc, GalNAc, and sialic acid [36,37], and binds to *S. aureus* and *S. typhimurium* LT2.

34.3 FICOLINS OF NONHUMAN SPECIES

34.3.1 Mouse

Two types of ficolins (ficolin-A and ficolin-B) are present in mice. Ficolin-A mRNA is highly expressed in the liver and spleen. In the liver, expression of ficolin-A mRNA increases during mouse development, peaking around birth but declining slightly in adulthood. Ficolin-A mRNA has been detected in macrophages obtained from the adult mouse liver and spleen. The ficolin-A protein is present in plasma as a GlcNAc-binding lectin and is a tetramer with 12 subunits, as with human L-ficolin [10]. cDNA cloning has revealed the presence of a ficolin-A variant that has a shorter collagen-like domain and a longer gap sequence that is generated from the ficolin A gene by alternative splicing [38]. Ficolin-B is expressed in bone marrow and spleen [10]. The expression of ficolin-B mRNA increases during mouse development, peaking 2 or 3 days before birth. Ficolin-B mRNA in the liver declines to undetectable levels by 4 weeks of age. The expression of ficolin-B switches from the embryonic liver to postnatal bone marrow (myeloid cell lineage) and spleen [39]. The ficolin-B protein has been reported to be localized in lysosomes of activated macrophages [40].

The ficolin-A gene is located some distance from the ficolin-B gene on chromosome 2A3. The ficolin-B gene is located in a region homologous to the human L-ficolin and M-ficolin genes locus. The ficolin-A gene consists of 10 exons (Figure 34.1A); the first exon encodes a signal peptide; exons 2 to 4 encode a region upstream of the collagen-like domain; the fifth exon encodes additional Gly-X-Y repeats and a neck domain; the sixth exon encodes an additional neck domain; and exons 7 to 10 encode the fibrinogen-like domain. The ficolin-B gene contains nine exons that are organized in a similar manner to ficolin-A, except that ficolin-B lacks the fifth exon that is found in the ficolin-A gene (Figure 34.1A). The exon organization of the ficolin-B gene is very similar to that of the M-ficolin gene. The mouse homolog of the H-ficolin gene exists as a pseudogene on chromosome 4 [41] Recombinant mouse ficolin-A, its splice variant, and ficolin-B exhibit a lectin activity toward GlcNAc and GalNAc, and, as with human M-ficolin, murine ficolin-B also recognizes sialic acid [38]. Recombinant mouse ficolin-A and ficolin-B bind to *S. aureus* and the binding is partially inhibited by GlcNAc. As with human L-ficolin, murine ficolin-A also binds to elastin [42].

34.3.2 Pig

As described in Section 34.1, ficolins were first discovered as a TGF-β1-binding protein isolated from porcine uterine membrane preparations [43]. cDNAs encoding two types of ficolin, ficolin α and ficolin β, have been reported [1]. The expression of ficolin α mRNA is high in tissue from the lung, liver, and bone marrow, but is low in uterus. Ficolin is present in porcine plasma as a GlcNAc-binding protein with binding specificity for *N*-acetyl groups [44–47]. There are two species of ficolin α in porcine plasma: one forms a tetramer of 12 subunits that has a high affinity for GlcNAc [47] and the other forms a dimer consisting of 24 subunits that exhibits a low to negligible affinity for GlcNAc. Based on the structural analyses of recombinant ficolin α, a model has been proposed in which two cysteine residues in the N-terminal are involved in interchain disulfide bridges, resulting in a tetrameric structure, as with human L-ficolin [48]. Ficolin α binds to

Actinobacillus pleuropneumoniae serotype 5B, *E. coli* O26, *S. aureus, Bacillus cereus, Bacillus subtilis, Enterococcus faecium*, and *Corynebacterum bovis* [44–46]. It also binds to lipopolysaccharides from Gram-negative bacteria (*E. coli*, *S. typhimurium*, *Salmonella enteritidis*, *Salmonella abortus equi*, *Pseudomonas aeruginosa*, *Shigella flexeneri*, *and Serratia marcescens*) and with lipoteichoic acid from Gram-positive bacteria (*Streptococcus sanguis, B. subtilis, Streptococcus pyogenes, and S. aureus*) [46]. Recombinant porcine ficolin binds to elastin, as does human L-ficolin and murine ficolin α [49]. Ficolin β mRNA is weakly expressed in bone marrow and neutrophils, but is not found in the uterus. Ficolin β is located in the membrane and cytoplasmic fractions of nonactivated neutrophils [50]. Upon activation with phorbol myristate acetate, the majority of ficolin β is secreted.

Based on the mRNA expression and the location of human, murine, and porcine ficolins, these proteins are classified into two types: the serum type and the nonserum type. Serum-type ficolins (human L-ficolin, H-ficolin, mouse ficolin-A, and porcine ficolin α) are expressed mainly in the liver and are secreted in the serum [51]. The nonserum-type ficolins (human M-ficolin, mouse ficolin B, and porcine ficolin β) are expressed mainly in nonhepatic tissues, such as peripheral leukocytes and bone marrow.

34.3.3 Hedgehog

Erinacin is a protein found in muscle extracts of the European hedgehog (*Erinaceus europaeus*) that consists of two types of subunits in the ratio 1:2 and has a molecular weight of approximately 100 kDa. Erinacin appears to be a member of the ficolin family, given preliminary sequence data, although the complete primary structure is not currently available [52]. Erinacin inhibits a metalloprotease in the venom of *Bothrops jararaca,* and therefore acts as an antihemorrhagic factor. It is unknown whether Erinacin has lectin activity.

34.3.4 Xenopus

Four ficolin cDNAs have been cloned from *Xenopus laevis* [53]. These ficolins, termed *Xenopus* ficolin (XeFCN) 1, 2, 3, and 4, show a 60%–83% identity between one another at the protein level. XeFCN1 is expressed mainly in tissues of the liver, spleen, and heart. Two types of XeFCN1 are found in the serum: one of 37 kDa and another of 40 kDa; each has different N-linked carbohydrates and both specifically recognize GlcNAc and GalNAc. XeFCN2 and XeFCN4 are mainly located in peripheral blood leukocytes, lung, and spleen tissues, whereas XeFCN3 is undetectable by Northern blotting.

34.3.5 Ascidian

cDNAs encoding four types of ficolins (AsFCNs) have been cloned from the hepatopancreas of the solitary ascidian, *Halocynthia roretzi*. The four ascidian ficolins show 63%–93% identity to each other at the protein level. They contain short collagen-like domains with five Gly-Xaa-Yaa repeats and have long segments between the collagen-like domain and the fibrinogen-like domain, when compared with mammalian ficolins. Two types of GlcNAc-binding lectins (p40 and p50) have been isolated from the hemolymph plasma [54]. p40 is a mixture of AsFCN1 and AsFCN2, while p50 is identical to AsFCN3. p40 recognizes *N*-acetyl groups in association with a pyranose ring such as GlcNAc and GalNAc, while p50 recognizes GlcNAc alone.

The gene localization, organization, and phylogenetic tree of the ficolin family suggests that the ficolin-B gene was the ancestral gene, and that the genes encoding the serum-type ficolins, L-ficolin and ficolin-A, evolved from common ficolin-B/M-ficolin lineages [41]. A putative scenario for the evolution of the ficolin family is that a gene duplication event, prior to the emergence of *Xenopus,* generated two prototypes: the H-ficolin and ficolin B/M-ficolin lineages from a common

ancestor. Thereafter, further divergent gene duplications of the ficolin B/M-ficolin lineage took place independently in individual species. As a result, the murine ficolin-A and -B genes and human L- and M-ficolin genes evolved.

34.4 ROLES OF FICOLINS IN INNATE IMMUNITY

Serum ficolins play a crucial role in innate immunity by binding to carbohydrates on the surface of a variety of bacteria in an antibody-independent manner. Human L-ficolin acts as an opsonin and enhances phagocytosis of neutrophils when bound to *S. typhimurium* TV119, an Ra chemotype strain that has a GlcNAc residue at the nonreducing terminus of its lipopolysaccharides [8,15]. Human and murine serum ficolins activate the complement system via the lectin pathway.

The lectin pathway was identified as the third mechanism of complement activation initiated by MBL [55]. MBL is a serum opsonic lectin with binding specificity for mannose, GlcNAc, glucose, and fucose. Human MBL is complexed with three types of serine protease: MASP-1, MASP-2, and MASP-3, and with sMAP (also called MAp19), which is a truncated form of MASP-2 [56]. The three types of MASP share the same domain structure (CUB1, EGF, CUB2, CCP1, CCP2, and serine protease domain). In the lectin pathway, upon binding of the MBL–MASP complexes to carbohydrates on the surface of microorganisms, MASP-2 within the complex becomes activated and acquires proteolytic activity against complement components C4 and C2 [57]. Functions of the other MASPs and sMAP in the complex remain to be elucidated.

As with MBL, L-ficolin and H-ficolin are present in a complex with the three types of MASP and sMAP in human serum [58,59]. MASP-1, MASP-2, MASP-3, and sMAP bind to L-ficolin in the presence of Ca^{2+}. The binding site on sMAP for L-ficolin has been mapped to the distal end of the CUB1 module [60]. L-ficolin–MASP complexes bind to *S. typhimurium* TV119 [58], β-(1.3)-D-glucan [21] and immobilized lipoteichoic acid from clinically important bacteria, including *S. pyogenes*, *Streptococcus agalactiae*, and *S. aureus*, leading to C4 activation [61]. L-ficolin–MASP complexes also bind to capsular polysaccharides of serotype III group B *streptococci*, a common cause of neonatal sepsis and meningitis, which leads to opsonophagocytic killing by neutrophils in the presence of human serum [62]. This opsonophagocytic killing is mediated by lectin pathway activation followed by alternative pathway activation. The H-ficolin–MASP complex binds to polysaccharides from *Aerococcus viridance*, which are coated on microplates, and activates C4 [59]. In addition, the H-ficolin–MASP complex exhibits *in vitro* bactericidal activity against *A. viridans* in the presence of human serum [63]. It has been shown that recombinant M-ficolin forms complexes with MASP-1 and MASP-2 and that M-ficolin–MASP complexes, which had been bound to GlcNAc or acetylated compounds, activates C4 [36]. Recombinant mouse ficolin-A and its variant, but not ficolin B, bind to mouse MASP-2 and sMAP, resulting in complement activation, suggesting that mouse ficolin-A and its variant act as recognition molecules of the lectin pathway.

34.5 ASSOCIATION OF FICOLINS WITH DISEASES

Ficolins might be associated with disease complications and inflammatory conditions. For example, a low serum concentration of L-ficolin has been reported to be associated with certain diseases: patients with hematological malignancies receiving chemotherapy have low L-ficolin levels when compared with healthy controls [16] and an increased incidence of low L-ficolin was found in children with respiratory infection, especially atopic disorders, when compared with healthy controls, suggesting a role for L-ficolin in protection from the microbes that may complicate allergic diseases [17]. Furthermore, L-ficolin concentrations were abnormally low in a handful of patients who experience recurrent miscarriages [14]. Glomerular deposition of L-ficolin along with MBL, MASPs, and C4d was observed in 25% of patients with IgA nephropathy (IgAN) and was associated with more pronounced histologic damage, such as increased mesangial proliferation and extracapillary proliferation, suggesting a

contribution of the lectin pathway in the progression of IgAN [64]. A significant difference in the frequency of −557 and −64 SNPs in the promoter of the *FCN2* gene have been found between HLA-B51 positive and HLA-B51 negative groups of patients with Behcet's disease [65]. Autoantibodies against H-ficolin were found in 4.3% of patients with systemic lupus erythematosus and in 0.3% of patients with other autoimmune diseases, such as chronic glomerulonephritis and primary biliary cirrhosis [66]. The pathophysiological implications of these autoantibodies remain to be elucidated. Lectin pathway activation by H-ficolin was observed in the peritubular capillary in the kidney allograft [67].

34.6 FICOLINS AND APOPTOSIS

Several lines of evidence suggest that ficolins might play a role in apoptosis. Serum L-ficolin has been shown to bind to apoptotic HL60, U937, and Jurkat T cells, while H-ficolin binds to apoptotic Jurkat T cells, but not to apoptotic HL60 and U937 [68]. Deposition of C4 and C3 on the apoptotic cells was observed. Purified L-ficolin and H-ficolin also bound to the apoptotic cells, and to cC1q receptor/calreticulin. Moreover, L-ficolin has been shown to bind to late apoptotic Jurkat T cells, apoptotic bodies, and necrotic cells, but not to early apoptotic cells [69] and the binding of L-ficolin to apoptotic cells was not inhibited by GlcNAc. L-ficolin has been shown to bind to DNA in a calcium-dependent manner and DNA inhibits the binding of L-ficolin to late apoptotic and necrotic cells, indicating that DNA but not GlcNAc is a ligand for L-ficolin. C4 deposition has been observed on necrotic cells and opsonization has been shown to enhance the attachment/uptake of necrotic cells by macrophages. Taken together, these findings suggest that ficolins participate in the clearance of dying host cells by activating complement to maintain homeostasis.

34.7 CONCLUSION AND FUTURE DIRECTIONS

Serum ficolins are important as pattern recognition molecules in innate immunity in both vertebrates and invertebrates. In addition to innate immunity, they may also play a role in apoptosis. Conversely, the functions of the nonserum ficolins are not fully understood. It is possible, however, that the major function of nonserum ficolins is associated with innate immunity. Although the binding specificity for carbohydrates varies within the ficolin family, it is likely that the *N*-acetyl group is a ligand for the majority of ficolins. GlcNAc, a carbohydrate that contains an *N*-acetyl group, is widely distributed on the surface of microbes. Therefore, ficolins are able to bind and eliminate a variety of microbial pathogens bearing GlcNAc by direct opsonic activity and activation of the lectin complement pathway. The function of serum ficolins also extends to MBL, suggesting that these proteins have a similar role in innate immunity.

REFERENCES

1. Ichijo, H. et al., Molecular cloning and characterization of ficolin, a multimeric protein with fibrinogen- and collagen-like domains, *J. Biol. Chem*., 268, 14505, 1993.
2. Lu, J. and Le, Y., Ficolins and the fibrinogen-like domain, *Immunobiology*, 199, 190, 1998.
3. Matsushita, M. and Fujita, T., Ficolins and the lectin complement pathway, *Immunol. Rev.*, 180, 78, 2001.
4. Matsushita, M. et al., Activation of the lectin complement pathway by ficolins, *Int. Immunopharmacol*., 1, 359, 2001.
5. Lu, J. et al., Collectins and ficolins: Sugar pattern recognition molecules of the mammalian innate immune system, *Biochim. Biophys. Acta*, 1572, 387, 2002.
6. Holmskov, U., Thiel, S. and Jensenius, J. C., Collections and ficolins: Humoral lectins of the innate immune defense, *Annu. Rev. Immunol*., 21, 547, 2003.
7. Fujita, T., Evolution of the lectin-complement pathway and its role in innate immunity, *Nat. Rev. Immunol*., 2, 346, 2002.

8. Matsushita, M. et al., A novel human serum lectin with collagen- and fibrinogen-like domains that functions as an opsonin, *J. Biol. Chem*., 271, 2448, 1996.
9. Le, Y. et al., Purification and binding properties of a human ficolin-like protein, *J. Immunol. Methods*, 204, 43, 1997.
10. Ohashi, T. and Erickson, H. P., Oligomeric structure and tissue distribution of ficolins from mouse, pig and human, *Arch. Biochem. Biophys*., 360, 223, 1998.
11. Hummelshoj, T. et al., Molecular organization of human Ficolin-2, *Mol. Immunol*., 44, 401, 2007.
12. Endo, Y. et al., Cloning and characterization of the human lectin P35 gene and its related gene, *Genomics*, 36, 515, 1996.
13. Le, Y. et al., Human L-ficolin: Plasma levels, sugar specificity, and assignment of its lectin activity to the fibrinogen-like (FBG) domain, *FEBS Lett*., 425, 367, 1998.
14. Kilpatrick, D. C., Fujita, T., and Matsushita, M., P35, an opsonic lectin of the ficolin family, in human blood from neonates, normal adults, and recurrent miscarriage patients, *Immunol. Lett*., 67, 109, 1999.
15. Taira, S. et al., Opsonic function and concentration of human serum ficolin/P35, *Fukushima J. Med. Sci*., 46, 13, 2000.
16. Kilpatrick, D. C. et al., No strong relationship between mannan binding lectin or plasma ficolins and chemotherapy-related infections, *Clin. Exp. Immunol*., 134, 279, 2003.
17. Atkinson, A. P. et al., L-ficolin in children with recurrent respiratory infections, *Clin. Exp. Immunol*., 138, 517, 2004.
18. Krarup, A. et al., Effect of capsulation of opportunistic pathogenic bacteria on binding of the pattern recognition molecules mannan-binding lectin, L-ficolin, and H-ficolin, *Infect. Immun*., 73, 1052, 2005.
19. Hummelshoj, T. et al., Polymorphisms in the FCN2 gene determine serum variation and function of Ficolin-2, *Hum. Mol. Genet*., 14, 1651, 2005.
20. Krarup, A. et al., L-ficolin is a pattern recognition molecule specific for acetyl groups, *J. Biol. Chem*., 279, 47513, 2004.
21. Ma, Y. G. et al., Human mannose-binding lectin and L-ficolin function as specific pattern recognition proteins in the lectin activation pathway of complement, *J. Biol. Chem*., 279, 25307, 2004.
22. Herpers, B. L. et al., Coding and non-coding polymorphisms in the lectin pathway activator L-ficolin gene in 188 Dutch blood bank donors, *Mol. Immunol*., 43, 851, 2006.
23. Harumiya, S. et al., EBP-37, a new elastin-binding protein in human plasma: Structural similarity to ficolins, transforming growth factor-beta 1-binding proteins, *J. Biochem. (Tokyo)*, 117, 1029, 1995.
24. Edgar, P. F., Hucolin, a new corticosteroid-binding protein from human plasma with structural similarities to ficolins, transforming growth factor-beta 1-binding proteins, *FEBS Lett*., 375, 159, 1995.
25. Gokudan, S. et al., Horseshoe crab acetyl group-recognizing lectins involved in innate immunity are structurally related to fibrinogen, *Proc. Natl. Acad. Sci. USA*, 96, 10086, 1999.
26. Gaboriauda, C. et al., X-ray structure of the trimeric fibrinogen-like domain of human L-ficolin and analysis of its carbohydrate binding site, *Mol. Immunol*., 41, 231, 2004.
27. Inaba, S. and Okochi, K., On a new precipitating antibody against normal human serum found in two patients with SLE (In Japanese), *Igaku No Ayumi* (in Japanese), 107, 690, 1978.
28. Yae, Y. et al., Isolation and characterization of a thermolabile beta-2 macroglycoprotein (thermolabile substance or Hakata antigen) detected by precipitating (auto) antibody in sera of patients with systemic lupus erythematosus, *Biochim. Biophys. Acta*, 1078, 369, 1991.
29. Sugimoto, R. et al., Cloning and characterization of the Hakata antigen, a member of the ficolin/opsonin p35 lectin family, *J. Biol. Chem*., 273, 20721, 1998.
30. Akaiwa, M. et al., Hakata antigen, a new member of the ficolin/opsonin p35 family, is a novel human lectin secreted into bronchus/alveolus and bile, *J. Histochem. Cytochem*., 47, 777, 1999.
31. Kuraya, M. et al., Expression of H-ficolin/Hakata antigen, mannose-binding lectin-associated serine protease (MASP)-1 and MASP-3 by human glioma cell line T98G, *Int. Immunol*., 15, 109, 2003.
32. Matsushita, M. and Fujita, T., The role of ficolins in innate immunity, *Immunobiology*, 205, 490, 2002.
33. Lu, J. et al., Human ficolin: cDNA cloning, demonstration of peripheral blood leucocytes as the major site of synthesis and assignment of the gene to chromosome 9, *Biochem. J*., 313, 473, 1996.
34. Lu, J. et al., Biosynthesis of human ficolin, an *Escherichia coli*-binding protein, by monocytes: Comparison with the synthesis of two macrophage-specific proteins, C1q and the mannose receptor, *Immunology*, 89, 289, 1996.
35. Teh, C. et al., M-ficolin is expressed on monocytes and is a lectin binding to *N*-acetyl-D-glucosamine and mediates monocyte adhesion and phagocytosis of *Escherichia coli*, *Immunology*, 101, 225, 2000.

36. Liu, Y. et al., Human M-ficolin is a secretory protein that activates the lectin complement pathway, *J. Immunol*., 175, 3150, 2005.
37. Frederiksen, P. D. et al., M-ficolin, an innate immune defence molecule, binds patterns of acetyl groups and activates complement, *Scand. J. Immunol*., 62, 462, 2005.
38. Endo, Y. et al., Carbohydrate-binding specificities of mouse ficolin A, a splicing variant of ficolin A and ficolin B and their complex formation with MASP-2 and sMAP, *Immunogenetics*, 57, 837, 2005.
39. Liu, Y. et al., Ficolin A and ficolin B are expressed in distinct ontogenic patterns and cell types in the mouse, *Mol. Immunol*., 42, 1265, 2005.
40. Runza, V. L. et al., Localization of the mouse defense lectin ficolin B in lysosomes of activated macrophages, *J. Endotoxin Res*., 12, 120, 2006.
41. Endo, Y. et al., Identification of the mouse H-ficolin gene as a pseudogene and orthology between mouse ficolins A/B and human L-/M-ficolins, *Genomics*, 84, 737, 2004.
42. Fujimori, Y. et al., Molecular cloning and characterization of mouse ficolin-A, *Biochem. Biophys. Res. Commun*., 244, 796, 1998.
43. Ichijo, H. et al., Purification of transforming growth factor-beta 1 binding proteins from porcine uterus membranes, *J. Biol. Chem*., 266, 22459, 1991.
44. Brooks, A. S., DeLay, J. P., and Hayes, M. A., Purification and binding properties of porcine plasma ficolin that binds *Actinobacillus pleuropneumoniae*, *Dev. Comp. Immunol*., 27, 835, 2003.
45. Brooks, A. S., DeLay, J. P. and Hayes, M. A., Characterization of porcine plasma ficolins that bind *Actinobacillus pleuropneumoniae* serotype 5B, *Immunobiology*, 207, 327, 2003.
46. Nahid, A. M. and Sugii, S., Binding of porcine ficolin-alpha to lipopolysaccharides from Gram-negative bacteria and lipoteichoic acids from Gram-positive bacteria, *Dev. Comp. Immunol*., 30, 335, 2006.
47. Ohashi, T. and Erickson, H. P., Two oligomeric forms of plasma ficolin have differential lectin activity, *J. Biol. Chem*., 272, 14220, 1997.
48. Ohashi, T. and Erickson, H. P., The disulfide bonding pattern in ficolin multimers, *J. Biol. Chem*., 279, 6534, 2004.
49. Harumiya, S. et al., Characterization of ficolins as novel elastin-binding proteins and molecular cloning of human ficolin-1, *J. Biochem. (Tokyo)*, 120, 745, 1996.
50. Brooks, A. S. et al., Expression and secretion of ficolin beta by porcine neutrophils, *Biochim. Biophys. Acta*, 1624, 36, 2003.
51. Endo, Y., Matsushita, M., and Fujita, T., Role of ficolin in innate immunity and its molecular basis. *Immunobiology*, 212, 371, 2007.
52. Omori-Satoh, T., Yamakawa, Y., and Mebs, D., The antihemorrhagic factor, erinacin, from the European hedgehog (*Erinaceus europaeus*), a metalloprotease inhibitor of large molecular size possessing ficolin/opsonin P35 lectin domains, *Toxicon*, 38, 1561, 2000.
53. Kakinuma, Y. et al., Molecular cloning and characterization of novel ficolins from *Xenopus laevis*, *Immunogenetics*, 55, 29, 2003.
54. Kenjo, A. et al., Cloning and characterization of novel ficolins from the solitary ascidian, *Halocynthia roretzi*, *J. Biol. Chem*., 276, 19959, 2001.
55. Matsushita, M., The lectin pathway of the complement system, *Microbiol. Immunol*., 40, 887, 1996.
56. Sorensen, R., Thiel, S., and Jensenius, J. C., Mannan-binding-lectin-associated serine proteases, characteristics and disease associations, *Springer Semin. Immunopathol*., 27, 299, 2005.
57. Fujita, T., Matsushita, M., and Endo, Y., The lectin-complement pathway—its role in innate immunity and evolution, *Immunol. Rev*., 198, 185, 2004.
58. Matsushita, M., Endo, Y. and Fujita, T., Cutting edge: complement-activating complex of ficolin and mannose-binding lectin-associated serine protease, *J. Immunol*., 164, 2281, 2000.
59. Matsushita, M. et al., Activation of the lectin complement pathway by H-ficolin (Hakata antigen), *J. Immunol*., 168, 3502, 2002.
60. Gregory, L. A. et al., The X-ray structure of human mannan-binding lectin-associated protein 19 (MAp19) and its interaction site with mannan-binding lectin and L-ficolin, *J. Biol. Chem*., 279, 29391, 2004.
61. Lynch, N. J. et al., L-ficolin specifically binds to lipoteichoic acid, a cell wall constituent of Gram-positive bacteria, and activates the lectin pathway of complement, *J. Immunol*., 172, 1198, 2004.
62. Aoyagi, Y. et al., Role of L-ficolin/mannose-binding lectin-associated serine protease complexes in the opsonophagocytosis of type III group B streptococci, *J. Immunol*., 174, 418, 2005.
63. Tsujimura, M. et al., Serum concentration of Hakata antigen, a member of the ficolins, is linked with inhibition of *Aerococcus viridans* growth, *Clin. Chim. Acta*, 325, 139, 2002.

64. Roos, A. et al., Glomerular activation of the lectin pathway of complement in IgA nephropathy is associated with more severe renal disease, *J. Am. Soc. Nephrol*., 17, 1724, 2006.
65. Chen, X. et al., Single nucleotide polymorphisms of Ficolin 2 gene in Behcet's disease, *J. Dermatol. Sci*., 43, 201, 2006.
66. Inaba, S., Okochi, Y. and Yae, Y., Serological studies of an SLE-associated antigen-antibody system discovered as a precipitation reaction in agarose gel: the HAKATA antigen-antibody system, *Fukuoka Acta Medica*, 81, 284, 1990.
67. Imai, N. et al., Immunohistochemical evidence of activated lectin pathway in kidney allografts with peritubular capillary C4d deposition, *Nephrol. Dial. Transplant*., 21, 2589, 2006.
68. Kuraya, M. et al., Specific binding of L-ficolin and H-ficolin to apoptotic cells leads to complement activation, *Immunobiology*, 209, 689, 2005.
69. Jensen, M. L. et al., Ficolin-2 recognizes DNA and participates in the clearance of dying host cells, *Mol. Immunol*., 44, 401, 2007.

35 Hemolytic Lectin in Marine Invertebrates

Tomomitsu Hatakeyama

CONTENTS

35.1 INTRODUCTION

A number of Ca^{2+}-dependent lectins have been found in various invertebrates. Many of them belong to C-type lectin family, which contain C-type carbohydrate recognition domains (CRDs) [1,2], and assumed to be involved in defense against foreign pathogens [3]. Four Ca^{2+}-dependent, Gal/GalNAc-specific lectins (CEL-I, II, III, and IV) have been isolated from the body fluid of the marine invertebrate (sea cucumber), *Cucumaria echinata* [4]. Among these, CEL-I and CEL-IV belong to the C-type lectin family [5,6], while CEL-III is a novel lectin that exhibits strong hemolytic and cytotoxic activities [7,8]. Based on several structural and biochemical analyses, it was subsequently found that CEL-III belongs to ricin-type (R-type) lectin family in spite of the Ca^{2+} dependency and its hemolytic activity is mediated by formation of ion-permeable pores composed of its oligomers in the target cell membranes [9]. Self-oligomerization of CEL-III in the membranes is mediated by conformational changes of the protein, triggered by binding to the cell surface carbohydrate chains. The three-dimensional structure of this lectin, along with its complex with specific carbohydrate, has recently been solved by x-ray crystallographic analysis and has provided important clues to speculate the mechanism of the hemolytic action. This chapter describes characteristic features of CEL-III as a hemolytic lectin and its putative mechanism of hemolytic activity on the basis of the structural information.

35.2 CARBOHYDRATE-BINDING ACTIVITY

Carbohydrate-binding specificity of CEL-III was originally estimated by inhibition of hemolysis by various carbohydrates [4]. The results were later confirmed by the solid-phase assay using

TABLE 35.1
Comparison of the Affinity of Carbohydrates for CEL-III Measured by the Solid-Phase Binding Assay and Inhibition of Hemolysis

Carbohydrate	Solid-Phase Method (K_i)[a] (mM)	Inhibition of Hemolysis by Carbohydrate[b] (%)
GalNAc	1.8	12.5
Lactose	2.0	18.8
Lactulose	2.1	15.6
Methyl β-galactoside	2.9	28.1
Methyl α-galactoside	6.2	39.1
Fucose	4.0	34.4
Galactose	4.2	40.6
Melibiose	5.5	46.9
Raffinose	4.7	56.3
Mannose	9.7	100
GlcNAc	13.0	—
Glucose	14.7	100

Note: All the carbohydrates used were of the D-configuration.

[a] K_i denotes carbohydrate concentration for 50%-inhibition of binding of CEL-III to lactose-coated microplate wells. (From Hatakeyama, T., Miyamoto, Y., Nagatomo, H., Sallay, I., and Yamasaki, N., *J. Biochem.,* 121, 63, 1997.)

[b] Hemolysis by CEL-III was measured using rabbit erythrocytes. (From Hatakeyama, T., Kohzaki, H., Nagatomo, H., and Yamasaki, N., *J. Biochem.,* 116, 209, 1994.)

carbohydrate-immobilized microplate (Table 35.1) [10]. CEL-III bound *N*-acetyl-D-galactosamine most strongly among simple carbohydrates tested, followed by lactose and lactulose. Preferential affinity for β-galactosides was found from the comparison between methyl α-galactoside and methyl β-galactoside as well as between lactose (Galβ1-4Glc) and melibiose (Galα1-6Glc), although affinity of β-galactosides was only about twice that of α-galactosides. Since fucose (6-deoxy-D-galactose) showed affinity comparable to galactose, the hydroxyl group at 6-position of galactose seems to have little contribution for interaction with CEL-III. This was subsequently confirmed from the three-dimensional structure of CEL-III/carbohydrate complex revealed by x-ray crystallography, in which 6-OH of bound GalNAc was oriented to the opposite direction from the binding site of CEL-III. Ca^{2+} dependency of CEL-III for carbohydrate-binding activity was also examined by the solid-phase carbohydrate-binding assay; the affinity for Ca^{2+} was found to be relatively low (K_d = 2.3 mM), compared with other Ca^{2+}-binding proteins with EF-hand structure, such as calmodulin, which show much higher affinity for Ca^{2+} (K_d = ~10^{-6} M).

35.3 HEMOLYTIC AND CYTOTOXIC ACTIVITIES

CEL-III exhibits strong hemolytic activity toward human and rabbit erythrocyte, whereas susceptibility to hemolysis was much less for horse and sheep erythrocytes, suggesting that the difference in its cell surface receptors (carbohydrate chains) may largely affect the susceptibility of the cells to the hemolytic action [4,7]. The hemolytic activity can be inhibited by specific carbohydrates containing galactose-related carbohydrates, such as GalNAc and lactose, indicating that binding to the carbohydrate chains on erythrocyte surface is a crucial step to the hemolytic action. CEL-III also exhibits strong cytotoxicity for some cultured cell lines. For example, Madin-Darby canine kidney (MDCK) cells showed high susceptibi-lity to cytotoxicity of CEL-III (LD_{50} = 53 ng/mL), whereas no significant cytotoxicity was observed with Chinese hamster ovary (CHO) cells up to 10 μg/mL [8]. The

experiments using FITC-labeled CEL-III demonstrated that the amount of CEL-III bound to MDCK cells was about twofold greater than that to CHO cells. This supports the idea that the activities of CEL-III are largely dependent on its lectin activity to bind to cell surface carbohydrates.

It has been originally assumed that hemolytic action of CEL-III might be caused by some enzymatic activity, such as phospholipase-like activity, after binding to the cell surface carbohydrate chains. However, analysis of the membrane lipids of the erythrocytes treated with CEL-III indicated no detectable difference with those of intact erythrocyte membrane. In addition, hemolytic action of CEL-III was most effective at lower temperatures around 10°C, and decreased markedly with increasing temperature. These suggest that CEL-III causes hemolysis through nonenzymatic action. On the other hand, hemolytic activity of CEL-III was apparently inhibited by soluble polymers, such as dextrans with molecular masses of more than 4000, while leakage of small molecules, such as ATP, was observed in such erythrocytes protected by dextran [7]. This suggests that the CEL-III forms ion-permeable membrane pores, and thereby leading to osmotic rupture of the cells. Other experiments using artificial lipid vesicles also indicated that CEL-III forms membrane pores in the lipid membranes containing glycolipids as receptors [9,11,12]. Such a membrane pore-forming activity has also been known for several bacterial protein toxins [13–15] as well as some invertebrate ones [16,17].

35.4 SELF-OLIGOMERIZATION OF CEL-III

There is a remarkable pH-dependent profile in hemolytic activity of CEL-III; the activity drastically increased with increasing pH from neutral to at pH 10, whereas almost no hemolysis was observed below pH 6.5 (Figure 35.1). In acidic region, only hemagglutination of erythrocytes was observed instead, indicating that the protein still retains carbohydrate-binding ability, but not pore-forming ability. When rabbit erythrocyte membranes treated with CEL-III at different pHs were subjected to immunoblotting analysis using anti-CEL-III antiserum, irreversible binding of the CEL-III oligomer of about 270 kDa was observed in alkaline region, particularly those treated at higher pHs (Figure 35.2) [9], demonstrating that higher pH significantly promoted irreversible binding of CEL-III to the membrane as an oligomeric form. This is in agreement with the increase in hemolytic activity of CEL-III at higher pH values. Self-oligomerization of CEL-III was also observed in the cell membranes

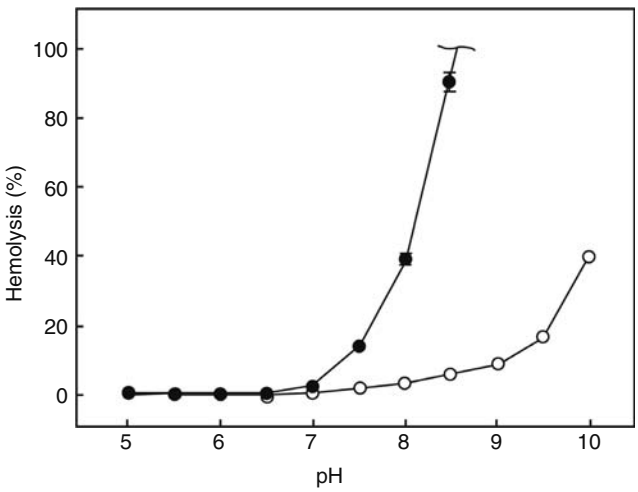


FIGURE 35.1 pH-dependence of the hemolytic activity of CEL-III. Activity measurements were performed in duplicate with 0.5 µg/mL (○) or 5 µg/mL (●) CEL-III and 5% rabbit erythrocytes in the presence of 0.15 M NaCl and 10 mM $CaCl_2$. The value for the erythrocytes lysed with 0.1% Triton X-100 was taken as 100%.

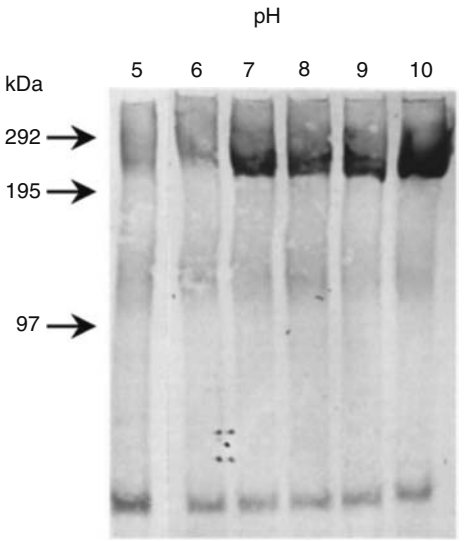


FIGURE 35.2 Immunoblotting of the erythrocyte membranes treated with CEL-III. The membranes were prepared from the CEL-III-treated erythrocytes at indicated pH values and subjected to SDS-PAGE. After electrophoretic transfer to a nitrocellulose membrane, CEL-III was detected with mouse anti-CEL-III antiserum.

of cultured cells that are susceptible to toxicity of CEL-III [18]. These observations indicate that the extent of hemolytic and cytotoxic activities of CEL-III is closely related with the amount of its oligomers formed in the target cell membranes. Therefore, it was concluded that the membrane pores were formed by membrane-inserted CEL-III oligomers, which were presumably composed of six monomers, as estimated from their size on SDS-PAGE. Although CEL-III oligomers form such strongly associated oligomers in the membrane once bound to the carbohydrate chains on susceptible cell, CEL-III monomer shows little tendency to oligomerize in buffer solution at neutral pH values. This suggests that oligomerization of CEL-III molecules is greatly promoted through some structural changes triggered by its binding to target cell surface.

35.5 OLIGOMERIZATION IN SOLUTION

Although CEL-III effectively oligomerizes upon binding to the susceptible cell surface in the alkaline region to lead hemolysis, self-oligomerization was also found to take place in aqueous solution upon binding of specific carbohydrates under high pH and high salt conditions, e.g., at pH 10 and in the presence of 1 M NaCl. The resulting CEL-III oligomer (carbohydrate-induced oligomer) showed a single band on SDS-PAGE, which was similar in size as that formed in the cell membranes treated by CEL-III [9]. This suggests that the formation of the carbohydrate-induced oligomers is based on essentially similar mechanism as that for the oligomer formed in the cell membranes (hemolytic oligomer). On the other hand, the mass of carbohydrate-induced oligomers was determined by small-angle x-ray scattering as more than 1000 kDa (more than 20-mer) [19], which was much larger than that estimated on SDS-PAGE (270 kDa). This suggests that the core oligomers of 270 kDa, which was observed on SDS-PAGE, further associate in aqueous solution by relatively weak hydrophobic interaction. In fact, increase in surface hydrophobicity during carbohydrate-induced oligomerization in solution was observed by the measurement using the fluorescent probe 8-anilino-1-naphthalenesulfonate [9]. Far-UV circular dichroism spectra revealed an increase in β-sheet content in this carbohydrate-induced oligomer, compared with its monomer, suggesting that carbohydrate-induced oligomerization proceeds through conformational change accompanied by an

increase in β-sheet content. Since the binding of specific carbohydrates, especially galactoside containing β1–4 linkage at their nonreducing ends are required for carbohydrate-induced oligomerization in solution, it seems likely that specific carbohydrate chains on the target cell surface also trigger conformational change of CEL-III to induce conformational changes, leading to its membrane insertion. Both carbohydrate-induced oligomer and hemolytic oligomer, which was solubilized using detergent from the erythrocyte membrane, retained carbohydrate-binding ability, whereas they no longer exhibited hemolytic activity. This suggests that insertion of CEL-III molecules into the cell membrane is coupled with self-association process, and once associated, the protein loses the ability to be inserted into cell membrane to form membrane pores.

35.6 AMINO ACID SEQUENCE OF CEL-III

Amino acid sequence of CEL-III was determined from its cDNA and partial amino acid sequences [20]. As shown in Figure 35.3, CEL-III is composed of a single polypeptide with 432 amino acid residues (47.5 kDa). From the database search, sequence homology with the B-chains of the plant toxic lectins, ricin and abrin, was found in its N-terminal two-thirds. The B-chains of ricin and abrin

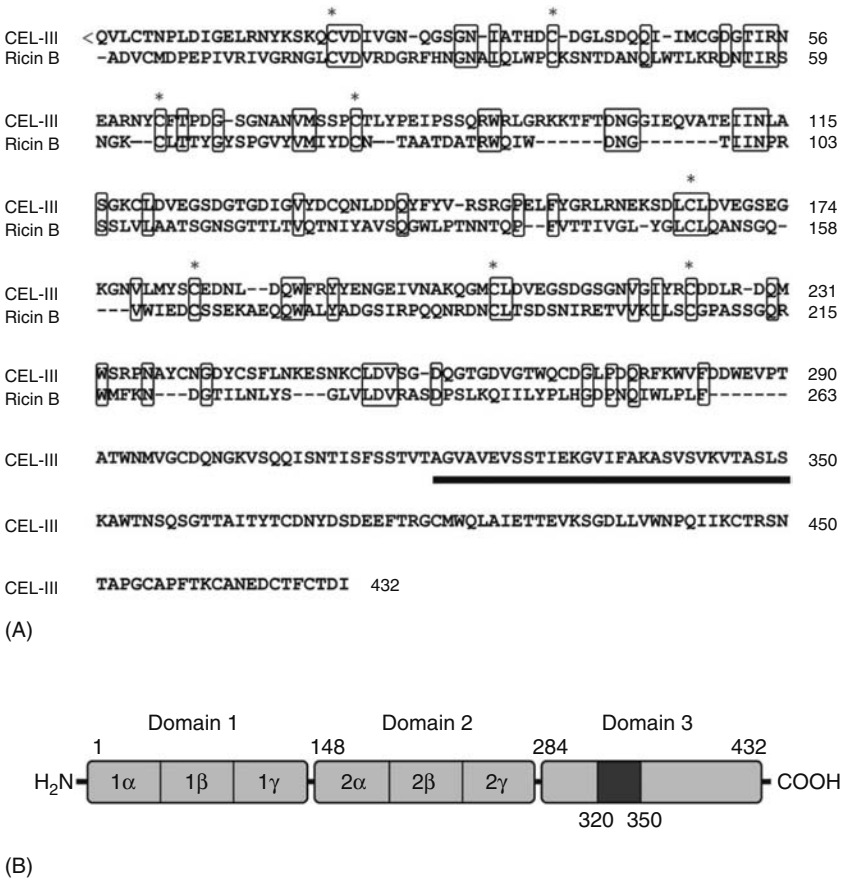


FIGURE 35.3 Comparison of the amino acid sequence of CEL-III with that of ricin B-chain (A), and domain structure of CEL-III (B). (A) Identical amino acids are enclosed in boxes. Half-cystines involved in internal disulfide bonds in ricin B-chain and corresponding residues in CEL-III are marked by asterisks. Hydrophobic region is indicated by a horizontal bar. N-terminal pyroglutamic acid of CEL-III is indicated by "<". (B) Domains 1 and 2 are carbohydrate-binding domains, each composed of three subdomains. Hydrophobic region (residues 320–350) in domain 3 is shown in dark gray.

are known to be lectin subunits, which are composed of two carbohydrate-binding domains, while their A-chains exert N-glycosidase activity and inactivate 60S subunits of ribosomes in eukaryotic cells, thereby exhibiting strong cytotoxicity. Sequence identity between domains 1 and 2 of CEL-III and ricin and abrin B-chains was only 24% and 29%, respectively. However, as shown in Figure 35.3A, similarity was particularly seen for eight half-cystines, which are involved in intrachain disulfide bonds in ricin B-chain, suggesting that they adopt similar tertiary structure in spite of the relatively low sequence identity. Ricin and abrin B-chains contain two carbohydrate-binding domains, which are known as β-trefoil domains, showing pseudo threefold symmetry. Each of these two domains is composed of three subdomains that share internal sequence similarity, and this is also the case for CEL-III (subdomains 1α-1γ and 2α-2γ) (Figure 35.3B). However, in contrast to ricin B-chain that shows only little sequence identity among its subdomains, those of CEL-III show fairly high sequence similarities with each other (up to 54% identity between subdomains 2α and 2β), suggesting their more symmetrical β-trefoil structure. This has been confirmed by x-ray crystallographic analysis as described below. In contrast to domains 1 and 2, the C-terminal region (domain 3) shows no homology with other known proteins. Since two distinct functions, i.e., carbohydrate-binding and oligomerization in the cell membrane, should be required for the hemolytic action and there is a hydrophobic segment of residues 320–350 in domain 3 (Figure 35.3A), it was inferred that domain 3 is primarily responsible for oligomerizing ability of CEL-III.

35.7 THREE-DIMENSIONAL STRUCTURE OF CEL-III

Three-dimensional structure of CEL-III has been solved by x-ray crystallographic analysis (Figure 35.4) [21]. The domain structure proposed based on the amino acid sequence was consequently manifested in the crystal structure. Domains 1 and 2 adopt β-trefoil fold, each of which contains three subdomains aligned to give pseudo threefold symmetry (Figure 35.5A). Each subdomain, except for subdomain 1β, contains one Ca^{2+} ion, and there is one Mg^{2+} ion at the center of the psuedo threefold axes. The presence of metal ions in domains 1 and 2 is the most conspicuous difference with other proteins having β-trefoil domains, which generally contain no metal ions (Figure 35.5B). As described below, Ca^{2+} ions are directly involved in binding with carbohydrates, while subdomain 1β, which has no Ca^{2+}, shows no carbohydrate-binding ability. This is in agreement with the fact that subdomain 1β lacks two aspartic acid residues responsible for Ca^{2+}-binding (Figure 35.6).

As seen in the comparison with ricin B-chain (Figure 35.5), pseudo threefold symmetry is more evident in CEL-III compared with that of ricin B-chain. This seems to be closely related to that CEL-III shows carbohydrate-binding ability at five of six subdomains [22], whereas ricin B-chain

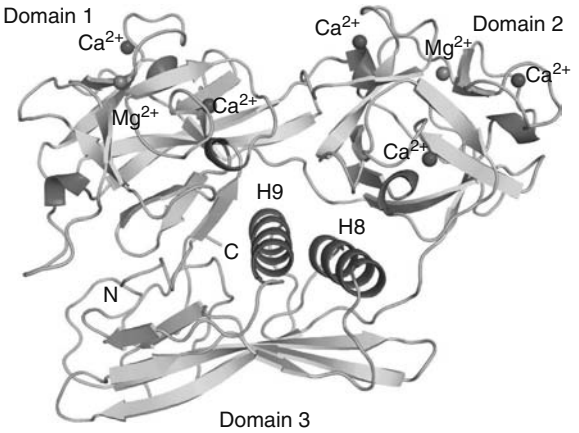


FIGURE 35.4 (See color insert following blank page 170. Also see CD for color figure.) Overall structure of CEL-III depicted as a ribbon model. Bound Mg^{2+} and Ca^{2+} ions are indicated by balls.

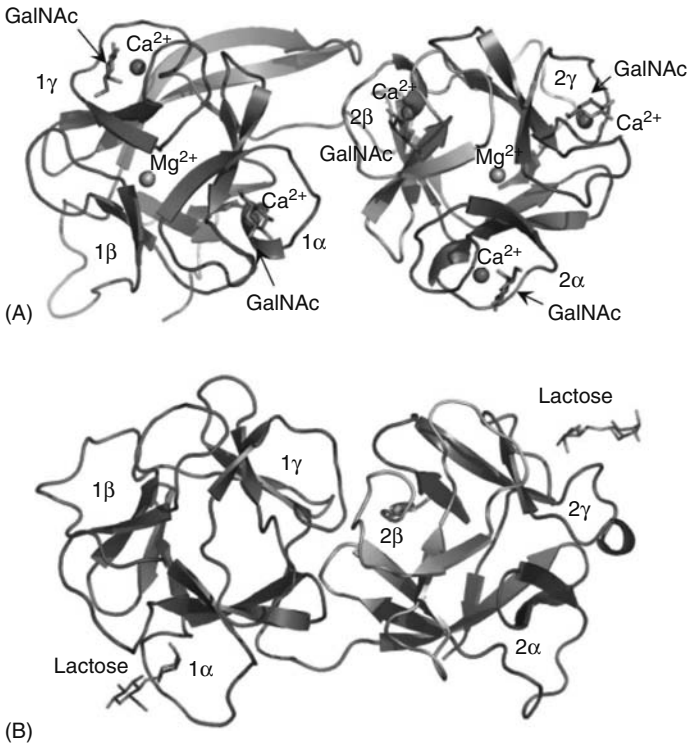


FIGURE 35.5 (See color insert following blank page 170. Also see CD for color figure.) Comparison of β-trefoil domains between CEL-III (A) and ricin B-chain (B). Bound carbohydrates are shown as stick models.

has only two or three active binding sites [23–25]. In contrast to N-terminal two carbohydrate-binding domains, C-terminal domain 3 showed a novel fold with extended β-sheet structure and characteristic two α-helices oriented nearly perpendicular to these β-strands (Figure 35.4). These α-helices correspond to the hydrophobic region of domain 3 (residues 320–350) and situated in the cleft formed between domains 1 and 2.

35.8 CARBOHYDRATE RECOGNITION MECHANISM

Figure 35.7A shows a carbohydrate-binding site (subdomain 1γ) of CEL-III/GalNAc complex [22]. As seen in this figure, 3-OH and 4-OH groups of bound GalNAc form coordinate bonds with a Ca^{2+}

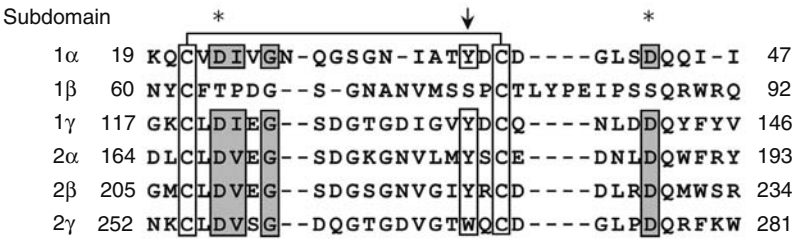


FIGURE 35.6 Sequence comparison of the subdomains of CEL-III. Amino acid residues involved in Ca^{2+}-binding are indicated by shading. Two aspartic acid residues, whose β-carboxyl groups form coordinate bonds with Ca^{2+} are indicated by asterisks. Aromatic residues that make stacking interaction with GalNAc are indicated by an arrow.

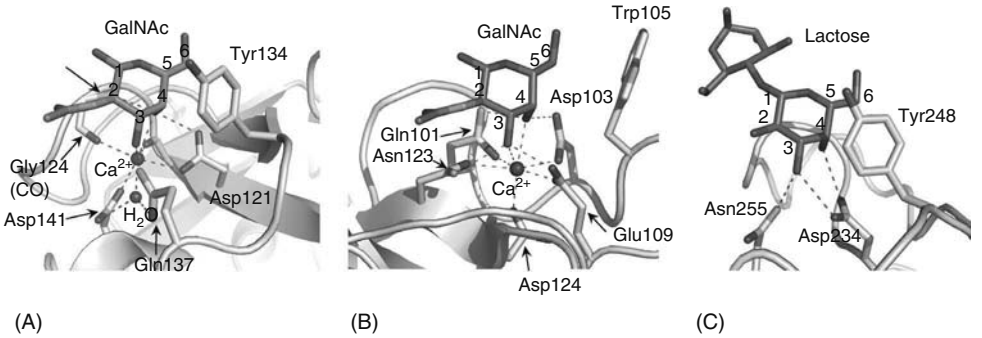


FIGURE 35.7 Comparison of carbohydrate-binding sites of CEL-III (A), a C-type lectin CEL-I (B) (From Sugawara, H., Kusunoki, M., Kurisu, G., Fujimoto, T., Aoyagi, H., and Hatakeyama, T., *J. Biol. Chem*., 279, 45219, 2004), and ricin B-chain (C) (From Rutenber, E., Katzin, B.J., Ernst, S., Collins, E.J., Mlsna, D., Ready, M.P., and Robertus, J.D., *Proteins*, 10, 240, 1991). Hydrogen and coordinate bonds are indicated by dashed lines.

ion. 3-OH group also makes a hydrogen bond with Gln137 which is fixed by a hydrogen bond with a water molecule coordinating the Ca^{2+} ion. On the other hand, 4-OH group forms hydrogen bonds with Asp121 and the amide NH of Gly124. In addition to the coordinate and hydrogen bonds, binding of GalNAc is further stabilized by stacking interaction between the hydrophobic face of GalNAc and side chain of Tyr134. These interactions between bound GalNAc and the protein, including stacking with aromatic residues, Try and Trp (Figure 35.6), are shared in all the binding sites in CEL-III/GalNAc complex. Carbohydrate-binding ability in five subdomains of CEL-III appears to be closely related with their relatively high sequence similarities [20]. This also suggests that CEL-III may conserve more primitive feature of β-trefoil domains, which have been evolved from an ancestral protein module of about 40 residues with carbohydrate-binding ability. The presence of multiple binding sites in a single protein molecule should be advantageous to make substantial binding affinity for carbohydrate chains on the cell surface. It has recently been suggested that the lipid rafts, which contain various glycolipids, function as receptors for several bacterial toxins [26,27]. Since lactosyl ceramide was found to be the most effective receptor glycolipid for CEL-III in human erythrocyte membrane [11], it seems likely that multiple carbohydrate-binding sites are advantageous to recognize glycolipid clusters existing in the lipid rafts on the target cell membrane.

Although carbohydrate-binding domains of CEL-III adopt β-trefoil fold, their Ca^{2+}-dependent carbohydrate recognition mechanism resembles those for C-type CRDs [28] in spite of the lack of structural similarity. In both cases, OH groups of bound carbohydrates form coordinate bonds with Ca^{2+} as well as hydrogen bond networks with the nearby amino acid side chains (Figure 35.7A and B). On the other hand, the orientation of the carbohydrate bound to CEL-III is very similar to that for ricin B-chain (Figure 35.7C). To our knowledge, CEL-III is the first lectin having the β-trefoil fold, which utilizes Ca^{2+} ion to recognize OH groups of specific carbohydrates by forming coordinate bonds. It seems possible that CEL-III acquired Ca^{2+}-mediated carbohydrate-recognition ability during evolution because of this animal's Ca^{2+}-rich environment in seawater.

35.9 MECHANISM OF HEMOLYTIC ACTIVITY

Chymotryptic digestion of carbohydrate-induced oligomer yielded the protease-resistant core fragments of about 91 kDa. From its N-terminal sequence analysis, this core fragments was found to be exclusively composed of domain 3 fragments, indicating that CEL-III molecules were strongly held in the oligomer through interactions between domain 3. On the other hand, partial digestion of intact CEL-III monomer using trypsin produced domain 2 and domain 3 fragments. The former exhibited

binding ability to the GalNAc-immobilized affinity column, while the latter showed no carbohydrate-binding activity and was isolated as an oligomerized form by gel filtration [29]. This result indicates that, once separated from other domains, domains 3 spontaneously proceeds to self-oligomerization, which may be triggered by exposure of the interface between domain 3 and two carbohydrate-binding domains to the solvent.

In the crystal structure of CEL-III, there are two α-helices (H8 and H9) in the cleft formed between the two carbohydrate-binding domains. They correspond to the hydrophobic region (residues 320–350) in domain 3 (Figure 35.4). Recently, we have found that the glutathione-S-transferase (GST)-fusion protein containing this α-helix region peptide at the C-terminus exhibits strong self-oligomerization tendency [30], which suggests that this region is closely related to oligomerization ability of CEL-III. There is a conspicuous feature in the amino acid sequence in this region that it contains several alternating hydrophobic and hydrophilic residues at every second positions (Figure 35.8A). Such sequences have also been found in the membrane-binding regions of several pore-forming toxins of bacterial origin [31], representatively α-hemolysin from *Staphylococcus aureus* (Figure 35.8B). In the heptameric structure of this toxin formed in the presence of detergent micelles, β-hairpins from each monomer are aligned to form a pore composed of β-barrel, in which hydrophobic and hydrophilic residues face exterior and interior side of the pore, respectively [13,14]. Therefore, it might be possible that the α-helix region of CEL-III also forms such an amphiphilic β-barrel when inserted into target cell membranes (Figure 35.9A). In fact, far-UV CD-spectra of chemically synthesized α-helix region peptides indicated that they adopt β-sheet structure in aqueous solution, which was particularly promoted in the presence of artificial lipid vesicles [30]. On the other hand, several synthetic peptides corresponding to different parts of this region exhibited antibacterial activity toward Gram-positive bacteria, *S. aureus* and *Bacillus subtilis*; the highest activity was observed with the peptide corresponding to residues 332–351 [32]. It was further revealed that

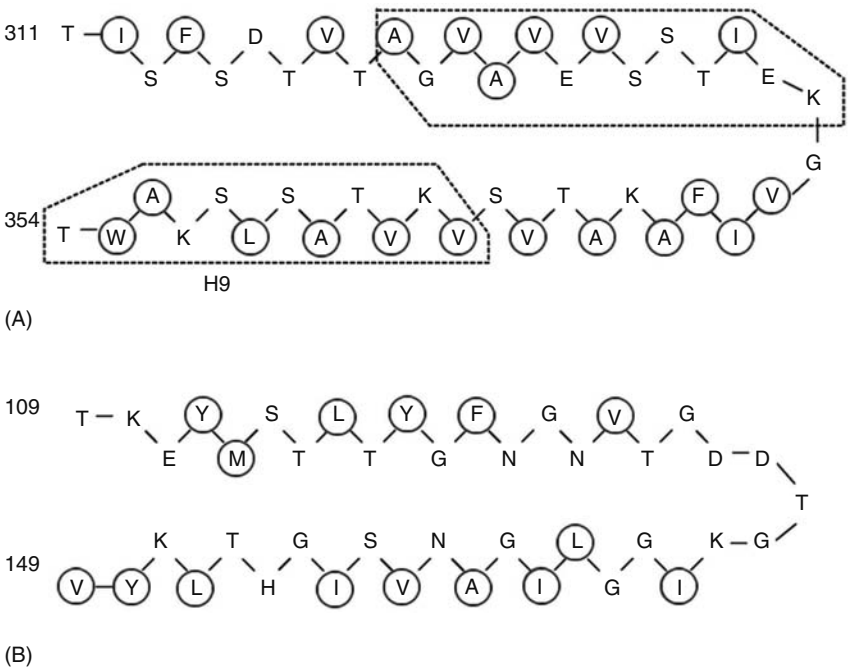


FIGURE 35.8 Amino acid sequences of the α-helix region of CEL-III (A) and the membrane-binding region of Staphylococcal α-hemolysin (B). Hydrophobic residues are enclosed in circles and two α-helices (H8 and H9) in domain 3 of CEL-III are indicated by dashed lines.

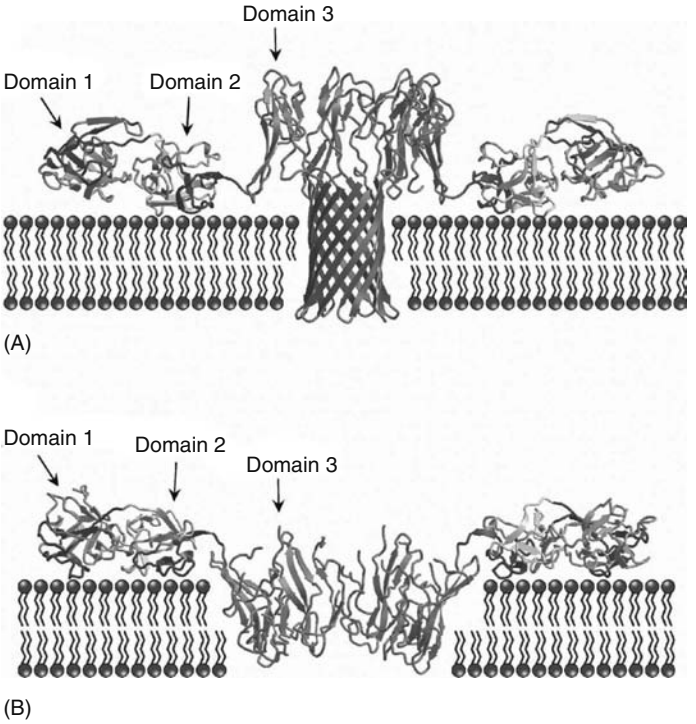


FIGURE 35.9 Two hypothetical models for membrane pore-formation by CEL-III, in which pores are formed by a β-barrel composed of the hydrophobic region (A), or whole domain 3 is inserted into the cell membrane (B).

antibacterial activity of these peptides was caused by a marked permeabilization of the bacterial cell membranes. These results also support that the α-helix region in domain 3 can interact with the cell membranes. Based on these experimental results, it was inferred that when exposed to solvent, the α-helix region may undergo structural transition to β-sheet, which is concurrently inserted into the cell membrane to form pores composed of β-barrel, like bacterial pore-forming toxins (Figure 35.9A). Alternatively, it seems also possible that larger portion of domain 3 can be inserted into membrane (Figure 35.9B), triggered by the structural change of the α-helix region, since CEL-III causes severe damages to the erythrocyte membrane after relatively long treatments [9] in contrast to bacterial pore-forming toxins, which generally produce stable membrane pores [33].

As mentioned above, there are five carbohydrate-binding sites in domains 1 and 2, and such multiple binding sites should contribute to the increase in the binding affinity for cell surface carbohydrate chains. In addition to this, binding to the glycolipids or glycoproteins at the multiple sites might cause a relative movement of domains 1 and 2, which in turn leads to a partial exposure of the α-helices located in the cleft between these domains, thereby promoting their structural transition from α-helix to β-sheet.

35.10 SIMILARITY WITH THE OTHER HEMOLYTIC LECTINS

There are some reports on lectins showing hemolytic activity [34–37]. Among them, the crystal structure of a lectin from the mushroom, *Laetiporus sulphureus* (LSL), has recently been solved [38]. This lectin is composed of N-terminal lectin module and C-terminal pore-forming module; the lectin module has a β-trefoil fold and the pore-forming module shows similarity with domains 3 and 4 of a pore-forming toxin aerolysin from *Aeromonas hydrophila* as well as ε-toxin from *Clostridium*

perfringens. Two of three subdomains (β and γ sites) in the N-terminal β-trefoil lectin domain show binding ability for *N*-acetyllactosamine in a similar manner as other homologous lectin domains. On the other hand, C-terminal pore-forming domains largely consist of β-sheets and responsible for the formation of oligomeric structure, which also assumed to be involved in membrane pore formation. Although there is no discernible sequence similarity between CEL-III and LSL, it is interesting that these lectins share a similar mode of action, i.e., binding to the cell surface carbohydrates with N-terminal β-trefoil lectin domains, followed by pore formation with C-terminal long β-strand-rich domains.

35.11 CONCLUSION AND FUTURE DIRECTIONS

It has been known that various marine invertebrates contain Ca^{2+}-dependent lectins [39–42]. One of the most probable functions of them is to act as humoral factors in self-defense system as innate immunity [3]. Although the peptides derived from the α-helix region of CEL-III were found to show antibacterial activity for two Gram-positive bacteria, *S. aureus* and *B. subtilis*, intact CEL-III does not act on these bacteria. Alternatively, CEL-III may have antimicrobial activity against some marine organisms having glactose-related carbohydrate chains. Interestingly, another *C. echinata* lectin CEL-I also shows relatively strong cytotoxicity through binding to specific carbohydrate chains on the cell membrane [43]. Therefore, it is also possible that *C. echinata* lectins, especially CEL-III, may play important roles in defense against attack of natural enemies or predators.

Several lines of evidence revealed that hemolytic action of CEL-III is brought about through oligomerization of its domain 3 to form pores in the target cell membrane. For this activity, binding via its domains 1 and 2 to cell surface carbohydrates is necessary to induce drastic conformational change of the protein, which leads to self-association of domain 3 in the cell membrane. The domains 1 and 2 adopt β-trefoil fold, which are often found in carbohydrate-binding modules in protein toxins, but interestingly, in the case of CEL-III, carbohydrate-binding is directly mediated via Ca^{2+} ions, like that of C-type CRDs, a structurally different protein family. Although hemolytic activity of CEL-III was found to be conducted by cooperative action of the carbohydrate-binding and oligomerizing domains, it still remains to be elucidated what structural changes make water-soluble CEL-III monomer to be inserted into the target cell membrane as strongly assembled oligomers. This process should include the exposure of hydrophobic face of some part of the protein to the solvent, and the α-helix region in domain 3 is most likely to be involved in it. Structural transition from water-soluble form to membrane-associated oligomers is one of the most interesting characteristics of CEL-III. Elucidation of its mechanism would provide important clues for not only hemolytic action of CEL-III, but also for that of various bacterial pore-forming toxins.

REFERENCES

1. Drickamer, K., Two distinct classes of carbohydrate-recognition domains in animal lectins, *J. Biol. Chem*., 263, 9557, 1988.
2. Drickamer, K., C-type lectin-like domains, *Curr. Opinion Struct. Biol*., 9, 585, 1999.
3. Vasta, G.R., Quesenberry, M., Ahmed, H., and O'Leary, N., C-type lectins and galectins mediate innate and adaptive immune functions: Their roles in the complement activation pathway, *Dev. Comp. Immunol*., 23, 401, 1999.
4. Hatakeyama, T., Kohzaki, H., Nagatomo, H., and Yamasaki, N., Purification and characterization of four Ca^{2+}-dependent lectins from the marine invertebrate, *Cucumaria echinata*, *J. Biochem*., 116, 209, 1994.
5. Hatakeyama, T., Ohuchi, K., Kuroki, M., and Yamasaki, N., Amino acid sequence of a C-type lectin CEL-IV from the marine invertebrate *Cucumaria echinata*, *Biosci. Biotechnol. Biochem*., 59, 1314, 1995.

6. Hatakeyama, T., Matsuo, N., Shiba, K., Nishinohara, S., Yamasaki, N., Sugawara, H., and Aoyagi, H., Amino acid sequence and carbohydrate-binding analysis of the *N*-acetyl-D-galactosamine-specific C-type lectin, CEL-I, from the Holothuroidea, *Cucumaria echinata*, *Biosci. Biotechnol. Biochem*., 66, 157, 2002.
7. Hatakeyama, T., Nagatomo, H., and Yamasaki, N., Interaction of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* with the erythrocyte membrane, *J. Biol. Chem*., 270, 3560, 1995.
8. Oda, T., Tsuru, M., Hatakeyama, T., Nagatomo, H., Muramatsu, T., and Yamasaki, N., Temperature- and pH-dependent cytotoxic effect of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* on various cell lines, *J. Biochem*., 121, 560, 1997.
9. Hatakeyama, T., Furukawa, M., Nagatomo, H., Yamasaki, N., and Mori, T., Oligomerization of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* induced by the binding of carbohydrate ligands, *J. Biol. Chem*., 271, 16915, 1996.
10. Hatakeyama, T., Miyamoto, Y., Nagatomo, H., Sallay, I., and Yamasaki, N., Carbohydrate-binding properties of the hemolytic lectin CEL-III from the holothuroidea *Cucumaria echinata* as analyzed using carbohydrate-coated microplate, *J. Biochem*., 121, 63, 1997.
11. Hatakeyama, T., Sato, T., Taira, E., Kuwahara, H., Niidome, T., and Aoyagi, H., Characterization of the interaction of hemolytic lectin CEL-III from the marine invertebrate, *Cucumaria echinata*, with artificial lipid membranes: involvement of neutral sphingoglycolipids in the pore-forming process, *J. Biochem*., 125, 277, 1999.
12. Kouriki-Nagatomo, H., Hatakeyama, T., Jelokhani-Niaraki, M., Kondo, M., Ehara, T., and Yamasaki, N., Molecular mechanism for pore-formation in lipid membranes by the hemolytic lectin CEL-III from marine invertebrate *Cucumaria echinata*, *Biosci. Biotechnol. Biochem*., 63, 1279, 1999.
13. Song, L., Hobaugh, M.R., Shustak, C., Cheley, S., Bayley, H., and Gouaux, J.E., Structure of staphylococcal α-hemolysin, a heptameric transmembrane pore, *Science*, 274, 1859, 1996.
14. Olson, R., Nariya, H., Yokota, K., Kamio, Y., and Gouaux, E., Crystal structure of staphylococcal LukF delineates conformational changes accompanying formation of a transmembrane channel, *Nat. Struct. Biol*., 6, 134, 1999.
15. Parker, M.W. and Feil, S.C., Pore-forming protein toxins: from structure to function, *Prog. Biophys. Mol. Biol*., 88, 91, 2005.
16. Kristan, K., Podlesek, Z., Hojnik, V., Gutierrez-Aguirre, I., Guncar, G., Turk, D., Gonzalez-Manas, J.M., Lakey, J.H., Macek, P., and Anderluh, G., Pore formation by equinatoxin, a eukaryotic pore-forming toxin, requires a flexible N-terminal region and a stable beta-sandwich, *J. Biol. Chem*., 279, 46509, 2004.
17. Yamaji, A., Sekizawa, Y., Emoto, K., Sakuraba, H., Inoue, K., Kobayashi, H., and Umeda, M., Lysenin, a novel sphingomyelin-specific binding protein, *J. Biol. Chem*., 273, 5300, 1998.
18. Oda, T., Shinmura, N., Nishioka, Y., Komatsu, N., Hatakeyama, T., and Muramatsu, T., Effect of the hemolytic lectin CEL-III from Holothuroidea *Cucumaria echinata* on the ANS fluorescence responses in sensitive MDCK and resistant CHO cells, *J. Biochem*., 125, 713, 1999.
19. Fujisawa, T., Kuwahara, H., Hiromasa, Y., Niidome, T., Aoyagi, H., and Hatakeyama, T., Small-angle X-ray scattering study on CEL-III, a hemolytic lectin from Holothuroidea *Cucumaria echinata*, and its oligomer induced by the binding of specific carbohydrate, *FEBS Lett*., 414, 79, 1997.
20. Nakano, M., Tabata, S., Sugihara, K., Kouzuma, Y., Kimura, M., and Yamasaki, N., Primary structure of hemolytic lectin CEL-III from marine invertebrate *Cucumaria echinata* and its cDNA: Structural similarity to the B-chain from plant lectin, ricin, *Biochim. Biophys. Acta*, 1435, 167, 1999.
21. Uchida, T., Yamasaki, T., Eto, S., Sugawara, H., Kurisu, G., Nakagawa, A., Kusunoki, M., and Hatakeyama, T., Crystal structure of the hemolytic lectin CEL-III isolated from the marine invertebrate *Cucumaria echinata*: Implications of domain structure for its membrane pore-formation mechanism, *J. Biol. Chem*., 135, 37133, 2004.
22. Hatakeyama, T., Unno, H., Kouzuma, Y., Uchida, T., Eto, S., Hidemura, H., Kato, N., Yonekura, M., and Kusunoki, M., C-Type lectin-like carbohydrate recognition of the hemolytic lectin CEL-III containing ricin-type β-trefoil folds, *J. Biol. Chem*. 282, 37826, 2007.
23. Zentz, C., Frenoy, J.P., and Bourrillon, R., Binding of galactose and lactose to ricin. Equilibrium studies, *Biochim. Biophys. Acta*, 536, 18, 1978.
24. Hatakeyama, T., Yamasaki, N., and Funatsu, G., Evidence for involvement of tryptophan residue in the low-affinity saccharide binding site of ricin D, *J. Biochem*., 99, 1049, 1986.

25. Frankel, A.E., Burbage, C., Fu, T., Tagge, E., Chandler, J., and Willingham, M.C., Ricin toxin contains at least three galactose-binding sites located in B chain subdomains 1α, 1β, and 2γ, *Biochemistry*, 9, 14749, 1996.
26. Fivaz, M., Abrami, L., and van der Goot, F.G., Landing on lipid rafts, *Trends Cell Biol*., 9, 212, 1999.
27. Abrami, L., Liu, S., Cosson, P., Leppla, S.H., and van der Goot, F.G., Anthrax toxin triggers endocytosis of its receptor via a lipid raft-mediated clathrin-dependent process, *J. Cell Biol*., 160, 321, 2003.
28. Dodd, R.B. and Drickamer, K., Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity, *Glycobiology*, 11, 71R, 2001.
29. Kouzuma, Y., Suzuki, Y., Nakano, M., Matsuyama, K., Tojo, S., Kimura, M., Yamasaki, T., Aoyagi, H., and Hatakeyama, T., Characterization of functional domains of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata*, *J. Biochem*., 134, 395, 2003.
30. Hisamatsu, K., Tsuda, N., Goda, S., and Hatakeyama, T., Characterization of the α-helix region in domain 3 of the haemolytic lectin CEL-III: Implications for self-oligomerization and haemolytic processes, *J. Biochem*., 143, 79, 2008.
31. Leppla S.H., *Bacillus anthracis* toxins. In: *The Comprehensive Sourcebook of Bacterial Protein Toxins*, 3rd ed., Alouf. J.E. and Popoff. M.R., Eds., Academic Press, New York, 2006, Chap. 18.
32. Hatakeyama, T., Suenaga, T., Eto, S., Niidome, T., and Aoyagi, H., Antibacterial activity of peptides derived from the C-terminal region of a hemolytic lectin, CEL-III, from the marine invertebrate *Cucumaria echinata*, *J. Biochem*., 135, 65, 2004.
33. Bhakdi, S., Weller, U., Walev, I., Martin, E., Jonas, D., and Palmer, M., A guide to the use of pore-forming toxins for controlled permeabilization of cell membranes, *Med. Microbiol. Immunol*., 182, 167, 1993.
34. Konska, G., Guillot, J., Dusser, M., Damez, M., and Botton, B., Isolation and characterization of an N-acetyllactosamine-binding lectin from the mushroom *Laetiporus sulfureus*, *J. Biochem*., 116, 519, 1994.
35. Armstrong, P.B., Swarnakar, S., Srimal, S., Misquith, S., Hahn, E.A., Aimes, R.T., and Quigley, J.P., A cytolytic function for a sialic acid-binding lectin that is a member of the pentraxin family of proteins, *J. Biol. Chem*., 271, 14717, 1996.
36. Satoh, F., Nakagawa, H., Yamada, H., Nagasaka, K., Nagasaka, T., Araki, Y., Tomihara, Y., Nozaki, M., Sakuraba, H., Ohshima, T., Hatakeyama, T., and Aoyagi, H., Fishing for bioactive substances from scorpionfish and some sea urchins, *J. Natur. Toxins*, 11, 297, 2002.
37. Tateno, H. and Goldstein, I.J., Molecular cloning, expression, and characterization of novel hemolytic lectins from the mushroom *Laetiporus sulphureus*, which show homology to bacterial toxins, *J. Biol. Chem*., 278, 40455, 2003.
38. Mancheno, J.M., Tateno, H., Goldstein, I.J., Martinez-Ripoll, M., and Hermoso, J.A., Structural analysis of the *Laetiporus sulphureus* hemolytic pore-forming lectin in complex with sugars, *J. Biol. Chem*., 60, 17251, 2004.
39. Suzuki, T., Takagi, T., Furukohri, T., Kawamura, K., and Nakauchi, M., A calcium-dependent galactose-binding lectin from the tunicate *Polyandrocarpa misakiensis*. Isolation, characterization, and amino acid sequence, *J. Biol. Chem*., 265, 1274, 1990.
40. Poget, S.F., Legge, G.B., Proctor, M.R., Butler, P.J., Bycroft, M., and Williams, R.L., The structure of a tunicate C-type lectin from *Polyandrocarpa misakiensis* complexed with D-galactose, *J. Mol. Biol*., 290, 867, 1999.
41. Nair, S.V., Burandt, M., Hutchinson, A., Raison, R.L., and Raftos, D.A., A C-type lectin from the tunicate, *Styela plicata*, that modulates cellular activity, *Comp. Biochem. Physiol. C. Toxicol. Pharmacol*., 129, 11, 2001.
42. Kakiuchi, M., Okino, N., Sueyoshi, N., Ichinose, S., Omori, A., Kawabata, S., Yamaguchi, K., and Ito, M., Purification, characterization, and cDNA cloning of α-*N*-acetylgalactosamine-specific lectin from starfish, *Asterina pectinifera*, *Glycobiology*, 12, 85, 2002.
43. Kuramoto, T., Uzuyama, H., Hatakeyama, T., Tamura, T., Nakashima, T., Yamaguchi, K., and Oda, T., Cytotoxicity of a GalNAc-specific C-type lectin CEL-I toward various cell lines, *J. Biochem*., 137, 41, 2005.

36 Structure–Function Relationship in Mammalian Chitinase-Like Lectins

Shaun Morroll, Sharon Turner, and Franco H. Falcone

CONTENTS

36.1 INTRODUCTION

Chitin is an unbranched polysaccharide comprising of multiple repetitive units of β(1,4)-linked *N*-acetyl-D-glucosamine. It is the third most abundant biopolymer on earth after cellulose, and is found mainly in insect exoskeletons, crustacean shells, and fungal cell walls. The enzymatic breakdown of chitin is mediated by chitinases, which hydrolyze the β-1,4-*N*-acetyl-D-glucosamine bonds. Chitinases are needed for ecdysis in growing insects and crustaceans [1], but also in fungi [2]. Chitinases are also known to have digestive functions and can reach very high concentrations in the pancreatic secretions of fish [3]. An acidic chitinase expressed in parts of the stomach with a low pH optimum has been more recently described in mammals [4,5]. Plant chitinases are prominent allergens and are thought to play a role in antifungal immunity [6].

From the biochemical point of view, chitinases are O-glycosyl hydrolases (EC 3.2.1.). According to carbohydrate active enzyme classification (CAZy, http://afmb.cnrs-mrs.fr/CAZY/acc.html), there are currently 97 different families of structurally related enzymes in this group. The tertiary structure is usually better conserved than the primary sequence, leading to a classification in clans. In the CAZy classification, chitinases are found in Families 18 and 19. Family 19 (also known as IA and IB or I and II) are mainly found in plants.

Although mammals and other higher vertebrates do not contain chitin and lack the enzymatic machinery to produce it, their genomes contain several clustered genes expressing chitinases as well as chitinase-like genes. Chitinases and chitinase-like genes in mammalians are all related to Family 18 chitinases.

Some of the encoded proteins are thought to have lost enzymatic activity due to the replacement of active site proteins essential for chitin hydrolysis [7,8]. The active site of chitinases has a DXX**D**XDX**E** consensus sequence, in which the last aspartic and the central glutamic acid are thought to be involved in the catalytic mechanism [9]. Mutation in any of these two amino acids usually leads to loss of enzymatic activity. These chitinase-like proteins, however, still have the ability to bind chitin and other polysaccharides, and therefore are referred to as chitinase-like lectins or "chi-lectins" (chilectins) in this analysis [10,11]. In humans, chitinases and chilectins related to CAZy Family 18 are clustered on two separate regions of chromosome 1 (Figure 36.1).

In mice, the chitinase/chilectin cluster is segregated on mouse chromosomes 3 and 1 (Figure 36.2). The corresponding areas on human chromosome are known to be in synteny, as can be seen on the human–mouse whole genome orthology map published by the mouse genome database (http://www.informatics.jax.org/) [12].

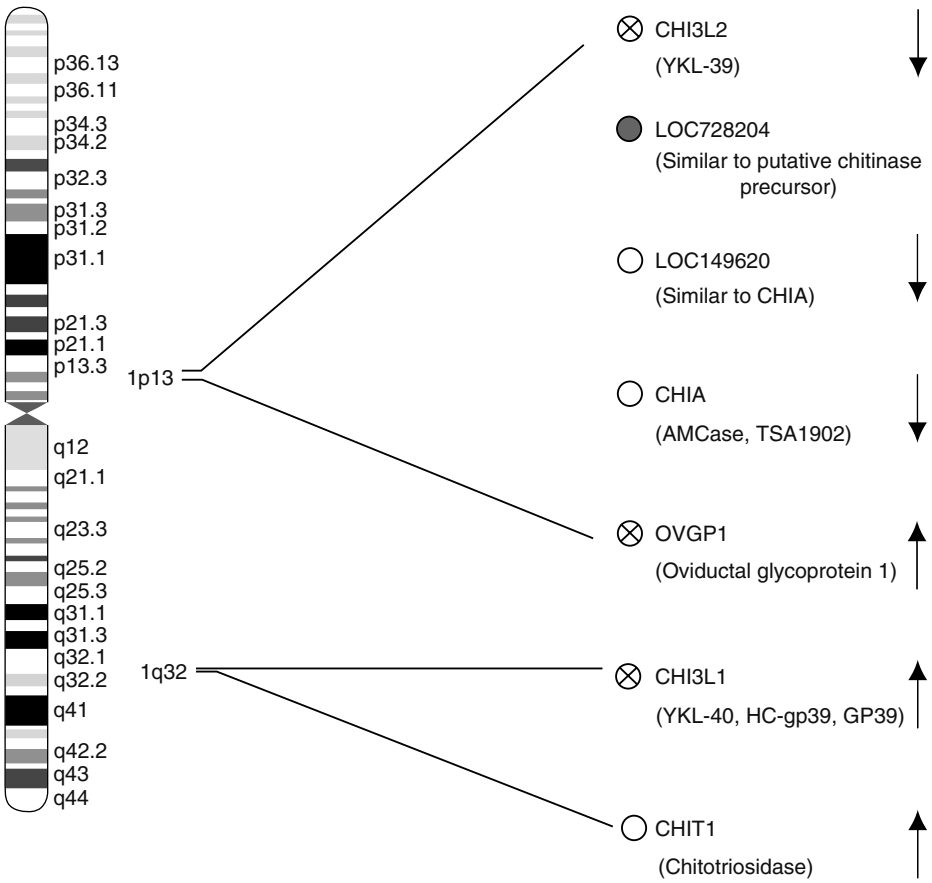


FIGURE 36.1 Chitinase/chilectin cluster on human chromosome 1 (obtained from NCBI Build 36.2, using the cytogenetic location map). Arrows indicate gene orientation, full white circles: known or potentially active chitinase; crossed circles: inactive chitinase (chilectin), grey circle: pseudogene.

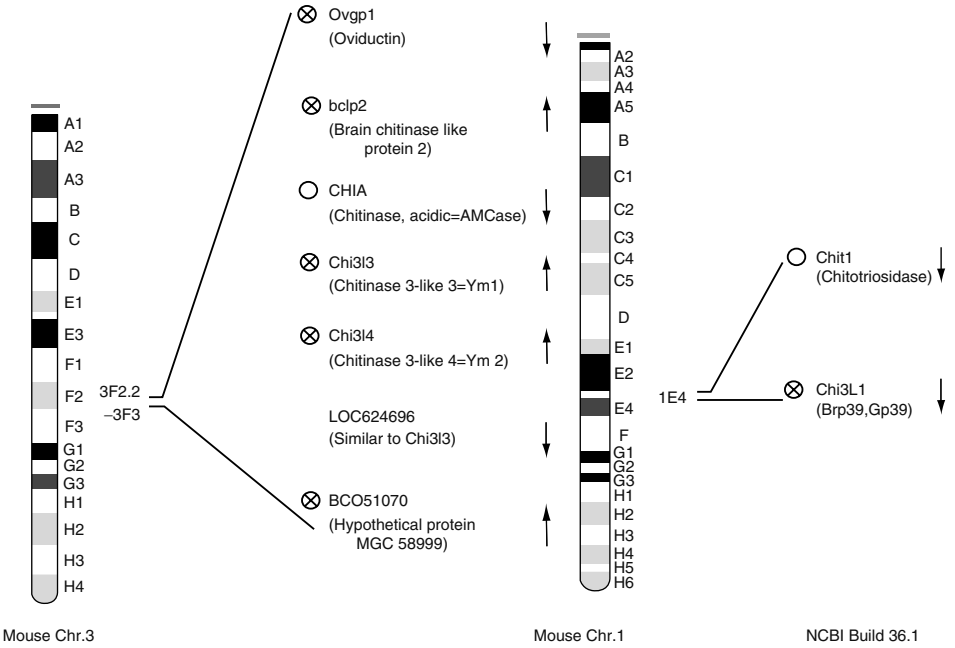


FIGURE 36.2 Chitinase/chilectin cluster on mouse chromosomes 1 and 3 (obtained from NCBI Build 36.1, using the cytogenetic location map). Arrows indicate gene orientation, full white circles: known or potential active chitinase; crossed circles: inactive chitinases (chilectins).

36.2 PHYLOGENETIC ANALYSIS OF THE CHITINASE/CHILECTIN CLUSTER

Our phylogenetic analysis, performed using most available mammalian chitinase and chilectin protein sequences, shown in Figure 36.3, results in five major groups being outlined: Ym1/Ym2, acidic chitinases, Chi3L1, chitotriosidase, and the oviductin group. The overall bootstrap support for the tree is very high. Whereas chitotriosidase (Chit1), acidic chitinase (AMCase), Chi3L1, and oviductin homologs are found in all analyzed mammalian gene sequences, the analysis also indicates that there seem to be no functional homologs for murine Ym1 and Ym2 in humans and other mammalians, except in other rodents. This lack of orthologs of Ym1 and Ym2 in humans, as demonstrated by our phylogenetic analysis, underpins the view recently expressed by Boot et al. [13]. It also suggests that Ym1 and Ym2 may be rodent-specific adaptations. Furthermore, even where putative orthologs appear to exist in several mammalian species (as is the case for chitotriosidase and AMCase), gene expression in specific organs and tissues may be controlled differently due to distinct promoter usage [5], and thus fulfill different roles in different species. There is therefore an important caveat when extrapolating chilectin and chitinase data derived from animal disease models to the human situation [5]).

36.3 CHITINASES AND CHI-LECTINS SHARE A COMMON STRUCTURE

The three-dimensional structure of chitinases and chilectins is similar. According to the Structural Classification of Proteins (scop: http://scop.mrc-lmb.cam.ac.uk/scop/) [14], both types of proteins are found in the alpha and beta proteins (α/β) structural class, and are characterized by the TIM β/α-barrel fold. The superfamily in scop is (trans)glycosidase. The TIM β/α -barrel fold is named after the first structure where this fold was first observed, the glycolytic pathway enzyme triosphosphate

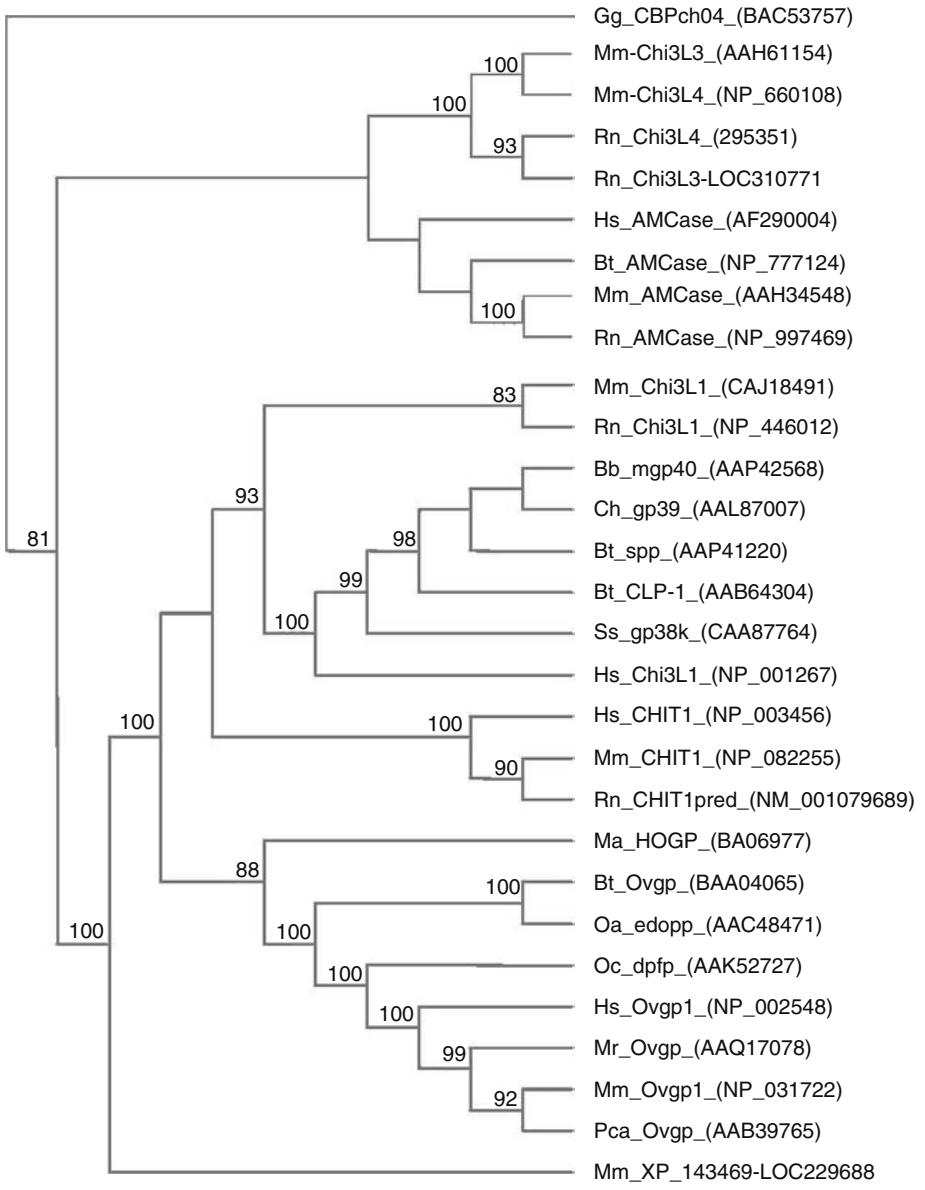


FIGURE 36.3 Phylogram and bootstrap analysis of currently known mammalian chitinases and chilectins. Abbreviations used are: Bb (*Bubalus bubalis,* buffalo); Bt (*Bos taurus*, true cattle); Ch (*Capra hircus*, domestic goat); Gg (*Gallus gallus*, red junglefowl); Hs (*Homo sapiens,* human); Ma (*Mesocricetus auratus;* golden hamster); Mm (*Mus musculus*, house mouse); Mmu (*Macaca mulatta*, rhesus monkey); Mr (*Macaca radiata,* bonnet macaque*);* Oa (*Ovis aries*, sheep); Oc (*Oryctolagus cuniculus,* European rabbit*)*; Pca (*Papiocynocephalus anubis,* olive baboon*);* Rn (*Rattus norvegicus,* rat*);* Ss (*Sus scrofa,* pig). Numbers indicated in brackets are GenBank Protein IDs. This phylogram was generated with the neighbor joining method using MacVector 7.2. Chicken CBPch04 was used as outgroup to root the tree. Human Chi3l2 (YKL-39) was not included, as its inclusion results in a polytomic tree. Chi3L2, however, does not group with the other "clades". Bootstrap values are for 1000 repeats.

isomerase [15]. This structure is characterized by eight parallel β strands (β1 to β8), organized into staves that resemble a barrel with a general 30° tilt about the barrel's axis. Surrounding the β-barrel structure are eight α helices (α1 to α8) with occasional turns introduced (Figure 36.4A through F). The TIM barrel structure of chitinases and chilectins forms the main portion of the tertiary structure of the proteins. At the C-terminus of the proteins between β strand 7 and α helix 7 is another domain, a structure referred to here as the α/β-insertion domain. This second domain contains a five-stranded antiparallel pseudoplanar β-sheet structure, which is intercalated by short β and α helical strands. The α/β-insertion domain is not found in plant chitinases (Figure 36.4C and D) [16]. Mammalian chilectins are similar to the chitinases, sharing the folding of the core domain and common van der Waals surfaces. Chitinases and chilectins have a common canyon (or substrate binding groove) that

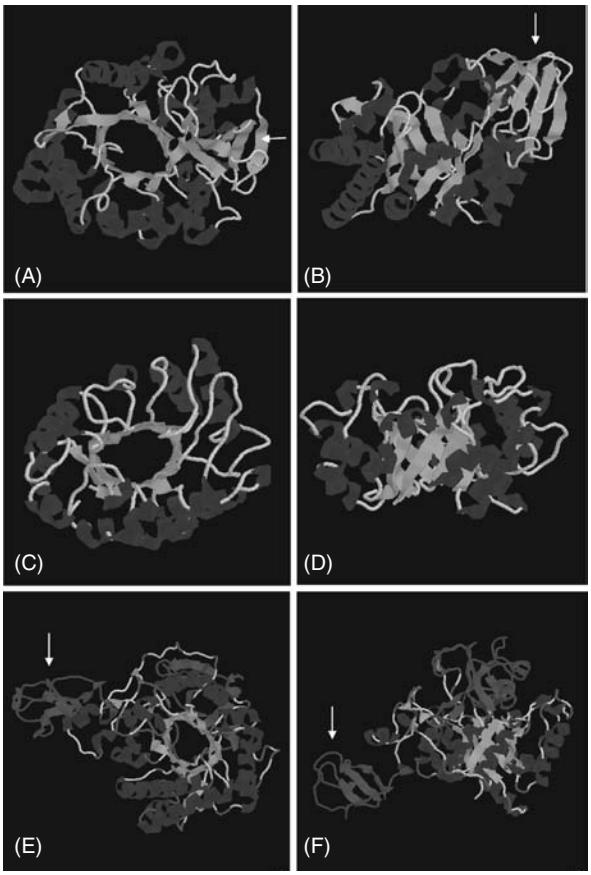


FIGURE 36.4 (See CD for color figure.) Three-dimensional structures of Type-II chitinases and chilectins. The top (A, C, E) and side views (B, D, F) highlight different characteristic features of chitinases and chilectins. Shown are murine Chi3L3 (Ym1) (A, B) (1VF8.pdb) (From Tsai, M.L., Liaw, S.H., and Chang, N.C., *J. Struct. Biol.,* 148, 290, 2004); hevamine (C, D) (2HVM.pdb) and the bacterial chitinase B *Serratia marcescens* (E, F) (106I.pdb, complexed with the catalytic intermediate mimic cyclic dipeptide Ci4 (From Houston, D.R. et al., *Biochem. J.,* 368, 23, 2002)). Structures were obtained from the indicated pdb files using RasMol 2.6 (From Sayle, R.A. and Milner-White, E.J., *Trends Biochem. Sci.,* 20, 374, 1995). In the center is the β-barrel structure (yellow); the general tilt of the barrel axis is best seen in the side view (B, D, F). Surrounding the β-barrel are the eight α-helices (pink). The α/β-insertion domain is also shown in yellow and highlighted by an arrow (A, B). This domain is not present in plant chitinases (C, D). Linked by a linker sequence, which is Ser/Gly-rich in homologues of AMCase, is a type-2 chitin binding domain (red and arrow in E and F).

FIGURE 36.5 Alignment of chitin-binding Type-2 domains from mammalian chitinases. These domains contain six conserved cysteines including two adjacent cysteines close to the carboxy-terminal end of the protein (Hs, *Homo sapiens*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Bt, *Bos taurus*; Gg, *Gallus gallus*; Cf, *Canis familiaris*).

separates the molecule into two hemispheres. This canyon is lined with aromatic residues, which interact with the chitin substrate in chitinases [17].

Mammalian chitinases also possess a separate chitin-binding domain called chitin-binding type-2 domain (Prosite PS50940; Pfam PF01607) (Figure 36.5). This domain (sometimes referred to as Peritrophin A-like domain) is found in animal and some baculovirus proteins, while a similar chitin-binding type-1 domain (Prosite PS50941) is found in bacterial and plant chitinases. The structure of chitin-binding type-2 domains has been solved for the small antimicrobial protein from horseshoe crab tachycitin (1dqc.pdb) [21]. The structure is very similar to the type-1 domain in hevein and consists of a distorted β-sandwich formed by a sheet of three antiparallel β-strands linked to a two β-strand sheet by a short connecting helical loop, starting with Cys40, which is also the start of the chitin-binding site (Cys-40 to Gly-60) in tachycitin.

Chitin-binding domains are not found in mammalian chilectins. As a result, their molecular weight is usually smaller than their active chitinase counterparts. Taken together, mammalian chilectins have several common structural features, which they share with chitinases. Although the tertiary structures of chilectins are now well understood, there is still a lack of knowledge and considerable debate regarding the physiological ligands and functions of mammalian chilectins. Although the folds appear to be highly conserved in different mammalian species, there are differences between the rodent and human chitinase/chilectin gene clusters, which point to a different set of functions. In the following sections, we will discuss what is known about the functions of the two currently best mammalian chilectins, Chi3L1 in humans and Chi3L3/Chi3L4 (Ym1/Ym2) in mice.

36.4 CHITINASE-3-LIKE-1 (CHI3L1, YKL-40)

Chi3L1, also known as YKL-40 or human cartilage glycoprotein-39 (HC-gp39), is the currently best studied human chilectin. Chi3L1 is a heparin and chitin-binding lectin without chitinase activity. The Chi3L1 gene consists of 10 exons and spans approximately 13 kb of genomic DNA. The single polypeptide chain consists of 383 amino acids with a molecular weight of 40476 Da [22]. This molecular mass (40 kDa), along with the single letter code for the three N-terminal amino acids of the mature protein, i.e., after cleavage of the signal sequence, tyrosine (Y), lysine (K), and leucine (L), resulted in the protein name YKL-40, a name which is generally used in the oncological literature. Chi3L1 is also present in several other mammalian species, e.g., pig (gp38k, 84% sequence identity) [23], mouse (Brp39, 73% identity) [24], goat (MGP-40, 83% sequence identity) [25], and cow (39 kDa whey protein, 83% sequence identity) [26]. The official gene name (HUGO Gene Nomenclature Committee, HGNC) is Chi3L1, it is highly recommended to use this name to avoid

confusion. Chi3L1 protein was identified in 1988 [26] in cows and five years later in humans [22], but still its exact physiological function is unknown. However many roles have been postulated, some of which are outlined below. Also it is probable, as Chi3L1 is a secreted protein, that its site or sites of action are extracellular although no cell surface or soluble receptors specific to Chi3L1 have yet been identified.

36.4.1 Functions of Chi3L1

36.4.1.1 Chi3L1 as Growth Factor

Chi3L1 may act a growth factor inducing the proliferation of connective tissue cells, including chondrocytes and synoviocytes [27,28]. Under normal physiological conditions, Chi3L1 protein secretion from these cell types is low [27,28]. However patients with pathophysiological inflammatory joint conditions, such as rheumatoid arthritis, have elevated levels of Chi3L1 (usually referred to in the rheumatological literature as HC-gp39) in their synovial fluid and sera [27], possibly resulting from increased Chi3L1 protein secretion from chondrocytes and synoviocytes themselves [28]. So it seems that Chi3L1 is involved in the homeostasis of connective tissues and the increased tissue remodeling seen in joint diseases may be associated with the growth factor properties of Chi3L1 present in high levels in patients with arthritic joint disease [27].

Chi3L1 is also found preformed in specific granules of human neutrophils isolated from the blood of healthy subjects [29] and is synthesized in large quantities by activated macrophages [30–32], thus Chi3L1 is also present in healthy subjects.

The proliferative capabilities of Chi3L1 can be explained by the initiation of the mitogen-activated protein kinase (MAPK) and phosphatidyl-inositol 3-kinase (PI-3K) signaling cascades [27], both of which are associated with the process of mitogenesis. Additionally, the activation of the these cytoplasmic signal transduction pathways reinforces the suggestion that Chi3L1 interacts with a receptor located in the plasma membrane of cells to exert its effects. The existence of a postulated receptor is consistent with the observation made by Houston et al. [10], who found that upon binding of the chitooligosaccharide GlcNAc8 (representing a soluble fragment of polymeric chitin as the presumed ligand), Chi3L1 undergoes a pronounced conformational change. This may point to a Chi3L1-dependent mechanism that enables the cells to "sense" the presence of specific extracellular matrix components and signal this to the cell, inducing an appropriate response, e.g., to sense and respond to changes in the extracellular environment, e.g., after injury or during disease.

Fusetti et al. have studied the carbohydrate binding regions of Chi3L1 and found that they are compatible with the binding of chitin (=*N*-acetyl-glucosamine) oligomers [11]. However, chitin does not occur in mammals, as they lack the enzymatic machinery to achieve its synthesis, and therefore the physiological ligand of Chi3L1 remains to be determined. A rather unexpected recent finding is that the bovine Chi3L1 ortholog from chondrocytes binds to Type I collagen, accelerating the rate of collagen fibril formation *in vitro* and appears to interfere with matrix metalloprotease-1 (MMP-1) (interstitial/fibroblast collagenase) degradation of type 1 collagen fibrils [33]. This not only points to a possible role in fibrosis, but also to the fact that the physiological ligands of Chi3L1 may be proteins rather than polysaccharides. How this observation fits with the known ability of Chi3L1 to bind chitin oligomers and heparin [34] remains to be seen.

36.4.1.2 Chi3L1 as Antiapoptotic Factor

The MAPK signal transduction pathway when activated links extracellular growth and differentiation signals with the induction of a number of transcription factors, which permit the progression of mitosis [35]. The activated PI-3K pathway plays a dual role in the cell; firstly it facilitates the progression of mitosis by promoting the entry of quiescent cells into the S phase of the cell cycle and secondly it has a major role in antiapoptotic cell survival by increasing the expression of several antiapoptotic proteins and downregulating proapoptotic proteins [36]. An antiapoptotic role for

human Chi3L1 is further indicated by the functioning of its murine ortholog, breast regression protein-39 (BRP-39), during mammary involution. BRP-39 is induced in murine mammary epithelial cells shortly after weaning occurs and during this time massive tissue remodeling with extensive programmed cell death takes place. It has been hypothesized that BRP-39 acts as a protective signaling factor to determine which cells survive this drastic tissue remodeling (Aronson N.N. Jr., unpublished data). A similar role is described for MGP-40, the probable ortholog in goats [25]. The first mammalian homolog of human Chi3L1 was discovered as a 39 kDa whey protein in mammary secretions collected during the nonlactating period [26].

36.4.1.3 Chi3L1 as Angiogenic Factor

Chi3L1 is a potent migration factor for endothelial cells, which stimulates the migration of vascular endothelial cells at a level comparable with the endothelial cell chemoattractant basic fibroblast growth factor (bFGF) [37]. The morphology of vascular endothelial cells is changed by the presence of Chi3L1, such that branching tubules are formed [37]. This is indicative of Chi3L1 functioning in angiogenesis by stimulating the migration and reorganization of vascular endothelial cells. Chi3L1 may also stimulate angiogenesis in hypoxic arthritic joints [38], as one of the earliest histological changes in rheumatoid arthritis is an increased number of blood vessels [39] and patients with this disease are known to have elevated Chi3L1 levels [27].

36.4.1.4 Chi3L1 and Tissue Remodeling

A further role for Chi3L1 in inflammatory responses or degradation and remodeling of the extracellular matrix was postulated after observing Chi3L1 expression in normal and pathologic disease states, including rheumatoid arthritis [27] and inflammatory bowel disease (IBD) [40]. Neutrophils and macrophages isolated in the synovial fluid of patients with rheumatoid arthritis [27] and in the inflamed intestine of patients with IBD [41] show high Chi3L1 expression. Chi3L1 or its peptide derivatives could be a target of the T-cell-mediated immune response, and this immune response may play a role in sustaining chronic inflammation [41].

Conversely, Chi3L1 can act to downregulate the degradative aspects of the inflammatory response. This is achieved by the modulation of the signaling pathways induced by the inflammatory cytokines tumor necrosis factor-α (TNF-α) and interleukin-1 (IL-1), namely the MAPK and stress-activated protein kinase/Jun N-terminal kinase (SAPK/JNK) pathways [42]. Since the pathological loss of extracellular matrix, particularly cartilage, seen in conditions such as inflammatory and degenerative arthritis, is largely driven by TNF-α and IL-1, Chi3L1 may play an important role in regulating the degenerative aspect of the inflammatory response of tissues by downregulating cytokine-induced MMP and chemokine production. More recently, Chi3L1 has also been shown to act as a positive regulator of chondrocyte differentiation and Type II collagen production [43] by inducing the expression of SOX9, a transcription factor thought to be involved in regulation of collage type–II synthesis [44].

36.4.2 Chi3L1 and Cancer

Thus Chi3L1 has numerous hypothesized roles including that of a growth factor and antiapoptotic protein, a migration and angiogenic factor and an autoantigen, resulting in inflammation and extracellular remodeling (Figure 36.6). A search of the Chi3L1 protein sequence against the dbest database at the NCBI using the BLAST algorithm [45,46] showed several types of cancer cells to overexpress the protein, including colorectal [45], ovarian [46], breast, uterine, prostate, kidney and lung carcinoma, glioblastoma, and germ cell tumors [47]. However, in the case of small cell lung carcinoma, it is not the tumor itself, but the peritumoral macrophages that provide the predominant source of Chi3L1 [48,49].

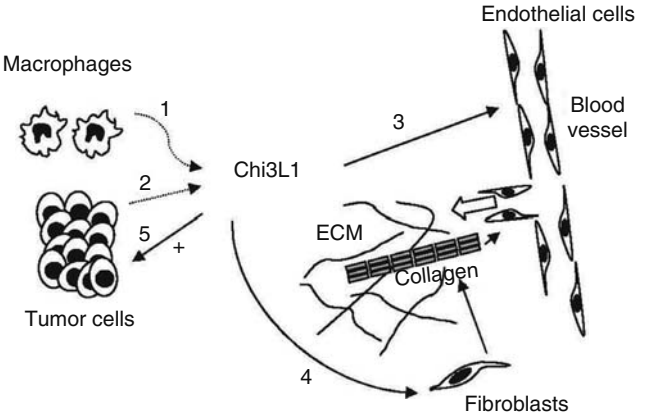


FIGURE 36.6 Potential roles of Chi3L1 in cancer development. Chi3L1 is produced either by the tumor itself (From Flach, J., Pilet, P.E., and Jolles, P., *Experientia*, 48, 701, 1992; Cintin, C. et al., *Cancer*, 95, 267, 2002; Jensen, B.V., Johansen, J.S., and Price, P.A., *Clin. Cancer Res*., 9, 4423, 2003.) or by infiltrating macrophages. (From Kramer, K.J. and Koga, D., *Insect Biochem*., 16, 851, 1986; Johansen, J.S. et al., *Lung Cancer*, 46, 333, 2004; Junker, N. et al., *Lung Cancer,* 48, 223, 2005.) The produced Chi3L1 could promote the growth of new blood vessels by inducing proliferation, migration, and tubule formation of endothelial cells. (From Fänge, R., Lundblad, G., and Lind, J., *Mar. Biol*., 36, 277, 1976; Malinda, K.M. et al., *Exp. Cell Res*., 250, 168, 1999.) The growing blood vessel is stabilized by synthesis of extracellular matrix (ECM) including basement membrane. The continued development of a tumor relies on an interaction between the tumor and its microenvironment, particularly the ECM. Consequently, a key modulator of cancer growth and metastasis is the correct balance of ECM synthesis and degradation. (From Boudreau, N. and Bissell, M.J., *Curr. Opin. Cell Biol*., 10, 640, 1998.) Tumor-associated macrophages and neutrophils isolated from the stroma surrounding breast and colorectal cancer cells produce growth and angiogenic factors, chemokines, chemotactic factors, MMPs, and other extracellular matrix digesting enzymes. (From Sunderkotter, C. et al., *J. Leukoc. Biol*., 55, 410, 1994.) They have also been shown to express Chi3L1 (From Kramer, K.J. and Koga, D., *Insect Biochem*., 16, 851, 1986.), which may act to promote extracellular matrix synthesis via its growth factor capabilities (From Boot, R.G. et al., *J. Biol. Chem*., 276, 6770, 2001; De Ceuninck, F. et al., *Biochem. Biophys. Res. Commun*., 285, 926, 2001; Jacques, C. et al., *Osteoarthritis Cartilage*, 15, 138, 2007.) or to suppress the actions of extracellular digesting enzymes. (From Ling, H. and Recklies, A.D., *Biochem. J*., 380, 651, 2004.) This would result in a disturbed balance between extracellular matrix synthesis and degradation, allowing the continued growth of the tumor. Finally, Chi3L1 may act on the tumor cells themselves by triggering antiapoptotic survival pathways (From Recklies, A.D., White, C., and Ling, H., *Biochem. J*., 365, 119, 2002; Franke, T.F., Kaplan, D.R., and Cantley, L.C., *Cell*, 88, 435, 1997.) via a yet to be identified cellular receptor. So it could be assumed that Chi3L1 expression from the tumor itself or from cells in the surrounding stroma may positively influence tumor growth, resulting in a highly malignant cancer, which corresponds with metastasis of the primary tumor and a poorer prognosis for the patient.

A number of case–control studies involving these cancer types have revealed that some patients had elevated serum Chi3L1 levels when compared to healthy controls. Studies relating to colorectal, [45,50], ovarian, [46] and recurrent breast carcinoma [51] and glioblastoma [52,53] have demonstrated a significant association between increased serum Chi3L1 levels and shorter survival times and shorter relapse periods. Human gliomas are well documented to express high levels of Chi3L1 [52,54] with Chi3L1 protein levels being reproducibly measured in the serum of patients with malignant primary brain tumors [54]. Additionally, investigations into metastatic breast cancer [47] indicate that the level of serum Chi3L1 reflects the aggressiveness of the cancer, with higher levels signifying more aggressive tumors. Consequently, Chi3L1 protein was identified as a potentially important marker and prognostic factor in these highly malignant human cancers. In the context of

breast cancer, it is worth stressing that the mouse homolog Brp39 was identified in neu/ras-initiated murine mammary tumors [24].

As the biological role of Chi3L1 in cancer is unknown, it is probable that a number of the functions elucidated for other cell types are applicable. So it follows that Chi3L1 may be involved in cancer cell proliferation and survival, in angiogenesis, in the inflammatory process around the tumor and in the remodeling of the extracellular matrix [38]. Chi3L1 may also act as a growth factor for cancer cells and protect them from apoptosis, via activation of the cell survival pathways MAPK and PI-3K, thus complying with one of the cornerstones of malignant transformation, a dysregulated increase in cell number [55].

For a tumor mass to grow beyond a certain size (2–3 mm^2), it must gain an adequate blood supply and the promigratory effect of Chi3L1 on vascular endothelial cells appears to implicate Chi3L1 in this process. Thus Chi3L1 may act as a positive regulator of angiogenesis surrounding the tumor and could play a role in primary and metastatic tumor growth [38]. However, it needs to be stressed that the migratory effect was demonstrated using the porcine homolog gp38k on human venous endothelial cells (HUVEC) [37] and therefore still needs to be confirmed with human Chi3L1, in an *in vivo* rather than *in vitro* setting.

Taken together, Chi3L1 is interesting on two levels: on one hand as a prognostic marker of disease and on the other hand also as a new target for cancer and rheumatoid arthritis therapy. It will therefore be of paramount importance to discover the identity of the physiological ligands and the nature of its putative cellular receptor.

36.5 CHI3L3 AND CHI3L4 (YM1/YM2)

36.5.1 Structure and Physiological Ligands

While Chi3L1 is the best studied chilectin in humans, Chi3L3 (Ym1), and to a lesser extent Chi3L4 (Ym2), which shares 95% of the amino acids with the former chilectin, have attracted significant attention since their original description by Jin et al. in 1998 [57]. As shown in Figure 36.7, the Glu^{139} in the chitinase active site consensus sequence DXX**D**XDX**E** is mutated to Gln^{139} in both genes, and Asp^{135} is mutated to Asn^{135}. As a result, the corresponding proteins have no enzymatic activity.

An unusual feature of Chi3L3 is its tendency to crystallize spontaneously around its isoelectric point (pH 5.7) under low salt conditions. The Chi3L3 crystal structure has been described at

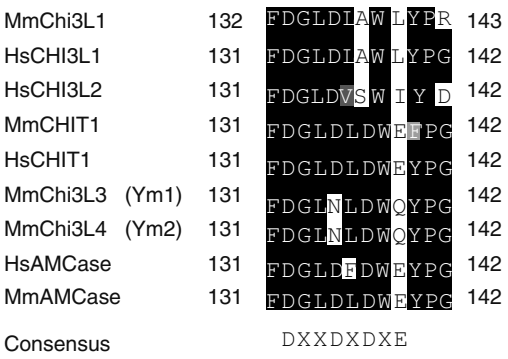


FIGURE 36.7 Multiple sequence alignment of the active site of mouse and human chitinases and chilectins. Inactive chitinases (chilectins) have a mutation in either two (MmChi3L3, MmChi3L4) or one (HsChi3L1, HsChi3L2, MmChi3L1) of the active site residues. (From Watanabe, T. et al., *J. Biol. Chem.*, 268, 18567, 1993.) Mammalian oviductins, not shown in this alignment, are also chilectins but have different overall protein sequence. Mm indicates mouse genes, Hs human genes.

2.5 Å resolution [58] and later refined at a high 1.31 Å resolution [18]. Chi3L3 shows similar structural fold to Chi3L1 but with a narrower ligand-binding pocket [18,58]. However, evidence for chitin binding is still controversial. Chang et al. demonstrated binding of NAG oligomers by surface plasmon resonance, but binding only occurred in the narrow pH range of 4–5, with no binding at physiological pH [59]. Tsai et al. failed to detect Chi3L3 binding to various NAG substrates [18], and crystal soaking experiments with various glucosamine or *N*-acetylglucosamine oligomers only yielded the uncomplexed protein. This is in accordance with our own inability to detect any significant binding of recombinant Chi3L3 to immobilized chitin oligomers (Falcone et al., 2005, unpublished data). As shown by Fusetti et al. [17], six solvent-exposed aromatic amino acid residues in human chitotriosidase, which line the substrate-binding groove, are important in interacting with the hydrophobic faces of the hexose rings of chitin oligomers by stacking forces. The refined structure of Ym1 shows that of these six aromatic amino acids, three (Tyr^{13}, Trp^{50}, and Trp^{191}) are replaced by charged amino acids Asp^{13}, Glu^{50}, and Lys^{97} [18]. Thus the lack of conservation in the substrate binding groove may account for the inability of Chi3L3 to bind chitin. Furthermore, as described in the introduction, Ym1, like the other chilectins, does not possess a Type-2 chitin-binding domain.

Since it appears that Chi3L3 can neither cleave nor possibly bind chitin, the question of what its physiological ligands are remains open. It is therefore likely that Chi3L3 will bind other carbohydrates, and indeed it is known that Chi3L3 binds heparin and heparan sulfate [59]. However it is also possible, as suggested by Tsai [18], that the physiological ligands of Chi3L3 are proteins rather than carbohydrates, and a similar situation has already been suggested for bovine Chi3L1, which as described earlier binds Type I collagen [33]. This of course would call into question the term chilectin for these proteins.

36.5.2 Suggested Functions of Chi3L3 and Chi3L4

The earliest suggested function for Chi3L3 comes from Owhashi et al. [60] who were studying the eosinophil chemotactic activity secreted by spleen cells of mice infected with the trematode parasite *Schistosoma japonicum*. A protein chemotactic for eosinophils and T-lymphocytes was purified by chromatography and from the N-terminal sequence, a full-length clone was obtained from a cDNA library. The protein was named ECF-L (eosinophil chemotactic factor from lymphocytes) and the sequence showed it to be nearly identical to Chi3L3. The recombinant protein was chemotactic *in vitro* and caused extravasation of eosinophils *in vivo*. It was soon shown that Chi3L3 is strongly induced in experimental nematode infection models such as lymphatic filariasis, where a 10,000-fold induction of mRNA expression occurs after implantation of adult *Brugia malayi* into the peritoneal cavity of mice [61], and also in more natural infection models such as oral infection with *Trichinella spiralis* [59], *Nippostrongylus brasiliensis*, and *Litomosoides sigmodontis* [62]. Our own work also indirectly supported the idea of a chemotactic factor, as Ym1 expression correlated with high levels of eosinophil recruitment to the peritoneal cavity [61]. However subsequent work by others failed to find the eosinophil chemotactic activity described originally [60], while others found the *in vitro* chemotactic effect of Chi3L3 to be weak compared with eotaxin, with no significant effect in the lung of Chi3L3/Chi3L4-treated mice [63]. It is also worth mentioning that Chi3L3 is expressed constitutively at high levels in the lung of noninfected healthy mice [64], also confirmed by tissue expression screening by Jin [57] and others [62], and therefore the link between eosinophil recruitment and Chi3L3 presence must be of a more complex nature. Thus it is currently not clear whether eosinophil chemotaxis can truly be ascribed as the physiological function of Chi3L3.

The main cellular source of Chi3L3 in the lungs of healthy and infected mice is the macrophage [61,64,65] and other antigen presenting cells [62]. The induction of Chi3L3 can be obtained by treating macrophages *in vitro* with IL-4 or IL-13, reflecting the Th2-skewed cytokine constellation known to occur in helminthic infection. Accordingly, Chi3L3 induction is now thought to be a

general feature of nematode (and possibly also trematode) infection [62], and this can be explained by the regulatory mechanism involving signal transducer and activator of transcription 6 (stat-6) described by Welch et al. [66], for which the Chi3L3 and Chi3L4 genes possess several canonical binding sites upstream of the 5′-translated region [66]. Thus Chi3L3 has been described as marker of alternative (i.e., Th2 cytokine–dependent, in contrast to classical IFN-γ–dependent) macrophage activation [65,67,68]. Raes has pointed out that this is not the case for human alternatively activated myeloid cells [69].

However, there is likely to be other currently unknown tiers of regulation beyond the stat-6 pathway. Ym1, but not Ym2, is induced, e.g., in the lung in parasitic infection [59,61,62], while Ym2 is induced in the lung in experimental asthma models [63,70]: both are situations in which a Th2 cytokine constellation is predominant, thus one would expect both genes to be induced via stat-6. However this does not seem to be the case, and Chi3L3 vs. Chi3L4 expression at the tissue level appear to be mutually exclusive with the exception of the thymus [57]. Finally, Chi3L3 has recently been implicated as osteoclast-stimulating factor [71] that acts by increasing LFA-1 and ICAM-1 in osteoclast precursors [28].

36.6 CONCLUSION AND FUTURE DIRECTIONS

Chitinases and chilectins share structural features and are found in clusters in mammalians (Figure 36.3), indicating that they may have arisen from gene duplication events before the different species diverged. Chilectins have lost the ability to cleave chitin due to mutations in the active site and do not possess a type 2 chitin-binding domain, but retain the ability to bind carbohydrates and possibly proteins. The physiological ligands of chilectins are still unclear bur appear to be different from chitin or chitin oligomers. Chilectins are expressed in chronic inflammatory conditions, nematode infections and in situations where extensive tissue remodeling occurs. The suggested functions range from growth and differentiation factors to neoangiogenesis, anti-inflammatory, and antiapoptotic factors. Because of their implication in rheumatoid arthritis, asthma, and cancer, chilectins may be novel therapeutic targets. Therefore, a better understanding of their physiological roles, ligands and putative cellular receptors is needed.

ACKNOWLEDGMENTS

S.T. was a recipient of a European Social Fund Oncology Training Award. S.M. and F.H.F. are funded by the Wellcome Trust (GR065978MA).

REFERENCES

1. Kramer, K.J. and Koga, D.: Insect chitin: Physical state, synthesis, degradation and metabolic regulation. *Insect Biochem*. 1986: 16: 851–877.
2. Flach, J., Pilet, P.E., and Jolles, P.: What's new in chitinase research? *Experientia*. 1992: 48: 701–716.
3. Fänge, R., Lundblad, G., and Lind, J.: Lysozyme and chitinase in blood and lymphomyeloid tissues of marine fish. *Mar Biol*. 1976: 36: 277–282.
4. Boot, R.G. et al.: Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J Biol Chem*. 2001: 276: 6770–6778.
5. Boot, R.G. et al.: Marked differences in tissue-specific expression of chitinases in mouse and man. *J Histochem Cytochem*. 2005: 53: 1283–1292.
6. Hoffmann-Sommergruber, K.: Pathogenesis-related (PR)-proteins identified as allergens. *Biochem Soc Trans*. 2002: 30: 930–935.
7. Bakkers, J., Kijne, J.W., and Spaink, H.P.: Function of chitin oligosaccharides in plant and animal development. *Exs*. 1999: 87: 71–83.
8. Bleau, G., Massicotte, F., Merlen, Y. and Boisvert, C.: Mammalian chitinaselike proteins. *Exs*. 1999: 87: 211–221.
9. Watanabe, T. et al.: Identification of glutamic acid 204 and aspartic acid 200 in chitinase A1 of *Bacillus circulans* WL-12 as essential residues for chitinase activity. *J Biol Chem*. 1993: 268: 18567–18572.

10. Houston, D.R. et al.: Structure and ligand-induced conformational change of the 39 kDa glycoprotein from human articular chondrocytes. *J Biol Chem*. 2003: 278: 30206–30212.
11. Fusetti, F. et al.: Crystal structure and carbohydrate-binding properties of the human cartilage glycoprotein-39. *J Biol Chem*. 2003: 278: 37753–37760.
12. Eppig, J.T. et al.: The mouse genome database (MGD): From genes to mice—a community resource for mouse biology. *Nucleic Acids Res*. 2005: 33: D4715.
13. Boot, R.G., Bussink, A.P., and Aerts, J.M.: Human acidic mammalian chitinase erroneously known as eosinophil chemotactic cytokine is not the ortholog of mouse ym1. *J Immunol*. 2005: 175: 2041–2042.
14. Murzin, A.G. et al.: SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol*. 1995: 247: 536–540.
15. Banner, D.W. et al.: Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature*. 1975: 255: 609–614.
16. Terwisscha van Scheltinga, A.C., Hennig, M. and Dijkstra, B.W.: The 1.8 Å resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. *J Mol Biol*. 1996: 262: 243–257.
17. Fusetti, F. et al.: Structure of human chitotriosidase. Implications for specific inhibitor design and function of mammalian chitinase-like lectins. *J Biol Chem*. 2002: 277: 25537–25544.
18. Tsai, M.L., Liaw, S.H. and Chang, N.C.: The crystal structure of Ym1 at 1.31 Å resolution. *J Struct Biol*. 2004: 148: 290–296.
19. Houston, D.R. et al.: The cyclic dipeptide CI-4 [cyclo-(L-Arg-d-Pro)] inhibits family 18 chitinases by structural mimicry of a reaction intermediate. *Biochem J*. 2002: 368: 23–27.
20. Sayle, R.A. and Milner-White, E.J.: RASMOL: Biomolecular graphics for all. *Trends Biochem Sci*. 1995: 20: 374.
21. Suetake, T. et al.: Chitin-binding proteins in invertebrates and plants comprise a common chitin-binding structural motif. *J Biol Chem*. 2000: 275: 17929–17932.
22. Hakala, B.E., White, C., and Recklies, A.D.: Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family. *J Biol Chem*. 1993: 268: 25803–25810.
23. Shackelton, L.M., Mann, D.M., and Millis, A.J.: Identification of a 38 kDa heparin-binding glycoprotein (gp38 k) in differentiating vascular smooth muscle cells as a member of a group of proteins associated with tissue remodeling. *J Biol Chem*. 1995: 270: 13076–13083.
24. Morrison, B.W. and Leder, P.: neu and ras initiate murine mammary tumors that share genetic markers generally absent in c-myc and int-2-initiated tumors. *Oncogene*. 1994: 9: 3417–3426.
25. Mohanty, A.K. et al.: Crystal structure of a novel regulatory 40 kDa mammary gland protein (MGP-40) secreted during involution. *J Biol Chem*. 2003: 278: 14451–14460.
26. Rejman, J.J. and Hurley, W.L.: Isolation and characterization of a novel 39 kilodalton whey protein from bovine mammary secretions collected during the nonlactating period. *Biochem Biophys Res Commun*. 1988: 150: 329–334.
27. Recklies, A.D., White, C., and Ling, H.: The chitinase 3-like protein human cartilage glycoprotein 39 (HC-gp39) stimulates proliferation of human connective-tissue cells and activates both extracellular signal-regulated kinase-and protein kinase B-mediated signalling pathways. *Biochem J*. 2002: 365: 119–126.
28. De Ceuninck, F. et al.: YKL-40 (cartilage gp-39) induces proliferative events in cultured chondrocytes and synoviocytes and increases glycosaminoglycan synthesis in chondrocytes. *Biochem Biophys Res Commun*. 2001: 285: 926–931.
29. Volck, B. et al.: YKL-40, a mammalian member of the chitinase family, is a matrix protein of specific granules in human neutrophils. *Proc Assoc Am Physicians*. 1998: 110: 351–360.
30. Rehli, M., Krause, S.W., and Andreesen, R.: Molecular characterization of the gene for human cartilage gp-39 (CHI3L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation. *Genomics*. 1997: 43: 221–225.
31. Kirkpatrick, R.B. et al.: Induction and expression of human cartilage glycoprotein 39 in rheumatoid inflammatory and peripheral blood monocyte-derived macrophages. *Exp Cell Res*. 1997: 237: 46–54.
32. Renkema, G.H. et al.: Chitotriosidase, a chitinase, and the 39 kDa human cartilage glycoprotein, a chitin-binding lectin, are homologues of family 18 glycosyl hydrolases secreted by human macrophages. *Eur J Biochem*. 1998: 251: 504–509.
33. Bigg, H.F. et al.: The mammalian chitinase-like lectin, YKL-40, binds specifically to type I collagen and modulates the rate of type I collagen fibril formation. *J Biol Chem*. 2006: 281: 21082–21095.
34. Nyirkos, P. and Golds, E.E.: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period. *Biochem J*. 1990: 269: 265–268.

35. Schaeffer, H.J. and Weber, M.J.: Mitogen-activated protein kinases: Specific messages from ubiquitous messengers. *Mol Cell Biol*. 1999: 19: 2435–2444.
36. Franke, T.F., Kaplan, D.R., and Cantley, L.C.: PI3K: Downstream AKTion blocks apoptosis. *Cell*. 1997: 88: 435–437.
37. Malinda, K.M. et al.: Gp38 k, a protein synthesized by vascular smooth muscle cells, stimulates directional migration of human umbilical vein endothelial cells. *Exp Cell Res*. 1999: 250: 168–173.
38. Johansen, J.S. et al.: Serum YKL-40, a new prognostic biomarker in cancer patients? *Cancer Epidemiol Biomarkers Prev*. 2006: 15: 194–202.
39. Tak, P.P. and Bresnihan, B.: The pathogenesis and prevention of joint damage in rheumatoid arthritis: Advances from synovial biopsy and tissue analysis. *Arthritis Rheum*. 2000: 43: 2619–2633.
40. Vos, K. et al.: Raised human cartilage glycoprotein-39 plasma levels in patients with rheumatoid arthritis and other inflammatory conditions. *Ann Rheum Dis*. 2000: 59: 544–548.
41. Vind, I. et al.: Serum YKL-40, a potential new marker of disease activity in patients with inflammatory bowel disease. *Scand J Gastroenterol*. 2003: 38: 599–605.
42. Ling, H. and Recklies, A.D.: The chitinase 3-like protein human cartilage glycoprotein 39 inhibits cellular responses to the inflammatory cytokines interleukin-1 and tumour necrosis factor-alpha. *Biochem J*. 2004: 380: 651–659.
43. Jacques, C. et al.: HC-gp39 contributes to chondrocyte differentiation by inducing SOX9 and type II collagen expressions. *Osteoarthritis Cartilage*. 2007: 15: 138–146.
44. Lefebvre, V. and de Crombrugghe, B.: Toward understanding SOX9 function in chondrocyte differentiation. *Matrix Biol*. 1998: 16: 529–540.
45. Cintin, C. et al.: High serum YKL-40 level after surgery for colorectal carcinoma is related to short survival. *Cancer*. 2002: 95: 267–274.
46. Hogdall, E.V. et al.: High plasma YKL-40 level in patients with ovarian cancer stage III is related to shorter survival. *Oncol Rep*. 2003: 10: 1535–1538.
47. Jensen, B.V., Johansen, J.S., and Price, P.A.: High levels of serum HER-2/neu and YKL-40 independently reflect aggressiveness of metastatic breast cancer. *Clin Cancer Res*. 2003: 9: 4423–4434.
48. Johansen, J.S. et al.: High serum YKL-40 level in patients with small cell lung cancer is related to early death. *Lung Cancer*. 2004: 46: 333–340.
49. Junker, N. et al.: Expression of YKL-40 by peritumoral macrophages in human small cell lung cancer. *Lung Cancer*. 2005: 48: 223–231.
50. Cintin, C. et al.: Serum YKL-40 and colorectal cancer. *Br J Cancer*. 1999: 79: 1494–1499.
51. Johansen, J.S. et al.: Serum YKL-40: A new potential marker of prognosis and location of metastases of patients with recurrent breast cancer. *Eur J Cancer*. 1995: 31A: 1437–1442.
52. Nutt, C.L. et al.: YKL-40 is a differential diagnostic marker for histologic subtypes of high-grade gliomas. *Clin Cancer Res*. 2005: 11: 2258–2264.
53. Pelloski, C.E. et al.: YKL-40 expression is associated with poorer response to radiation and shorter overall survival in glioblastoma. *Clin Cancer Res*. 2005: 11: 3326–3334.
54. Tanwar, M.K., Gilbert, M.R., and Holland, E.C.: Gene expression microarray analysis reveals YKL-40 to be a potential serum marker for malignant character in human glioma. *Cancer Res*. 2002: 62: 4364–4368.
55. Boudreau, N. and Bissell, M.J.: Extracellular matrix signaling: Integration of form and function in normal and malignant cells. *Curr Opin Cell Biol*. 1998: 10: 640–646.
56. Sunderkotter, C. et al.: Macrophages and angiogenesis. *J Leukoc Biol*. 1994: 55: 410–422.
57. Jin, H.M. et al.: Genetic characterization of the murine Ym1 gene and identification of a cluster of highly homologous genes. *Genomics*. 1998: 54: 316–322.
58. Sun, Y.J. et al.: The crystal structure of a novel mammalian lectin, Ym1, suggests a saccharide binding site. *J Biol Chem*. 2001: 276: 17507–17514.
59. Chang, N.C. et al.: A macrophage protein, Ym1, transiently expressed during inflammation is a novel mammalian lectin. *J Biol Chem*. 2001: 276: 17497506.
60. Owhashi, M., Arita, H., and Hayai, N.: Identification of a novel eosinophil chemotactic cytokine (ECF-L) as a chitinase family protein. *J Biol Chem*. 2000: 275: 1279–1286.
61. Falcone, F.H. et al.: A *Brugia malayi* homolog of macrophage migration inhibitory factor reveals an important link between macrophages and eosinophil recruitment during nematode infection. *J Immunol*. 2001: 167: 5348–5354.
62. Nair, M.G. et al.: Chitinase and Fizz family members are a generalized feature of nematode infection with selective upregulation of Ym1 and Fizz1 by antigen-presenting cells. *Infect Immun*. 2005: 73: 385–394.

63. Webb, D.C., McKenzie, A.N., and Foster, P.S.: Expression of the Ym2 lectin-binding protein is dependent on interleukin (IL)-4 and IL-13 signal transduction: Identification of a novel allergy-associated protein. *J Biol Chem*. 2001: 276: 41969–41976.
64. Nio, J. et al.: Cellular expression of murine Ym1 and Ym2, chitinase family proteins, as revealed by in situ hybridization and immunohistochemistry. *Histochem Cell Biol*. 2004: 121: 473–482.
65. Nair, M.G., Cochrane, D.W., and Allen, J.E.: Macrophages in chronic type 2 inflammation have a novel phenotype characterized by the abundant expression of Ym1 and Fizz1 that can be partly replicated in vitro. *Immunol Lett*. 2003: 85: 173–180.
66. Welch, J.S. et al.: TH2 cytokines and allergic challenge induce Ym1 expression in macrophages by a STAT6-dependent mechanism. *J Biol Chem*. 2002: 277: 42821–42829.
67. Raes, G. et al.: FIZZ1 and Ym as tools to discriminate between differentially activated macrophages. *Dev Immunol*. 2002: 9: 151–159.
68. Loke, P. et al.: IL-4 dependent alternatively-activated macrophages have a distinctive in vivo gene expression phenotype. *BMC Immunol*. 2002: 3: 7.
69. Raes, G. et al.: Arginase-1 and Ym1 are markers for murine, but not human, alternatively activated myeloid cells. *J Immunol*. 2005: 174: 6561; author reply 6561–6562.
70. Fajardo, I. et al.: Increased levels of hypoxia-sensitive proteins in allergic airway inflammation. *Am J Respir Crit Care Med*. 2004: 170: 477–484.
71. Oba, Y. et al.: Eosinophil chemotactic factor-L (ECF-L): A novel osteoclast stimulating factor. *J Bone Miner Res*. 2003: 18: 1332–1341.

Index

A

B

C

D

E

F

G

H

I

L

M

N

O

P

R

S

T

V

X

Y

Z

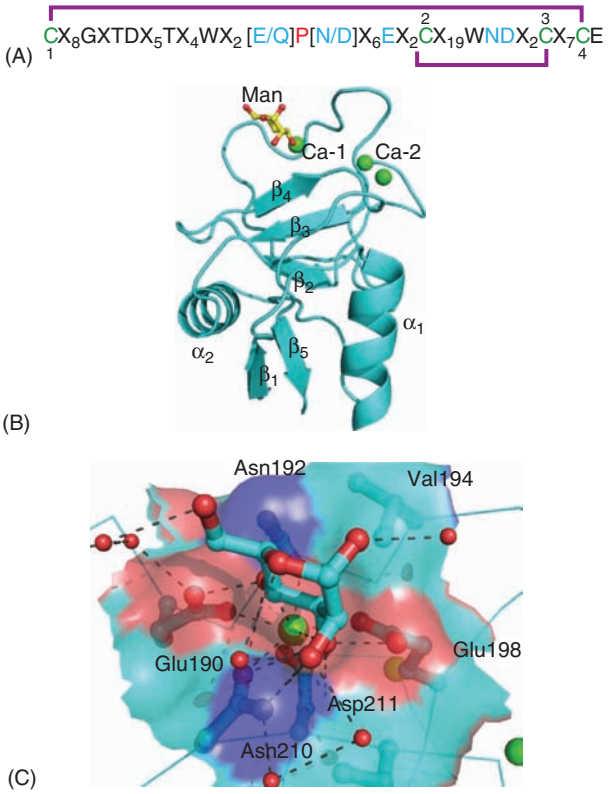


FIGURE 2.1

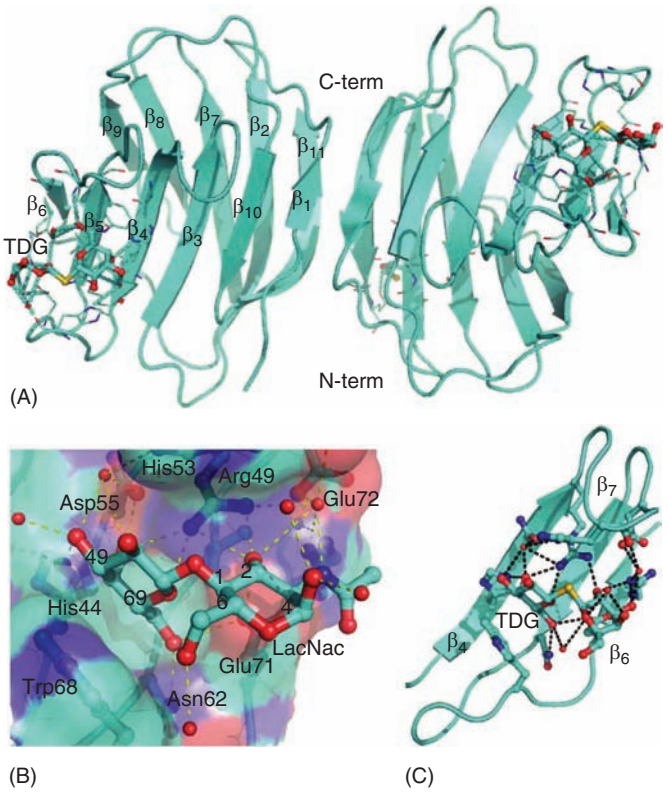


FIGURE 2.2

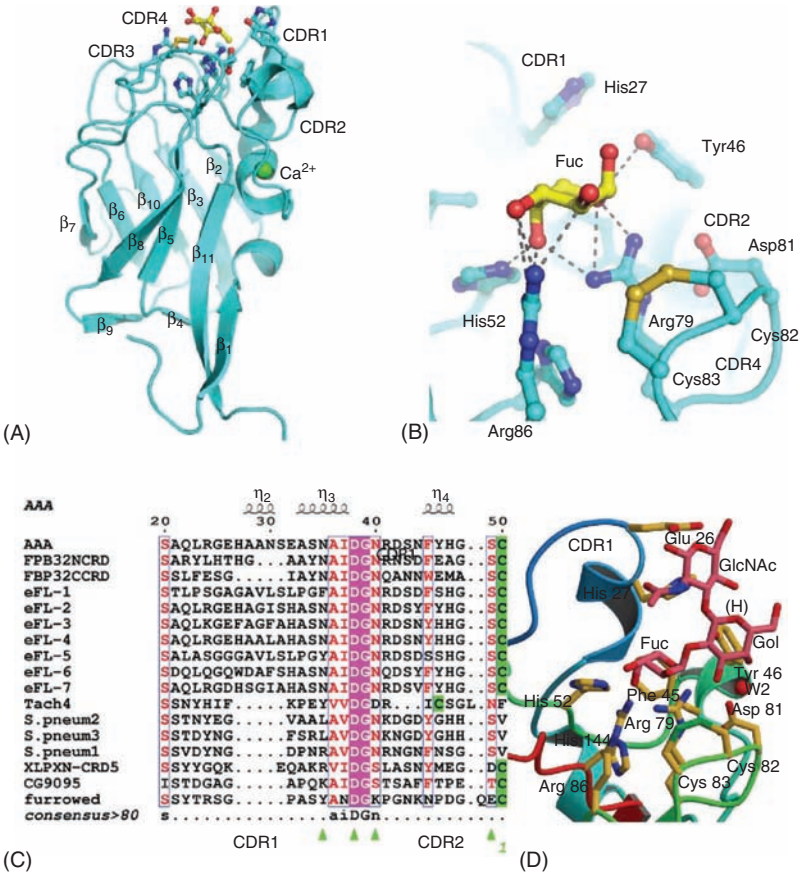


FIGURE 2.4

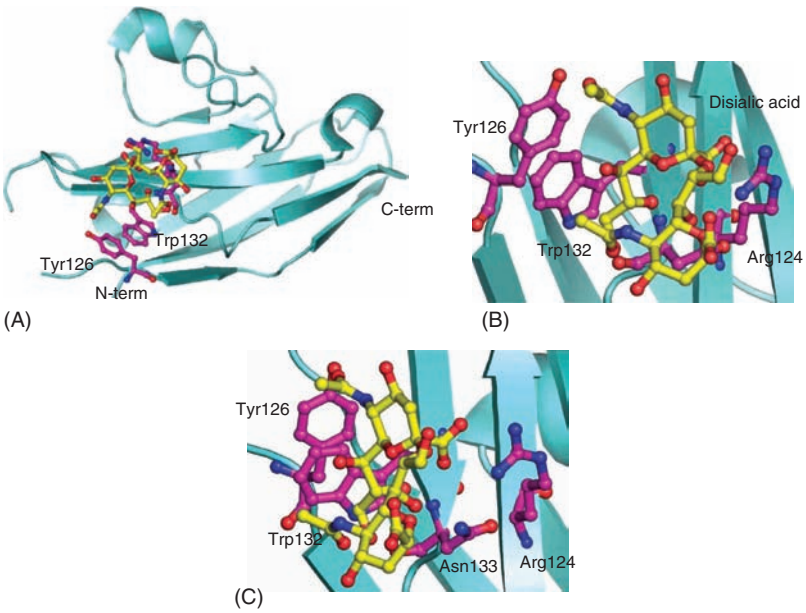


FIGURE 2.5

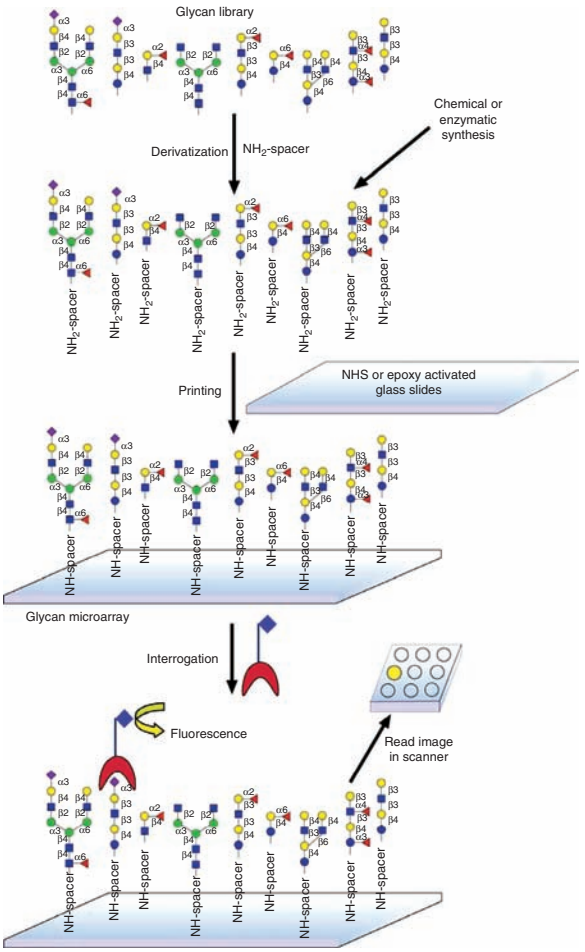


FIGURE 4.3

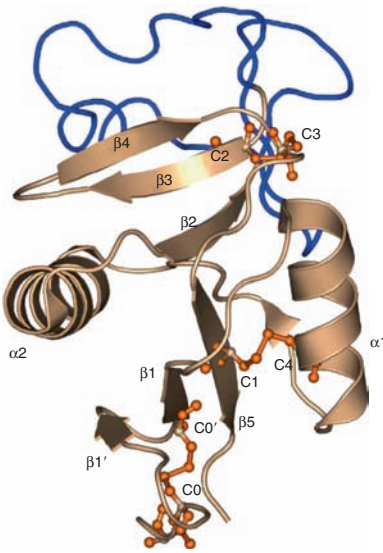
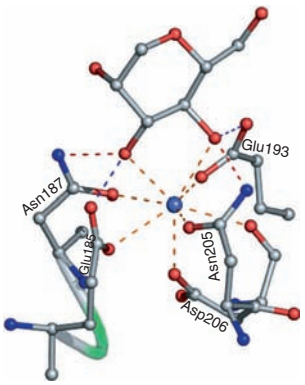
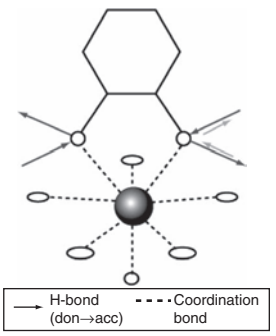


FIGURE 6.1



(A)



(B)

FIGURE 6.3

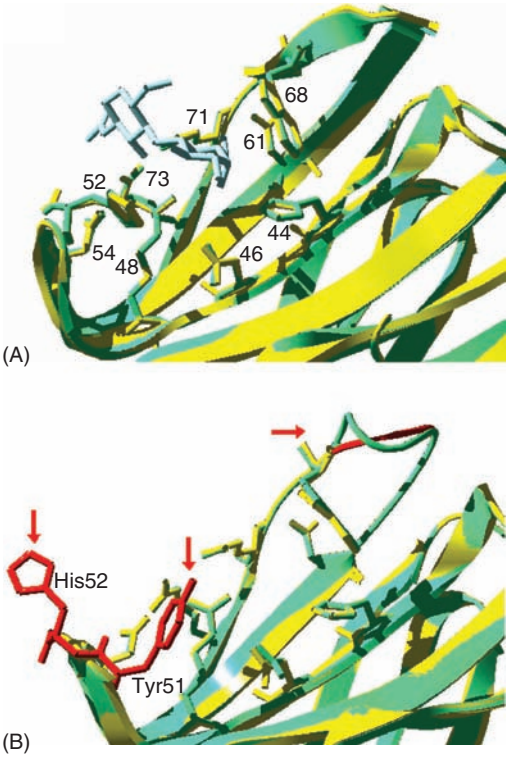


FIGURE 7.1

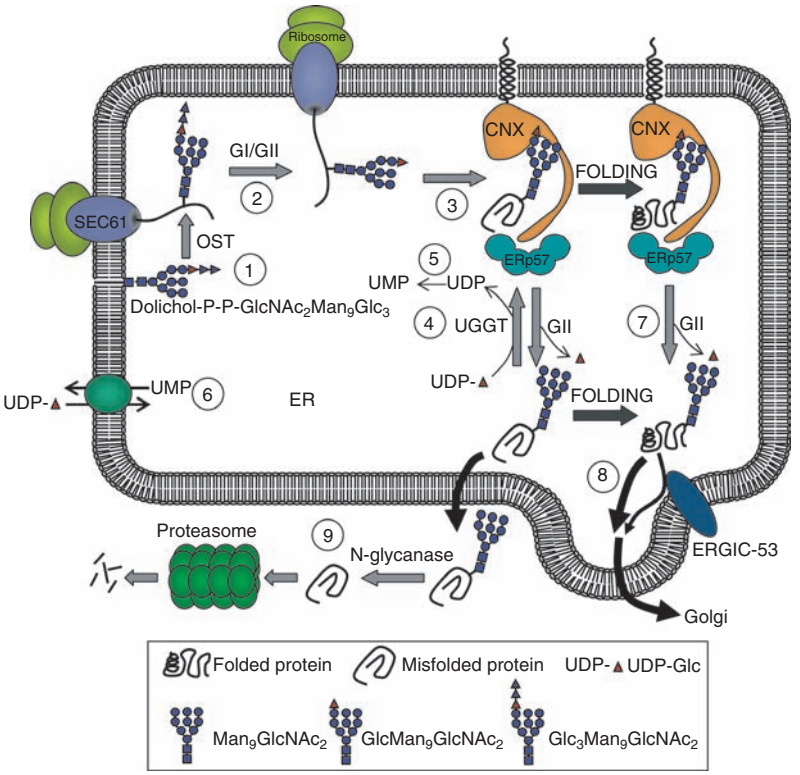


FIGURE 8.3

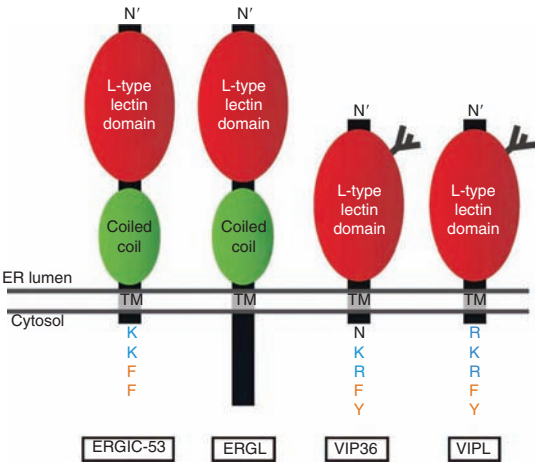


FIGURE 9.1

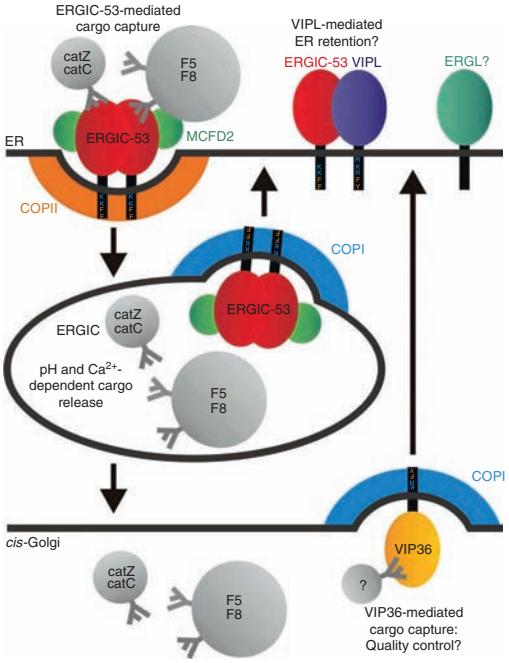


FIGURE 9.2

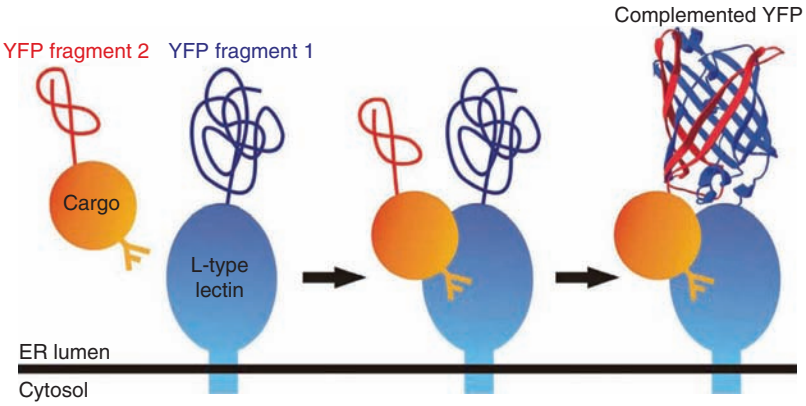


FIGURE 9.3

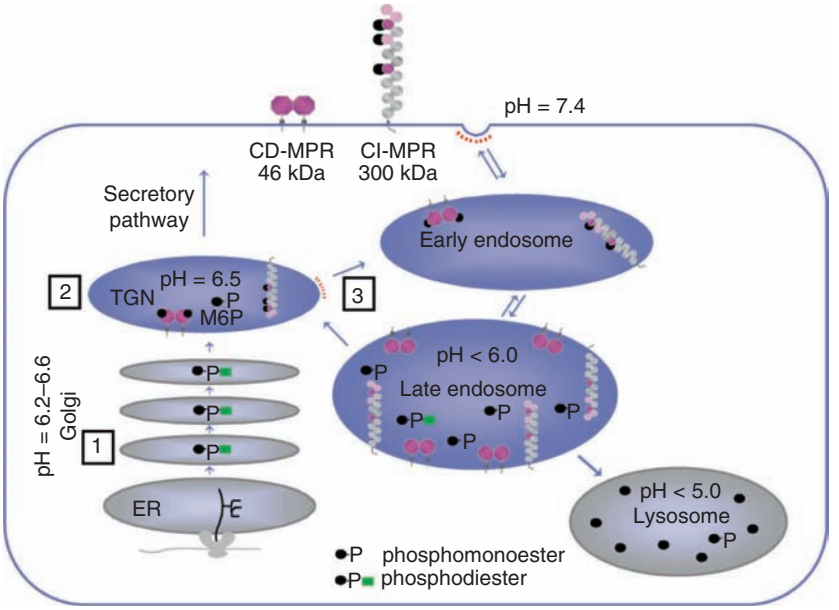


FIGURE 10.3

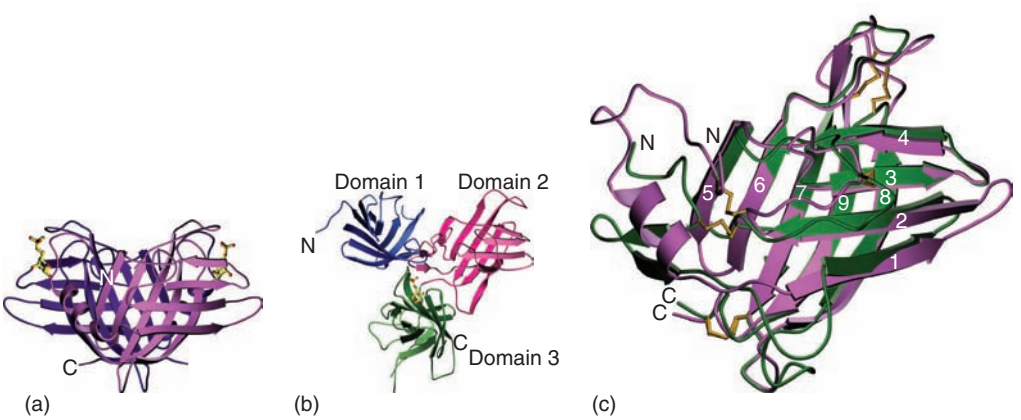


FIGURE 10.4

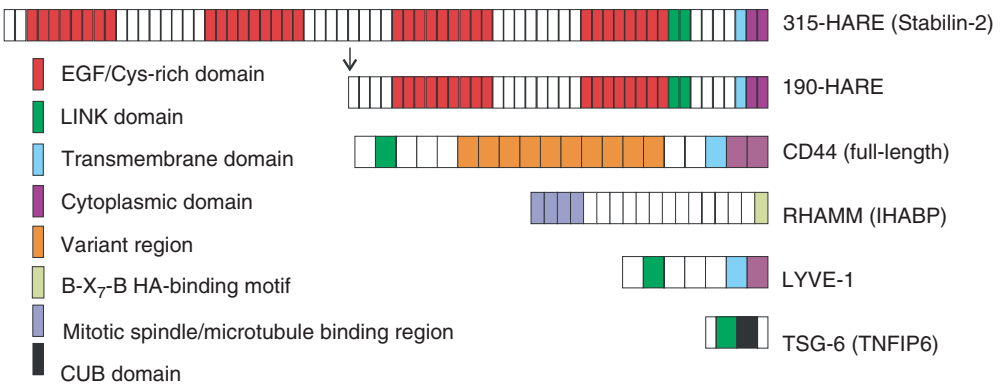


FIGURE 12.2

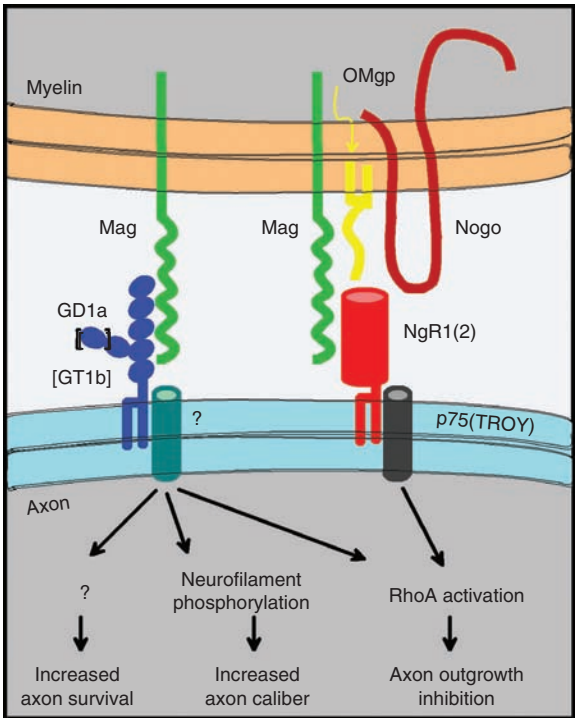


FIGURE 14.5

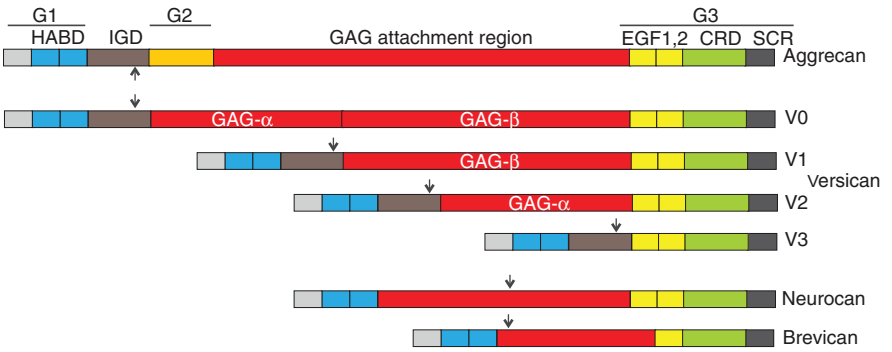
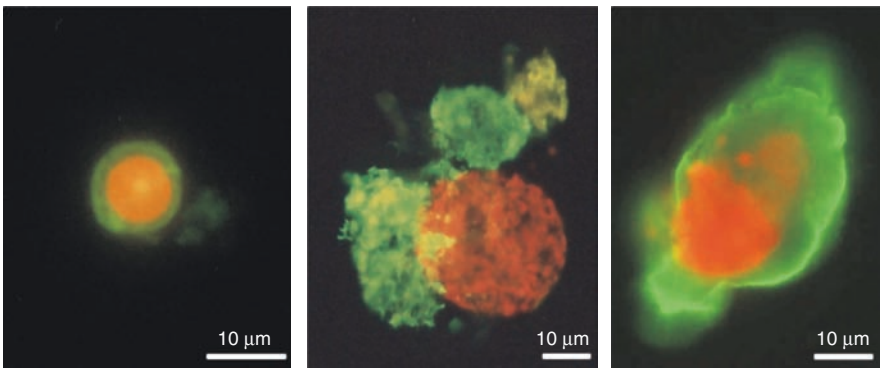
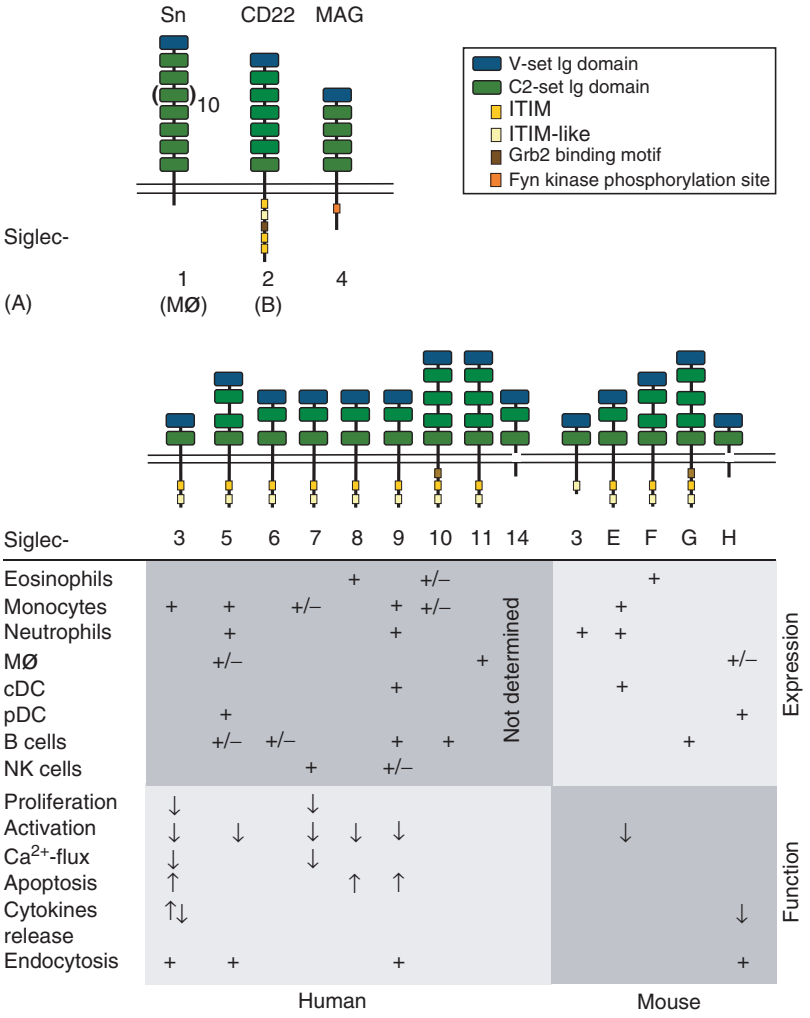


FIGURE 15.1



(A) (B) (C)

FIGURE 16.4



(B) CD33-related Siglecs

FIGURE 17.1

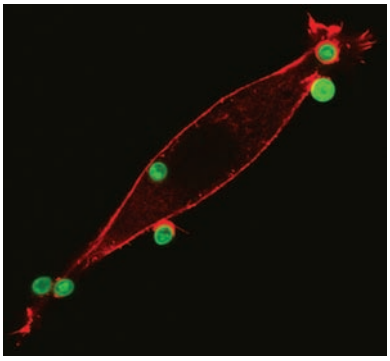


FIGURE 18.3

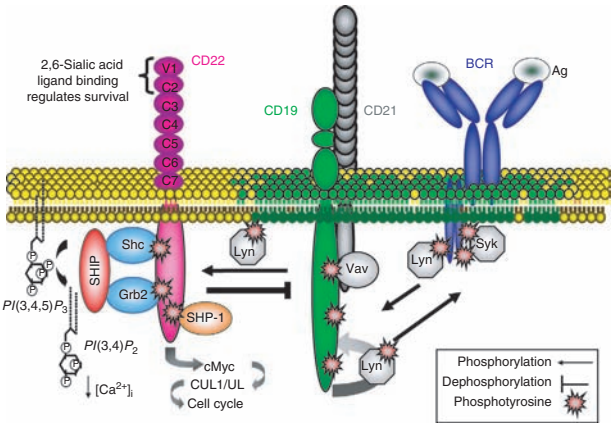


FIGURE 19.1

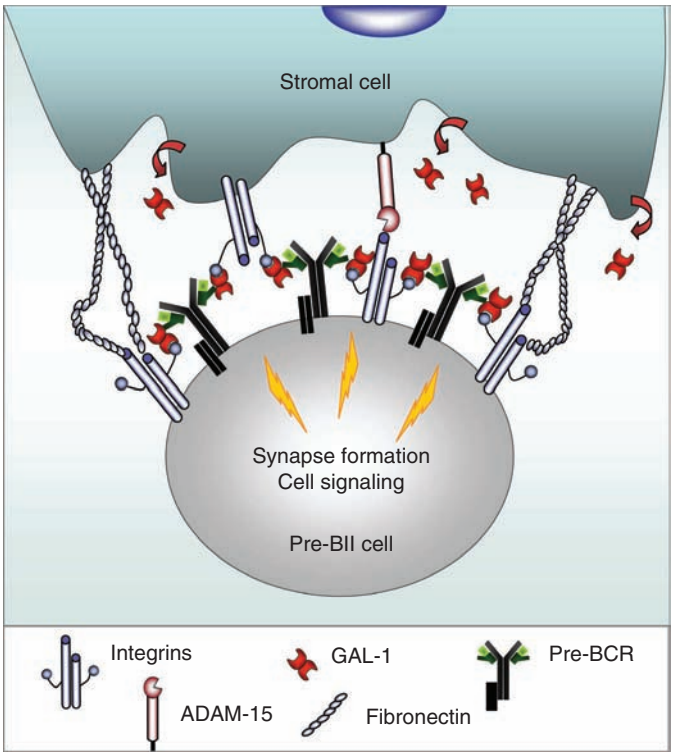


FIGURE 20.2

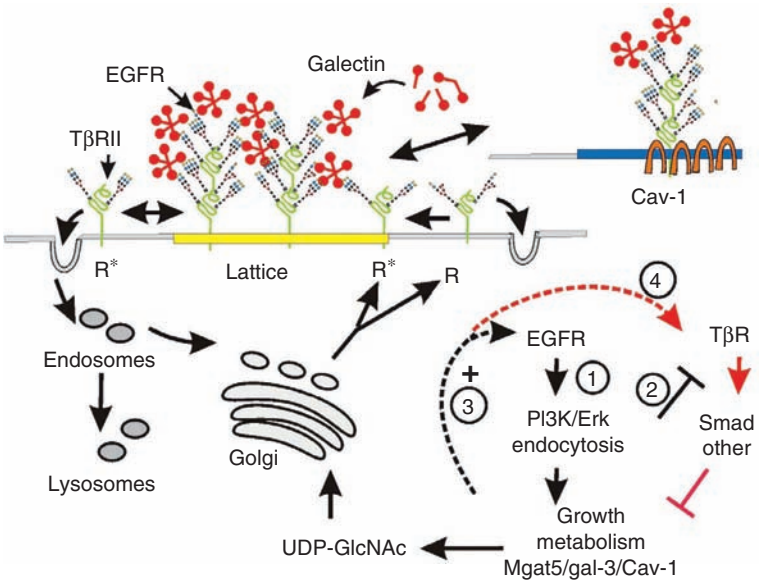


FIGURE 21.3

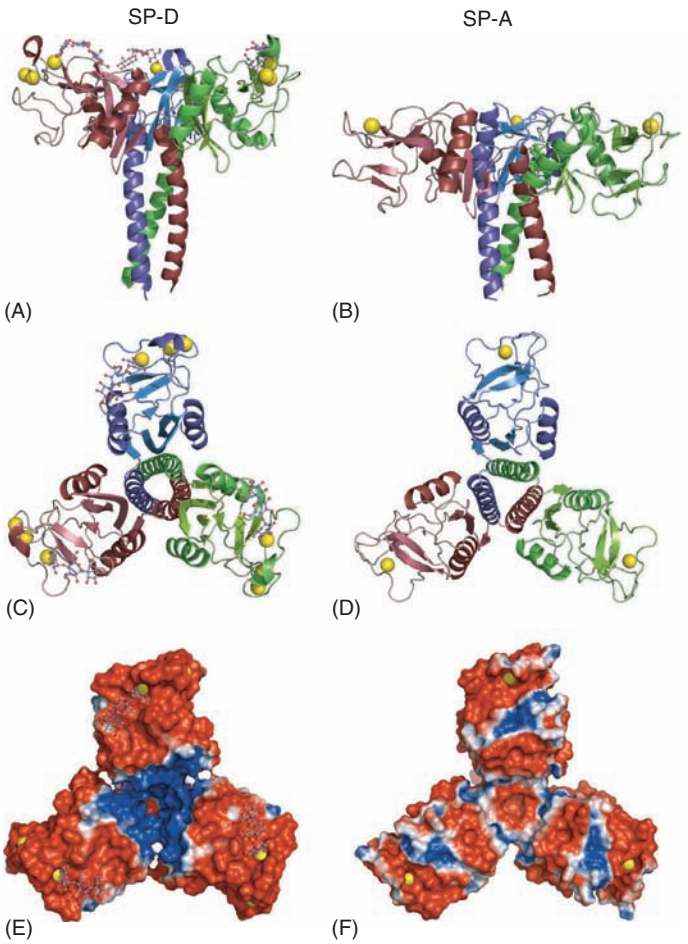


FIGURE 23.3

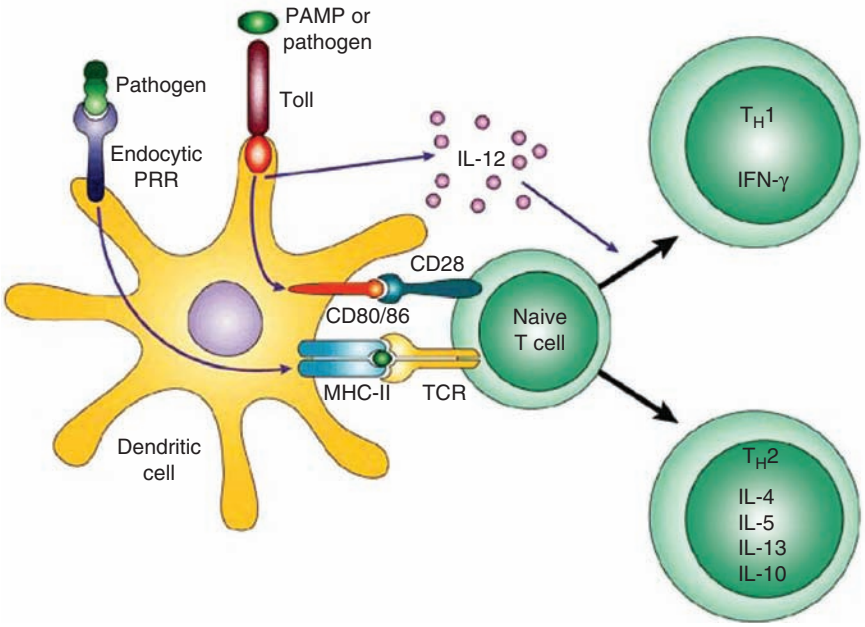


FIGURE 24.1

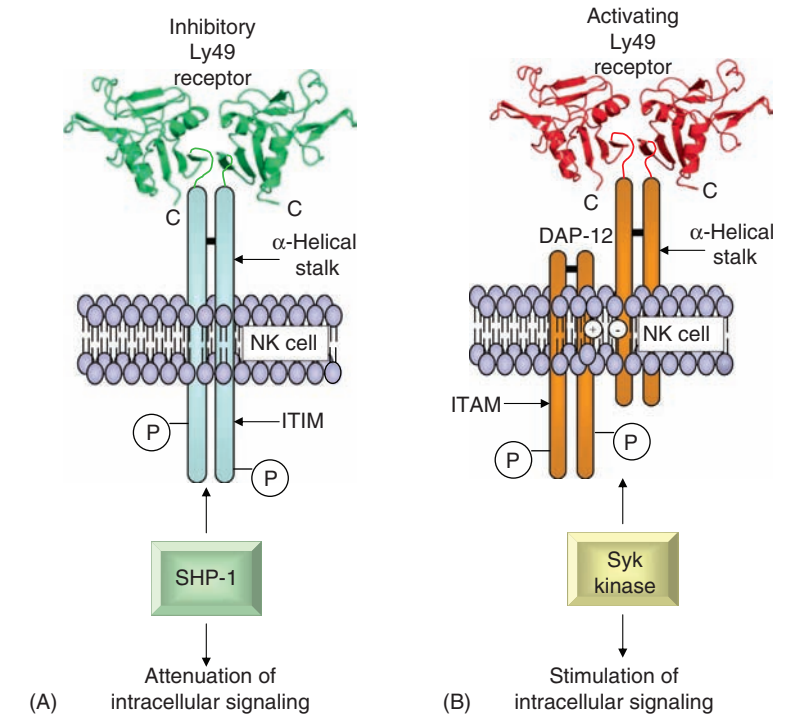


FIGURE 25.2

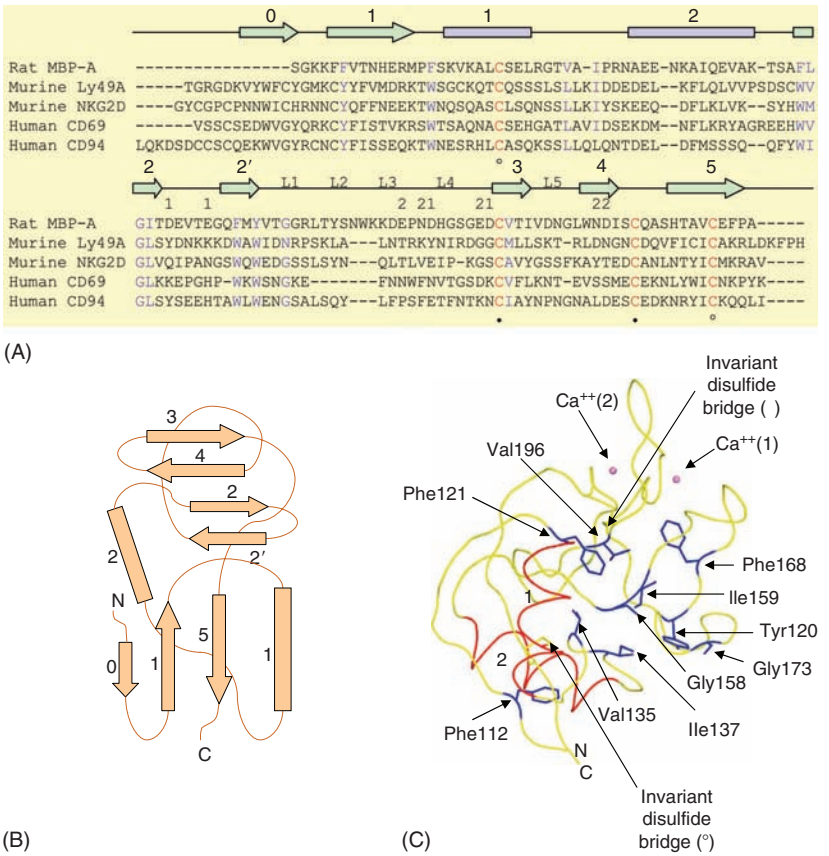


FIGURE 25.3

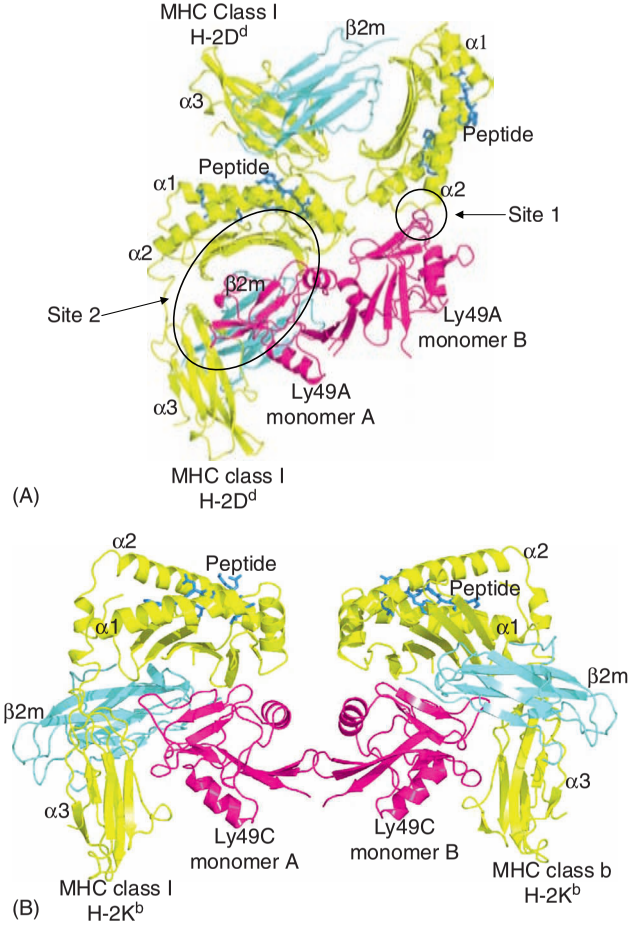


FIGURE 25.6

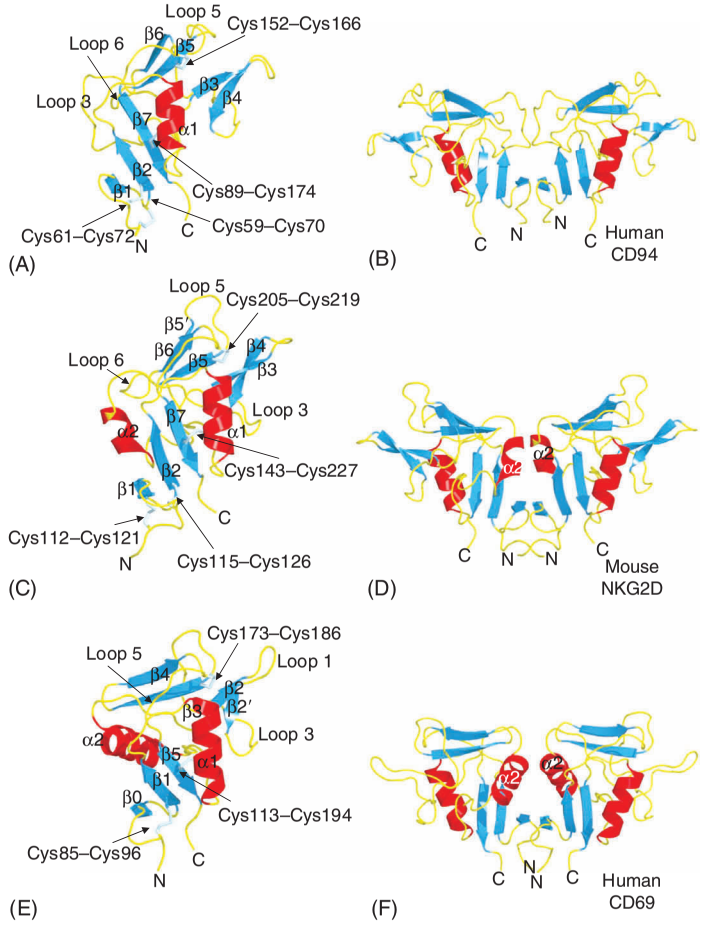


FIGURE 25.7

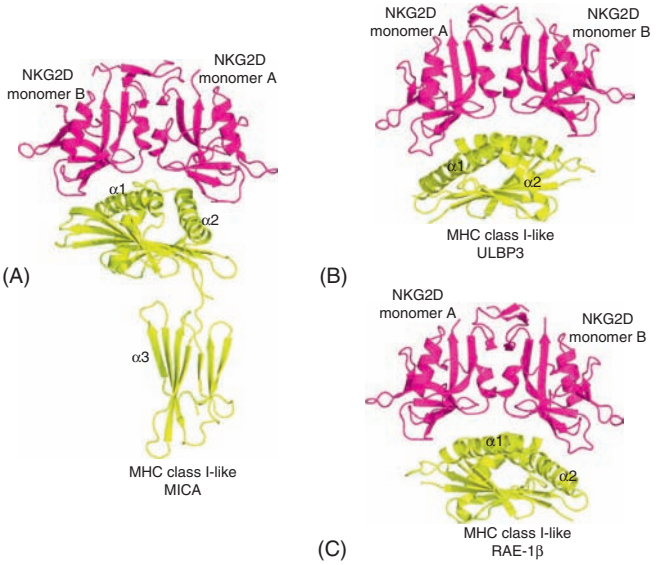


FIGURE 25.8

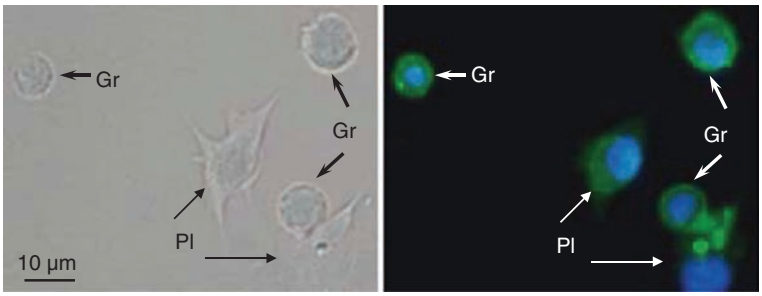


FIGURE 26.4

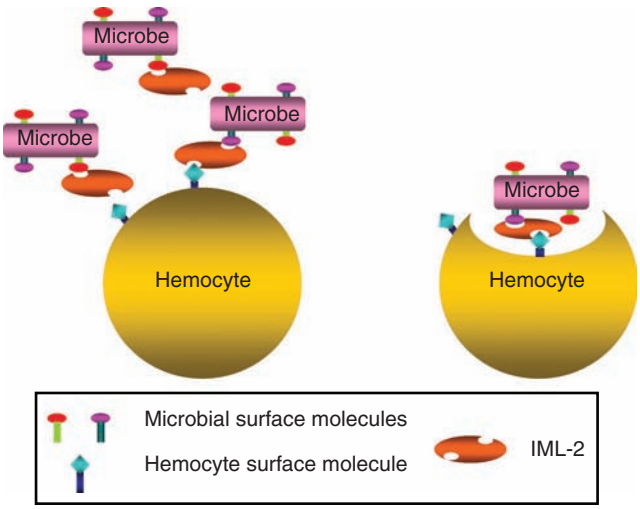


FIGURE 26.5

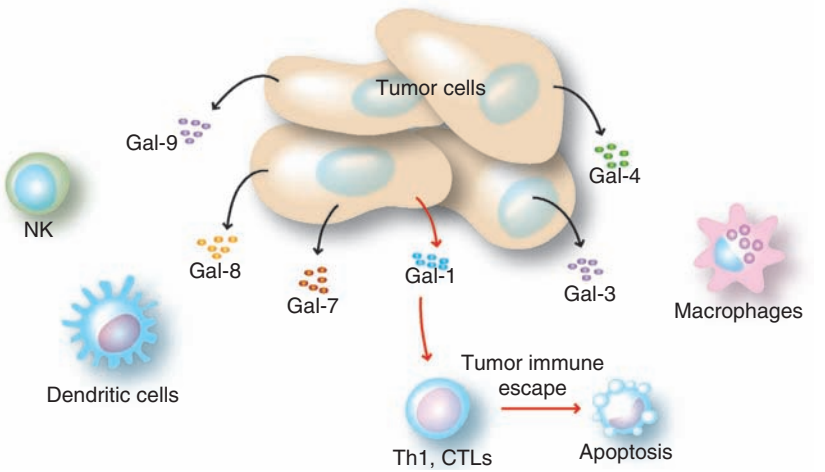


FIGURE 27.3

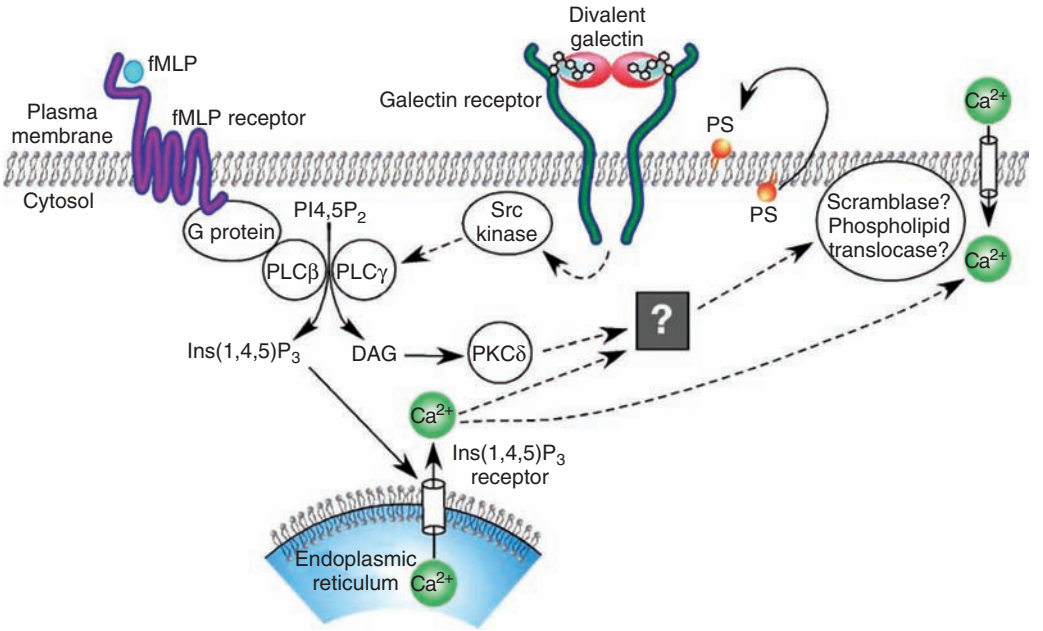


FIGURE 28.1

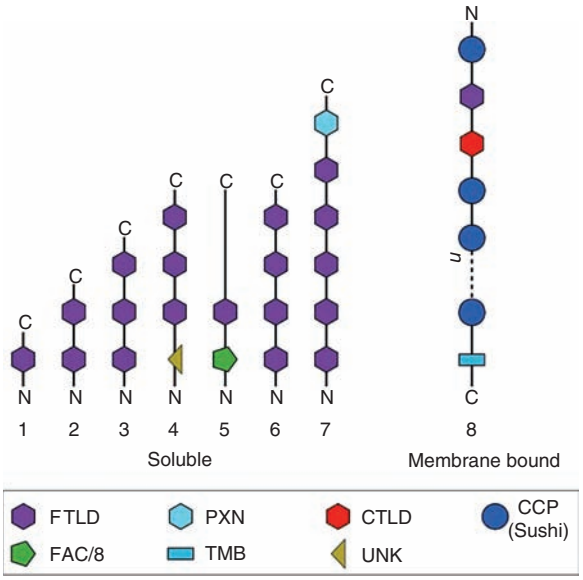


FIGURE 31.2

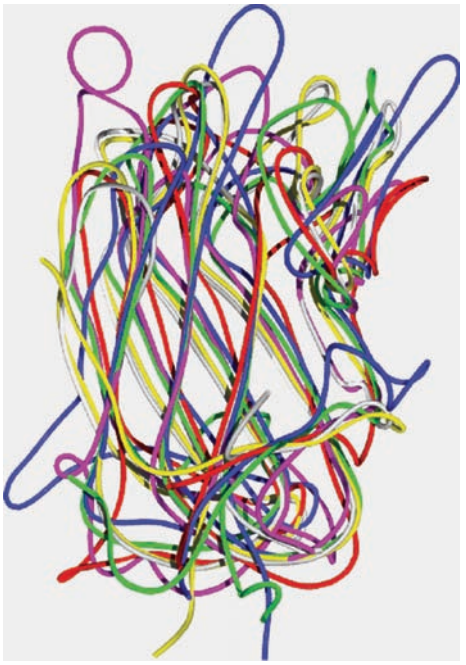


FIGURE 31.3

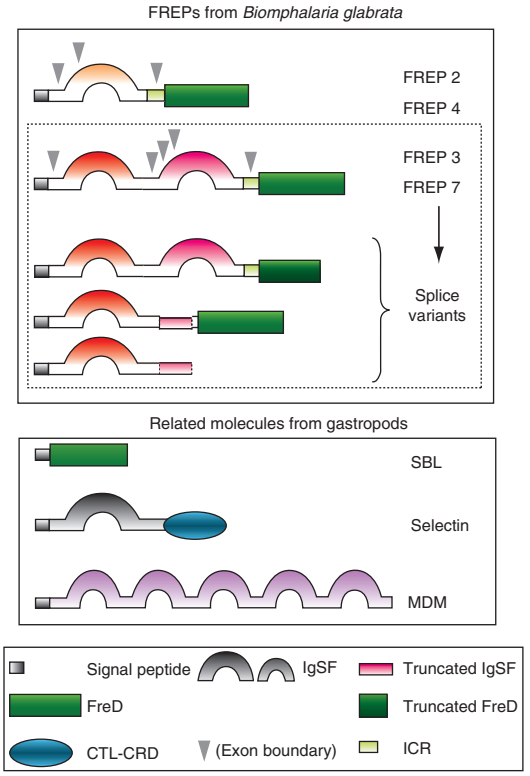


FIGURE 32.4

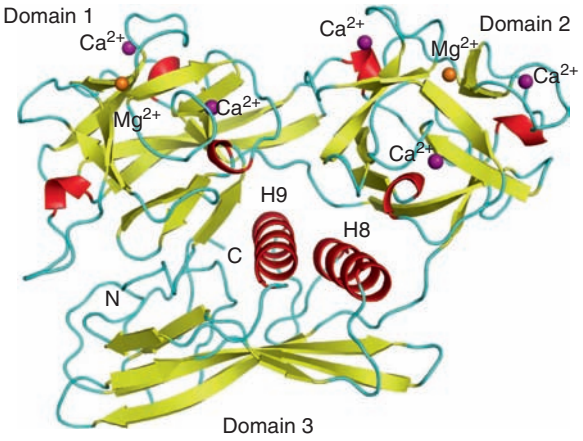


FIGURE 35.4

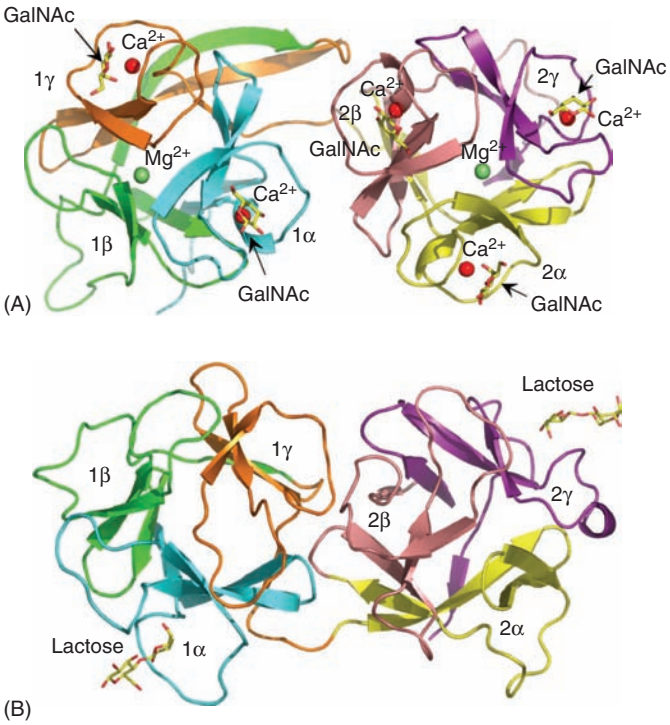


FIGURE 35.5

BIOLOGICAL SCIENCES

ANIMAL LECTINS

A Functional View

Lectins and their ligands are under a heavy microscope due to their potential applications to pharmacology, immunology, cancer therapy, and agriculture. With growing interest in the glycobiology field, the body of research related to lectin roles has grown at an explosive rate, particularly in the past 20 years. ***Animal Lectins: A Functional View*** presents the most up-to-date analysis of these carbohydrate-binding, and potentially lifesaving, proteins in one comprehensive volume.

This standard-setting resource presents new insights into the biological roles of most animal lectins, including their function in infection prevention through innate immunity. It also discusses such lectin behaviors as:

- Glycoprotein folding
- Sorting and targeting
- Cell adhesion
- Embryogenesis
- Cell–cell interactions and signaling

Under the editorial guidance of two leading protein chemistry experts and with contributions from well-recognized field authorities, this book logically and systematically discusses both intracellular and extracellular lectin functions. Describing the structural basis of protein–carbohydrate interactions, this book:

- Examines lectins that participate in glycoprotein folding, sorting and secretion, targeting, degradation, and clearance
- Includes a CD with more than 70 color illustrations
- Studies lectins that are involved in cell–cell interactions, signaling, and transport
- Contains a 16-page color insert
- Addresses lectins that are involved in the recognition and effector functions in innate and adaptive immunity
- Presents recently discovered lectins and lectin families

This text is an essential springboard for future field research and a clear and concise knowledge base that glycobiology, biochemistry, and immunology researchers cannot afford to be without.

7269

ISBN: 978-0-8493-7269-8
90000
9 780849 372698

www.crcpress.com

CRC Press
Taylor & Francis Group
an **informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK