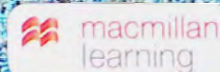


LEHNINGER

Principles of BIOCHEMISTRY

Eighth Edition

David L. Nelson • Michael M. Cox



PGR2026-00021

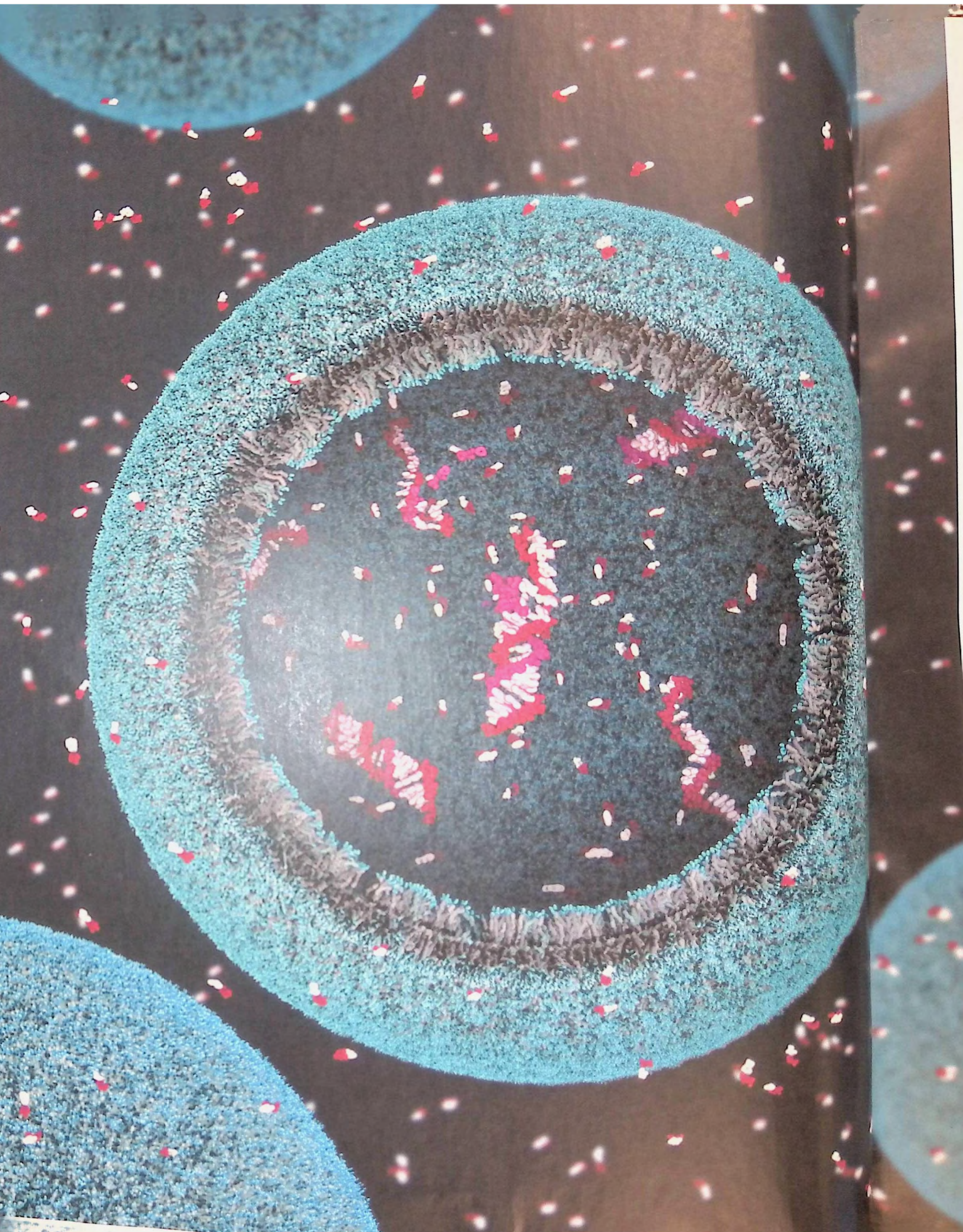
Sun Pharmaceutical Industries v. Biofrontera Bioscience GMBH

Sun's Exhibit 1013

EX1013, Page 1

Lehninger

Principles of Biochemistry



Lehninger Principles of Biochemistry

EIGHTH EDITION

David L. Nelson

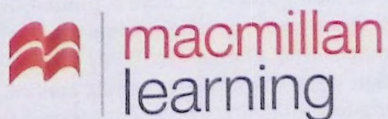
*Professor Emeritus of Biochemistry
University of Wisconsin–Madison*

Michael M. Cox

*Professor of Biochemistry
University of Wisconsin–Madison*

Aaron A. Hoskins

*Associate Professor of Biochemistry
University of Wisconsin–Madison*



Austin • Boston • New York • Plymouth

ANTIBODIES

Senior Vice President, STEM: Daryl Fox
Executive Program Director: Sandra Lindelof
Program Manager, Biochemistry: Elizabeth Simmons
Senior Marketing Manager: Maureen Rachford
Executive Content Development Manager, STEM: Debbie Hardin
Development Editor: Catherine Murphy
Executive Project Manager, Content, STEM: Katrina Mangold
Editorial Project Manager: Karen Misler
Director of Content, Life and Earth Sciences: Jennifer Driscoll Hollis
Executive Media Editor: Amy Thorne
Media Editors: Cassandra Korsvik, Kelsey Hughes
Editorial Assistant: Nathan Livingston
Marketing Assistant: Morgan Psiuk
Director of Content Management Enhancement: Tracey Kuehn
Senior Managing Editor: Lisa Kinne
Senior Content Project Manager: Vivien Weiss
Senior Workflow Project Manager: Paul W. Rohloff
Production Supervisor: Robert Cherry
Director of Design, Content Management: Diana Blume
Design Services Manager: Natasha Wolfe
Cover Designer: John Callahan
Text Designer: Maureen McCutcheon
Art Managers: Janice Donnola, Matthew McAdams
Illustrations: Emiko Paul, H. Adam Steinberg
Director of Digital Production: Keri deManigold
Media Project Manager: Brian Nobile
Permissions Manager: Michael McCarty
Media Permissions Manager: Christine Buese
Photo Researcher: Jennifer Atkins
Composition: Lumina Datamatics, Inc.
Printing and Binding: LSC Communications
Cover Image, Title Page, and Part Openers: Janet Iwasa, University of Utah

Library of Congress Control Number: 2020942138

ISBN-13: 978-1-319-22800-2

ISBN-10: 1-319-22800-3

© 2021, 2017, 2013, 2008 by W. H. Freeman and Company

All rights reserved.

Printed in the United States of America

1 2 3 4 5 6 25 24 23 22 21 20

Macmillan Learning
One New York Plaza
Suite 4600
New York, NY 10004-1562
www.macmillanlearning.com



w. h. freeman
Macmillan Learning

In 1946, William Freeman founded W. H. Freeman and Company and published Linus Pauling's *General Chemistry*, which revolutionized the chemistry curriculum and established the prototype for a Freeman text. W. H. Freeman quickly became a publishing house where leading researchers can make significant contributions to mathematics and science. In 1996, W. H. Freeman joined Macmillan and we have since proudly continued the legacy of providing revolutionary, quality educational tools for teaching and learning in STEM.

To Our Teachers

Paul R. Burton

Albert Finholt

Jeff Gelles

William P. Jencks

Eugene P. Kennedy

Homer Knoss

Arthur Kornberg

I. Robert Lehman

Andy LiWang

Patti LiWang

Melissa J. Moore

Douglas A. Nelson

Wesley A. Pearson

David E. Sheppard

JoAnne Stubbe

Harold B. White

CARBOHYDRATES AND GLYCOBIOLOGY

- 7.1 Monosaccharides and Disaccharides 230
- 7.2 Polysaccharides 241
- 7.3 Glycoconjugates: Proteoglycans, Glycoproteins, and Glycolipids 247
- 7.4 Carbohydrates as Informational Molecules: The Sugar Code 254
- 7.5 Working with Carbohydrates 258

Carbohydrates are the most abundant biomolecules on Earth. Each year, photosynthesis converts more than 100 billion metric tons of CO_2 and H_2O into cellulose and other plant products. The carbohydrates in these plant products are a dietary staple in most parts of the world, and the oxidation of carbohydrates is the central energy-yielding pathway in most nonphotosynthetic cells. Carbohydrate polymers called glycans serve as structural and protective elements in the cell walls of bacteria, fungi, and plants, and in the connective tissues of animals. Other carbohydrate polymers lubricate skeletal joints and participate in cell-cell recognition and adhesion. In addition, some complex carbohydrate polymers covalently attached to proteins or lipids act as signals that determine the intracellular destination or metabolic fate of these hybrid molecules, called glycoconjugates.

Carbohydrates are aldehydes or ketones with at least two hydroxyl groups, or substances that yield such compounds on hydrolysis. Many, but not all, carbohydrates have the empirical formula $(\text{CH}_2\text{O})_n$; some also contain nitrogen, phosphorus, or sulfur. There are three major size classes of carbohydrates: monosaccharides, oligosaccharides, and polysaccharides (the word “saccharide” is derived from the Greek *sakcharon*, meaning “sugar”). **Monosaccharides**, or simple sugars, consist of a single polyhydroxy aldehyde or ketone unit. The most abundant monosaccharide in nature is the six-carbon sugar D-glucose, sometimes referred to as dextrose.

Oligosaccharides consist of short chains of monosaccharide units, or residues, joined by characteristic

linkages called glycosidic bonds. The most abundant are the **disaccharides**, with two monosaccharide units. Sucrose (table sugar), for example, consists of the six-carbon sugars D-glucose and D-fructose. All common monosaccharides and disaccharides have names ending with the suffix “-ose.” In cells, most oligosaccharides consisting of three or more units do not occur as free entities but are joined to nonsugar molecules (lipids or proteins) in glycoconjugates.

The **polysaccharides** are sugar polymers containing more than 10 monosaccharide units; some have hundreds or thousands of units. Some polysaccharides, such as cellulose, are linear chains; others, such as glycogen, are branched. Both cellulose and glycogen consist of recurring units of D-glucose, but they differ in the type of glycosidic linkage and consequently have strikingly different properties and biological roles.

This chapter introduces the major classes of carbohydrates and glycoconjugates and provides examples of their many structural and functional roles. Learning the structures and chemical properties of biomolecules is essential, because they are the vocabulary and the grammar of biochemistry. As you read about carbohydrates, note how specific cases illustrate these general principles that underlie all of biochemistry.

- P1** Carbohydrates can have multiple chiral carbons; the configuration of groups around each carbon atom determines how the compound interacts with other biomolecules. As we saw for L-amino acids in proteins, with rare exceptions, biological evolution selected one stereochemical series (D-series) for sugars.
- P2** Monomeric subunits, monosaccharides, serve as the building blocks of large carbohydrate polymers. The specific sugar, the way the units are linked, and whether the polymer is branched determine its properties and thus its function.
- P3** Storage of low molecular weight metabolites in polymeric form avoids the very high osmolarity that would result from storing them as individual monomers. If the glucose in liver glycogen were monomeric, the glucose

concentration in liver would be so high that cells would swell and lyse from the entry of water by osmosis.

- P4** The sequences of complex polysaccharides are determined by the intrinsic properties of the biosynthetic enzymes that add each monomeric unit to the growing polymer. This is in contrast with DNA, RNA, and proteins, which are synthesized on templates that direct their sequence.
- P5** Polysaccharides assume three-dimensional structures with the lowest-energy conformations, determined by covalent bonds, hydrogen bonds, charge interactions, and steric factors. Starch folds into a helical structure stabilized by internal hydrogen bonds; cellulose assumes an extended structure in which intermolecular hydrogen bonds are more important.
- P6** Molecular complementarity is central to function. The recognition of oligosaccharides by sugar-binding proteins (lectins) results from a perfect fit between lectin and ligand.
- P7** An almost infinite variety of discrete structures can be built from a small number of monomeric subunits. Even short polymers, when arranged in different sequences, joined through different linkages, and branched to specific degrees, present unique faces recognized by their molecular partners.

7.1 Monosaccharides and Disaccharides

The simplest of the carbohydrates, the monosaccharides, are either aldehydes or ketones with two or more hydroxyl groups; the six-carbon monosaccharides glucose and fructose have five hydroxyl groups. Many of the carbon atoms to which the hydroxyl groups are attached are chiral centers, which give rise to the many sugar stereoisomers found in nature. Stereoisomerism in sugars is biologically significant because the enzymes that act on sugars are strictly

stereospecific, typically preferring one stereoisomer to another by three or more orders of magnitude, as reflected in K_m values or binding constants. It is as difficult to fit the wrong sugar stereoisomer into an enzyme's binding site as it is to put your left glove on your right hand.

We begin by describing the families of monosaccharides with backbones of three to seven carbons—their structure, their stereoisomeric forms, and the means of representing their three-dimensional structures on paper. We then discuss several chemical reactions of the carbonyl groups of monosaccharides. One such reaction, the addition of a hydroxyl group from within the same molecule, generates cyclic forms with four or more backbone carbons (the forms that predominate in aqueous solution).

P1 This ring closure creates a new chiral center, adding further stereochemical complexity to this class of compounds. The nomenclature about each carbon atom in a cyclic form and the means of representing these structures on paper are described in some detail; this information will be useful as we discuss the metabolism of monosaccharides in Part II. We also introduce here some important monosaccharide derivatives that will be examined closely in later chapters.

The Two Families of Monosaccharides Are Aldoses and Ketoses

Monosaccharides are colorless, crystalline solids that are freely soluble in water but insoluble in nonpolar solvents. Most have a sweet taste (**Box 7-1**). The backbones of common monosaccharides are unbranched carbon chains in which all the carbon atoms are linked by single bonds. In this open-chain form, one of the carbon atoms is double-bonded to an oxygen atom to form a carbonyl group; each of the other carbon atoms has a hydroxyl group. If the carbonyl group is at an end of the carbon chain (that is, in an aldehyde group), the monosaccharide is an **aldose**; if the carbonyl group is at any other position (in a ketone group), the monosaccharide is a **ketose**. The simplest monosaccharides are the two three-carbon trioses: glyceraldehyde, an aldotriose, and dihydroxyacetone, a ketotriose (**Fig. 7-1a**).

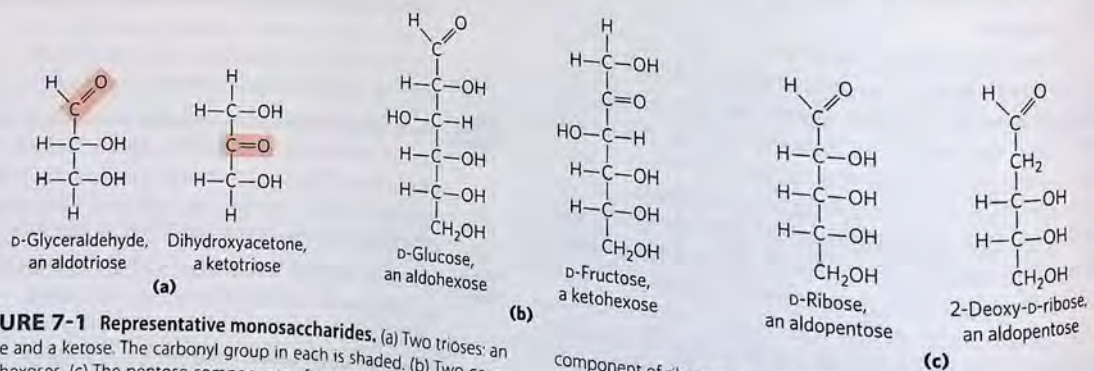


FIGURE 7-1 Representative monosaccharides. (a) Two trioses: an aldose and a ketose. The carbonyl group in each is shaded. (b) Two common hexoses. (c) The pentose components of nucleic acids. D-Ribose is a

component of ribonucleic acid (RNA), and 2-deoxy-D-ribose is a component of deoxyribonucleic acid (DNA).



What Makes Sugar Sweet?

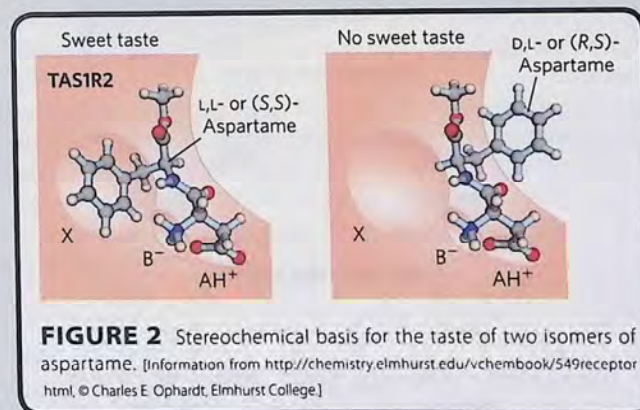
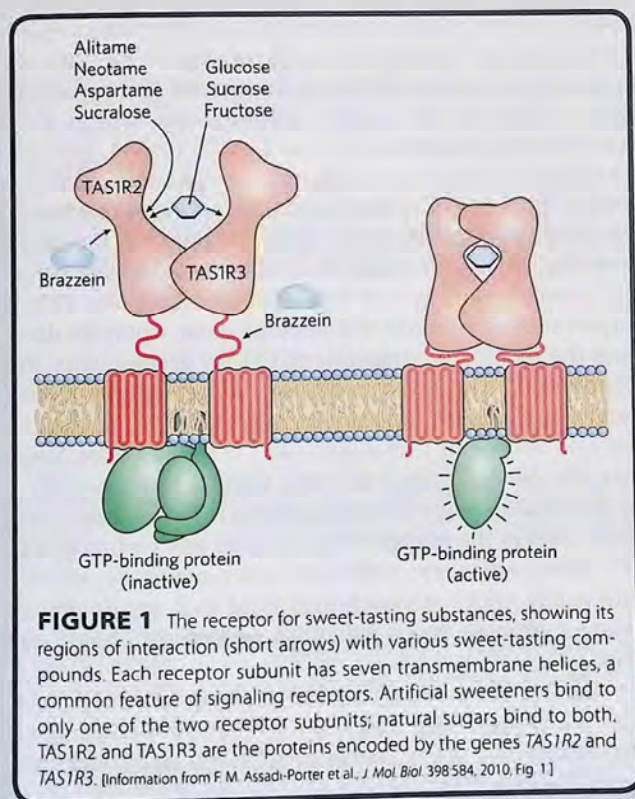
Sweetness is one of the five basic flavors that humans can taste; the others are sour, bitter, salty, and umami. Sweetness is detected by protein receptors in the plasma membranes of cells in the taste buds on the surface of the tongue. In humans, two closely related genes (*TAS1R2* and *TAS1R3*) encode sweet-taste receptors (Fig. 1). When a molecule with a compatible structure binds a sweet-taste receptor's extracellular domain, it triggers a series of events in the cell (including activation of a GTP-binding protein; see Fig. 12-20) that generates an electrical signal to the brain that is interpreted as "sweet."

During evolution, there has probably been selection for the ability to taste compounds found in foods containing important nutrients, such as the carbohydrates that are major fuels for most organisms. Most sugars, including sucrose, glucose, lactose, and fructose, taste sweet, but other classes of compounds also bind sweet-taste receptors. The amino acids glycine, alanine, and serine are mildly sweet and harmless; nitrobenzene and ethylene glycol have a strong sweet taste

but are toxic. (See Box 18-2 for a remarkable medical mystery involving ethylene glycol poisoning.) Several natural products are extraordinarily sweet. Stevioside, a sugar derivative isolated from the leaves of the stevia plant (*Stevia rebaudiana* Bertoni), is several hundred times sweeter than an equivalent amount of sucrose. The small protein brazzein (54 amino acids), isolated from berries of the Oubli vine (*Pentadiplandra brazzeana* Baillon) in Gabon and Cameroon, is 17,000 times sweeter than sucrose on a molar basis. Presumably, the sweetness of the berries encourages their consumption by animals that then disperse the seeds so that new plants are established.

In societies where obesity is a major health problem, compounds that give foods a sweet taste without adding the calories found in sugars are common food additives. The artificial sweetener aspartame demonstrates the importance of stereochemistry in biology (Fig. 2). One simple model of binding to the sweet-taste receptor involves three sites: AH^+ , B^- , and X. Site AH^+ has a group (an alcohol or an amine) that binds the partially negative oxygen of the carboxylic acid of (*S,S*)-aspartame. Site B^- has a partially negative oxygen available to hydrogen-bond with the amine nitrogen of (*S,S*)-aspartame. Site X is oriented perpendicular to the other two groups and can accommodate the hydrophobic benzene ring of (*S,S*)-aspartame.

When the steric match is correct, as for (*S,S*)-aspartame, on the left in Figure 2, the sweet receptor is stimulated and the signal "sweet" is conducted to the brain. When the match is not correct, as for (*R,S*)-aspartame, the sweet receptor is not stimulated; in fact, (*R,S*)-aspartame stimulates a separate receptor specifically for bitterness. Stereoisomerism really matters!



Monosaccharides with four, five, six, and seven carbon atoms in their backbones are called, respectively, tetroses, pentoses, hexoses, and heptoses. There are aldoses and ketoses of each of these chain lengths: aldotetroses and ketotetroses, aldopentoses and ketopentoses, and so on. The hexoses, which include the aldohexose D-glucose

and the ketohexose D-fructose (Fig. 7-1b), are the most common monosaccharides in nature — the products of photosynthesis and key intermediates in the central energy-yielding reaction sequence in most organisms. The aldopentoses D-ribose and 2-deoxy-D-ribose (Fig. 7-1c) are components of nucleotides and nucleic acids (Chapter 8).

Monosaccharides Have Asymmetric Centers

P1 All the monosaccharides except dihydroxyacetone contain one or more asymmetric (chiral) carbon atoms and thus occur in optically active isomeric forms (pp. 16–17). The simplest aldose, glyceraldehyde, contains one chiral center (the middle carbon atom) and therefore has two different optical isomers, or **enantiomers** (Fig. 7-2).

KEY CONVENTION One of the two enantiomers of glyceraldehyde is, by convention, designated the *D* isomer; the other is the *L* isomer. As for other biomolecules with chiral centers, the absolute configurations of sugars are known from x-ray crystallography. To represent three-dimensional sugar structures on paper, we often use **Fischer projection formulas** (Fig. 7-2). In these projections, bonds drawn horizontally indicate bonds that project out of the plane of the paper, toward the reader; bonds drawn vertically project behind the plane of the paper, away from the reader. ■

In general, a molecule with n chiral centers can have 2^n stereoisomers. Glyceraldehyde has $2^1 = 2$; the aldohexoses, with four chiral centers, have $2^4 = 16$. The

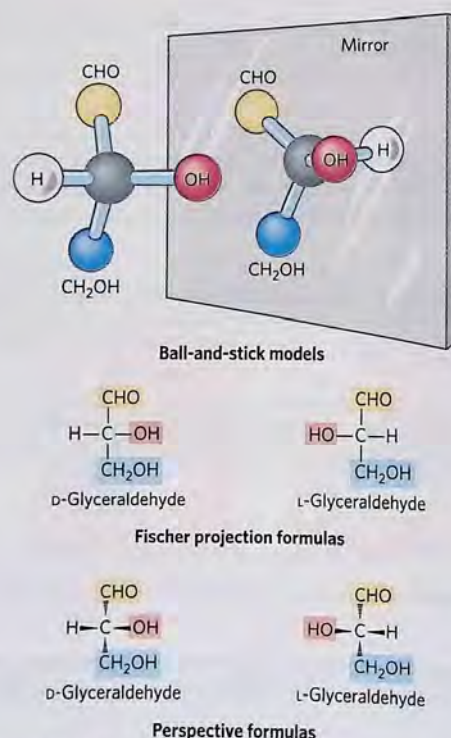
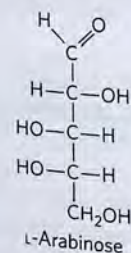


FIGURE 7-2 Three ways to represent the two enantiomers of glyceraldehyde. The enantiomers are mirror images of each other. Ball-and-stick models show the actual configuration of molecules. In Fischer projections, vertical lines point behind the plane of the page, and horizontal lines project above the page. Recall (see Fig. 1-17) that in perspective formulas, the wide end of a solid wedge projects out of the plane of the paper, toward the reader; a dashed wedge extends behind.

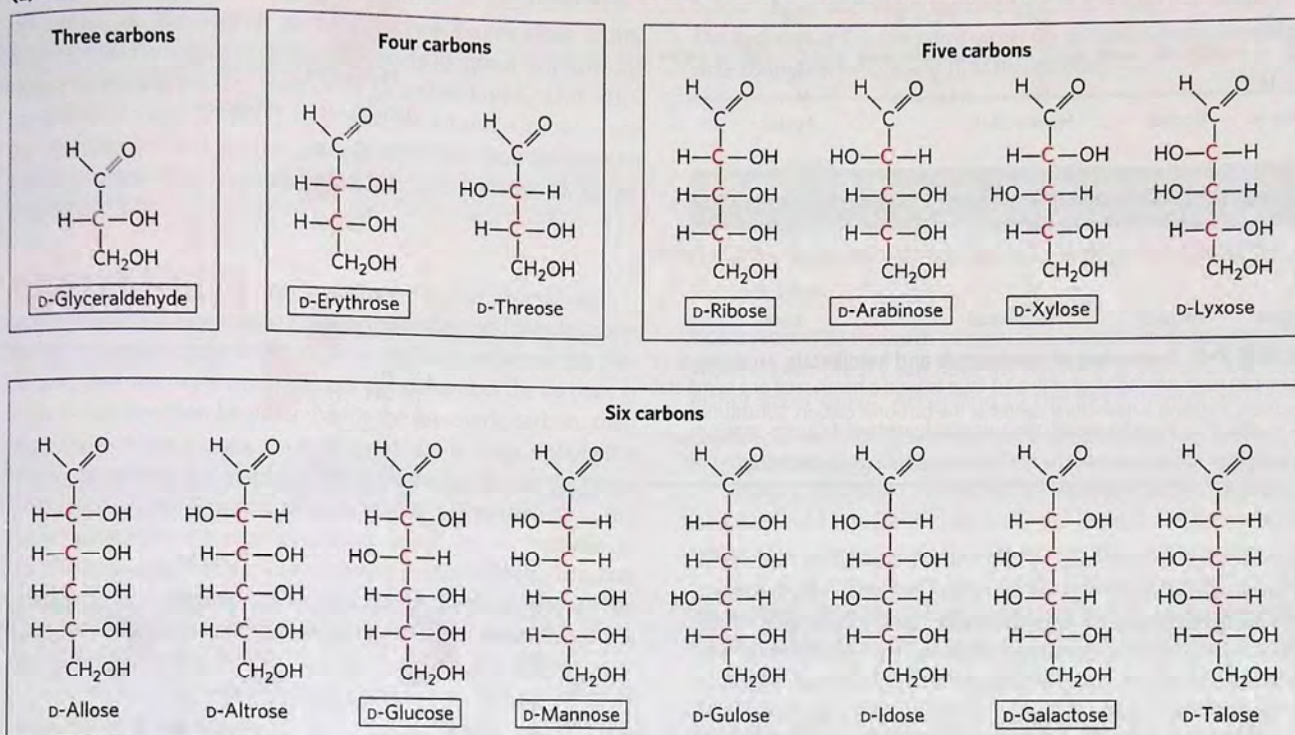
stereoisomers of monosaccharides of each carbon-chain length can be divided into two groups that differ in the configuration about the chiral center *most distant* from the carbonyl carbon. Those in which the configuration at this reference carbon is the same as that of *D*-glyceraldehyde are designated *D* isomers, and those with the same configuration as *L*-glyceraldehyde are *L* isomers. In other words, when the hydroxyl group on the reference carbon is on the right (*dextro*) in a projection formula that has the carbonyl carbon at the top, the sugar is the *D* isomer; when the hydroxyl group is on the left (*levo*), it is the *L* isomer. Of the 16 possible aldohexoses, eight are *D* forms and eight are *L*. **P1** Most of the hexoses of living organisms are *D* isomers. Why *D* isomers? An interesting and unanswered question. Recall that all of the amino acids found in proteins are exclusively one of two possible stereoisomers, *L* (p. 73). The basis for this initial preference for one isomer during evolution is unknown; however, once one isomer became prevalent, evolving enzymes able to use that isomer efficiently would have a selective advantage.

Figure 7-3 shows the structures of the *D* stereoisomers of all the aldoses and ketoses having three to six carbon atoms. The carbons of a sugar are numbered beginning at the end of the chain nearest the carbonyl group. Each of the eight *D*-aldohexoses, which differ in the stereochemistry at C-2, C-3, or C-4, has its own name: *D*-glucose, *D*-galactose, *D*-mannose, and so forth (Fig. 7-3a). The four- and five-carbon ketoses are designated by inserting “ul” into the name of a corresponding aldose; for example, *D*-ribulose is the ketopentose corresponding to the aldopentose *D*-ribose. (The importance of ribulose will become clear when we discuss the fixation of atmospheric CO_2 by green plants, in Chapter 20.) The ketohexoses are named otherwise: for example, fructose (from the Latin *fructus*, “fruit”; fruits are one source of this sugar) and sorbose (from *Sorbus*, the genus of mountain ash, which has berries rich in the related sugar alcohol sorbitol). Two sugars that differ only in the configuration around one carbon atom are called **epimers**; *D*-glucose and *D*-mannose, which differ only in the stereochemistry at C-2, are epimers, as are *D*-glucose and *D*-galactose (which differ at C-4) (Fig. 7-4).

Some sugars occur naturally in their *L* form; examples are *L*-arabinose and the *L* isomers of some sugar derivatives that are common components of glycoconjugates (Section 7.3).



(a) D-Aldoses



(b) D-Ketoses

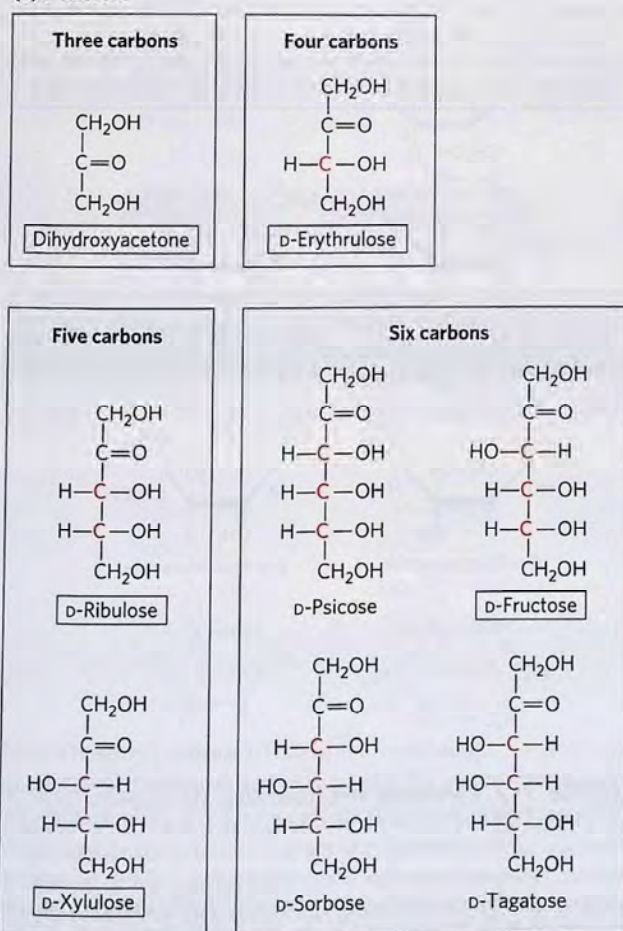


FIGURE 7-3 Aldoses and ketoses. The series of (a) D-aldoses and (b) D-ketoses having from three to six carbon atoms, shown as projection formulas. The carbon atoms in red are chiral centers. In all of these D isomers, the chiral carbon *most distant from the carbonyl carbon* has the same configuration as the chiral carbon in D-glyceraldehyde. Shown are the most common sugars in nature; you will encounter these again in this and later chapters.

The Common Monosaccharides Have Cyclic Structures

For simplicity, we have thus far represented the structures of aldoses and ketoses as straight-chain molecules (Figs 7-3, 7-4). In fact, in aqueous solution, aldotetroses and all monosaccharides with five or more carbon atoms in the backbone occur predominantly as cyclic (ring) structures in which the carbonyl group has formed a covalent bond with the oxygen of a hydroxyl group in the same sugar molecule. The formation of these ring

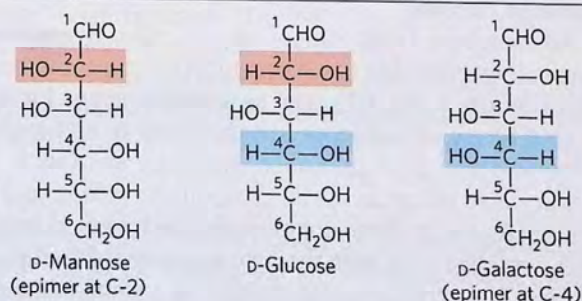


FIGURE 7-4 Epimers. D-Glucose and two of its epimers are shown as projection formulas. Each epimer differs from D-glucose in the configuration at one chiral center (shaded light red or blue).

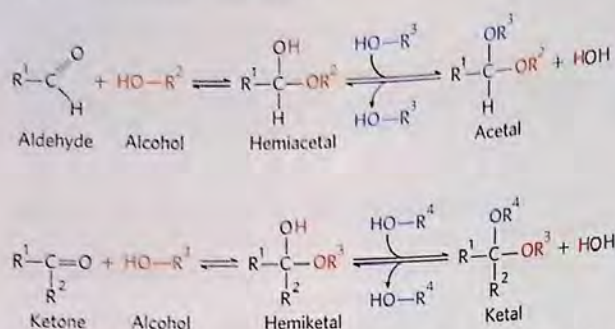


FIGURE 7-5 Formation of hemiacetals and hemiketals. An aldehyde or a ketone can react with an alcohol in a 1:1 ratio to yield a hemiacetal or a hemiketal, respectively, creating a new chiral center at the carbonyl carbon. Substitution of a second alcohol molecule produces an acetal or a ketal. When the second alcohol is part of another sugar molecule, the bond produced is a glycosidic bond.

structures is the result of a general reaction between alcohols and aldehydes or ketones to form derivatives called **hemiacetals** or **hemiketals**. Two molecules of an alcohol can add to a carbonyl carbon; the product of the first addition is a hemiacetal (for addition to an aldose) or a hemiketal (for addition to a ketose). If the —OH and carbonyl groups are on the same molecule, a five- or six-membered ring results. Addition of the second molecule of alcohol produces the full **acetal** or **ketal** (Fig. 7-5), and the linkage formed is a glycosidic bond. When the two molecules that react are monosaccharides, the acetal or ketal formed is a disaccharide.

The reaction with the first molecule of alcohol creates an additional chiral center (the carbonyl carbon). Because the alcohol can add in either of two ways, attacking either the “front” or the “back” of the carbonyl carbon, the reaction can produce either of two stereoisomeric configurations, denoted α and β . For example, D-glucose exists in solution as an intramolecular hemiacetal in which the free hydroxyl group at C-5 has reacted with the aldehydic C-1, rendering the latter carbon asymmetric and producing two possible stereoisomers, designated α and β (Fig. 7-6). Isomeric forms of monosaccharides that differ only in their configuration about the hemiacetal or hemiketal carbon atom are called **anomers**, and the carbonyl carbon atom is called the **anomeric carbon**.

Six-membered ring compounds are called **pyranoses** because they resemble the six-membered ring compound pyran (shown in Fig. 7-7). The systematic names for the two ring forms of D-glucose are therefore α -D-glucopyranose and β -D-glucopyranose. Ketohexoses (such as fructose) also occur as cyclic compounds with α and β anomeric forms. In these compounds, the hydroxyl group at C-5 (or C-6) reacts with the keto group at C-2 to form a **furanose** (or pyranose) ring containing a hemiketal linkage (Fig. 7-7). D-Fructose readily forms the furanose ring (Fig. 7-7); the more common anomer of this sugar in combined forms or in derivatives is β -D-fructofuranose.

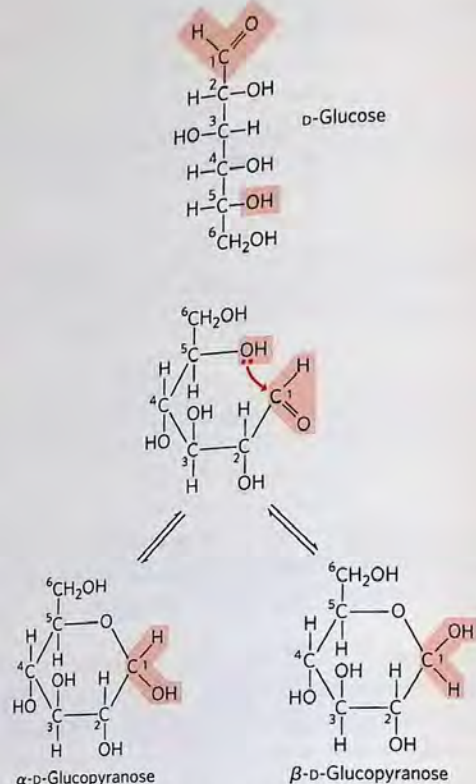


FIGURE 7-6 Formation of the two cyclic forms of D-glucose. Reaction between the aldehyde group at C-1 and the hydroxyl group at C-5 forms a hemiacetal linkage, producing either of two stereoisomers, the α and β anomers, which differ only in the stereochemistry around the hemiacetal carbon. This reaction is reversible. The interconversion of α and β anomers is called mutarotation.

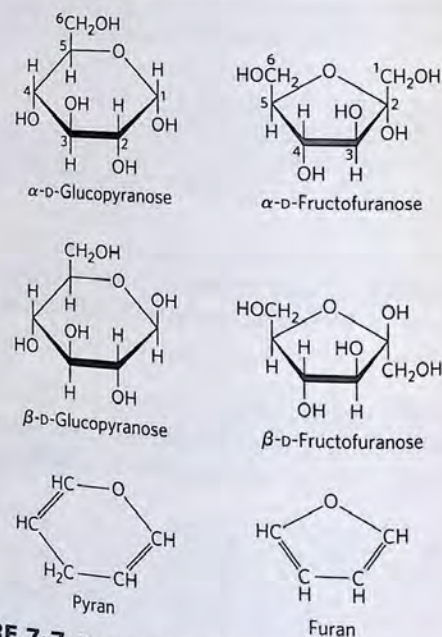
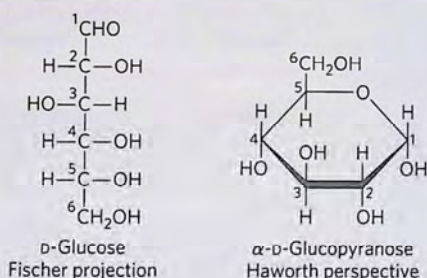


FIGURE 7-7 Pyranoses and furanoses. The pyranose forms of D-glucose and the furanose forms of D-fructose are shown here as Haworth perspective formulas. The edges of the ring nearest the reader are represented by bold lines. Hydroxyl groups below the plane of the ring in these Haworth perspectives would appear at the right side of a Fischer projection (compare with Fig. 7-6). Pyran and furan are shown for comparison.

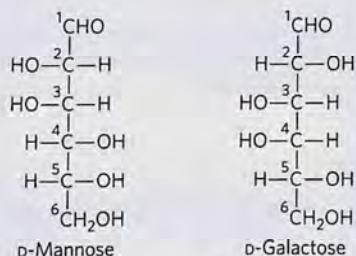
Cyclic sugar structures are more accurately represented in **Haworth perspective formulas** than in the Fischer projections commonly used for linear sugar structures. In Haworth perspectives, the six-membered ring is tilted to make its plane almost perpendicular to that of the paper, with the bonds closest to the reader drawn thicker than those farther away, as in Figure 7-7.

KEY CONVENTION To convert the Fischer projection formula of any linear D-hexose to a Haworth perspective formula showing the molecule's cyclic structure, draw the six-membered ring (five carbons, and one oxygen at the upper right), number the carbons in a clockwise direction beginning with the anomeric carbon, then place the hydroxyl groups. If a hydroxyl group is to the right in the Fischer projection, it is placed pointing down (i.e., below the plane of the ring) in the Haworth perspective; if it is to the left in the Fischer projection, it is placed pointing up (i.e., above the plane) in the Haworth perspective. The terminal $\text{—CH}_2\text{OH}$ group projects upward for the D enantiomer, downward for the L enantiomer. The hydroxyl on the anomeric carbon can point up or down. When the anomeric hydroxyl of a D-hexose is on the same side of the ring as C-6, the structure is by definition β ; when it is on the opposite side from C-6, the structure is α .



WORKED EXAMPLE 7-1 Conversion of Fischer Projection to Haworth Perspective Formulas

Draw the Haworth perspective formulas for D-mannose and D-galactose.



SOLUTION: Pyranoses are six-membered rings, so start with six-membered Haworth structures with the oxygen atom at the top right. Number the carbon atoms clockwise, starting with the aldose carbon. For mannose, place the hydroxyls on C-2, C-3, and C-4 above, above, and below the ring, respectively (because in the Fischer projection they are on the left, left, and right sides of the mannose backbone). For D-galactose, the hydroxyls are oriented

below, above, and above the ring for C-2, C-3, and C-4, respectively. The hydroxyl at C-1 can point either up or down; there are two possible configurations, α and β , at this carbon.

WORKED EXAMPLE 7-2 Drawing Haworth Perspective Formulas of Sugar Isomers

Draw the Haworth perspective formulas for α -D-mannose and β -L-galactose.

SOLUTION: The Haworth perspective formula of D-mannose from Worked Example 7-1 can have the hydroxyl group at C-1 pointing either up or down. According to the Key Convention, for the α form, the C-1 hydroxyl is pointing down when C-6 is up, as it is in D-mannose.

For β -L-galactose, use the Fischer representation of D-galactose (see Worked Example 7-1) to draw the correct Fischer representation of L-galactose, which is its mirror image: the hydroxyls at C-2, C-3, C-4, and C-5 are on the left, right, right, and left sides, respectively. Now draw the Haworth perspective, a six-membered ring in which the —OH groups on C-2, C-3, and C-4 are oriented up, down, and down, respectively, because in the Fischer representation they are on the left, right, and right sides. Because it is the β form, the —OH on the anomeric carbon points down (same side as C-5).

The α and β anomers of D-glucose interconvert in aqueous solution by a process called **mutarotation**, in which one ring form (say, the α anomer) opens briefly into the linear form, then closes again to produce the β anomer (Fig. 7-6). Thus, a solution of β -D-glucose and a solution of α -D-glucose eventually form identical equilibrium mixtures having identical optical properties. This mixture consists of about one-third α -D-glucose, two-thirds β -D-glucose, and very small amounts of the linear form and the five-membered ring (glucofuranose) form.

Haworth perspective formulas like those in Figure 7-7 are commonly used to show the stereochemistry of ring forms of monosaccharides. However, the six-membered pyranose ring is not planar, as Haworth perspectives suggest, but tends to assume either of two "chair" conformations (Fig. 7-8). Recall from Chapter 1 that **P5** two conformations of a molecule are interconvertible without the breakage of covalent bonds, whereas two configurations can be interconverted only by breaking a covalent bond. To interconvert α and β configurations, the bond involving the ring oxygen atom has to be broken, but interconversion of the two chair forms (which are conformers) does not require bond breakage and does not change configurations at any of the ring carbons. The specific three-dimensional structures of the monosaccharide units are important in determining the biological properties and functions of some polysaccharides, as we shall see.

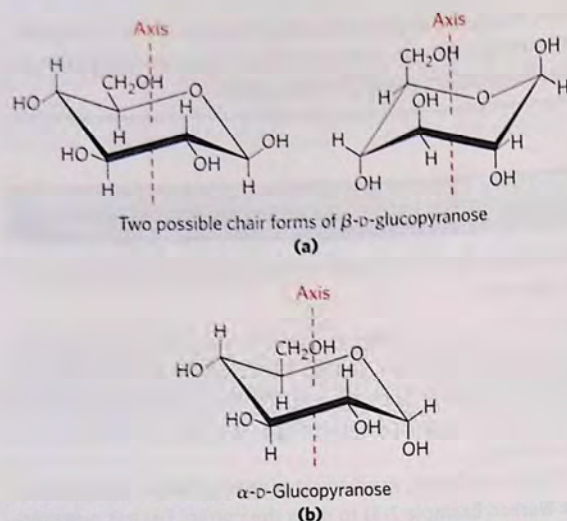
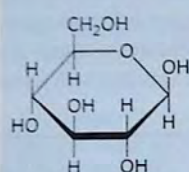


FIGURE 7-8 Conformational formulas of pyranoses. (a) Two chair forms of the pyranose ring of β -D-glucopyranose. Two conformers such as these are not readily interconvertible; an input of about 46 kJ of such energy per mole of sugar is required to force the interconversion of chair forms. Another conformation, the "boat" (not shown), is seen only in derivatives with very bulky substituents. (b) The preferred chair conformation of α -D-glucopyranose.

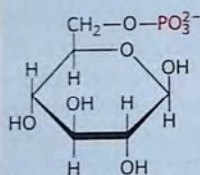
Organisms Contain a Variety of Hexose Derivatives

In addition to simple hexoses such as glucose, galactose, and mannose, there are many sugar derivatives in which a hydroxyl group in the parent compound is replaced with another substituent, or a carbon atom is replaced with another substituent, or a carbon atom is replaced with another substituent (Fig. 7-9). In glucosamine, oxidized to a carboxyl group (Fig. 7-9). In glucosamine, galactosamine, and mannosamine, the hydroxyl at C-2 of the parent compound is replaced with an amino group.

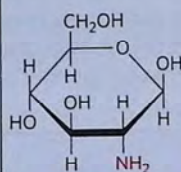
Glucose family



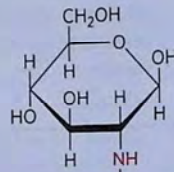
β -D-Glucose



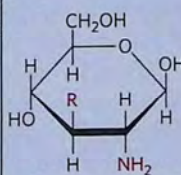
β -D-Glucose 6-phosphate



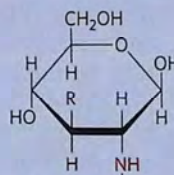
β -D-Glucosamine



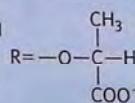
N-Acetyl- β -D-glucosamine



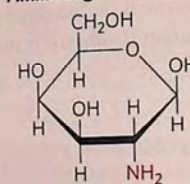
Muramic acid



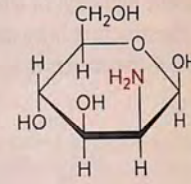
N-Acetylmuramic acid



Amino sugars

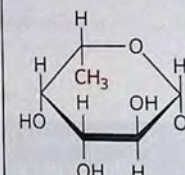


β -D-Galactosamine

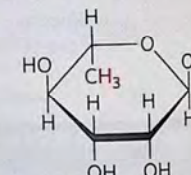


β -D-Mannosamine

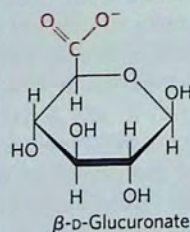
Deoxy sugars



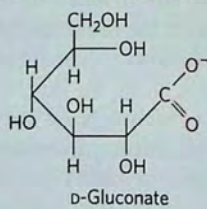
β -L-Fucose



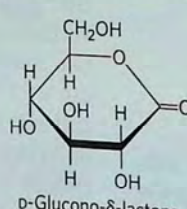
α -L-Rhamnose



β -D-Glucuronate

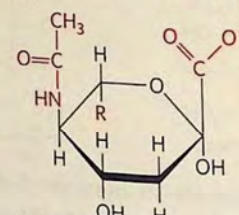


D-Gluconate



D-Glucono- δ -lactone

Acidic sugars



N-Acetylneuraminic acid (a sialic acid)

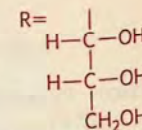


FIGURE 7-9 Some hexose derivatives important in biology.

In amino sugars, an $-\text{NH}_2$ group replaces one of the $-\text{OH}$ groups in the parent hexose. Substitution of $-\text{H}$ for $-\text{OH}$ produces a deoxy sugar, note that the deoxy sugars shown here occur in nature as the L isomers. The

acidic sugars contain a carboxylate group, which confers a negative charge at neutral pH. D-Glucono- δ -lactone results from formation of an ester linkage between the C-1 carboxylate group and the C-5 (also known as the δ carbon) hydroxyl group of D-gluconate.

The amino group is commonly condensed with acetic acid, as in *N*-acetylglucosamine. This glucosamine derivative is part of many structural polymers, including those of the bacterial cell wall. Substitution of a hydrogen for the hydroxyl group at C-6 of L-galactose or L-mannose produces L-fucose or L-rhamnose, respectively. L-Fucose is found in the complex oligosaccharide components of glycoproteins and glycolipids; L-rhamnose is found in plant polysaccharides.

Oxidation of the carbonyl (aldehyde) carbon of glucose to the carboxyl level produces gluconic acid, used in medicine as an innocuous counterion when administering positively charged drugs (such as quinine) or ions (such as Ca^{2+}). Other aldoses yield other **aldonic acids**. Oxidation of the carbon at the other end of the carbon chain — C-6 of glucose, galactose, or mannose — forms the corresponding **uronic acid**: glucuronic, galacturonic, or mannuronic acid. Both aldonic acids and uronic acids form stable intramolecular esters called lactones (Fig. 7-9, lower left). The sialic acids are a family of sugars with the same nine-carbon backbone. One of them, *N*-acetylneuraminic acid (often referred to simply as “sialic acid”), is a derivative of *N*-acetylmannosamine that occurs in many glycoproteins and glycolipids on animal cell surfaces, providing sites of recognition by other cells or extracellular carbohydrate-binding proteins. The carboxylic acid groups of the acidic sugar derivatives are ionized at pH 7, and the compounds are therefore correctly named as the carboxylates — glucuronate, galacturonate, and so forth.

In the synthesis and metabolism of carbohydrates, the intermediates are very often not the sugars themselves but their phosphorylated derivatives. Condensation of phosphoric acid with one of the hydroxyl groups of a sugar forms a phosphate ester, as in glucose 6-phosphate (Fig. 7-9), the first metabolite in the pathway by which most organisms oxidize glucose for energy. Sugar phosphates are relatively stable at neutral pH and bear a negative charge. One effect of sugar phosphorylation within cells is to trap the sugar inside the cell; most cells do not have plasma membrane transporters for phosphorylated sugars. Phosphorylation also activates sugars for subsequent chemical transformation. Several important phosphorylated derivatives of sugars are components of nucleotides (discussed in the next chapter).

Sugars That Are, or Can Form, Aldehydes Are Reducing Sugars

Free aldehyde groups in sugars undergo a characteristic redox reaction with Cu^{2+} under alkaline conditions. As the sugar is oxidized from aldehyde to carboxylic acid, Cu^{2+} is reduced to Cu^+ , which forms a brick-red precipitate. This reaction defines **reducing sugars**, which include, for example, glucose, galactose, mannose, ribose, and glyceraldehyde. The reaction occurs only with a *free* aldose; but because cyclic aldoses are in equilibrium with

their linear forms, which do have free aldehyde groups (Fig. 7-6), all aldose monosaccharides are reducing sugars. Ketoses that can rearrange (tautomerize) to form aldehydes are also reducing sugars; fructose and ribulose are reducing sugars, for example.

The metabolism of glucose in people with diabetes can be monitored by measuring urinary glucose with a simple qualitative assay for reducing sugar, or by measuring quantitatively the nonenzymatic reaction between glucose and the hemoglobin in blood (**Box 7-2**).

Disaccharides (such as maltose, lactose, and sucrose) consist of two monosaccharides joined covalently by an **O-glycosidic bond**, which is formed when a hydroxyl group of one sugar molecule, typically in its cyclic form, reacts with the anomeric carbon of the other (Fig. 7-10). This reaction represents the formation of an acetal from a hemiacetal (such as glucopyranose) and an alcohol (a hydroxyl group of the second sugar molecule) (Fig. 7-5), and the resulting compound is called a glycoside. Glycosidic bonds are readily hydrolyzed by acid but resist cleavage by base. Thus disaccharides can be hydrolyzed to yield their free monosaccharide components by boiling with dilute acid.

When the anomeric carbon is involved in a glycosidic bond, the easy interconversion of linear and cyclic forms shown in Figure 7-6 is prevented. Formation of a glycosidic bond therefore renders a sugar nonreducing. In describing disaccharides or polysaccharides, the end of a chain with a free anomeric carbon (one not involved in a glycosidic bond) is called the **reducing end**.

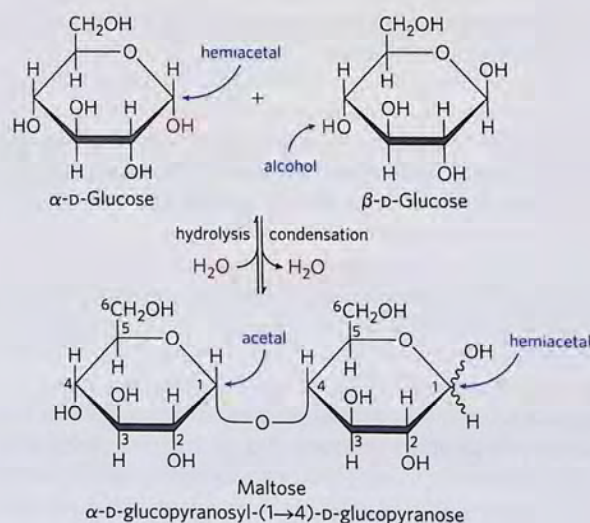


FIGURE 7-10 Formation of maltose. A disaccharide is formed from two monosaccharides (here, two molecules of D-glucose) when an —OH (alcohol) of one monosaccharide molecule (right) condenses with the intramolecular hemiacetal of the other (left), with elimination of H_2O and formation of a glycosidic bond. The reversal of this reaction is hydrolysis — attack by H_2O on the glycosidic bond. The maltose molecule, shown here, retains a reducing hemiacetal at the C-1 not involved in the glycosidic bond. Because mutarotation interconverts the α and β forms of the hemiacetal, the bonds at this position are sometimes depicted with wavy lines to indicate that the structure may be either α or β .

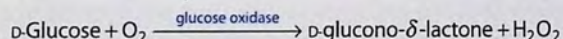


Blood Glucose Measurements in the Diagnosis and Treatment of Diabetes

Glucose is the principal fuel for the brain. When the amount of glucose reaching the brain is too low, the consequences can be dire: lethargy, coma, permanent brain damage, and death. Complex hormonal mechanisms have evolved to ensure that the concentration of glucose in the blood remains high enough (about 5 mM) to satisfy the brain's needs—but not too high, because elevated blood glucose can also have serious physiological consequences.

Individuals with insulin-dependent diabetes mellitus do not produce sufficient insulin, the hormone that normally serves to reduce blood glucose concentration. If the diabetes is untreated, blood glucose levels may rise to several-fold higher than normal. These high glucose levels are believed to be at least one cause of the serious long-term consequences of untreated diabetes—kidney failure, cardiovascular disease, blindness, and impaired wound healing—so one goal of therapy is to provide just enough insulin (by injection) to keep blood glucose levels near normal. To maintain the correct balance of exercise, diet, and insulin, individuals with diabetes must measure their blood glucose concentration several times a day, and adjust the amount of insulin injected appropriately.

The concentrations of glucose in blood can be determined by a simple assay for reducing sugar. A single drop of blood is added to a test strip containing the enzyme glucose oxidase, which catalyzes the reaction:



A second enzyme, a peroxidase, catalyzes reaction of the H_2O_2 with a colorless compound to create a colored product, which is quantified with a simple photometer that reads out the blood glucose concentration.

Because blood glucose levels change with the timing of meals and exercise, single-time measurements do not reflect the *average* blood glucose over hours and days, so dangerous increases may go undetected. The average glucose concentration can be assessed by looking at its effect on hemoglobin, the oxygen-carrying protein in erythrocytes (p. 153). Transporters in the erythrocyte membrane equilibrate intracellular and plasma glucose concentrations, so hemoglobin is constantly exposed to glucose at whatever concentration is present in the blood. A series of relatively slow *nonenzymatic* reactions (Fig. 1) occurs between glucose and primary amino groups in hemoglobin (either the amino-terminal Val or the ϵ -amino groups of Lys residues) (Fig. 2). The rate of this process is proportional to the concentration of glucose, so the reaction can be used to estimate the average blood glucose

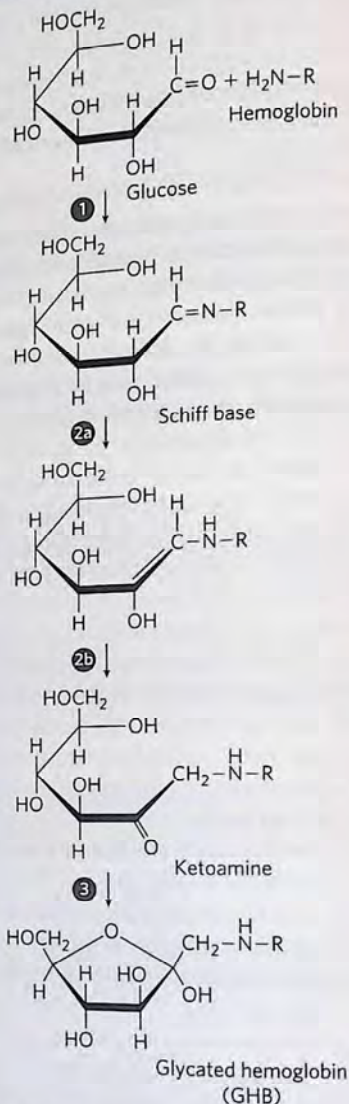
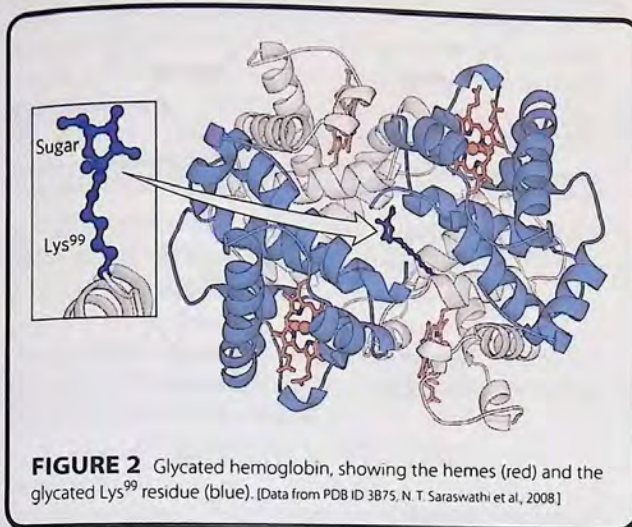


FIGURE 1 The nonenzymatic reaction of glucose with a primary amino group in hemoglobin begins with ① formation of a Schiff base, which ② undergoes a rearrangement to generate a stable product; ③ this ketoamine can further cyclize to yield GHB.

level over weeks. The amount of glycated hemoglobin (GHB) present at any time reflects the average blood glucose concentration over the circulating lifetime of the erythrocyte (about 120 days), although the concentration in the two weeks before the test is the most important in setting the level of GHB.

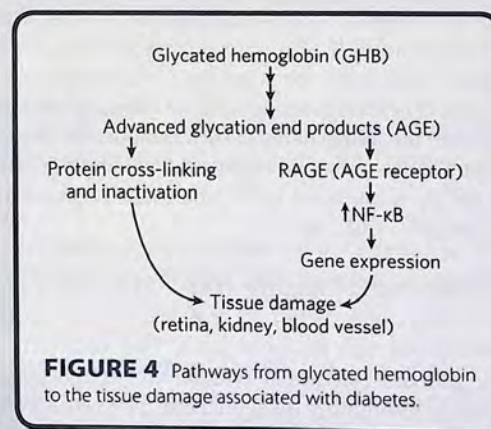
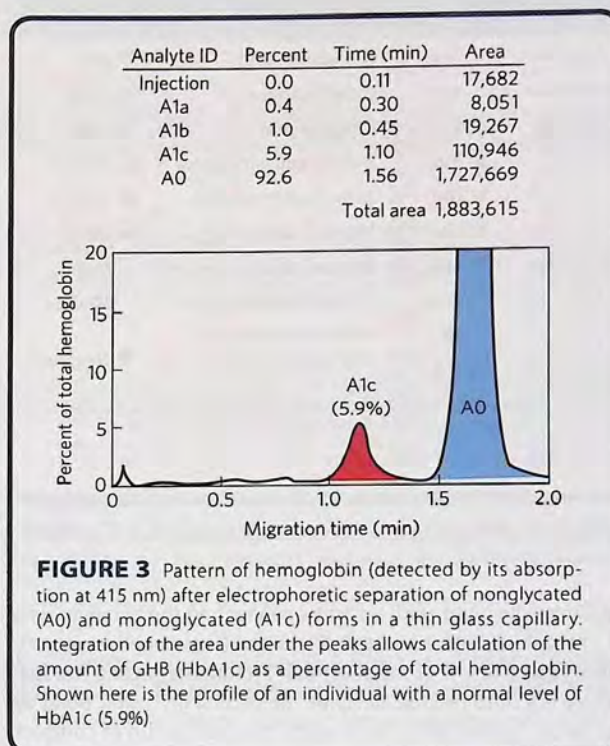
The disaccharide maltose (Fig. 7-10) contains two D-glucose residues joined by a glycosidic linkage between C-1 (the anomeric carbon) of one glucose residue and C-4 of the other. Because the disaccharide retains a free anomeric carbon (C-1 of the glucose

residue on the right in Fig. 7-10), maltose is a reducing sugar. The configuration of the anomeric carbon atom in the glycosidic linkage is α . The glucose residue with the free anomeric carbon is capable of existing in α - and β -pyranose forms.



The extent of **hemoglobin glycation** (so named to distinguish it from glycosylation, the *enzymatic* transfer of glucose to a protein) is measured clinically by extracting hemoglobin from a small sample of blood and separating GHB from unmodified hemoglobin electrophoretically (Fig. 3), taking advantage of the charge difference resulting from modification of the amino group(s). Normal values of the monoglycated hemoglobin referred to as HbA1c are about 5% of total hemoglobin (corresponding to an average blood glucose level of 120 mg/100 mL). In people with untreated diabetes, however, this value may be as high as 13%, indicating an average blood glucose level of about 300 mg/100 mL — dangerously high. One criterion for success in an individual program of insulin therapy (the timing, frequency, and amount of insulin injected) is maintaining HbA1c values at about 7%.

Glycated hemoglobin undergoes a series of rearrangements, oxidations, and dehydrations of the carbohydrate moiety to produce a heterogeneous mixture of advanced glycation end products (AGE), such as ϵ -N-carboxymethyl-lysine and methylglyoxal. AGE are believed to be responsible for at least some of the pathology of diabetes. AGE can leave the erythrocyte and form covalent cross-links between proteins, interfering with normal protein function (Fig. 4). The accumulation of relatively high concentrations of AGE in people with diabetes may, by cross-linking critical proteins, cause the damage to the kidneys, retinas, and cardiovascular system that characterizes the disease. Some AGE also can act



through membrane receptors for AGE (RAGE), triggering intracellular responses that include activation of a transcription factor (NF- κ B) and consequent changes in gene expression. This pathogenic process is a potential target for drug action.

KEY CONVENTION To name reducing disaccharides such as maltose unambiguously, and especially to name more complex oligosaccharides, several rules are followed. By convention, the name describes the compound written with its nonreducing end to the left, and we can “build up” the name in the following

order. (1) Give the configuration (α or β) at the anomeric carbon joining the first monosaccharide unit (on the left) to the second. (2) Name the nonreducing residue; to distinguish five- and six-membered ring structures, insert “furano” or “pyrano” into the name. (3) Indicate in parentheses the two carbon atoms joined by

TABLE 7-1 Symbols and Abbreviations for Common Monosaccharides and Some of Their Derivatives

Abequose	Abe	Glucuronic acid	◊ GlcA
Arabinose	Ara	Galactosamine	◻ GalN
Fructose	Fru	Glucosamine	◻ GlcN
Fucose	▲ Fuc	N-Acetylglucosamine	◻ GalNAc
Galactose	◊ Gal	N-Acetylglucosamine	◻ GlcNAc
Glucose	● Glc	Iduronic acid	◊ IdOA
Mannose	● Man	Muramic acid	Mur
Rhamnose	Rha	N-Acetylmuramic acid	Mur2Ac
Ribose	Rib	N-Acetylneuraminic acid (a sialic acid)	◆ Neu5Ac
Xylose	★ Xyl		

Note: In a commonly used convention, hexoses are represented as circles, N-acetylhexosamines as squares, and hexosamines as squares divided diagonally. All sugars with the 'gluco' configuration are blue, those with the 'galacto' configuration are yellow, and 'manno' sugars are green. Other substituents can be added as needed: sulfate (S), phosphate (P), O-acetyl (OAc), or O-methyl (OMe).

the glycosidic bond, with an arrow connecting the two numbers; for example, (1→4) shows that C-1 of the first-named sugar residue is joined to C-4 of the second. (4) Name the second residue. If there is a third residue, describe the second glycosidic bond by the same conventions. (To shorten the description of complex polysaccharides, three-letter abbreviations or colored symbols for the monosaccharides are often used, as given in Table 7-1.) Following this convention for naming oligosaccharides, maltose is α -D-glucopyranosyl-(1→4)-D-glucopyranose. Because most sugars encountered in this book are the D enantiomers and the pyranose form of hexoses predominates, we generally use a shortened version of the formal name of such compounds, giving the configuration of the anomeric carbon and naming the carbons joined by the glycosidic bond. In this abbreviated nomenclature, maltose is Glc(α 1→4)Glc. ■

The disaccharide lactose (Fig. 7-11), which yields D-galactose and D-glucose on hydrolysis, occurs naturally in milk and gives milk its sweetness. The anomeric carbon of the glucose residue is available for oxidation, and thus lactose is a reducing disaccharide. The reducing end of this disaccharide, which by convention is drawn on the right, is glucose, and the disaccharide is named as a derivative of glucose. Its abbreviated name is Gal(β 1→4)Glc, with the linkage shown in parentheses; the anomeric carbon of galactose is β , and C-1 of Gal is linked to C-4 of Glc.

The enzyme lactase—absent in lactose-intolerant individuals—begins the digestive process in the small intestine by splitting the (β 1→4) bond of lactose into monosaccharides, which can be absorbed from the small intestine. Lactose, like other disaccharides, is not absorbed from the small intestine, and in lactose-intolerant individuals the undigested lactose passes into the large intestine. Here, the increased osmolarity due

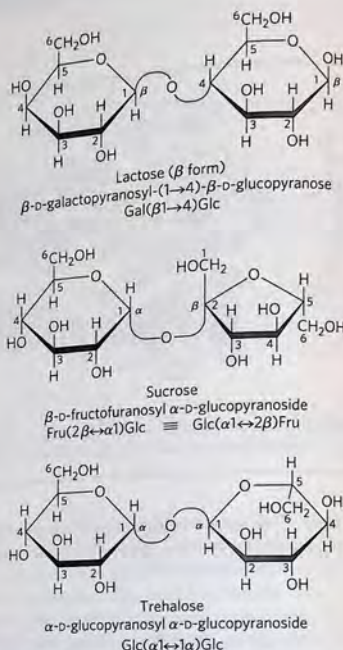


FIGURE 7-11 Three common disaccharides. Like maltose in Figure 7-10, these disaccharides are shown as Haworth perspectives. The common name, full systematic name, and abbreviation are given. Formal nomenclature for sucrose names glucose as the parent glycoside, although it is typically depicted as shown, with glucose on the left. The two abbreviated symbols shown for sucrose are equivalent (■).

to dissolved lactose opposes the absorption of water from the intestine into the bloodstream, causing watery, loose stools. In addition, fermentation of the lactose by intestinal bacteria produces large volumes of CO_2 , which leads to the bloating, cramps, and gas associated with lactose intolerance.

Sucrose (table sugar) is a disaccharide of glucose and fructose. It is formed by plants but not by animals. In contrast to maltose and lactose, sucrose contains no free anomeric carbon atom; the anomeric carbons of both monosaccharide units are involved in the glycosidic bond (Fig. 7-11). Sucrose is therefore a nonreducing sugar, and its stability—its resistance to oxidation—makes it a suitable molecule for the storage and transport of energy in plants. In the abbreviated nomenclature, a double-headed arrow connects the symbols specifying the anomeric carbons and their configurations. Thus the abbreviated name of sucrose is either Glc(α 1→2 β)Fru or Fru(2 β ↔1 α)Glc. Sucrose is either a major intermediate product of photosynthesis; in many plants it is the principal form in which sugar is transported from the leaves to other parts of the plant body.

Trehalose, $\text{Glc}(\alpha 1 \leftrightarrow 1 \alpha)\text{Glc}$ (Fig. 7-11) — a disaccharide of D-glucose that, like sucrose, is a nonreducing sugar — is a major constituent of the circulating fluid (hemolymph) of insects. It serves as an energy-storage compound, and it serves as antifreeze in some invertebrate organisms.

SUMMARY 7.1 Monosaccharides and Disaccharides

- Sugars (saccharides) contain an aldehyde or ketone group and two or more hydroxyl groups.
- Monosaccharides generally contain several chiral carbons and therefore exist in a variety of stereochemical forms, which may be represented on paper as Fischer projections. Epimers are sugars that differ in configuration at only one carbon atom.
- Monosaccharides with at least six carbons commonly form cyclic structures, in which the aldehyde or ketone group is joined to a hydroxyl group of the same molecule. The cyclic structure can be represented as a Haworth perspective formula. The carbon atom originally found in the aldehyde or ketone group (the anomeric carbon) can assume either of two configurations, α and β , which are interconvertible by mutarotation. In the linear form of the monosaccharide, which is in equilibrium with the cyclic forms, the anomeric carbon is easily oxidized, making the compound a reducing sugar.
- Naturally occurring hexoses include some that have $-\text{NH}_2$ at C-2 (amino sugars), often with the amino group acetylated. Oxidation of the carbonyl carbon of glucose and other aldoses yields aldonic acids (gluconic acid); oxidation at C-6 produces uronic acids (glucuronate). Some sugar intermediates are phosphate esters (for example, glucose 6-phosphate).
- A hydroxyl group of one monosaccharide can add to the anomeric carbon of a second monosaccharide to form an acetal called a glycoside. Oligosaccharides are short polymers of several different monosaccharides joined by glycosidic bonds. At one end of the chain, the reducing end, is a monosaccharide residue with its anomeric carbon not involved in a glycosidic bond, and therefore available to be oxidized.
- Disaccharides and oligosaccharides are named as derivatives of the sugar at the reducing end. Their names provide the order of monosaccharide units, the configuration at each anomeric carbon, and the carbon atoms involved in the glycosidic linkage(s).

7.2 Polysaccharides

Most carbohydrates found in nature occur as polysaccharides, polymers of medium to high molecular weight ($M_r > 20,000$). **P2** Polysaccharides, also called **glycans**, differ from each other in the identity of their recurring monosaccharide units, in the length of their chains, in

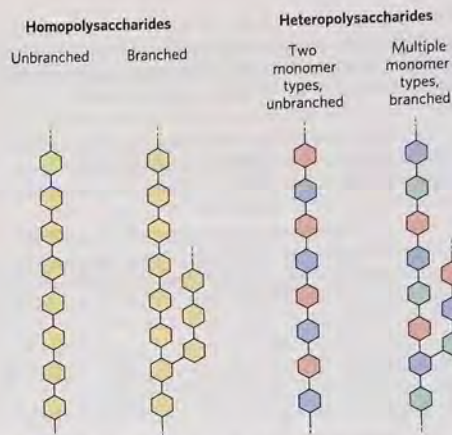


FIGURE 7-12 Homopolysaccharides and heteropolysaccharides. Polysaccharides may be composed of one, two, or several different monosaccharides, in straight or branched chains of varying length.

the types of bonds linking the units, and in the degree of branching. **Homopolysaccharides** contain only a single monomeric sugar species; **heteropolysaccharides** contain two or more kinds of monomers (Fig. 7-12). Some homopolysaccharides serve as storage forms of monosaccharides that are used as fuels; starch and glycogen are homopolysaccharides of this type. Other homopolysaccharides (cellulose and chitin, for example) serve as structural elements in plant cell walls and animal exoskeletons. Heteropolysaccharides provide extracellular support for organisms of all kingdoms. For example, the rigid layer of the bacterial cell envelope (the peptidoglycan) is composed in part of a heteropolysaccharide built from two alternating monosaccharide units (see Fig. 6-32). In animal tissues, the extracellular space is occupied by several types of heteropolysaccharides, which form a matrix that holds individual cells together and provides protection, shape, and support to cells, tissues, and organs.

P4 Unlike proteins, polysaccharides generally do not have defined lengths or molecular weights. This difference is a consequence of the mechanisms of assembly of the two types of polymer. As we shall see in Chapter 27, proteins are synthesized on a template (messenger RNA) of defined sequence and length, by enzymes that follow the template exactly. For polysaccharide synthesis there is no template; rather, the program for polysaccharide synthesis is intrinsic to the enzymes that catalyze the polymerization of the monomeric units, and there is no specific stopping point in the synthetic process; the products thus vary in length.

Some Homopolysaccharides Are Storage Forms of Fuel

The most important storage polysaccharides are starch in plant cells and glycogen in animal cells. Both

polysaccharides occur intracellularly as large clusters or granules. Starch and glycogen molecules are heavily hydrated, because they have many exposed hydroxyl groups available to hydrogen-bond with water. Most plant cells have the ability to form starch, and starch storage is especially abundant in tubers (underground stems), such as potatoes, and in seeds.

Starch contains two types of glucose polymer, amylose and amylopectin (Fig. 7-13). Amylose consists of long, unbranched chains of D-glucose residues connected by (α 1 $\rightarrow$ 4) linkages (as in maltose). Such chains vary in molecular weight from a few thousand to more than a million. Amylopectin is even larger (M_r up to 200 million) but unlike amylose is highly branched. The glycosidic linkages joining successive glucose residues in amylopectin chains are (α 1 $\rightarrow$ 4); the branch points (occurring every 24 to 30 residues) are (α 1 $\rightarrow$ 6) linkages.

Glycogen is the main storage polysaccharide of animal cells. Like amylopectin, glycogen is a polymer of (α 1 $\rightarrow$ 4)-linked glucose subunits, with (α 1 $\rightarrow$ 6)-linked branches, but glycogen is more extensively branched (on average, a branch every 8 to 12 residues) and more compact than starch. Glycogen is especially abundant in the liver, where it may constitute as much as 7% of the wet

weight; it is also present in skeletal muscle. In hepatocytes, glycogen is found in large α granules (see Fig. 15-1), which are clusters of smaller β granules, each of which is a cluster of smaller glycogen molecule with an average single, highly branched glycogen molecule with an average molecular weight of several million. The large α glycogen granules also contain, in tightly bound form, the enzymes responsible for the synthesis and degradation of glycogen (see Fig. 15-16).

Because each branch of glycogen ends with a nonreducing sugar unit, a glycogen molecule with n branches has $n + 1$ nonreducing ends, but only one reducing end. When glycogen is used as an energy source, glucose units are removed one at a time from the nonreducing ends. Degradative enzymes that act only at nonreducing ends can work simultaneously on the many branches, speeding the conversion of the polymer to monosaccharides.

Why not store glucose in its monomeric form?

P3 Hepatocytes in the fed state store glycogen equivalent to a glucose concentration of 0.4 M. The actual concentration of glycogen, which contributes little to the osmolarity of the cytosol, is about 0.01 μ M. If the cytosol contained 0.4 M glucose, the osmolarity would be threateningly elevated, leading to osmotic entry of water that might rupture the cell (see Fig. 2-12).

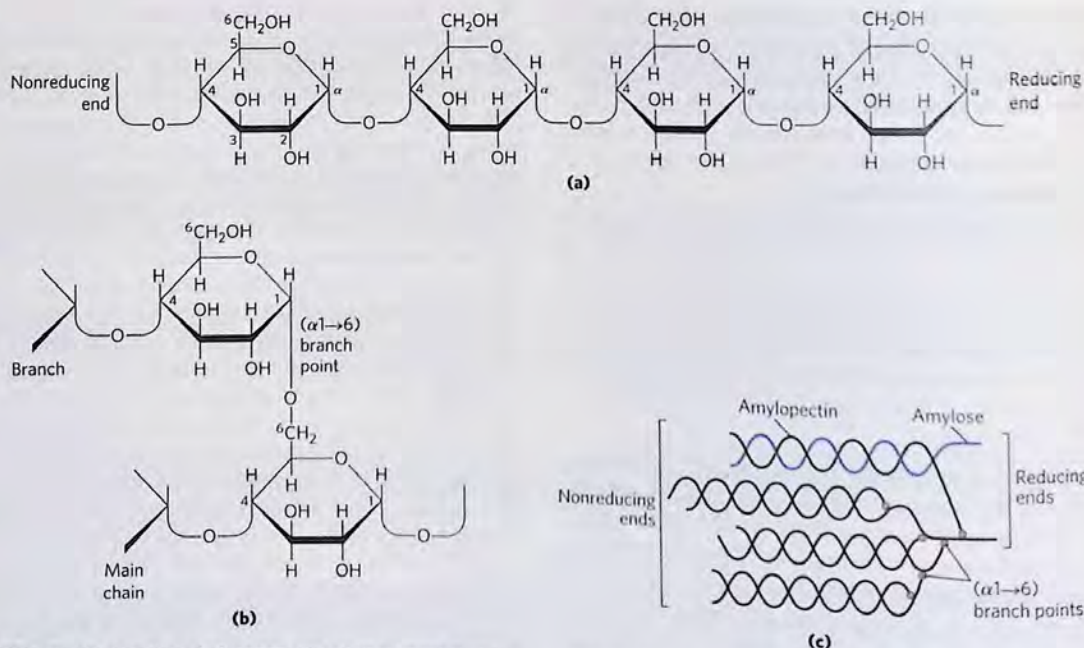


FIGURE 7-13 Starch and glycogen. (a) A short segment of a linear polymer of D-glucose residues in (α 1 $\rightarrow$ 4) linkage, shown here as Haworth perspectives. This is the basic structure of amylose of starch (in plants) and glycogen (in animals). A single chain of amylose can contain several thousand glucose residues. (b) An (α 1 $\rightarrow$ 6) branch point of amylopectin or glycogen. (c) A cluster of amylose and amylopectin like that believed to occur in starch

granules. Strands of amylopectin (black) form double-helical structures with each other or with amylose strands (blue). Amylopectin has frequent (α 1 $\rightarrow$ 6) branches are removed enzymatically during the mobilization of starch for energy production. Glycogen has a similar structure but is more highly branched and more compact.

Some Homopolysaccharides Serve Structural Roles

Cellulose, a tough, fibrous, water-insoluble substance, is found in the cell walls of plants, particularly in stalks, stems, trunks, and all the woody portions of the plant body. Cellulose constitutes much of the mass of wood, and cotton is almost pure cellulose. Like amylose, the cellulose molecule is a linear, unbranched homopolysaccharide, consisting of 10,000 to 15,000 D-glucose units. But there is a very important difference: **P5** in cellulose the glucose residues have the β configuration (Fig. 7-14), whereas in amylose the glucose is in the α configuration. The glucose residues in cellulose are linked by (β 1 $\rightarrow$ 4) glycosidic bonds, in contrast to the (α 1 $\rightarrow$ 4) bonds of amylose. This difference causes individual molecules of cellulose and amylose to fold differently in space, giving them very different macroscopic structures and physical properties.

The tough, fibrous nature of cellulose makes it useful in such commercial products as cardboard and insulation material, and it is a major constituent of cotton and linen fabrics. Cellulose is also the starting material for the commercial production of cellophane, rayon, and lyocell.

Chitin is a linear homopolysaccharide composed of N-acetylglucosamine residues in (β 1 $\rightarrow$ 4) linkage (Fig. 7-15). The only chemical difference from cellulose is the replacement of the hydroxyl group at C-2 with an acetylated amino group, which makes chitin more hydrophobic and water-resistant than cellulose. Chitin forms

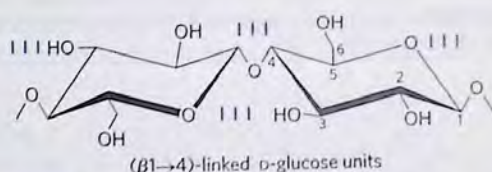


FIGURE 7-14 Cellulose. Two units of a cellulose chain; the D-glucose residues are in (β 1 $\rightarrow$ 4) linkage. The rigid chair structures can rotate relative to one another.

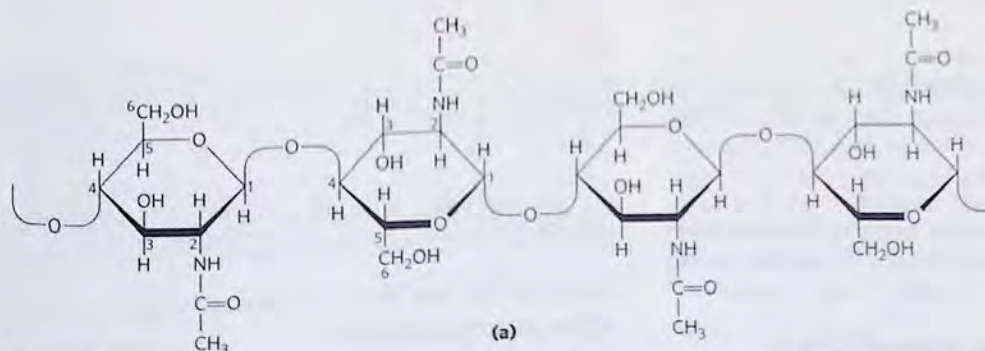


FIGURE 7-15 Chitin. (a) A short segment of chitin, a homopolymer of N-acetyl-D-glucosamine units in (β 1 $\rightarrow$ 4) linkage. (b) A spotted June beetle (*Pelidnota punctata*), showing its surface armor (exoskeleton) of chitin.

extended fibers similar to those of cellulose, and like cellulose cannot be digested by vertebrates. Chitin is the principal component of the hard exoskeletons of nearly a million species of arthropods—insects, lobsters, and crabs, for example—and is probably the second most abundant polysaccharide, next to cellulose, in nature; an estimated 1 billion tons of chitin are produced in the biosphere each year.

Steric Factors and Hydrogen Bonding Influence Homopolysaccharide Folding

P5 The folding of polysaccharides in three dimensions follows the same principles as those governing polypeptide structure: subunits with a more-or-less rigid structure dictated by covalent bonds form three-dimensional macromolecular structures that are stabilized by weak interactions within or between molecules, such as hydrogen bonds, interactions due to the hydrophobic effect, van der Waals interactions, and, for polymers with charged subunits, electrostatic interactions. Because polysaccharides have so many hydroxyl groups, hydrogen bonding has an especially important influence on their structure. Glycogen, starch, cellulose, and chitin are composed of pyranoside (six-membered ring) subunits, as are the oligosaccharides of glycoproteins and glycolipids, to be discussed later. Such molecules can be represented as a series of rigid pyranose rings connected by an oxygen atom bridging two carbon atoms (the glycosidic bond). There is, in principle, free rotation about both C—O bonds linking the residues, but as in polypeptides (see Figs 4-2, 4-8), rotation about each bond is limited by steric hindrance by substituents. The three-dimensional structures of these molecules can be described in terms of the dihedral angles, ϕ and ψ , about the glycosidic bond (Fig. 7-16). The bulkiness of the pyranose ring and its substituents, along with electronic effects at the anomeric carbon, place constraints on the angles ϕ and ψ ; thus, certain conformations are much more stable than others.



tle (*Pelidnota punctata*), showing its surface armor (exoskeleton) of chitin. [© Paul Whitten/Science Source]

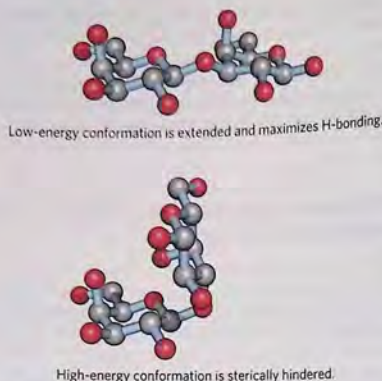


FIGURE 7-16 Different energetic conformations of a disaccharide.

The torsion angles ϕ (phi) and ψ (psi), which define the spatial relationship between adjacent rings, can in principle have any value from 0° to 360° . In fact, some torsion angles give conformations that maximize hydrogen bonding and minimize energy, whereas others give conformations that are sterically hindered, as we can see in two conformers of the disaccharide Gal(β 1 $\rightarrow$ 3)Gal at opposite energetic extremes.

PS The most stable three-dimensional structure for the (α 1 $\rightarrow$ 4)-linked chains of starch and glycogen is a tightly coiled helix (**Fig. 7-17**), stabilized by interchain hydrogen bonds. The average plane of each residue along the amylose chain forms a 60° angle with the average plane of the preceding residue, so the helical structure has six residues per turn. For amylose, the core of the helix is of precisely the right dimensions to accommodate iodine as complex ions (I_2 and I_3^-). This interaction gives an intensely blue product, making it a common qualitative test for amylose.

PS For cellulose, the most stable conformation is that in which each chair is turned 180° relative to its neighbors, yielding a straight, extended chain. All —OH groups are available for hydrogen bonding with neighboring chains. With several chains lying side by side, a stabilizing network of interchain and intrachain hydrogen bonds produces straight, stable supramolecular fibers of great tensile strength (**Fig. 7-18**). This property of cellulose has made it useful to civilizations for millennia. Many manufactured products, including papyrus, paper, cardboard, rayon, insulating tiles, and a variety of other useful materials, are derived from cellulose. The water content of these materials is low because extensive interchain hydrogen bonding between cellulose molecules satisfies their capacity for hydrogen-bond formation.

Peptidoglycan Reinforces the Bacterial Cell Wall

The rigid component of bacterial cell walls (peptidoglycan) is a heteropolymer of alternating (β 1 $\rightarrow$ 4)-linked

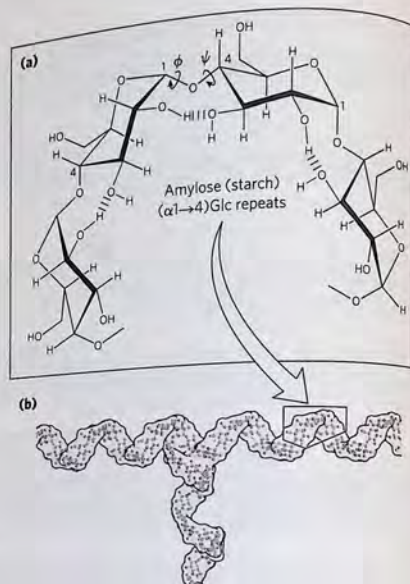


FIGURE 7-17 Helical structure of starch (amylose). (a) Four glucose units of starch (amylose), showing both the free rotation possible around the ψ and ϕ angles of the (α 1 $\rightarrow$ 4) glycosidic bonds and the hydrogen bonding possible between adjacent rings. In this most stable conformation, with adjacent rigid chairs at 60° to one another, the polysaccharide chain is curved. (b) A model of a segment of amylopectin (branched amylose). The conformation of (α 1 $\rightarrow$ 4) linkages in amylose, amylopectin, and glycogen causes these polymers to assume tightly coiled helical structures. [Data from www.biotechnology.co.uk/fsmol/amylopectin.html]

N-acetylglucosamine and *N*-acetylmuramic acid residues (see Fig. 6-32). The linear polymers lie side by side in the cell wall, cross-linked by short peptides, the exact structure of which depends on the bacterial species. The peptide cross-links weld the polysaccharide chains into a strong sheath (peptidoglycan) that envelops the entire cell and prevents cellular swelling and lysis due to the osmotic entry of water.

The enzyme lysozyme kills bacteria by hydrolyzing the (β 1 $\rightarrow$ 4) glycosidic bond between *N*-acetylglucosamine and *N*-acetylmuramic acid. The enzyme is found in human tears, where it is presumably a defense against bacterial infections of the eye, and in hen's eggs, where it prevents bacterial infection of the developing embryo. Penicillin and related antibiotics kill bacteria by preventing synthesis of the peptidoglycan cross-links, leaving the cell wall too weak to resist osmotic lysis (p. 211).

Glycosaminoglycans Are Heteropolysaccharides of the Extracellular Matrix

The extracellular space in the tissues of multicellular animals is filled with a gel-like material, the **extracellular matrix (ECM)**, which holds the cells together and

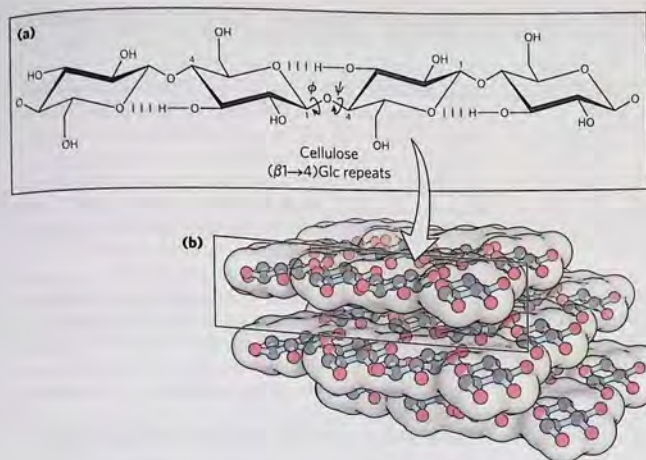


FIGURE 7-18 The linear structure of cellulose chains. (a) Four glucose units of cellulose. In the most stable conformation of cellulose, the $\beta(1 \rightarrow 4)$ glycosidic linkages put the adjacent rigid chairs at 180° relative to each other. (b) A model of straight, unbranched cellulose segments, held together in layers stabilized by interchain hydrogen bonds. (b) Data from Cornell, B. 2016. Sugar Polymers. Available at: <http://fb.bioninja.com.au>. (Accessed 11 March 2020).

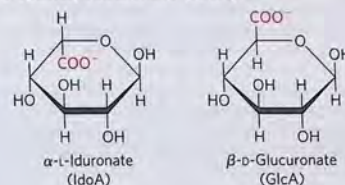
provides a porous pathway for the diffusion of nutrients and oxygen to individual cells. The ECM is composed of an interlocking meshwork of heteropolysaccharides (also called ground substance) and fibrous proteins such as fibrillar collagens, elastins, and fibronectins. The basement membrane is a specialized ECM that underlies epithelial cells; it consists of specialized collagens, laminins, and heteropolysaccharides.

These heteropolysaccharides, the **glycosaminoglycans**, are a family of linear polymers composed of repeating disaccharide units (Fig. 7-19). They are unique to animals and bacteria and are not found in plants. One of the two monosaccharides is always either *N*-acetylglucosamine or *N*-acetylgalactosamine; the other is in most cases a uronic acid, usually D-glucuronic or L-iduronic acid. Some glycosaminoglycans contain esterified sulfate groups. **P2** The combination of sulfate groups and the carboxylate groups of the uronic acid residues gives glycosaminoglycans a very high density of negative charge. To minimize the repulsive forces among neighboring charged groups, these molecules assume an extended conformation in solution, forming a rodlike helix in which the negatively charged carboxylate groups occur on alternate sides of the helix (as shown for heparin in Fig. 7-19). The extended rod form also provides maximum separation between the negatively charged sulfate groups. The specific patterns of sulfated and nonsulfated sugar residues in glycosaminoglycans allow specific recognition by a variety of protein ligands that bind electrostatically to these molecules. The sulfated glycosaminoglycans are attached to extracellular proteins to form proteoglycans (Section 7.3).

The glycosaminoglycan **hyaluronan** (hyaluronic acid) contains alternating residues of D-glucuronic acid and *N*-acetylglucosamine (Fig. 7-19). With up to 50,000 repeats of the basic disaccharide unit, hyaluronan has a

molecular weight of several million; it forms clear, highly viscous, noncompressible solutions that serve as lubricants in the synovial fluid of joints and give the vitreous humor of the vertebrate eye its jellylike consistency (the Greek *hyalos* means "glass"; hyaluronan can have a glassy or translucent appearance). Hyaluronan is also a component of the ECM of cartilage and tendons, to which it contributes tensile strength and elasticity as a result of its strong noncovalent interactions with other components of the matrix. Hyaluronidase, an enzyme secreted by some pathogenic bacteria, can hydrolyze the glycosidic linkages of hyaluronan, rendering tissues more susceptible to bacterial invasion. In many animal species, a similar enzyme in sperm hydrolyzes the outer glycosaminoglycan coat around an ovum, allowing sperm penetration.

Other glycosaminoglycans differ from hyaluronan in three respects: they are generally much shorter polymers, they are covalently linked to specific proteins (proteoglycans), and one or both monomeric units differ from those of hyaluronan. **Chondroitin sulfate** (Greek *chondros*, "cartilage") contributes to the tensile strength of cartilage, tendons, ligaments, heart valves, and the walls of the aorta. **Dermatan sulfate** (Greek *derma*, "skin") contributes to the pliability of skin and is also present in blood vessels and heart valves. In this polymer, many of the glucuronate residues present in chondroitin sulfate are replaced by their C-5 epimer, L-iduronate (IdoA).



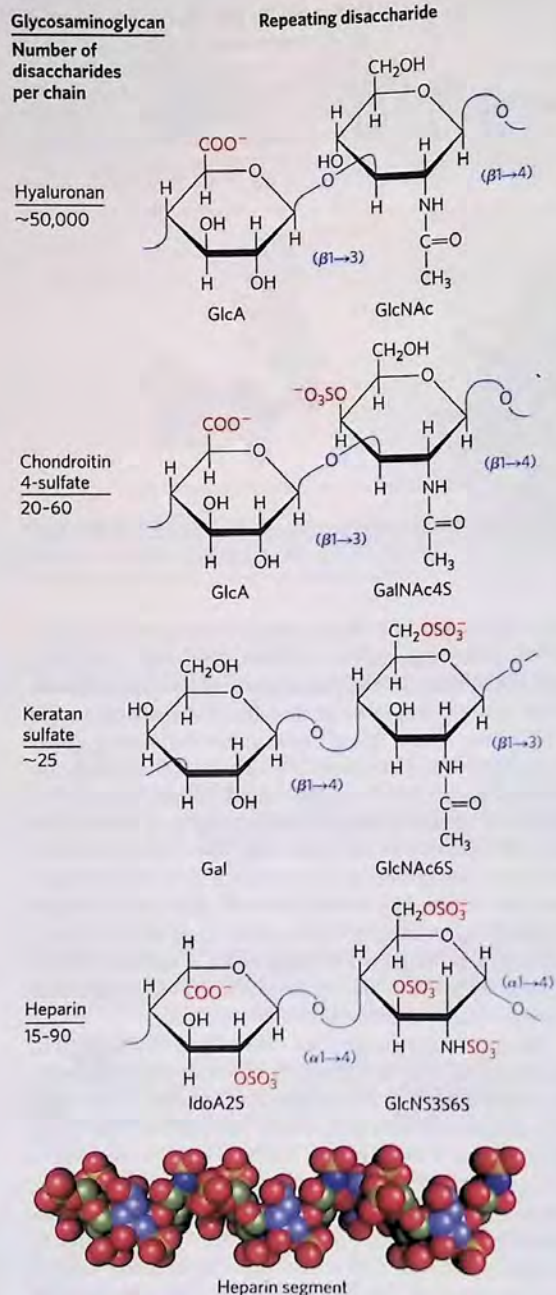


FIGURE 7-19 Repeating units of some common glycosaminoglycans of extracellular matrix. The molecules are copolymers of alternating uronic acid and amino sugar residues (keratan sulfate is the exception), with sulfate esters in any of several positions, except in hyaluronan. The ionized carboxylate and sulfate groups give these polymers their characteristic high negative charge. Therapeutic heparin contains primarily iduronic acid (IdoA) and a smaller proportion of glucuronic acid (not shown) and is generally highly sulfated and heterogeneous in length. The space-filling model shows a heparin segment as its structure in solution, as determined by NMR spectroscopy. Carbons in iduronic acid sulfate are colored blue; those in glucosamine sulfate are green. Oxygen and sulfur are shown in red and yellow, respectively. Hydrogen atoms are not shown (for clarity). [Data for molecular model from PDB ID 1HPN, B. Mulloy et al., *Biochem. J.* 293:849, 1993.]

Keratan sulfates (Greek *keras*, "horn") have no uronic acid, and their sulfate content is variable. They are present in cornea, cartilage, bone, and a variety of horny structures formed from dead cells: horn, hair, hoofs, nails, and claws. **Heparan sulfate** (Greek *hēpar*, "liver"; it was originally isolated from dog liver) contains variable, nonrandom arrangements of sulfated and non-sulfated sugars. The exact sequence of sulfated residues gives the molecule the ability to interact specifically with a large number of proteins, including growth factors and ECM components, as well as various enzymes and factors present in plasma. Heparin is a highly sulfated, intracellular form of heparan sulfate produced primarily by mast cells (a type of leukocyte, or immune cell). Its physiological role is not yet clear, but purified heparin is used as a therapeutic agent to inhibit coagulation of blood through its capacity to bind the protease inhibitor antithrombin (see Fig. 7-24).

Table 7-2 summarizes the composition, properties, roles, and occurrence of the polysaccharides described in Section 7.2.

SUMMARY 7.2 Polysaccharides

- Homopolysaccharides contain only a single monomeric sugar species; heteropolysaccharides contain two or more kinds of monomers.
- The homopolysaccharides starch and glycogen are storage fuels in plant, animal, and bacterial cells. They consist of D-glucose units with (α 1 $\rightarrow$ 4) linkages, and both contain some branches.
- The homopolysaccharides cellulose, chitin, and dextran serve structural roles. Cellulose, composed of (β 1 $\rightarrow$ 4)-linked D-glucose residues, lends strength and rigidity to plant cell walls. Chitin, a polymer of (β 1 $\rightarrow$ 4)-linked N-acetylglucosamine, strengthens the exoskeletons of arthropods.
- Homopolysaccharides assume stable conformations dictated by weak interactions. The chair form of the pyranose ring is essentially rigid, so the conformation of the polymers is determined by rotation about the C—O bonds of the glycosidic linkage. Starch and glycogen form helical structures with intrachain hydrogen bonding; cellulose and chitin form long, straight strands that interact with neighboring strands.
- Bacterial cell walls are strengthened by peptidoglycan in which the repeating disaccharide is GlcNAc(β 1 $\rightarrow$ 4)Mur2Ac.
- Glycosaminoglycans are extracellular heteropolysaccharides in which one of the two monosaccharide units is a uronic acid (keratan sulfate is an exception) and the other is an N-acetylated amino sugar. The high density of negative charge on these molecules forces them to assume extended conformations. These polymers (hyaluronan, chondroitin sulfate, dermatan sulfate, and

TABLE 7-2 Structures and Roles of Some Polysaccharides

Polymer	Type ^a	Repeating unit ^b	Size (number of monosaccharide units)	Roles/significance
Energy storage: in plants				
Amylose	Homo-	($\alpha 1 \rightarrow 4$)Glc, linear	50–5,000	
Amylopectin	Homo-	($\alpha 1 \rightarrow 4$)Glc, with ($\alpha 1 \rightarrow 6$)Glc branches every 24–30 residues	Up to 10^6	
Glycogen	Homo-	($\alpha 1 \rightarrow 4$)Glc, with ($\alpha 1 \rightarrow 6$)Glc branches every 8–12 residues	Up to 50,000	Energy storage: in bacteria and animal cells
Cellulose	Homo-	($\beta 1 \rightarrow 4$)Glc	Up to 15,000	Structural: in plants, gives rigidity and strength to cell walls
Chitin	Homo-	($\beta 1 \rightarrow 4$)GlcNAc	Very large	Structural: in insects, spiders, crustaceans, gives rigidity and strength to exoskeletons
Dextran	Homo-	($\alpha 1 \rightarrow 6$)Glc, with ($\alpha 1 \rightarrow 3$) branches	Wide range	Structural: in bacteria, extracellular adhesive
Peptidoglycan	Hetero-; peptides attached	4)Mur2Ac($\beta 1 \rightarrow 4$)GlcNAc($\beta 1$)	Very large	Structural: in bacteria, gives rigidity and strength to cell envelope
Hyaluronan (a glycosaminoglycan)	Hetero-; acidic	4)GlcA($\beta 1 \rightarrow 3$)GlcNAc($\beta 1$)	Up to 100,000	Structural: in vertebrates, extracellular matrix of skin and connective tissue; viscosity and lubrication in joints

^aEach polymer is classified as a homopolysaccharide (homo-) or heteropolysaccharide (hetero-).

^bThe abbreviated names for the peptidoglycan, agarose, and hyaluronan repeating units indicate that the polymer contains repeats of this disaccharide unit. For example, in peptidoglycan, the GlcNAc of one disaccharide unit is ($\beta 1 \rightarrow 4$)-linked to the first residue of the next disaccharide unit.

keratan sulfate) provide viscosity, adhesiveness, and tensile strength to the extracellular matrix.

7.3 Glycoconjugates: Proteoglycans, Glycoproteins, and Glycolipids

In addition to their important roles as fuel stores (starch, glycogen, dextran) and as structural materials (cellulose, chitin, peptidoglycans), polysaccharides and oligosaccharides are information carriers. Some provide communication between cells and their extracellular surroundings; others label proteins for transport to and localization in specific organelles, or for destruction when the protein is malformed or superfluous; and others serve as recognition sites for extracellular signal molecules (growth factors, for example) or extracellular parasites (bacteria or viruses). On almost every eukaryotic cell, specific oligosaccharide chains attached to components of the plasma membrane form a carbohydrate layer (the glycocalyx), several nanometers thick, that serves as an information-rich surface that the cell shows to its surroundings. These oligosaccharides are central players in cell-cell recognition and adhesion, cell migration during development, blood clotting, the immune response, wound healing, and other cellular processes. In most of these cases, the informational carbohydrate is covalently joined to a protein or a lipid to form a **glycoconjugate**, which is the biologically active molecule (Fig. 7-20).

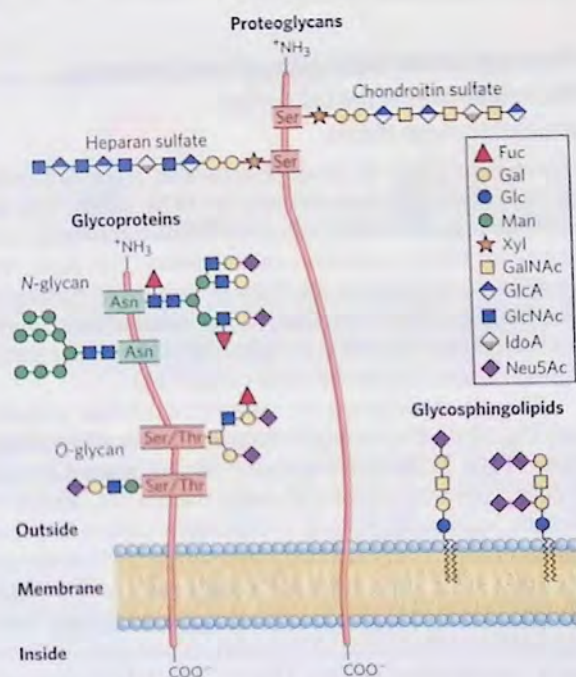


FIGURE 7-20 Glycoconjugates. The structures of some typical proteoglycans, glycoproteins, and glycosphingolipids described in the text.

Proteoglycans are macromolecules of the cell surface or ECM in which one or more sulfated glycosaminoglycan chains are joined covalently to a membrane protein or

a secreted protein. The glycosaminoglycan chain can bind to extracellular proteins through electrostatic interactions between the protein and the negatively charged sugar moieties on the proteoglycan. Proteoglycans are major components of all extracellular matrices.

Glycoproteins have one or several oligosaccharides of varying complexity joined covalently to a protein. They are usually found on the outer face of the plasma membrane (as part of the glycocalyx), in the ECM, and in the blood. Inside cells, they are found in specific organelles such as Golgi complexes (where the oligosaccharide moieties are added to the proteins), secretory granules, and lysosomes.

P6 The oligosaccharide portions of glycoproteins are very heterogeneous and, like glycosaminoglycans, are rich in information, forming highly specific sites for recognition and high-affinity binding by carbohydrate-binding proteins called lectins. Some cytosolic and nuclear proteins can be glycosylated as well.

Glycolipids are plasma membrane components in which the hydrophilic head groups are oligosaccharides.

Glycosphingolipids are a class of glycolipids with a specific backbone structure that will be considered in more detail in Chapter 10. They have complex oligosaccharide components that act as specific sites for recognition by lectins, as they do in glycoproteins. Neurons are rich in glycosphingolipids, which act in nerve conduction and myelin formation. Glycosphingolipids also play a role in signal transduction in cells.

Proteoglycans Are Glycosaminoglycan-Containing Macromolecules of the Cell Surface and Extracellular Matrix

Mammalian cells can produce dozens of types of proteoglycans. Many are secreted into the ECM, where they act as tissue organizers and influence cellular activities, such as growth factor activation and adhesion. The basic proteoglycan unit consists of a "core protein" with covalently attached glycosaminoglycan(s). The point of attachment is a Ser residue, to which the glycosaminoglycan is joined through a tetrasaccharide linker (Fig. 7-21).

Some proteoglycans are integral membrane proteins (see Fig. 11-9). For example, the basal lamina, a specialized layer of the ECM that separates two organized groups of cells, contains a family of core proteins (M_r 20,000 to 40,000), each with several covalently attached heparan sulfate chains. There are two major families of membrane heparan sulfate proteoglycans. **Syndecans** have a single transmembrane domain and an extracellular domain bearing three to five chains of heparan sulfate and, in some cases, chondroitin sulfate (Fig. 7-22a). **Glypicans** are attached to the membrane by a GPI anchor, a glycosylated derivative of the membrane lipid phosphatidylinositol (see Fig. 11-16). Both syndecans and glypicans can be shed into the extracellular space. A protease in the ECM that cuts proteins close to the membrane surface releases syndecan ectodomains (domains outside the plasma membrane),

and a phospholipase that breaks the connection to the GPI anchor releases glypicans. These mechanisms provide a way for a cell to change its surface features quickly. Shedding is highly regulated and is activated in proliferating cells, such as cancer cells. Proteoglycan shedding

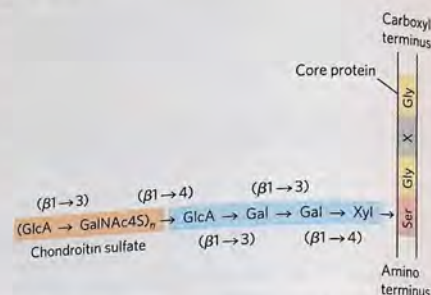


FIGURE 7-21 Proteoglycan structure, showing the tetrasaccharide bridge. A typical tetrasaccharide linker (blue) connects a glycosaminoglycan—in this case, chondroitin sulfate (orange)—to a Ser residue in the core protein. The xylose residue at the reducing end of the linker is joined by its anomeric carbon to the hydroxyl of the Ser residue.

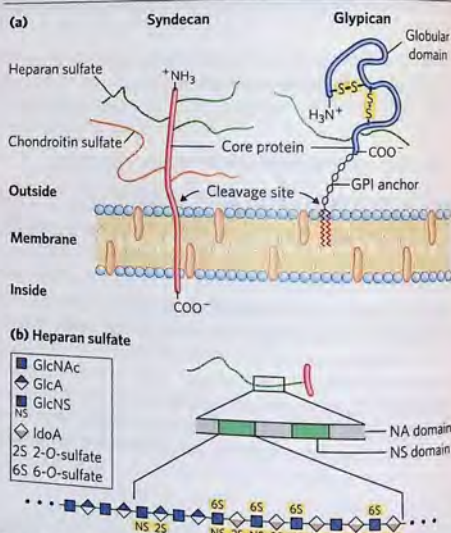


FIGURE 7-22 Two families of membrane proteoglycans. (a) Syndecans and glypicans can be shed from the membrane by enzymatic cleavage near the outer membrane surface. (b) Along a heparan sulfate chain, regions rich in sulfated sugars, the NS domains (green), alternate with regions with mostly unmodified residues of GlcNAc and GlcA, the NA domains (gray). One of the NS domains is shown in more detail, revealing a high density of modified residues: GlcNAc (N-sulfoglucosamine) and both GlcA and IdoA, with a sulfate ester at C-2. (a) Information from U. Häcker et al., *Nature Rev. Mol. Cell Biol.* 6:530, 2005. (b) Information from J. Turnbull et al., *Trends Cell Biol.* 11:75, 2001.]

is involved in cell-cell recognition and adhesion, and in the proliferation and differentiation of cells. Numerous chondroitin sulfate and dermatan sulfate proteoglycans also exist, some as membrane-bound entities, others as secreted products in the ECM.

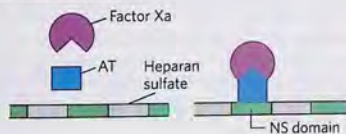
The glycosaminoglycan chains in proteoglycans can bind to a variety of extracellular ligands and thereby modulate the ligands' interaction with specific receptors of the cell surface. Detailed studies of heparan sulfate demonstrate a domain structure that is not random; some domains (typically three to eight disaccharide units long) differ from neighboring domains in sequence and in ability to bind to specific proteins. Highly sulfated domains (called NS domains) alternate with domains having unmodified GlcNAc and GlcA residues (*N*-acetylated, or NA, domains) (Fig. 7-22b). The exact pattern of sulfation in the NS domain depends on the particular proteoglycan; given the number of possible modifications of the GlcNAc-IdoA (iduronic acid) dimer, at least 32 different disaccharide units are possible. Furthermore, the same core protein can display different heparan sulfate structures when synthesized in different cell types.

Heparan sulfate molecules with precisely organized NS domains bind specifically to extracellular proteins and signaling molecules to alter their activities. The change in activity may result from a conformational change in the protein that is induced by the binding (Fig. 7-23a), or it may be due to the ability of adjacent domains of heparan

sulfate to bind to two different proteins, bringing them into close proximity and enhancing protein-protein interactions (Fig. 7-23b). A third general mechanism of action is the binding of extracellular signal molecules (growth factors, for example) to heparan sulfate, which increases their local concentrations and enhances their interaction with growth factor receptors on the cell surface; in this case, the heparan sulfate acts as a coreceptor (Fig. 7-23c). For example, fibroblast growth factor (FGF), an extracellular protein signal that stimulates cell division, first binds to heparan sulfate moieties of syndecan molecules in the target cell's plasma membrane. Syndecan presents FGF to the FGF plasma membrane receptor, and only then can FGF interact productively with its receptor to trigger cell division. Finally, in another type of mechanism, the NS domains interact—electrostatically and otherwise—with a variety of soluble molecules outside the cell, maintaining high local concentrations at the cell surface (Fig. 7-23d).

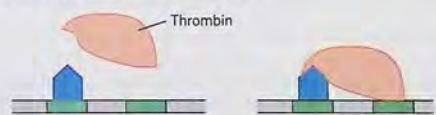
The protease thrombin, essential to blood coagulation (see Fig. 6-44), is inhibited by another blood protein, antithrombin, which prevents premature blood clotting. Antithrombin does not bind to or inhibit thrombin in the absence of heparan sulfate. In the presence of heparan sulfate or heparin, the binding affinity of thrombin for antithrombin increases 2,000-fold, and thrombin is strongly inhibited. When thrombin and antithrombin are crystallized in the presence of a short (16 residue) segment of

(a) Conformational activation



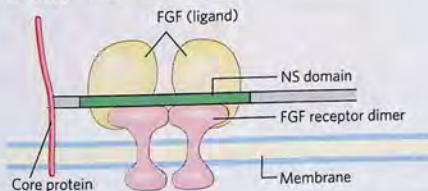
A conformational change induced in the protein antithrombin (AT) on binding a specific pentasaccharide NS domain allows its interaction with blood clotting factor Xa, preventing clotting.

(b) Enhanced protein-protein interaction



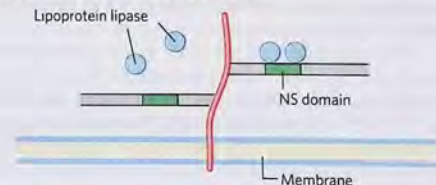
Binding of AT and thrombin to two adjacent NS domains brings the two proteins into close proximity, favoring their interaction, which inhibits blood clotting.

(c) Coreceptor for extracellular ligands



NS domains interact with both the fibroblast growth factor (FGF) and its receptor, bringing the oligomeric complex together and increasing the effectiveness of a low concentration of FGF.

(d) Cell surface localization/concentration



The high density of negative charges in heparan sulfate attracts positively charged lipoprotein lipase molecules and holds them by electrostatic and sequence-specific interactions with NS domains.

FIGURE 7-23 Four types of protein interactions with NS domains of heparan sulfate. [Information from J. Turnbull et al., *Trends Cell Biol.* 11:75, 2001.]

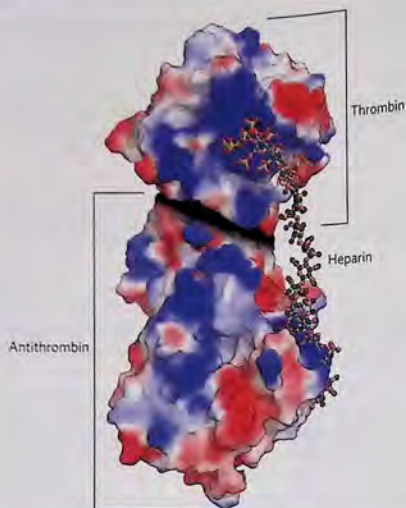


FIGURE 7-24 Molecular basis for heparan sulfate enhancement of the binding of thrombin to antithrombin. In this crystal structure of thrombin, antithrombin, and a 16 residue heparan sulfate-like polymer, all crystallized together, the binding sites for heparan sulfate in both proteins are rich in Arg and Lys residues. These positively charged regions, shown in blue in this electrostatic representation of the proteins, allow strong electrostatic interaction with multiple negatively charged sulfates and carboxylates of the heparan sulfate. Consequently, the affinity of antithrombin for thrombin is three orders of magnitude greater in the presence of heparan sulfate than in its absence. [Data from PDB ID 1TB6, W. Li et al., *Nat. Struct. Mol. Biol.* 11:857, 2004.]

heparan sulfate, the negatively charged heparan sulfate mimic is seen to bridge positively charged regions of the two proteins, causing an allosteric change that inhibits thrombin's protease activity (Fig. 7-24). The binding sites for heparan sulfate and heparin in both proteins are rich in Arg and Lys residues; the amino acids' positive charges interact electrostatically with the sulfates of the glycosaminoglycans. Antithrombin also inhibits two other blood coagulation proteins (factors IXa and Xa) in a heparan sulfate-dependent process.

The importance of correctly synthesizing sulfated domains in heparan sulfate is demonstrated in mutant ("knockout") mice lacking the enzyme that sulfates the C-2 hydroxyl of iduronate (IdoA). These animals are born without kidneys and with severe developmental abnormalities of the skeleton and eyes. Other studies demonstrate that membrane proteoglycans are important in the liver for clearing lipoproteins from the blood. Finally, there is growing evidence that proteoglycans containing heparan sulfate and chondroitin sulfate provide directional cues for axon outgrowth, influencing the path taken by developing axons in the nervous system.

The functional importance of proteoglycans and the glycosaminoglycans associated with them can also be seen in the effects of mutations that block the synthesis or degradation of these polymers in humans (Box 7-3).

Some proteoglycans can form **proteoglycan aggregates**, enormous supramolecular assemblies of many core proteins all bound to a single molecule of hyaluronan (also called hyaluronate). Aggrecan core protein ($M_r \sim 250,000$) has multiple chains of chondroitin sulfate and keratan sulfate, joined to Ser residues in the core protein through trisaccharide linkers, to give an aggrecan monomer of $M_r \sim 2 \times 10^6$. When a hundred or more of these "decorated" core proteins bind a single, extended molecule of hyaluronan (Fig. 7-25), the resulting proteoglycan aggregate ($M_r > 2 \times 10^8$) and its associated water of hydration occupy a volume about equal to that of a bacterial cell! Aggrecan interacts strongly with collagen in the ECM of cartilage, contributing to the development, tensile strength, and resilience of this connective tissue.

Interwoven with these enormous extracellular proteoglycans are fibrous matrix proteins such as collagen, elastin, and fibronectin, forming a cross-linked meshwork that gives the whole ECM strength and resilience. Some of these proteins are multiadhesive, a single protein having binding sites for several different matrix molecules. **Fibronectin**, for example, has separate domains that bind fibrin, heparan sulfate, and collagen. Fibronectin and a number of other proteins in the extracellular matrix contain the conserved RGD sequence (Arg-Gly-Asp), through which they bind to a family of proteins called **integrins**. Integrins mediate signaling between the cell interior and the network of molecules in the ECM. The overall picture of cell-matrix

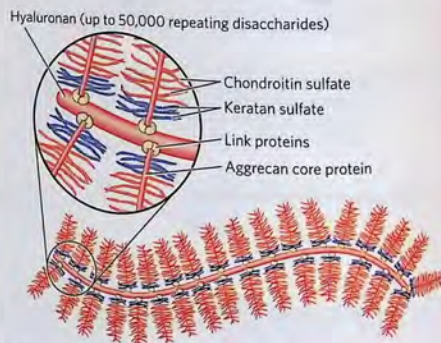


FIGURE 7-25 Proteoglycan aggregate of the extracellular matrix. Schematic drawing of a proteoglycan with many aggrecan molecules. One very long molecule of hyaluronan is associated noncovalently with about 100 molecules of the core protein aggrecan. Each aggrecan molecule contains many covalently bound chondroitin sulfate and keratan sulfate chains. Link proteins at the junction between each core protein and the hyaluronan backbone mediate the core protein-hyaluronan interaction.



Defects in the Synthesis or Degradation of Sulfated Glycosaminoglycans Can Lead to Serious Human Disease

Glycosaminoglycan synthesis requires enzymes that activate monomeric sugars, transport them across membranes, condense the activated sugars into polysaccharides, and add sulfates. Mutations in any of these enzymes in humans can lead to structural defects in the glycosaminoglycan (or in the proteoglycans formed from them). The result can be any of a wide variety of defects in cell signaling, cell proliferation, tissue morphogenesis, or interactions with growth factors (Fig. 1). For example, failure to extend the disaccharide unit GlcNAc-GlcA leads to a bone abnormality in which multiple, large bone spurs develop (Fig. 2).

When the defect occurs in degradative enzymes, the accumulation of incompletely degraded glycosaminoglycans can produce diseases ranging from moderate, as in Scheie

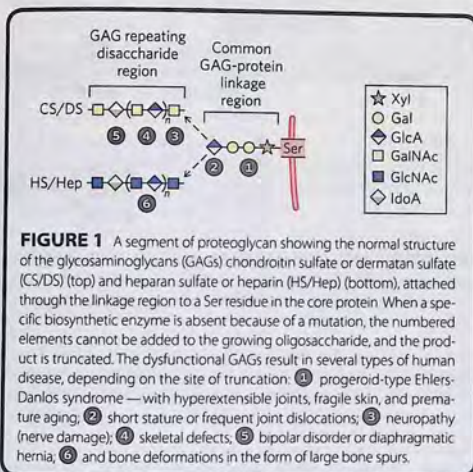


FIGURE 2 Bone deformation characteristic of multiple hereditary exostoses, a disease resulting from a genetic inability to add the GlcNAc-GlcA disaccharide to the growing heparan sulfate or heparin chain (see ⑥ in Fig. 1). The extra bone growth is artificially colored red in this x-ray of the humerus (upper arm bone). [CNRI/Science Photo Library/Science Source.]

syndrome, with joint stiffening but normal intelligence and life span, to severe, as in Hurler syndrome, with enlarged internal organs, heart disease, dwarfism, intellectual disability, and early death. Glycosaminoglycans were formerly called mucopolysaccharides, and diseases caused by genetic defects in their breakdown are often still called mucopolysaccharidoses.

interactions that emerges (Fig. 7-26) shows an array of interactions between cellular and extracellular molecules. These interactions serve not merely to anchor cells to the ECM, providing the strength and elasticity of skin and joints. **P7** They also provide paths that direct the migration of cells in developing tissue and serve to convey information in both directions across the plasma membrane.

Glycoproteins Have Covalently Attached Oligosaccharides

Glycoproteins are carbohydrate-protein conjugates in which the glycans are branched and are much smaller and more structurally diverse than the huge glycosaminoglycans of proteoglycans. The carbohydrate is attached at its anomeric carbon through a glycosidic link to the —OH of a Ser or Thr residue (O-linked), or through an

N-glycosyl link to the amide nitrogen of an Asn residue (N-linked) (Fig. 7-27). **N-glycosyl bonds** join the anomeric carbon of a sugar to a nitrogen atom in glycoproteins and nucleotides (see Fig. 8-1). Some glycoproteins have a single oligosaccharide chain, but many have more than one; the carbohydrate may constitute from 1% to 70% of the glycoprotein by mass. About half of all proteins of mammals are glycosylated.

As we shall see in Chapter 11, the external surface of the plasma membrane has many membrane glycoproteins with arrays of covalently attached oligosaccharides of varying complexity. **Mucins** are secreted or membrane glycoproteins that can contain large numbers of O-linked oligosaccharide chains. Mucins are present in most secretions; they are what gives mucus its characteristic slipperiness. Many of the proteins secreted by eukaryotic

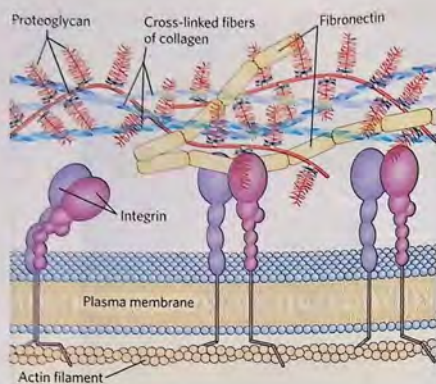


FIGURE 7-26 Interactions between cells and the extracellular matrix. The association between cells and the proteoglycan of the extracellular matrix is mediated by a membrane protein (integrin) and by an extracellular protein (fibronectin, with its RGD motif) with binding sites for both integrin and the proteoglycan. Note the close association of collagen fibers with the fibronectin and proteoglycan.

cells are glycoproteins, including most of the proteins of blood. For example, immunoglobulins (antibodies) and certain hormones, such as follicle-stimulating hormone, luteinizing hormone, and thyroid-stimulating hormone, are glycoproteins. Some of the proteins secreted by the pancreas are glycosylated, as are most of the proteins contained in lysosomes. Many milk proteins, including the major whey protein α -lactalbumin, are glycosylated.

Glycomics is the systematic characterization of all carbohydrate components of a given cell or tissue, including those attached to proteins and to lipids. For glycoproteins, this also means determining which proteins are glycosylated and where in the amino acid sequence each oligosaccharide is attached. This is a challenging undertaking, but worthwhile because of the potential insights it offers into normal patterns of glycosylation and the ways in which they are altered during development or in genetic diseases or cancer. Current methods of characterizing the entire carbohydrate complement of cells depend heavily on sophisticated application of nuclear magnetic resonance and mass spectrometry.

The structures of a large number of *O*- and *N*-linked oligosaccharides from a variety of glycoproteins are known; Figures 7-20 and 7-27 show a few typical examples. In Chapter 27, we consider the mechanisms by which specific proteins acquire specific oligosaccharide moieties.

What are the biological advantages of adding oligosaccharides to proteins? The very hydrophilic carbohydrate clusters alter the polarity and solubility of

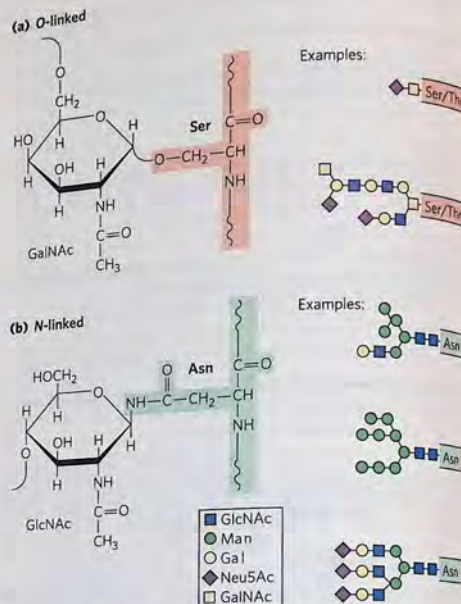


FIGURE 7-27 Oligosaccharide linkages in glycoproteins.

(a) *O*-linked oligosaccharides have a glycosidic bond to the hydroxyl group of Ser or Thr residues (light red), illustrated here with GalNAc as the sugar at the reducing end of the oligosaccharide. One simple chain and one complex chain are shown. (b) *N*-linked oligosaccharides have an *N*-glycosyl bond to the amide nitrogen of an Asn residue (green), illustrated here with GlcNAc as the terminal sugar. Three common types of oligosaccharide chains that are *N*-linked in glycoproteins are shown. A complete description of oligosaccharide structure requires specification of the position and stereochemistry (α or β) of each glycosidic linkage.

the proteins with which they are conjugated. **P7** Oligosaccharide chains that are attached to newly synthesized proteins in the endoplasmic reticulum (ER) and elaborated in the Golgi complex serve as destination labels and also act in protein quality control, targeting misfolded proteins for degradation (see Figs. 27-38, 27-39). When numerous negatively charged oligosaccharide chains are clustered in a single region of a protein, the charge repulsion among them favors the formation of an extended, rodlike structure in that region. The bulkiness and negative charge of oligosaccharide chains also protect some proteins from attack by proteolytic enzymes.

Beyond these global physical effects on protein structure, there are also more specific biological effects of oligosaccharide chains in glycoproteins (Section 7.4). The importance of normal protein glycosylation is clear from at least 40 different genetic disorders of glycosylation

having been found in humans, all causing severely defective physical or mental development; some of these disorders are fatal.

Glycolipids and Lipopolysaccharides Are Membrane Components

Glycoproteins are not the only cellular components that bear complex oligosaccharide chains; some lipids, too, have covalently bound oligosaccharides. **Gangliosides** are membrane lipids of eukaryotic cells in which the polar head group, the part of the lipid that forms the outer surface of the membrane, is a complex oligosaccharide containing a sialic acid (Fig. 7-9) and other monosaccharide residues. Some of the oligosaccharide moieties of gangliosides, such as those that determine human blood groups (see Fig. 10-13), are identical with those found in certain glycoproteins, which therefore also contribute to blood group type. Like the oligosaccharide moieties of glycoproteins, those of membrane lipids are generally, perhaps always, found on the outer face of the plasma membrane.

Lipopolysaccharides are the dominant surface feature of the outer membrane of gram-negative bacteria such as *Escherichia coli* and *Salmonella typhimurium*. These molecules are prime targets of the antibodies produced by the vertebrate immune system in response to bacterial infection and are therefore important determinants of the serotype of bacterial strains. (Serotypes are strains that are distinguished on the basis of antigenic properties.) The lipopolysaccharides of *S. typhimurium* contain six fatty acids bound to two glucosamine residues, one of which is the point of attachment for a complex oligosaccharide (Fig. 7-28). *E. coli* has similar but unique lipopolysaccharides. The lipid A portion of the lipopolysaccharides of some bacteria is called endotoxin; its toxicity to humans and other animals is responsible for the dangerously lowered blood pressure that occurs in toxic shock syndrome resulting from gram-negative bacterial infections. ■

SUMMARY 7.3 Glycoconjugates: Proteoglycans, Glycoproteins, and Glycolipids

- Proteoglycans are huge molecules in which one or more large glycans, called sulfated glycosaminoglycans (heparan sulfate, chondroitin sulfate, dermatan sulfate, or keratan sulfate), are covalently attached to a core protein. Bound to the outside of the plasma membrane through a peptide or lipid, proteoglycans provide points of adhesion, recognition, and information transfer between cells, or between a cell and the extracellular matrix. In the extracellular matrix, enormous proteoglycan aggregates form, bound to a long hyaluronan molecule.
- Glycoproteins contain oligosaccharides covalently linked to Asn, Ser, or Thr residues. The oligosaccharides

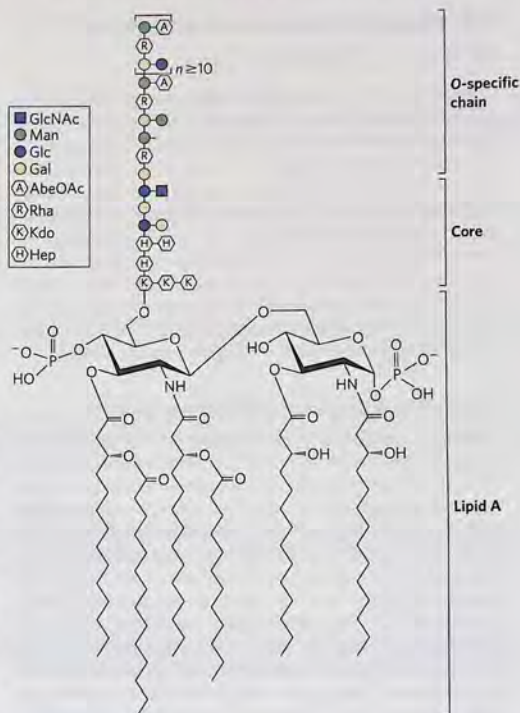


FIGURE 7-28 Bacterial lipopolysaccharides. Schematic diagram of the lipopolysaccharide of the outer membrane of *S. typhimurium*. Kdo is 3-deoxy-D-manno-octulosonic acid (previously called ketodeoxyoctonic acid); Hep is L-glycero-D-manno-heptose; AbeOAc is abequeose (a 3,6-dideoxyhexose) acetylated on one of its hydroxyls. Different bacterial species have subtly different lipopolysaccharide structures, but they have in common a lipid A region, composed of six fatty acid residues and two phosphorylated glucosamines, a core oligosaccharide, and an "O-specific" chain, which is the principal determinant of the serotype (immunological reactivity) of the bacterium. The outer membranes of the gram-negative bacteria *S. typhimurium* and *E. coli* contain so many lipopolysaccharide molecules that the cell surface is almost completely covered with O-specific chains.

are typically branched and smaller than glycosaminoglycans. Many cell surface or extracellular proteins are glycoproteins, as are most secreted proteins. The covalently attached oligosaccharides influence the folding and stability of the proteins, provide critical information about the targeting of newly synthesized proteins, and allow specific recognition by other proteins.

- Glycolipids and glycosphingolipids in plants and animals and lipopolysaccharides in bacteria are components of the cell envelope, with covalently attached oligosaccharide chains exposed on the cell's outer surface.

EX1013, Page 32

neuraminidase in vitro, and reintroducing it into the circulation, the treated erythrocytes disappear from the bloodstream within a few hours; erythrocytes with intact oligosaccharides (withdrawn and reintroduced without neuraminidase treatment) continue to circulate for days.

Selectins are a family of plasma membrane lectins that mediate cell-cell recognition and adhesion in a wide range of cellular processes. One such process is the movement of immune cells (leukocytes) through the capillary wall, from blood to tissues, at sites of infection or inflammation (Fig. 7-29). At an infection site, P-selectin on the surface of capillary endothelial cells interacts with a specific oligosaccharide of the surface glycoproteins of circulating leukocytes. This interaction slows the leukocytes as they roll along the endothelial lining of the capillaries. A second interaction, between integrin molecules in the leukocyte plasma membrane and an adhesion protein on the endothelial cell surface, now stops the leukocyte and allows it to move through the capillary wall into the infected tissues to initiate the immune attack. Two other selectins participate in this "lymphocyte homing": E-selectin on the endothelial cell and L-selectin on the leukocyte bind their cognate oligosaccharides on the leukocyte and endothelial cell, respectively. Several of the selectins essential to lymphocyte homing bind specifically to the tetrasaccharide Neu5Ac-(α 2 $\rightarrow$ 3)-D-Gal-(β 1 $\rightarrow$ 4)(α -L-Fuc[1 $\rightarrow$ 4])-D-GlcNAc, called sialyl Lewis x or sialyl Le^x.

Human selectins mediate the inflammatory responses in rheumatoid arthritis, asthma, psoriasis, multiple sclerosis, and the rejection of transplanted organs,

and thus there is great interest in developing drugs that inhibit selectin-mediated cell adhesion. Many carcinomas express sialyl Lewis x, which, when shed into the circulation, facilitates tumor cell survival and metastasis. Carbohydrate derivatives that mimic the sialyl Lewis x portion of sialoglycoproteins and compete for specific selectin-binding sites, or that alter the biosynthesis of the oligosaccharide, might prove effective as selectin-specific drugs for treating chronic inflammation or metastatic disease.

Several animal viruses, including the influenza virus, attach to their host cells through interactions with oligosaccharides displayed on the host cell surface. The lectin of the influenza virus, known as the HA (hemagglutinin) protein, is essential for viral entry and infection. After the virus has entered a host cell and has been replicated, the newly synthesized viral particles bud out of the cell, wrapped in a portion of its plasma membrane. A viral sialidase (neuraminidase) trims the terminal sialic acid residue from the host cell's oligosaccharides, releasing the viral particles from their interaction with the cell and preventing their aggregation with one another. Another round of infection can now begin. The antiviral drugs oseltamivir (Tamiflu) and zanamivir (Relenza) are used clinically in the treatment of influenza. These drugs are sugar analogs; they inhibit the viral sialidase by competing with the host cell's oligosaccharides for binding (Fig. 7-30). This prevents the release of viruses from the infected cell by sialidase, and also causes viral particles to aggregate, both of which block another cycle of infection.

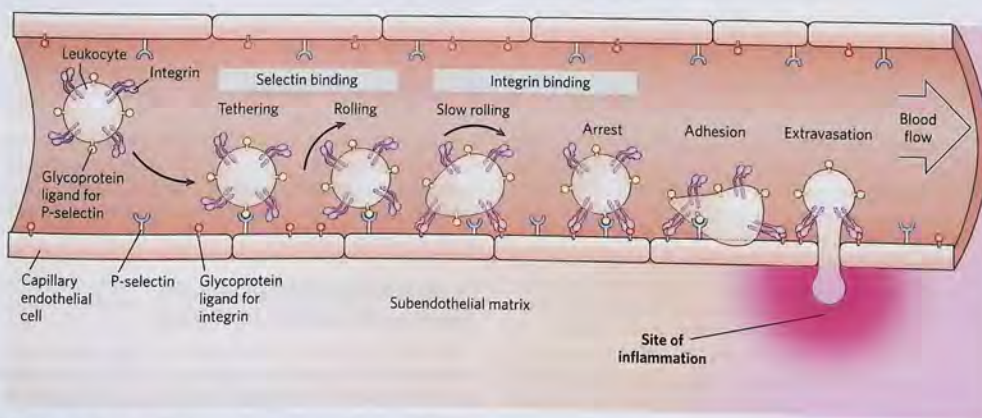


FIGURE 7-29 Role of lectin-ligand interactions in leukocyte movement to the site of an infection or injury. A leukocyte circulating through a capillary is slowed by transient interactions between P-selectin molecules in the plasma membrane of the capillary endothelial cells and glycoprotein ligands for P-selectin on the leukocyte surface. As the leukocyte interacts with successive P-selectin molecules, it rolls along

the capillary surface. Near a site of inflammation, stronger interactions between integrin in the leukocyte surface and its ligand in the capillary surface lead to tight adhesion. The leukocyte stops rolling and, under the influence of signals from the site of inflammation (such as sialyl Lewis x), escapes through the capillary wall as it moves toward the region of inflammation.

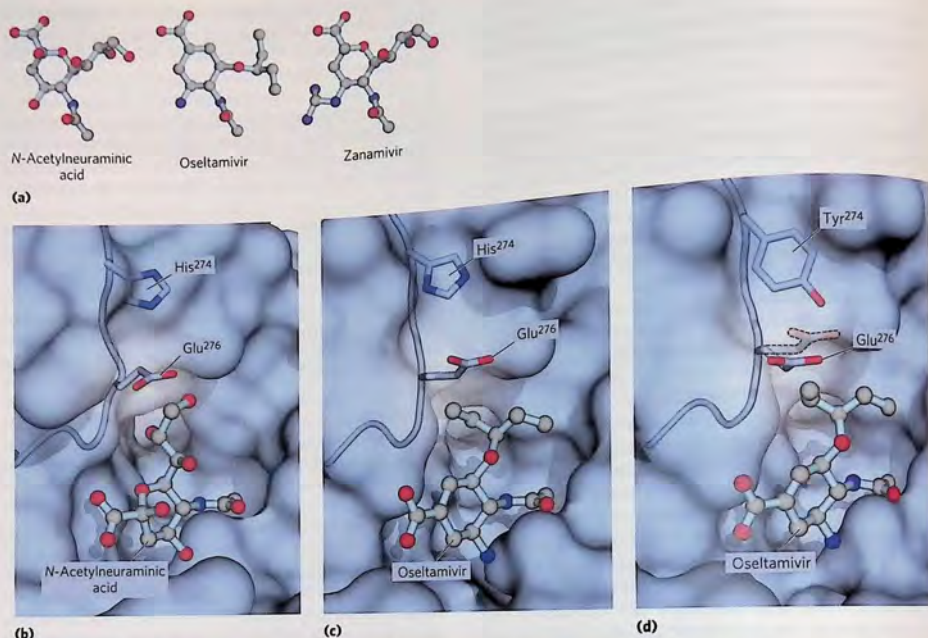


FIGURE 7-30 Binding site on influenza neuraminidase for *N*-acetylneuraminic acid and an antiviral drug. (a) The normal binding ligand for influenza neuraminidase is a sialic acid, *N*-acetylneuraminic acid. The drugs oseltamivir and zanamivir occupy the same site on the enzyme, competitively inhibiting it and blocking viral release from the host cell. (b) The normal interaction with *N*-acetylneuraminic acid in the binding site. (c) Oseltamivir can fit into this site by pushing a nearby Glu residue

out of the way. (d) A mutation in the influenza virus's gene for neuraminidase replaces a His near this Glu residue with the larger side chain of a Tyr. Now, oseltamivir is not as effective at pushing the Glu out of its way, and the drug binds much less well to the binding site, making the mutant virus effectively resistant to oseltamivir. [Data from (b) PDB ID 2BAT, J. N. Varghese et al., *Proteins* 14:327, 1992; (c) PDB ID 2HU4, R. J. Russell et al., *Nature* 443:45, 2006; (d) PDB ID 3C12, J. Collins et al., *Nature* 453:1258, 2008.]

Some of the most devastating of the human parasitic diseases, widespread in much of the developing world, are caused by eukaryotic microorganisms that display unusual surface oligosaccharides, which in some cases are known to be protective for the parasites. These organisms include the trypanosomes, responsible for African sleeping sickness (see Box 6-1) and Chagas disease; *Plasmodium falciparum*, the malaria parasite; and *Entamoeba histolytica*, the causative agent of amoebic dysentery. The prospect of finding drugs that interfere with the synthesis of these unusual oligosaccharide chains, and therefore with the replication of the parasites, has inspired much recent work on the biosynthetic pathways of these oligosaccharides. ■

Lectins also act intracellularly, in sorting proteins for transportation to specific cellular compartments (see

Chapter 27). For example, **P7** an oligosaccharide containing mannose 6-phosphate, recognized by a lectin, acts as a molecular "ZIP code" that tags newly synthesized proteins in the Golgi complex for transfer to the lysosome (see Fig. 27-41).

Lectin-Carbohydrate Interactions Are Highly Specific and Often Multivalent

The high density of information in the structure of oligosaccharides provides a sugar code with an essentially unlimited number of unique "words" small enough to be read by a single protein. In their carbohydrate-binding sites, lectins have a subtle molecular complementarity that allows interaction only with their correct carbohydrate binding partners. Often a divalent metal ion such as Ca^{2+} or Mn^{2+}

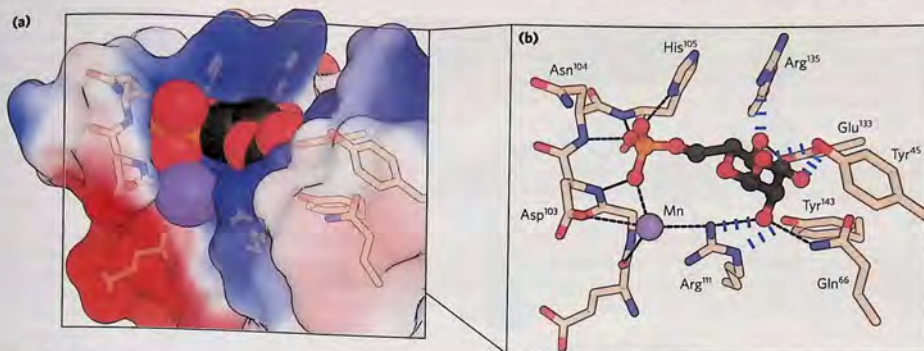


FIGURE 7-31 Details of a lectin-carbohydrate interaction. (a) Structure of the bovine mannose 6-phosphate receptor complexed with mannose 6-phosphate. The protein is represented as a surface contour image, showing the surface as predominantly negatively charged (red) or positively charged (blue). Mannose 6-phosphate is shown as a space-filling model; a manganese ion is shown as a violet sphere. (b) An enlarged view of the binding site. Mannose

6-phosphate is hydrogen-bonded to Arg¹¹¹ and coordinated with the manganese ion (shown smaller than its van der Waals radius, for clarity). Each hydroxyl group of mannose is hydrogen-bonded to the protein. The His¹⁰⁵ hydrogen-bonded to a phosphate oxygen of mannose 6-phosphate may be the residue that, when protonated at low pH, causes the receptor to release mannose 6-phosphate into the lysosome. [Data from PDB ID 1M6P; D. L. Roberts et al., Cell 93:639, 1998.]

is part of the binding site. Lectin-ligand interactions can have extraordinarily high specificity. The affinity between an oligosaccharide and an individual carbohydrate binding domain (CBD) of a lectin is sometimes modest (micromolar to millimolar K_d values), but the effective affinity is often greatly increased by lectin multivalency, in which a single lectin molecule has multiple CBDs. In a cluster of oligosaccharides—as is commonly found on a membrane surface, for example—each oligosaccharide can engage one of the lectin's CBDs, strengthening the interaction. When cells express multiple lectin receptors, the avidity of the interaction can be very high, enabling highly cooperative events such as cell attachment and rolling (Fig. 7-29).

X-ray crystallographic studies of the structure of the lectin that is the receptor for mannose 6-phosphate reveal details that explain the specificity of the binding and the role of a divalent cation in the lectin-sugar interaction (Fig. 7-31). When the protein tagged with mannose 6-phosphate reaches the lysosome (which has a lower internal pH than the Golgi complex), the receptor loses its affinity for mannose 6-phosphate and the tagged protein is released into the lysosomal matrix.

In addition to such highly specific interactions, there are more general interactions that contribute to the binding of many carbohydrates to their lectins. For example, many sugars have a more polar side and a less polar side (Fig. 7-32); the more polar side hydrogen-bonds with the lectin, while the less polar side undergoes interactions with nonpolar amino acid residues through the hydrophobic effect. The sum of all these interactions produces high-affinity binding and high specificity of lectins for their carbohydrate ligands. This represents

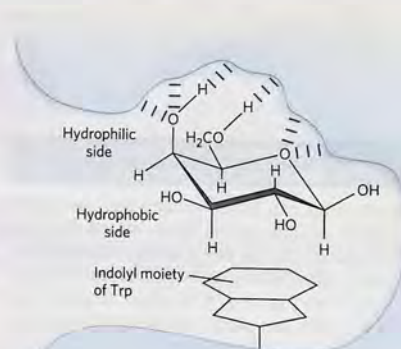


FIGURE 7-32 Interactions of sugar residues due to the hydrophobic effect. Sugar units such as galactose have a more polar side (the top of the chair as shown here, with the ring oxygen and several hydroxyls), available to hydrogen-bond with the lectin, and a less polar side that can interact with nonpolar side chains in the protein, such as the indole ring of Trp residues, through the hydrophobic effect. [Information from a figure provided by Dr. C.-W. von der Lieth, Heidelberg; H.-J. Gabius, *Naturwissenschaften* 87:108, 2000, Fig. 6.]

a kind of information transfer that is clearly central in many processes within and between cells. **Figure 7-33** summarizes some of the biological interactions mediated by the sugar code.

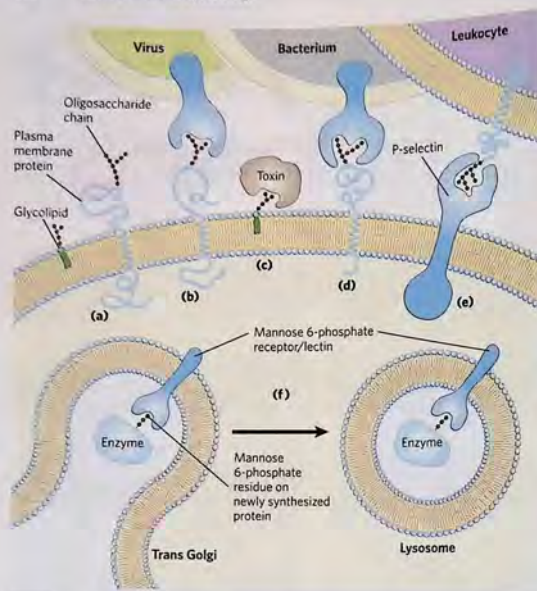


FIGURE 7-33 Role of oligosaccharides in recognition events at the cell surface and in the endomembrane system. (a) Oligosaccharides with unique structures (represented as strings of red hexagons) are components of a variety of glycoproteins or glycolipids on the outer surface of plasma membranes. Their oligosaccharide moieties are bound by extracellular lectins with high specificity and affinity. (b) Viruses that infect animal cells, such as the influenza virus, bind to cell surface glycoproteins as the first step in infection. (c) Bacterial toxins, such as the cholera and pertussis toxins, bind to a surface glycolipid before entering a cell. (d) Some bacteria adhere to and then colonize or infect animal cells. (e) Selectins (lectins) in the plasma membrane of certain cells mediate cell-cell interactions, such as those of leukocytes with the endothelial cells of the capillary wall at an infection site. (f) The mannose 6-phosphate receptor/lectin of the trans Golgi complex binds to the oligosaccharide of lysosomal enzymes, targeting them for transfer into the lysosome. [Information from N. Sharon and H. Lili, *Sci. Am.* 268 (January) 82, 1993]

SUMMARY 7.4 Carbohydrates as Informational Molecules: The Sugar Code

- Monosaccharides can be assembled into an almost limitless variety of oligosaccharides, which differ in the stereochemistry and position of glycosidic bonds, the type and orientation of substituent groups, and the number and type of branches. Glycans are far more information-dense than nucleic acids or proteins.
- The sugar specificity of many plant lectins makes them powerful laboratory reagents in glycobiology. Receptor/lectins in the surface of hepatocytes recognize glycoproteins that have lost their terminal sialic acid residue and mediate the normal uptake (by hepatocytes) and destruction of blood cells and of certain circulating glycoprotein hormones.
- Bacterial and viral pathogens and some eukaryotic parasites adhere to their animal cell targets through binding of lectins on the pathogens to oligosaccharides on the target cell surface, or vice versa. Interactions are often polyvalent.
- Structural studies of lectin-sugar complexes show the detailed complementarity between the two molecules, which accounts for the strength and specificity of lectin interactions with carbohydrates.

7.5 Working with Carbohydrates

Oligosaccharide analysis is complicated by the fact that, unlike nucleic acids and proteins, oligosaccharides can be branched and are joined by a variety of linkages. The high charge density of many oligosaccharides and polysaccharides, and the relative lability of the sulfate esters in glycosaminoglycans, present further difficulties.

For simple, linear polymers such as amylose, the positions of the glycosidic bonds are determined by the classical method of exhaustive methylation: treating the intact polysaccharide with methyl iodide in a strongly basic medium to convert all free hydroxyls to acid-stable methyl ethers, then hydrolyzing the methylated polysaccharide in acid. The only free hydroxyls in the monosaccharide derivatives so produced are those that were involved in glycosidic bonds. To determine the sequence of monosaccharide residues, including any branches that are present, exoglycosidases of known specificity are used to remove residues one at a time from the nonreducing end(s). The known specificity of these exoglycosidases often allows deduction of the position and stereochemistry of the linkages.

For analysis of the oligosaccharide moieties of glycoproteins and glycolipids, the oligosaccharides are released by purified enzymes—glycosidases that specifically cleave O- or N-linked oligosaccharides, or lipases

that remove lipid head groups. Alternatively, *O*-linked glycans can be released from glycoproteins by treatment with hydrazine.

The resulting mixtures of carbohydrates are resolved into their individual components by a variety of methods (Fig. 7-34), including the same techniques used in protein and amino acid separation: fractional precipitation by solvents, and ion-exchange and size-exclusion chromatography (see Fig. 3-17). Highly purified lectins, attached covalently to an insoluble support, can be used in affinity chromatography of carbohydrates.

Hydrolysis of oligosaccharides and polysaccharides in strong acid yields a mixture of monosaccharides, which can be identified and quantified by chromatographic techniques to yield the overall composition of the polymer.

Oligosaccharide analysis relies heavily on mass spectrometry (see Figs 3-28 and 3-29) and high-resolution NMR spectroscopy (see Fig. 4-31). NMR analysis alone,

especially for oligosaccharides of moderate size, can yield much information about sequence, linkage position, and anomeric carbon configuration. Automated procedures and commercial instruments are used for the routine determination of oligosaccharide structure, but the sequencing of branched oligosaccharides joined by more than one type of bond remains a far more formidable task than determining the linear sequences of proteins and nucleic acids.

Another important tool in working with carbohydrates is chemical synthesis, which has proved to be a powerful approach to understanding the biological functions of glycosaminoglycans and oligosaccharides. The chemistry involved in such syntheses is difficult, but carbohydrate chemists can now synthesize short segments of almost any glycosaminoglycan, with correct stereochemistry, chain length, and sulfation pattern, and oligosaccharides significantly more complex than those

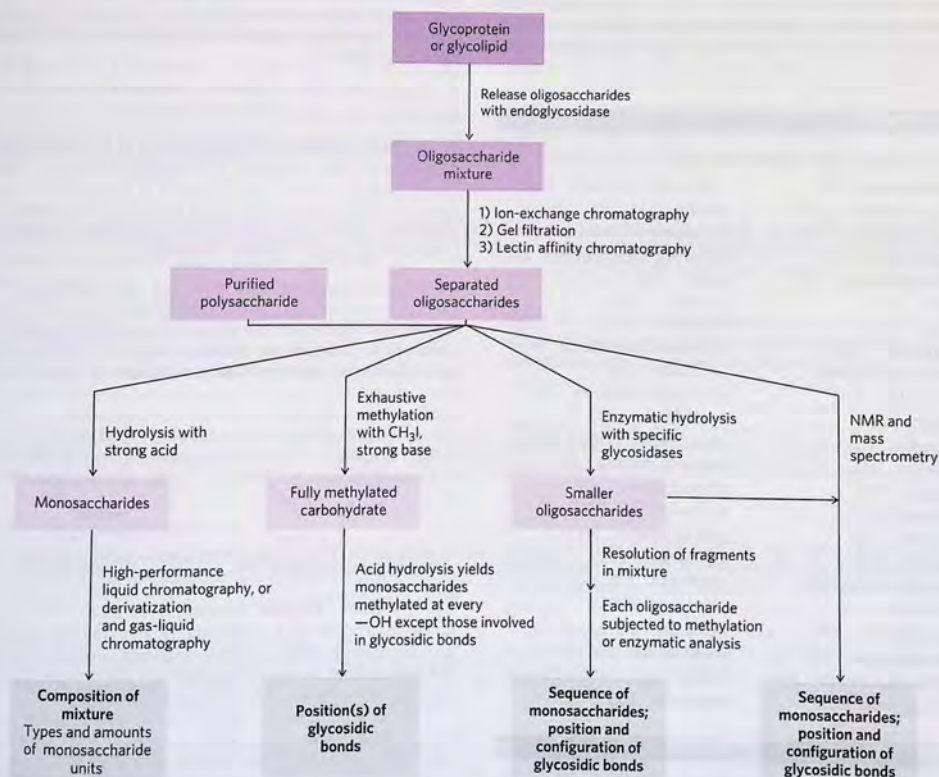


FIGURE 7-34 Methods of carbohydrate analysis. A carbohydrate purified in the first stage of the analysis often requires all four analytical routes for its complete characterization.

shown in Figure 7-27. Solid-phase oligosaccharide synthesis is based on the same principles (and has the same advantages) as peptide synthesis (see Fig. 3-30), but requires a set of tools unique to carbohydrate chemistry: blocking groups and activating groups that allow the synthesis of glycosidic linkages with the correct hydroxyl group. Synthetic approaches of this type currently represent an area of great interest, because it is difficult to purify defined oligosaccharides in adequate quantities from natural sources.

SUMMARY 7.5 Working with Carbohydrates

■ Establishing the complete structure of oligosaccharides and polysaccharides is a more complex problem than protein and nucleic acid analysis. Traditional chemical and enzymatic approaches, as well as mass spectrometry and high-resolution NMR spectroscopy, applicable to small samples of carbohydrate, yield essential information about sequence, configuration at anomeric and other carbons, and positions of glycosidic bonds. Solid-phase synthetic methods yield defined oligosaccharides that are of great value in exploring lectin-oligosaccharide interactions and may prove clinically useful.

KEY TERMS

Terms in bold are defined in the glossary.

carbohydrate 229	hemoglobin glycation 239
monosaccharide 229	glycan 241
oligosaccharide 229	homopolysaccharide 241
disaccharide 229	heteropolysaccharide 241
polysaccharide 229	starch 242
aldose 230	glycogen 242
ketose 230	cellulose 243
enantiomer 232	extracellular matrix (ECM) 244
Fischer projection formulas 232	glycosaminoglycan 245
epimers 232	hyaluronan 245
hemiacetal 234	chondroitin sulfate 245
hemiketal 234	heparan sulfate 246
anomers 234	glyconjugate 247
anomeric carbon 234	proteoglycan 247
pyranose 234	glycoprotein 248
furanose 234	glycolipid 248
Haworth perspective formulas 235	syndecan 248
mutarotation 235	glypican 248
reducing sugar 237	glycomics 252
O-glycosidic bonds 237	ganglioside 253
reducing end 237	lectin 254
	selectins 255

PROBLEMS

1. **Sugar Alcohols** In the monosaccharide derivatives known as sugar alcohols, the carbonyl oxygen is reduced to a hydroxyl group. For example, D-glyceraldehyde can be reduced to glycerol. Why can the sugar alcohol glycerol no longer be designated D or L?

2. **Recognizing Epimers** Using Figure 7-3, identify the epimers of (a) D-allose, (b) D-glucose, and (c) D-ribose at C-2, C-3, and C-4.

3. **Configuration and Conformation** Which bond(s) in α -D-glucose must be broken to change its configuration to β -D-glucose? Which bond(s) must be broken to convert D-glucose to D-mannose? Which bond(s) must be broken to convert one "chair" form of D-glucose to the other?

4. **Sugar Structures** Compare and contrast the structural features of each pair: (a) Cellulose and glycogen (b) D-glucose and D-fructose (c) Maltose and sucrose

5. **Haworth Structures** Draw the Haworth perspective formulas for α -D-mannose and β -D-galactose.

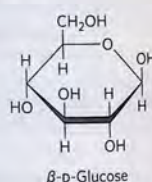
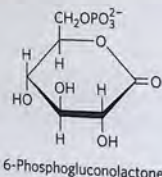
6. **Reducing Sugars** Draw the structural formula for α -D-glucosyl-(1 $\rightarrow$ 6)-D-mannosamine, and circle the part of this structure that makes the compound a reducing sugar.

7. **Hemoglobin Glycation** The measurement of glycated hemoglobin (the HbA1c level) to monitor the regulation of blood glucose (Box 7-2) is not reliable in individuals with certain conditions or diseases. Can you suggest an explanation for a lower-than-normal HbA1c level in a patient who definitely is not diabetic?

8. **Hemiacetal and Glycosidic Linkages** Explain the difference between a hemiacetal and a glycoside.

9. **A Taste of Honey** The sweetness of honey gradually decreases at a high temperature. Also, high-fructose corn syrup (a commercial product in which much of the glucose in corn syrup is converted to fructose) is used for sweetening cold drinks but not hot drinks. What chemical property of fructose could account for both of these observations?

10. **Glucuronolactone and Glucose Oxidation States** The cyclic glucose derivative 6-phosphogluconolactone is an intermediate in the pentose phosphate pathway (discussed in Chapter 14). Compare the oxidation state of C-1 for the cyclic form of both glucuronolactone and β -D-glucose.



11. **Invertase "Inverts" Sucrose** As sweet as sucrose is, an equimolar mixture of its constituent monosaccharides, D-glucose and D-fructose, is sweeter. Besides enhancing sweetness, fructose has hygroscopic properties that improve the texture of foods reducing crystallization and increasing moisture.

In the food industry, hydrolyzed sucrose is called invert sugar, and the yeast enzyme that hydrolyzes it is called invertase. The hydrolysis reaction is generally monitored by measuring the specific rotation of the solution, which is positive (+66.4°) for sucrose but becomes negative (inverts) as more D-glucose (specific rotation = +52.7°) and D-fructose (specific rotation = -92°) form.

From what you know about the chemistry of the glycosidic bond, how would you hydrolyze sucrose to invert sugar nonenzymatically in a home kitchen?

12. Manufacture of Liquid-Filled Chocolates The manufacture of chocolates containing a liquid center is an interesting application of enzyme engineering. The flavored liquid center consists largely of an aqueous solution of sugars rich in fructose to provide sweetness. The technical dilemma is the following: the chocolate coating must be prepared by pouring hot, melted chocolate over a solid (or almost solid) core, yet the final product must have a liquid, fructose-rich center. Suggest a way to solve this problem. (Hint: Sucrose is much less soluble than a mixture of glucose and fructose.)

13. Anomers of Sucrose? Lactose exists in two anomeric forms, but no anomeric forms of sucrose have been reported. Why?

14. Gentiobiose Gentiobiose (β -Glc(β 1 $\rightarrow$ 6) β -Glc) is a disaccharide found in some plant glycosides. Draw the Haworth structure of gentiobiose based on its abbreviated name. Is it a reducing sugar? Does it undergo mutarotation?

15. Identifying Reducing Sugars Is *N*-acetyl- β -D-glucosamine (Fig. 7-9) a reducing sugar? What about D-glucosamine? Is the disaccharide GlcN(α 1 $\rightarrow$ 1 α)Glc a reducing sugar?

16. Physical Properties of Cellulose and Glycogen The almost pure cellulose obtained from the seed threads of *Gossypium* (cotton) is tough, fibrous, and completely insoluble in water. In contrast, glycogen obtained from muscle or liver disperses readily in hot water to make a turbid solution. Despite their markedly different physical properties, both substances are (1 $\rightarrow$ 4)-linked D-glucose polymers of comparable molecular weight. What structural features of these two polysaccharides underlie their different physical properties? Suggest possible biological advantages of their respective properties.

17. Dimensions of a Polysaccharide Compare the dimensions of a molecule of cellulose and a molecule of amylose, each of M_r 200,000.

18. Growth Rate of Bamboo The stems of bamboo, a tropical grass, can grow at the phenomenal rate of 0.3 m/day under optimal conditions. Given that the stems are composed almost entirely of cellulose fibers oriented in the direction of growth, calculate the number of sugar residues per second that must be added enzymatically to growing cellulose chains to account for the growth rate. Each D-glucose unit contributes $\sim$ 0.5 nm to the length of a cellulose molecule.

19. Glycoproteins versus Proteoglycans Which characteristics describe glycoproteins and which describe proteoglycans?

- (a) Exclusively located at the cell surface and in the extracellular matrix
- (b) May contain *N*-linked glycosidic bonds
- (c) Found in Golgi complexes, secretory granules, and lysosomes
- (d) Include the heparan sulfate family
- (e) Sulfated glycosaminoglycan chains can only be covalently linked to a Ser residue
- (f) Form highly specific sites for recognition and high-affinity binding by lectins

20. Relative Stability of Two Conformers Explain why the two structures shown in Figure 7-16 are so different in energy (stability). Hint: See Figure 1-21.

21. Volume of Chondroitin Sulfate in Solution One critical function of chondroitin sulfate is to act as a lubricant in skeletal joints by creating a gel-like medium that is resilient to friction and shock. This function seems to be related to a distinctive property of chondroitin sulfate: the volume occupied by the molecule is much greater in solution than in the dehydrated solid. Why is the volume so much larger in solution?

22. Heparin Interactions Heparin, a highly negatively charged glycosaminoglycan, is used clinically as an anticoagulant. It acts by binding several plasma proteins, including antithrombin III, an inhibitor of blood clotting. The 1:1 binding of heparin to antithrombin III seems to cause a conformational change in the protein that greatly increases its ability to inhibit clotting. What amino acid residues of antithrombin III are likely to interact with heparin?

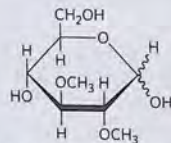
23. Permutations of a Trisaccharide Three different hexoses (A, B, and C) can be combined to form a large number of trisaccharides. What structural features of trisaccharides allow so many permutations and combinations?

24. Effect of Sialic Acid on SDS Polyacrylamide Gel Electrophoresis Suppose you have four forms of a protein, all with identical amino acid sequence but containing zero, one, two, or three oligosaccharide chains, each ending in a single sialic acid residue. Draw the gel pattern you would expect when a mixture of these four glycoproteins is subjected to SDS polyacrylamide gel electrophoresis (see Fig. 3-18) and stained for protein. Identify any bands in your drawing.

25. Information Content of Oligosaccharides The carbohydrate portion of some glycoproteins may serve as a cellular recognition site. To perform this function, the oligosaccharide(s) must have the potential to exist in a large variety of forms. Which can produce a greater variety of structures: oligopeptides composed of five different amino acid residues, or oligosaccharides composed of five different monosaccharide residues? Explain.

26. Castor Bean Toxin The seed of the castor bean (*Ricinus communis*) contains large amounts of ricin, a poison that is deadly to animals, including humans. One of the two subunits of this toxin is a lectin that binds terminal *N*-acetyl-galactosamine residues on glycoproteins on the surface of eukaryotic cells, allowing the other subunit to enter the cell and kill it by preventing proteins from being made. Suggest a possible antidote to prevent or reverse ricin-mediated entry of the toxin.

27. Determination of the Extent of Branching in Amylopectin A biochemist wants to determine the amount of branching in amylopectin, defined by the number of (α 1 $\rightarrow$ 6) glycosidic bonds present. First, she treats the sample with methyl iodide, a methylating agent that replaces the hydrogen of every sugar hydroxyl with a methyl group, converting —OH to —OCH_3 . She then hydrolyzes all the glycosidic bonds in the treated sample in aqueous acid and measures the amount of 2,3-di-O-methylglucose formed.



2,3-Di-O-methylglucose

In this representation of 2,3-di-O-methylglucose, the wavy bonds at C-1 indicate that the structure represents both anomers (α and β).

(a) Explain the basis of this procedure for determining the number of (α 1 $\rightarrow$ 6) branch points in amylopectin. What happens to the unbranched glucose residues in amylopectin during the methylation and hydrolysis procedure?

(b) A 258 mg sample of amylopectin that was treated by the methylation and hydrolysis procedure described yielded 12.4 mg of 2,3-di-O-methylglucose. Determine what percentage of the glucose residues in the amylopectin contained an (α 1 $\rightarrow$ 6) branch. (Assume that the average molecular weight of a glucose residue in amylopectin is 162 g/mol and the molecular weight of 2,3-di-O-methylglucose is 208 g/mol.)

DATA ANALYSIS PROBLEM

28. Determining the Structure of ABO Blood Group Antigens

The human ABO blood group system was discovered in 1901, and in 1924 this trait was shown to be inherited at a single gene locus with three alleles. In 1960, W. T. J. Morgan published a paper summarizing what was known at that time about the structure of the ABO antigen molecules. When the paper was published, the complete structures of the A, B, and O antigens were not yet known; this paper is an example of what scientific knowledge looks like "in the making."

In any attempt to determine the structure of an unknown biological compound, researchers must deal with two fundamental problems: (1) If you don't know what it is, how do you know if it is pure? (2) If you don't know what it is, how do you know that your extraction and purification conditions have not changed its structure? Morgan addressed problem 1 through several methods. One method is described in his paper as observing "constant analytical values after fractional solubility tests" (p. 312). In this case, "analytical values" are measurements of chemical composition, melting point, and so forth.

(a) Based on your understanding of chemical techniques, what could Morgan mean by "fractional solubility tests"?

(b) Why would the analytical values obtained from fractional solubility tests of a pure substance be constant, and those of an impure substance not be constant?

Morgan addressed problem 2 by using an assay to measure the immunological activity of the substance present in different samples.

(c) Why was it important for Morgan's studies, and especially for addressing problem 2, that this activity assay be quantitative (measuring a level of activity) rather than simply qualitative (measuring only the presence or absence of a substance)?

The structure of the blood group antigens is shown in Figure 10-13. In his paper, Morgan listed several properties of the three antigens, A, B, and O, that were known at that time (p. 314):

1. Type B antigen has a higher content of galactose than A or O.
2. Type A antigen contains more total amino sugars than B or O.

3. The glucosamine:galactosamine ratio for the A antigen is roughly 1:2; for B, it is roughly 2:5.

(d) Which of these findings is (are) consistent with the known structures of the blood group antigens?

(e) How do you explain the discrepancies between Morgan's data and the known structures?

In later work, Morgan and his colleagues used a clever technique to obtain structural information about the blood group antigens. Enzymes had been found that would specifically degrade the antigens. However, these were available only as crude enzyme preparations, perhaps containing more than one enzyme of unknown specificity. Degradation of the blood type antigens by these crude enzymes could be inhibited by the addition of particular sugar molecules to the reaction. Only sugars that would cause this inhibition. One found in the blood type antigens would cause this inhibition. One enzyme preparation, isolated from the protozoan *Trichomonas foetus*, would degrade all three antigens and was inhibited by the addition of particular sugars. The results of these studies are summarized in the table below, showing the percentage of substrate remaining unchanged when the *T. foetus* enzyme acted on the blood group antigens in the presence of sugars.

Sugar added	Unchanged substrate (%)		
	A antigen	B antigen	O antigen
Control—no sugar	3	1	1
L-Fucose	3	1	100
D-Fucose	3	1	1
L-Galactose	3	1	3
D-Galactose	6	100	1
N-Acetylglucosamine	3	1	1
N-Acetylgalactosamine	100	6	1

For the O antigen, a comparison of the control and L-fucose results shows that L-fucose inhibits the degradation of the antigen. This is an example of product inhibition, in which an excess of reaction product shifts the equilibrium of the reaction, preventing further breakdown of substrate.

(f) Although the O antigen contains galactose, N-acetylglucosamine, and N-acetylgalactosamine, none of these sugars inhibited the degradation of this antigen. Based on these data, is the enzyme preparation from *T. foetus* an endoglycosidase or an exoglycosidase? (Endoglycosidases cut bonds between interior residues; exoglycosidases remove one residue at a time from the end of a polymer.) Explain your reasoning.

(g) Fucose is also present in the A and B antigens. Based on the structure of these antigens, why does fucose fail to prevent their degradation by the *T. foetus* enzyme? What structure would be produced?

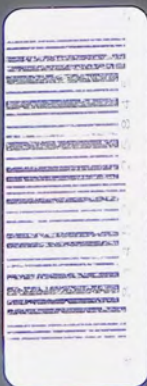
(h) Which of the results in (f) and (g) are consistent with the structures shown in Figure 10-13? Explain your reasoning.

Reference

Morgan, W.T.J. 1960. The Croonian Lecture: a contribution to human biochemical genetics; the chemical basis of blood-group specificity. *Proc. R. Soc. Lond. B Biol. Sci.* 151:308-347.

ON THE COVER:

In this artistic rendering, a protocell composed of a bilayer of fatty acids surrounds an aqueous compartment containing an RNA ribozyme and nucleic acids. Protocells are used by biochemists to explore reactions that could mimic those that led to the origin of life on Earth. Billions of years ago, hydrophobic molecules may have formed cell-like assemblies like protocells in which molecules capable of replicating themselves could be concentrated and protected from the external environment. These likely provided the origins of the membranes, nucleic acids, and catalysts found in today's cells.



Cover image: Janet Iwasa, University of Utah



macmillan
learning

macmillanlearning.com

