15 Proteins

The number and variety of carbon compounds far exceeds the number and variety of compounds based on any other element. Three extremely important groups of organic compounds are **proteins, carbohydrates and synthetic polymers**.

These all display properties, including strength, density and reactivity, that can be explained by considering their primary, secondary and tertiary structures. In the next few chapters we will be looking at the properties and structures of some examples of these compounds.

From your junior science you are aware that proteins are essential in your diet and that they are the building blocks of life. Every cell contains proteins. They are needed for growth and repair of organisms because they are present everywhere in the body structure – in skin, hair, nails, muscles, tendons, cartilage. Haemoglobin which carries oxygen around the body is a protein; muscle fibres which allow us to move are proteins. Proteins make up the enzymes of the body – so they catalyse chemical reactions in cells and organs. Proteins also help in the fight against disease as they are present as antibodies. But proteins are not all helpful to us – they are also present in toxins and venoms such as snake venom.

The **structures of proteins** are as varied as their functions – they can form soft, flexible structures; rigid structures; soluble crystals or insoluble fibres. Their functions include strengthening and protecting body structures, transport, contraction, defence and regulation of metabolic activities.

However, one thing all proteins have in common, is that they all consist of a series of small molecules called **amino acids**. Each protein consists of a chain of amino acids.

Amino acids

Amino acids are made of carbon, hydrogen, oxygen, nitrogen, and some also contain sulfur. Amino acids are carboxylic acids with an attached amine ($-NH_2$) group, so each amino acid has, along its length, at least one amino (NH_2) group and one carboxyl (COOH) group.

These amino acids are soluble in water because the NH_2 or the COOH group will ionise forming NH_3^+ and COO^- and also hydrogen bonds will form with the water molecules.

The biologically important amino acids, are called 2-amino acids. This is because in a carboxylic acid, we number the carbons starting from the C in COOH. So the next C – with the COOH and the NH_2 attached to it is the second C.

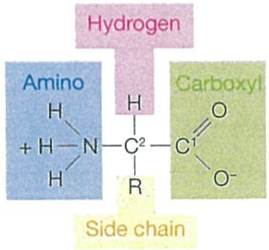


Figure 15.1 Structure of a 2-amino acid.

In some texts you will see 2-amino acids called **α-amino acids**. This is because the carbon with a carboxyl group attached can also be called the alpha carbon atom (α-carbon). The amino group and carboxyl group are both bonded to this α-carbon, so the amino acids are called α-amino acids.

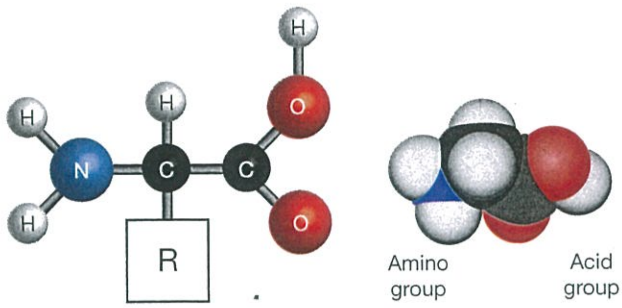


Figure 15.2 Two models of an α-amino acid when it is not ionised.

Amino acids are amphoteric – they react as both an acid and a base. They have both acidic and basic properties because the carboxyl group is acidic and the amine group is basic. The equations show amino acids acting as an acid and as a base.

Reacting as an acid:

$$R-CH(COO^-)-NH_3^+ + OH^- \rightleftharpoons R-CH(COO^-)-NH_2 + H_2O$$

Reacting as a base:

$$R-CH(COO^-)-NH_3^+ + H^+ \rightleftharpoons R-CH(COOH)-NH_3^+$$

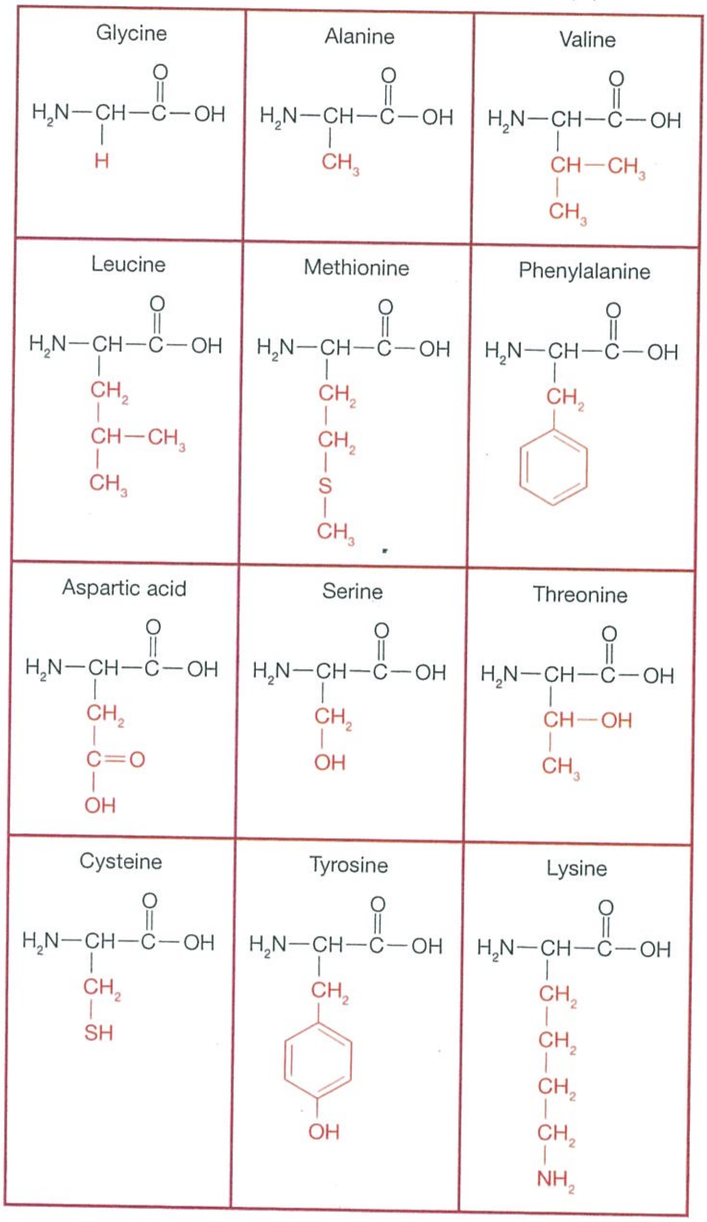
Amino acids are also described as **amphiprotic** – they can accept and also donate protons. Depending on pH, the $-NH_2$ group and the $-COOH$ group of amino acids carry charge as they accept or donate protons.

Low pH (acidic)	Neutral pH	High pH (alkaline)
$^+H_3N-CH(R)-COOH$	$^+H_3N-CH(R)-COO^-$	$H_2N-CH(R)-COO^-$

In very acidic solutions the amino group attracts hydrogen ions and the amino acid becomes positively charged. In very alkaline solutions, the carboxyl group ionises and the amino acid develops an overall negative charge.

Twenty different amino acids have been isolated from the proteins of organisms and some of these are shown in Table 15.1. You will notice that the simplest amino acid is glycine.

Table 15.1 Some of the 20 amino acids that make up proteins.



Peptide bonds

Each protein is made of long chains of amino acids, in a definite sequence that is determined by the DNA code. The amino acids are held together by special bonds called peptide bonds which occur between the α-amino group and the α-carboxyl group of the amino acids. A peptide bond forms whenever two amino acids join together.

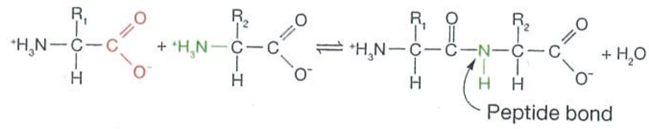


Figure 15.3 Forming a peptide bond.

The 20 different amino acids can bond to make millions of different combinations. Each protein has its own special sequence of amino acid molecules, if even one amino acid is in the wrong order, the protein will not work properly.

Two amino acids joined by a peptide bond is called a dipeptide, three is a tripeptide, but proteins have many peptide bonds and so they are polypeptides.

Polypeptides

A polypeptide is a long, linear chain of many amino acids, joined by peptide bonds. The 20 amino acids can combine in many different ways to make millions of different polypeptides. This is just like the 26 letters of the alphabet combining in different ways to make millions of different words.

Proteins consist mostly of one polypeptide chain, although some proteins have more than one. The number and sequence of the amino acids determines the properties of the protein and this affects its functions.

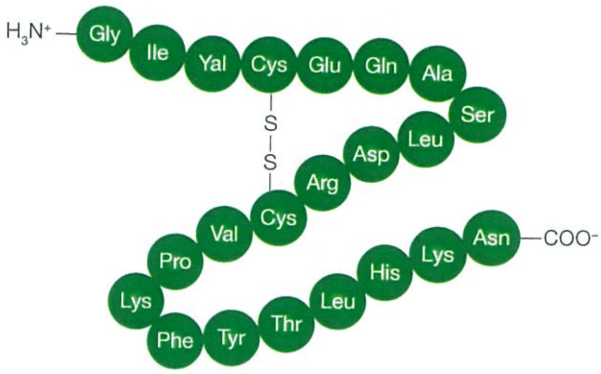
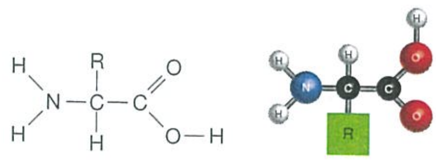


Figure 15.4 A polypeptide model.

QUESTIONS

1. (a) Name the elements that make up amino acids.
   (b) Label the amino acid formula and model below to show the amino group, the carboxyl group and the 2-carbon atom.



   (c) Explain why the amino acids that make up proteins in living organisms are called 2-amino acids.

16 Protein Structure

Each protein is made of one or more polypeptide chains which become twisted into a spiral shape (a helix) or else pleated. Different parts of the chain then become linked together. This forms the final protein structure which determines its properties.

Primary structure

The primary structure of proteins is the number and sequence of the amino acids in the polypeptide chain, the identity of each amino acid, its order and also its orientation.

An average polypeptide is about 300 amino acids in length, but some have thousands of amino acids.

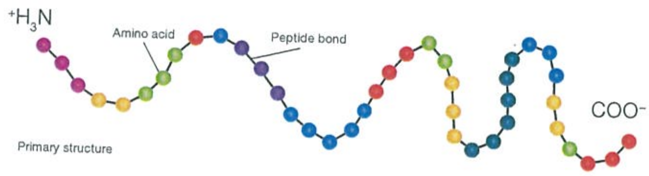


Figure 16.1 Primary structure of proteins.

Interactions between side chains and the peptide bonds cause the polypeptide chains to be arranged into a complex 3-dimensional structure which determines the 3-dimensional shape and the functions and properties of each protein.

Secondary structure

The secondary structure of proteins refers to the arrangement of the polypeptide chains which may be into the shape of a helix or a pleated sheet. Hydrogen bonds form and help to stabilise these structures.

Two types of secondary structures are a pleated sheet and an alpha helix.

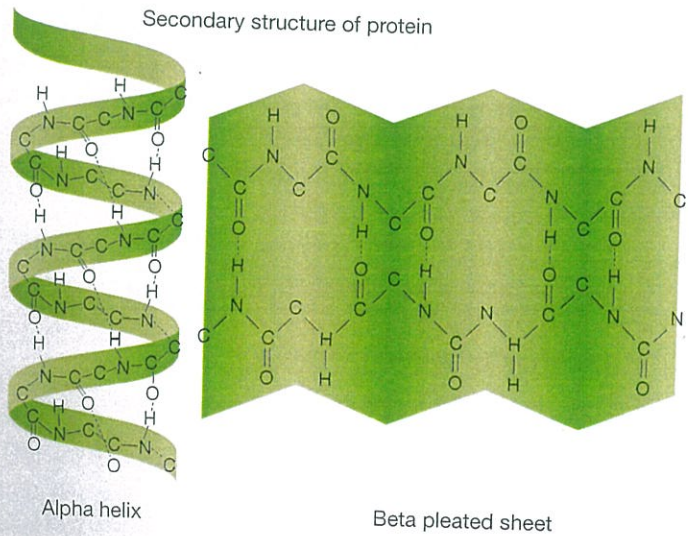


Figure 16.2 Secondary structure of proteins.

The pleated sheet arrangement produces a protein that is flexible but not elastic, e.g. keratin in silk. The helix shape can lead to greater elasticity of a protein as the coils stretch and recoil, e.g. elastin in elastic connective tissue of the body.

Tertiary structure

The tertiary structure of a protein is the further folding and curling of the whole polypeptide, including the secondary structures. For example, **myoglobin** is the pigment in muscles that contains iron and carries oxygen. Its polypeptide is organised like a helix with bends in it, folding back on itself in folds that are not symmetrical.

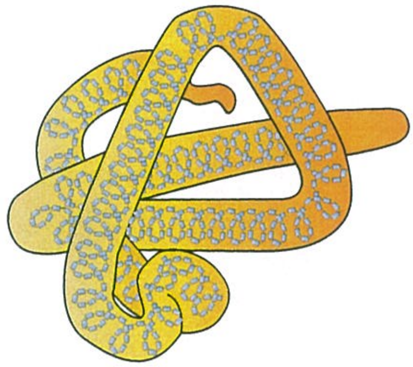


Figure 16.3 Tertiary structure – myoglobin.

Tertiary structure can be broken down by slight changes in temperature or pH. This alters their properties, for example when you cook egg white. Protein structure is important for correct function. Misfolded proteins are thought to be involved in Parkinson's and Huntington's diseases.

Quaternary structure

The quaternary structure of proteins is determined by the way proteins with two or more polypeptide chains fit their chains together. Haemoglobin, many enzymes, hormones such as insulin and antibodies have quaternary structure.

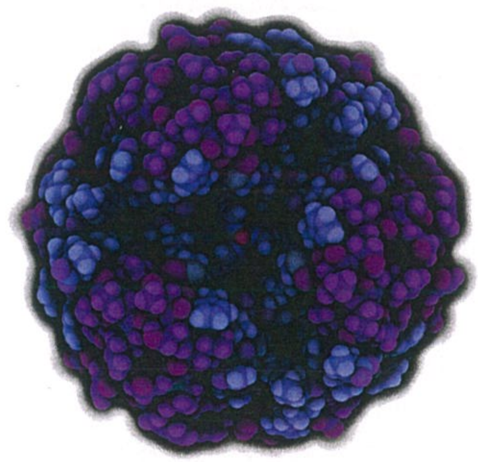


Figure 16.4 Quaternary protein structure.

Proteins are classified as fibrous or globular depending on their structure. Fibrous proteins are tough, rigid, insoluble in water and form structures such as nails and hair. Globular proteins are soluble in water and roughly spherical.

Bonding in proteins

Covalent peptide bonds hold together the amino acids in the primary structure.

Non-covalent bonds primarily hold the secondary, tertiary and quaternary structures together.

Hydrogen bonds form between an oxygen atom in a carbonyl (–C=O) group and a hydrogen atom in an amide group of the polypeptide chain. These hydrogen bonds hold together the chains in the secondary helix or pleated sheet structure.

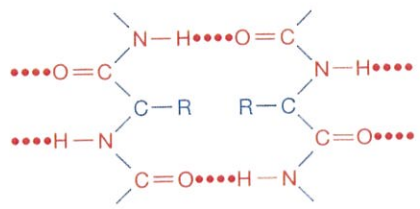


Figure 16.5 Hydrogen bonds in protein.

Side chain interactions

Side chain interactions are responsible for the folding of the protein into its tertiary and quaternary structures.

In Figure 16.6 you can see the backbone of the chain with R groups of the amino acids within the chain projecting from it. These R groups are the side chains.

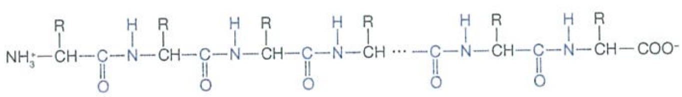


Figure 16.6 Side chains.

Depending on the nature of the side chains (R groups), dispersion forces, ionic bonds and/or hydrogen bonds can form between them.

Some amino acids have non-polar side chains which tend to lie towards the centre of the folded protein structure as they are hydrophobic – they do not readily bond with water.

Disulfide bridges (links) are cross-links between chains that are formed when a polypeptide chain folds and coils, and two cysteine side groups end up near each other. These can react and form an extra covalent link called a disulfide bridge which provides extra stability. Figure 16.7 below shows disulfide bridges in insulin.

All of these bonds and forces can help hold the protein structure together and help determine the strength, density and biodegradability of individual proteins. It is the combined effect of all the amino acid side chains that determines the final 3-D structure of proteins and their chemical reactivity.

Spider web silk is made of proteins. One type of spider can spin seven different types of silk from its spinnerets, and its strongest silk is stronger than steel, with only one tenth of the weight. This silk has a tensile strength similar to the polymer Kevlar, which is used to make such items as bulletproof clothing.

Modelling molecules

A **model** is a representation of reality. It can be a diagram, a physical model, a description, a map or a flow chart.

Making models of molecules, especially complex molecules like proteins, helps to summarise what is known, to visualise a molecule more accurately and to predict its 3-D structure and chemical behaviour. The ability to do this is essential in areas such as the design of drugs and in biotechnology.

Models of small molecules can be made by hand, from physical materials. Ball and stick models can be extremely useful, but they do have their limitations, e.g. atoms are not coloured and they do not have hard surfaces, and bonds between atoms are not shaped like a little stick.

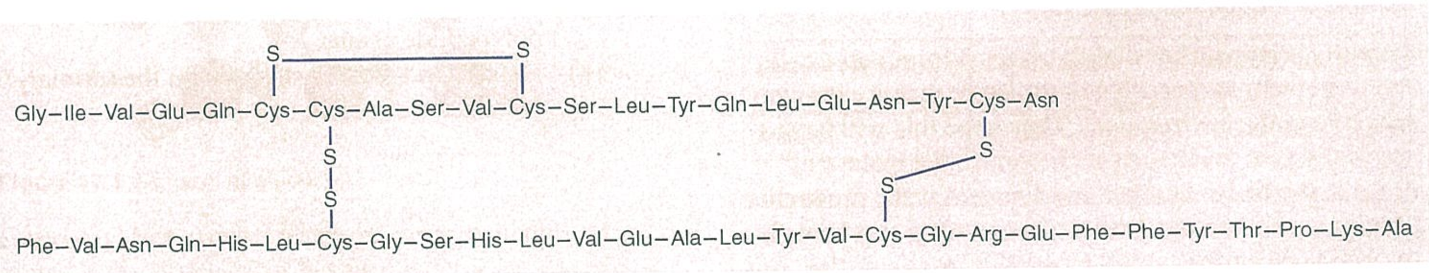


Figure 16.7 Disulfide links in insulin.