PROTEINS

SIZE, COMPOSITION, STRUCTURE AND FUNCTION

AUGUST 6, 2020. **BIO 101**

Protein

A protein is a functional biological molecule that is made up of one or more polypeptides that are folded/coiled into a specific structure. Proteins are important macromolecules that serve as structural elements, transportation channels, signal receptors and transmitters, and enzymes. Proteins are linear polymer that are built up of the monomer units called amino acids. There are 20 different amino acids and they are connected by a peptide bond between the carboxyl group and the amino group in a linear chain called a polypeptide. Each protein has different side chains or the "R" groups. Proteins have many different active functional groups attached to them to help define their properties and functions. Proteins cover a wide range of functions, ranging from very rigid structural elements to transmitting information between cells. Each person has several hundred thousands of different proteins in their body. Proteins are made from amino acids and therefore always contain the element **Carbon, Hydrogen, Oxygen** and **Nitrogen**, and in some cases **Sulphur**. Some proteins form complexes with other molecules containing **Phosphorus, Iron, Zinc** and **Copper.** Proteins are macromolecules of high M, (relative formula mass or molecular mass), typically between several thousands and several millions, consisting of chains of amino acids. All proteins have primary, secondary and tertiary structures but quaternary structures only arise when a protein is made up of two or more polypeptide chains. The folding of proteins is also driven and reinforced by the formation of many bonds between different parts of the chain. The formation of these bonds depends on the amino acid sequence. The study of their structures is important because proteins are essential for every activity in the human body as well as they are the key components of biological materials. Proteins are the most abundant organic molecules to be found in cells and form over 50% of their total dry mass. They are an essential component of the diet of animals and may be converted to both fat and carbohydrate by the cells. Their diversity enables them to display a great range of structural and metabolic activities within the organism.

Size of Protein

Simple peptides containing two, three or four amino acids residues are called di-, tri- and tetrapeptides respectively. Polypeptides are chains of many amino acid residues (up to several thousand -table 3.6). A protein may possess one or more polypeptide chains.

Table 3.6 Sizes of some proteins.

Protein	M_r *(molecular mass)*	*Number of amino acids*	*Number of polypeptide chains*
Ribonuclease	12 640	124	1
Lysozyme	13 930	129	1
Myoglobin	16 890	153	1
Haemoglobin	64 500	574	4
TMV (tobacco mosaic virus)	about 400 000 000	about 336 500	2130

The largest protein complexes are found in viruses where M_rs of over 400 000 000 are commonly found.

Classification of Proteins

Because of the complexity of protein molecules and their diversity of function, it is very difficult to classify them in a single, well-defined fashion. Three alternative methods are given in tables 3.7(according to structure), 3.8(according to composition) and 3.9(according to function)

Table 3.7 Classification of proteins according to structure.

Type	*Nature*	*Function*
Fibrous 	Secondary structure most important (little or no tertiary structure) Insoluble in water Physically tough Long parallel polypeptide chains cross-linked at intervals forming long fibres or sheets	Perform structural functions in cells and organisms e.g. **collagen** (tendons, bone, connective tissue), **myosin** (in muscle), **silk** (spiders' webs), **keratin** (hair, horn, nails, feathers)
Globular 	Tertiary structure most important Polypeptide chains tightly folded to form spherical shape Easily soluble	Form enzymes, antibodies and some hormones, e.g. **insulin** Other important roles
Intermediate	Fibrous but soluble	e.g. **fibrinogen** – forms insoluble fibrin when blood clots

Classification of proteins according to composition

1. Simple proteins: Only amino acids form their structure
2. Conjugated proteins: complex compounds consisting of globular proteins and tightly-bound non-protein material, the non-protein material is called prosthetic group.

Table 3.8 Classification of proteins according to composition (Conjugated proteins)

Name	**Prosthetic Group**	**Location**
Phosphoprotein	Phosphoric acid	Casein of milk, Vitellin of egg yolk
Glygoprotein	Carbohydrate	Membrane structure, mucin (component of saliva)
Nucleoprotein	Nucleic acid	Component of viruses, chromosomes, ribosome structure.
Chromoprotein	Pigment	Haemoglobin-haem (iron-containing pigment) Phytochrome (plant pigment) Cytochrome (respiratory pigment)
Lipoprotein	Lipid	Membrane structure Lipid transported in blood as lipoprotein
Flavoprotein	Flavine adenine dinucleotide (FAD)	Important in electron transport chain in respiration
Metal Protein	Metal	e.g nitrate reductase, the enzyme in plant which convert nitrate to nitrite.

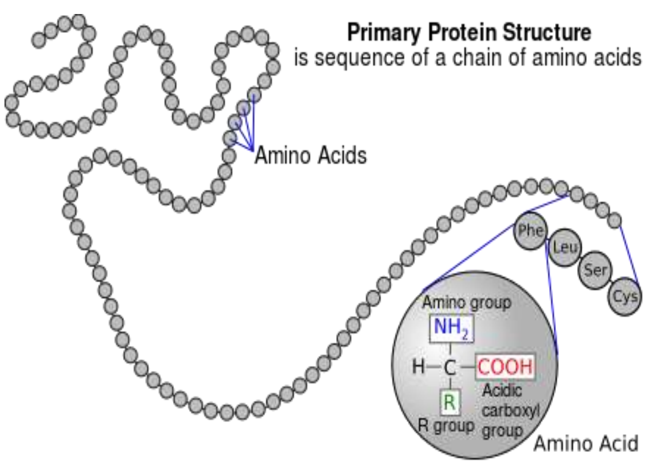
Table 3.9 Classification of proteins according to according to function

Type	**Examples**	**Occurrence/Function**
Structural	Collagen Keratin Elastin Viral coat proteins	Competent of connective tissue, bone, tendons, cartilage Skin, feathers, nails, hairs, horn Elastic connective tissue (ligament) Wraps up nucleic acid of virus
Enzymes	Trypsin Ribulose-bisphosphate carboxylase Glutamine synthetase	Catalyses hydrolysis protein Catalyses carboxylation (addition CO_2) of ribulose bisphosphate in photosynthesis Calalyses synthesis of amino acid glutamine from glutamic acid + ammonia
Hormones	Insulin Glucagon ACTH	Insulin and glucagon help to regulate glucose metabolism Stimulates growth and activity of the adrenal cortex
Respiratory pigment	Haemoglobin Myoglobin	Transport O_2 in vertebrate blood Store O_2 in muscles
Transport	Serum albumin	Transport fatty acids and lipids in blood
Protective	Antigens Fibrinogen Thrombin	Form complexes with foreign protein Form fibrin in blood clothing Involved in blood clotting mechanism
Contractile	Myosin Actin	Moving filaments in myofibrils of muscle Stationary filaments in myofibrils of muscle
Storage	Ovalbumin Casein	Egg while protein Milk protein
Toxins	Snake venom Diphtheria toxin	Enzymes Toxin made by diphtheria bacteria

Structure of Proteins

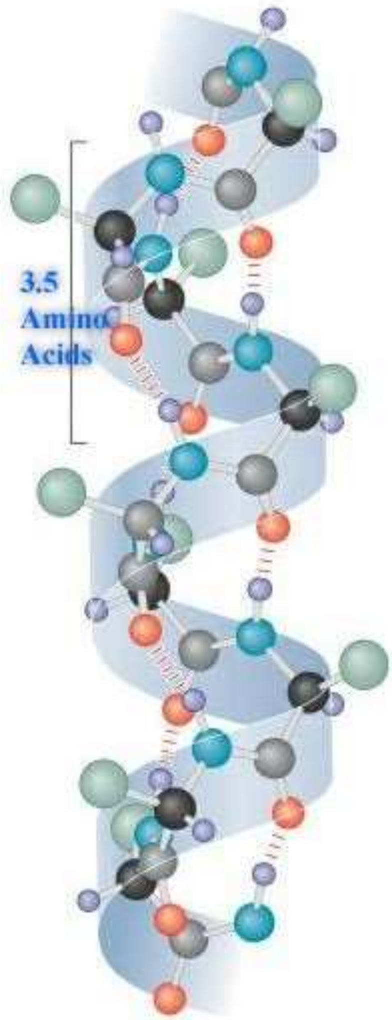
Each protein possesses a characteristic three-dimensional shape, its conformation. There are four separate levels of structure and organization of as follows:

Primary: The primary structure of a protein is the level of protein structure which refers to the specific sequence of amino acids. When two amino acids are in such a position that the carboxyl groups of each amino acid are adjacent to each other, they can be combined by undergoing a dehydration reaction which results in the formation of a peptide bond. Amino acids in a polypeptide (protein) are linked by peptide bonds that begin with the N-terminal with a free amino group and ends at C-terminal with a free carboxyl group. The peptide bond is planar and cannot rotate freely due to a partial double bond character. While there is a restricted rotation about peptide bond, there are two free rotations on (N-C) bond and (C-C) bond, which are called torsion angles, or more specifically the phi and psi angles. The freedoms of rotation of these two bonds are also limited due to steric hindrance. Genes carry the information to make polypeptides with a defined amino acid sequence. An average polypeptide is about 300 amino acids in length, and some genes encode polypeptides that are a few thousand amino acids long. It's important to know the primary structure of the protein because the primary structure encodes motifs that are of functional importance in their biological function; structure and function are correlated at all levels of biological organization.

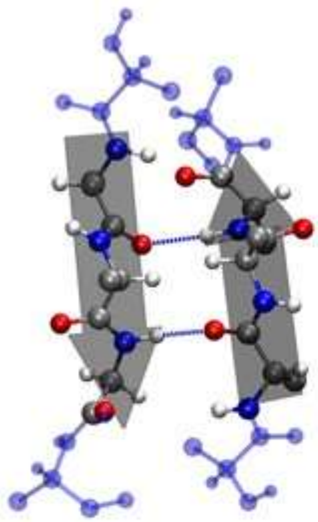


A picture of primary structure of protein.

Secondary: The amino acid sequence of a polypeptide, together with the laws of chemistry and physics, cause a polypeptide to fold into a more compact structure. Amino acids can rotate around bonds within a protein. This is the reason proteins are flexible and can fold into a variety of shapes. Folding can be irregular or certain regions can have a repeating folding pattern. The coils and folds that result from the hydrogen bonds between the repeating segments of the polypeptide backbone are called secondary structures. Although the individual hydrogen bonds are weak, they are able to support a specific shape for that part of the protein due to the fact that they are repeated many times over a long part of the chain. Secondary structures of a protein are proposed by Pauling and Corey. Its structures are formed by amino acids that are located within short distances of each other. Because of the planar nature of the peptide bonds, only certain types of secondary structure exist. The three important secondary structures are α-helix, β-sheets, and β-turns. Also, the beta sheets can be parallel, antiparallel, or mixed. Antiparallel beta sheets are more stable because the hydrogen bonds are at a ninety-degree angle. The a-helix is a coiled structure stabilized by intrachain hydrogen bonds.



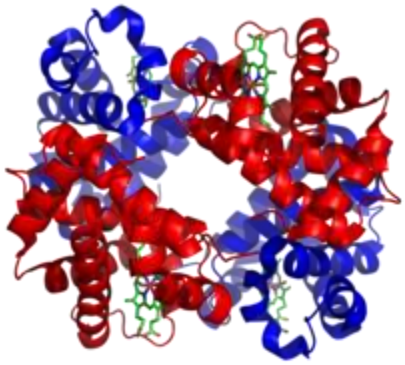
One type of secondary structure, an alpha helix.



Another type of secondary structure, a beta sheet.

Tertiary: As the secondary structure becomes established due to the primary structure, a polypeptide folds and refolds upon itself to assume a complex three-dimensional shape called the protein tertiary structure. Tertiary structure is the overall shape of a polypeptide. Tertiary structure results from the interactions between the side chains (R groups) of the various amino acids. This three-dimensional structure is due to intramolecular interactions between the side groups along the polypeptide chain. Its domain typically contains 300 – 400 amino acids, and it adopts a stable tertiary structure when it is isolated from their parent protein. As a polypeptide folds into its functional shape, amino acids that have hydrophobic side chains tend to end up clustered at the core of the protein so that they are out of contact with water. Covalent bonds called disulphide bridges can also affect the shape of a protein. Disulphide bridges form where two amino acids containing sulfhydryl groups on their side chains are brought close together by how the protein is folding. For some proteins, such as ribonuclease, the tertiary structure is the final structure of a functional protein. Other proteins are composed of two or more polypeptides and adopt a quaternary structure.

Quaternary: While all proteins contain primary, secondary and tertiary structures, quaternary structures are reserved for proteins composed of two or more polypeptide chains. Proteins that have quaternary structures contain more than one polypeptide and each adopt a tertiary structure and then assemble with each other via intermolecular interactions. The quaternary structure of a protein is the overall structure that is the result of the addition of these polypeptide subunits. The individual polypeptides are called protein subunits, which means different polypeptides folded separately. Subunits may be identical polypeptides or they may be different. When proteins consist of more than one polypeptide chain, they are said to have quaternary structure and are also known as multimeric proteins, meaning proteins consisting of many parts. Quaternary structures can also define as when more than one protein come together to create either a dimer, trimer, tetramer, etc…. Haemoglobin is an example of a quaternary structure that is composed of two alpha subunits and two beta subunits.



A picture of Haemoglobin, one of the most well-known quaternary structure of protein.