Glycoscience

Glycoscience:

The Intriguing Area to Understand Human Biology, Health and Disease

By

Gherman Wiederschain

Cambridge Scholars Publishing

Glycoscience: The Intriguing Area to Understand Human Biology, Health and Disease

By Gherman Wiederschain

This book first published 2024

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

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

ISBN: 978-1-0364-1425-2
ISBN (Ebook): 978-1-0364-1426-9

This book is dedicated to the memory of my teachers, mentors, collaborators and friends all over the world, who have inspired me or helped in some way during my long career in science, particularly in the Glycobiology field.

Orechovich V. N., Rosenfeld E. L., Troitsky G. V., Tzvetkova I. V., Lukomskaya I. S., Belenki D. M., Michailov V. I., Ushakova N., Kozlova I. K., Gorodetsky V. K., Shibaev V., Kochetkov N. K., Derevitzkaja V. A., Chizov O.; Chorlin A. Ya., Stepanenko B. N., Bochkov A. F., Bakinovsky L. V. Vasiljev Y. M., Grachev M. A. – all of them from the USSR, Russia, Moscow.

Victor Ginsburgh (USA), Harry Schachter (Canada), Henri-Géry Hers (Belgium), Roger W. Jeanloz (USA), Jean Montreuil (France), Bergelson L. D. (USSR, Israel), Sharon N. (Israel), Richard Kuhn (Germany), Gauhe A. (Germany), Sen-I. Hakomori (USA), R. McCluer (USA), Lees M. (USA), Jungalwala F. B. (USA), Raghavan S. (USA), Gelfand I. M. (USSR, USA).

“Nature is nowhere accustomed more openly to display her secret mysteries than in cases where she shows traces of her workings apart from the beaten path; nor is there any better way to advance the proper practice of medicine than to give our minds to the discovery of the usual law of nature by careful investigation of cases of rarer forms of disease”.

—William Harvey
(1 April, 1578 – 3 June, 1657)
Discovered Systemic Blood Circulation

TABLE OF CONTENTS

PREFACE

The idea for writing this book was clear after teaching glycobiology for several semesters to Boston College students. Most of them were seniors and by the time of our sessions and seminars, they had already passed through several courses including general biochemistry, molecular and cell biology and other fundamental areas of biology and biochemistry. However, as I understood from talks with students, the area of glycobiology, and specifically, the relationship of this field of science with human diseases, were unclear for practically all of the participants of my seminars. During our seminars about glycoscience, and after reading recommended original papers, reviews and books, students were able to not only accept the basics of glycobiology, but prepared presentations of various topics of glycobiology and posters that were presented and discussed at Annual Hamilton Symposium of Biology Department at Boston College.

The context of this book assumes background knowledge in biochemistry and cell biology at an undergraduate level. My goal was to give readers the most important features of glycobiology, focusing on glycosylation in mammalians and starting from the structure of simple and complex carbohydrates to severe diseases developed through hereditary defects of biosynthesis and degradation of glycoconjugates. I am trying to keep all of this volume of information at a very simple level, while keeping it comprehensive on a molecular level. A variety of human diseases related to the glycodysfunctions are considered here as well because some of these abnormalities very often are useful for diagnosis and strategy of treatment.

The book is divided into II parts. Part I presents basic knowledge about glycans structure, subcellular processes of biosynthesis, degradation, and biological role of glycoconjugates. Part II considers the variety of human diseases that develop due to hereditary deficiency of enzymatic systems related to the biosynthesis and degradation of glycoconjugates and also some diseases that characterize some abnormalities in structure of glycoconjugates and are important markers of certain diseases, including cancer, cardiovascular, gastrointestinal diseases, and others. Part II also summarizes comprehensive data related to the informatics, biotechnology and pharmaceutical industry. I hope the book will be useful for college,

university, and medical school students, and for scientists who are involved in various projects related to glycobiology.

There is also one important aspect of this book. The author dedicated this book to the memory of many of his colleagues, teachers and mentors over the world who are not with us anymore. Their influence, discussion and advice were extremely helpful for my personal experience during various periods of my time in the glycobiology field over more than 60 years. The names of these scientist are shown on the first special page of this book.

I want to thank the many students at the Biology Department of Boston College, who attended my course on Glycobiology and Human Diseases, because by answering my questions and asking me questions they helped to clarify a lot of concepts. There are of course too many people to thank for all the help, support, mentorship, comfort and forgiveness they provided for me over the decades of my career. I am beyond grateful.

Gherman Wiederschain, MD, PhD, Doc.Sci
July 2024, Boston, MA

INTRODUCTION

Glycobiology is focused on understanding the chemical structure, metabolism, and biological function of simple and complex carbohydrates (glycans) and their glycoconjugates: glycoproteins, glycosaminoproteoglycans, and glycolipids. The last three groups of biopolymers, are the final products of glycosylation, the major post-translational modification of proteins and lipids, and play a pivotal role in cellular homeostasis. Protein folding, intercellular communication, embryonic development, the immune system, various types of cell/molecular recognition, and many other processes are characterized by close involvement of glycoconjugates (Varki, 2017).

Glycomics is much younger than genomics and proteomics, but it allows understanding of not only many unknown processes in the healthy state, but has also significantly expanded our knowledge of the mechanisms of various human diseases. This book will present data on unique structural features of simple and complex carbohydrates, cellular and subcellular localization of glycoconjugates, biosynthesis and degradation of glycans and glycoconjugates, as well as their role in determining the fate of molecules, the functionality of the immune system, cell malignancy, and receptor systems. Numerous hereditary disorders developed due to the faulty degradation or biosynthesis of glycans by a certain enzyme in a complex metabolic pathway will also be considered.

However, this book cannot describe and analyze all aspects of glycobiology in a deep level. Many scientific papers, reviews, and monographs highlighting various trends in glycobiology have been published within the past several decades. Along with recently published encyclopedic handbooks in glycobiology with almost 3000 pages (Barchi J., 2021), there are comparatively brief, but rather informative textbooks on glycobiology for students and junior researchers (Kilcoyne and Joshi, 2023; Wiederschain, 2016, 2017). One of these books, Essentials of Glycobiology (Varki and Cummings, 2022), published as the fourth enlarged edition (859 pages), contains contemporary data on many aspects of glycobiology and is a very useful handbook. Some important materials presented in these books related to the glycoscience and human disease were included to our monograph in the brief versions.

Because of the limited volume of the presented book, we will give full titles of the recently published reviews, monographs and collective compendiums summarizing data of many scientific groups in glycobiology with a great variety of approaches. Some of these publications have been used as comprehensive resources for this book, and readers can find these sources in the corresponding references. The purpose of this book is to discuss several trends and progresses, problems, and prospects for the glycoscience as a comparatively young area of research that is already keeps important place in general, cell and molecular biology and with significant impacts to understanding human health and diseases.

This book, as was already mentioned in the Preface, was built base on lectures and seminars course for senior students of Biology Department of Boston College. It's intended to serve the needs the starting researches in the field, but we believe that it will also provide a useful knowledge for more advanced scientists including biochemists, cell biologists, immunologists and virologists, who encounter glycoconjugates in their research.

While keeping human health in the forefront, this book integrates a thorough discussion of glycobiology fundamentals with its growing areas of perspectives from the health, biopharmaceutical, and diagnostic sciences.

Bibliography

Barchi J. (Editor-in Chief), Comprehensive Glycoscience, 2nd edition, V. 1-5, Elsevier, 2021, pp. 2964.

Kilcoyne M. and Joshi L. (Eds.), 2023, Translational Glycobiology in Human Health and Disease, pp.480.

Varki, A. 2017. Biological roles of glycans. Glycobiology, vol. 27, no. 1, 3–49, doi: 10.1093/glycob/cww086.

Varki, A., Cummings, R. D. (Eds. et al.) Essentials of Glycobiology, 4th Edition by Spring Harbor Laboratory Press, 2022, pp. 859.

Wiederschain, G. (Ed.), Glycobiology and Human Diseases, CRC Press, 2016, pp.324.

Wiederschain, G. 2017. Glycobiology, In: Encyclopedia of Physical Organic Chemistry, Zerong Wang (Ed.). John Wiley & Sons, Inc. ISBN 978-1-118-468588-6.

PART I.

FUNDAMENTAL PROCESSES OF STRUCTURE, BIOSYNTHESIS, DEGRADATION, AND BIOLOGICAL ROLE OF GLYCOCONJUGATES

CHAPTER 1

CARBOHYDRATES, MONOSACCHARIDES, GLYCAN, GLYCOCONJUGATES, AND GLYCOSIDIC LINKAGE

1-1. Peculiarity of glycan structure.

In the 19th century, sugar–based substances were referred to as carbohydrates, or "hydrate of carbon", that are based on the general formula Cx(H2O)n that also possess a carbonyl group, either an aldehyde or a ketone. The *n* is integer ranging from 3 to 9. Monosaccharides are the simplest of these polyhydroxylated carbonyl compounds. The name of "saccharide' was derived from the Greek word for sugar or sweetness.

The monosaccharides are join together to give rise to oligosaccharides that usually consist about 20 monosaccharides residues connected by glycosidic linkages. The polysaccharides contain typically more than 20 linear or branched monosaccharide residues, such as cellulose or starch in the plant or glycogen in the animals.

The term "glycoconjugate" is used to describe a macromolecule that contains monosaccharides covalently linked to proteins or lipids (Wiederschain 2016, 2017; Varki 2017; Varki and Cummings, 2022).

Monosaccharides exist in solution as an equilibrium mixture of acyclic and cyclic forms. The percentage of each form depends on the sugar structure. The cyclic form of a monosaccharide is characterized by a hemiacetal group formed by the reaction of one of the hydroxyl groups with the C-1 aldehyde or ketone. For reasons of chemical stability, five- and six-membered rings are most commonly formed from acyclic monosaccharides. Hexoses (six – carbon aldoses) and hexuloses (six-carbon ketoses) form six-membered rings via a C-1—O—C-5 ring closure; they form five-membered rings through a C-1—O—C-4 ring closure. A five-membered cyclic hemiacetal is labeled a "furanose" and a six-membered cyclic hemiacetal is called a "pyranose". Pentoses can form both pyranose and furanose forms

(Seeberger, 2022; Taylor and Drickamer, 2011).

As discriminated from nucleic acids and proteins, which are linear polymers with the same type of bonds between monomers (phosphodiester bonds for nucleic acids and peptide bonds for proteins) within the structures, monosaccharides as a unit of carbohydrate chains can be D- or L-sugars that are connected together via a variety of glycosidic linkages between the carbohydrate units. In addition, the individual carbohydrate moiety, in the case of hexoses, can be in furanoside or pyranoside form that often binds them to each other with α- or β-glycosidic linkage between the most reactive hemiacetal hydroxyl group at the C1 of one monosaccharide and one of the hydroxyl groups at C2, C3, C4, or C6 of another hexose or occasionally also with the C1 atom of the second monosaccharide as shown in trehalose.

Just like polypeptides having amino and carboxyl termini and polynucleotides with 5 and 3 termini, oligosaccharides, as a more complicated carbohydrate chains, have a direction that is defined by their *reducing* and *nonreducing termini*, as illustrated by a disaccharide (e.g., cellobiose) in Fig. 1-1-1.

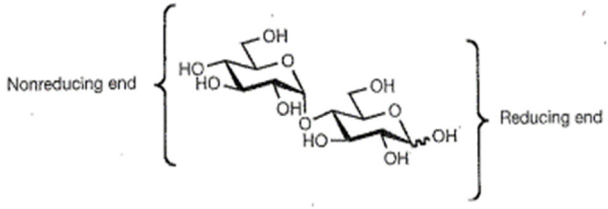

Fig. 1-1-1. Nonreducing and reducing ends

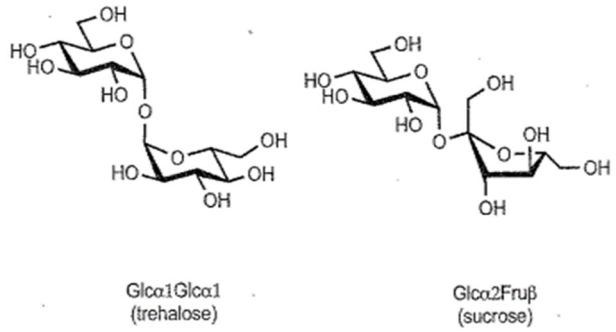

Fig. 1-1-2. Nonreducing disaccharides

Aldoses and ketoses are reducing sugars because they responded positively in a chemical test that effected oxidation of their aldehyde and hydroxyketone functionalities, respectively.

Two monosaccharide units can be joined together by a glycosidic bond – this is fundamental linkage among the monosaccharides units found in all oligo- and polysaccharides.

The glycosidic bond is formed between the anomeric carbon of one monosaccharide and a hydroxyl group of another. In chemical terms, a hemiacetal group reacts with an alcohol group to form an acetal.

It is important that glycosidic bonds can be formed with virtually any hydroxylated compounds, including simple alcohol such methanol or hydrohy amino acids such a serine, threonine, and tyrosine. Indeed, glycosidic linkages are formed between sugars and these amino acids within proteins to form glycoproteins. Like the hemiacetal, the acetal or glycosidic linkage can exist in two stereoisomeric forms: α- or β. But unlike the hemiacetal, the acetal is configurationally stable under most conditions.

When a glycosidic bond is formed, its configuration is maintained indefinitely. The glycosidic linkage is the most flexible part of a disaccharide structure. While the chair conformation of the constituent monosaccharides is relatively rigid, the torsional angles around the glycosidic bond can vary. Thus, disaccharide of well-defined primary structure can adopt multiple conformations in solution that differ in the relative orientation of the two monosaccharides.

The combination of structural rigidity and flexibility is typical of complex carbohydrates and, more than likely, essential to their biological functions. For example, two glucose residues can be joined together to form up to 11 different disaccharides through the variation of the type of glycosidic bond and the linkage between these two glucose units, as illustrated in Fig. 1-1-3 by the structures of eight different disaccharides (e.g., maltose (Glcα1→4Glc) and gentibiose (Glcβ1→6Glc) and others.

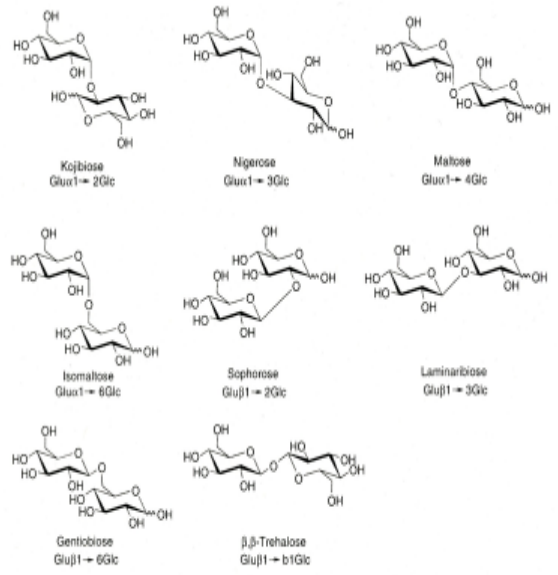

Fig. 1-1-3. The structures of various disaccharides containing only D-glucose

Owing to the difference in the order and type of linkages between the carbohydrate units, oligosaccharides consisting of the same base carbohydrate moieties would have very different three-dimensional structures and biological activities. Unlike the bivalent amino acid and nucleotide that link its preceding and following monomers to form higher order of biomolecules (peptide and oligonucleotide, respectively), a monosaccharide can be involved in more than two glycosidic linkages, thus serving as a branch point. The common occurrence of the branched sequences (as opposed to the linear sequences that are found in almost all peptides and oligonucleotides) is unique to glycans, the polymers of monosaccharides, and contributes to their structural diversity.

In addition, glycans are often linked to other biomolecules, such as lipids or polypeptides, through glycosidic linkages to form glycoconjugates. In these cases, a glycan is often referred to as the *glycone* of glycoconjugates and the noncarbohydrate component (i.e., lipid or polypeptide) is named the *aglycone*.

Glycosidic bonds are the major types of linkage between monosaccharide constituents from disaccharides to oligosaccharides. For example, in lactose (abbreviated as Galβ1→4Glc), a glycosidic linkage has been formed between the C1 of galactose and 4-hydroxyl group of glucose, and the anomeric C1 of galactose is in the β configuration.

The glycosidic linkage is the most flexible part of a disaccharide structure. While the chair conformation of the constituent monosaccharides is relatively rigid, the torsional angles around the glycosidic bond can vary. Thus, disaccharide of well-defined primary structure can adopt multiple conformations in solution that differ in the relative orientation of the two monosaccharides. The combination of structural rigidity and flexibility is typical of complex carbohydrates and, more than likely, essential to their biological functions.

The glucose residue retains the aldehyde function in the form of hemiacetal. Because of the ability of this group to reduce inorganic ions such as Cu2+, it is referred to as the reducing end of the disaccharide. The galactose residue in lactose constitutes the nonreducing end. There is no reducing end for trehalose and sucrose because one of these disaccharides consists anomeric α1-1 linkage (trehalose) between two glucose components and the other (sucrose, table sugar) is O-α-D-glucopyranosyl-(1-2)-β-D-fructofuranoside, as shown in Fig. 1-1-2.

It should be pointed out that the hydroxyl groups of both monosaccharides and oligosaccharides can be biochemically (by enzymes) or chemically modified without affecting the glycosidic linkages. The often applied chemical modifications of hydroxyl groups include phosphorylation, methylation, sulfation, esterification, oxidation, and deoxygenation.

The chemical modifications on hydroxyl groups of sugars are commonly involved during the synthesis of mono- and oligosaccharides. For example, deoxygenation of carbohydrate is to replace the hydroxyl groups of carbohydrate moiety with hydrogen atoms to form the deoxysugars, as represented by the natural L-fucose, and L-rhamnose. Some of the carbohydrates arising from the representative modifications are depicted in Fig.1-1-4.

The limited availability of structurally well-defined diverse glycans remains a major obstacle for deciphering biological functions as well as biomedical applications of carbohydrates. Despite tremendous progress that has been made in past decades, the synthesis of structurally well-defined complex

glycans still represents one of the most challenging topics in synthetic chemistry. Chemical synthesis of glycans is a time-consuming and labor-intensive process that requires elaborate planning and skilled personnel. In contrast, glycosyltransferase-catalyzed enzymatic synthesis provides a more efficient, convenient, low-cost, and sustainable alternative to affording diverse and complex glycans. However, the existing methods are still insufficient to fulfill the increasing demand for specific synthetic glycan libraries necessary for functional glycomics research. This is mainly attributed to the inherent character of the glycan biosynthetic pathway. In nature, there are too many glycosyltransferases involved in the in vivo glycan synthesis, but only a small number of them are available for in vitro enzymatic synthesis. For instance, humans have over 200 glycosyltransferases, but only a few of them could be produced from the conventional bacterial expression system, and most of these membrane-associated enzymes could be overexpressed only in eukaryotic cells. Moreover, the glycan biosynthetic pathway is an on template-driven process, which eventually ends up with heterogeneous glycan product mixtures. Therefore, it is not a practical solution for the in vitro enzymatic synthesis of complex glycans by simply copying the glycan biosynthetic pathway. In the past decade, we have tried to develop a simplified and transformable approach to the enzymatic modular assembly of a human glycan library. Despite the structural complexity of human glycans, the glyco informatics analysis based on the known glycan structure database and the human glycosyltransferase database indicates that there are approximately 56 disaccharide patterns present in the human glycome and only16 disaccharide linkages are required to account for over 80%of the total disaccharide fragments, while 35 disaccharide linkages are sufficient to cover over 95%of all disaccharide fragments of human glycome. Regardless of the substrate specificity, if one glycosyltransferase could be used for the synthesis of all of the same glycosidic linkages in human glycome, it will require only a few dozen glycosyltransferases for the assembly of entire human glycans. According to the glyco bioinformatics analysis results, we rationally designed about two dozen enzyme modules for the synthesis of over 20 common glycosidic linkages in human glycome, in which each enzyme module contains a glycosyltransferase and a group of enzymes for the in situ generation of a nucleotide-activated sugar donor. By sequential glycosylation using orchestrated enzyme modules, we have completed the synthesis of over 200 structurally well-defined complex human glycans including blood group antigens, O-mannosyl-glycans, human milk oligosaccharides, and others. To overcome the product microheterogeneity problem of enzymatic synthesis in the nontemplate-

driven glycan biosynthetic pathway, we developed several substrate engineering strategies to control or manipulate the outcome of glycosyltransferase-catalyzed reactions for the precise synthesis of structurally well-defined isomeric complex glycans (Liu et al. 2024).

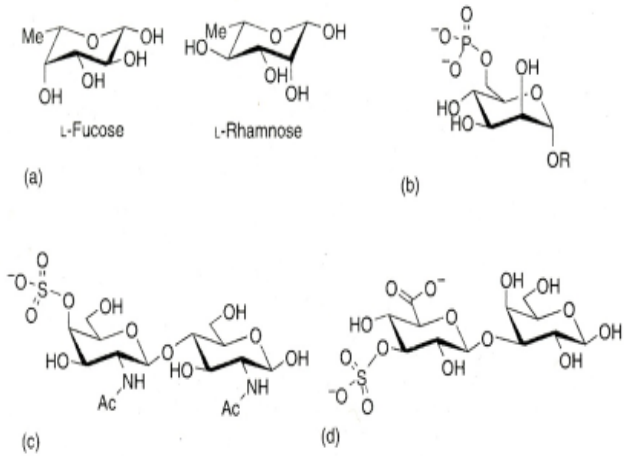

Fig. 1-1-4. Illustration of sugar modifications: (a) deoxygenation, (b) phosphorylation, (c) sulfation, and (d) oxidation.

Bibliography

Liu, C-C.; Ye, J. and Cao, H., 2024.Chemical Evolution of Enzyme-Catalyzed Glycosylation. Chem. Res., 57, p. 636−647.

Seeberger, P., 2022. Monosaccharide diversity. In: Varki A., Cummings R. C. (Eds. et al.) Essentials of Glycobiology, 4th Edition by Spring Harbor Laboratory Press, p. 21-32.

Taylor, M. and Drickamer K., 2011. Introduction to glycobiology. 3rd ed., Oxford University Press, Oxford, New York, pp. 283.

Varki, A., Cummings R. D. (Eds. et al.), Essentials of Glycobiology, 4th Edition by Spring Harbor Laboratory Press, 2022, pp. 859.

Wiederschain, G. (Ed.), Glycobiology and Human Diseases, CRC Press, 2016, pp. 324.

Wiederschain, G., Glycobiology, In: Encyclopedia of Physical Organic Chemistry. Wang, Z. (Ed.), 2017, John Wiley & Sons, Inc. ISBN 978-1-118-468588-6.

CHAPTER 2

UNIQUE INFORMATIONAL FEATURES OF SIMPLE AND COMPLEX CARBOHYDRATE MOLECULES AND THEIR STRUCTURAL SIGNIFICANCE AS A GLYCOCODE FOR BIOLOGICAL MOLECULAR DIVERSITY

Due to the nature of multifunctionality within a single carbohydrate moiety, all possible connections between monosaccharides lead to the formation of versatile carbohydrate structures.

For example, three identical monosaccharides of hexoses (e.g., β-glucopyranoside) can theoretically be jointed to produce 176 different trisaccharides, and three different monosaccharides (XYZ) can be connected to produce 1056 possible trisaccharides. These many possible trisaccharides include not only the ones arising from different glycosidic linkages between the monosaccharides (i.e., 1→2, 1→3, 1→4, and 1→6) but also the carbohydrates containing different type of glycoside bonds at the anomeric positions (i.e., αα, αβ, or ββ). In contrast, three identical amino acid units can only be linked together to form one tripeptide with the same peptide bond, whereas three different amino acids can form only six different tripeptides (i.e., 3! =6, ABC, ACB, BCA, BAC, CAB, and CBA).

The longer carbohydrate chains, including those of simple glycans, polysaccharides, and glycoconjugates, can also be branched. Such combinations significantly increase the number of possible carbohydrate structures having the same monosaccharide sequence. Thus, a pentasaccharide consisting of five different hexose units has 2 144 640 possible structural isomers, and a hexasaccharide consisting of six different hexoses has more than a trillion possible structural combinations (Sharon, 1975; Varki and Sharon, 2009). It seems that the spectrum of carbohydrate structures is expanding and yielding new combinations under the influence of various factors, including disease-associated changes in the body. The multiple monosaccharide building blocks can be linked to various

regiochemistries and stereochemistries, and the resulting oligosaccharides can be assembled on protein or lipid scaffolds. Glycoconjugates therefore comprise an "information-rich" system capable of participating in a wide range of biological functions. Fortunately, the natural glycans are not so tremendously diversified. From approximately 200 known carbohydrate chains of numerous glycoconjugates isolated from vertebrates, we can identify only 10–13 sugars and their derivatives as shown in Table 1 and Fig. 2-1.

TABLE 1 The Simple Carbohydrates Isolated from Glycoconjugates of Vertebrates

Carbohydrate	Abbreviation	Carbohydrate	Abbreviation
Abequose	Abe	Mannose	Man
Arabinose	Ara	Muramic acid	Mur
Fructose	Fru	*N*-Acetylgalactosamine	GalNAc
Fucose	Fuc	*N*-Acetylglucosamine	GlcNAc
Galactosamine	GalN	*N*-Acetylmuramic acid	Mur2Ac
Galactose	Gal	*N*-Acetylneuraminic acid (a sialic acid)	Neu5Ac
Glucosamine	GlcN	Rhamnose	Rha
Glucose	Glc	Ribose	Rib
Glucuronic acid	GlcA	Xylose	Xyl
Iduronic acid	IdoA		

Among these sugars, *N*-acetylneuraminate (*N*-acetylneuraminic and Glycolylneuraminic acids, also called sialic acids) is often found as a terminal residue of oligosaccharide chains in glycoproteins. Sialic acid imparts a negative charge to glycoproteins and glycolipids because its carboxyl group tends to dissociate a proton at the physiological pH.

While the comparatively limited number of monosaccharide constituents potentially can provide extremely huge number of carbohydrate chains in glycoconjugates, when these basic carbohydrates are linked together to form polysaccharides, the resulting structures are unexpectedly simpler than those found in glycoconjugates. Below are the examples of three polysaccharides of various biological functions that consist of only one monosaccharide, D-glucose.

Polysaccharides such as starch, glycogen, and cellulose are polymers D-glucose, containing different types of glycosidic linkages (α or β glycosidic

bonds) between the glucose units at the anomeric positions of each sugar unit and different degrees of carbohydrate chain branching through the 1→6 linkages, leading to discrete conformation and properties of each of these three homopolysaccharides.

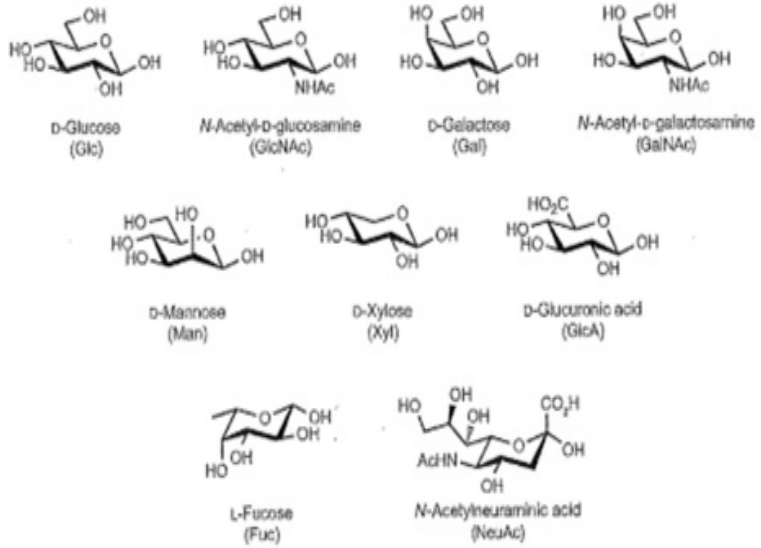

Fig. 2-1-1. The structures of common monosaccharides found in vertebrates.

Starch itself is composed of both linear and branched homopolysaccharides of D-glucose, that is, amylose and amylopectin. In amylose, the glucoses are linked by α (1→4) glycosidic bonds, meaning that the preceding glucose unit is connected with the carbon atom at position 1 to the carbon atom at position 4 of the following glucose unit with α configuration at the anomeric position, and the ending glucose unit is in a reducing hemiacetal format. This produces an unbranched chain of glucose that then folds up to form a coil or helix. By contrast, in addition to the linear chain of glucose like that in amylose, amylopectin contains branches of d-glucose connecting to the linear chain with α (1→6) linkages, roughly about every 20–24 glucose units on the amylose chain.

Starch forms a hollow helical structure with about 20–25 glucose residues of amylose and is the most suitable energy storing material for green plants.

Glycogen, a high-molecular-weight polymer of glucose with a starch like helical structure, has similar sets of bonds, α (1→4) bonds in the linear moiety of the molecule and α (1→6) bonds at the branching points that occur every 4–8 glucose residues. Glycogen is the main energy storing material not only in animal cells, but also in insects, fungi, and yeast.

The strongest fibers of plant cell walls are composed of cellulose, which is a linear homopolysaccharide of glucose jointed through β (1→4) glycosidic bonds and is unrelated to any of the energy storage polymers (Dumitriu, 2005). Cellulose is a linear polymer of β-(1→4)-d-glucopyranose in 4C1 conformation. The fully equatorial glycosidic bonds of β-linked glucopyranose residues allow the formation of the most stable chair conformation of the six-membered cyclic structures, minimizing its flexibility (e.g., relative to the slightly more flexible α-linked glucopyranose residues in starch). Along the homopolysaccharide chain, every other glucose is flipped over, due to the β-glycosidic linkages, promoting the formation of hydrogen bonds between the two consecutive glucose units, so-called the intra-chain hydrogen bonds. This promotes intra-chain and inter-chain hydrogen bonds and van der Waals interactions that cause cellulose chains to be straight and rigid, and pack with a crystalline arrangement in thick bundles – microfibrils. The role of cellulose is to provide strength and rigidity to plant cell walls, which can withstand high gradients of hydrostatic pressure and prevent osmotic swelling of plant cells.

It should be pointed out that even a small difference in chemical bonds could have a great impact on the overall physical properties and chemical reactivity of the resulting polymers, as demonstrated in the case of starch and cellulose, in which starch contains the α-glycosidic bonds, whereas cellulose carries only β-glycosidic bonds to connect the same monomers, that is, D-glucose. As a result, human beings can digest starch but cannot break down cellulose because our body contains enzymes that break starch down into glucose to fuel the body, whereas our body does not have the enzymes that can break down the β-glycosidic bonds between the carbohydrate units. Similarly, even the disaccharide consisting of two sugar units with β-glycosidic bond would cause health problem for certain people, as shown in the case of lactose intolerance symptoms, owing to the lack of lactase generated in the intestine to break down the lactose in the dairy products, causing stomach cramps and diarrhea.

Pectins, the other group of plant polysaccharides, are biological modulators of variety of human physiological reactions. The primary plant cell wall is

composed of a mixture of cellulose, hemicellulose, and pectin, while the middle lamella can be considered as an extension of this matrix material from which the cellulose are lacking. Pectic substances make up about 30 percentage of the dry matter of the primary cell wall and are the primary macromolecules of the middle lamella. In the plant cell wall several dozen chains of cellulose interconnected by hydrogen and van der Waals forces form linear microfibrils. Molecules of hemicelluloses, which include xyloglucan, xylans, glucurono-arabino-xylans, gluco-, galacto-, and galacto-glucomannans are situated between these microfibrils forming network with the molecules of cellulose. Pectin has promising pharmaceutical uses due to its low toxicity and water solubility. Pectin hydrogels have been used in tablet formulations as a binding agent and have been used in controlled-release matrix tablet formulations (Popov 2016). The modern lead pectin-based drug delivery product appears to be a fentanyl nasal spray formulation, which has successfully met the primary objective in a pivotal Phase III clinical study as treatment for breakthrough pain that is induced by chemoradiotherapy in head and neck cancer. Pectin has been found to be useful in the prevention and therapy of a variety of diseases. Pectins have been shown to possess hypolipidemic, hypoglycemic, anti-cancer, anti-infective, anti-ulcer, immunostimulating, anti-inflammatory, antioxidant and other effects. The enormous literature on the physiological activity of pectin includes clinical and epidemiological observations as well as experimental studies with different animal models and cell test systems.

The effects based on the ability of the pectin to induce a viscous or gelled gastro-intestinal content appear to represent the most studied effects of pectin on humans. Many studies demonstrate that pectin affects lipid metabolism, enhances glucose tolerance, increases satiety and promotes gastrointestinal health. Importantly, most of the data mentioned in the review were obtained using commercial pectins. Although the majority of plant tissues contain pectins, their industrial production is almost entirely based on only a few sources, such as citrus peel and apple pomace. The structure of common commercial pectins appears to be rather simple. The conditions of the industrial production of pectin are optimized to improve the content of the linear PG fragments, which are necessary for the gelation of pectins. Commercial apple and citrus pectins are considered to be devoid of the side-branched sugar chains. In view of the present knowledge on the complexity of the structure of the entire pectin macromolecule, it is timely and important to study the role of pectins that contain both linear and branch areas in human health. Therefore, the first problem to be solved is to investigate the physiological effects of native pectins of different plants. The second item that requires investigation is the physiological effects of

pectin at lower doses. Indeed, most of the data summarized here were obtained using an intake of at least 10 g of pectin. It is not clear whether low pectin doses that failed to increase the viscosity of the gastrointestinal content could affect any other physiological processes. Paradoxically, clinical and epidemiological data for a protective role for pectin in human cancer and inflammation are weak. Only a few human trials reported on the anti-cancer, anti-inflammatory, and anti-allergic effects of ingested pectin. The mechanism of these effects is not clear. The proposed interaction of galactose residues of MCP with galectin-3 has not been demonstrated in humans yet. Therefore, the pectins as biological modulators of human physiological reactions should be to conduct extensive human studies on the anti-cancer and anti-inflammatory effects of pectin. The mechanism of the physiological effects of pectin is thought to be realized inside the gastrointestinal tract because pectin is not digested and is not absorbed into the blood stream. The main physiological processes that are affected by pectin appear to be enterohepatic recirculation, gastric emptying and the absorption of nutrients. Indirectly, pectin is considered to influence the body's metabolism through having an effect on the gut microbiota by increasing the population of healthy microflora and stimulating the production of SCFA. The MeOH-mediated regulation of gene transcriptional activity after pectin ingestion appears to represent a new possible mechanism for the physiological activity of pectin. In addition to digestive and gastrointestinal processes, other physiological functions, such as cognitive functions, could be proposed to be influenced by pectin. Another indirect possible mechanism that should be studied is the production by the human immune system of autoantibodies against pectin polysaccharide chains. These antibodies could be involved in both physiological and pathological processes. Studies of the effects of pectin on immunity are of special interest because of the importance of the immune system in human health. It is suggested that future research on the effect of pectin on physiological functions will provide valuable insight into undefined mechanisms and could lead to new strategies to derive the greatest benefit from the rational use of dietary pectin (Popov 2016).

Some animals do carry enzymes to digest cellulose so that they can live on cellulose, like termites to eat wood, and cattle to eat grass. In addition, due to the intrachain hydrogen bonds between the glucose unit, cellulose is a much stronger than starch. Starch is practically useless as a material, but cellulose is strong enough to make fibers, and hence rope, clothing, and so on. Cellulose does not dissolve in water the way starch will and does not break down as easily. Otherwise, if cellulose is as soluble in water as starch, then a good soaking rain would wash away all the wooden houses, park

benches, and playground equipment.

Still, more structurally diverse compounds exist by combining carbohydrate moieties into other molecules, where glycoproteins, proteoglycans, and glycolipids are the most important biological glycoconjugates. The structures of these complex compounds, their nontemplate syntheses under the influence of highly specific glycosyltransferases and degradations under the influence of glycosidases, are being studied in many research laboratories with remarkable progress (Ban et al. 2012; Fraser-Reid et al., 2008; Lairson et al. 2008; Rini et al., 2022). One of the most fascinating aspects of glycosylation of biopolymers is the phenomenon of microheterogeneity. This term indicates that at any attachment site for glycan on a given protein and lipid synthesized by a particular cell type, a range of variations can be found in the glycosylated structures.

The extent of this microheterogeneity can vary considerably from one glycosylation site to another, even in the structures of the same type of molecules synthesized in one cell or between various cells. Such microheterogeneity is partially associated with the difficulty in establishing the relationship between the structures of glycoconjugates and their biological diversity and complexity. It also makes the study of structure of glycans more difficult with respect to the structure of nucleic acids and proteins.

Overall, carbohydrates are ideal substrates for generating explicit informational properties since the numbers of combinations for the permutations of linkages in carbohydrates are larger than that can be achieved by amino acids in peptides or proteins and are unique in biological polymers and glycoconjugates, where branching is possible. Compared to amino acids in proteins and nucleotides in nucleic acids, the versatility of carbohydrate to form structural isomers is unsurpassed.

The resulting highly dense coding capacity of carbohydrate chains is established by the variability in (i) anomeric status, (ii) linkage positions, (iii) ring size, (iv) the type of branching, and (v) introduction of site-specific substituents.

The original concept of the Central Dogma by Francis Crick ignored posttranslational modifications (such as protein glycosylation) that greatly magnify the functions of a single protein encoded by a particular gene. The detail analysis and significant data for supporting this point of view already summarized in the book “The Sugar Code. Fundamentals of Glycoscience” (Gabius, 2009).

It should be pointed out that the oligosaccharide units are not flexible but exhibit highly specific structures with only limited degrees of freedom. The different building blocks of glycans (the alphabet of the sugar code) and the way they can be linked to form oligo- and polysaccharides are much more informative with some conformational peculiarity that transforms two-dimensional biological code to three-dimensional structures and significantly increase the spectrum of biological specificity of glycoconjugates (Cummings 2024).

Bibliography

Ban, L.; Pettit, N.; Li, L. et al., 2012. Discovery of glycosyltransferases using carbohydrate arrays and mass spectrometry. Nat. Chem. Biol., **8**, 769–773.

Cummings, R. D. 2024. Evolution and diversity of glycomolecules from unicellular organisms to human. BioCosmos 1, 1-35.

Dumitriu, N., (Ed.), 2005. Polysaccharides. Structural Diversity and Functional Versatility, 2nd ed., Marcel Dekker, New York.

Fraser-Reid, B.O., Tatsuta, K., and Thiem, J., 2008. Glycoscience. Chemistry and Chemical Biology, 2nd ed., Springer-Verlag, Tokyo.

Gabius, H.-J. (Ed.), 2009. The Sugar Code: Fundamentals of Glycoscience, Wiley-VCH Verlag, GmbH & Co, KGaA, Weinheim.

Hu, Y. and Walker, S., 2002. Remarkable structural similarities between diverse glycosyltransferases. Chem. Biol., **9**, 1287–1296.

Kamerling, J.P., Boons, G.-J., Lee, Y.C. et al. 2007. Comprehensive Glycoscience. From Chemistry to System Biology, vol. 1-4, Elsevier, Amsterdam.

Lairson, L.L, Henrissat, B., Davies, G.J., and Withers, S.G., 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem., **77**, 521–555.

Lennarz, W. and Hart, G., 1994. Methods in Enzymology, Guide to Techniques in Glycobiology, vol. 230, Academic Press, San Diego.

Popov, S. V., 2016. Pectins as biological modulators of human physiological reactions. In: "Glycobiology and Human Diseases", (Wiederschain, G., Ed.), CRC Press, Taylor & Francis Group, p. 202-222.

Rini, J.; Moremen K.; Davis B., and Esko, J.D. 2022. Glycosyltransferases and glycan-processing enzymes. In: Essentials of Glycobiology, 4th ed., p. 67-78. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sansom, S. and Markman, O., 2007. Glycobiology, Scion Publishing Ltd.