

Biological Molecules

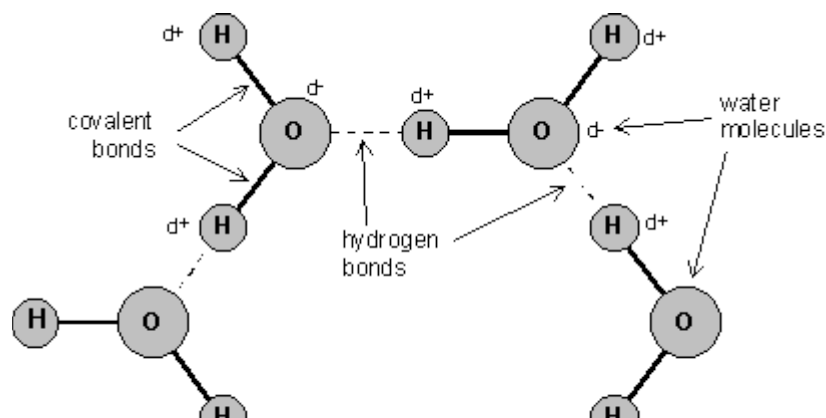
Life on Earth evolved in the water, and all life still depends on water. At least 80% of the mass of living organisms is water, and almost all the chemical reactions of life take place in aqueous solution. The other chemicals that make up living things are mostly organic macromolecules belonging to the four groups proteins, nucleic acids, carbohydrates or lipids. These macromolecules are made up from specific monomers as shown in the table below. Between them these four groups make up 93% of the dry mass of living organisms, the remaining 7% comprising small organic molecules (like vitamins) and inorganic ions.

Group name	monomers	polymers	% dry mass
Proteins	amino acids	polypeptides	50
nucleic acids	nucleotides	polynucleotides	18
carbohydrates	monosaccharides	polysaccharides	15
Group name	components	largest unit	% dry mass
lipids	fatty acids + glycerol	Triglycerides	10

The first part of this unit is about each of these groups. We'll look at each of these groups in detail, except nucleic acids, which are studied in module 2.

Water (additional information for your own interest)

Water molecules are charged, with the oxygen atom being slightly negative (δ^-) and the hydrogen atoms being slightly positive (δ^+). These opposite charges attract each other, forming hydrogen bonds. These are weak, long distance bonds that are very common and very important in biology.

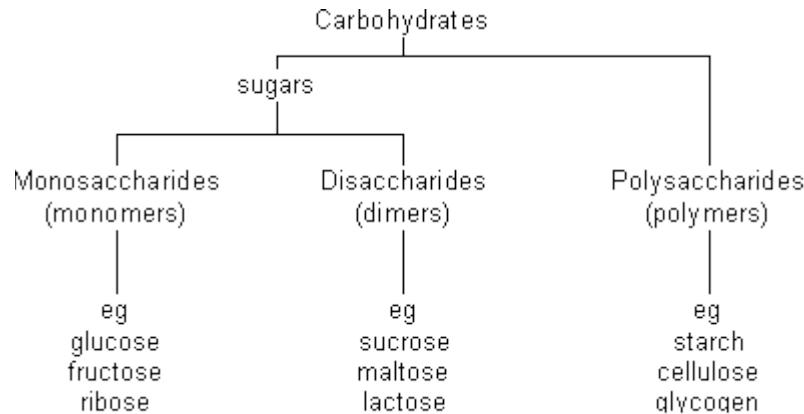


Water has a number of important properties essential for life. Many of the properties below are due to the hydrogen bonds in water:

- **Solvent.** Because it is charged, water is a very good solvent. Charged or polar molecules such as salts, sugars, amino acids dissolve readily in water and so are called hydrophilic ("water loving"). Uncharged or non-polar molecules such as lipids do not dissolve so well in water and are called hydrophobic ("water hating").
- **Specific heat capacity.** Water has a specific heat capacity of $4.2 \text{ J g}^{-1} \text{ }^{\circ}\text{C}^{-1}$, which means that it takes 4.2 joules of energy to heat 1 g of water by 1°C . This is unusually high and it means that water does not change temperature very easily. This minimises fluctuations in temperature inside cells, and it also means that sea temperature is remarkably constant.
- **Latent heat of vaporisation.** Water requires a lot of energy to change state from a liquid into a gas, and this is made use of as a cooling mechanism in animals (sweating and panting) and plants (transpiration). As water evaporates it extracts heat from around it, cooling the organism.
- **Latent heat of fusion.** Water also requires a lot of heat to change state from a solid to a liquid, and must lose a lot of heat to change state from a liquid to a solid. This means it is difficult to freeze water, so ice crystals are less likely to form inside cells.
- **Density.** Water is unique in that the solid state (ice) is less dense than the liquid state, so ice floats on water. As the air temperature cools, bodies of water freeze from the surface, forming a layer of ice with liquid water underneath. This allows aquatic ecosystems to exist even in sub-zero temperatures.
- **Cohesion.** Water molecules "stick together" due to their hydrogen bonds, so water has high cohesion. This explains why long columns of water can be sucked up tall trees by transpiration without breaking. It also explains surface tension, which allows small animals to walk on water.
- **Ionisation.** When many salts dissolve in water they ionise into discrete positive and negative ions (e.g. $\text{NaCl} \rightarrow \text{Na}^{+} + \text{Cl}^{-}$). Many important biological molecules are weak acids, which also ionise in solution (e.g. acetic acid \rightarrow acetate $^{-}$ + H^{+}). The names of the acid and ionised forms (acetic acid and acetate in this example) are often used loosely and interchangeably, which can cause confusion. You will come across many examples of two names referring to the same substance, e.g.: phosphoric acid and phosphate, lactic acid and lactate, citric acid and citrate, pyruvic acid and pyruvate, aspartic acid and aspartate, etc. The ionised form is the one found in living cells.
- **pH.** Water itself is partly ionised ($\text{H}_2\text{O} \rightleftharpoons \text{H}^{+} + \text{OH}^{-}$), so it is a source of protons (H^{+} ions), and indeed many biochemical reactions are sensitive to pH ($-\log[\text{H}^{+}]$). Pure water cannot buffer changes in H^{+} concentration, so is not a buffer and can easily be any pH, but the cytoplasm and tissue fluids of living organisms are usually well buffered at about neutral pH (pH 7-8).

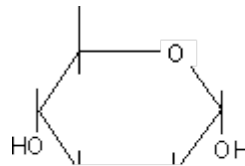
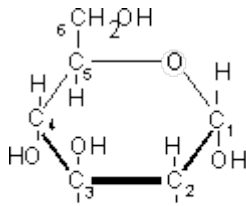
Carbohydrates

Carbohydrates contain only the elements carbon, hydrogen and oxygen. The group includes monomers, dimers and polymers, as shown in this diagram:

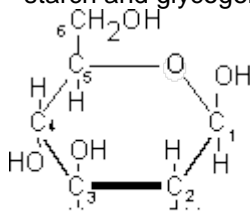


Monosaccharides (simple sugars)

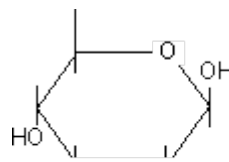
These all have the formula $(\text{CH}_2\text{O})_n$, where n can be 3-7. The most common and important monosaccharide is glucose, which is a six-carbon or hexose sugar, so has the formula $\text{C}_6\text{H}_{12}\text{O}_6$. Its structure is:



α -glucose (used to make starch and glycogen)



or more simply



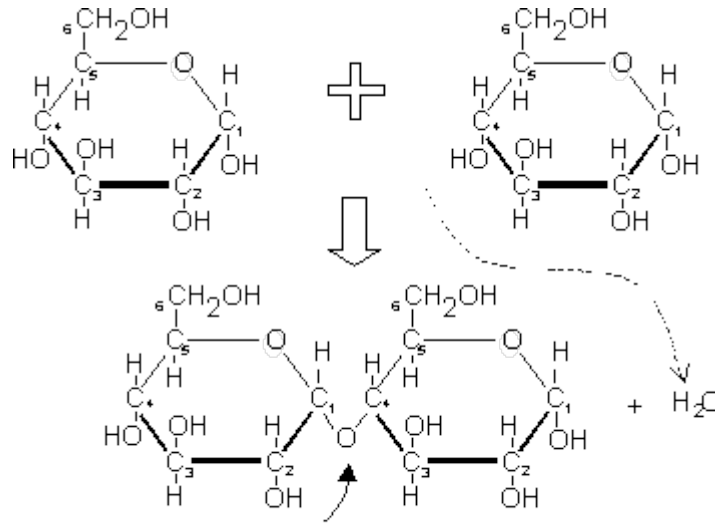
β -glucose (used to make cellulose)

Glucose forms a six-sided ring, although in three-dimensions it forms a structure that looks a bit like a chair. The six carbon atoms are numbered as shown, so we can refer to individual carbon atoms in the structure. In animals glucose is the main transport sugar in the blood, and its concentration in the blood is carefully controlled. There are many isomers of glucose, with the same chemical formula ($\text{C}_6\text{H}_{12}\text{O}_6$), but different structural formulae. These isomers include fructose and galactose.

Common five-carbon, or pentose sugars (where $n = 5$, $\text{C}_5\text{H}_{10}\text{O}_5$) include ribose and deoxyribose (found in nucleic acids and ATP) and ribulose (which occurs in photosynthesis).

Disaccharides (double sugars) [\[back to top\]](#)

Disaccharides are formed when two monosaccharides are joined together by a glycosidic bond. The reaction involves the formation of a molecule of water (H₂O):



This shows two glucose molecules joining together to form the disaccharide maltose. Because this bond is between carbon 1 of one molecule and carbon 4 of the other molecule it is called a 1-4 glycosidic bond. Bonds between other carbon atoms are possible, leading to different shapes, and branched chains.

This kind of reaction, where H₂O is formed, is called a condensation reaction. The reverse process, when bonds are broken by the addition of water (e.g. in digestion), is called a hydrolysis reaction.

- In general:
- polymerisation reactions are condensations
 - breakdown reactions are hydrolyses

There are three common disaccharides:

- Maltose (or malt sugar) is glucose 1-4 glucose. It is formed on digestion of starch by amylase, because this enzyme breaks starch down into two-glucose units. Brewing beer starts with malt, which is a maltose solution made from germinated barley. Maltose is the structure shown above.
- Sucrose (or cane sugar) is glucose 1-2 fructose. It is common in plants because it is less reactive than glucose, and it is their main transport sugar. It is the common table sugar that you put in your tea.
- Lactose (or milk sugar) is galactose 1-4 glucose. It is found only in mammalian milk, and is the main source of energy for infant mammals.

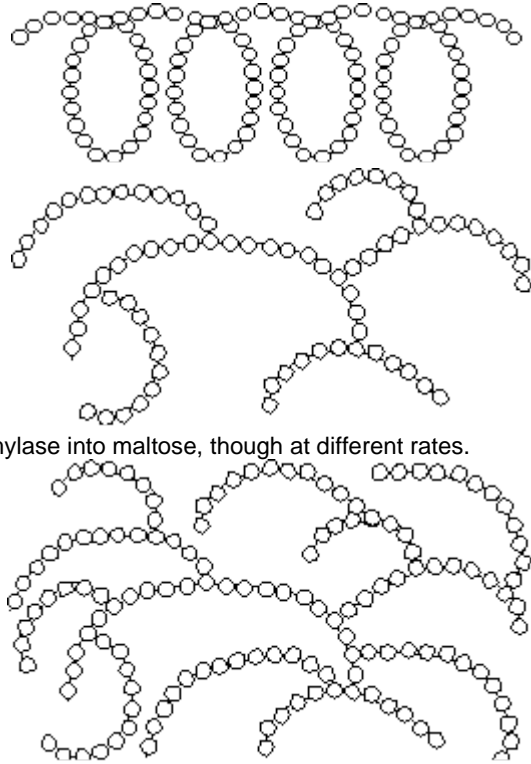
Polysaccharides

Polysaccharides are long chains of many monosaccharides joined together by glycosidic bonds. There are three important polysaccharides:

- **Starch** is the plant storage polysaccharide. It is insoluble and forms starch granules inside many plant cells. Being insoluble means starch does not change the water potential of cells, so does not cause the cells to take up water by osmosis (more on osmosis later). It is not a pure substance, but is a mixture of amylose and amylopectin.

Amylose is simply poly-(1-4) glucose, so is a straight chain. In fact the chain is floppy, and it tends to coil up into a helix.

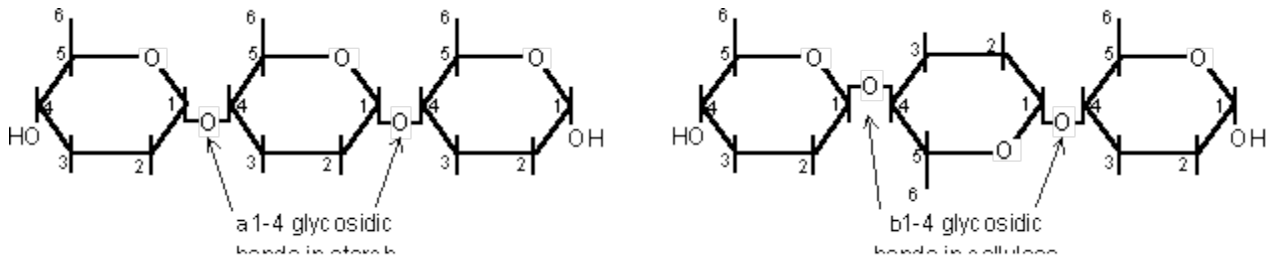
Amylopectin is poly(1-4) glucose with about 4% (1-6) branches. This gives it a more open molecular structure than amylose. Because it has more ends, it can be broken more quickly than amylose by amylase enzymes.



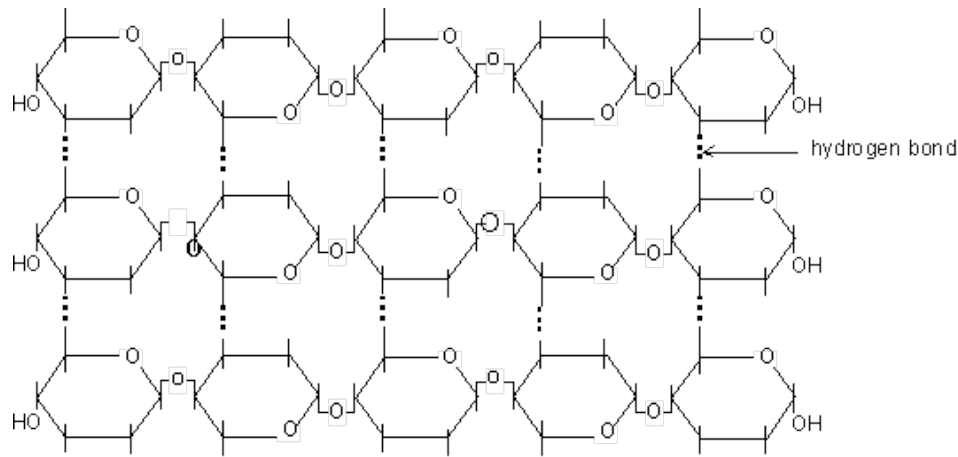
Both amylose and amylopectin are broken down by the enzyme amylase into maltose, though at different rates.

- **Glycogen** is similar in structure to amylopectin. It is poly (1-4) glucose with 9% (1-6) branches. It is made by animals as their storage polysaccharide, and is found mainly in muscle and liver. Because it is so highly branched, it can be mobilised (broken down to glucose for energy) very quickly.

- **Cellulose** is only found in plants, where it is the main component of cell walls. It is poly (1-4) glucose, but with a different isomer of glucose. Starch and glycogen contain α-glucose, in which the hydroxyl group on carbon 1 sticks down from the ring, while cellulose contains β-glucose, in which the hydroxyl group on carbon 1 sticks up. This means that in a chain alternate glucose molecules are inverted.



This apparently tiny difference makes a huge difference in structure and properties. While the α 1-4 glucose polymer in starch coils up to form granules, the β 1-4 glucose polymer in cellulose forms straight chains. Hundreds of these chains are linked together by hydrogen bonds to form cellulose microfibrils. These microfibrils are very strong and rigid, and give strength to plant cells, and therefore to young plants and also to materials such as paper, cotton and sellotape.



The β -glycosidic bond cannot be broken by amylase, but requires a specific cellulase enzyme. The only organisms that possess a cellulase enzyme are bacteria, so herbivorous animals, like cows and termites whose diet is mainly cellulose, have mutualistic bacteria in their guts so that they can digest cellulose. Humans cannot digest cellulose, and it is referred to as fibre.

- Other polysaccharides that you may come across include:
 - Chitin (poly glucose amine), found in fungal cell walls and the exoskeletons of insects.
 - Pectin (poly galactose uronate), found in plant cell walls.
 - Agar (poly galactose sulphate), found in algae and used to make agar plates.
 - Murein (a sugar-peptide polymer), found in bacterial cell walls.
 - Lignin (a complex polymer), found in the walls of xylem cells, is the main component of wood.

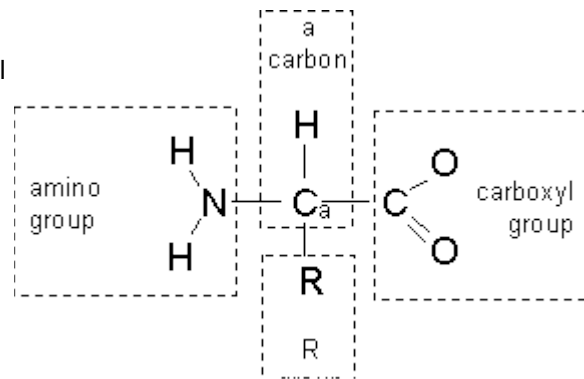
Proteins

Proteins are the most complex and most diverse group of biological compounds. They have an astonishing range of different functions, as this list shows.

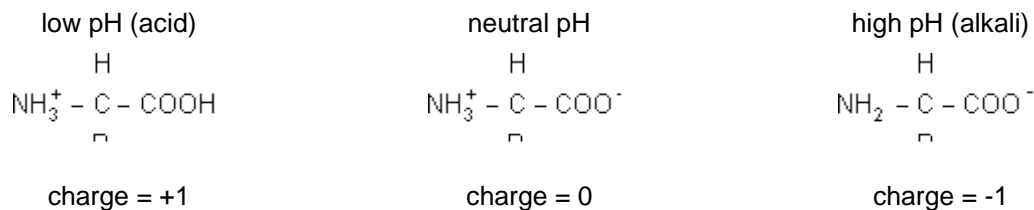
structure	e.g. collagen (bone, cartilage, tendon), keratin (hair), actin (muscle)
enzymes	e.g. amylase, pepsin, catalase, etc (>10,000 others)
transport	e.g. haemoglobin (oxygen), transferrin (iron)
pumps	e.g. Na^+K^+ pump in cell membranes
motors	e.g. myosin (muscle), kinesin (cilia)
hormones	e.g. insulin, glucagon
receptors	e.g. rhodopsin (light receptor in retina)
antibodies	e.g. immunoglobulins
storage	e.g. albumins in eggs and blood, caesin in milk
blood clotting	e.g. thrombin, fibrin
lubrication	e.g. glycoproteins in synovial fluid
toxins	e.g. diphtheria toxin
antifreeze	e.g. glycoproteins in arctic flea
and many more!	

Proteins are made of amino acids. Amino acids are made of the five elements C H O N S. The general structure of an amino acid molecule is shown on the right. There is a central carbon atom (called the "alpha carbon"), with four different chemical groups attached to it:

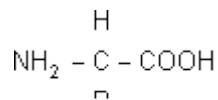
- a hydrogen atom
- a basic amino group
- an acidic carboxyl group
- a variable "R" group (or side chain)



Amino acids are so-called because they have both amino groups and acid groups, which have opposite charges. At neutral pH (found in most living organisms), the groups are ionised as shown above, so there is a positive charge at one end of the molecule and a negative charge at the other end. The overall net charge on the molecule is therefore zero. A molecule like this, with both positive and negative charges is called a zwitterion. The charge on the amino acid changes with pH:

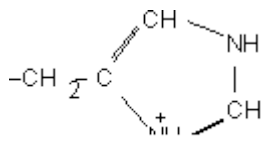
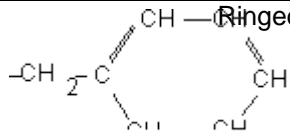
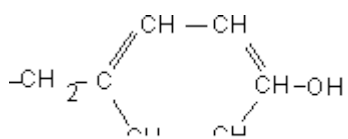
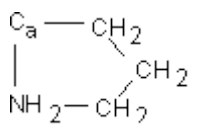
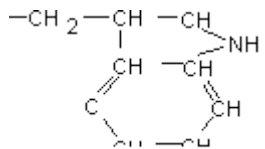


It is these changes in charge with pH that explain the effect of pH on enzymes. A solid, crystallised amino acid has the uncharged structure (below), but this form never exists in solution, and therefore doesn't exist in living things (although it is the form usually given in textbooks).



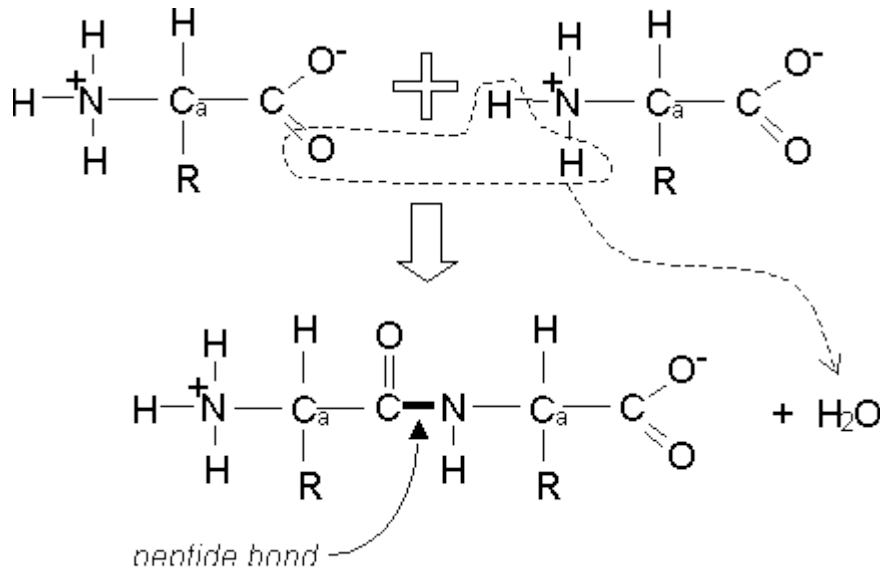
There are 20 different R groups, and so 20 different amino acids. Since each R group is slightly different, each amino acid has different properties, and this in turn means that proteins can have a wide range of properties. The following table shows the 20 different R groups, grouped by property, which gives an idea of the range of properties. You do not need to learn these, but it is interesting to see the different structures, and you should be familiar with the amino acid names. You may already have heard of some, such as the food additive monosodium glutamate, which is simply the sodium salt of the amino acid glutamate. Be careful not to confuse the names of amino acids with those of bases in DNA, such as cysteine (amino acid) and cytosine (base), threonine (amino acid) and thymine (base). There are 3-letter and 1-letter abbreviations for each amino acid.

The Twenty Amino Acid R-Groups

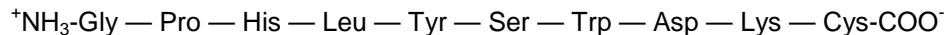
Simple R groups		Basic R groups	
<p>Glycine Gly G</p> <p>Alanine Ala A</p> <p>Valine Val V</p> <p>Leucine Leu L</p> <p>Isoleucine Ile I</p>	H $ $ $-\text{CH}$ $ $ H	<p>Lysine</p> <p>Arginine</p> <p>Histidine</p> <p>Asparagine</p> <p>Glutamine</p>	$-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{NH}_2^+$ $-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{NH}-\text{C}(=\text{NH}_2)^+$  $-\text{CH}_2-\text{C}(=\text{O})\text{NH}_2^+$ $-\text{CH}_2-\text{CH}_2-\text{C}(=\text{O})\text{NH}_2^+$
<p>Serine Ser S</p> <p>Threonine Thr T</p>	H H $ $ $ $ $-\text{CH}$ $-\text{OH}$ $ $ H	<p>Aspartate</p> <p>Glutamate</p>	<p style="text-align: center;">Acidic R groups</p> $-\text{CH}_2-\text{C}(=\text{O})\text{O}^-$ $-\text{CH}_2-\text{CH}_2-\text{C}(=\text{O})\text{O}^-$
<p>Cysteine Cys C</p> <p>Methionine Met M</p>	H H $ $ $ $ $-\text{CH}$ $-\text{OH}$ $ $ H	<p>Phenylalanine</p> <p>Tyrosine</p>	<p style="text-align: center;">Ringed R groups</p>  
<p>Proline Pro P</p>	<p style="text-align: center;">Cyclic R group</p> 	<p>Tryptophan</p>	

Polypeptides [\[back to top\]](#)

Amino acids are joined together by peptide bonds. The reaction involves the formation of a molecule of water in another condensation polymerisation reaction:



When two amino acids join together a dipeptide is formed. Three amino acids form a tripeptide. Many amino acids form a polypeptide. e.g.:



In a polypeptide there is always one end with a free amino (NH₃) group, called the N-terminus, and one end with a free carboxyl (CO₂) group, called the C-terminus.

In a protein the polypeptide chain may be hundreds of amino acids long. Amino acid polymerisation to form polypeptides is part of protein synthesis. It takes place in ribosomes, and is special because it requires an RNA template. The sequence of amino acids in a polypeptide chain is determined by the sequence of the genetic code in DNA. Protein synthesis it studied in detail in module 2.

Protein Structure

Polypeptides are just a string of amino acids, but they fold up to form the complex and well-defined three-dimensional structure of working proteins. To help to understand protein structure, it is broken down into four levels:

1. Primary Structure

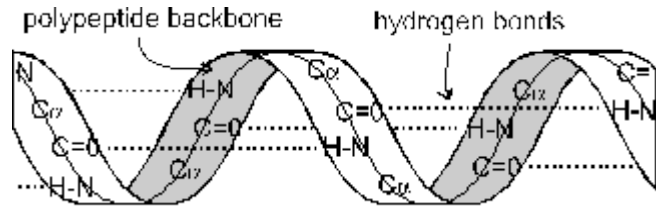
This is just the sequence of amino acids in the polypeptide chain, so is not really a structure at all. However, the primary structure does determine the rest of the protein structure. Finding the primary structure of a protein is called protein sequencing, and the first protein to be sequenced was the protein hormone insulin, by the Cambridge biochemist Fredrick Sanger, for which work he got the Nobel prize in 1958.

2. Secondary Structure

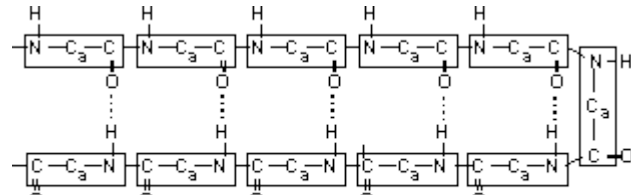


This is the most basic level of protein folding, and consists of a few basic motifs that are found in all proteins. The secondary structure is held together by hydrogen bonds between the carboxyl groups and the amino groups in the polypeptide backbone. The two most common secondary structure motifs are the α -helix and the β -sheet.

The α -helix. The polypeptide chain is wound round to form a helix. It is held together by hydrogen bonds running parallel with the long helical axis. There are so many hydrogen bonds that this is a very stable and strong structure. Do not confuse the α -helix of proteins with the famous double helix of DNA. Helices are common structures throughout biology.



The β -sheet. The polypeptide chain zig-zags back and forward forming a sheet of antiparallel strands. Once again it is held together by hydrogen bonds.



The α -helix and the β -sheet were discovered by Linus Pauling, for which work he got the Nobel prize in 1954. There are a number of other secondary structure motifs such as the β -bend, the triple helix (only found in collagen), and the random coil.

3. Tertiary Structure

This is the compact globular structure formed by the folding up of a whole polypeptide chain. Every protein has a unique tertiary structure, which is responsible for its properties and function. For example the shape of the active site in an enzyme is due to its tertiary structure. The tertiary structure is held together by bonds between the R groups of the amino acids in the protein, and so depends on what the sequence of amino acids is. There are three kinds of bonds involved:

- hydrogen bonds, which are weak.
- ionic bonds between R-groups with positive or negative charges, which are quite strong.
- sulphur bridges - covalent S-S bonds between two cysteine amino acids, which are strong.

So the secondary structure is due to backbone interactions and is thus largely independent of primary sequence, while tertiary structure is due to side chain interactions and thus depends on the amino acid sequence.

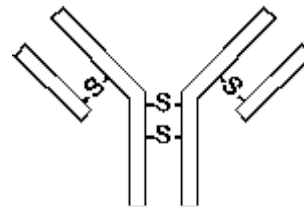
4. Quaternary Structure

This structure is found in proteins containing more than one polypeptide chain, and simply means how the different polypeptide chains are arranged together. The individual polypeptide chains are usually globular, but can arrange themselves into a variety of quaternary shapes. e.g.:

Haemoglobin, the oxygen-carrying protein in red blood cells, consists of four globular subunits arranged in a tetrahedral (pyramid) structure. Each subunit contains one iron atom and can bind one molecule of oxygen.



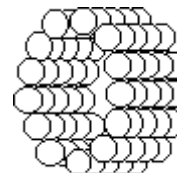
Immunoglobulins, the proteins that make antibodies, comprise four polypeptide chains arranged in a Y-shape. The chains are held together by sulphur bridges. This shape allows antibodies to link antigens together, causing them to clump.



Actin, one of the proteins found in muscles, consists of many globular subunits arranged in a double helix to form long filaments.



Tubulin is a globular protein that polymerises to form hollow tubes called microtubules. These form part of the cytoskeleton, and make cilia and flagella move.



These four structures are not real stages in the formation of a protein, but are simply a convenient classification that scientists invented to help them to understand proteins. In fact proteins fold into all these structures at the same time, as they are synthesised.

The final three-dimensional shape of a protein can be classified as globular or fibrous.

globular structure



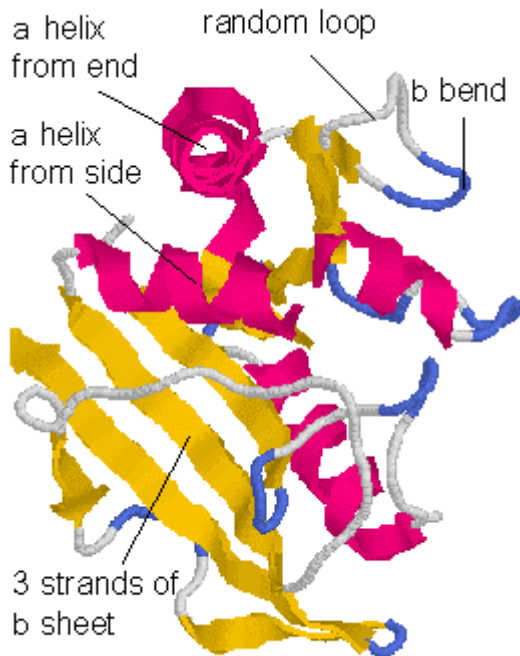
fibrous (or filamentous) structure



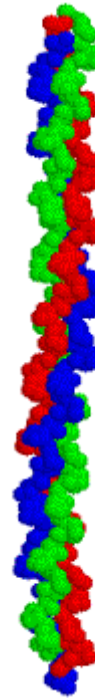
The vast majority of proteins are globular, including enzymes, membrane proteins, receptors, storage proteins, etc. Fibrous proteins look like ropes and tend to have structural roles such as collagen (bone), keratin (hair), tubulin

(cytoskeleton) and actin (muscle). They are usually composed of many polypeptide chains. A few proteins have both structures: the muscle protein myosin has a long fibrous tail and a globular head, which acts as an enzyme.

This diagram shows a molecule of the enzyme dihydrofolate reductase, which comprises a single polypeptide chain. It has been drawn to highlight the different secondary structures.

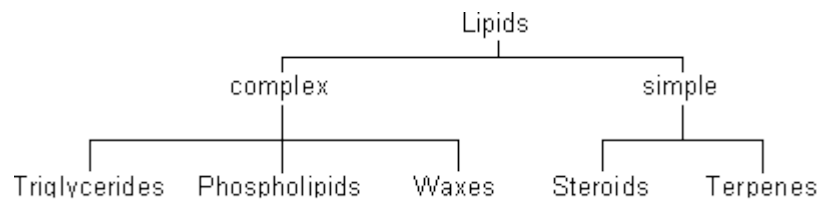


This diagram shows part of a molecule of collagen, which is found in bone and cartilage. It has a unique, very strong triple-helix structure.



Lipids

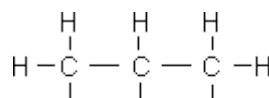
Lipids are a mixed group of hydrophobic compounds composed of the elements carbon, hydrogen and oxygen.

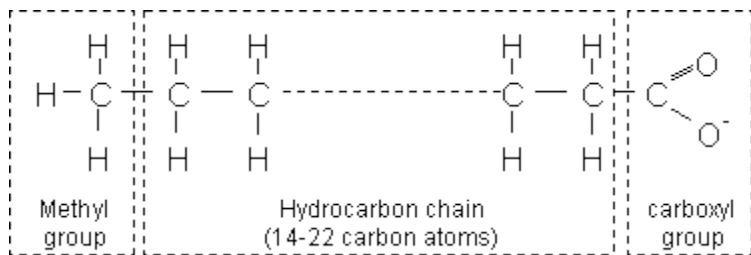


Triglycerides

Triglycerides are commonly called fats or oils. They are made of glycerol and fatty acids.

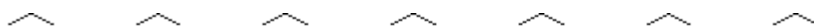
Glycerol is a small, 3-carbon molecule with three alcohol groups.





Fatty acids are long molecules with a polar, hydrophilic end and a non-polar, hydrophobic "tail". The hydrocarbon chain can be from 14 to 22 CH₂ units long, but it is always an even number because of the way fatty acids are made. The hydrocarbon chain is sometimes called an R group, so the formula of a fatty acid can be written as R-COO⁻.

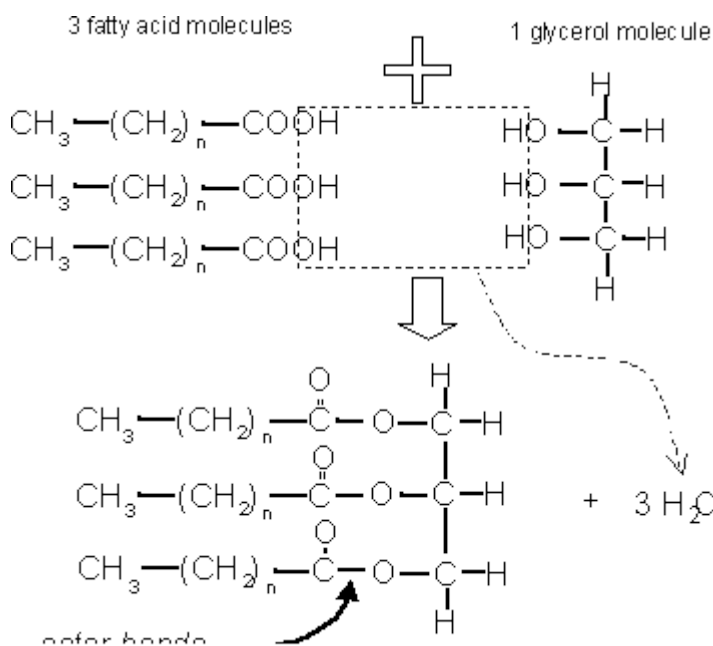
- If there are no C=C double bonds in the hydrocarbon chain, then it is a saturated fatty acid (i.e. saturated with hydrogen). These fatty acids form straight chains, and have a high melting point.



- If there are C=C double bonds in the hydrocarbon chain, then it is an unsaturated fatty acid (i.e. unsaturated with hydrogen). These fatty acids form bent chains, and have a low melting point. Fatty acids with more than one double bond are called poly-unsaturated fatty acids (PUFAs).



One molecule of glycerol joins together with three fatty acid molecules to form a triglyceride molecule, in another condensation polymerisation reaction:

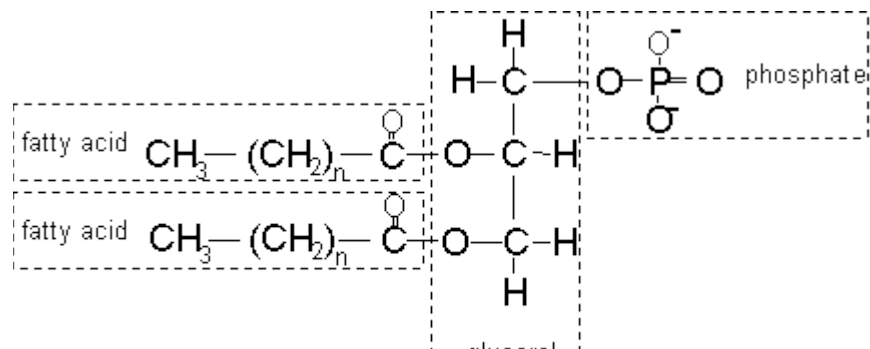


Triglycerides are insoluble in water. They are used for storage, insulation and protection in fatty tissue (or adipose tissue) found under the skin (sub-cutaneous) or surrounding organs. They yield more energy per unit mass than other compounds so are good for energy storage. Carbohydrates can be mobilised more quickly, and glycogen is stored in muscles and liver for immediate energy requirements.

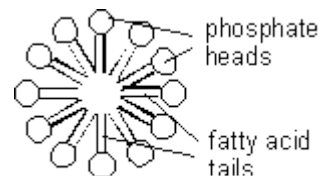
- Triglycerides containing saturated fatty acids have a high melting point and tend to be found in warm-blooded animals. At room temperature they are solids (fats), e.g. butter, lard.
- Triglycerides containing unsaturated fatty acids have a low melting point and tend to be found in cold-blooded animals and plants. At room temperature they are liquids (oils), e.g. fish oil, vegetable oils.

Phospholipids

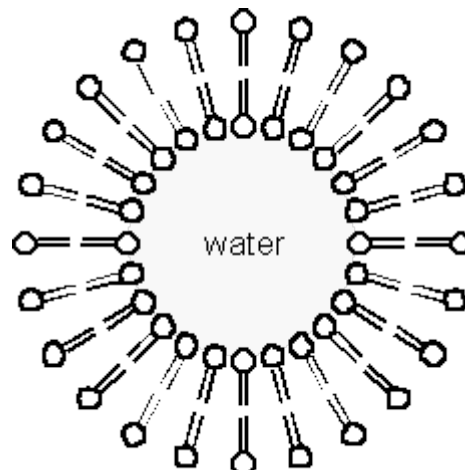
Phospholipids have a similar structure to triglycerides, but with a phosphate group in place of one fatty acid chain. There may also be other groups attached to the phosphate. Phospholipids have a polar hydrophilic "head" (the negatively-charged phosphate group) and two non-polar hydrophobic "tails" (the fatty acid chains). This mixture of properties is fundamental to biology, for phospholipids are the main components of cell membranes.



When mixed with water, phospholipids form droplet spheres with the hydrophilic heads facing the water and the hydrophobic tails facing each other. This is called a micelle.



Alternatively, they may form a double-layered phospholipid bilayer. This traps a compartment of water in the middle separated from the external water by the hydrophobic sphere. This naturally-occurring structure is called a liposome, and is similar to a membrane surrounding a cell.



Waxes

Waxes are formed from fatty acids and long-chain alcohols. They are commonly found wherever waterproofing is needed, such as in leaf cuticles, insect exoskeletons, birds' feathers and mammals' fur.

Steroids

Steroids are small hydrophobic molecules found mainly in animals. They include:

- cholesterol, which is found in animals cell membranes to increase stiffness
- bile salts, which help to emulsify dietary fats
- steroid hormones such as testosterone, oestrogen, progesterone and cortisol
- vitamin D, which aids Ca^{2+} uptake by bones.

Terpenes

Terpenes are small hydrophobic molecules found mainly in plants. They include vitamin A, carotene and plant oils such as geraniol, camphor and menthol.

Chromatography

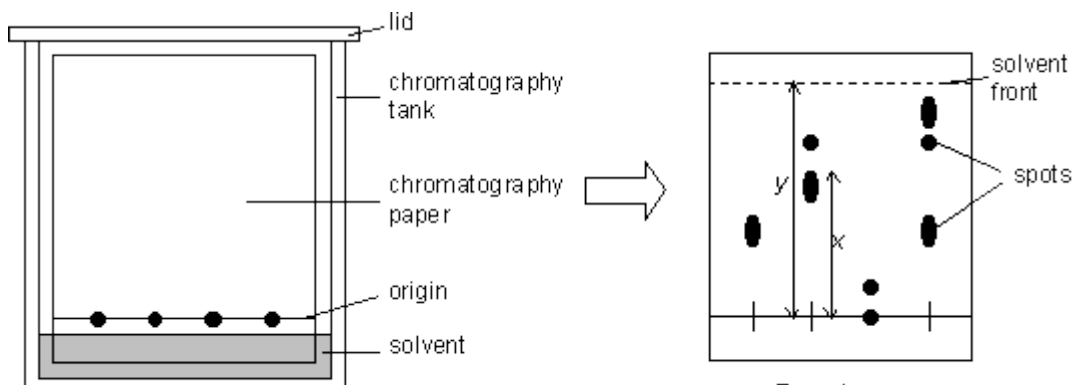
Chromatography is used to separate pure substances from a mixture of substances, such as a cell extract. It is based on different substances having different solubilities in different solvents. A simple and common form of chromatography uses filter paper.

1. Pour some solvent into a chromatography tank and seal it, so the atmosphere is saturated with solvent vapour. Different solvents are suitable for different tasks, but they are usually mixtures of water with organic liquids such as ethanol or propanone.
2. Place a drop of the mixture to be separated onto a sheet of chromatography paper near one end. This is the origin of the chromatogram. The spot should be small but concentrated. Repeat for any other mixtures. Label the spots with pencil, as ink may dissolve.
3. Place the chromatography sheet into the tank so that the origin is just above the level of solvent, and leave for several hours. The solvent will rise up the paper by capillary action carrying the contents of the mixture with it. Any solutes dissolved in the solvent will be partitioned between the organic solvent (the moving phase) and the water, which is held by the paper (the stationary phase). The more soluble a solute is in the solvent the further up the paper it will move.
4. When the solvent has nearly reached the top of the paper, the paper is removed and the position of the solvent front marked. The chromatogram may need to be developed to make the spots visible. For example amino acids stain purple with ninhydrin.

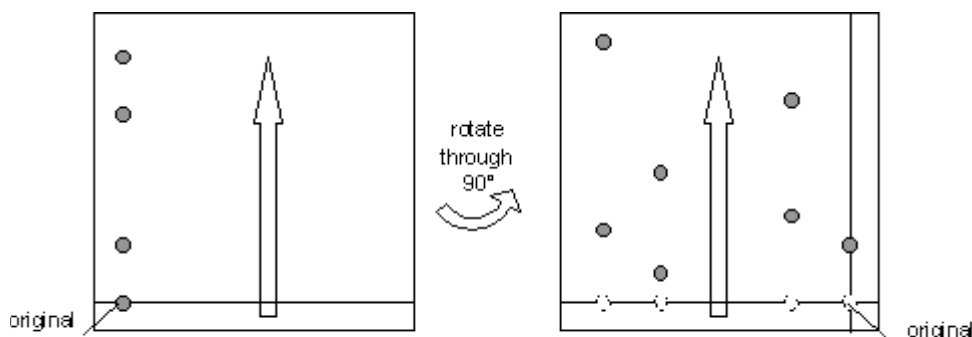
5. The chromatogram can be analysed by measuring the distance travelled by the solvent front, and the distance from the origin to the centre of each spot. This is used to calculate the R_f (relative front) value for each spot:

$$R_f = \frac{\text{distance moved by spot}}{\text{distance moved by solvent}}$$

An R_f value is characteristic of a particular solute in a particular solvent. It can be used to identify components of a mixture by comparing to tables of known R_f values.



Sometimes chromatography with a single solvent is not enough to separate all the constituents of a mixture. In this case the separation can be improved by two-dimensional chromatography, where the chromatography paper is turned through 90° and run a second time in a second solvent. Solutes that didn't separate in one solvent will separate in another because they have different solubilities.



There are many different types of chromatography.

- Paper chromatography is the simplest, but does not always give very clean separation.
- Thin layer chromatography (tlc) uses a thin layer of cellulose or silica coated onto a plastic or glass sheet. This is more expensive, but gives much better and more reliable separation.
- Column chromatography uses a glass column filled with a cellulose slurry. Large samples can be pumped through the column and the separated fractions can be collected for further experiments, so this is preparative chromatography as opposed to analytical chromatography.

- High performance liquid chromatography (HPLC) is an improved form of column chromatography that delivers excellent separation very quickly.
- Electrophoresis uses an electric current to separate molecules on the basis of charge. It can also be used to separate on the basis of molecular size, and as such is used in DNA sequencing.

Biochemical Tests

These five tests identify the main biologically important chemical compounds. For each test take a small amount of the substance to test, and shake it in water in a test tube. If the sample is a piece of food, then grind it with some water in a pestle and mortar to break up the cells and release the cell contents. Many of these compounds are insoluble, but the tests work just as well on a fine suspension.

- **Starch** (iodine test). To approximately 2 cm³ of test solution add two drops of iodine/potassium iodide solution. A blue-black colour indicates the presence of starch as a starch-polyiodide complex is formed. Starch is only slightly soluble in water, but the test works well in a suspension or as a solid.
- **Reducing Sugars** (Benedict's test). All monosaccharides and most disaccharides (except sucrose) will reduce copper (II) sulphate, producing a precipitate of copper (I) oxide on heating, so they are called reducing sugars. Benedict's reagent is an aqueous solution of copper (II) sulphate, sodium carbonate and sodium citrate. To approximately 2 cm³ of test solution add an equal quantity of Benedict's reagent. Shake, and heat for a few minutes at 95°C in a water bath. A precipitate indicates reducing sugar. The colour and density of the precipitate gives an indication of the amount of reducing sugar present, so this test is semi-quantitative. The original pale blue colour means no reducing sugar, a green precipitate means relatively little sugar; a brown or red precipitate means progressively more sugar is present.
- **Non-reducing Sugars** (Benedict's test). Sucrose is called a non-reducing sugar because it does not reduce copper sulphate, so there is no direct test for sucrose. However, if it is first hydrolysed (broken down) to its constituent monosaccharides (glucose and fructose), it will then give a positive Benedict's test. So sucrose is the only sugar that will give a negative Benedict's test before hydrolysis and a positive test afterwards. First test a sample for reducing sugars, to see if there are any present before hydrolysis. Then, using a separate sample, boil the test solution with dilute hydrochloric acid for a few minutes to hydrolyse the glycosidic bond. Neutralise the solution by gently adding small amounts of solid sodium hydrogen carbonate until it stops fizzing, then test as before for reducing sugars.
- **Lipids** (emulsion test). Lipids do not dissolve in water, but do dissolve in ethanol. This characteristic is used in the emulsion test. Do not start by dissolving the sample in water, but instead shake some of the test sample with about 4 cm³ of ethanol. Decant the liquid into a test tube of water, leaving any undissolved substances behind. If there are lipids dissolved in the ethanol, they will precipitate in the water, forming a cloudy white emulsion. The test can be improved by adding the dye Sudan III, which stains lipids red.
- **Protein** (biuret test). To about 2 cm³ of test solution add an equal volume of biuret solution, down the side of the test tube. A blue ring forms at the surface of the solution, which disappears on shaking, and the solution turns lilac-purple, indicating protein. The colour is due to a complex between nitrogen atoms in the peptide chain and Cu²⁺ ions, so this is really a test for peptide bonds.