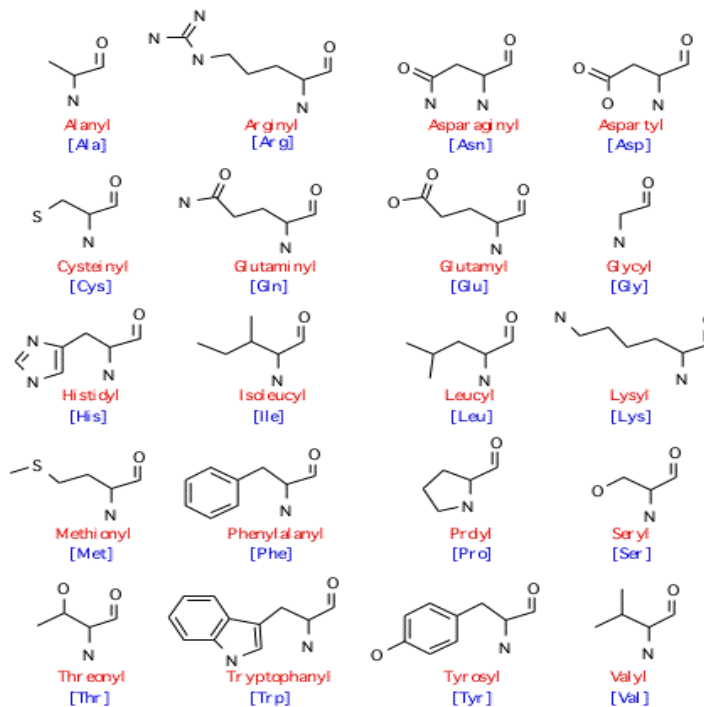


Biology

How does the primary sequence of a protein dictate its structure and function?

- Explain all the different types of structures.
- Use a range of proteins stating their functions linking back to their primary structure.
- Use diagrams.

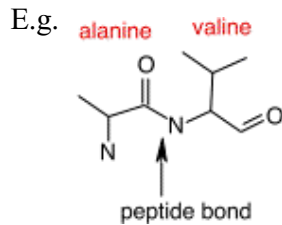
First discovered in 1838, proteins are complex organic compounds containing the elements carbon, hydrogen, oxygen, nitrogen and sometimes sulphur. Other proteins can also form complexes with other molecules containing phosphorus, iron, zinc and copper. These are known as prosthetic groups. Proteins are macromolecules of high relative molecular mass (between several thousand and million) consisting of chains of amino acids. There are twenty different types of amino acids, as shown below with their chemical structure.



These can be classified into three groups. These are polar, non polar and charged. Polar and charged amino acids will often be found on the surface of the protein interacting with the surrounding water (hydrophilic). Non polar (hydrophobic) amino acids will bury themselves in the interior. Amino acids are primary amines that contain an alpha carbon that is connected to an amino (NH_2) group, a carboxyl group (COOH) and a variable side group (R) – amino acids shown above. This side group gives each amino acid a distinctive property and function.



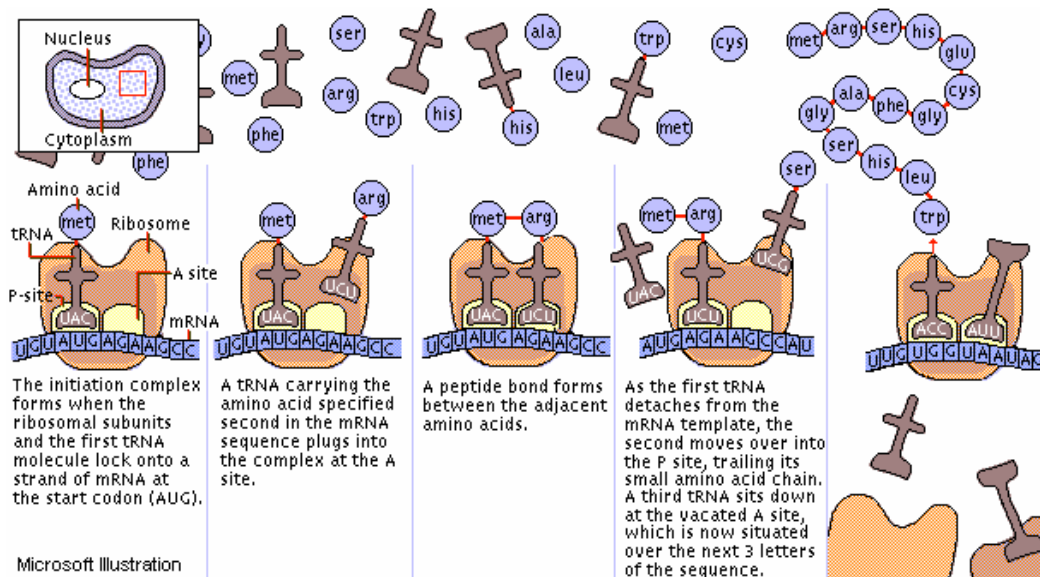
Linking an amino group to a carboxyl group on another amino acid then creates polymers. This is known as a peptide bond forming polypeptide chains by a condensation reaction.



The almost numberless combinations in which the acids line up and the helical and globular shapes into which the strands coil also explain the tasks that proteins perform. Each protein possesses a characteristic three-dimensional shape known as its conformation. The structure consists of four different levels. These are: -

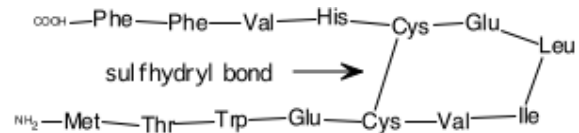
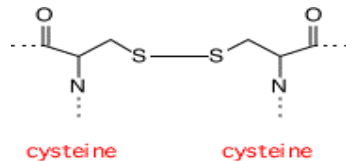
- Primary structure,
- Secondary structure,
- Tertiary structure and
- Quaternary structure.

The primary structure is the number and sequence of amino acids held together by peptide bonds in a polypeptide chain. There are over 10000 proteins in the human body all composed of different arrangements of the 20 amino acids. The sequence of amino acids of a protein dictates its biological function. In turn, this sequence is strictly controlled by the sequence of bases in DNA. Peptides and proteins are formed when a ribosome link numerous amino acids (10 – 10000) together in a long polymer. This long chain is known as the primary sequence. It is formed in the following way by protein synthesis: -



The production of proteins takes place in the cytoplasm of a cell. There are three main steps. At the start (*far left*) all of the necessary parts of the process are brought together by a small molecule called a ribosome. During elongation, amino acids, the building blocks of proteins, are joined to one another in a long chain. The sequence in which the amino acids are added is determined by the messenger RNA (mRNA), a transcribed copy of the DNA in every cell's nucleus. Termination (*far right*), takes place when the mRNA sequence contains one of several "stop" codons. A release factor binds to the mRNA at these sequences and triggers the breakup of the ribosome complex. The released chain is called the primary structure of a protein.

During and after synthesis the primary sequence will change to a more stable structure for the protein (secondary structure). How the protein folds is largely dictated by the primary sequence of amino acids. Each amino acid will in the sequence will join with other amino acids to conserve the most energy. Numerous bonds stabilize this structure. The first is hydrogen bonds that will form between the CO and NH groups. The H atom of the NH group of one amino acid is bonded to the O atom of the CO group three amino acids away. The second are sulfhydryl linkages. These are covalent bonds between cysteine groups. Cysteine has a sulphur group available for binding to other groups. Often in proteins, adjacent sulfhydryl groups on cysteins will form a covalent link in a protein, which makes them crucial for a protein to perform its function.



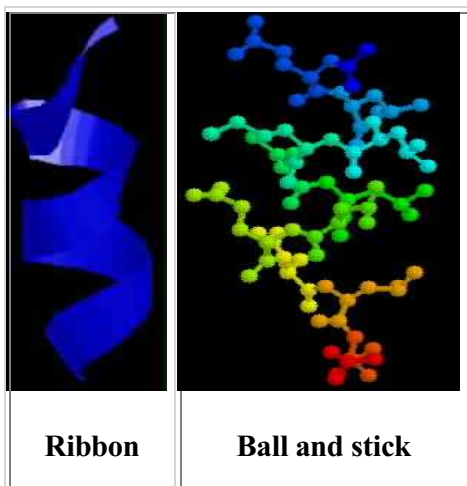
The chemical structure of a sulphydryl bond

A sulphydryl bond in a peptide

Others include hydrophobic interactions and ionic interactions.

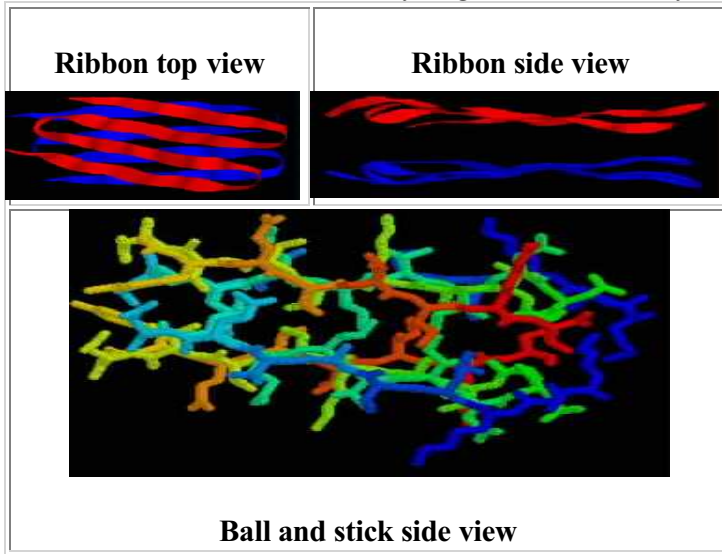
Proteins will often have stretches of amino acids that will show two common secondary structures. These are the alpha helix and the beta (pleated) sheets. The hydrogen bonding and the hydrophobic interactions between amino acids in the protein decide the formation of these structures.

The alpha helix is shown as a ribbon (ball and stick diagram) of amino acids. This structure is very stable and flexible and is often seen in parts of a protein that may need to bend or move.



A protein that is entirely alpha-helical is also called fibrous.

In the beta pleated sheet, two planes of amino acids will form, lining up in such a way that hydrogen bonds can form between facing amino acids in each sheet. The beta sheet is different than the alpha helix in those far distant amino acids in the protein can come together to form this structure. This structure tends to be rigid and less flexible. Most proteins are globular molecules in that they contain beta sheets and have irregular structures. This is mainly due to the interference in hydrogen bonding in certain R groups, the occurrence of disulphide bridges between different parts of the same chain and the lack of hydrogen bonds made by the amino acid proline.



Usually the polypeptide chain bends and folds forming a more compact, globular shape. This is the proteins tertiary structure and is maintained by the ionic, hydrogen, disulphide bonds and hydrophobic interactions.

Many enzymes and structures actually have more than one polypeptide chain. The arrangement of these chains is known as the quaternary structure. Proteins can contain several copies of an identical protein or they may consist of a number of polypeptides in different ratios.

Proteins can be classified according to many categories. One of them is fibrous, globular and intermediate. Fibrous proteins tend to be structural, carrying out numerous functions in cells and organisms. They tend to have long parallel polypeptide chains cross-linked at intervals forming long fibres or sheets. These include collagen, keratin, fibrinogen, fibroin and muscle proteins.

Collagen, which makes up bone, skin, tendons and cartilage is the most abundant protein found in vertebrates. The molecule usually contains three very long polypeptide chains each with about 1,000 amino acids that twist into a regularly triple helix. They give tendons and skin their great tensile strength. The complete triple helix compound is called tropocollagen. When boiling denatures long collagen fibrils, their chains are shortened to form gelatin.

Keratin, which makes up the outermost layer of the human skin, hair and nails and the scales, hooves and feathers of animals, twists into a regularly repeating alpha helix. It protects the body from the environment and is completely insoluble in water.

Its many disulphide bonds make it an extremely stable protein, being able to resist proteolytic (protein-hydrolysing) enzymes.

Fibrinogen is a blood plasma protein responsible for blood clotting. With the catalytic action of thrombin, fibrinogen is converted into molecules of the insoluble protein fibrin, which link together to form clots.

Fibroin is the protein used by silkworms when spinning their cocoon threads. The protein consists of a number of chains that are longer than the alpha helices. They run in opposite directions to each other and are arranged in parallel order. They are joined together by hydrogen bonds formed between the C=O and NH groups of one chain and the NH and C=O groups of opposite chains. This makes the structure very stable. This is called beta-configuration, and the whole structure is known as a beta-pleated sheet. This sheet has a high tensile strength and can't be stretched, but the arrangement of the polypeptides makes the silk very supple.

A muscle protein such as myosin, the protein responsible for muscle contraction, combines with actin, another muscle protein, forming actomyosin. The different filaments of actomyosin shorten, causing the contracting action of the muscle.

Globular proteins are spherical and highly soluble. They are part of the body's metabolism. Examples include albumin, globulin, casein, haemoglobin, all the enzymes and protein hormones.

All the enzymes are globular proteins that combine with other substances, called substrate, to catalyze the chemical reactions in the body. They are responsible for metabolism and its regulation, these molecules have catalytic sites on which substrate fits in a lock-and-key to trigger and control metabolism throughout the body.

Protein hormones come from the endocrine glands and do not act as enzymes. Instead they stimulate target organs that which control important activities, for example, the rate of metabolism and the production of digestive enzymes and milk. Insulin, secreted by the islets of Langerhans, regulates carbohydrate metabolism by controlling blood glucose levels. Thyroglobulin, from the thyroid gland, regulates overall metabolism and calcitonin, also from the thyroid, lowers blood calcium levels.

Antibodies (immunoglobulins) make up the thousands of different proteins that are generated in the blood serum in reaction to antigens (body-invading substances or organisms). A single antigen may bring out the production of many antibodies, which combine with different sites on the antigen molecule, neutralize it, and cause it to precipitate from the blood.

Myoglobin and haemoglobin are known as conjugated as they have a prosthetic group (non protein) group attached to them.

Haemoglobin consists of four separate polypeptide chains of two types, each with an individual prosthetic group. Generally it has two alpha chains and two beta chains. The two alpha chains each contain 141 amino acids and the two beta chains each contain 146 amino acids. The function of haemoglobin is transport oxygen to vertebrate blood.

Myoglobin consists only of one polypeptide chain containing one prosthetic group – iron. Myoglobin is another oxygen carrier found mainly in muscles.

