Essentials of Glycobiology

Edited by

Ajit Varki
Richard Cummings
Jeffrey Esko
Hudson Freeze
Gerald Hart
Jamey Marth

With Contributions from

Maarten Chrispeels
Ole Hindsgaul
James C. Paulson
John Lowe
Adriana Manzi
Leland Powell
Herman van Halbeek

COLD SPRING HARBOR LABORATORY PRESS
Cold Spring Harbor, New York

ESSENTIALS OF GLYCOBIOLOGY

Published by Cold Spring Harbor Laboratory Press,
Cold Spring Harbor, New York

Printed in the United States of America

Editorial Editor: Kaaren Janssen
Editorial Coordinators: Mary Cozza, Inez Sialiano
Production Editor: Dorothy Brown
Desktop Editor: Danny deBruin
Book Design: Denise Weiss
Cover Design: Tony Urgo

Front Cover (printed hardcover): A ribbon diagram of the bovine cation-dependent Man-6-P receptor (CD-MPR). The two monomers (purple ribbon and blue ribbon) of the dimer as well as the ligand, Man-6-P (gold ball-and-stick model), are shown. (Modified, with permission, from Roberts et al., *Cell 93:* 639–648 [1998], Figure 4a.)

Back Cover (printed hardcover) High-mannose-type N-glycans can be processed down a variety of biosynthetic pathways, two of which are shown, using the symbol nomenclature recommended in this book. The diphosphorylated glycan shown is the optimal ligand for the mannose 6-phosphate receptors (see front cover).

Library of Congress Cataloging-in-Publication Data

The essentials of glycobiology / edited by Ajit Varki ... et al.].
p. cm.
Includes bibliographical references and index.
ISBN 0-87969-559-5 (cloth). -- ISBN 0-87969-560-9 (printed hardcover)
1. Glycoproteins. 2. Glycosylation. I. Varki, Ajit.
QP552.G59E87 1999
572' .68--dc21 99-17198
CIP

10 9 8 7 6 5 4 3 2 1

All Cold Spring Harbor Laboratory Press publications may be ordered directly from Cold Spring Harbor Laboratory Press, 10 Skyline Drive, Plainview, New York 11803. (Phone: 1-800-843-4388 in Continental U.S. and Canada.) All other locations: (516) 349-1946. E-mail: cshpress@cshl.org. For a complete catalog of all Cold Spring Harbor Laboratory Press publications, visit our World Wide Web Site http://www.cshl.org/

FOREWORD

THE *ESSENTIALS OF GLYCOBIOLOGY* could not appear at a more opportune time, for the field is in a period of enormous progress and the prospects for future advances are even greater. Glycobiology has its roots in the nineteenth century, when chemists first began to analyze sugars and polysaccharides. Perhaps the first glycoprotein to be studied was the "glycogenous matter" of liver which the famous French physiologist Claude Bernard identified in 1855 as a storage form of glucose. (Interestingly, the evidence that glycogen is a glycoprotein was not obtained until more than 100 years later.) Advances in this area continued at a steady rate during most of this century, but the past 20 years has witnessed an unparalleled explosion of new knowledge that has transformed the field. There are many reasons for this acceleration of progress, including great technical advances in establishing oligosaccharide structures, but by far the most important is the application of recombinant DNA technology to the field. This has resulted in the molecular characterization of the enzymes involved in the assembly, processing, and degradation of oligosaccharides and proteoglycans. It has also allowed the identification of numerous families of plant and animal lectins that recognize carbohydrate structures. The surprising finding to emerge is the vast number of enzymes and proteins that are devoted to glycoprotein and glycolipid synthesis and function. The understanding of the biologic roles of glycans has also increased to a great extent, and we now know that these molecules serve multiple functions, ranging from assisting the folding of nascent proteins to determining the trafficking of lymphocytes and granulocytes in the circulation. The important role of glycans is underscored by the growing list of human diseases that are the result of defects in glycan assembly. The challenge for the future is to further define the biologic functions played by glycans. In this regard, recombinant DNA technology has provided another valuable tool for the glycobiologist: the ability to disrupt genes of interest in mice and other organisms. This presents an unparalleled opportunity for the scientist interested in elucidating the biologic roles of sugars. Without a doubt, the future in this area is the brightest it has ever been.

Essentials of Glycobiology provides an ideal entry into the field. It contains the basic information needed to understand this area along with the most current work at the forefront of the field. The authors are to be commended for assembling a broad, comprehensive, well-organized overview of this burgeoning field. They have also been successful in conveying the excitement in this area of research.

Stuart Kornfeld
Washington University School of Medicine
March 1999

PREFACE

EMERGING FROM ITS ROOTS in classical carbohydrate chemistry and biochemistry, glycobiology has become a vibrant, expanding, and important extension of modern molecular biology. Over the years, many outstanding monographs and books have documented important advances in this area and summarized critical methods and concepts (see listing following this preface). These volumes continue to serve as excellent resources for those interested in glycobiology. Why then should one publish an additional book on the subject? Most of these prior volumes have been directed at the specialist, assuming a substantial level of technical sophistication and expertise and a working knowledge of the relevant jargon. We present here a book that seeks to fulfill a somewhat different need: to summarize the current state of the art for the expert and yet serve as a resource for the novice wishing to explore the essentials of glycobiology.

This book had its origins from some independent lines of effort. For several years, some of us have been teaching a short elective course in glycobiology for graduate students at the University of California, San Diego. With the recent arrival of additional faculty with expertise in this field, it was decided to present a more comprehensive course on the subject, to be supplemented by a course book that could be then converted into a formal text. Meanwhile, other experts elsewhere in the country had put forward independent proposals to fill the perceived need for a basic textbook in glycobiology. Following a discussion over a beer after a glycobiology conference, we decided to pool all our efforts in this direction.

Since a major goal was to produce a text that would be accessible to students and other trainees, we used the 1998 *UCSD Spring Quarter Graduate Course in Glycobiology* as the basis for creating the text. By recruiting several additional experts as lecturers, we could present a comprehensive course that covered most aspects of the field. Each lecturer was asked to provide handouts for the students that were essentially the first drafts of chapters for the book. In turn, each student was required to provide anonymous critiques of some chapters as a part of the course requirement. This approach ensured not only that the draft chapters were written early on, but also that they underwent in-depth evaluation by bright young minds with an expressed interest in the field. Additional rounds of internal review by the group of six editors served to produce what we hope will be a valuable resource not only for the expert in the field, but also for the novice who wants to learn about glycobiology. We have tried to be as accurate and up to date as possible and to present a balanced point of view on controversial subjects. Given the current breadth of knowledge, it was not possible to do full justice to all aspects of the field, nor to comprehensively reference the extensive literature that exists. The relative emphasis on vertebrate biology bespeaks the greater volume of information currently available in this area of glycobiology.

The Editors are indebted to many others who made this book possible. Besides the students who took the course for credit, several other trainees audited the course and provided very useful feedback. Although the editors wrote the majority of the chapters in the

book, the efforts of the other lecturer/authors were crucial in assuring the depth of expertise needed to cover the field effectively. Special thanks are due to John Inglis and Kaaren Janssen at Cold Spring Harbor Laboratory Press for realizing the potential of this book, and for putting up with our many demands and idiosyncrasies. We also thank the Press staff, Jan Argentine, Inez Sialiano, Mary Cozza, Denise Weiss, Dotty Brown, and Danny deBruin, who deserve much credit for keeping us on track and converting our efforts into an attractive product. Last but not least, we acknowledge our families, lab members, and administrative assistants who supported us through all of the hard work needed to create this text. It now remains for the reader to decide if we have achieved our goals in producing this book.

Ajit Varki
for
The Editors

BOOKS AND MONOGRAPH RESOURCES

Gottschalk A., ed. 1960. *The chemistry and biology of sialic acids and related substances.* Cambridge University Press, Cambridge, United Kingdom.

Ginsburg V. and Neufeld E., eds. 1966. Complex carbohydrates, part A. *Methods Enzymol.*, vol. 8. Academic Press, San Diego, California.

Whistler R., ed. 1968–80. *Methods in carbohydrate chemistry*, vols. I–VIII. Academic Press, San Diego, California.

Ginsburg V., ed. 1972. Complex carbohydrates, part B. *Methods Enzymol.*, vol 28. Academic Press, San Diego, California.

Gottschalk A., ed. 1972. *Glycoproteins: Their composition, structure and function.* Elsevier, New York.

Rosenberg A. and Schengrund C.-L., eds. 1976. *Biological roles of sialic acids.* Plenum Press, New York.

Ginsburg V., ed. 1978. Complex carbohydrates, part C. *Methods Enzymol.*, vol. 50. Academic Press, San Diego, California.

Sweeley C.C., ed. 1979. *Cell surface glycolipids.* American Chemical Society, Washington, D.C.

Lennarz W.J., ed. 1980. *The biochemistry of glycoproteins and proteoglycans.* Plenum Press, New York.

Ginsburg V. and Robbins P., eds. 1981. *Biology of carbohydrates*, vol. 1. Wiley, New York.

Ginsburg V., ed. 1982. Complex carbohydrates, part D. *Methods Enzymol.*, vol. 83. Academic Press, San Diego, California.

Horowitz M. and Pigman W., eds. 1982. *The glycoconjugates.* Academic Press, New York.

Schauer R., ed. 1982. *Sialic acids, chemistry, metabolism, and function.* Springer-Verlag, New York.

Ivatt R.J., ed. 1984. *The biology of glycoproteins.* Plenum Press, New York.

Ginsburg V. and Robbins P., eds. 1985. *Biology of carbohydrates*, vol. 2. Wiley, New York.

Beeley J.G., ed. 1985. *Glycoprotein and proteoglycan techniques.* Elsevier, Amsterdam, The Netherlands.

Liener I.E., Sharon N., and Goldstein I.J., eds. 1986. *The lectins: Properties, functions, and applications in biology and medicine.* Academic Press, Orlando, Florida.

Chaplin M.F. and Kennedy J.F., eds. 1987. *Carbohydrate analysis: A practical approach.* IRL Press, Oxford, United Kingdom.

Ginsburg V., ed. 1987. Complex carbohydrates, part E. *Methods Enzymol*, vol. 138. Academic Press, San Diego, California.

Feizi T. 1989. *Carbohydrate recognition in cellular function.* Ciba Foundation Symposium, vol. 145. Wiley, New York.

Evered D. and Whelan J., eds. 1989. *The biology of hyaluronan.* Ciba Foundation Symposium, vol. 143. Wiley, New York.

Ginsburg V., ed. 1989. Complex carbohydrates, part F. *Methods Enzymol.*, vol. 179. Academic Press, San Diego, California.

Greiling H. and Scott J.E., eds. 1989. *Keratan sulphate: Chemistry, biology, chemical pathology*. The Biochemical Society, London, United Kingdom.

Margolis R.U. and Margolis R.K., eds. 1989. *Neurobiology of glycoconjugates*. Plenum Press, New York.

Sharon N. and Lis H., eds. 1989. *Lectins*. Chapman and Hall, London, United Kingdom.

Lane D.G. and Lindahl U., eds. 1990. *Heparin: Chemical and biological properties*. CRC Press, Boca Raton, Florida.

Ginsburg V. and Robbins P., eds. 1991. *Biology of carbohydrates*, vol. 3. Wiley, New York.

Fukuda M., ed. 1992. *Cell surface carbohydrates and cell development*. CRC Press, Boca Raton, Florida.

Allen H.J. and Kisailus E.C., eds. 1992. *Glycoconjugates: Composition, structure, and function*. Dekker, New York.

Roth J., Rutishauser U., and Troy F., eds. 1992. *Polysialic acids*. Birkhauser Verlag, Basel, Switzerland.

Fukuda M., ed. 1992. *Glycobiology: A practical approach*. IRL Press, Oxford, United Kingdom.

Lennarz W.J. and Hart G.W., eds. 1993. Guide to techniques in glycobiology. *Methods Enzymol.*, vol. 230. Academic Press, San Diego, California.

Roberts D.D. and Mecham R.P., eds. 1993. *Cell surface and extracellular glycoconjugates: Structure and function*. Academic Press, San Diego.

Bock K. and Clausen H., eds. 1994. *Complex carbohydrates in drug research: Structural and functional aspects*. Munksgaard, Copenhagen, Denmark.

Fukuda M. and Hindsgaul O., eds. 1994. *Molecular glycobiology*. Oxford University Press, New York.

Alavi A. and Axford J.S. 1995. *Advances in experimental medicine and biology*, vol. 376, *Glycoimmunology*. Plenum Press, New York.

Montreuil J., Vliegenthart J.F.G., and Schachter H., eds. 1995. *Glycoproteins*. Elsevier, New York.

Rosenberg A., ed. 1995. *Biology of the sialic acids*. Plenum Press, New York.

Verbert A., ed. 1995. *Methods on glycoconjugates: A laboratory manual*. Harwood Academic Publishers, Switzerland.

Montreuil J., Vliegenthart J.F.G., and Schachter H., eds. 1996. *Glycoproteins and disease*. Elsevier, New York.

Townsend R.R. and Hotchkiss A.T., eds. 1997. *Techniques in glycobiology*. Marcel Dekker, New York.

Gabius H.J. and Gabius S., eds. 1997. *Glycosciences: Status and perspectives*. Chapman and Hall, New York.

Brockhausen I. and Kuhns W. 1997. *Glycoproteins and human disease*. R.G. Landes, Austin.

Montreuil J., Vliegenthart J.F.G., and Schachter H., eds. 1997. *Glycoproteins II*. Elsevier, New York.

Conrad. H.E., ed. 1998. *Heparin-binding proteins*. Academic Press, San Diego, California.

Hounsell E.F., ed. 1998. *Methods in molecular biology*, vol. 76, *Glycoanalysis protocols*. Humana Press, Totowa, New Jersy.

CONTRIBUTORS

Maarten Chrispeels, Professor of Biology
Member, Glycobiology Research and Training Center
University of California, San Diego, La Jolla, California

Richard Cummings, Professor of Biochemistry
Ed Miller Endowed Professor of Biochemistry and Molecular Biology
The University of Oklahoma, Oklahoma City, Oklahoma

Jeffrey Esko, Professor of Medicine
Associate Director, Glycobiology Research and Training Center
University of California, San Diego, La Jolla, California

Hudson Freeze, Professor of Glycobiology
The Burnham Institute, La Jolla, California

Gerald Hart, Director and Professor of Biological Chemistry
Johns Hopkins University School of Medicine, Baltimore, Maryland

Ole Hindsgaul, Professor of Chemistry
University of Alberta, Edmonton, Canada

John Lowe, Professor of Pathology, HHMI Investigator
University of Michigan, Ann Arbor, Michigan

Adriana Manzi, Associate Director, Analytical Development/Quality Control
Nextran, A Division of Baxter Healthcare, San Diego, California

Jamey Marth, Professor of Medicine, HHMI Associate Investigator
Member, Glycobiology Research and Training Center
University of California, San Diego, La Jolla, California

James C. Paulson, Professor of Molecular Biology
The Scripps Research Institute, La Jolla, California

Leland Powell, Adjunct Professor of Medicine
University of California, Los Angeles, California

Herman van Halbeek, Director, Glycotechnology Core Resource
Member, Glycobiology Research and Training Center
University of California, San Diego, La Jolla, California

Ajit Varki, Professor of Medicine
Director, Glycobiology Research and Training Center
University of California, San Diego, La Jolla, California

ABBREVIATIONS

ADP	adenosine diphosphate
AIDS	acquired immunodeficiency syndrome
ASGPR	asialoglycoprotein receptor
ATP	adenosine triphosphate
bFGF	basic fibroblast growth factor
β-HCG	β-human chorionic gonadotrophin
CCSD	complex carbohydrate structure database
CDA	congenital dyserythropoietic anemia
CD-MPR	cation-dependent Man-6-P receptor
CDGS	carbohydrate-deficient glycoprotein syndrome
CHO	Chinese hamster ovary
CI-MPR	cation-independent Man-6-P receptor
CL-43	collectin-43
CMV	cytomegalovirus
CMP	cytidine monophosphate
CNS	central nervous system
CoA	coenzyme A
ConA	concanavalin A
CRD	carbohydrate-recognition domain
CS	chondroitin sulfate
CTD	carboxy-terminal domain
Dol	dolichol
DS	dermatan sulfate
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGF	epidermal growth factor
eIF-2	elongation factor 2
Endo H	endo-β-N-acetylglucosaminidase
EPO	erythropoietin
ER	endoplasmic reticulum
ES	embryonic stem (cells)
FBS	fetal bovine serum
FGF	fibroblast growth factor
FITC-WGA	fluorescein isothiocyanate–wheat-germ agglutinin
Fru	fructose
Fuc	fucose
Gal	galactose
GAG	glycosaminoglycan
GalNAc	N-acetylgalactosamine

GC-MS	gas-liquid chromatography/mass spectrometry
GDP	guanine diphosphate
GlcA	glucuronic acid
Glc	glucose
Glc-Cer	glucosylceramide
GlcNAc	N-acetylglucosamine
GlcN	glucosamine
GM-CSF	granulocyte macrophage–colony-stimulating factor
GMP	guanine monophosphate
GPI	glycosylphosphatidylinositol
HA	hyaluronic acid
HDL	high-density lipoprotein
HEMPAS	hereditary erythrocytic multinuclearity with positive acidified serum test
HexA	Hexuronic acid
HexNAc	N-acetylhexosamine
HIV-1	human immunodeficiency virus type 1
HPLC	high-pressure liquid chromatography
HS	heparan sulfate
HSPG	heparan sulfate proteoglycan
ICE	interleukin converting enzyme
IdoA	iduronic acid
IGF-II	insulin-like growth factor-II
IgG	immunoglobulin G
IgSF	immunoglobulin superfamily
KDO	keto-deoxyoctulosonic acid
KDN	keto-deoxynonulosonic acid
KS	keratan sulfate
LAMPs	lysosome-associated membrane proteins
LC-MS	liquid chromatography coupled to mass spectrometry
LCO	lipochitooligosaccharides
LDL	low-density lipoprotein
Le^a	Lewis A
Le^b	Lewis B
Le^x	Lewis X
LIF	leukemia-inhibitory factor
LLO	lipid-linked oligosaccharide
LPG	lipophosphoglycan
LPS	lipopolysaccharide
MAG	myelin-associated glycoprotein
Man	mannose
MBP	mannose-binding protein
MDO	membrane-derived oligosaccharides
MHC	major histocompatibility complex
MPR	mannose-6-phosphate receptor
MurNAc	N-acetylmuramic acid
NAD	nicotinamide adenine dinucleotide, oxidized
NADP	nicotinamide adenine dinucleotide phosphate, oxidized
NADPH	nicotinamide adenine dinucleotide phosphate, reduced

N-CAM	neural cell adhesion molecule
NeurAc	N-acetyl neuraminic acid
NGF	nerve growth factor
NK	natural killer cells
NMR	nuclear magnetic resonance
OST	oligosaccharyltransferase
PAPS	3′ phosphoadenyl-5′phosphosulfate
PCR	polymerase chain reaction
PDGF	platelet-derived growth factor
PDMP	d/l-threo-1-phenyl-2-decanoylamino-3-morpholino-1-propanol
PHA	phytohemagglutinin
PMM	phosphomannomutase
PMI	phosphomannose isomerase
PNA	peanut agglutinin
PNGase	peptide: N-glycosidase
PNH	paroxysmal nocturnal hemoglobinuria
PNS	peripheral nervous system
PSA	polysialic acid
PSGL-1	P-selectin glycoprotein ligand-1
Rib	ribose
RCA	*Ricinus communis* agglutinin (I or II)
RP-HPLC	reverse-phase high-pressure liquid chromatography
SAPs	sphingolipid activator proteins
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
Sia	sialic acid
SMP	Schwann cell myelin protein
SNA	*Sambucus nigra* agglutinin
SSEA	stage-specific embryonic antigens
SV40	simian virus 40
sVSG	soluble-variant surface glycoprotein
TF antigen	Thomsen-Friedenreich antigen
TGF-β	transforming growth factor-β
TH1	T-helper-1 (cells)
TLC	thin-layer chromatography
t-PA	tissue plasminogen activator
UDP	uridine diphosphate
UMP	uridine monophosphate
VSG	variant surface glycoprotein
WGA	wheat-germ agglutinin
Xyl	xylose

CONTENTS

GENERAL PRINCIPLES

BIOSYNTHESIS, METABOLISM, AND FUNCTION

PROTEINS THAT RECOGNIZE GLYCANS

GLYCANS IN GENETIC DISORDERS AND DISEASE

METHODS AND APPLICATIONS

CHAPTER 1

Historical Background and Overview

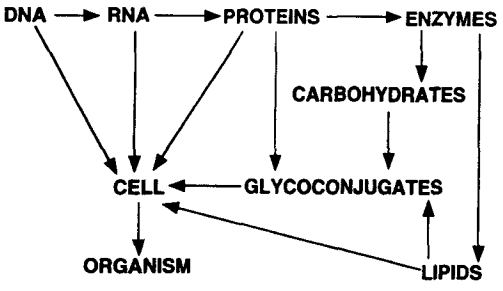
THIS CHAPTER PROVIDES HISTORICAL BACKGROUND to the emergence of the field of glycobiology, as well as an overview of this book. General terms and definitions found throughout the volume are also considered. The common monosaccharide units of glycoconjugates are mentioned and a uniform symbol nomenclature used for structural depictions throughout the book is presented. The general oligosaccharide classes to be discussed in the book are mentioned, and an overview of the general pathways for their biosynthesis is provided. Topological issues relevant to biosynthesis and function are also considered.

WHAT IS GLYCOBIOLOGY? (1–4)

The central paradigm of modern molecular biology is that biological information flows from DNA to RNA to protein. The power of this concept lies not only in its template-driven precision, but also in the ability to manipulate any one class of molecules based on knowledge of another, and in the patterns of sequence homology and relatedness that predict function and reveal evolutionary relationships. With the upcoming completion of the genomic sequences of humans and several other commonly studied model organisms, even more spectacular gains in the understanding of biological systems are anticipated. However, there is often a tendency to assume the following extension of the central paradigm:

DNA → RNA → PROTEIN → CELL → ORGANISM

In actual fact, creating a cell requires two other major classes of molecules: lipids and carbohydrates. These molecules can serve as intermediates in generating energy, as signaling molecules, or as structural components. The structural roles of carbohydrates become particularly important in constructing complex multicellular organs and organisms, which requires interactions of cells with one another and with the surrounding matrix. Indeed, all cells and many macromolecules in nature carry a dense and complex array of covalently attached sugar chains (called oligosaccharides or glycans). In some instances, these glycans can also be free-standing entities. Since most glycans are on the outer surface of cellular and secreted macromolecules, they are in a position to modulate or mediate a wide variety of events in cell-cell and cell-matrix interactions crucial to the development and function of a complex multicellular organism. They are also in a position to mediate interactions between organisms (e.g., between host and parasite). In addition, simple, highly dynamic protein-bound glycans are abundant in the nucleus and cytoplasm, where they appear to serve as regulatory switches. An extended paradigm of molecular biology can thus be rendered as follows:



In the first part of this century, the chemistry, biochemistry, and biology of carbohydrates were very prominent matters of interest. However, during the initial phase of the modern revolution in molecular biology, studies of glycans lagged far behind those of other major classes of molecules. This was in large part due to their inherent structural complexity, the difficulty in easily determining their sequence, and the fact that their biosynthesis could not be directly predicted from the DNA template. The development of a variety of new technologies for exploring the structures of these sugar chains has opened up a new frontier of molecular biology which has been called glycobiology. This word was first coined in 1988 by Rademacher, Parekh, and Dwek to recognize the coming together of the traditional disciplines of carbohydrate chemistry and biochemistry with modern understanding of the cellular and molecular biology of glycans. The term glycobiology has

gained wide acceptance, with a major biomedical journal, a growing scientific society, and a Gordon Research Conference now bearing this name.

Defined in the broadest sense, glycobiology is then the study of the structure, biosynthesis, and biology of saccharides (sugar chains or glycans) that are widely distributed in nature. It is one of the more rapidly growing fields in the biomedical sciences, with relevance to basic research, biomedicine, and biotechnology. Indeed, several biotechnology, pharmaceutical, and laboratory supply companies have invested heavily in the area. The field ranges from the chemistry of carbohydrates and the enzymology of glycan-modifying proteins to the functions of glycans in complex biological systems, and their manipulation by a variety of techniques. Research in glycobiology requires a foundation not only in the nomenclature, biosynthesis, structure, chemical synthesis, and functions of complex glycans, but also in the general disciplines of molecular genetics, cellular biology, physiology, and protein chemistry. This volume provides an overview of the field of glycobiology, with a particular emphasis on the glycans of higher animal systems, about which the greatest amount is currently known. It is assumed that the reader has a basic background in graduate-level chemistry, biochemistry, and cell biology.

MONOSACCHARIDES ARE THE BASIC STRUCTURAL UNITS OF GLYCANS (5)

Carbohydrates are defined as polyhydroxyaldehydes or polyhydroxyketones, or larger compounds that can be hydrolyzed into such units (for examples, see below and for more details, see Chapter 2). A monosaccharide is a carbohydrate that cannot be hydrolyzed into a simpler unit. It has a potential carbonyl group at the end of the carbon chain (an aldehyde group) or at an inner carbon (a ketone group). These two types of monosaccharides are therefore named aldoses and ketoses. Free monosaccharides can exist in open chain or ring forms (Figure 1.1).

Ring forms of the monosaccharides are the rule in oligosaccharides, which are branched or linear chains of monosaccharides attached to one another via glycosidic linkages (the term polysaccharide is typically reserved for large glycans that are composed of repeating oligosaccharide motifs). The ring form of a monosaccharide generates a chiral (anomeric) center (at C-1 for aldo sugars or at C-2 for keto sugars) (for details, see Chapter 2). A glycosidic linkage involves the attachment of a monosaccharide to another residue, typically via the hydroxyl group of this anomeric center, which can be α linkages or β link-

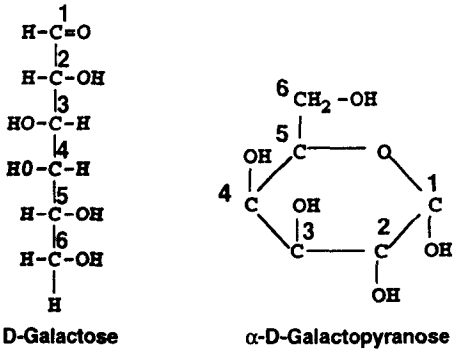


FIGURE 1.1. Open chain and ring forms of galactose. Changes in the orientation of hydroxyl groups around specific carbon atoms result in new molecules that have a distinct biology and biochemistry (e.g., glucose is the 4-epimer of galactose).

ages depending on the relationship of the oxygen to the anomeric carbon (see Chapter 2). It is important to realize that these two types of linkages confer very different structural properties and biological functions upon sequences that are otherwise identical in composition. A glycoconjugate is a compound in which one or more monosaccharide or oligosaccharide units (the glycone) are covalently linked to a noncarbohydrate moiety (the aglycone). An oligosaccharide that is unattached to an aglycone usually retains the potential reducing power of the aldehyde or ketone in its terminal monosaccharide component. This end of a sugar chain is therefore often called the reducing terminus or reducing end (this term tends to be used even when the sugar chain is attached to an aglycone and thus has actually lost its reducing power). Correspondingly, the outer end of the chain tends to be called the nonreducing end (note the analogy to the 5′ and 3′ ends of nucleotide chains or the amino and carboxyl termini of polypeptides).

GLYCANS CAN CONSTITUTE A MAJOR PORTION OF A GLYCOCONJUGATE (2)

In naturally occurring glycoconjugates, the portion of the molecule comprising the glycans can vary greatly, from being very minor in amount to being the dominant component. Indeed, it is striking that sugar chains make up a substantial portion of the mass of most glycoconjugates (for a typical example, see Figure 1.2). For this reason, the surfaces of most types of cells (which are heavily decorated with different kinds of glycoconjugates) are effectively covered with a dense coating of sugars, giving rise to the so-called glycocalyx. This cell surface structure was first observed by electron microscopists many years ago as an anionic layer external to the plasmalemma, which could be decorated with polycationic reagents like cationized ferritin (for an example, see Figure 1.3).

MONOSACCHARIDES GENERATE MORE LINKAGE VARIATION THAN AMINO ACIDS OR NUCLEOTIDES (1,6)

Nucleotides and proteins are linear polymers that can each have only one basic type of linkage. In contrast, each monosaccharide can theoretically generate an α or a β linkage to any one of several positions on another monosaccharide in a chain or to another type of molecule. Thus, it has been pointed out that although three nucleotide bases or amino

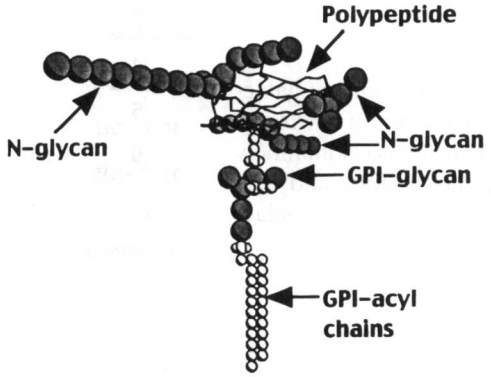


FIGURE 1.2. Schematic representation of the Thy-1 glycoprotein including the three N-glycans and a glycophospholipid (GPI-glycan) anchor whose acyl chains would normally be embedded in the membrane bilayer. Note that the polypeptide represents only a relatively small portion of the total mass of the protein. (Modified, with permission, from [2] Rademacher et al. 1988 [© Annual Reviews].)

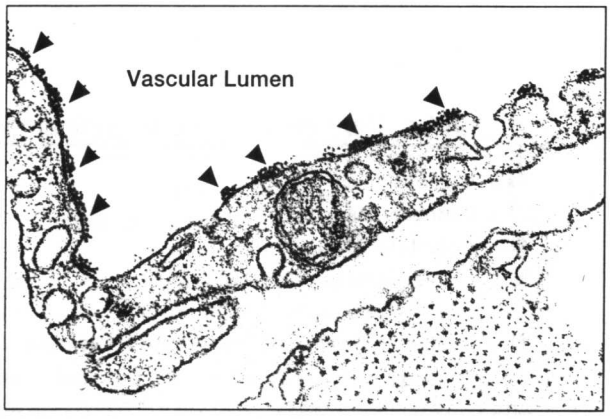


FIGURE 1.3. Electron micrograph of endothelial cells from a blood capillary in the diaphragm muscle of a rat, showing the lumenal plasmalemma of the cells (facing the blood) decorated with particles of (pI 8.4) cationized ferritin (see arrowheads). These particles are binding to acidic residues (sialic-acid-containing glycans and sulfated glycosaminoglycans) contained in the cell surface "glycocalyx." Note that the particles are several layers deep, indicating the remarkable thickness of this layer of glycoconjugates. (Courtesy of George E. Palade.)

acids can only generate six variations, three hexoses could produce (depending on which factors are considered) anywhere from 1,056 to 27,648 unique trisaccharides. As the number of units in the polymer increases, this difference in complexity becomes even greater. For example, a hexasaccharide with six hexoses could have more than 1 trillion possible combinations. Thus, an almost unimaginable number of possible saccharide units could be theoretically present in biological systems. Fortunately, for the student of glycobiology, naturally occurring biological macromolecules contain relatively few of the possible monosaccharide units in a limited number of combinations.

COMMON MONOSACCHARIDE UNITS OF ANIMAL GLYCOCONJUGATES (5,7)

The common monosaccharides found in higher animal oligosaccharides are listed below, along with their standard abbreviations (for more details regarding their structures, see Chapter 2).

- *Sialic Acids:* Family of nine-carbon acidic sugars (generic abbreviation is Sia), of which the most common is N-acetyl neuraminic acid (Neu5Ac, also sometimes called NeuNAc, NeuAc, or NANA) (for more details, see Chapter 15).
- *Hexoses:* Six-carbon neutral sugars, including glucose (Glc), galactose (Gal), mannose (Man).
- *Hexosamines:* Hexose with an amino group at the 2-position, which can be either free or, more commonly, N-acetylated: N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc).
- *Deoxyhexoses:* Six-carbon neutral sugar without the hydroxyl group at the 6-position, fucose (Fuc).
- *Pentoses:* Five-carbon sugar, xylose (Xyl).
- *Uronic Acids*: Hexose with a negatively charged carboxylate at the 6-position, glucuronic acid (GlcA) and iduronic acid (IdA).