Carbohydrates are the most abundant biomolecules on Earth. Each year, photosynthesis converts more than 100 billion metric tons of CO₂ and H₂O into cellulose and other plant products. Certain carbohydrates (sugar and starch) 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. Insoluble carbohydrate polymers serve as structural and protective elements in the cell walls of bacteria and plants and in the connective tissues of animals. Other carbohydrate polymers lubricate skeletal joints and participate in recognition and adhesion between cells. More complex carbohydrate polymers covalently attached to proteins or lipids act as signals that determine the intracellular location or metabolic fate of these hybrid molecules, called glycoconjugates. This chapter introduces the major classes of carbohydrates and glycoconjugates and provides a few examples of their many structural and functional roles.

Carbohydrates are predominantly cyclized polyhydroxy aldehydes or ketones, 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 \(\text{sakcharon}\), 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. Typical is sucrose, or cane sugar, which 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 having three or more units do not occur as free entities but are joined to nonsugar molecules (lipids or proteins) in glycoconjugates.

Sugar polymers occur in a continuous range of sizes. Those containing more than about 20 monosaccharide units are generally called polysaccharides. Polysaccharides may have hundreds or thousands of monosaccharide units. Some polysaccharide molecules, such as cellulose, are linear chains, whereas others, such as glycogen, are branched chains. The plant products starch and cellulose both 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.
Monosaccharides and Disaccharides

The simplest of the carbohydrates, the monosaccharides, are either aldehydes or ketones with one or more hydroxyl groups; the six-carbon monosaccharides glucose and fructose have five hydroxyl groups. The carbon atoms to which hydroxyl groups are attached are often chiral centers, which give rise to the many sugar stereoisomers found in nature. We begin by describing the families of monosaccharides with backbones of three to seven carbons—their structure and 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 the cyclic forms of five- and six-carbon sugars (which predominate in aqueous solution) and creates a new chiral center, adding further stereochemical complexity to this class of compounds. The nomenclature for unambiguously specifying the configuration about each carbon atom in a cyclic form and the means of representing these structures on paper are therefore described in some detail; these will be useful later in our discussion of the metabolism of monosaccharides.

We also introduce here some important monosaccharide derivatives that will be encountered 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. The backbones of common monosaccharide molecules are unbranched carbon chains in which all the carbon atoms are linked by single bonds. In the 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 (i.e., 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. 9–1a).

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. 9–1b), are the most common monosaccharides in nature. The aldopentoses d-ribose and 2-deoxy-d-ribose (Fig. 9–1c) are components of nucleotides and nucleic acids (Chapter 10).
Monosaccharides Have Asymmetric Centers

All the monosaccharides except dihydroxyacetone contain one or more asymmetric (chiral) carbon atoms and thus occur in optically active isomeric forms (pp. 58–61). The simplest aldose, glyceraldehyde, contains one chiral center (the middle carbon atom) and therefore has two different optical isomers, or enantiomers (Fig. 9–2). By convention, one of these two forms is designated the D isomer, the other 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. 9–2).

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 \( 2^4 = 16 \) stereoisomers. The 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. When the hydroxyl group on the reference carbon is on the right in the projection formula, the sugar is the D isomer; when on the left, it is the L isomer. Of the 16 possible aldohexoses, eight are D forms and eight are L. Most of the hexoses of living organisms are D isomers.

Figure 9–3 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. 9–3a). The four- and five-carbon ketoses are designated by
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 these D isomers, the chiral carbon most distant from the carbonyl carbon has the same configuration as the chiral carbon in D-glyceraldehyde. The sugars named in boxes are the most common in nature; we shall encounter these again in this and later chapters.

**Figure 9-3**
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 these D isomers, the chiral carbon most distant from the carbonyl carbon has the same configuration as the chiral carbon in D-glyceraldehyde. The sugars named in boxes are the most common in nature; we shall encounter these again in this and later chapters.
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 red).

The Common Monosaccharides Have Cyclic Structures

For simplicity, we have thus far represented the structures of various aldoses and ketoses as straight-chain forms (Figs. 9–3, 9–4). 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 along the chain. The formation of these ring structures is the result of a general reaction between aldehydes or ketones and alcohols to form derivatives called hemiacetals or hemiketals (Fig. 9–5), which contain an additional asymmetric carbon atom and thus can exist in two stereoisomeric forms. For example, D-glucose exists in solution as an intramolecular hemiacetal in which the free hydroxyl group at C-5 has reacted

![Figure 9–5](image-url)

**Figure 9–5**

Formation of hemiacetals and hemiketals. An aldehyde or ketone can react with an alcohol in a 1:1 ratio to yield a hemiacetal or hemiketal, respectively, creating a new chiral center at the carbonyl carbon. Substitution of a second alcohol molecule produces an acetal or ketal. When the second alcohol is part of another sugar molecule, the bond produced is a glycosidic bond (p. 301).
Part II  Structure and Catalysis

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 \( \alpha \) and \( \beta \) anomers, which differ only in the stereochemistry around the hemiacetal carbon. The interconversion of \( \alpha \) and \( \beta \) anomers is called mutarotation.

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. 9–6). Pyran and furan are shown for comparison.
Isomeric forms of monosaccharides that differ only in their configuration about the hemiacetal or hemiketal carbon atom are called **anomers**. The hemiacetal or carbonyl carbon atom is called the **anomeric carbon**. The α and β anomers of D-glucose interconvert in aqueous solution by a process called **mutarotation**. Thus a solution of α-D-glucose and a solution of β-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 and five-membered ring (glucofuranose) forms.

Ketohexoses also occur in α and β anomeric forms. In these compounds the hydroxyl group on C-5 (or C-6) reacts with the keto group at C-2, forming a furanose (or pyranose) ring containing a hemiketal linkage (Fig. 9–5). D-Fructose forms predominantly the furanose ring (Fig. 9–7); the more common anomer in combined forms or derivatives is β-D-fructofuranose.

**Haworth perspective formulas** like those in Figure 9–7 are commonly used to show the 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. 9–8). Recall from Chapter 3 that two **conformations** of a molecule are interconvertible without the breakage of covalent bonds, but two **configurations** can be interconverted only by breaking a covalent bond—for example, in the case of α and β configurations, the bond involving the ring oxygen atom. The specific three-dimensional conformations of the monosaccharide units are important in determining the biological properties and functions of some polysaccharides, as we shall see.

**Organisms Contain a Variety of Hexose Derivatives**

In addition to simple hexoses such as glucose, galactose, and mannose, there are a number of sugar derivatives in which a hydroxyl group in the parent compound is replaced with another substituent, or a carbon atom is oxidized to a carboxyl group (Fig. 9–9). In glucosamine, galactosamine, and mannosamine, the hydroxyl at C-2 of the parent compound is replaced with an amino group. The amino group is nearly always condensed with acetic acid, as in N-acetylglicosamine. This glucosamine derivative is part of many structural polymers, including those of the bacterial cell wall. Bacterial cell walls also contain a derivative of glucosamine called N-acetylmuramic acid, in which lactic acid (a three-carbon carboxylic acid) is ether-linked to the oxygen at C-3 of N-acetylglicosamine. The substitution of a hydrogen for the hydroxyl group at C-6 of L-galactose or L-mannose produces L-fucose or L-rhamnose, respectively; these deoxy sugars are found in plant polysaccharides and in the complex oligosaccharide components of glycoproteins and glycolipids described later.

When the carbonyl (aldehyde) carbon of glucose is oxidized to the carboxyl level, gluconic acid is produced; other aldoses yield other **aldonic acids**. Oxidation of the carbon at the other end of the carbon chain—C-6

---

**figure 9–8**

Conformational formulas of pyranoses. (a) Two chair forms of the pyranose ring. Substituents on the ring carbons may be either axial (ax), projecting parallel with the vertical axis through the ring, or equatorial (eq), projecting roughly perpendicular to this axis. Generally, substituents in the equatorial positions are less sterically hindered by neighboring substituents, and conformations with their bulky substituents in equatorial positions are favored. Another conformation, the “boat” (not shown), is only seen in derivatives with very bulky substituents. (b) A chair conformation of α-D-glucopyranose.
**Glucose family**

<table>
<thead>
<tr>
<th>β-D-Glucose</th>
<th>β-D-Glucosamine</th>
<th>N-Acetyl-β-D-glucosamine</th>
</tr>
</thead>
<tbody>
<tr>
<td>β-D-Glucose 6-phosphate</td>
<td>Muramic acid</td>
<td>N-Acetylmuramic acid</td>
</tr>
</tbody>
</table>

**Amino sugars**

<table>
<thead>
<tr>
<th>β-D-Glucosamine</th>
<th>β-D-Mannosamine</th>
</tr>
</thead>
</table>

**Deoxy sugars**

<table>
<thead>
<tr>
<th>β-L-Fucose</th>
<th>α-L-Rhamnose</th>
</tr>
</thead>
</table>

**Acidic sugars**

<table>
<thead>
<tr>
<th>β-D-Glucuronate</th>
<th>D-Glucuronate</th>
<th>D-Glucono-δ-lactone</th>
</tr>
</thead>
</table>

| N-Acetylneuraminic acid (sialic acid) |

---

**Figure 9–9**

Some hexose derivatives important in biology. In amino sugars, an —NH₂ group replaces one of the —OH groups in the parent hexose. Substitution of —H for —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.

Of glucose, galactose, or mannose—forms the corresponding uronic acid: glucuronic, galacturonic, or mannnuronic acid. Both aldonic and uronic acid form stable intramolecular esters called lactones (Fig. 9–9, lower left). In addition to these acidic hexose derivatives, one nine-carbon acidic sugar deserves mention: N-acetylneuraminic acid (sialic acid), a derivative of N-acetylmansnosamine, is a component of many glycoproteins and glycolipids in animals. The carboxylic acid groups of the acidic sugar derivatives are ionized at pH 7, and the compounds are therefore correctly named as 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. 9–9). Sugar phosphates are relatively stable at neutral pH and bear a negative charge, which is important for their biological functions.
charge. One effect of sugar phosphorylation within cells is to trap the sugar inside the cell; cells do not in general have plasma membrane transporters for phosphorylated sugars. Phosphorylation also activates sugars for subsequent chemical transformation. Several important phosphorylated derivatives of sugars are discussed in the next chapter.

**Monosaccharides Are Reducing Agents**

Monosaccharides can be oxidized by relatively mild oxidizing agents such as ferric (Fe$^{3+}$) or cupric (Cu$^{2+}$) ion (Fig. 9–10a). The carbonyl carbon is oxidized to a carboxyl group. Glucose and other sugars capable of reducing ferric or cupric ion are called **reducing sugars**. This property is the basis of Fehling’s reaction, a qualitative test for the presence of reducing sugar. By measuring the amount of oxidizing agent reduced by a solution of a sugar, it is also possible to estimate the concentration of that sugar. For many years, this test was used to detect and measure elevated glucose levels in blood and urine in the diagnosis of diabetes mellitus. Now, more sensitive methods for measuring blood glucose employ an enzyme, glucose oxidase (Fig. 9–10b).

![Monosaccharide and Disaccharide Structures](image)

**Disaccharides Contain a Glycosidic Bond**

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 reacts with the anomeric carbon of the other (Fig. 9–11). This reaction represents the formation of an acetal from a hemiacetal (such as glucopyranose) and an alcohol (a hydroxyl group of a second sugar molecule) (Fig. 9–5). When an anomeric carbon participates
in a glycosidic bond, it cannot be oxidized by cupric or ferric ion. The sugar containing the anomeric carbon atom cannot exist in linear form and no longer acts as a reducing sugar. In describing disaccharides or polysaccharides, the end of a chain with a free anomeric carbon (i.e., not involved in a glycosidic bond) is commonly called the reducing end. 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. Another type of glycosidic bond joins the anomeric carbon of a sugar to a nitrogen atom in glycoproteins (see Fig. 9–25). These N-glycosyl bonds are also found in all nucleotides (Chapter 10).

The disaccharide maltose (Fig. 9–11) 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 free anomeric carbon (C-1 of the glucose residue on the right in Fig. 9–11) can be oxidized, maltose is a reducing disaccharide. 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.

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 with its nonreducing end to the left, and the name is “built up” in the following order. (1) The configuration (α or β) at the anomeric carbon joining the first monosaccharide unit (on the left) to the second is given. (2) The nonreducing residue is named. To distinguish five- and six-membered ring structures, “furano” or “pyrano” is inserted into the name. (3) The two carbon atoms joined by the glycosidic bond are indicated in parentheses, 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) The second residue is named. If there is a third residue, the second glycosidic bond is described next, by the same conventions. (To shorten the description of complex polysaccharides, three-letter abbreviations for each monosaccharide are often used, as given in Table 9–1.) 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 will 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(α1→4)Glc.

**table 9–1**

<table>
<thead>
<tr>
<th>Abbreviations for Common Monosaccharides and Some of Their Derivatives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe</td>
</tr>
<tr>
<td>Abequose</td>
</tr>
<tr>
<td>Arabinose</td>
</tr>
<tr>
<td>Fucose</td>
</tr>
<tr>
<td>Galactose</td>
</tr>
<tr>
<td>Glucose</td>
</tr>
<tr>
<td>Mannose</td>
</tr>
<tr>
<td>Rhamnose</td>
</tr>
<tr>
<td>Ribose</td>
</tr>
<tr>
<td>Xylose</td>
</tr>
</tbody>
</table>
The disaccharide lactose (Fig. 9–12), which yields D-galactose and D-glucose on hydrolysis, occurs naturally only in milk. The anomeric carbon of the glucose residue is available for oxidation, and thus lactose is a reducing disaccharide. Its abbreviated name is Gal(\(\beta_1\rightarrow4\))Glc. Sucrose (table sugar) is a disaccharide of glucose and fructose. It is formed by plants but not by higher 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. 9–12). Sucrose is therefore not a reducing sugar. Nonreducing disaccharides are named as glycosides; the positions joined are the anomeric carbons. In the abbreviated nomenclature, a double-headed arrow connects the symbols specifying the anomeric carbons and their configurations. For example, the abbreviated name of sucrose is either Glc(\(\alpha_1\rightarrow2\))Fru or Fru(\(\beta_2\rightarrow1\alpha\))Glc. Sucrose is 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, Glc(\(\alpha_1\rightarrow1\alpha\))Glc (Fig. 9–12), is a disaccharide of D-glucose that, like sucrose, is a nonreducing sugar. It is a major constituent of the circulating fluid (hemolymph) of insects, in which it serves as an energy storage compound.

**Polysaccharides**

Most carbohydrates found in nature occur as polysaccharides, polymers of medium to high molecular weight. Polysaccharides, also called glycans, differ from each other in the identity of their recurring monosaccharide units, in the length of their chains, in the types of bonds linking the units, and in the degree of branching. **Homopolysaccharides** contain only a single type of monomer; **heteropolysaccharides** contain two or more different kinds (Fig. 9–13). Some homopolysaccharides serve as storage forms of monosaccharides 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. 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.

Unlike proteins, polysaccharides generally do not have definite molecular weights. This difference is a consequence of the mechanisms of assembly of the two types of polymers. 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.

**Starch and Glycogen Are Stored Fuels**

The most important storage polysaccharides are starch in plant cells and glycogen in animal cells. Both polysaccharides occur intracellularly as large clusters or granules (Fig. 9–14). 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, but it is especially abundant in tubers, such as potatoes, and in seeds, such as those of maize (corn).

**Starch** contains two types of glucose polymer, amylose and amylpectin (Fig. 9–15). The former consists of long, unbranched chains of D-glucose residues connected by (α1→4) linkages. Such chains vary in molecular weight from a few thousand to over a million. Amylopectin also has a high molecular weight (up to 100 million) but unlike amylose is highly branched. The glycosidic linkages joining successive glucose residues in amylpectin chains are (α1→4); the branch points (about one per 24 to 30 residues) are (α1→6) linkages.

**Glycogen** is the main storage polysaccharide of animal cells. Like amylpectin, glycogen is a polymer of (α1→4)-linked subunits of glucose, with (α1→6)-linked branches, but glycogen is more extensively branched (on average, one branch per 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 granules (Fig. 9–14b), which are themselves clusters of smaller granules composed of single, highly branched glycogen molecules with an average molecular weight of several million. Such glycogen granules also contain, in tightly bound form, the enzymes responsible for the synthesis and degradation of glycogen.

---

**figure 9–14**

Electron micrographs of starch and glycogen granules. (a) Large starch granules in a single chloroplast. Starch is made in the chloroplast from D-glucose formed photosynthetically. (b) Glycogen granules in a hepatocyte. These granules form in the cytosol and are much smaller (~0.1 μm) than starch granules (~1.0 μm).
Because each branch in glycogen ends with a nonreducing sugar (one without a free anomeric carbon), a glycogen molecule has as many nonreducing ends as it has branches, 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? It has been calculated that hepatocytes store glycogen equivalent to a glucose concentration of 0.4 M. The actual concentration of glycogen, which is insoluble and contributes little to the osmolarity of the cytosol, is about 0.01 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. 4–11). Furthermore, with an intracellular glucose concentration of 0.4 M and an external concentration of about 5 mM (the concentration in the blood of a mammal), the free-energy change for glucose uptake into cells against this very high concentration gradient would be prohibitively large.
The three-dimensional structure of starch is shown in Figure 9–16, and it is compared with the structure of cellulose below.

**Cellulose and Chitin Are Structural Homopolysaccharides**

Cellulose, a fibrous, tough, 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 and the main chains of amylopectin and glycogen, 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: in cellulose the glucose residues have the \( \beta \) configuration (Fig. 9–17), whereas in amylose, amylopectin, and glycogen the glucose is in the \( \alpha \) configuration. The glucose residues in cellulose are linked by \((\beta 1\rightarrow 4)\) glycosidic bonds. This difference gives cellulose and amylose very different three-dimensional structures and physical properties.

The three-dimensional structure of carbohydrate macromolecules is based on the same principles as those underlying 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. Because polysaccharides have so many hydroxyl groups, hydrogen bonding has an especially important influence on their structures. Polymers of \( \beta \)-D-glucose, such as cellulose, can be represented as a series of rigid pyranose rings connected by an oxygen atom bridging two carbon atoms (the glycosidic bond). There is free rotation about both C—O bonds linking the two residues (Fig. 9–17a). The most stable conformation for the polymer is that in which each chair is turned 180° relative to its neighbors, yielding a straight, extended chain. 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. 9–17b). The tensile strength of cellulose has made it a useful substance to civilizations for millenia. Many manufactured products, including paper, cardboard, rayon, insulating tiles, and other packing and building 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.
In contrast to the straight fibers produced by (β1→4)-linked polymers such as cellulose, the most favorable conformation for (α1→4)-linked polymers of D-glucose, such as starch and glycogen, is a tightly coiled helical structure stabilized by hydrogen bonds (Fig. 9–16). glycogen and starch ingested in the diet are hydrolyzed by α-amylases, enzymes in saliva and intestinal secretions that break (α1→4) glycosidic bonds between glucose units. Most animals cannot use cellulose as a fuel source because they lack an enzyme to hydrolyze the (β1→4) linkages. Termites readily digest cellulose (and therefore wood), but only because their intestinal tract harbors a symbiotic microorganism, Trichonympha, that secretes cellulase, an enzyme that hydrolyzes (β1→4) linkages between glucose units. Wood-rot fungi and bacteria also produce cellulase. The only vertebrates able to use cellulose as food are cattle and other ruminants (sheep, goats, camels, giraffes). The extra stomach compartment (rumen) of a ruminant teems with bacteria and protists that secrete cellulase.

Chitin is a linear homopolysaccharide composed of N-acetyl-D-glucosamine residues in β linkage (Fig. 9–18). The only chemical difference from cellulose is the replacement of the hydroxyl group at C-2 with an acetylated amino group. Chitin forms extended fibers similar to those of cellulose, and like cellulose is indigestible by vertebrate animals. 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.

**Bacterial Cell Walls Contain Peptidoglycans**

The rigid component of bacterial cell walls is a heteropolymer of alternating (β1→4)-linked N-acetylglucosamine and N-acetylmuramic acid residues (Fig. 9–19). 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 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 $(\beta_1\rightarrow4)$ glycosidic bond between N-acetylg glucosamine and N-acetylmuramic acid. Lysozyme is notably present in tears, presumably as a defense against bacterial infections of the eye. It is also produced by certain bacterial viruses to ensure their release from the host bacterial cell, an essential step of the viral infection cycle. Penicillin and related antibiotics kill bacteria by preventing the synthesis of cross-links, leaving the cell wall too weak to resist osmotic lysis (see Box 20–1).

**Glycosaminoglycans Are Components of the Extracellular Matrix**

The extracellular space in the tissues of multicellular animals is filled with a gel-like material, the **extracellular matrix**, also called ground substance, which holds the cells together and provides a porous pathway for the diffusion of nutrients and oxygen to individual cells. The extracellular matrix is composed of an interlocking meshwork of heteropolysaccharides and fibrous proteins such as collagen, elastin, fibronectin, and laminin. The heteropolysaccharides, called **glycosaminoglycans**, are a family of linear polymers composed of repeating disaccharide units (Fig. 9–20). One of the two monosaccharides is always either $N$-acetylg glucosamine or $N$-acetyl-
galactosamine; the other is in most cases a uronic acid, usually d-glucuronic or L-iduronic acid. In some glycosaminoglycans, one or more of the hydroxyls of the amino sugar is esterified with sulfate. 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. The specific patterns of sulfated and nonsulfated sugar residues in glycosaminoglycans provide for specific recognition by a variety of protein ligands that bind electrostatically to these molecules. Glycosaminoglycans are attached to extracellular proteins to form proteoglycans (discussed below).

The glycosaminoglycan hyaluronic acid (hyaluronate at physiological pH) contains alternating residues of d-glucuronic acid and N-acetylgalactosamine (Fig. 9–20). With up to 50,000 repeats of the basic disaccharide unit, hyaluronates have molecular weights greater than 1 million; they form clear, highly viscous 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”; hyaluronates can have a glassy

<table>
<thead>
<tr>
<th>Glycosaminoglycan</th>
<th>Repeating disaccharide</th>
<th>Number of disaccharides per chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyaluronate</td>
<td>GlicA</td>
<td>GlcNAc</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>COO^-</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>COO^-</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Chondroitin 4-sulfate</td>
<td>GlicA</td>
<td>GalNAc4SO3</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>COO^-</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>COO^-</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td>Keratan sulfate</td>
<td>Gal</td>
<td>GlcNAc6SO3</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>COO^-</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td>OH</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>H</td>
</tr>
</tbody>
</table>

**Figure 9–20**
Repeating units of some common glycosaminoglycans of extracellular matrix. The molecules are copolymers of alternating uronic acid and amino sugar residues, with sulfate esters in any of several positions. The ionized carboxylate and sulfate groups (red) give these polymers their characteristic highly negative charge.
or translucent appearance). Hyaluronate is also an essential component of the extracellular matrix of cartilage and tendons, to which it contributes tensile strength and elasticity as a result of strong interactions with other components of the matrix. Hyaluronidase, an enzyme secreted by some pathogenic (disease-causing) bacteria, can hydrolyze the glycosidic linkages of hyaluronate, rendering tissues more susceptible to bacterial invasion. In many organisms, a similar enzyme in sperm hydrolyzes an outer glycosaminoglycan coat around the ovum, allowing sperm penetration.

Other glycosaminoglycans differ from hyaluronate in two respects: they are generally much shorter polymers and they are covalently linked to specific proteins (proteoglycans, discussed below). Chondroitin sulfate (Greek chondros, “cartilage”) contributes to the tensile strength of cartilage, tendons, ligaments, 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. 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 of dead cells: horn, hair, hoofs, nails, and claws. Heparin (Greek hépar, “liver”) is a natural anticoagulant made in mast cells and released into the blood, where it inhibits coagulation by binding to and stimulating the anticoagulant protein antithrombin III. The interaction is strongly electrostatic; heparin has the highest negative charge density of any known biological macromolecule. Purified heparin is routinely added to blood samples obtained for clinical analysis, and to blood donated for transfusion, to prevent clotting.

Table 9–2 summarizes the composition, properties, roles, and occurrence of the polysaccharides described in this section.

<table>
<thead>
<tr>
<th>Polymer</th>
<th>Type*</th>
<th>Repeating unit†</th>
<th>Size (number of monosaccharide units)</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starch</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amylose</td>
<td>Homo-</td>
<td>(α1→4)Glc, linear</td>
<td>50–5,000</td>
<td>Energy storage: in plants</td>
</tr>
<tr>
<td>Amylopectin</td>
<td>Homo-</td>
<td>(α1→4)Glc, with (α1→6)Glc branches every 24 to 30 residues</td>
<td>Up to 10^6</td>
<td></td>
</tr>
<tr>
<td>Glycogen</td>
<td>Homo-</td>
<td>(α1→4)Glc, with (α1→6)Glc branches every 8 to 12 residues</td>
<td>Up to 50,000</td>
<td>Energy storage: in bacteria and animal cells</td>
</tr>
<tr>
<td>Cellulose</td>
<td>Homo-</td>
<td>(β1→4)Glc</td>
<td>Up to 15,000</td>
<td>Structural: in plants, gives rigidity and strength to cell walls</td>
</tr>
<tr>
<td>Chitin</td>
<td>Homo-</td>
<td>(β1→4)GlcNAc</td>
<td>Very large</td>
<td>Structural: in insects, spiders, crustaceans, gives rigidity and strength to exoskeletons</td>
</tr>
<tr>
<td>Peptidoglycan</td>
<td>Hetero-; peptides attached</td>
<td>4)Mur2Ac(β1→4)GlcNAc(β1)</td>
<td>Very large</td>
<td>Structural: in bacteria, gives rigidity and strength to cell envelope</td>
</tr>
<tr>
<td>Hyaluronate (a glycosaminoglycan)</td>
<td>Hetero-; acidic</td>
<td>4)GlcA(β1→3)GlcNAc(β1)</td>
<td>Up to 100,000</td>
<td>Structural: in vertebrates, extracellular matrix of skin and connective tissue; viscosity and lubrication in joints</td>
</tr>
</tbody>
</table>

* Each polymer is classified as a homopolysaccharide (homo-) or heteropolysaccharide (hetero-).
† The abbreviated names for the peptidoglycan and hyaluronate repeating units indicate that the polymer contains repeats of this disaccharide unit, with the GlcNAc of one disaccharide unit linked β(1→4) to the first residue of the next disaccharide unit.
Glycoconjugates: Proteoglycans, Glycoproteins, and Glycolipids

In addition to their important roles as stored fuels (starch, glycogen) and as structural materials (cellulose, chitin, peptidoglycan), polysaccharides (and oligosaccharides) are information carriers: they serve as destination labels for some proteins and as mediators of specific cell-cell interactions and interactions between cells and the extracellular matrix. Specific carbohydrate-containing molecules act in cell-cell recognition and adhesion, cell migration during development, blood clotting, the immune response, and wound healing, to name but a few of their many roles. In most of these cases, the informational carbohydrate is covalently joined to a protein or a lipid to form a glycconjugate, which is the biologically active molecule.

Proteoglycans are macromolecules of the cell surface or extracellular matrix in which one or more glycosaminoglycan chains are joined covalently to a membrane protein or a secreted protein. The glycosaminoglycan moiety commonly forms the greater fraction (by mass) of the proteoglycan molecule, dominates the structure, and is often the main site of biological activity. In many cases the biological activity is the provision of multiple binding sites, rich in opportunities for hydrogen bonding and electrostatic interactions with other proteins of the cell surface or the extracellular matrix. Proteoglycans are major components of connective tissue such as cartilage, in which their many noncovalent interactions with other proteoglycans, proteins, and glycosaminoglycans provide strength and resilience.

Glycoproteins have one or several oligosaccharides of varying complexity joined covalently to a protein. They are found on the outer face of the plasma membrane, in the extracellular matrix, and in the blood. Inside cells they are found in specific organelles such as Golgi complexes, secretory granules, and lysosomes. The oligosaccharide portions of glycoproteins are less monotonous than the glycosaminoglycan chains of proteoglycans; they are rich in information, forming highly specific sites for recognition and high-affinity binding by other proteins.

Glycolipids are membrane lipids in which the hydrophilic head groups are oligosaccharides, which, as in glycoproteins, act as specific sites for recognition by carbohydrate-binding proteins.

Glycobiology, the study of the structure and function of glycoconjugates, is one of the most active and exciting areas of biochemistry and cell biology. Our discussion uses just a few examples to illustrate the diversity of structure and the range of biological activity of the glycoconjugates. In Chapter 20 we discuss the biosynthesis of polysaccharides, including the glycosaminoglycans, and in Chapter 27, the assembly of oligosaccharide chains on glycoproteins.

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

The basic proteoglycan unit consists of a “core protein” with covalently attached glycosaminoglycan(s). For example, the sheetlike extracellular matrix (basal lamina) that separates organized groups of cells contains a family of core proteins (M, 20,000 to 40,000), each with several covalently attached heparan sulfate chains. (Heparan sulfate is structurally similar to heparin but has a lower density of sulfate esters.) The point of attachment is commonly a Ser residue, to which the glycosaminoglycan is joined through a trisaccharide bridge (Fig. 9–21). The Ser residue is generally in the sequence –Ser–Gly–X–Gly– (where X is any amino acid residue), although not every protein with this sequence has an attached glycosaminoglycan. Many proteoglycans are secreted into the extracellular matrix, but

\[
\begin{align*}
\text{GlcA} &\rightarrow \text{GalNAc} \\
\text{GlcA} &\rightarrow \text{Gal} \\
\text{Gal} &\rightarrow \text{Xyl} \\
\text{Xyl} &\rightarrow \text{Ser}
\end{align*}
\]

**Figure 9–21**

Proteoglycan structure, showing the trisaccharide bridge. A typical trisaccharide linker (blue) connects a glycosaminoglycan—in this case chondroitin sulfate (orange)—to a Ser residue (red) in the core protein. The xylose residue at the reducing end of the linker is joined by its anomic carbon to the hydroxyl of the Ser residue.
some are integral membrane proteins. For example, syndecan core protein (Mr 56,000) has a single transmembrane domain and an extracellular domain bearing three chains of heparan sulfate and two of chondroitin sulfate, each attached to a Ser residue (Fig. 9–22). The heparan sulfate moieties bind a variety of extracellular ligands and thereby modulate the ligands’ interaction with specific receptors of the cell surface.

Some proteoglycans can form proteoglycan aggregates, enormous supramolecular assemblies of many core proteins all bound to a single molecule of hyaluronate. Aggrecan core protein (Mr 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 Mr \(-2 \times 10^6\). When a hundred or more of these “decorated” core proteins bind a single, extended molecule of hyaluronate (Fig. 9–23), the resulting proteoglycan aggregate (Mr \(>2 \times 10^8\)) and its associated water of hydration occupy a volume about equal to that of a bacterial cell! Aggrecan interacts strongly with collagen in the extracellular matrix of cartilage, contributing to its development and tensile strength.

Intertwined with these enormous extracellular proteoglycans are fibrous matrix proteins such as collagen, elastin, and fibronectin, forming a cross-linked meshwork that gives the whole extracellular matrix 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, collagen, and a family of plasma membrane proteins called integrins that mediate signaling between the cell interior and the extracellular matrix (see Fig. 12–19). Integrins, in turn, have binding sites for a number of other extracellular macromolecules. The picture of cell-matrix interactions that emerges (Fig. 9–24) shows an array of interactions between cellular and extracellular molecules. These interactions serve not merely to anchor cells to the extracellular matrix but also to provide paths that direct the migration of cells

**figure 9–22**
Proteoglycan structure of an integral membrane protein. This schematic diagram shows syndecan, a core protein of the plasma membrane. The amino-terminal domain on the extracellular side of the membrane is covalently attached (by trisaccharide linkers such as those in Fig. 9–21) to three heparan sulfate chains and two chondroitin sulfate chains.

**figure 9–23**
A proteoglycan aggregate of the extracellular matrix. One very long molecule of hyaluronate 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 situated at the junction between each core protein and the hyaluronate backbone mediate the core protein–hyaluronate interaction.
in developing tissue and, through integrins, to convey information in both
directions across the plasma membrane.

Matrix proteoglycans are essential in the response of cells to certain ex-
tracellular growth factors. For example, fibroblast growth factor (FGF), an
extracellular protein signal that stimulates cell division, first binds to he-
paran sulfate moieties of syndecan molecules in the target cell’s plasma
membrane. Syndecan then “presents” FGF to the specific FGF plasma
membrane receptor, and only then can FGF interact productively with its
receptor to trigger cell division.

**Glycoproteins Are Information-Rich Conjugates Containing Oligosaccharides**

Glycoproteins are carbohydrate-protein conjugates in which the carbohy-
drate moieties are smaller and more structurally diverse than the gly-
cosaminoglycans of proteoglycans. The carbohydrate is attached at its
anomic 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. 9–25). Some glycoproteins have a single
oligosaccharide chain, but many have more than one; the carbohydrate may
constitute from 1% to 70% or more of the glycoprotein by mass. The struc-
tures of a large number of O- and N-linked oligosaccharides from a variety
of glycoproteins are known; Figure 9–25 shows a few typical examples.

As we will see in Chapter 12, the external surface of the plasma mem-
brane has many membrane glycoproteins with arrays of covalently attached
oligosaccharides of varying complexity. One of the best-characterized mem-
brane glycoproteins is glycophorin A of the erythrocyte membrane (see Fig.
12–10). It contains 60% carbohydrate by mass in the form of 16 oligosac-
charide chains (totalling 60 to 70 monosaccharide residues) covalently at-
tached to amino acid residues near the amino terminus of the polypeptide
chain. Fifteen of the oligosaccharide chains are O-linked to Ser or Thr
residues, and one is N-linked to an Asn residue.

**figure 9–24**
Interactions between cells and extracellular matrix. The
association between cells and the proteoglycan of extra-
cellular matrix is mediated by a membrane protein (inte-
grin) and by an extracellular protein (fibronectin in this
example) with binding sites for both integrin and the pro-
teoglycan. Note the close association of collagen fibers
with the fibronectin and proteoglycan.

**figure 9–25**
Oligosaccharide linkages in glycoproteins. (a) O-linked
oligosaccharides have a glycosidic bond to the hydroxyl
group of Ser or Thr residues (shaded pink), illustrated
here with GalNAc as the sugar at the reducing end of the
oligosaccharide. One simple chain and one complex chain
are shown. (Recall that Neu5Ac—N-acetylneuraminic
acid—is commonly called sialic acid.) (b) N-linked
oligosaccharides have an N-glycosyl bond to the amide
nitrogen of an Asn residue (shaded green), illustrated here
with GlcNAc as the terminal sugar. Three common types
of oligosaccharide chains that are N-linked in glycopro-
teins are shown. A complete description of oligosaccharide
structure requires specification of the position and stereo-
chemistry (α or β) of each glycosidic linkage.
Many of the proteins secreted by eukaryotic 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. Many milk proteins, including lactalbumin, and some of the proteins secreted by the pancreas (e.g., ribonuclease) are glycosylated, as are most of the proteins contained in lysosomes.

The biological advantages of adding oligosaccharides to proteins are not fully understood. The very hydrophilic clusters of carbohydrate alter the polarity and solubility of the proteins with which they are conjugated. Oligosaccharide chains attached to newly synthesized proteins in the Golgi complex may also influence the sequence of polypeptide-folding events that determine the tertiary structure of the protein (see Fig. 27–36); steric interactions between peptide and oligosaccharide may preclude one folding route and favor another. 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 more specific biological effects of oligosaccharide chains in glycoproteins. We noted earlier the difference between the information-rich linear sequences of nucleic acids and proteins and the monotonous regularity of homopolysaccharides such as cellulose (see Fig. 3–23). The oligosaccharides attached to glycoproteins are generally not monotonous but enormously rich in structural information. Consider the oligosaccharides in Figure 9–25, typical of those found in many glycoproteins. The most complex of those shown contains 14 monosaccharide residues of four different kinds, variously linked as (1→2), (1→3), (1→4), (1→6), (2→3), and (2→6), some with the α and some with the β configuration. The number of possible permutations and combinations of monosaccharide types and glycosidic linkages in an oligosaccharide of this size is astronomical. Each of the oligosaccharides in Figure 9–25 therefore presents a unique face, recognizable by the enzymes and receptors that interact with it. A number of cases are known in which the same protein produced in two tissues has different glycosylation patterns. The human protein interferon IFN-β1 produced in ovarian cells, for example, has oligosaccharide chains that differ from those of the same protein produced in breast epithelial cells. The biological significance of these tissue glycoforms is not understood, but in some way the oligosaccharide chains represent a tissue-specific mark.

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 sialic acid (Fig. 9–9) and other monosaccharide residues. Some of the oligosaccharide moieties of gangliosides, such as those that determine human blood groups (see Fig. 11–12), are identical with those found in certain glycoproteins, which therefore also contribute to blood group type determination. 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 E. coli and Salmonella typhimurium. They are prime targets of the antibodies produced by the ver-
tebrate 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. 9–26).

*E. coli* has similar but unique lipopolysaccharides. The lipopolysaccharides of some bacteria are toxic to humans and other animals; for example, they are responsible for the dangerously lowered blood pressure that occurs in toxic shock syndrome resulting from gram-negative bacterial infections.

**Oligosaccharide-Lectin Interactions Mediate Many Biological Processes**

Lectins, found in all organisms, are proteins that bind carbohydrates with high affinity and specificity. A number of critically situated hydrogen-bonding partners in the carbohydrate recognition domain of each type of protein bind to specific oligosaccharides, and thus lectins easily distinguish between closely similar sugars (Table 9–3). Lectins serve in a wide variety of cell-cell recognition and adhesion processes. Here we discuss just a few examples of the roles of lectins in nature. In the laboratory, purified lectins are useful reagents for detecting and separating glycoproteins with different oligosaccharide moieties.

**figure 9–26**

**Bacterial lipopolysaccharides.** (a) Schematic diagram of the lipopolysaccharide of the outer membrane of *Salmonella typhimurium*. Kdo is 3-deoxy-α-manno-octonic acid, also called ketodeoxyoctulosonic acid; Hep is D-glycero-α-mannoheptose; AbeOAc is abequose (a 3,6-dideoxyhexose) acetylated on one of its hydroxyls. There are six fatty acids in the lipid A portion of the molecule. Different bacterial species have subtly different lipopolysaccharide structures, but they have in common a lipid region (lipid A), a core oligosaccharide, and an “O-specific” chain, which is the principal determinant of the serotype (immunological reactivity) of the bacterium. The outer membrane of the gram-negative bacteria *S. typhimurium* and *E. coli* contains so many lipopolysaccharide molecules that the cell surface is virtually covered with O-specific chains. (b) Space-filling molecular model of the lipopolysaccharide from *E. coli*. 
The sialic acid (Neu5Ac) residues situated at the ends of the oligosaccharide chains of many plasma glycoproteins (Fig. 9–25) protect the proteins from uptake and degradation in the liver. For example, ceruloplasmin, a copper-transporting glycoprotein, has several oligosaccharide chains ending in sialic acid. Removal of sialic acid residues by the enzyme sialidase is one way in which the body marks “old” proteins for destruction and replacement. The plasma membrane of hepatocytes has lectin molecules (asialoglycoprotein receptors; “asialo-” indicating “without sialic acid”) that specifically bind oligosaccharide chains with galactose residues no longer “protected” by a terminal sialic acid residue. Receptor-ceruloplasmin interaction triggers endocytosis and destruction of the ceruloplasmin.

A similar mechanism is apparently responsible for removing old erythrocytes from the mammalian bloodstream. Newly synthesized erythrocytes have several membrane glycoproteins with oligosaccharide chains that end in sialic acid. When sialic acid residues are removed by withdrawing a sample of blood, treating it with sialidase in vitro, and reintroducing it into the circulation, the treated erythrocytes disappear from the bloodstream within a few hours, whereas those with intact oligosaccharides (erythrocytes withdrawn and reintroduced without sialidase treatment) continue to circulate for days.

Some peptide hormones that circulate in the blood have oligosaccharide moieties that strongly influence their circulatory lifetime. Luteinizing hormone and thyrotropin (polypeptide hormones produced in the adrenal cortex) have N-linked oligosaccharides that end with the disaccharide GalNAc4SO3(β1→4)GlcNAC, which is recognized by a lectin (receptor) of hepatocytes. (GalNAc4SO3 is N-acetylgalactosamine sulfated on the —OH group of C-4.) Receptor-hormone interaction mediates the uptake and destruction of luteinizing hormone and thyrotropin, reducing their concentra-

<table>
<thead>
<tr>
<th>Lectin family and lectin</th>
<th>Abbreviation</th>
<th>Ligand(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concanavalin A</td>
<td>ConA</td>
<td>Manα1-→OCH3</td>
</tr>
<tr>
<td>Griffonia simplicifolia</td>
<td>GS4</td>
<td>Lewis b (Le³) tetrasaccharide</td>
</tr>
<tr>
<td>Wheat germ agglutinin</td>
<td>WGA</td>
<td>Neu5Ac(α2→3)Gal(β1→4)Glc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GlcNAC(β1→4)GlcNAC</td>
</tr>
<tr>
<td>Ricin</td>
<td></td>
<td>Gal(β1→4)Glc</td>
</tr>
<tr>
<td>Animal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galectin-1</td>
<td></td>
<td>Gal(β1→4)Glc</td>
</tr>
<tr>
<td>Mannose-binding protein A</td>
<td>MBP-A</td>
<td>High-mannose octasaccharide</td>
</tr>
<tr>
<td>Viral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Influenza virus hemagglutinin</td>
<td>HA</td>
<td>Neu5Ac(α2→6)Gal(β1→4)Glc</td>
</tr>
<tr>
<td>Polyoma virus protein 1</td>
<td>VP1</td>
<td>Neu5Ac(α2→3)Gal(β1→4)Glc</td>
</tr>
<tr>
<td>Bacterial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterotoxin</td>
<td>LT</td>
<td>Gal</td>
</tr>
<tr>
<td>Cholera toxin</td>
<td>CT</td>
<td>GM1 pentasaccharide</td>
</tr>
</tbody>
</table>

tion in the blood. Thus the blood levels of these hormones undergo a periodic rise (due to secretion by the adrenal cortex) and fall (due to destruction by hepatocytes).

Selectins are a family of lectins, found in plasma membranes, that mediate cell-cell recognition and adhesion in a wide range of cellular processes. One such process is the movement of immune cells (T lymphocytes) through the capillary wall, from blood to tissues, at sites of infection or inflammation (Fig. 9–27). At an infection site, P-selectin on the surface of capillary endothelial cells interacts with a specific oligosaccharide of the glycoproteins of circulating T lymphocytes. This interaction slows the T cells as they adhere to and roll along the endothelial lining of the capillaries. A second interaction, between integrin molecules in the T-cell plasma membrane and an adhesion protein on the endothelial cell surface, now stops the T cell and allows it to move through the capillary wall into the infected tissues to initiate the immune attack. Two other selectins participate in this process: E-selectin on the endothelial cell and L-selectin on the T cell bind their cognate oligosaccharides on the other cell.

Some microbial pathogens have lectins that mediate bacterial adhesion to host cells or toxin entry into cells. The bacterium believed responsible for most gastric ulcers, *Helicobacter pylori*, adheres to the inner surface of the stomach by interactions between bacterial membrane lectins and specific oligosaccharides of membrane glycoproteins of the gastric epithelial cells (Fig. 9–28). Among the binding sites recognized by *H. pylori* is the oligosaccharide Le^b^ when it is a part of the type O blood group determinant. This observation helps to explain the severalfold greater incidence of gastric ulcers in people of blood type O than those of type A or B. Chemically synthesized analogs of the Le^b^ oligosaccharide may prove useful in treating this type of ulcer. Administered orally, they could prevent bacterial adhesion (and thus infection) by competing with the gastric glycoproteins for binding to the bacterial lectin.

The cholera toxin molecule (produced by *Vibrio cholerae*) triggers diarrhea after entering intestinal cells responsible for water absorption from the intestine. It attaches to its target cell through the oligosaccharide of ganglioside GM1, a membrane phospholipid (see Box 11–2, Fig. 1), in the surface of intestinal epithelial cells. Similarly, the pertussis toxin produced by *Bordetella pertussis*, the bacterium that causes whooping cough, enters target cells only after interacting with an oligosaccharide (or perhaps several oligosaccharides) with a terminal sialic acid residue. Understanding the details of the oligosaccharide-binding sites of these toxins (lectins) may allow the development of genetically engineered toxin analogs for use in vaccines. Toxin analogs engineered to lack the carbohydrate binding site would be harmless because they could not bind to and enter cells, but they might elicit an immune response that would protect the recipient if exposed to the natural toxin.

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 is the HA protein, which we will describe in Chapter 12. The HA protein is essential for viral entry and infection (see pp. 405–406 and Fig. 12–21a).

**figure 9–27**
Role of lectin-ligand interactions in lymphocyte movement to the site of an infection or injury. A T lymphocyte circulating through a capillary is slowed by transient interactions between P-selectin molecules in the plasma membrane of capillary endothelial cells and glycoprotein ligands for P-selectin on the T-cell surface. As it interacts with successive P-selectin molecules, the T cell rolls along the capillary surface. Near a site of inflammation, stronger interactions between integrin in the capillary surface and its ligand in the T-cell surface lead to tight adhesion. The T-cell stops rolling and, under the influence of signals sent out from the site of the inflammation, begins extravasation—escape through the capillary wall—as it moves toward the site of inflammation.

**figure 9–28**
*Helicobacter pylori* cells adhering to the gastric surface. This bacterium causes ulcers by interactions between a bacterial surface lectin and the Le^b^ oligosaccharide (a blood group antigen) of the gastric epithelium.
Figure 9–29 summarizes the types of biological interactions that are mediated by carbohydrate binding.

**Analysis of Carbohydrates**

The growing appreciation of the importance of oligosaccharide structure in biological recognition has been the driving force behind the development of methods for analyzing the structure and stereochemistry of complex oligosaccharides. 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. For such analysis, oligosaccharides are generally removed from their protein or lipid conjugates, then subjected to stepwise degradation with specific reagents that reveal bond position or stereochemistry. Mass spectrometry and NMR spectroscopy have also become invaluable in deciphering oligosaccharide structure.

The oligosaccharide moieties of glycoproteins or glycolipids can be released by purified enzymes—glycosidases that specifically cleave O- or N-linked oligosaccharides or lipases that remove lipid head groups. Mixtures of carbohydrates are resolved into their individual components (Fig. 9–30) by some of the same techniques useful in protein and amino acid separation: fractional precipitation by solvents, and ion-exchange and gel filtration (size-exclusion) chromatography (see Fig. 5–18). Highly purified lectins, attached covalently to an insoluble support, are commonly used in affinity chromatography of carbohydrates (see Fig. 5–18c). Hydrolysis of oligosaccharides and polysaccharides in strong acid yields a mixture of monosaccharides, which, after conversion to suitable volatile derivatives, may be separated, identified, and quantified by gas-liquid chromatography (see p. 385) to yield the overall composition of the polymer. For simple, linear polymers such as amylose, the positions of the glycosidic bonds are determined.
by 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 present in the monosaccharide derivatives produced are those that were involved in glycosidic bonds. To determine the sequence of monosaccharide residues, including branches if they are present, exoglycosidases of known specificity are used to remove residues one at a time from the nonreducing end(s). The specificity of these exoglycosidases often allows deduction of the position and stereochemistry of the linkages.

Oligosaccharide analysis relies increasingly on mass spectrometry and high-resolution NMR spectroscopy. NMR analysis alone, especially for oligosaccharides of moderate size, can yield much information about sequence, linkage position, and anomic carbon configuration. Polysaccharides and large oligosaccharides can be treated chemically or with endoglycosidases.
to split specific internal glycosidic bonds, producing several smaller, more easily analyzable oligo-saccharides. 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 problem than determining the linear sequences of proteins and nucleic acids, which have monomers joined by a single bond type.

Carbohydrates are predominantly cyclized poly-hydroxy aldehydes or ketones, which occur in nature as monosaccharides (aldoses or ketoses), oligosaccharides (several monosaccharide units), and polysaccharides (large linear or branched molecules containing many monosaccharide units). Monosaccharides have at least one asymmetric carbon atom and thus exist in stereoisomeric forms. Most common, naturally occurring sugars, such as ribose, glucose, fructose, and mannose, are of the D series. Simple sugars having four or more carbon atoms may exist in the form of closed-ring hemiacetals or hemiketals, either furanoses (five-membered ring) or pyranoses (six-membered ring). Furanoses and pyranoses occur in anomeric α and β forms, which are interconverted in the process of mutarotation. Sugars with free, oxidizable anomeric carbons are called reducing sugars. Many derivatives of the simple sugars are found in living cells, including amino sugars and their acetylated derivatives, aldonic acids, and uronic acids. The hydroxyl groups of monosaccharides can also form phosphate and sulfate esters. Disaccharides consist of two monosaccharides joined by a glycosidic bond.

Polysaccharides (glycans) contain many monosaccharide residues in glycosidic linkage. Some (starch and glycogen) function as storage forms of carbohydrate, high molecular weight, branched polymers of glucose having (α1→4) linkages in the main chains and (α1→6) linkages at the branch points. Other polysaccharides play a structural role in cell walls and exoskeletons. Cellulose (in plants) has β-glucose units in (β1→4) linkage; chitin (in insect exoskeletons) is a linear polymer of N-acetylglucosamine joined by (β1→4) linkages. Bacterial cell walls contain peptidoglycans, linear polysaccharides of alternating N-acetylmuramic acid and N-acetylglucosamine residues, cross-linked by short peptide chains. The extracellular matrix surrounding cells in animal tissues contains a variety of glycosaminoglycans, linear polymers of alternating amino sugar and uronic acid units. These include hyaluronate, a high molecular weight polymer of alternating d-glucuronic acid and N-acetyl-d-glucosamine residues, and a variety of shorter, sulfated, very acidic heteropolysaccharides such as heparan sulfate, chondroitin sulfate, keratan sulfate, and dermatan sulfate.

Many of the oligosaccharides in a cell are in glycoconjugates—hybrid molecules in which the carbohydrate moiety is covalently bound to a protein or lipid. Proteoglycans, proteins (such as syndecan and aggrecan core proteins) with one or more covalently attached glycosaminoglycan chains, are generally found on the outer surface of cells or in the extracellular matrix. In most tissues, the extracellular matrix contains a variety of proteoglycans and multiadhesive proteins such as fibronectin, as well as huge proteoglycan aggregates consisting of many proteoglycans non-covalently bound to a single large hyaluronate molecule.

Glycoproteins contain covalently linked oligosaccharides that are smaller but more structurally complex, and therefore more information-rich, than glycosaminoglycans. Many cell surface or extracellular proteins are glycoproteins, as are most secreted proteins. The carbohydrate moieties serve as biological labels, which are “read” by lectins, proteins that bind oligosaccharide chains with high affinity and selectivity. Interactions between lectins and specific oligosaccharides are central to cell-cell recognition and adhesion. In vertebrates, oligosaccharide tags govern the rate of degradation of certain peptide hormones, circulating proteins, and blood cells. Selectins are plasma membrane lectins that bind carbohydrate chains in the extracellular matrix or on the surfaces of other cells, thereby mediating the flow of information between cell and matrix or between cells. The adhesion of bacterial and viral pathogens (including Helicobacter, the cholera and pertussis toxins, and the influenza virus) to their animal
cell targets occurs through binding of lectins in the pathogens to oligosaccharides of the target cell surface.

The structure of oligosaccharides and polysaccharides is investigated by a combination of methods: specific enzymatic hydrolysis to determine stereochemistry and produce smaller fragments for further analysis; methylation analysis to locate the glycosidic bonds; mass spectrometry, stepwise degradation, and NMR spectroscopy to determine sequences; and high-resolution NMR spectroscopy to establish configurations at anomeric carbons.

further reading

General Background on Carbohydrate Chemistry


A comprehensive text at the graduate level.


Thorough, advanced treatment of the chemistry and biology of cell surface carbohydrates. Good chapters on lectins, carbohydrate recognition in cell-cell interactions, and chemical synthesis of oligosaccharides.


The fundamentals of carbohydrate chemistry and biology, presented at a level suitable for advanced undergraduates and graduate students.


Chapters 34 and 35 cover the structure, stereochemistry, nomenclature, and chemical reactions of carbohydrates.


Comprehensive treatise on carbohydrate chemistry.

Glycosaminoglycans and Proteoglycans


An advanced review.


A review focusing on the most recent genetic and molecular biological studies of the matrix proteoglycans. The structure-function relationships of some paradigmatic proteoglycans are discussed in depth and novel aspects of their biology are examined.


An advanced review of the chemistry and biology of glycosaminoglycans.

Glycoproteins


A review that considers the genesis of glycoforms, functional roles for glycosylation, and structure-function relationships for several glycoproteins.


Lectins, Recognition, and Adhesion


A look at the role of the oligosaccharides that determine blood type in the adhesion of this microorganism to the stomach lining, producing ulcers.

Analogs of recognition oligosaccharides are used to block adhesion of a pathogen to its host cell target.


Evidence for roles of sulfated oligosaccharides in peptide hormone half-life, symbiont interactions in nitrogen-fixing legumes, and lymphocyte homing.


Article on the role of integrins in cell-cell adhesion, and possible roles in arthritis, heart disease, stroke, osteoporosis, and the spread of cancer.


This review briefly describes several recently determined structures of cell adhesion molecules, summarizes some of the main findings about each structure, and highlights common features of different cell adhesion systems.


A short review that focuses on the interaction of selectins with their carbohydrate ligands.


Chemical basis and biological roles of carbohydrate recognition.


Good treatment of the chemical basis of carbohydrate-protein interactions.

### Analysis of Carbohydrates


Very useful manual for analysis of all types of sugar-containing molecules—monosaccharides, polysaccharides and glycosaminoglycans, glycoproteins, proteoglycans, and glycolipids.


Excellent survey of the uses of NMR, mass spectrometry, and enzymatic reagents to determine oligosaccharide structure.


A how-to manual for the isolation and characterization of the oligosaccharide moieties of glycoproteins, using the whole range of modern techniques.


Thorough description of methylation analysis of carbohydrates.


Practical guide to working with oligosaccharides.


On the use of purified enzymes in analysis of structure and stereochemistry.


### problems

1. **Determination of an Empirical Formula** An unknown substance containing only C, H, and O was isolated from goose liver. A 0.423 g sample produced 0.620 g of CO₂ and 0.254 g of H₂O after complete combustion in excess oxygen. Is the empirical formula of this substance consistent with its being a carbohydrate? Explain.

2. **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. However, this sugar alcohol is no longer designated D or L. Why?

3. **Melting Points of Monosaccharide Osazone Derivatives** Many carbohydrates react with phenylhydrazine (C₆H₅NHNH₂) to form bright yellow crystalline derivatives known as osazones:
The melting temperatures of these derivatives are easily determined and are characteristic for each osazone. This information was used to help identify monosaccharides before the development of HPLC or gas-liquid chromatography. Listed below are the melting points (MPs) of some aldose-osazone derivatives:

<table>
<thead>
<tr>
<th>Monosaccharide</th>
<th>MP of anhydrous monosaccharide (°C)</th>
<th>MP of osazone derivative (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>146</td>
<td>205</td>
</tr>
<tr>
<td>Mannose</td>
<td>132</td>
<td>205</td>
</tr>
<tr>
<td>Galactose</td>
<td>165–168</td>
<td>201</td>
</tr>
<tr>
<td>Talose</td>
<td>128–130</td>
<td>201</td>
</tr>
</tbody>
</table>

As the table shows, certain pairs of derivatives have the same melting points, although the underivatized monosaccharides do not. Why do glucose and mannose, and galactose and talose, form osazone derivatives with the same melting points?

4. Interconversion of d-Galactose Forms A solution of one stereoisomer of a given monosaccharide rotates plane-polarized light to the left (counterclockwise) and is called the levorotatory isomer, designated (−); the other stereoisomer rotates plane-polarized light to the same extent but to the right (clockwise) and is called the dextrorotatory isomer, designated (+). An equimolar mixture of the (+) and (−) forms does not rotate plane-polarized light.

The optical activity of a stereoisomer is expressed quantitatively by its optical rotation, the number of degrees by which plane-polarized light is rotated on passage through a given path length of a solution of the compound at a given concentration. The specific rotation \([\alpha]_D\) of an optically active compound is defined thus:

\[
[\alpha]_D^{\text{obs}} = \frac{\text{observed optical rotation (°)}}{\text{optical path length (dm)} \times \text{concentration (g/mL)}}
\]

The temperature and the wavelength of the light employed (usually the D line of sodium, 589 nm) must be specified in the definition.

A freshly prepared solution of \(\alpha\)-d-galactose (1 g/mL in a 10 cm cell) shows an optical rotation of +150.7°. Over time, the observed rotation of the solution gradually decreases and reaches an equilibrium value of +80.2°. In contrast, a freshly prepared solution (1 g/mL) of \(\beta\)-d-galactose shows an optical rotation of only +52.8°. Moreover, the rotation increases over time to an equilibrium value of +80.2°, identical to that reached by \(\alpha\)-d-galactose.

(a) Draw the Haworth perspective formulas of the \(\alpha\) and \(\beta\) forms of d-galactose. What feature distinguishes the two forms?

(b) Why does the optical rotation of a freshly prepared solution of the \(\alpha\) form gradually decrease with time? Why do solutions of the \(\alpha\) and \(\beta\) forms (at equal concentrations) reach the same optical rotation at equilibrium?

(c) Calculate the percentage composition of the two forms of d-galactose at equilibrium.

5. A Taste of Honey The fructose in honey is mainly in the \(\beta\)-d-pyranose form. This is one of the sweetest substances known, about twice as sweet as glucose. The \(\beta\)-d-furanose form of fructose is much less sweet. 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 but not hot drinks. What chemical property of fructose could account for both of these observations?

6. Glucose Oxidase in Determination of Blood Glucose The enzyme glucose oxidase isolated from the mold \(\text{Penicillium notatum}\) catalyzes the oxidation of \(\beta\)-d-glucose to \(\beta\)-glucono-\(\delta\)-lactone. This enzyme is highly specific for the \(\beta\) anomer of glucose and does not affect the \(\alpha\) anomer. In spite of this specificity, the reaction catalyzed by glucose oxidase is commonly used in a clinical assay for total blood glucose—that is, for solutions consisting of a mixture of \(\beta\) and \(\alpha\)-d-glucose. How is this possible? Aside from allowing the detection of smaller quantities of glucose, what advantage does glucose oxidase offer over Fehling’s reagent for the determination of blood glucose?

7. Invertase “Inverts” Sucrose The hydrolysis of sucrose (specific rotation +66.5°) yields an equimolar mixture of d-glucose (specific rotation +52.5°) and d-fructose (specific rotation −92°).

(a) Suggest a convenient way to determine the rate of hydrolysis of sucrose by an enzyme preparation extracted from the lining of the small intestine.

(b) Explain why an equimolar mixture of d-glucose and d-fructose formed by hydrolysis of sucrose is called invert sugar in the food industry.

(c) The enzyme invertase (its preferred name is now sucrase) is allowed to act on a solution of sucrose until the optical rotation of the solution becomes zero. What fraction of the sucrose has been hydrolyzed?

8. 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.)

9. Anomers of Sucrose? Although lactose exists in two anomeric forms, no anomeric forms of sucrose have been reported. Why?

10. 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. Although they have markedly different physical properties, both substances are composed of (1→4)-linked D-glucose polymers of comparable molecular weight. What structural features of these two polysaccharides underlie their different physical properties? Explain the biological advantages of their respective properties.

11. 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 in the cellulose molecule is about 0.45 nm long.

12. Glycogen as Energy Storage: How Long Can a Game Bird Fly? Since ancient times it has been observed that certain game birds, such as grouse, quail, and pheasants, are easily fatigued. The Greek historian Xenophon wrote, “The bustards...can be caught if one is quick in starting them up, for they will fly only a short distance, like partridges, and soon tire; and their flesh is delicious.” The flight muscles of game birds rely almost entirely on the use of glucose 1-phosphate for energy, in the form of ATP (Chapter 15). In game birds, glucose 1-phosphate is formed by the breakdown of stored muscle glycogen, catalyzed by the enzyme glycogen phosphorylase. The rate of ATP production is limited by the rate at which glycogen can be broken down. During a “panic flight,” the game bird’s rate of glycogen breakdown is quite high, approximately 120 μmol/min of glucose 1-phosphate produced per gram of fresh tissue. Given that the flight muscles usually contain about 0.35% glycogen by weight, calculate how long a game bird can fly.

13. 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 appears 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 occupied by the molecule so much larger in solution?

14. 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 appears 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?

15. Information Content of Oligosaccharides The carbohydrate portion of some glycoproteins may serve as a cellular recognition site. In order to perform this function, the oligosaccharide moiety of glycoproteins 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.

16. Determination of the Extent of Branching in Amylopectin The extent of branching (number of (α1→6) glycosidic bonds) in amyllopectin can be determined by the following procedure. A sample of amyllopectin is exhaustively treated with a methylating agent (methyl iodide) that replaces all the hydrogens of the sugar hydroxyls with methyl groups, converting −OH to −OCH3. All the glycosidic bonds in the treated sample are then hydrolyzed in aqueous acid. The amount of 2,3-di-O-methylglucose in the hydrolyzed sample is determined.

(a) Explain the basis of this procedure for determining the number of (α1→6) branch points in amyllopectin. What happens to the unbranched glucose residues in amyllopectin during the methylation and hydrolysis procedure?

(b) A 258 mg sample of amyllopectin treated as described above yielded 12.4 mg of 2,3-di-O-methylglucose. Determine what percentage of the glucose residues in amyllopectin contain an (α1→6) branch.

17. Structural Analysis of a Polysaccharide A polysaccharide of unknown structure was isolated, subjected to exhaustive methylation, and hydrolyzed. Analysis of the products revealed three methylated sugars in the ratio 20:1:1. The sugars were 2,3,4-tri-O-methyl-D-glucose; 2,4-di-O-methyl-D-glucose; and 2,3,4,6-tetra-O-methyl-D-glucose. What is the structure of the polysaccharide?