PROGRESS IN

MOLECULAR BIOLOGY AND TRANSLATIONAL SCIENCE

VOLUME 93

GLYCOSAMINOGLYCANS IN DEVELOPMENT, HEALTH AND DISEASE

EDITED BY
LIJUAN ZHANG

PROGRESS IN

Molecular Biology and Translational Science

Volume 93

This page intentionally left blank

PROGRESS IN

Molecular Biology and Translational Science

Glycosaminoglycans in Development, Health and Disease

edited by

Lijuan Zhang

Department of Pathology and Immunology
Washington University Medical School
St. Louis, Missouri, USA

Volume 93

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier

Academic Press is an imprint of Elsevier
32 Jamestown Road, London, NW1 7BY, UK
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

This book is printed on acid-free paper. ♾

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

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

ISBN: 978-0-12-381282-7
ISSN: 1877-1173

For information on all Academic Press publications
visit our website at elsevierdirect.com

Printed and Bound in the USA

10 11 12 13 10 9 8 7 6 5 4 3 2 1

Contents

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

Rafael S. Aquino, Division of Respiratory Diseases, Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA (373)

Jacques U. Baenziger, Department of Pathology, Washington University School of Medicine, St. Louis, Missouri, USA (289)

Lubor Borsig, Institute of Physiology, and Zürich Center for Integrative Human Physiology, University of Zürich, Switzerland (335)

Ariane de Agostini, Department of Gynaecology and Obstetrics, Laboratory of Reproductive Biology, and Department of Genetic and Laboratory Medicine, Service of Clinical Pathology, Geneva University Hospitals and University of Geneva, Geneva, Switzerland (153)

Jeffrey D. Esko, Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla, California, USA (213)

Erin M. Foley, Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, University of California, San Diego, La Jolla, California, USA (213)

Mark M. Fuster, Department of Medicine, Division of Pulmonary and Critical Care, University of California San Diego, and VA San Diego Healthcare System, La Jolla, California, USA (179)

Hiroko Habuchi, Research Complex for the Medicine Frontiers, Aichi Medical University, Nagakute, Aichi, Japan (79)

Masaru Inatani, Sanford Children's Health Research Center, Sanford-Burnham Medical Research Institute, La Jolla, California, USA, and Department of Ophthalmology and Visual Science, Kumamoto University Graduate School of Medical Sciences, Kumamoto, Japan (133)

Fumitoshi Irie, Sanford Children's Health Research Center, Sanford-Burnham Medical Research Institute, La Jolla, California, USA (133)

Tomomi Izumikawa, Department of Biochemistry, Kobe Pharmaceutical University, Kobe, Japan (19)

Andreas R. Janecke, Department of Pediatrics II, Innsbruck Medical University, Anichstrasse 35, Innsbruck, Austria (289)

Koji Kimata, Research Complex for the Medicine Frontiers, Aichi Medical University, Nagakute, Aichi, Japan (79)

Hiroshi Kitagawa, Department of Biochemistry, Kobe Pharmaceutical University, Kobe, Japan (19)

Lena Kjellén, Department of Medical Biochemistry and Microbiology, The Biomedical Center, Uppsala University, Uppsala, Sweden, and Department of Biomedicine, University of Bergen, Bergen, Norway (35)

Michael Klüppel, Human Molecular Genetics Program, Children's Memorial Research Center, and Department of Pediatrics; and Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Feinberg School of Medicine, Chicago, Illinois, USA (113)

Takashi Kobayashi, Department of Medicine, Dartmouth Medical School, Hanover, New Hampshire, USA (153)

Eui Seung Lee, Division of Respiratory Diseases, Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA (373)

Jin-Ping Li, Department of Medical Biochemistry and Microbiology, University of Uppsala, Uppsala, Sweden (59, 309)

Hong Lu, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (395, 423, 443, 473)

Thomas Müller, Department of Pediatrics II, Innsbruck Medical University, Anichstrasse 35, Innsbruck, Austria (289)

Yoshihiro Matsumoto, Sanford Children's Health Research Center, Sanford-Burnham Medical Research Institute, La Jolla, California, USA, and Department of Orthopaedic Surgery, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan (133)

Junko Ogawa, Sanford Children's Health Research Center, Sanford-Burnham Medical Research Institute, La Jolla, California, USA (133)

Jing Pan, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (395, 423, 443, 473)

Pyong Woo Park, Division of Respiratory Diseases, Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA (373)

Yi Qian, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (395, 423, 443,473)

Maria Ringvall, Department of Medical Biochemistry and Microbiology, The Biomedical Center, Uppsala University, Uppsala, Sweden (35)

Barbara P. Schick, Department of Medicine, Cardeza Foundation for Hematologic Research, Thomas Jefferson University, Philadelphia, Pennsylvania, USA (235)

Fei F. Shih, Department of Pediatrics, Washington University Medical School, St. Louis, Missouri, USA (443)

Nicholas W. Shworak, Department of Medicine, and Department of Pharmacology and Toxicology, Dartmouth Medical School, Hanover, New Hampshire, USA (153)

Nicole C. Smits, Department of Medicine, Dartmouth Medical School, Hanover, New Hampshire, USA (153)

Daniel R. Studelska, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (443, 473)

Douglas M. Tollefsen, Hematology Division, Washington University Medical School, St. Louis, Missouri, USA (351)

Lianchun Wang, Department of Biochemistry and Molecular Biology, Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia, USA (179)

Peter Weiser, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (395, 423, 443, 473)

Yu Yamaguchi, Sanford Children's Health Research Center, Sanford-Burnham Medical Research Institute, La Jolla, California, USA (133)

Lijuan Zhang, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (1, 289, 395, 423, 443, 473)

Xiao Zhang, Department of Public Health and Caring Sciences, Molecular Geriatrics, University of Uppsala, Uppsala, Sweden (309)

Xiaodong Zhou, Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA (395, 423, 443, 473)

Preface

All animal cells produce two major types of sulfated linear polysaccharide glycosaminoglycans (GAGs): heparan sulfate and chondroitin sulfate. Due to their structural diversity, GAGs have been claimed to be the most information-dense biopolymers found in nature. GAGs bind to numerous protein ligands and receptors involved in diverse biological processes, including cell division, growth control, signal transduction, cell adhesion, hemostasis, and lipid metabolism (Chapters 1 and 2). Emerging from its roots in classical chemistry and biochemistry, the progress in the field of GAGs has greatly accelerated during the past 20 years through genetic analysis. Transgenic and knockout animal data (Chapters 2–13) provide compelling evidence that this structural diversity is a component of a sugar/sulfation-GAG code. The sugar/sulfation-GAG code imparts unique and specific biological functions during development, health, and disease (Chapters 2–16), which makes GAG an essential component of modern molecular biology and human physiology.

Heparin, the mostly sulfated heparan sulfate made by mast cells, has the highest anticoagulation activity by inhibiting thrombin generation and thrombin activities among different GAGs and has been used as an anticoagulant drug for over 70 years. Crude heparin isolated from animal tissues consists of $\sim$50% heparin and $\sim$50% less sulfated GAGs including heparan sulfate, chondroitin sulfate, and dermatan sulfate. Heparin manufacturers purchase crude heparin from small vendors and produce heparin from crude heparin by removing the low-sulfated GAGs using sophisticated techniques. In 2007 and 2008, contaminated heparin was associated with hundreds of anaphylactic reactions and at least 149 anaphylactic reaction-associated deaths in the United States. It turns out that the heparin contaminants are chemically sulfated/modified low-sulfated GAGs.

The contaminated heparin associated anaphylactic reactions are a result of up-regulated immune system by oversulfated GAGs through contact system activation. The contact system was first discovered as an *in vitro* thrombin generation system where artificial surfaces induce thrombin generation and clotting. Injury-associated thrombin generation in animals instantly leads to immune system up-regulation. However, the *in vivo* role of contact system activation-generated thrombin and immune system up-regulation has been overlooked during the past 50 years. Chapters 17–20 demonstrate that contact

system activation-generated thrombin and immune system up-regulation induced by abnormal GAG/protein aggregates are the outcomes of different autoimmune diseases, including Lupus, rheumatoid arthritis, psoriasis, heparin-induced thrombocytopenia/thrombosis (HIT), and different kinds of human cancers.

The Editor is indebted to many others who made this book possible. Special thanks are due to Dr. Michael Conn, the editor of *Progress in Molecular Biology and Translational Science* book series, who projected this book volume; and to Lisa Tickner, Delsy Retchagar, Malathi Samayan, and Sunita Sundararajan at Elsevier for keeping us on track and converting our efforts into a product. Last but not least, we acknowledge the support and hard work of our families and lab members in producing each chapter in the book. It now remains for the reader to decide whether we have achieved our goals in compiling the GAG book.

LIJUAN ZHANG
Missouri, USA

Glycosaminoglycan (GAG) Biosynthesis and GAG-Binding Proteins

Lijuan Zhang

Department of Pathology and Immunology, Washington University Medical School, St. Louis, Missouri, USA

Two major types of glycosaminoglycan (GAG) polysaccharides, heparan sulfate and chondroitin sulfate, are polymerized and modified by enzymes that are encoded by more than 40 genes in animal cells. Because of the expression repertoire of the GAG assembly and modification enzymes, each heparan sulfate and chondroitin sulfate chain has a sulfation pattern, chain length, and fine structure that is potentially unique to each animal cell. GAGs interact with hundreds of proteins. Such interactions protect growth factors, chemokines, and cytokines against proteolysis. GAGs catalyze protease (such as thrombin) inhibition by serpins. GAGs regulate multiple signaling pathways including, but not limited to, fibroblast growth factor (FGF)/FGFR, hepatocyte growth factor (HGF)/c-Met, glial cell line-derived neurotrophic factor (GDNF)/c-Ret/GFRα1, vascular endothelial growth factor (VEGF)/VEGFR, platelet derived growth factor (PDGF)/PDGFR, BAFF/TACI, Indian hedgehog, Wnt, and BMP signaling pathways,where genetic studies have revealed an absolute requirement for GAGs in these pathways. Most importantly, protein/GAG aggregates induce thrombin generation and immune system upregulation by activating the contact system. Abnormal protein/GAG aggregates are associated with a variety of devastating human diseases including, but not limited to, Alzheimer's, diabetes, prion or transmissible spongiform encephalopathies, Lupus, heparin-induced thrombocytopenia/thrombosis, and different kinds of cancers. Therefore, GAGs are essential components of modern molecular biology and human physiology. Understanding GAG structure and function at

Progress in Molecular Biology and Translational Science, Vol. 93
DOI: 10.1016/S1877-1173(10)93001-9

1877-1173/10 $35.00

molecular level with regard to development and health represents a unique opportunity in combating different kinds of human diseases.

Abbreviations: CS, chondroitin sulfate; GAG, glycosaminoglycan; HS, heparan sulfate

I. Glycosaminoglycans (GAGs)

GAGs are linear polysaccharides that are made by all animal cells. Two major types of GAGs are heparan sulfate (HS) and chondroitin sulfate (CS). HS and CS comprise repeating hexosamine-uronic acid disaccharides that are sulfated to varying degrees[1,2] (Fig. 1). Heparin is the most highly sulfated HS made by mast cells. DS is one type of CS containing IdoA residues and is made by many types of animal cells.

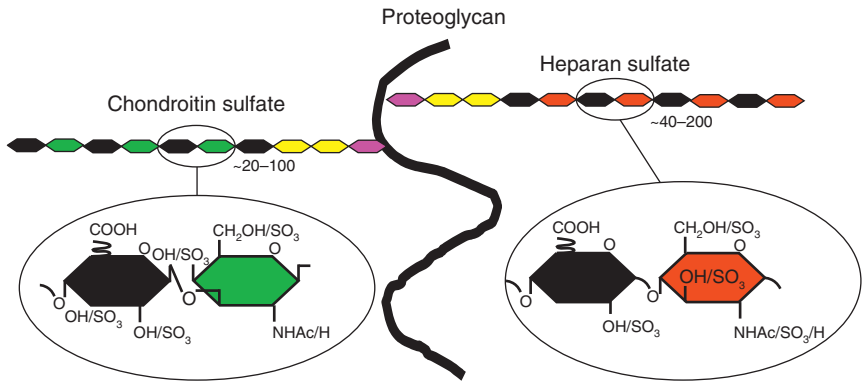

FIG. 1. Diagrammatic representation of GAG assembly on a proteoglycan core protein. Both HS and CS are attached to specific serine residues of proteoglycan core protein by the linkage tetrasaccharide GlcA (black)-Gal (yellow)-Gal (yellow)-Xyl (pink). Biosynthesis starts with the transfer of xylose from UDP-xylose to a serine residue of a core protein catalyzed by two xylosyltransferases. The linkage region is then synthesized by the sequential addition of two galactose residues (by galactosyltransferase I and II), and glucuronic acid (by glucuronosyltransferase I) from the corresponding UDP-sugars. After completion of the linkage tetrasaccharides, the addition of first HexNac residue occurs. Addition of GalNAc from UDP-GalNAc by *N*-acetylgalactosaminyl transferase I to the nonreducing terminal GlcA commits the intermediate to CS sysnthesis, which occurs subsequently through alternating addition of GlcA and GalNAc (green) by chondroitin synthase. If GlcNAc is added to the linkage tetrasaccharide instead by *N*-acetylglucosaminyltransferase I, HS synthesis occurs. Alternating GlcA and GlcNAc (red) residues are then added by HS copolymerases (EXT-1 and EXT-2) from their corresponding UDP-sugars. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this chapter.)

HS and CS are abundantly produced ($\sim 10^5$–10^6 copies on the cell surface, $\sim$ mg/ml concentrations in the extracellular matrix) by animal cells.[3] This synthesis occurs in the *Golgi* in the form of HS, CS, or HS/CS hybrid proteoglycans. In proteoglycans, anywhere between one (e.g., decorin) to over one hundred GAG chains (e.g., aggrecan) can be assembled on a proteoglycan core protein (Fig. 1). The number of repeating disaccharides of HS and CS varies depending on the source of cells or tissues used for GAG isolation. GAGs adopt an extended helical coil structure with a length ranging from 40 to 160 nm. Such abundance and size implies that GAGs are a dominant feature of the cell surface and are an important feature of the extracellular matrix.

II. Proteoglycan

All mammalian cells produce proteoglycans and can secrete them into the extracellular matrix, insert them into the plasma membrane, or store them in secretory granules. Over 50 proteoglycan cDNAs have been cloned (Table I). Some of the proteoglycan gene products were known as important functional proteins long before they were known to be proteoglycans. It is typical for a cell to express multiple types of HS and CS proteoglycans. For example, at least 11 different proteoglycans are expressed by human lung fibroblasts[4–10] and at least 23 types are expressed in the nervous system.[11] Most core proteins not only serve as GAG carriers but also have their own functional protein domains. In some instances, they may also contain GAG-binding domains. The core proteins usually determine the number of GAG chains, the type (HS or CS), and the ultimate destination (apical, luminal, intra- or extracellular) of the finished proteoglycan. GAG chains carried by proteoglycans are sometimes cleaved before they become biologically active. Extracellular heparanases (heparan sulfate endoglucuronidases),[12] sulfatases,[13,14] and free GAG chains have been extensively reported.[15–18]

The structural variation of proteoglycans in different cells or tissues is due to a number of factors. First, over 50 core proteins have been identified, and these can be substituted with CS or both HS and CS. Another source of variability lies in the stoichiometry of GAG chain substitution. For example, syndecan-1 has five GAG attachment sites, but not all of the sites are used equally. Thus, a preparation of syndecan-1 represents a diverse population of syndecan-1 molecules. Other proteoglycans, such as thrombomodulin, can be a "part time proteoglycan" that is, they may exist with or without a GAG chain or with only a truncated GAG chain. These characteristics, typical of all proteoglycans, create diversity that may facilitate the formation of binding sites of variable density and affinity for different ligands in a cell- and tissue-specific manner, beyond its GAG-binding specificity.

TABLE I
CLONED PROTEOGLYCANS

HS proteoglycans	CS proteoglycans		
Syndecan	Syndecan-1 (TM)	Lectican	Aggrecan
	Syndecan-2 (TM)		Versican/PG-M
	Syndecan-3 (TM)		Neurocan
	Syndecan-4 (TM)		Brevican (soluble or GPI-linked)
Glypican	Glypican-1 (GPI-Linked)	IPM	IPM150 (interphotoreceptor matrix PG, TM)
	Glypican-2 (GPI-Linked)		IPM 200 (TM)[92]
	Glypican-3 (GPI-Linked)		SPACRCAN (TM)[93]
	Glypican-4 (GPI-Linked)	SLRP	Decorin (small leucine-rich proteoglycans)
	Glypican-5 (GPI-Linked)		Biglycan
	Glypican-6 (GPI-Linked)		PG-lb
Testican	Testican-1	Others	CD44 (CS splicing form, TM)
	Testican-2		Thrombomodulin (TM)
	Testican-3		Invariant chain
Others	Betaglycan (TM)		APP (amyloid precursor protein)
	Tyrosine kinase receptor (TM)		APLP2
	PRPg (proline rich PG)		Lepecan[94]
	Perlecan		Bicunin
	Agrin		Chromogranin A
	Collagen XVIII		α2(IX) collagen
	CD44 (HS splicing form, TM)		Tanasin-C
	Serglycin		Laminin α4[95]
			CSF (colony stimulating factor)
			Phosphocan (3 splicing forms, soluble or TM)
			Claustrin
			NG2 (TM)
			Neuroglycan-C (TM)
			MCSP (melanoma-associated CSPG)[96]
			PRG4 (megakaryocyte stimulating factor)[97]
			Endoglycan[98]
			Endocan[99]
			Bamacan[100]

Accession numbers of nonreferenced genes can be found in Refs. 11,22.
TM, transmembrane protein.

III. GAG Biosynthesis

HS and CS are elaborated after the synthesis of the GAG-protein linkage region, GlcA-Gal-Gal-Xyl, which is attached to specific Ser residues on the proteoglycan core protein (Fig. 1). The synthesis of this region is initiated by the addition of a Xyl to Ser followed by the addition of two Gal residues and is completed by the addition of GlcA. The pathways of HS and CS synthesis diverge after formation of the linkage region. The addition of GlcNAc to the linkage region commits the intermediate proteoglycan molecule to the assembly of HS. Similarly, the addition of a GalNAc commits it to CS synthesis.

CS assembly on the linkage region represents a default pathway.[19–21] HS assembly requires amino acid determinants proximal to the linkage tetrasaccharide. The enzyme necessary for initiating HS synthesis prefers sites containing multiple acidic residues, one or more hydrophobic residues, and repetitive Ser-Gly units.[22] Later on, it was discovered that certain structural determinants exist in glypican-1 that prevent CS formation. These may act by blocking access of the initiating β-GalNAc transferase. The proportion of HS to CS carried on a HS proteoglycan is cell type- (or tissue-) dependent. For example, the only "absolute" HS proteoglycan, glypican-1, carries 90% HS and 10% CS when expressed in COS cells and 80% HS and 20% CS when expressed in CHO cells.[23] Very low percentages of HS have been detected on certain CS proteoglycans, for example biglycan, and aggrecan.[24,25] The enzymes responsible for making essential GAG building blocks and linkage region are summarized in Table II.

The fine structure of HS is determined by the step-wise action of multiple enzymes and enzyme isoforms (Table III). The two subunits of the HS copolymerase are each encoded by genes that are known to be tumor suppressor genes (EXT-1 and EXT-2).[26,27] EXT genes in humans are associated with multiple hereditary exostoses, a syndrome characterized by neonatal growth plate tumors.

As shown in Table III, several HS modification enzymes are present in multiple isoforms. The expression of each isoform, including those of sulfotransferases NST, 6-OST and 3-OST, is tissue specific.[1] This is thought to be important for generating HS with unique saccharide sequences. In this regard, it is instructive to point out that 3-OST-1 modified HS binds antithrombin and has anticoagulant activity,[1,28] whereas 3-OST-3 modified HS binds to herpes envelope protein gD and serves as an entry receptor for the herpes simplex virus-1.[29,30] In contrast, 3-OST-5 modified HS binds to both antithrombin and herpes envelope protein gD and has both anticoagulant and entry receptor properties.[31]

In summary, at least 20 genes are involved in HS polymerization and modification. Combinatorial expression of these genes means that the structure of HS can be altered to modulate a wide variety of tissue-specific functions.

TABLE II

Cloned Enzymes for Making GAG Building Blocks and for GAG Initiation

DTDST sulfate transporter 1	UDP-GlcA transporter	UDP-glucose dehydrogenase	Galactosyltransferase I
PAPS synthetase-1	UDP-GlcNAc transporter	UDP-GlcA decarboxylase	Galactosyltransferase II
PAPS synthetase-2	UDP-GalNAc transporter	Xylosyltransferase I	Glucuronosyltransferase I (knockout mice; chapter "Mice Deficient in Glucuronyltransferase-I")
PAPS transporter	UDP-Gal transporter	Xylosyltransferase II	Glucuronosyltransferase II

Accession numbers of corresponding human genes can be found in Ref. 1. Three glucuronosyltransferases for linkage tetrasaccharide biosynthesis in *Drosophila* have been identified.[101]

TABLE III
CLONED ENZYMES FOR HS FORMATION

αGlcNAcTIA (EXTL1)	NDST-1	Epimerase	6-OST-3	3-OST-3B
αGlcNAcTIB (EXTL2)	NDST-2	2-OST	3-OST-1	3-OST-4
Copolymerase EXT1	NDST-3	6-OST-1	3-OST-2	3-OST-5
Copolymerase EXT2	NDST-4	6-OST-2	3-OST-3A	3-OST-6

Accession numbers of corresponding human genes can be found in Ref. 1.

EXT1 conditional knock out: chapter "Roles of Heparan Sulfate in Mammalian Brain Development: Current Views Based on the Findings from Ext1 Conditional Knockout Studies."

NDST-1 knockout and conditional knockout mice: chapters "Mice Deficient in Heparan Sulfate *N*-Deacetylase/*N*-Sulfotransferase 1, Endothelial Heparan Sulfate in Angiogenesis, and Hepatic Heparan Sulfate Proteoglycans and Endocytic Clearance of Triglyceride-Rich Lipoproteins."

Epimerase knockout mice: chapter "Mice Deficient in Glucuronyl C5-Epimerase—An Enzyme Converting Glucuronic Acid to Iduronic Acid in Heparan Sulfate/Heparin Biosynthesis."

6-OST-1 knockout mice: chapter "Mice Deficient in Heparan Sulfate 6-O-Sulfotransferase-1."

3-OST-1 knockout mice: chapter "Anticoagulant Heparan Sulfate: To Clot or Not."

TABLE IV
TYPES OF CS DISACCHARIDES

Major disaccharides	Other disaccharides found	
CS-A: GlcA-GalNAc4S	GlcA-GalNAc	GlcA2S-GalNAc
CS-B: IdoA-GalNAc4S	IdoA2S-GalNAc4S	GlcA3S-GalNAc
CS-C: GlcA-GalNAc6S	IdoAGalNAc4S6S	GlcA3S-GalNAc4S
CS-D: GlcA2S-GalNAc6S	IdoA2SGalNAc4S6S	GlcA3S-GalNAc4S6S
CS-E: GlcA-GalNAc4S6S	IdoA2S-GalNAc	GlcA3S-GalNAc6S

CS has been classified as CS-A, CS-B (dermatan sulfate), CS-C, CS-D, and CS-E (Table IV) according to the major constituent of the repeating disaccharides. However, all CSs are hybrid structures that contain more than two types of disaccharides. CS-B, or dermatan sulfate, is distinguished because it contains IdoA residues.

The fine structure of CS also depends on the temporal and tissue-specific expression of a variety of modifying enzymes and enzyme isoforms (Table V).

Some sulfation appears to be specific to CS or HS. GlcA 3-O-sulfation has been detected only in CS.[32,33] Conversely, N-sulfation occurs only in HS but not in CS. Because the HS epimerase requires N-sulfated residues for its activity and CS is devoid of N-sulfation, epimerization of GlcA to IdoA residues of CS are catalyzed by two distinct CS epimerases.[34]

TABLE V
ENZYMES FOR CS FORMATION

GalNAc I	GlcA 3-OST	4OST3[102]	6OST3
GalNAc II	Epimerase I and II	IdoA-GalNAc 4-OST	6OST4
Chondroitin synthase	4-OST1	6OST1	GalNAc4S 6-OST
GlcA/IdoA 2-OST	4-OST2	6OST2	IdoA-GalNAc-IdoA 6OST

Accession numbers of non-referenced genes can be found in Ref. 2.

4-OST1 knockout mice: chapter "Roles of Heparan Sulfate in Mammalian Brain Development: Current Views Based on the Findings from Ext1 Conditional Knockout Studies."

IdoA-GalNAc 4-OST deficiency in human: chapters "Congenital Disorders of Glycosylation with Emphasis on Loss of Dermatan-4-Sulfotransferase."

In contrast to CS, which tend to have long tracts of fully modified disaccharides, the modification reactions in HS biosynthesis occur in clusters along the chain, with regions devoid of sulfation separating the modified tracts. This arrangement gives rise to segments referred to as *N*-acetylated (NA), *N*-sulfated (NS), and mixed domains (NA/NS). In general, the sulfation reactions fail to go to completion, resulting in tremendous chemical heterogeneity within the modified regions in CS and HS.

In summary, HS and CS are characterized by a linear chain of 20–400 disaccharide units. The disaccharide repeat unit in HS can be modified by N- and O-sulfation (6-O- and 3-O-sulfation of the glucosamine and 2-O-sulfation of the uronic acid) and by epimerization of the glucuronic acid to iduronic acid. The disaccharide repeat unit in CS can be modified by 4-O- and 6-O-sulfation of the galactosamine and 2-O- and 3-O-sulfation of the uronic acid and by epimerization of glucuronic acid to iduronic acid. Together, the five different modifications for disaccharides in HS and CS give rise to $2^5 = 32$ combinations. With 23 disaccharides found in CS and 24 found in HS,[1] a HS or CS hexasaccharide could have several thousand possible sequences, thereby making HS and CS not only the most acidic but also the most information-dense biopolymers found in nature. Understanding how and in what order the cells assemble specific HS and CS sequences is one of the important areas in research on GAG.

IV. GAG-Binding Proteins

GAGs participate in a variety of physiological processes such as binding, activation, or immobilization of various protein ligands, such as growth factors, cytokines, chemokines, extracellular matrix components, proteases, protease inhibitors, and lipoprotein lipase.[1,3,35–42] These interactions depend, to a large

extent, on the composition and fine structure of the GAG chains, which in turn depend on the substrate specificity of the various biosynthetic enzymes and regulatory factors.

Heparin interacts with 23% of the plasma proteins.[43] More than 200 GAG-binding proteins have been described in literature (Fig. 2). To a large extent, these studies have focused on protein interactions with heparin. This bias may reflect the commercial availability of heparin, which is frequently used for fractionation studies and heparin-Sepharose affinity chromatography. The binding of protein ligands to heparin is thought to mimic the physiological interaction of proteins with the HS that is present on cell surfaces and in the extracellular matrix.

Most animal tissues and organs contain more CS than HS.[44] CS acts as a biological activator in a number of instances, in apparent independence of its co-existence with HS. Examples include modulation of axon growth,[45,46] wound healing,[47] NFκB transcriptional activation of endothelial cells,[48] circumscribed release of short-lived kinin hormones from precursor depots to regulate local blood pressure and inflammatory responses,[49] heparin cofactor II inhibition of localized coagulation,[50] and low density lipoprotein (LDL) binding that affects intramural retention of atherogenic lipoproteins.[51] A variety of bacterial-, viral-, and parasite-proteins also bind to GAGs (chapter "Diverse Functions of Glycosaminoglycans in Infectious Diseases").

GAGs regulate many growth factor signaling pathways[52] including, but not limited to, fibroblast growth factor (FGF)/FGFR,[53] hepatocyte growth factor (HGF)/c-Met,[54–57] glial cell line-derived neurotrophic factor (GDNF)/c-Ret/GFRα1,[58,59] vascular endothelial growth factor (VEGF)/VEGFR,[60,61] platelet derived growth factor (PDGF)/PDGFR,[62] BAFF/TACI,[63] Indian hedgehog (Ihh), Wnt, and BMP signaling pathways, where genetic studies revealed an absolute requirement for GAGs.[64,65]

Heparin has been used for preventing and treating thromboembolic disorders for over 70 years as it inhibits thrombin generation and thrombin activities. Thromboembolic disorders are the leading cause of disabilities and deaths in a variety of unrelated human diseases, such as coronary heart disease,[66] cancer,[67] diabetes,[68] kidney failure,[69] autoimmune diseases,[70] and heparin-induced thrombocytopenia and thrombosis (HITT).[71] Thrombin is the only known enzyme that causes thrombus formation. Thrombin also plays multiple roles in development, tissue repair, inflammation, and hemostasis.[72] Conditional loss of prothrombin leads to the rapid death of adult mice,[73] suggesting that thrombin is a key survival factor, which is continuously generated in blood circulation.

In 2007 and 2008, contaminated heparin was associated with hundreds of anaphylactic reactions and at least 149 anaphylactic reaction-associated deaths in the United States. Heparin is contaminated with chemically sulfated or

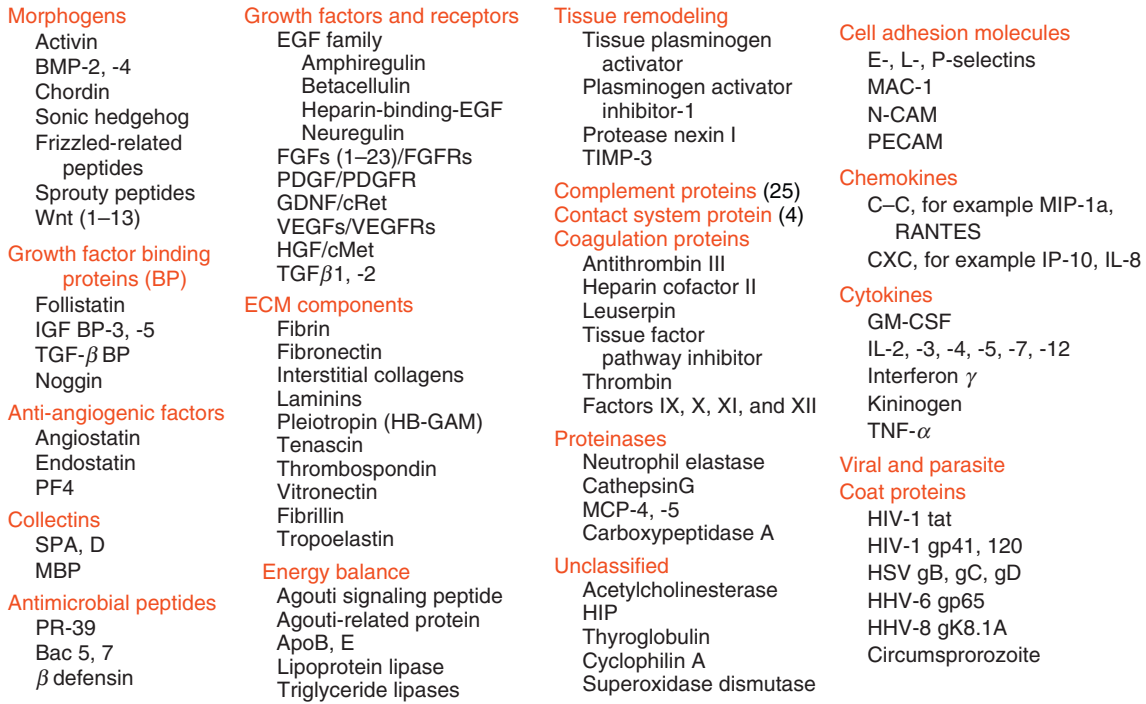

FIG. 2. GAG-binding proteins. The GAG-binding proteins reported in literature were originally cataloged and compiled by Dr. Merton Bernfield and were updated by the author.