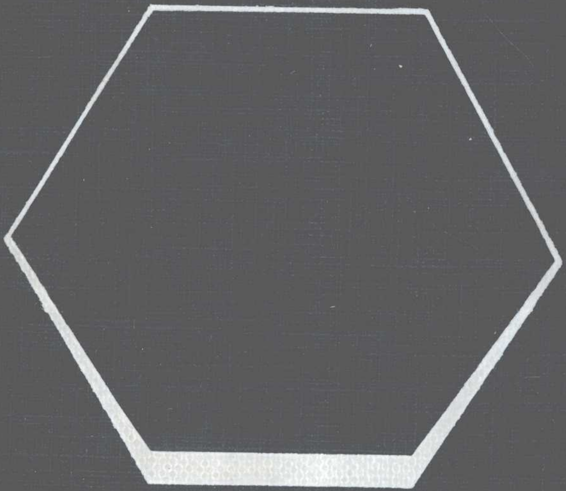
Carbohydrate Chemistry

Monosaccharides and Their Oligomers



Hassan S. El Khadem

Carbohydrate Chemistry

Monosaccharides and Their Oligomers

Hassan S. El Khadem

The American University
Washington D.C.

ACADEMIC PRESS, INC.
Harcourt Brace Jovanovich, Publishers
San Diego New York Berkeley Bc
London Sydney Tokyo Toronto

ACADEMIC PRESS, INC.
1250 Sixth Avenue
San Diego, California 92101

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD.
24-28 Oval Road, London NW1 7DX

Library of Congress Cataloging-in-Publication Data

El Khadem, Hassan Saad, Date
Carbohydrate chemistry.

Bibliography: p.
Includes index.
1. Monosaccharides. 2. Carbohydrates. I. Title.
QD321.E48 1988 547.7′813 87-47760
ISBN 0-12-236870-3 (alk. paper)

PRINTED IN THE UNITED STATES OF AMERICA
88 89 90 91 9 8 7 6 5 4 3 2 1

Preface

Carbohydrate Chemistry: Monosaccharides and Their Oligomers was designed as a classroom textbook for college carbohydrate chemistry courses. It is primarily intended for use by undergraduate and graduate students enrolled in chemistry, biochemistry, and pre-med curricula. It is also intended for students in the colleges of arts and sciences, pharmacy, agriculture, and medicine, who are engaged in research in the fields of carbohydrates and natural product chemistry. Its format fills a gap between large, multivolume reference books designed mainly for research and elementary, sometimes superficial books.

Review articles dealing with topics discussed in the various chapters are listed in chapter order in the bibliography at the end of the book. In addition, references for a small number of recent articles, not found in these reviews, appear throughout the text. Some problems on NMR spectroscopy and other topics have also been added to Chapters 3–7. Answers to the problems are in the appendix at the back of the book.

The book is divided into two parts. The first, "Monosaccharides," deals with monomeric carbohydrates, whereas oligosaccharides and oligonucleotides are discussed in the second. Chapter 1 is an introduction that outlines the importance of carbohydrates as a major group of naturally occurring compounds. After a short historical section on the discovery of sugars, the classification of carbohydrates is discussed in detail.

The four chapters that follow deal with the chemistry of monosaccharides. Thus, Chapter 2 discusses the determination of the structure, configuration, and conformation of monosaccharides. Although the elucidation of the structure of monosaccharides was achieved at a time when sophisticated instrumental techniques were not available, infrared and nuclear magnetic resonance (NMR) spectroscopy, as well as mass spectrometry, are extensively discussed in this chapter. Nomenclature is discussed next, in a short chapter that illustrates the proper use of the rules adopted by the International Committee on Carbohydrate Nomenclature and by *Chemical Abstracts*.

The two longest chapters in Part I are Chapter 4, "Physical Properties Used in Structure Elucidation," and Chapter 5, "Reactions of Monosaccharides." The first deals with spectroscopic methods, such as NMR (^{1}H, ^{13}C, and ^{15}N) and molecular and electronic spectroscopy, as well as mass spectrometry. These, together with the optical properties of monosaccharides (optical rotation, optical rotatory dispersion, and circular dichroism), are extensively used today in structure determinations in the field of carbohydrates. Chapter 5 deals with the reactions of monosaccharides, starting with addition reactions of the carbonyl group. These are followed by nucleophilic substitution reactions at the anomeric carbon atom and at the less reactive nonanomeric carbon atoms. Oxidations and reductions of monosaccharides are then discussed, followed by a section on how to plan retrosynthetic schemes.

Part II starts with Chapter 6 on the methods used for the structure elucidation of oligosaccharides. These include wet chemical methods suitable for the study of oligomers available in gram quantities and nondestructive physical methods capable of handling milligram quantities of rare oligomers. The synthesis of oligosaccharides and their chemical modification is treated in Chapter 7, which includes some novel approaches to the synthesis of these oligomers.

The author is deeply indebted to R. S. Tipson, who read the entire manuscript and made several useful suggestions. Thanks are also due to Derek Horton and Stephen Hanessian, who read a second draft of the manuscript and made valuable suggestions. In addition, D. L. Swartz offered suggestions and helped with the operation of a word processor. Many figures and diagrams were made available by authors for use in the text; for example, photographs of neutron diffraction analyses were presented by G. A. Jeffrey, and of NMR spectra by B. Coxon and A. C. Allen. Their help and that of other friends is greatly appreciated.

Contents

I. Monosaccharides

II. Oligomeric Saccharides: Oligosaccharides and Nucleotides

List of Tables

1

Carbohydrates

I. HISTORICAL BACKGROUND

The origins of carbohydrate chemistry can be traced back to the civilizations of antiquity. Thus, for example, the manufacture of beer and wine by alcoholic fermentation of grain starch and grape sugar is well documented on the walls of ancient Egyptian tombs. The isolation of cellulose fibers by retting flax was widely used by the civilizations of the Far East and the Near East and was introduced by the Greeks to Europe. It is also known that gums and resins were valued commodities at the dawn of the Christian era.

The isolation of sucrose from the juice of sugarcane marks an important milestone in sugar chemistry. In the Far East, where this plant grew, sugar was isolated as a yellowish syrupy concentrate which crystallized, on standing, into a brown mass, and a number of Chinese recipes dating from the fourth century describe in detail how the sugarcane juice was concentrated. It is interesting that in Sanskrit the word *sugar* means *sweet sand,* which aptly describes the properties of crushed raw sugar. Significant progress in the manufacture of sugar occurred after the French Revolution, when Europe, under blockade, had to rely on the sugar beet for manufacture of this valued commodity. Improved methods of isolation and refining were developed, including treating the syrup with lime to precipitate the calcium complex, regenerating the sugar with sulfur dioxide, and decolorizing it with animal charcoal.

Another important industrial process in use today that may be traced to Far Eastern origins is the manufacture of paper from wood pulp. This was achieved around A.D. 600 in China and constituted a marked improvement over the then-available Egyptian papyri. The paper produced was, however, brittle and tore when folded, because the lignin had not been removed. The art of papermaking was brought by Marco Polo to Europe, where methods for removing lignin from the pulp were devised, first using alkali and then sulfite.

II. IMPORTANCE OF CARBOHYDRATES

In addition to the manufacture of sucrose and paper, which are today important industries, carbohydrates play a major role in a number of other industries. These include (a) the food industry, which uses huge amounts of starch in various degrees of purity in the manufacture of baked goods and pastas, of gums in food processing, and of mono- and oligosaccharides as sweeteners, and employs them in fermentation to make beer and wine; (b) the textile industry, which, despite the advent of synthetics, is still dependent to a large extent on cellulose; (c) the pharmaceutical industry, particularly in the areas of antibiotics, intravenous solutions, and vitamin C; and (d) the chemical industry, which produces and markets several pure sugars and their derivatives.

Carbohydrates also play a key role in the process of life; thus, the *master molecule* DNA is a polymer made up of repeating units composed of four nucleotides of 2-deoxy-D-*erythro*-pentofuranose (''2-deoxy-D-ribose''), whose sequence constitutes the coded template responsible for replication and transcription. The science of molecular biology is based on the properties of the polymers of these nucleotides. Saccharide derivatives also form part of many vital enzyme systems, specifically as coenzymes. They are also responsible for cell recognition and are therefore of great importance in immunology.

The best known use of carbohydrates is undoubtedly in nutrition, as members of a major food type which is metabolized to produce energy. Although the average percentage of carbohydrates consumed by humans in comparison to other food types differs from country to country, and figures are often inaccurate or unavailable, the value for the world as a whole has been estimated at over 80%. The exothermic reactions that produce energy in the cell are the outcome of a number of complex enzyme cycles that originate with hexoses and end up with one-, two-, or three-carbon units. The energy released from these reactions is stored in the cell in the form of a key sugar intermediate, adenosine triphosphate

(ATP), and released when needed by its conversion into the di- or monophosphate. Finally, it should be realized that carbohydrates are the most abundant organic components of plants (more than 50% of the dry weight) and therefore constitute the major part of our renewable fuels and the starting material from which most of our fossil fuels were engendered.

III. DEFINITIONS AND SCOPE

In the nineteenth century, when the empirical formulas of many organic compounds were being determined, it was discovered that all of the sugars known at the time had the formula $C_x(H_2O)_y$. They were accordingly thought to be hydrated carbons and so were called *carbohydrates*. This name was applied not only to the soluble sugars but also to polysaccharides, such as starch, that were thought, because of their formulas, to be hydrated carbons. Today's usage of the word *carbohydrate* applies to a large number of organic compounds, monomeric, oligomeric, and polymeric in nature, which do not necessarily have their hydrogen and oxygen atoms in the molecular ratio of 2 : 1 but which can be either synthesized from or hydrolyzed to monosaccharides.

Because the word *sugar* has been used to mean (a) sucrose by the general public, (b) glucose by the medical profession, and (c) mono- to oligosaccharides by chemists, it is slowly being replaced by the less ambiguous term *saccharide,* meaning sugarlike, which is usually prefixed by mono-, di-, oligo-, or poly- to designate the degree of polymerization of the specific compound under discussion. Thus, today one speaks of the chemistry of *carbohydrates* and the reactions of *monosaccharides*.

IV. CLASSIFICATION OF CARBOHYDRATES

Carbohydrates are classified, according to their degree of polymerization, into monomeric carbohydrates, which include monosaccharides and their derivatives, and polymeric carbohydrates, which comprise oligosaccharides, polysaccharides, DNA, and RNA. These polymeric carbohydrates differ in the type of bridge that links their monosaccharide units. Thus, oligosaccharides and polysaccharides are polyacetals, linked by acetal oxygen bridges, whereas DNA and RNA are poly(phosphoric esters), linked by phosphate bridges. In addition to these well-defined groups of carbohydrates, there exist a number of derivatives, for example, antibiotics, that are best studied as a separate group, because some of their members may be monomeric, whereas others are oligomeric.

A. Monosaccharides

Monosaccharides are chiral polyhydroxyalkanals or polyhydroxyalkanones which often exist in cylic hemiacetal forms. Monosaccharides are divided into two major groups according to whether their acyclic forms possess an aldehyde group or a keto group, that is, into aldoses or ketoses (glyculoses). These, in turn, are each classified, according to the number of carbon atoms in the monosaccharide chain (usually 3–10), into trioses, tetroses, pentoses, hexoses, etc. By prefixing aldo- to these names one may define more closely a group of aldoses, for example, aldopentoses, whereas for ketoses it is customary to use the ending -ulose, as in hexuloses. Finally, monosaccharides may be grouped according to the size of their rings into five-membered furanoses and six-membered pyranoses. It should be noted that, in order to form a furanose ring, four carbon atoms and one oxygen atom are needed, so only aldotetroses and higher aldoses and 2-pentuloses and higher ketoses can cyclize in this ring form. Similarly, in order to form a pyranose ring, five carbon atoms and one oxygen atom are required, so only aldopentoses and 2-hexuloses as well as their higher analogs can cyclize in this form. Ultimately, by combining the ring type with the names used above (for example, aldopentose or hexulose), such combination names as aldopentofuranoses and hexulopyranoses can be formed, which define without ambiguity the group to which a monosaccharide belongs (see Table I).

As their name denotes, monosaccharides are monomeric in nature and, unlike the oligosaccharides and polysaccharides, which will be discussed later, they cannot be depolymerized by hydrolysis to simpler sugars. Monosaccharides and oligosaccharides are soluble in water and their solutions in water are often sweet-tasting, which is why they are referred to as sugars.

B. Oligosaccharides

Oligosaccharides and polysaccharides are polyacetals which respectively have, as their names denote (from the Greek *oligos*, few; and *poly*, many), a low (DP = 2–10) or a high (DP >10) degree of polymerization. They are composed of a number of monosaccharides linked together by acetal oxygen bridges, and, upon depolymerization (hydrolysis), they yield one or more types of monosaccharide. Oligosaccharides and polysaccharides are further grouped into (a) simple (true) oligosaccharides and polysaccharides, which are oligomers and polymers of monosaccharides that yield on *complete* hydrolysis *only* monosaccharides; and (b) conjugate oligosaccharides and polysaccharides, which are oligomers and poly-

TABLE I
Monosaccharides

Aldoses				
Monosaccharide	Aldose[a]	No. of chiral C	Aldofuranose	Aldopyranose
Triose	Aldotriose	1	—	—
Tetrose	Aldotetrose	2	Tetrofuranose; aldotetrofuranose	— —
Pentose	Aldopentose	3	Pentofuranose; aldopentofuranose	Pentopyranose; aldopentopyranose
Hexose	Aldohexose	4	Hexofuranose; aldohexofuranose	Hexopyranose; aldohexopyranose
Heptose	Aldoheptose	5	Heptofuranose; aldoheptofuranose	Heptopyranose; aldoheptopyranose
Octose	Aldo-octose	6	Octofuranose; aldo-octofuranose	Octopyranose; aldo-octopyranose
Nonose	Aldononose	7	Nonofuranose; aldononofuranose	Nonopyranose; aldononopyranose
Decose	Aldodecose	8	Decofuranose; aldodecofuranose	Decopyranose; aldodecopyranose

Ketoses (Glyculoses)				
Monosaccharide	Ketose[b]	No. of chiral C	Ketofuranose	Ketopyranose
Tetrose	Tetrulose	1	—	—
Pentose	Pentulose	2	Pentulofuranose	—
Hexose	Hexulose	3	Hexulofuranose	Hexulopyranose
Heptose	Heptulose	4	Heptulofuranose	Heptulopyranose
Octose	Octulose	5	Octulofuranose	Octulopyranose
Nonose	Nonulose	6	Nonulofuranose	Nonulopyranose
Decose	Deculose	7	Deculofuranose	Deculopyranose

[a] Although an achiral aldobiose (glycolaldehyde) exists, it is not considered to be a saccharide because, by definition, a saccharide must contain at least one asymmetric carbon atom.

[b] Although an achiral triulose (1,3-dihydroxyacetone) exists, it is not considered to be a saccharide because it lacks an asymmetric carbon atom.

mers of monosaccharides linked to a nonsaccharide, such as a lipid or a peptide. One can further classify simple oligosaccharides, according to degree of polymerization, into disaccharides, trisaccharides, tetrasaccharides, etc., and according to whether or not the oligomer chain has at one end a hemiacetal function (a latent aldehyde or keto group). Such terminal groups are readily converted into carboxylic groups by mild oxidants and, accordingly, oligosaccharides possessing these groups are referred to as *reducing,* in contradistinction to those which resist such oxidation and are designated *nonreducing*. Thus, whereas all monosaccharides are reducing, there are reducing and nonreducing disaccharides, trisaccharides, etc. Because of their sweet taste, monosaccharides and lower oligosaccharides have been called sugars. It should be noted, however, that sweetness decreases with increasing DP of oligosaccharides, and beyond a DP of 4 the oligomer is tasteless. Table II shows the classification of oligosaccharides and gives the names of the most common ones.

C. Polysaccharides

Polysaccharides and oligosaccharides are polymeric in nature and are structurally similar (both are polyacetals having oxygen bridges linking the monosaccharide monomers), but they may differ markedly in degree

TABLE II
Oligosaccharides[a]

Reducing		Nonreducing	
Homo-oligosaccharides	Hetero-oligosaccharides	Homo-oligosaccharides	Hetero-oligosaccharides
Simple oligosaccharides			
I. Disaccharides			
Maltose	Lactose	Trehalose	Sucrose
4-α-D-Glc*p*-D-Glc	4-β-D-Gal*p*-D-Glc	α-D-Glc*p*-α-D-Glc*p*	β-D-Fru*f*-α-D-Glc*p*
Cellobiose	Lactulose	Isotrehalose	Isosucrose
4-β-D-Glc*p*-D-Glc	4-β-D-Gal*p*-D-Fru	β-D-Glc*p*-β-D-Glc*p*	α-D-Fru*f*-β-D-Glc*p*
Isomaltose	Melibose		
6-α-D-Glc*p*-D-Glc	6-α-D-Gal*p*-D-Glc		
Gentiobiose	Turanose		
6-β-D-Glc*p*-D-Glc	3-α-D-Glc*p*-D-Fru		
II. Trisaccharides			
Maltotriose	Manninotriose		Raffinose
4-α-D-Glc*p*-maltose	6-α-D-Gal*p*-melibiose		6-α-D-Gal*p*-sucrose
III. Tetrasaccharides			
Maltotetraose			Stachyose
4-α-D-Glc*p*-maltotriose			6-α-D-Gal*p*-raffinose
IV. Pentasaccharides			
V. Hexasaccharides			
VI. Heptasaccharides			
VII. Octasaccharides			
VIII. Nonasaccharides			
IX. Decasaccharides			
Conjugate oligosaccharides			
			Glycolipids

[a] Glc, Glucose; Gal, galactose; Fru, fructose; *f*, furanose; *p*, pyranose.

of polymerization; the polysaccharides may reach a DP of 10^5, whereas, by definition, the maximum DP for oligosaccharides is 10. Although, by convention, compounds having a degree of polymerization of 11 or more are designated polysaccharides, the differences between the properties of lower polysaccharides with a DP of 11 and those of higher (DP = 10) oligosaccharides can hardly be detected. However, most polysaccharides have a much higher degree of polymerization than oligosaccharides, which renders quite significant the sum of the gradual changes that occur in their physical properties with increasing DP. For example, because the solubility decreases and the viscosity increases with a rise in DP, some higher polysaccharides, such as cellulose, are completely insoluble in water (all oligosaccharides are soluble), while the increase in viscosity may cause the solutions of other polysaccharides to set and gel.

There are several ways in which to classify polysaccharides. A common one is to group them according to their sources, that is, into plant and animal polysaccharides, and then subdivide the former into skeletal polysaccharides (cellulose, etc.), reserve polysaccharides (starch, etc.), gums and mucilages, algal polysaccharides, bacterial polysaccharides, and so forth. The disadvantage of this classification is that it tells us very little about the chemistry of these polymers.

The classification used in chemistry texts distinguishes between (a) simple (true) polysaccharides, which afford on depolymerization only mono- and oligosaccharides or their derivatives (esters or ethers), and (b) conjugate polymers made up of a polysaccharide linked to another polymer, such as a peptide or a protein (to form a glycopeptide or glycoprotein). Polysaccharides are, in turn, grouped into two major classes: (i) homopolysaccharides, which are simple polymers having as a repeating unit (monomer) one type of monosaccharide, and (ii) heteropolysaccharides, which are made up of more than one type of monosaccharide. Because the shape of polymers significantly influences their physical properties, each of these types of polymer is further divided into linear and branched polysaccharides. Table III shows the classes to which some of the common polysaccharides belong.

D. DNA, RNA, Nucleotides, and Nucleosides

Unlike oligo- and polysaccharides, which are polyacetals linked by oxygen bridges, DNA and RNA are polyesters linked by phosphate bridges. DNA is the largest known polymer; its DP exceeds 10^{12} in human genes and decreases as the ladder of evolution is descended. This giant molecule plays a key role in replication and in transcription. The latter is achieved by doubling one of the DNA strands with a smaller polymer,

TABLE III
Polysaccharides

Homopolysaccharides		Heteropolysaccharides	
Linear	Branched	Linear	Branched
Simple polysaccharides			
Amylose (α-D-glucan)	Amylopectin	Mannans	Gums
Cellulose (β-D-glucan)	Glycogen	Xylans	Mucilages
Chitin (D-glucosaminan)			Pectins
			Algin
			Agar
			Bacterial polysaccharides
Conjugate polysaccharides			
			Peptidoglycans
			Glycoproteins
			Lectins

mRNA, which in turn binds with a string of oligomers, tRNA, to form the peptide chain. The monomers of DNA and RNA are made up of phosphorylated 2-deoxy-D-*erythro*-pentofuranosyl- and D-ribofuranosyl-purine and -pyrimidine bases, designated nucleotides. The latter can undergo hydrolysis of their phosphoric ester groups to afford simpler monomers, the nucleosides. Thus, it is apparent that this group may also be divided, according to the degree of polymerization, into monomers (nucleosides and nucleotides), oligomers (tRNA), and polymers (DNA and mRNA).

E. Other Saccharide Derivatives

Grouped under this broad heading are a number of important carbohydrate derivatives which are best studied together, rather than being artificially divided to conform to a certain classification. For example, carbohydrate-containing antibiotics constitute a group of complex organic molecules of therapeutic importance. Some of their members are monosaccharide derivatives and, if grouped under this class, would be separated from closely related oligosaccharide derivatives.

I

Monosaccharides