The Institute of Biology's
Studies in Biology no. 125

Lipids and Polysaccharides in Biology

Anna J. Furth

The Institute of Biology's
Studies in Biology no. 125

Lipids and Polysaccharides in Biology

Anna J. Furth
Lecturer in Biology,
The Open University

Edward Arnold

First published 1980
by Edward Arnold (Publishers) Limited
41 Bedford Square, London WC1 3DQ

British Library Cataloguing in Publication Data

Furth, Anna J.
Lipids and polysaccharides in biology. – (Institute of Biology. Studies in biology; no. 125
ISSN 0537-9024).
1. Lipids 2. Polysaccharides
I. Title II. Series
574.1'9247 QP751

ISBN 0-7131-2805-4

Printed and bound in Great Britain at
The Camelot Press Ltd, Southampton

General Preface to Series

Because it is no longer possible for one textbook to cover the whole field of biology while remaining sufficiently up to date the Institute of Biology has sponsored this series so that teachers and students can learn about significant developments. The enthusiastic acceptance of 'Studies in Biology' shows that the books are providing authoritative views of biological topics.

The features of the series include the attention given to methods, the selected list of books for further reading and, wherever possible, suggestions for practical work.

Readers' comments will be welcomed by the Education Officer of the Institute.

1980

Institute of Biology
41 Queen's Gate
London SW7 5HU

Preface

Until recently lipids and polysaccharides were regarded simply as the food storage and structural support molecules of biology. However, they have as much potential for subtle variation in structure as the 'traditional' informational molecules – nucleic acids and proteins – and recent research has revealed their highly specific interactions in membranes and at cell surfaces. While describing the familiar food storage and support roles, this book aims also to provide the background knowledge essential for following new developments.

The relationship between chemical structure and biological role is constantly emphasized, as the informational role of cell surface polysaccharides, the water-trapping role of gums and mucins and the distributive role of serum lipoproteins and bile salts.

I am greatly indebted to Dr David Ellar, Miss Susan Foote and Professor Charles Phelps for their invaluable comments and suggestions, and to my long-suffering family for their support. This book is dedicated to A.H.F.

Milton Keynes, 1980

A. J. F.

Contents

1 Lipids, Polysaccharides and Macromolecules

1.1 What are lipids and polysaccharides?

The bulk of the solid material in living organisms is made up of protein, lipid and polysaccharide. A fourth class, present in relatively small amounts, contains the nucleic acids. Until recently lipids and polysaccharides were thought of as the least interesting of these biological molecules, but here I hope to convince you that this is a mistaken and outmoded view. Lipids and polysaccharides are as carefully adapted for their biological role as the nucleic acids and proteins and like them, have plenty of potential for carrying information.

Lipids can be broadly divided into three groups – *neutral fats* which are exemplified by butter and by the fat deposits under the skin, *phospholipids* which are all-important constituents of membranes, and *steroids* which include cholesterol, fat-soluble vitamins and sex hormones. The lipids shown in Fig. 1–1 may appear as a rather motley collection of molecules but they share one feature – a large proportion of CH_2 groups.

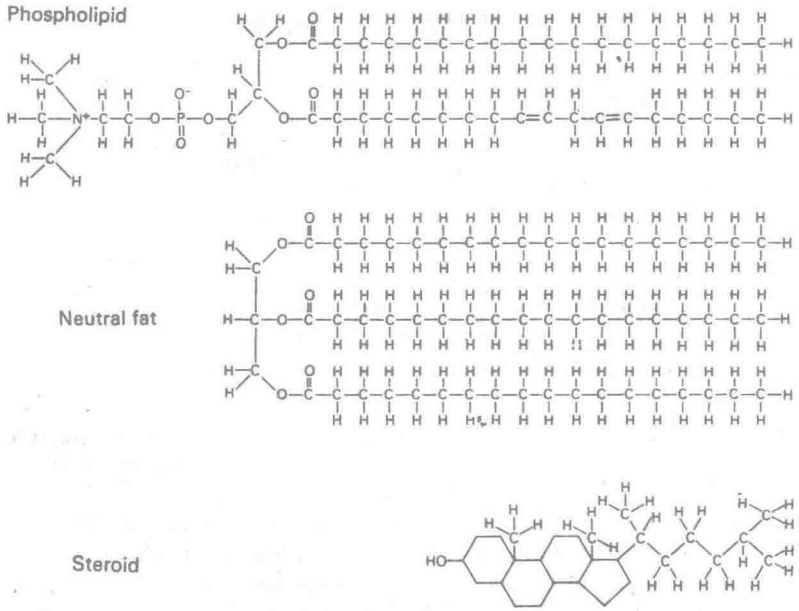


Fig. 1–1 Some lipids.

This provides their characteristic property, a high solubility in nonpolar solvents (such as chloroform or carbon tetrachloride) and a low solubility in water.

This contrasts with glucose and sucrose (see Fig. 1–2) which are water-soluble and readily available for metabolic reactions in the cell (e.g. glycolysis). These sugars are examples of mono- or oligosaccharides, whereas in this book we are chiefly concerned with the water-insoluble polysaccharides produced by monosaccharide polymerization. (*Oligo* means few; *poly* means many; *saccharide* means sugar.) These include water-trapping polysaccharides like the plant gums, as well as more familiar molecules like starch and cellulose. The short polysaccharide chains that form the *glyco* part of glycoproteins are also described here.

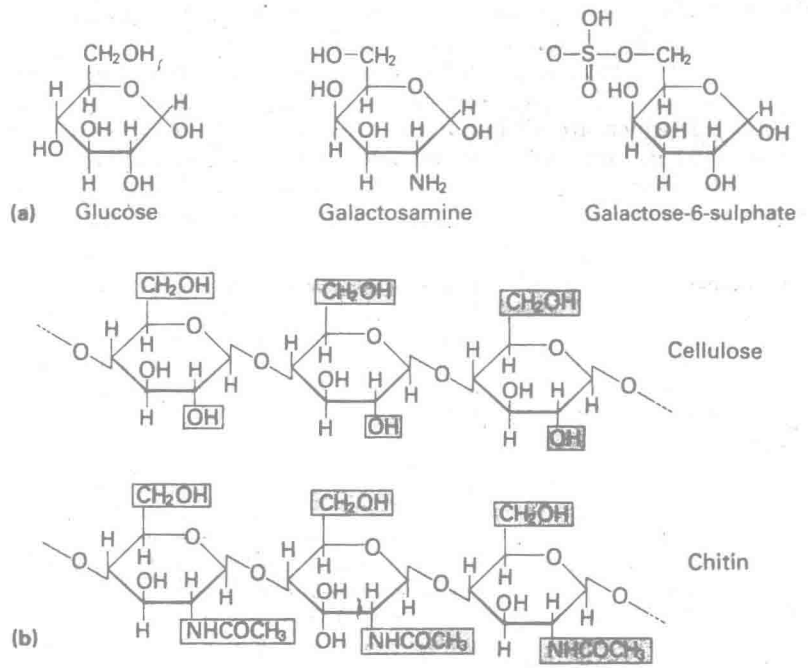


Fig. 1–2 Some mono- and polysaccharides. At each junction in the ring is a carbon atom. (a) Monosaccharides and (b) polysaccharides.

Traditionally the monosaccharides from which polysaccharides are built up, were defined in terms of a basic formula, $C_nH_{2n}O_n$. However, this definition excludes the many nitrogen and sulphur containing sugars, and also fails to emphasize one of the most important characteristics of a sugar – that it normally exists in the form of a ring.

1.2 Biological role

Both glycogen (a polysaccharide) and triglyceride (a lipid) are *food storage molecules*, and represent different ways in which energy-rich molecules can be accumulated in the cell. Both satisfy the basic requirements of a food storage molecule, that is, they can be stored in a concentrated form that does not upset the osmotic balance of the cell, yet each can be readily mobilized.

Apart from bone, some of the most rigid *structural support* molecules in biology are composed of polysaccharide. An extreme example is chitin, the major component of insect exoskeleton. Another well known but less brittle support polysaccharide is cellulose. These are both *fibrous polysaccharides*, but recently there has been considerable interest in a rather different class of structural molecule, the *gelling polysaccharides*. These operate in quite a different way, by enmeshing enormous quantities of water in a three-dimensional network of polysaccharide fibres. I shall be describing the molecular requirements for both these approaches to structural support.

The structural support role of lipids is seen in membranes, where the key requirement is for flexibility or *fluidity*, a concept that has radically altered our thinking on membrane biochemistry. Membranes are no longer seen as static structures. Considerable movement may take place, particularly in a direction parallel to the cell surface, and, as we shall see in Chapter 3, it is the peculiar molecular properties of the lipid components that make this movement possible.

An *informational role* for polysaccharides and lipids is a relatively new idea. Until recently it was thought to be confined to the nucleic acids and the globular proteins, both of which can be enormously selective in the molecules with which they interact. Enzyme–substrate interaction (see WYNN, 1979) and antibody–antigen interaction are examples of this molecular specificity. It is now realized that sugars, long known for their variety of structure, may also contribute to the information content of a macromolecule. This has important consequences in the regulation of tissue growth (malignant and otherwise), in graft rejection and in many intercellular control mechanisms. Even if polysaccharide-mediated messages have lower information content and are less detailed than those relayed by proteins or nucleic acids, this is still an exciting area where polysaccharide chemistry is expanding fast.

Lipids too have great potential for variation in structure. It is now clear that membranes vary in their lipid make-up, depending on location within the cell (e.g. mitochondrial or lysosomal), on growth conditions and other factors. However, examples of highly specific lipid interactions are so far rather rare, and this again is a new concept.

The biological role of any molecule is very much linked to its structure. This is particularly true of the macromolecules, which include both

polysaccharides and – as I shall argue here – the lipids. A firm understanding of macromolecular structure is therefore essential for what follows.

1.3 Weak bonding in macromolecules

The structure of macromolecules is heavily dependent on weak bonds, and this is what lies behind many of their unique properties. Unlike the vast majority of covalent bonds, these can be readily broken and remade without the need for enzymic catalysis. As you can see from Table 1, weak bond energy is a whole order of magnitude lower than that of covalent bonds, as found C—H or C—C. In biology the majority of weak bonds are hydrogen bonds, hydrophobic bonds or – less frequently – ionic bonds.

Table 1 Bond energies in macromolecules.

Bond type	*Bond energy* kJ* mol^{-1}
Covalent	200–400
Hydrogen	up to 20
Van der Waals	up to 4
Hydrophobic	4–8
Ionic	up to 4

* 1 kcal = 4.18 kJ

Hydrogen bonds form when H atoms attached to electronegative atoms such as oxygen or nitrogen lie close enough to interact with other electronegative atoms – usually O or N again (see Fig. 1–3). An essential feature of a hydrogen bond is its directionality, and strong hydrogen

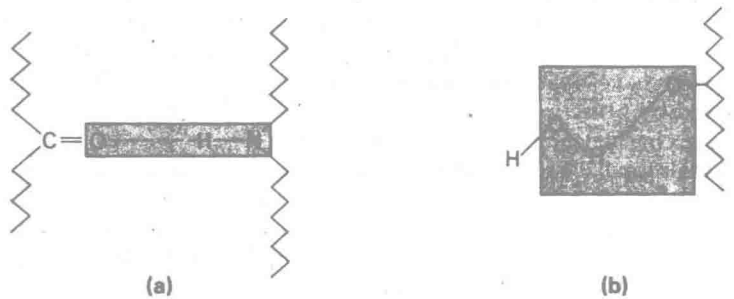


Fig. 1–3 Hydrogen bonds (a) strong and (b) weak. Dashed lines represent hydrogen bonds. Zig-zag lines represent parts of a macromolecular chain.

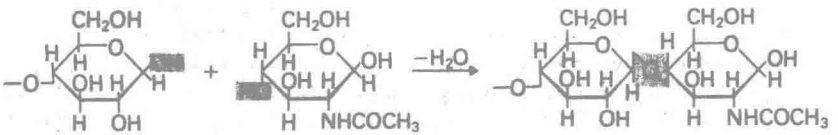
bonds form only if the three combining atoms are colinear (compare Fig. 1–3 a and b). Hence there is a large pay-off in terms of energy if the macromolecule chain can be folded to bring potential H-bonding atoms directly in line with one another. In sugars, the commonest H-bonding residues are the ubiquitous OH groups, the NH_2 groups of amino sugars and the ring oxygens (see Fig. 1–2a). In lipids the H-bonding residues are mostly O atoms, and tend to be clustered at one end of the molecule (see Fig. 1–1).

The concept of *hydrophobic bonding* is tied up with polarity or rather nonpolarity. Where a covalent bond joins two atoms of equal electronegativity (equal electron attracting power), the shared electrons will be evenly distributed between the combining nuclei. Hence a nonpolar molecule like CH_4 is electrically neutral, in contrast to a highly polarized molecule like $\overrightarrow{H—Cl}$ (the arrow points in the direction of electron displacement). Nonpolar residues (also called hydrophobic or 'water hating') include the CH_2 and CH_3 groups which, as you can see from Fig. 1–1, are particularly abundant in lipids. Nonpolar molecules, or parts of molecules, tend to be drawn together by a force known as van der Waals bonding. This produces a nonpolar region – like the interior of an oil drop – from which polar residues are expelled. These polar residues then tend to form hydrogen bonds with surrounding H_2O molecules. It is this combination of binding forces – van der Waals and hydrogen – that goes to make up hydrophobic bonding.

Finally, we come to *ionic bonding*, which is the electrostatic attraction between oppositely charged ions such as SO_3^- (from ionization of SO_3H) and NH_3^+ (from ionization of NH_2). Ionic groups occur chiefly in polysaccharides; phosphate residues provide the ionic groups in lipids.

1.4 Structure of macromolecules

The structural hierarchy on which a macromolecule is built up was originally elucidated from work on proteins. It is now clear that the same format can be applied to nucleic acids, polysaccharides and even to lipids. Figure 1–4 outlines the main features of this hierarchy, with globular protein as the example. *Primary structure* describes the order in which monomer building blocks (e.g. amino acid or monosaccharide) are assembled. Each is linked to its neighbour in a reaction which formally removes the elements of water. The equation below shows such a reaction between two monosaccharide residues.



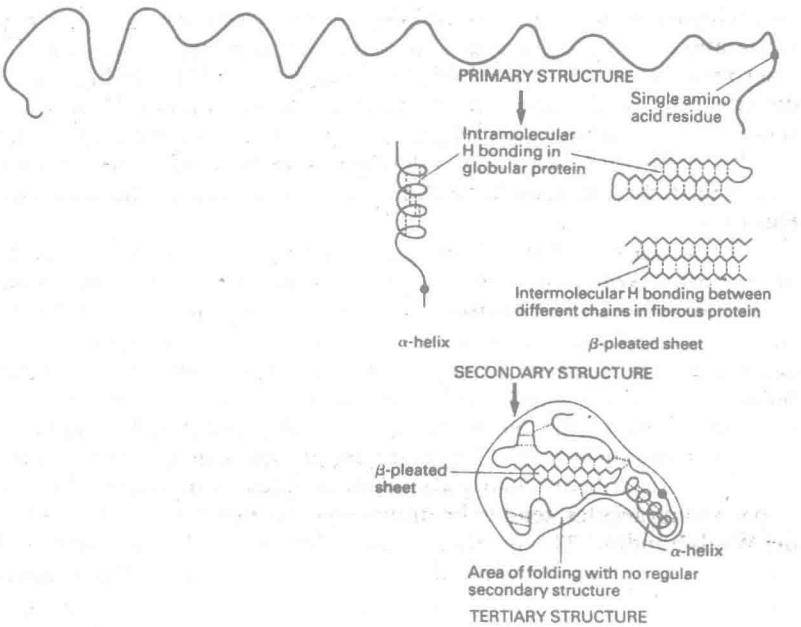


Fig. 1–4 Higher order structure in a globular protein. Dashed lines represent stabilizing H bonds. (From The Open University (1977). S322 *Biochemistry and Molecular Biology*, Unit 1, The Open University Press.)

In a pure protein, the combining units of the polymer are all amino acids; in a pure polysaccharide they are all sugar. But many of the more interesting macromolecules are hybrids, like the glycoproteins to be described in Chapter 2. These contain anything from 10% to 90% carbohydrate, often in the form of short chains projecting from the polypeptide backbone.

The molecular weight of the chain-like macromolecule may run to several millions, so clearly it cannot be accommodated in the cell without careful folding. Its folding pattern is known as *higher order structure* and can be described in terms of secondary, tertiary and quaternary structure.

Pure *secondary structure* is found in fibrous proteins like silk, and in support polysaccharides like cellulose. In the cell, these molecules are either twisted into helices or pleated into ribbons, and therefore considerably shorter than you would expect from their primary structures. Both twisting and pleating are highly regular, and stabilized by hydrogen bonds that run either along the length of the chain (*intra*molecularly) or at right angles to it (*inter*molecularly). The protein α-helix (Fig. 1–4) and the polysaccharides cellulose and chitin (Fig. 2–15) are all stabilized by intramolecular hydrogen bonding. Intermolecular H

bonding is seen where similar molecules pack together to form fibres.

All fibrous molecules are fairly extended, and their folding patterns may be fully described in terms of secondary structure. But in more compact, approximately spherical molecules like the globular proteins the folding pattern is more complex. It includes both the regular H-bonded elements that make up secondary structure, and the irregularly folded (but precisely defined) stretches that lie in between. Together these folding patterns make up the overall shape or *conformation* of the globular macromolecule, and this is what is known as *tertiary structure*. The surface of the folded 'globule' is by no means smooth, but covered with carefully defined bumps and dents – rather like the surface of a walnut. This surface geometry is extremely important for the functioning of macromolecules in the cell.

Like secondary structure, tertiary structure is stabilized solely by weak bonding. Figure 1–4 shows some of these bonds, running between adjacent lengths of folded chain. All types of weak bonding may contribute to higher order structure but, apart from disulphide bonds in proteins, the only covalently-linked part of a macromolecule is the string of component building blocks that go to make up its primary structure. This preponderance of weak bonding in the structure of macromolecules has far-reaching consequences.

I have now described how macromolecules are built up, and indicated how the structural hierarchy can apply to polysaccharides as well as proteins. However, you may remark that we have said very little about lipids. Some people are reluctant to consider these as macromolecules at all and there is some justification for this view if you consider the sizes of the lipid molecules in Fig. 1–1. One such as tristearin, a neutral fat with sixteen CH_2 groups in each of its three sidechains, has a molecular weight of 891. This falls far short of the molecular weight of a medium sized protein such as ribonuclease (molecular weight 14 000). In biology, however, we are interested in molecules as they occur in the cell, and here lipids are invariably found not in isolation but clumped together into specific aggregates. A membrane (see Fig. 3–11) is one such aggregate. There is, however, one major difference between lipid aggregates and the collection of amino acids or sugars that go to make up conventional macromolecules. The lipid aggregate is held together by non-covalent bonding and in contrast to the covalently linked macromolecules, it is a very fluid structure. Fluidity is one of the unique characteristics of lipids, to which we shall return in Chapter 3.

You might expect there to be an almost infinite number of ways in which a macromolecular chain could fold up in the cell. But the astonishing fact is that every molecule of a given primary structure invariably folds to the same shape. The reason for this reproducibility is that in solution, a *molecule spontaneously folds to the conformation of minimum free energy* – or to a conformation very closely related to it. Under a given

set of conditions such as temperature and pH, and for a given primary structure, there is only one such minimum-free-energy conformation. Thus the higher order structure of a macromolecule is determined by its primary structure. This extremely important principle was originally demonstrated with proteins, but it governs the folding and aggregation patterns of all macromolecules in the cell.

1.5 Characteristic properties of macromolecules – specificity and flexibility

The structural hierarchy of a macromolecule has two important consequences – it permits enormous specificity in the interaction of a macromolecule with other molecules, and it produces a flexibility that would be quite impossible on a rigid covalently-bonded structure.

The enzyme active site is one example of a *specific interaction site* built into the surface of a folded macromolecule. Other examples are shown in Fig. 1–5. The specificity with which interaction takes place at these sites depends on the closeness of fit between the surfaces of the two interacting molecules. This in turn depends on the nature of the monomer building blocks lining the interaction site. Where the macromolecule is a protein, there are twenty-two different amino acids to choose from. If it is a glycoprotein this number is augmented by an almost infinite pool of sugar residues. If it is a polysaccharide or a protein with covalently-bound lipid, the number grows even more. You should now begin to see how the informational role of a macromolecule depends on its structure. Without the structural hierarchy there could be none of the precise folding patterns needed to produce this wide range of specific interaction sites.

Specificity originates in the gene, since this controls primary and hence higher order structure in all macromolecules. Small changes in primary

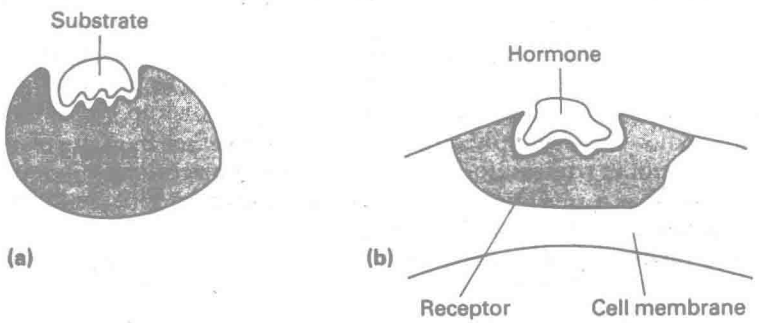


Fig. 1–5 Specific interaction sites in the higher order structure of a macromolecule. (a) Substrate binding site (active site) in an enzyme. (b) Hormone binding site in a membrane-bound receptor molecule.

structure have to be accommodated in a different higher order structure. We shall see examples of this in Chapter 2, where new sugars appearing in informational polysaccharides may radically alter the geography of specific interaction sites, and hence the biological function of the macromolecule. Protein and nucleic acid primary structures are directly dependent on the gene, but polysaccharide and lipid primary structures are one step removed. They depend on gene products, the enzymes responsible for their synthesis.

Our second characteristic is *flexibility*, known in lipids as fluidity. The weak bonds that stabilize higher order structure (see Table 1) can be readily broken and remade, permitting subtle changes of shape known as *conformational changes*. It is these that, at the molecular level, lie behind many of the control mechanisms of biology and they occur not only in proteins (where they were first demonstrated) but in polysaccharides and lipids as well.

Now that we have described the characteristics of macromolecules, we can see how far they fit the behaviour of polysaccharides and lipids in the cell.

2 Polysaccharides

2.1 Sugar building blocks

2.1.1 *Structural isomerism*

The basic formula of a monosaccharide is $C_nH_{2n}O_n$. Most of the sugars described in this book have n values of 6, or less frequently, 5. How does this simple molecular formula produce so many variants? In sugar chemistry, most of the variety comes from isomerism – subtle distinctions between compounds which have the same *molecular* formula (e.g. $C_nH_{2n}O_n$) but different arrangements of their component atoms. Molecular formulae describe simply the relative proportions of constituent atoms; they give no indication of how these are linked together into recognizable groups such as hydroxyl (OH) or carbonyl (C=O) residues. This information comes from *structural* formulae like that of glucose shown in Fig. 2–1a. By convention, carbon 1 of the sugar chain

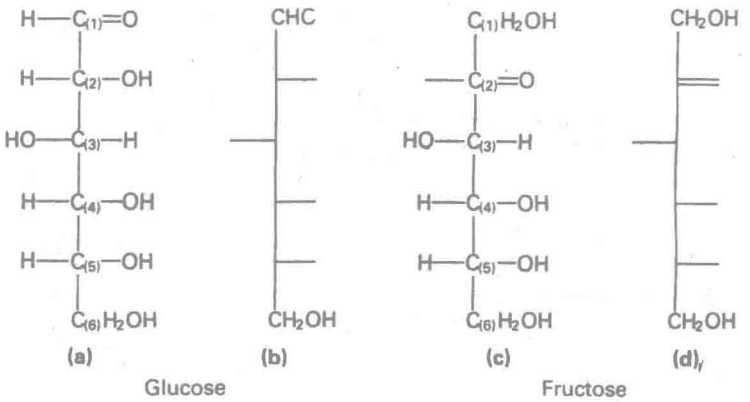


Fig. 2–1 Full and simplified formulae of (a) and (b) glucose and (c) and (d) fructose in the open chain configurations.

is either that carrying the carbonyl group (as in glucose) or else the carbon at the end of the chain nearest to this group. It is usually abbreviated to C-1 or $C_{(1)}$. Figure 2–1b shows a simplified version of the glucose formula, omitting both single H atoms and C symbols for carbons 2–5. The OH groups on all except C-6 are shown simply as straight lines.

Sugars with the carbonyl group on C-1 belong to the *aldose* family shown in full in Fig. 2–2. We may imagine this being built up from glyceraldehyde (shown at the top) by stepwise addition of CHOH groups.

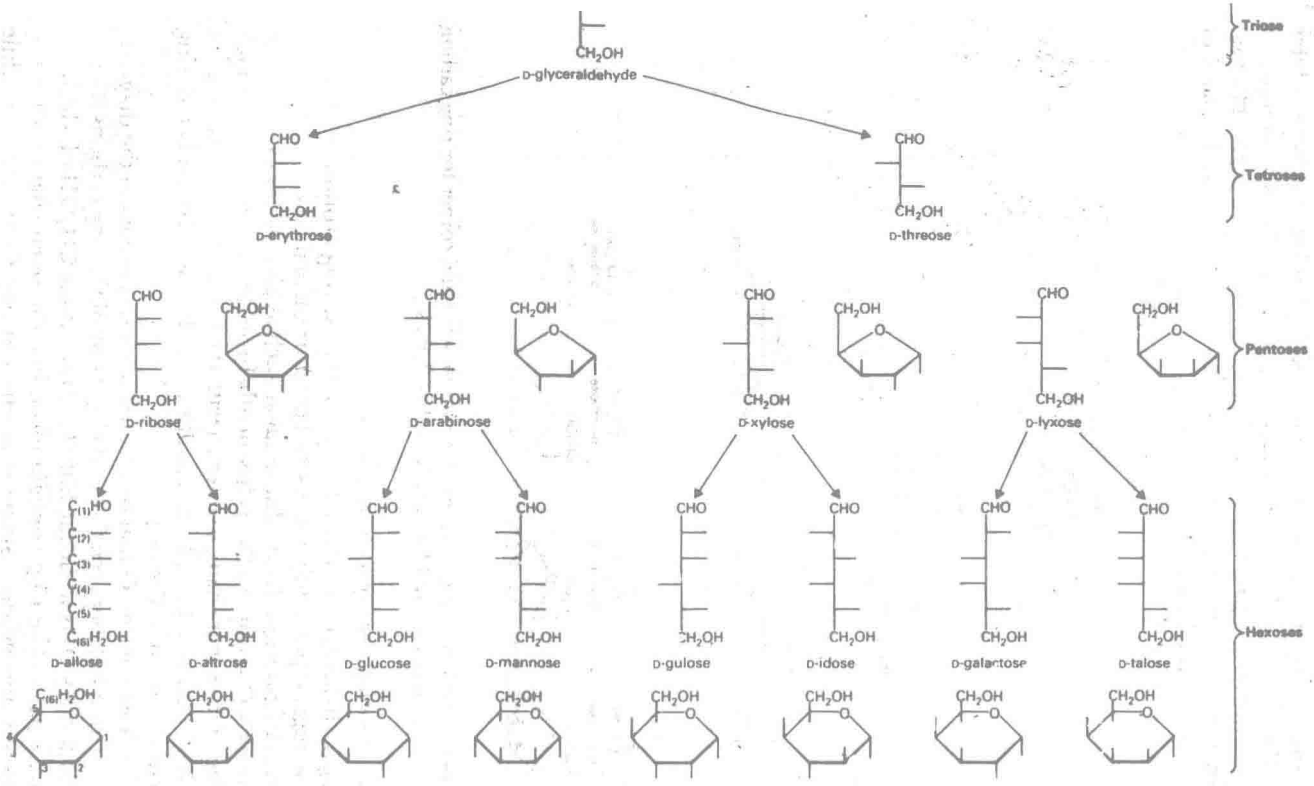


Fig. 2–2 The aldose sugar family. Open chain and ring configurations are shown. Carbon numbering is shown in the bottom left-hand corner of the figure, where $C_{(1)}$ or just the number 1 represents the first carbon of the chain. In the ring configuration, each junction represents a carbon atom (see Fig. 2–4). In the open chain configuration, horizontal bars represent OH groups (see Fig. 2–1).

A second family of sugars are the *ketoses* shown in Fig. 2–3, which differ from the aldoses in having the carbonyl group on C-2. Now compare a ketose 6-carbon sugar, fructose with a 6-carbon aldose sugar, glucose. These two are *structural isomers*. They have the same molecular but different structural formulae, and are clearly different molecules.

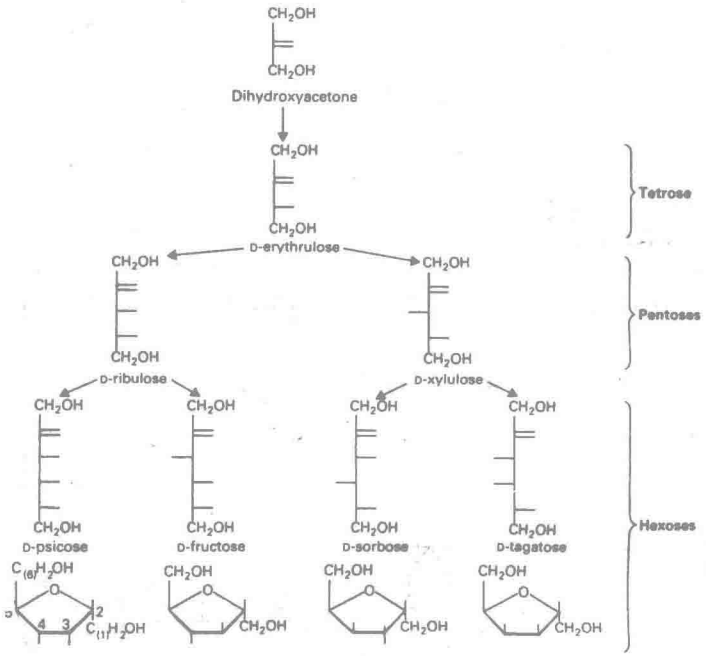


Fig. 2–3 The ketose sugar family. See bottom left-hand corner for ring carbon numbering and Fig. 2–2 for other details.

2.1.2 *Forming ring sugars from open chains*

Sugar chains are not rigid, but constantly waving around. As a result, simple sugars spend only a very small proportion of their time in the open chain configuration. In an aldose sugar such as glucose for example, the reactive C=O group on C-1 is frequently brought close enough, through flexing of the sugar chain, to interact with the OH on C-5. An exchange of covalent bonds takes place, resulting in the two 6-membered ring structures shown in Fig. 2–4.

Note how the ring O that originated from C-5 (tinted in the diagram) now forms a link between C-1 and C-5. Four of the ring carbons (C-1, 2, 3, 4) still carry OHs while the fifth (C-5) carries the CH_2OH of which C-6 forms a part. (The ring configuration, like the open chain formula, is frequently abbreviated. Single H atoms and ring Cs are omitted, while