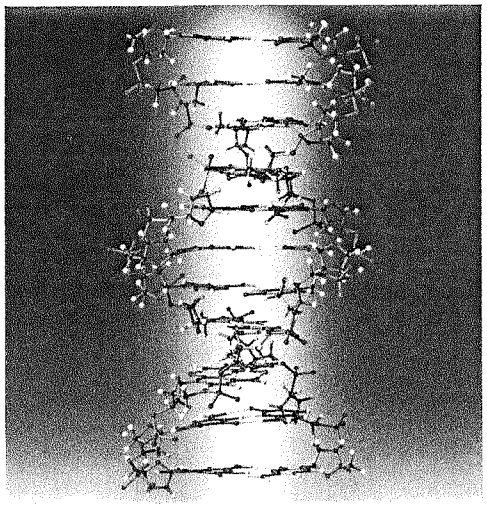
The characteristics and function of DNA are inherently linked to its structure. You can learn about the the structure of DNA and its related properties by building and exploring models using Orbit™ or Minit™ molecular modelling sets. These can be purchased assembled or as kits with detailed instructions and include explanation of the structure and related properties.

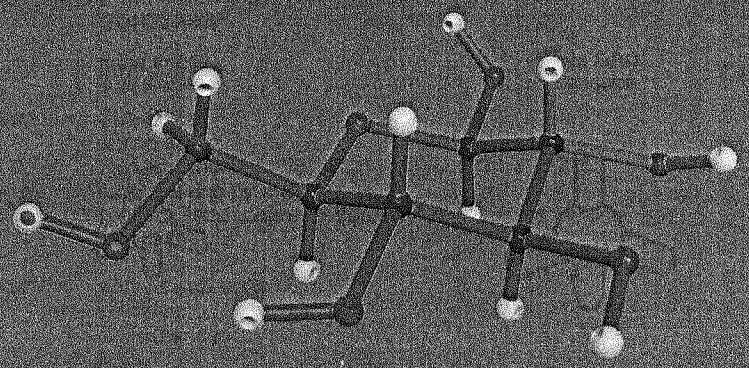


The Orbit Proview™ DNA model

Also available are models for diamond, graphite, sodium chloride, iron, ice, fluorite, zinc blende and lead iodide, wurtzite, rutile, caesium chloride, metal hexagonal close packing and metal cubic close packing.

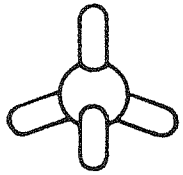
Modelling Biochemistry

Biochemistry activities with the Orbit™ and Minit™ molecular building systems.

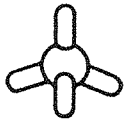


Cochranes of Oxford Ltd

Orbit atom
actual size



Minit atom
actual size



Published by Cochranes of Oxford Ltd.
Grove Farm Barns, Shipton-under-Wychwood,
Chipping Norton OX7 6DG, UK

Website: www.cochranes.co.uk
Email: sales@cochranes.co.uk

Cover picture: β-D-glucose (Orbit System)

Printed in the UK

Biochemistry

Instructions and notes on the use of the Minit and Orbit molecular building systems by **R S Lowrie** MA, D Phil, FRIC
Editorial consultant **Patrick Riley**, MB, PhD

CONTENTS

Introduction

The Minit and Orbit molecular building systems are designed to enable you to build molecular and crystal structures simply, accurately and at low cost. Molecules and crystals are represented by atom centres linked by tubes. This leads to a framework or skeletal representation of the molecule that is particularly good at showing the geometric structure of the substance. You can also perform 'reactions' with the models and examine different representations of valency and bonding.

The atoms consist of plastic centres having prongs set at the correct bond angles. Minit atom centres are 6mm in diameter and Orbit atom centres are 10 mm in diameter.

The centres are coloured according to the element or group of elements. In this booklet, an abbreviation is assigned to each colour that is usually the chemical symbol of the principal element that it represents. The principal elements and their colour coding are shown in Table I.

Table I Colour coding of atom centres

Colour of atom centre	*Principal element*	*Colour code*	*Colour of atom centre*	*Principal element*	*Colour code*
White	Hydrogen	H	Yellow	Sulphur	S
Black	Carbon	C	Purple	Phosphorus	P
Blue	Nitrogen	N	Green	Chlorine	Cl
Red	Oxygen	O	Silver/Grey	Metals	M

The centres have one or more prongs sticking out of them. The number of prongs and angles between them determine each centre's shape. For reference, each shape has a code letter as shown in Table II. The shape code letter is used in this book as a superscript after the colour code. Thus a red (oxygen) divalent 110° centre is denoted O^d whilst a red tetrahedral centre is referred to as O^k.

Table II

Code letters demoting the shapes of atom centres for biochemistry

Type of atom	*Code letter*	*Shape of atom model and bond angles*	*Type of atom*	*Code letter*	*Shape of atom model and bond angles*
Univalent	a		Trivalent	g	
Diunivalent	b			h	
	c			i	
	d			j	
	e		Tetrahedral	k	
			Octahedral	l	

In the Minit System the atom centre shapes d, g, h and i can be distinguished and orientated by the presence of one or more bars moulded on them, as shown in Table II. For instance, if shape g is used at the interface of five- and six-membered planar rings the 108° angle to be used in the five membered ring is marked with the bar and the 120° angle is the smaller of the remaining two. In the Orbit System atom centre shapes e – j have the relevant angles in degrees engraved on them.

Introduction to the bonds

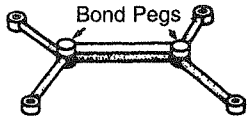
Plastic tubes or straws representing bonds join the atom centres. Covalent bonds are normally represented by green or grey straws; white straws can be used for other purposes, such as hydrogen bonds.

The length of the straw determines the bond length of the model. Tables III and IV give straw lengths to the nearest 0.5 cm for common bonds and shows how these are calculated. The straw length is less than the model bond length to allow for the radii of the two atom centres (Orbit 0.5cm each, Minit 0.3 cm each). To study the concepts in this booklet modelling with three bond lengths is sufficient. You can use the straws provided in your set, choosing lengths nearest to those given in table III.

For more accurate modelling you can cut straws to your desired length. Straws pre-cut to other lengths may also be available from your supplier. For demonstration models you may choose to use a larger scale.

When substituting one atom or group for another on a model, it is wise to remove the associated bond, as the replacement group may require a bond of a different length.

Multiple bonds may be simply represented by a single straw (not showing the multiplicity) or shown using multiple flexible tubes. Some sets include pegs for building multiple bonds in a different form from that shown in the text. The advantage of these pegs is that they show the presence of the multiple bond and restrict rotation around it without taking up a lot of space.



Double bonds have on occasion been omitted from drawings of models. This has occurred where rotation cannot occur around the double bond eg planar ring structures, and where rotation does not alter the shape of the molecule eg

Table III Minit System

Calculation of straw lengths to nearest half centimetre in the scale 100 pm (1Å) = 2 cm for covalent bonds

Your set contains straws cut to 1.5, 2.0 and 2.5 cm, together with longer straws which you can cut to your own lengths.

Bond	*Sum of covalent radii (pm)*	*internuclear distance (pm)*	*Converted to cm*	*Straw Length to nearest ½cm*
O–H	67 + 30	97	1.94	1.5
N–H	70 + 30	100	2.00	1.5
C–H	75 + 30	105	2.10	1.5
C=O carbonyl	67 + 55	122	2.44	2.0
S–H	100 + 30	130	2.60	2.0
C=N aromatic	70 + 62	132	2.64	2.0
C=C	67 + 67	134	2.68	2.0
C=C aromatic	70 + 70	140	2.80	2.0
C–O	77 + 67	144	2.88	2.5
C–N aliphatic	70 + 67	147	2.94	2.5
C–C	77 + 77	154	3.08	2.5
C–S	77 + 100	177	3.54	3.0
P–O	110 + 67	177	3.54	3.0
S–S	100 + 100	200	4.00	3.5
Hydrogen bond	–	–	–	3.5

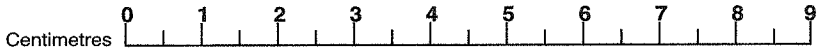


Table IV Orbit System

Calculation of straw lengths to nearest half centimetre in the scale 100 pm (1Å) = 3 cm for covalent bonds
The Orbit set contains straws cut to 2.0, 3.0, 3.5 cm and some longer straws which you can cut to your own lengths.

Bond	*Sum of covalent radii (pm)*	*internuclear distance (pm)*	*Converted to cm*	*Straw Length to nearest ½cm*
O–H	67 + 30	97	2.91	2.0
N–H	70 + 30	100	3.00	2.0
C–H	75 + 30	105	3.15	2.0
C=O carbonyl	67 + 55	122	3.66	2.5
S–H	100 + 30	130	3.90	3.0
C=N aromatic	70 + 62	132	3.96	3.0
C=C	67 + 67	134	4.02	3.0
C=C aromatic	70 + 70	140	4.20	3.0
C–O	77 + 67	144	4.32	3.5
C–N aliphatic	70 + 67	147	4.41	3.5
C–C	77 + 77	154	4.62	3.5
C–S	77 + 100	177	5.31	4.5
P–O	110 + 67	177	5.31	4.5
S–S	100 + 100	200	6.00	5.0
Hydrogen bond	–	–	–	5.0

Use of this booklet
The instructions given here may be used alongside any conventional textbook on structural biochemistry.

For the larger models, more than one kit may be needed. This is indicated in the text.

1. Sub-units

AMINO ACIDS
Amino acids are the sub-units, or building blocks, from which protein molecules are made. Only 20 occur commonly in nature, and these are all α-amino acids, that is, the amino group is attached to the carbon atom adjacent to the carboxyl group (the α-carbon atom). The general formula of α-amino acids is $NH_2.CHR.COOH$.

Glycine, aminoacetic acid, $NH_2.CH_2.COOH$
Glycine is the simplest amino acid.

1.1 *What is the formulation of the group R of the general formula, in glycine?*

In aqueous solution, and in the solid, glycine exists as a dipolar ion or 'zwitterion', $NH_3^+.CH2.COO^-$.

1.2 *Explain why a zwitterion forms, in terms of your knowledge of the chemical properties of the amino group and the carboxyl group.*

Figure 1 shows how to construct a model of the glycine zwitterion. Use the atomic centres shown and the closest available straw lengths to those in tables III and IV.

To make the non-ionic form of glycine, replace the N^k centre of figure 1 by an N^i centre and attach two hydrogen atoms to N^i. Remove one of the O^a centres from the carboxyl group, and replace it with a hydroxyl group, O^dH^a.

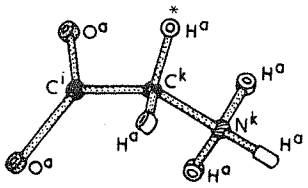


Figure 1 Glycine zwitterion

Alanine

Construct a second model of the un-ionized form of glycine: remove one of the hydrogen atoms (the one marked * in figure 1) from the α-carbon atom and replace with a methyl group $-CH_3$. You have made a model of L-alanine, (figure 2).

1.3 *What special property is possessed by the α-carbon atom in alanine?*

To help answer question 1.3, try exchanging the positions of the CH_3-group and the H-atom attached to the α-carbon centre, and observe whether the struction you obtain is the same or not.

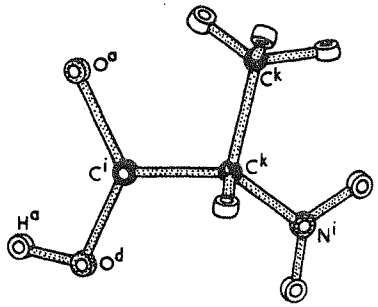


Figure 2 L-alanine

All the main amino acids that occur in proteins are asymmetric, except for glycine, and moreover they occur in the L-configuration in almost every case. Figure 2 represents the L-configuration of alanine.

1.4 *Predict what will happen to alanine if (a) aqueous hydrogen ions (b) aqueous hydroxide ions are added to its aqueous solution.*

The 'corn' rule

The rule for recognizing the L-isomer of an amino-acid is as follows: view the α-carbon atom along the H—C bond with the H atom towards you. In a clockwise direction the groups read CO - R - N, (CO of the carboxyl group, R the substituted side chain, N the amino group). Figure 2 shows this viewpoint.

Amino acids with two carboxyl groups

If the group R (in the general formula) itself contains a carboxyl group, the whole molecule will contain two such, groups. Examples are aspartic acid (R = $CH_2.COOH$, figure 3) and glutamic acid (R = $CH_2.CH_2.COOH$, figure 4).

Construct two models of the un-ionized form of glycine. Make models of the groups shown in figures 3 and 4. Hence construct the complete models of the two amino acids by substituting R for H in each case. Be careful to attach group R in the L-configuration (refer to figure 2).

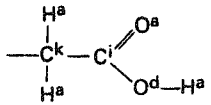


Figure 3 Side chain for aspartic acid

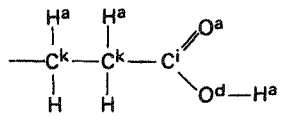


Figure 4 Side chain for glutamic acid

Amino acids with two amino groups

If the group R itself contains an amino group, the whole amino acid will contain two amino groups. The simplest naturally occurring example is L-lysine, where R = $CH_2.CH_2.CH_2.CH_2.NH_2$. Construct a model of the group R as in figure 5, and proceed as before to make a model of lysine in the L-configuration.

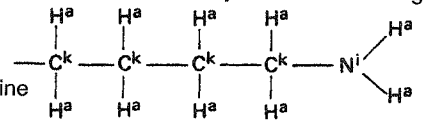
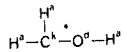
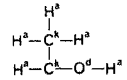
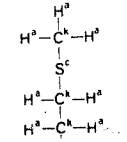
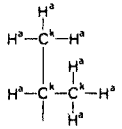
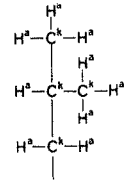
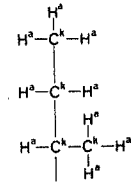


Figure 5 Side chain for lysine

The next section contains details for the construction of the remaining amino acids commonly found in nature. When constructing models of the non-ionic forms use N^i for the amino groups and C^i for the carboxyl groups throughout.

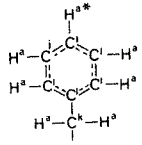
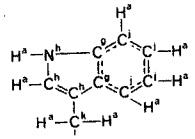
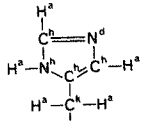
Other amino acids

The amino acids are known by three-letter (occasionally five-letter) abbreviations which are given in the tables below: these abbreviations are necessary when more complex structures are referred to in later sections.

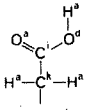
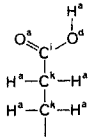
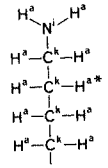
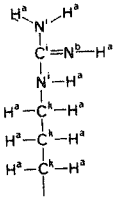
Name	*Abbrev-iation*	*Structural formula of R†*
a) Aliphatic		
glycine	gly	
alanine	ala	
cysteine	cys	
serine	ser	
threonine	thr	
methionine	met	
valine	val	
leucine	leu	
isoleucine	ile	

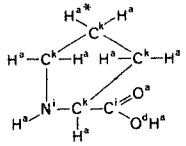
Use N^i for the amino group throughout, and C^i for the carboxyl group

†(The formulae include code letters for atomic centres)

Name	*Abbrev-iation*	*Structural formula of R†*
b) Aromatic		
phenylalanine	phe	
tyrosine	tyr	
as above, with O^dH^a in place of H^{a*}		
c) Heterocyclic		
tryptophan	try	
histidine	his	

Centres C^g, C^h, N^h:
See Table II for a guide to the angles and markings on these atom centres. Where shapes g or h is not available shapes j or i may be used in their place with the straws bending a little.

Name	*Abbrev-iation*	*Structural formula of R†*
d) Acidic		
aspartic acid	asp	
glutamic acid	glu	
e) Basic		
lysine	lys	
hydroxylysine As lysine, with O^dH^a in place of H^{a*}	hylys	
arginine	arg	

Name	Abbrev-iation	Structural formula of R†
f) Imino-acids		
proline	pro	
hydroxyproline As proline with O^dH^a in place of H^{a*}	hypro	

†*(The formulae include code letters for atomic centres)*

MONOSACCHARIDES

Monosaccharides are the sub-units from which more complex carbohydrates are built up. Monosaccharides are sugars, and their names end in -ose. They can be classified in terms of the number of carbon atoms, eg pentoses are simple sugars containing five carbon atoms, hexoses contain six carbon atoms etc. Monosaccharides are hydroxy-aldehydes or ketones, the simplest aldehyde sugar, or aldose, is the triose glyceraldehyde.

Glyceraldehyde, $C_3H_6O_3$

Figure 6 shows the details required for constructing a model of the simplest possible monosaccharide, glyceraldehyde. Note that the centre carbon atom is asymmetric: the diagram shows D-glyceraldehyde (sugars with the L-configuration are rare in nature).

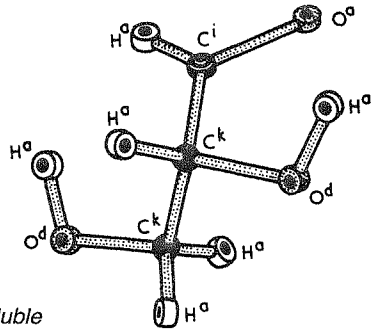


Figure 6 D-glyceraldehyde

1.5 *The general formula of a carbohydrate is $C_X(H_2O)_y$ Write the formula of glyceraldehyde in this form.*

1.6 *Predict whether glyceraldehyde will be soluble in water, from your knowledge of the physical properties of the groups present in glyceraldehyde.*

1.7 *What will be the expected action of glyceraldehyde on Fehling's solution?*

Convention used in text-books for asymmetric carbon atoms

It is easy to make accurate models of sugars from text-book formulae, provided the following universally used convention is remembered: bonds drawn horizontally are assumed to come up from the carbon centre towards the reader, whereas bonds drawn vertically are assumed to project down from the carbon centre away from the reader. Hence D-glyceraldehyde is show thus:

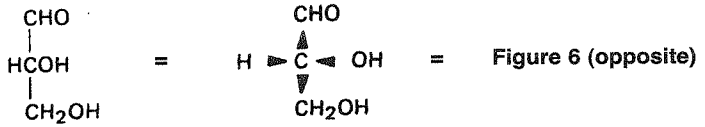


Figure 6 (opposite)

Pentoses and hexoses

Tetroses [general formula $C_4(H_2O)_4$] are rare in nature, but pentoses [$C_5(H_2O)_5$] and hexoses [$C_6(H_2O)_6$] are common. D-ribose is a pentose of biochemical importance. Figure 7 shows its chemical formula written according to the convention described. The numbers on the carbon atoms will be needed on page 15.

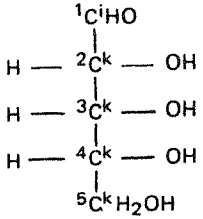


Figure 7 D-ribose

1.8 *How many of the carbon atoms in D-ribose are asymmetric?*

1.9 *How many isomers does this give rise to?*

One of the commonest hexoses is D-glucose. Figure 8 shows the conventional formula of D-glucose.

Construct models of D-ribose and D-glucose as depicted in figures 7 and 8; keep these models for the next exercise.

Ketoses have the carbonyl group in the chain instead of at one end, they have one less asymmetric carbon atom than the corresponding aldoses. A common ketose is the ketohexose fructose. Figure 8a shows the conventional formula of D-fructose.

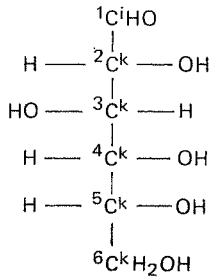


Figure 8 D-glucose

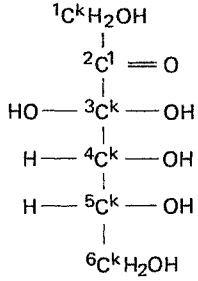


Figure 8a D-fructose

FURANOSE AND PYRANOSE RINGS

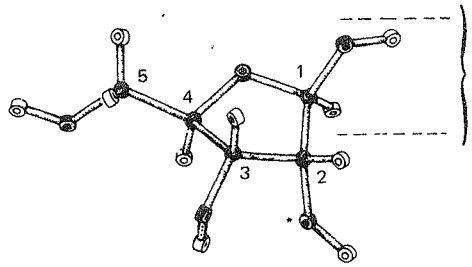
The straight-chain forms of the pentoses and hexoses are able to cyclise, ie form five-membered (furanose) or six-membered (pyranose) rings. To convert D-ribose to its furanose form proceed as follows:

(a) replace carbon centre number 1 (figure 7) with a C^k centre, and replace its oxygen centre with a two-pronged centre, O^d;

(b) replace the hydrogen atom attached to carbon atom number 1; leave the fourth C^k prong free;

(c) remove the hydrogen centre belonging to the OH group on carbon centre number 4, and attach it to the spare prong of the O^d centre attached to carbon atom number 1;

(d) complete the five-membered ring by forming a bond between carbon atom number 1 (spare prong) and the oxygen atom attached to carbon atom number 4. The result should appear as in figure 9, but there is one possible difference that it is important to notice. It is possible that the H and OH groups attached to carbon atom number 1 may come out in exchanged positions. The alternative that you obtain depends entirely on the choice of 'spare prong' made in (b) above, and that choice was quite arbitrary. You have introduced an extra asymmetric centre on cyclisation, at carbon atom number 1, and this accounts for the possibility of two isomers related to this centre.



This configuration represents the β isomer of D-ribofuranose

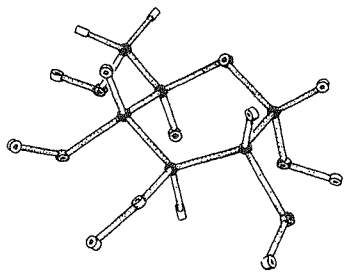
Figure 9 β-D-Ribofuranose

Deoxyribose

An important monomer unit is 2-deoxyribose. It occurs in DNA (page 39). Construct a model of it by replacing the hydroxyl group marked * in figure 9 with a single hydrogen atom.

α-and β-glucose

Glucose forms a pyranose ring by cyclisation between carbon atoms 1 and 5 (figure 8). To convert your straight-chain model of D-glucose into its pyranose form, proceed exactly as with ribose, but use carbon atom number 5 instead of number 4 in parts (c) and (d) of the instructions. Carbon atom number 1 has now become a new asymmetric centre and it has two possible configurations giving two isomers of D-glucose, called α-D-glucose and β-D-glucose. (Figure 10).



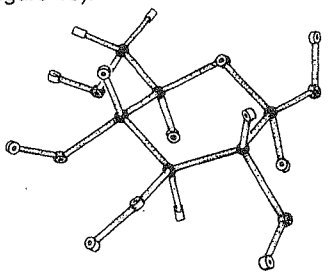


Figure 10 (a) α-D-glucose (b) β-D-glucose

1.10 *What is the total number of asymmetric carbon atoms in α-D-glucose?*

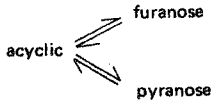
1.11 *How many enantiomers of α-D-glucose are possible?*

Notice that the 6-membered ring is flexible. Two common forms are the chair and boat shown in Figure 10(c), both these can exist in two forms illustrated for β-D-glucose in Figures 10(d) and 10(e). Considerable steric repulsion occurs in the boat forms and the chair forms are more stable.

The two forms of the chair are interchangeable by holding the model lightly but firmly at two opposite ring atoms, firm pressure in an up and down direction will flip the ring from chair to boat to the other chair form. As the rings change from one chair form to the other so the groups which were perpendicular to the ring plane (axial groups marked ax) swivel so that they are approximately horizontal to the plane of the ring (equatorial groups marked eq).

Steric hindrance occurs with large groups in axial positions so the most favoured form is that with the most bulky groups in equatorial positions. β-D-glucose is the only D-hexose capable of having all of its large groups (-OH and $-CH_2OH$) equatorial in the pyranose form. It is significant that D-glucose plays such a major role in biochemistry.

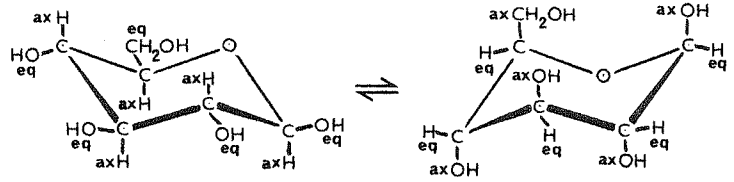
The cyclisation of monosaccharides is an equilibrium reaction



and more than one conformation may be present in solution depending on their relative stabilities. In an aqueous solution of D-glucose there is approximately 64% β-pyranose and 36% α-pyranose, less than 1% furanose. Aqueous D-ribose however has approximately 56% β-pyranose, 20% α-pyranose, 16% β-furanose and 6% α-furanose.

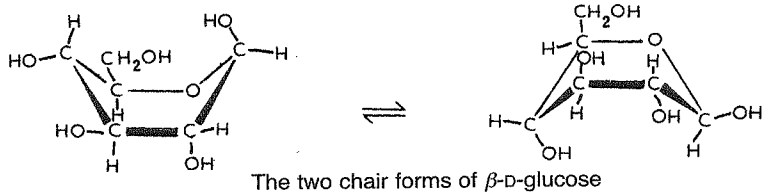
The acyclic form is never found in appreciable quantities but can form and react to give typical aldehyde reactions such as the reduction of Fehlings solution.

Figure 10(c)



The two chair forms of β-D-glucose

Figure 10(d)



The two chair forms of β-D-glucose

Figure 10(e)

α-and β-fructose

Fructose forms a furanose ring by cyclisation between carbon atoms 2 and 5 (Figure 8a). To convert your straight chain model of d-fructose into its furanose form proceed as for glucose but change carbon atom number 2 instead of

number 1 and replace the -CH_2OH group instead of the hydrogen atom in (b). Carbon atom number 2 is now asymmetric and α- and β-isomers are possible.

GLYCEROL, FATTY ACIDS AND STEROIDS

The purpose of this section is to introduce some of the molecular sub-units which together build up the class of compounds known as lipids. A lipid is any water insoluble material that can be extracted from cells by the action of an organic solvent. There are many diverse functions. The classification can be summarized as follows:

Class		Containing the sub-units
Lipids	fats and oils	glycerol, fatty acids
	waxes	long-chain alcohols, fatty acids
	phospholipids	glycerol, fatty acids, phosphate organic bases
	steroids	substituted polycyclic hydrocarbons

Glycerol, propane-1,2,3-triol, $C_3H_8O_3$

Construct a model of glycerol as in figure 11.

1.12 *Is glycerol a carbohydrate?*

1.13 *From your knowledge of the bonds present, estimate the boiling point and water-solubility of glycerol.*

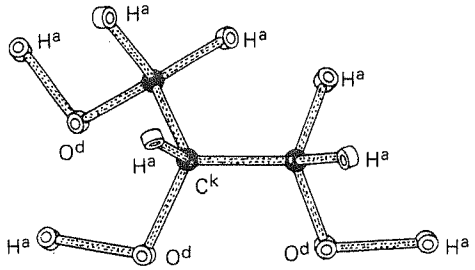


Figure 11 Glycerol

Fatty acids $CH_3(CH_2)_nCOOH$

The simplest fatty acid is acetic acid, CH_3COOH ($n = 0$). If $n = 14$ the acid is palmitic acid; $n = 16$ is stearic acid. Models of these acids may be constructed using C^i for the carboxyl group, and C^k for all the other carbon centres. (Figure 12).

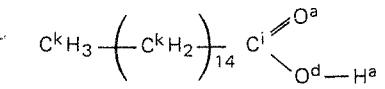


Figure 12 Palmitic acid

In addition, unsaturated fatty acids such as oleic acid, $CH_3(CH_2)_7CH = CH(CH_2)_7COOH$, are common constituents of fats and oils such as olive oil. Construct a model of oleic acid using Ck centres throughout the carbon chain. Two of the carbon atoms are linked by a double bond (figure 13) and this can be made if flexible straws are used as shown. Note that the model should show the double bond in the *cis* configuration. *Trans* acids are rare.

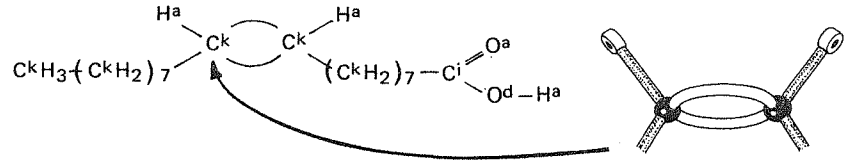


Figure 13 Oleic acid

It should be noted that, although double bonds occur frequently in biochemical molecules, there is no point in using these flexible straws to show them in the model unless they are essential to obtain the correct shape of the molecule. For this reason, the C = O group is best shown as 'C^i—O^a since rotation about the bond has no effect on shape. Similarly, the bonds in aromatic rings and peptide links are shown by single straws even though their bond order is greater than unity.

Steroids

Steroids are compounds based upon the polycyclic nucleus depicted in figure 14. The numbers are those conventionally used for identifying carbon atoms when naming steroids. Figure 15 shows a drawing of a model of the basic steroid skeleton. Note that only carbon atoms 5 and 6 are double-bonded: these should

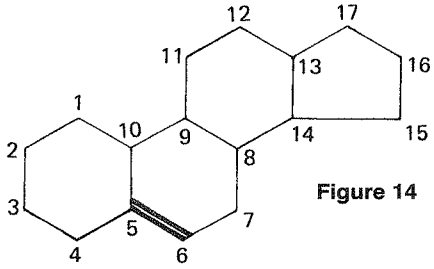


Figure 14

be represented by C^k centres in the model, but flexible straws should be used for the links.

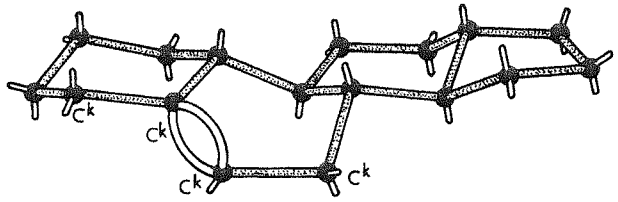


Figure 15 The basic steroid skeleton

To construct a model of cholesterol, attach a hydroxyl group to C_3, methyl groups to C_{10} and C_{13}, and the group—$CH(CH_3).CH_2.CH_2.CH_2.CH(CH_3)_2$ to C_{17}. Figure 16).

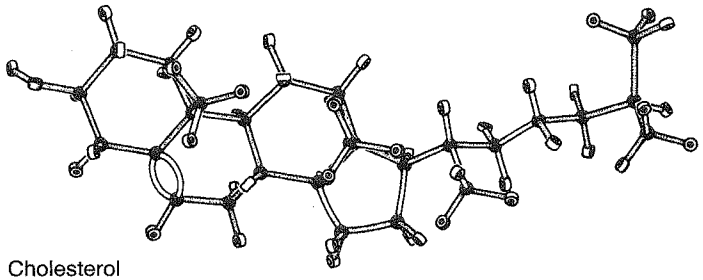


Figure 16 Cholesterol

1.14 *How many asymmetric centres are there in cholesterol?*

1.15 *How many isomers are therefore theoretically possible?*

It is interesting to note that only one form of cholesterol is known in nature, namely that drawn in figure 16. Attempt to copy this diagram exactly when constructing the model.

ORGANIC BASIS

Purine and pyrimidine bases are important sub-units of nucleic acids, the substances that carry genetic information.

Pyrimidine and its derivatives

Pyrimidine may be regarded as similar to benzene with nitrogen atoms in place of two of the carbon atoms. To construct a model build a hexagonal ring as shown in figure 17 and add hydrogen atoms to the carbon atoms as shown. There is no point in attempting to show double bonds as these are delocalised in a similar manner to the electron system in benzene. The ring will remain planar without the assistance of double-bond straws. Use N^J and C^J atoms in the ring.

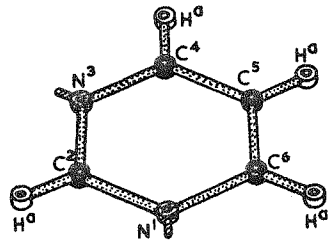
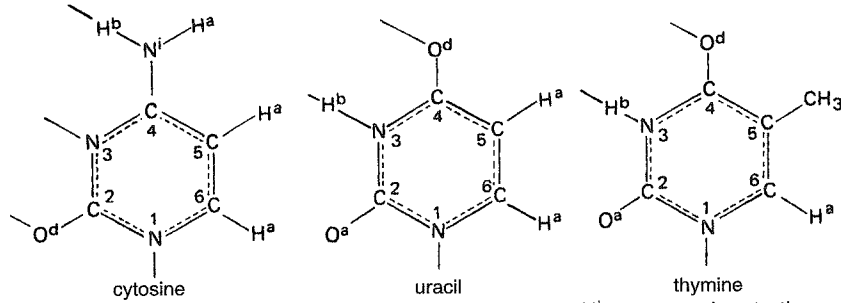


Figure 17 Pyrimidine

Now construct models of the three main pyrimidine derivatives that occur in nature. They are:



Use O^d centres to denote the oxygen atoms and a planar N^j centre to denote the nitrogen atom in the amino-group. Use three-pronged centres N^j for both nitrogen atoms in the ring, and leave the third prong of N_1 spare. The reason for this will be explained on page 39.

Purine and its derivatives
Figure 18 shows a model of purine. Where atom shapes g and h are used take care of their orientation. If the atom centre shapes g or h is not available shapes j or i may be used with the straws bending a little.

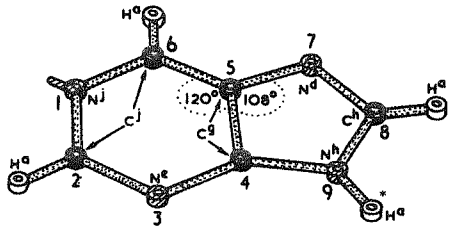


Figure 18 Purine

Now construct models of the two purine derivatives that occur in nature. They are:

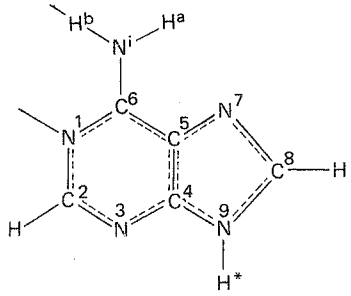


Figure 19 Adenine

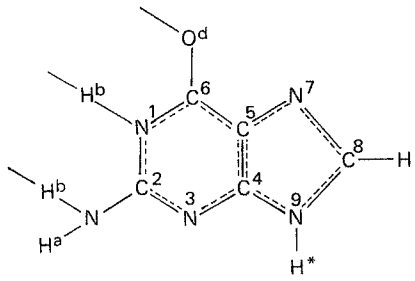


Figure 20 Guanine

1.16 *What type of chemical force can exist between molecules containing hydroxy-groups or amino-groups?*

The answer to question 1.16 is given on page 39 and is important to the function of the purine and pyrimidine bases in nucleic acids.

(*Hydrogen atoms marked thus in figures 18, 19 and 20 are not present in nucleic acids).

Another organic base
A common base in phospholipid structures is choline, $[(CH_3)_3N.CH_2CH_2OH]^+OH^-$. Construct a model of the cation, using an N^k centre to represent the nitrogen atom, and C^k centres for all the carbon atoms. (Figure 21).

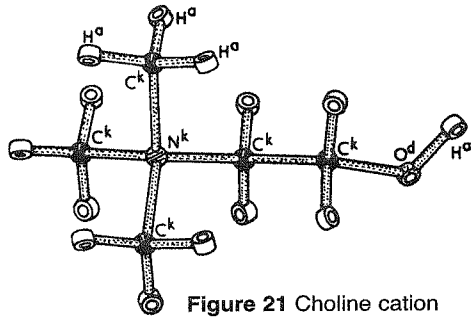


Figure 21 Choline cation

2. Linked sub-units

PEPTIDES—THE PEPTIDE LINK
The peptide link,—CO.NH—can be regarded as resulting from the condensation reaction between two amino acids:

$$-COOH + H_2N- \longrightarrow -CO.NH- + H_2O$$

The molecule resulting from the linkage of two amino acids is termed a dipeptide, three amino acids give a tripeptide, and so on. The nitrogen atom in the peptide link is planar, not pyramidal, and the link is constructed as in figure 22. Alternative (a) should be used in simple models of peptides; alternative (b) will be required when depicting secondary structure (page 33) involving hydrogen bonds between adjacent peptide chains.

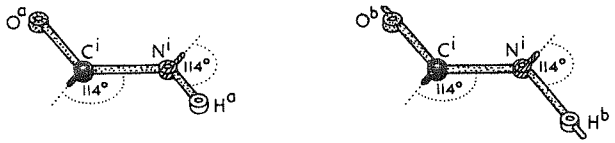


Figure 22 (a) Peptide link unit for a simple peptide (b) Peptide link modified to accept hydrogen bonds.

2.1 *As the nitrogen atom in the peptide link is planar, what does this tell you about the lone lone pairs of electrons on the nitrogen atom?*

Glutathione (glu-cys-gly)
Glutathione is a tripeptide that plays an important role in cell function. Construct separate models of glutamic acid (glu), cysteine (cys) and glycine (gly) as described on page 10. Throughout use planar nitrogen centres, N^{l}, for the amino group nitrogen atoms, and C^{l} centres for the carboxyl group carbon atoms.

Arrange the models in the above order reading from left to right, each with its amino group on the left and its carboxyl group on the right. Now form peptide links between the molecules as described in the previous paragraph. Figure 23 shows the structure that should result.

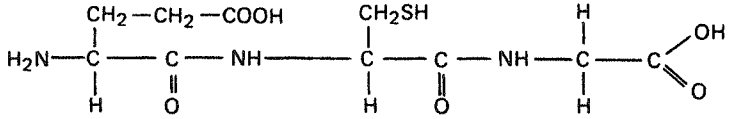


Figure 23 Glutathione

Cysteine bridges and cyclic peptides
Two cysteine residues in separate parts of a polypeptide chain may link together by oxidation (removal of two hydrogen atoms):

$$-SH + HS- \longrightarrow -S-S- + 2[H]$$

This reaction leads to cross-linking and cyclisation. To illustrate the effect with a simple example, make a model of part of the hormone oxytocin. Refer to page 10 for the amino acid structures of this peptide:

pro—cys—asp—glu—ile—tyr—cys

Remember to have the amino groups on the left each time, linking the residues in the above order reading from left to right. Proline therefore has a free amino group, and the end cysteine residue a free carboxyl group.

Now coil the model round so that the SH-groups of the two cysteine residues are adjacent. Remove the two hydrogen atoms and form a sulphur bridge, (figure 24).

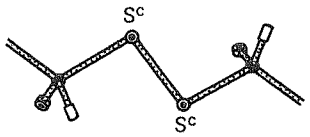
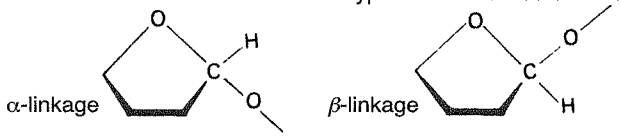


Figure 24 Disulphide

2.2. *The cysteine bridge is one way in which cross-links can be established between adjacent peptide chains what other forms of cross-bridge linking force can exist?*

DISACCHARIDES
α and β linkages between monosaccharides
On page 14 are described the different configurations, α and β, of the groups attached to the carbon atom where the ring is formed. The configuration is important when considering the formation of links between monosaccharide sub-units to form disaccharides. The different types of link can be summarized thus:



Maltose (glucose-1α, 4-glucose), $C_{12}H_{22}O_{11}$
First construct two models of α-D-glucose (page 15). Now form a link between carbon atom no 1 of one model (the atom which determines the α-configuration) and carbon atom no 4 of the other model. Figure 25 shows the structural formulae written using the convention that most text-books adopt—the hexagonal ring is simplified by imagining it to be planar, and the upward and downward bonds projecting from the ring are imagined to be exactly vertical. Figure 26 shows the actual shape of the model that results.

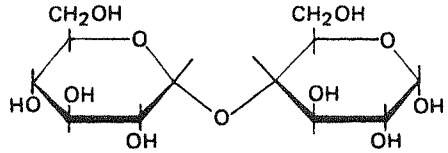


Figure 25 Maltose

Cellobiose (glucose-1β, 4-glucose), $C_{12}H_{22}O_{11}$

Construct two models of, β-D-glucose (page 15). Form a link between carbon atom no 1 of one model (the atom which determines the β-configuration) and carbon atom no 4 of the other model. Figure 27 shows the conventional structural formula, and figure 28 shows the actual shape of the model that results.

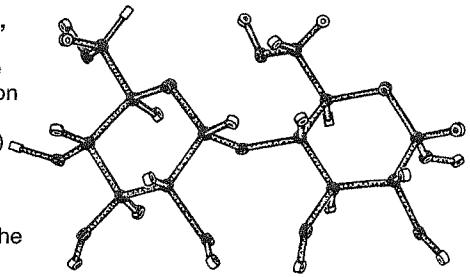


Figure 26 Maltose

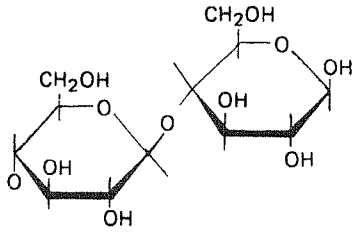


Figure 27 Cellobiose

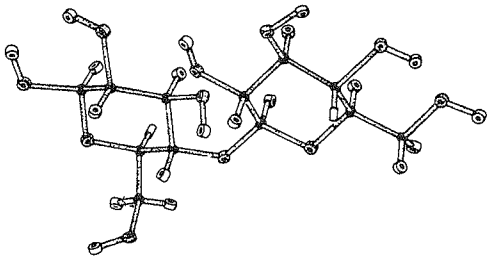


Figure 28 Cellobiose

Note that the α and β-links can permit free rotation. Rotate the right-hand glucose half of the model depicted in figure 28, through 180° with respect to the left-hand half. The configuration of cellobiose that results is the basis of the polysaccharide cellulose (page 36).

The type of linkage, whether α or β, is important in determining the configuration of the long chain that results when many monosaccharides are linked together.

2.3 *Why is a given enzyme often specific for one particular type of linkage?*

Sucrose, (glucose-1α, 2β-fructose), $C_{12}H_{22}O_{11}$

Sucrose is common sugar, ie 'cane-sugar or 'beet sugar'. Construct a model of β-D-fructose (furanose form) as shown in figure 29. Construct a model of α-D-glucose and then link it up with β-D-fructofuranose as illustrated in figure 30 and 31.

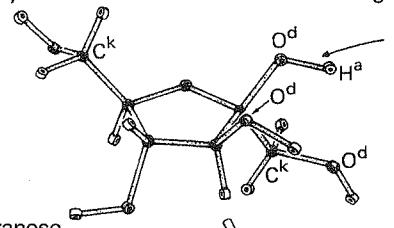


Figure 29 β-D-fructofuranose

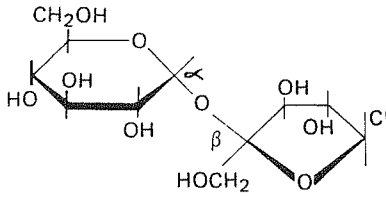


Figure 30 Sucrose

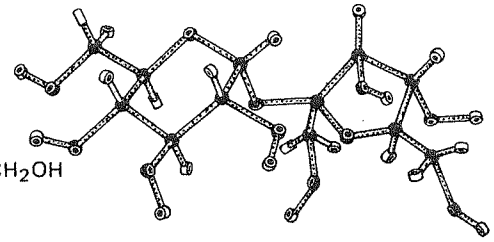


Figure 31 Sucrose

LIPIDS

Glycerides

Fats are esters of the trihydric alcohol glycerol (propane-1, 2, 3-triol, page 18) and saturated or unsaturated fatty acids (page 19).

Construct models of glycerol and oleic acid as detailed on page 18. Now detach the OH-group from the carboxyl end of the oleic acid molecule, and remove the hydrogen centre from the 1-hydroxyl group of glycerol. Form an ester link between the two parts, and form a water molecule from the detached fragments. You have made a model of a monoglyceride, glyceryl mono-oleate (figure 32).

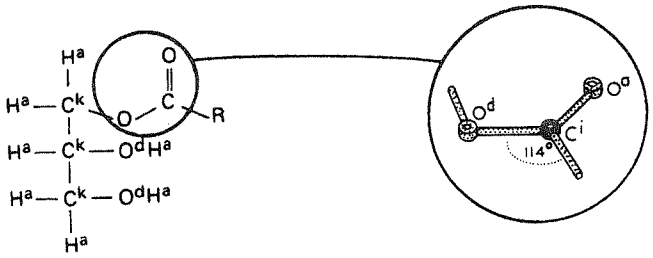


Figure 32 Glyceryl mono-oleate [R=CH3(CH2)7CH=CH(CH2)7—, see page 18]

All three hydroxyl groups of glycerol are able to participate in esterification. If desired, you may construct a model of glyceryl trioleate, the chief constituent of olive oil (figure 33), but you will need parts from more than one set.

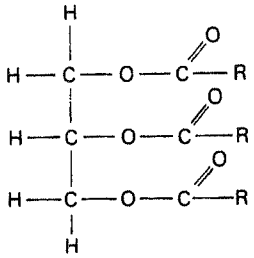


Figure 33 Glyceryl trioleate

Triglycerides are important energy stores in living cells, their 'combustion' to carbon dioxide and water being an energy releasing process.

Phospholipids

Phospholipids are fat-like substances closely related to glycerides, but containing phosphorus in the form of a phosphate group.

Construct a phosphate group, HPO_4, as shown in figure 34. Note that three of the oxygen centres are O^d and one O^a. Two of the O^d centres have spare prongs to accept further groups.

Construct a model of the choline cation, $[(CH_3)_3NCH_2CH_2OH]^+$ (page 23). Remove the OH-group from it, and attach it to the phosphate group using one of the spare O^d prongs on the phosphate group (figure 35).

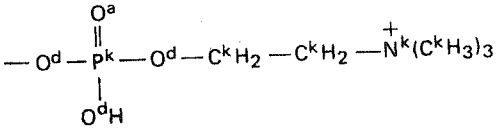


Figure 35

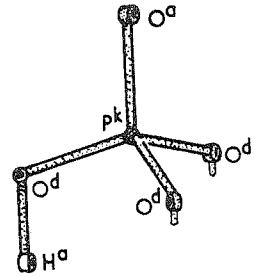


Figure 34

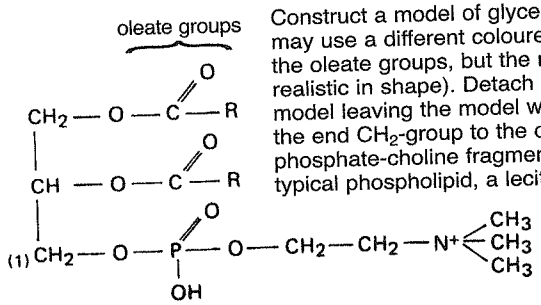


Figure 36 A lecithin

Construct a model of glyceryl trioleate (to save time, you may use a different coloured centre, eg Hal^a to depict the oleate groups, but the resultant model will not be realistic in shape). Detach the 1-oleate group from this model leaving the model with an end CH_2-group. Attach the end CH_2-group to the other spare O^d prong of the phosphate-choline fragment. The model now depicts a typical phospholipid, a lecithin (figure 36).

NUCLEOSIDES AND NUCLEOTIDES

A nucleoside is a linked sub-unit composed of sugar + base.
A nucleotide is a linked sub-unit composed of sugar + base + phosphate group.

Nucleosides

Construct models of the five bases described on pages 21 and 22, and also the furanose form of ribose (page 15). To construct a nucleotide. it is necessary to detach the OH-group from the C_1 atom of the ribose ring, and form a link with either the N_9 atom of adenine or guanine, or the N_1 atom of cytosine, uracil or thymine, using the spare nitrogen prong left for that purpose. Figures 37 and 38 show models of the nucleosides that result from combining ribose with adenine and cytosine respectively.

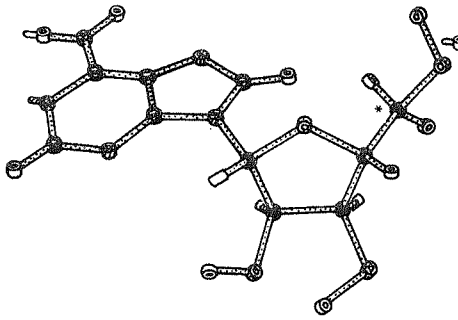


Figure 37 Adenine + ribose = adenosine

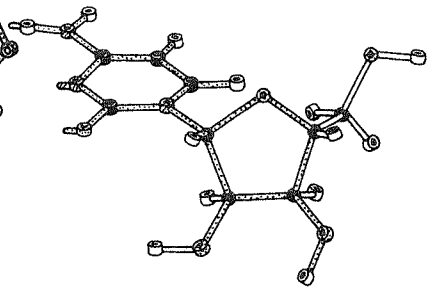
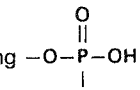


Figure 38 Cytosine + ribose = cytidine

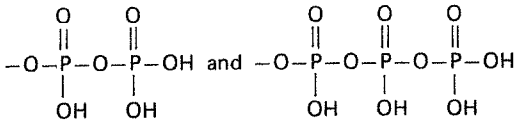
Nucleotides

Construct a phosphate grouping  in the manner described on page 29.

Detach the OH-group from the C_5 atom of the ribose ring in adenosine (marked * in figure 37) and attached the phosphate grouping. You have now made a model of adenosine monophosphate, AMP.

ADP and ATP

Construct models of the groups



using P^k centres for the phosphorus atoms. Detach the monophosphate group from your model of adenosine monophosphate, AMP. Attach each of the groups depicted above in turn to the atom marked * in figure 37. You have made models of adenosine diphosphate, ADP, and adenosine triphosphate, ATP.

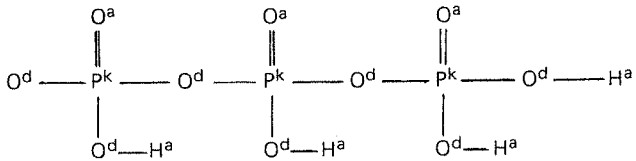


Figure 39 Triphosphate group for ATP

ATP is a major carrier of energy in cells.

Nicotinamide adenine dinucleotide, NAD

It is possible for two nucleotides to link together via their phosphate groups, and such substances are of importance in the electron exchange processes that occur during cellular respiration.

Construct a model of nicotinamide (figure 40) and its monophosphate nucleotide (figure 41). Detach the end OH-group from your model of AMP made earlier, and detach a hydrogen centre from one of the OH-groups in the phosphate part of the nicotinamide nucleotide. Figure 42 shows the structure of the complex dinucleotide that results.

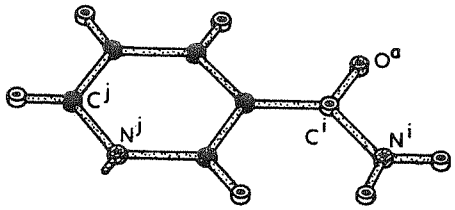


Figure 40 Nicotinamide

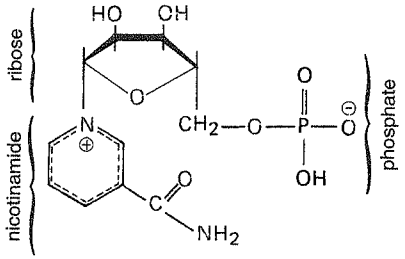


Figure 41

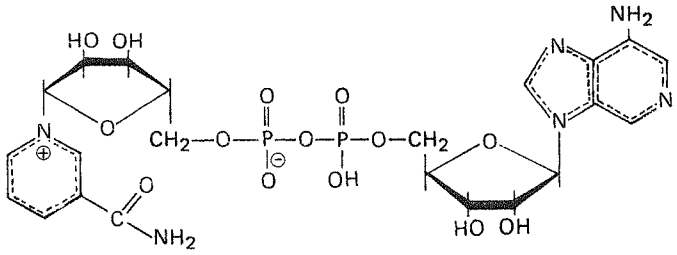


Figure 42 NAD

3. Complex molecules

PROTEINS

Primary structure of a protein

A protein is a high molecular weight polypeptide (page 23). Construct a number of models (about ten) a simple amino acids (page 7). It does no particularly matter which are chosen, but for ease of construction a simple derivative of glycine such as L-alanine is recommended: take care that all the amino acid models are in the L-form. Omit the terminal hydrogens and OH group so the amino acids are ready to join together.

Arrange your models in a row, with the amino groups on the left. Now connect them in the form of peptide links as described in figure 22 (b). You have now made a short length of a protein molecule. Note that H^b and O^b centres are required when forming the peptide links.

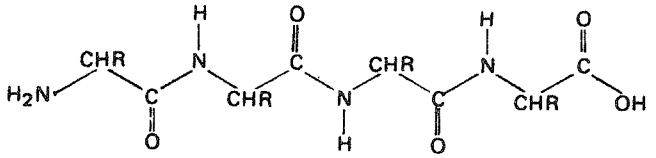
3.1 Write down the structure of the protein you have made, using standard three letter abbreviations.

3.2 Polypeptides are classed as proteins when their molecular weight is greater than about 5000. Estimate approximately how many amino acid residues are required to make a protein of this molecular weight, assuming a fairly random mixture of amino acids.

The information in your answer to question 3.1 gives us the primary structure of the protein-it tells us nothing about the overall shape of the molecule.

Secondary structure-sheets

Divide your primary protein into two approximately equal lengths and arrange each with its backbone in an extended zig-zag of this type.



Now place the two sections adjacent to one another so that the N — H bonds in one chain match up with the C = O bonds in the other chain, and vice versa. Using 3.5 cm (Orbit 5 cm) white straws to represent hydrogen bonds make connections between the groups.

The result is described as a β-pleated sheet. Figure 43 shows an example of the configuration that results, using L-alanine as the monomer throughout.

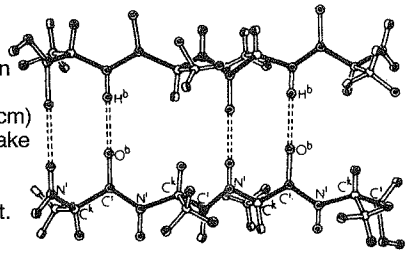


Figure 43 β-pleated protein sheet

Secondary structure—the alpha helix
The alpha-helix is a right-handed spiral structure in which adjacent coils are spaced apart by hydrogen bonds similar to those just described. Return your protein to one long chain and twist the backbone of your model carefully into a spiral form, so that the N — H bond in peptide link number 1, pointing upwards, matches up with the C = O bond in peptide link number 4, pointing downwards. Insert a straw to represent a hydrogen bond. Continue in like manner, aligning the upward-pointing N — H bond of peptide link number 2 with the downward point C = O bond of peptide link number 5, etc, inserting hydrogen bond straws as you go along. Figure 44 shows a drawing of the model.

3.3 *Explain the effects of (a) gentle heating and (b) change of pH upon the secondary structure of a protein.*

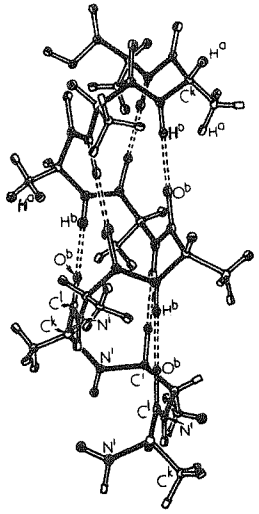


Figure 44 Model of the alpha-helix, using L-alanine as the monomer throughout

Tertiary structure
Superimposed upon the secondary (eg alpha-helix) structure of a protein is a higher order of structure called 'tertiary structure'. The alpha-helix coils may themselves be coiled, cf the 'coiled-coil' filament of a light bulb.

Adjacent coils of the super helix may be held together partly by hydrogen bonds, and partly by cysteine bridges. To illustrate tertiary structure make two short alpha-helices each containing a cysteine residue, and make a bridge between the two sulphur atoms.

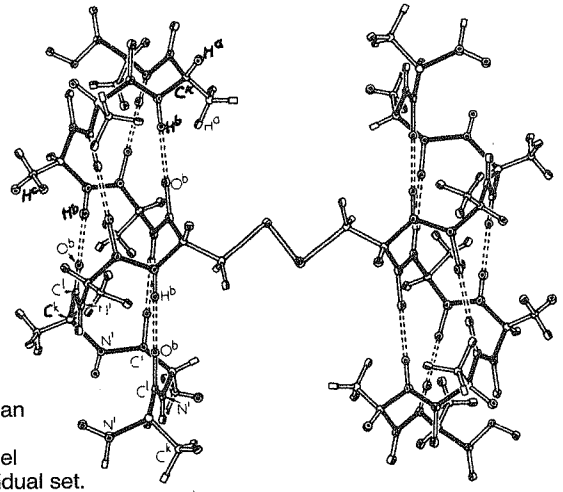


Figure 45 Model to illustrate an element of tertiary structure (disulphide bridge). This model requires more than one individual set.

POLYSACCHARIDES

Cellulose
Cellulose consists of long chain molecules each containing several hundred β-D-glucose residues linked at their 1 and 4 positions.

Construct three models of β-D-glucose (page 15). Link two together as described on page 26 to form cellobiose. Link the third glucose residue and note the straight chain configuration of the resultant polymer model. Note also that a molecule of water is eliminated as each link is formed. The polymer formed is called cellulose.

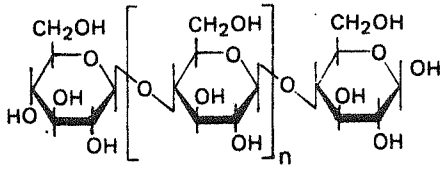


Figure 46 Cellulose (the β-links are shown in this way for convenience in drawing). (cf figure 27)

Amylose
Amylose comprises some 35% of starch (the remainder being amylopectin, see below). Make a model as follows. Make several models of α-D-glucose (page 15) and connect them at their 1 and 4 positions in the form of a chain. Each link is identical to that of maltose (page 26). Note that, unlike cellulose, amylose polymer has a tendency to form coils rather than straight chains.

3.4 *Explain in terms of structure why cellulose is a fibrous structure whereas amylose is not.*

Amylopectin
Amylopectin, the other main constituent of starch, differs from amylose in having branched chains in its structure.

Take your model of amylose, and form a 1α-6 link, between carbon atom number 6 of one of the α-glucose residues in the amylose and carbon atom number 1 of another α-D-glucose model. Figure 47 shows diagrammatically how the link should be formed.

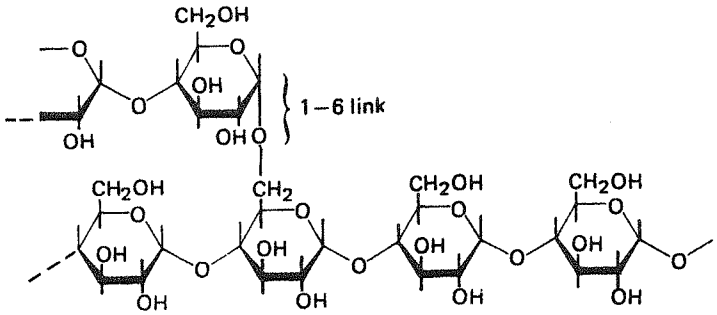


Figure 47 Amylopectin, side-chain formation

3.5 *Starch can be broken down into a disaccharide by means of an enzyme, amylase. Which disaccharide will be produced?*

Glycogen is a polysaccharide not unlike amylose, but more highly branched. It is the main energy store for animals.

RIBONUCLEIC ACID, RNA
Ribonucleic acid, or RNA, is a polynucleotide. It is a single strand consisting of a sugar-phosphate 'backbone' with the individual bases (adenine, guanine, cytosine and uracil, page 21) protruding. There are alternative procedures for building up a model, depending on whether or not you have a supply of individual nucleotide models to start with.

If you have no nucleotide models, construct first the sugar-phosphate backbone. Make four models of β-D-ribofuranose (page 15). Detach the OH-group from carbon atom number 5, and also from carbon atom number 3 (figure 48).

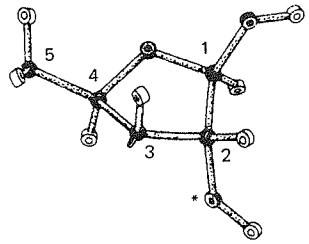
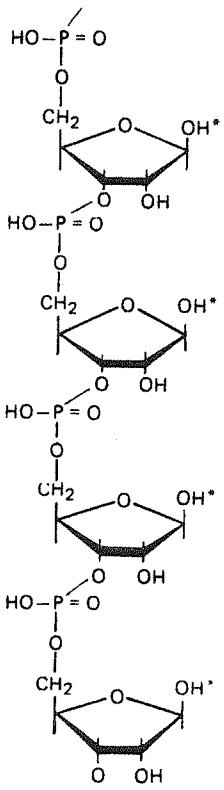
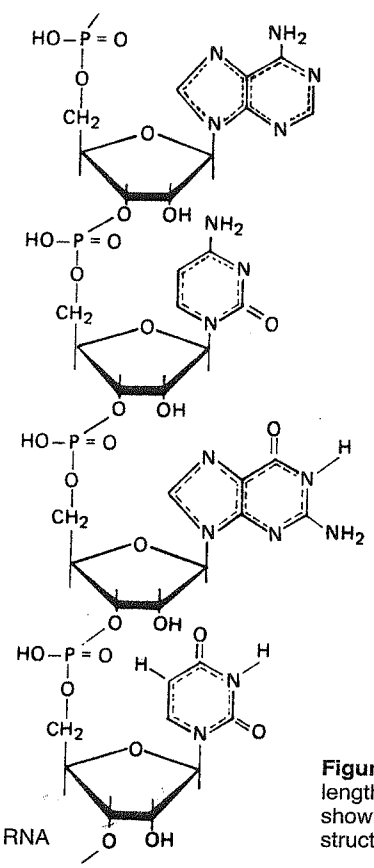


Figure 48 Ribose unit for constructing RNA backbone (cf figure 9)





adenine

cytosine

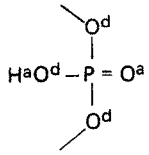
guanine

uracil

Figure 49 The ribose phosphate 'backbone' for RNA

Figure 50 A short length of RNA showing primary structure

Construct four phosphate groups



Link phosphate and ribose groups alternately as illustrated in figure 49.

Construct models of the four bases (page 21). Remove the OH-groups from the carbon atoms number 1 (marked * in figure 49) and attach the bases via the correct nitrogen atom. The order of attachment of the bases along the backbone does not matter for the purposes of this model.

Figure 50 shows a diagram which may be used as a guide. If starting with ready made nucleotides, link them as in figure 50, always ensuring that the phosphate group is attached to carbon atom number 3 of the ribose ring.

DEOXYRIBONUCLEIC ACID, DNA

The primary structure of DNA differs from that of RNA in that its backbone has deoxyribose (page 15) in place of ribose. Also the base thymine is present whereas uracil is absent.

The secondary structure of DNA is a double helix, with hydrogen bonds bridging the two strands. The configuration is such that base-pairs form, this pairing off being an important factor in the self-replication of DNA.

First construct two short 'backbones', each containing two phosphate and two deoxy-ribose residues, linking the residues as described for RNA.

Next form 'base-pairs' thymine-adenine using two hydrogen bonds (figure 51) and cytosine-guanine using three hydrogen bonds (figure 52).

Take one of the base-pairs, and attach one of the short backbones to it (figure 53). The mode of attachment is exactly as for RNA, but this time be careful to have the backbone angled in the general direction indicated in the diagram. Next add the second 'backbone' to the other base of the base-pair, (figure 54). Finally connect the second base-pair across the other two sugar residues in the backbone (figure 55).

For a larger model of DNA, it is necessary to provide some support, and a clamp stand is necessary. Each base-pair is mounted on the stand using wire supports, with a distance of 6.5 cm (Orbit 10 cm) between the planes. The angle between the horizontal axes of adjacent base pairs is 36°. The sequence of base-pairs forms a right handed helix.

Figure 51 Adenine-thymine

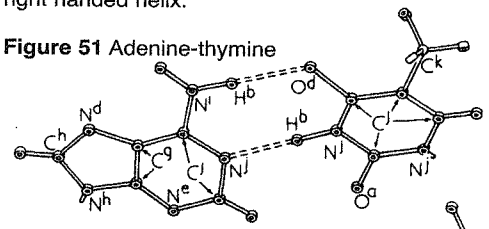
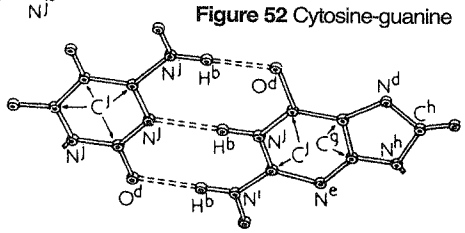


Figure 52 Cytosine-guanine



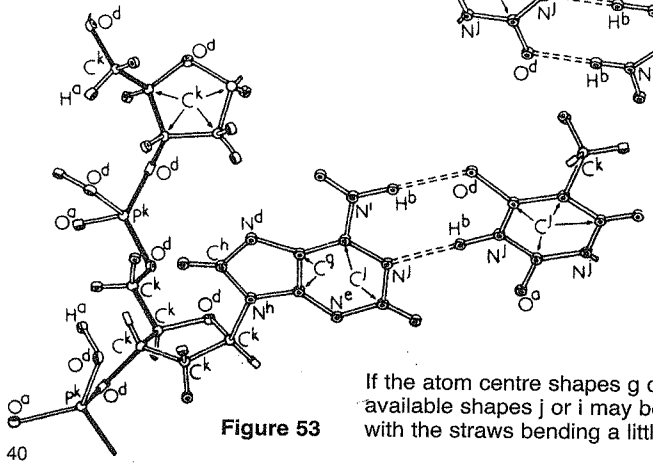


Figure 53

If the atom centre shapes g or h is not available shapes j or i may be used in its place with the straws bending a little.

Figure 54

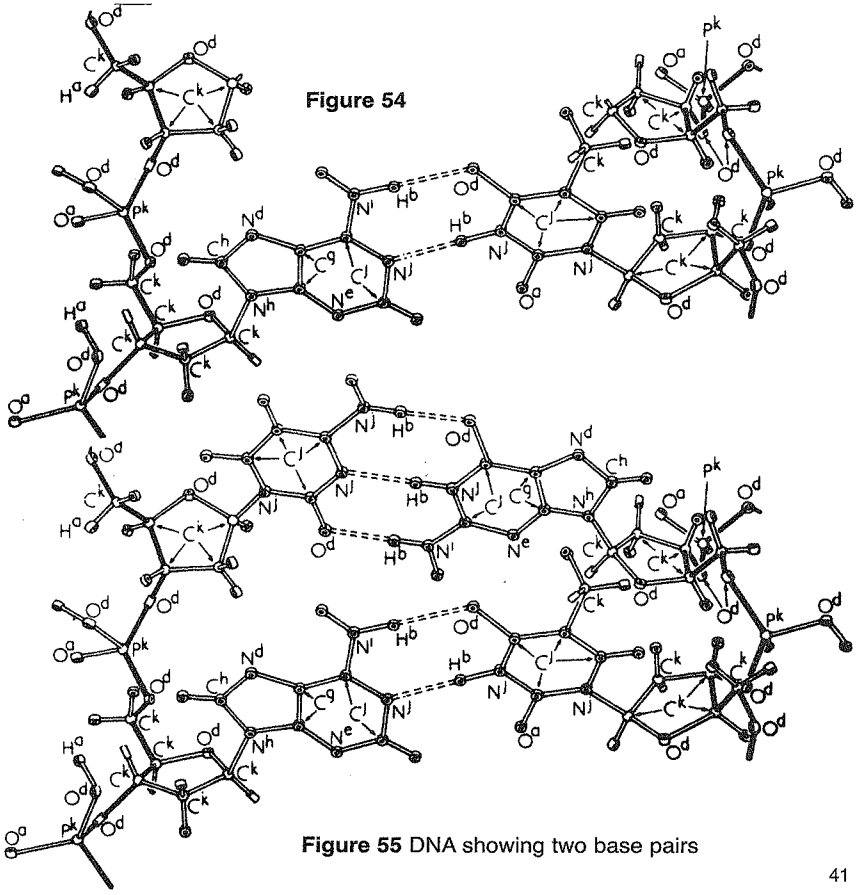


Figure 55 DNA showing two base pairs

Notes and answers to questions

Chapter 1 Sub-units

1.1 R = H

1.2 The amino group is basic (proton acceptor) and the carboxyl group is acidic (proton donor). Hence proton transfer can occur

1.3 It is asymmetric. Note that throughout the booklet it is important that you have the α-carbon atom in the L-configuration

1.4 a) the $-COO^-$ end will become -COOH; the $-NH_3^+$ end will be unchanged
b) the $-NH_3^+$ end will become $-NH_2$; the $-COO^-$ end will be unchanged

1.5 $C_3(H_2O)_3$, ie x = 3 and y = 3

1.6 Extensive hydrogen bonds with hydroxyl groups, therefore water-soluble

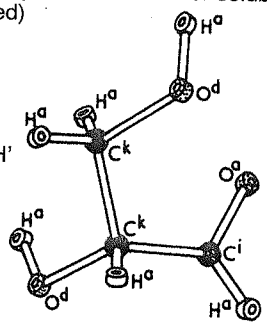
1.7 Reducing agent (copper(1) oxide precipitated)

1.8 Three

1.9 $2^3 = 8$ isomers

Page 12—note:

D-glyceraldehyde, figure 6, shown here viewed down the H—C bond of the asymmetric carbon. This view demonstrates the 'CO—R—OH' rule for recognizing D-isomers, cf the corn rule page 9. Reading anti-clockwise the groups go CO carboxyl group, R the group CH_2OH, OH the hydroxyl group



1.10 Five

1.11 $2^5 = 32$. There are eight aldohexoses, each having a D and an L form; each D and L form can cyclise in the α or β configuration

1.12 No

1.13 One would expect it to have high boiling point and be water soluble. The actual boiling point is 214°C. Extensive hydrogen bonding is the reason

1.14 Eight (ie numbers 3, 8, 9, 10, 13, 14, 17, and the first carbon atom of the side-chain)

1.15 $2^8 = 256$. Fortunately for you the exact configuration of each carbon atom is not critical from the point of view of the resultant model! Even so, only one isomer is known in nature

1.16 Hydrogen bonds

Page 21—note:

The heterocyclic bases are shown in the keto form which is considered to be the state which they preferentially take up. However, the tautomeric enol forms are frequently depicted in text books

Chapter 2 Linked sub-units

2.1 There can be no free lone pairs (they are in fact delocalised)

2.2 Electrostatic forces between ionic side-chain groups; hydrogen bonds

Page 23—note:

The delocalisation of the electrons in the peptide bond account for the partial double bond character of the C—N bond which is responsible for the *planarity* of the peptide link as well as increasing the hydrogen bonding potential of the NH and C = O groups

2.3 Because the linkage determines the configuration of the substrate molecule. It may critically affect its ability to fit the active side of the enzyme

Page 32—note:

The apparent discrepancy in the formula for nicotinamide and its structure figures 40 and 41, are due to the essential planarity of the ring nitrogen despite the tetrahedral character introduced by the protonation. If preferred tetrahedral nitrogen can be substituted without serious distortion of the ring

Chapter 3 Complex molecules

3.1 Commence with the residue having the free amino group

3.2 Assuming an amino acid to have an average molecular weight of about 125, 40 is about the minimum number of residues for a protein. In fact, an accurate calculation should take into account the fact that the quantities of each amino acid residue are not evenly distributed

3.3 Gently heating breaks hydrogen bonds, and hence disrupts secondary structure; it does not disturb primary structure. Alteration of pH may add or remove protons from groups participating in hydrogen bonding or it may neutralise the electrostatic force arising from a charged side chain group (by adding or removing a proton). In either case secondary structure will be altered, though primary structure may be unaffected. Coagulation of a protein (eg the hardening of white of egg in acids or on boiling) is an example of the destruction of secondary structure by these means.

Page 33—note:
Perhaps insufficient emphasis has been placed on the complexity of protein structure. The elements which determine the conformation of proteins are illustrated but the exceedingly elaborate structures found in proteins are barely hinted at

3.4 Cellulose is fibrous because its chains are linear. In fact bundles of cellulose molecules bond together to form strands called micelles; bundles of micelles twisted together form microfibrils; bundles of microfibrils form the fibre

3.5 Maltose

Page 36—note:
The iodine extinction reaction obtained with amylose is due to the molecule wrapping helically round the iodine molecules. This is not possible for carbohydrate polymers linked in a different way

Page 39—note:
In nucleosides and nucleotides and nucleic acids, it is important to note that the angle between the plane of the bases and the plane of the sugar ring is approximately a right angle. This is important in enabling the polymers to be correctly constructed. Also note that the acidic properties of nucleic acids are due to the ionisation of one of the phosphate hydroxyls and that the negatively charged oxygen atoms are directed away from each other in the double stranded structures. It should also be noted that both RNA and DNA can exist in either single stranded or double stranded states and that tertiary structure may supervene in the form of super helices, etc.

Further sets in this series

LATTICES

The provision of 8 and 12 coordinated atom centres allows complex models to be built without difficulty.
The booklet covers the following:

The seven crystal systems
Forms of carbon-diamond
Forms of carbon-graphite
Metals
Sodium chloride and 6:6 cubic lattices
Zinc blende and wurtzite, 4:4 structures
Rutile-Ax_2 compounds with 6:3 coordination
Ice
Layer structures

ORGANIC AND INORGANIC CHEMISTRY

The booklet covers the following:

Principles of molecular shape
Saturated hydrocarbons
Unsaturated hydrocarbons
Cyclic hydrocarbons
Functional groups in organic chemistry
Benzene and its derivatives
Polymerization
Optical isomerism
Simple organic molecules
Complex ions