**PROTEINS**

Proteins are the most versatile molecules present in living organisms, from lower prokaryotes to higher eukaryotes, where they perform functions essential for life. The term “protein” derives from the Greek word “protos”, that means primary or preeminent. They are a class of molecules present in all living organisms and in all compartments of the cell; in animal cells they may constitute more than 50% of their dry weight.

Proteins are involved in virtually all biological processes. Thus, most chemical reactions occurring in life forms are catalysed by specific proteins called enzymes that are able to increase reaction rates by several orders of magnitude. Proteins can also transport and store a wide array of ions and small molecules as well as electrons. They possess hormonal activity and, in the form of antibodies which distinguish between self and nonself, they defend organisms against intruders. They participate in the reception and transmission of signals and stimuli at both intracellular and extracellular levels. They play crucial roles in the regulation of the expression of genetic information, at either transcription or translation, in connection with the control of growth and differentiation of cells. They are also necessary for providing the mechanical support and filamentous architecture within and between cells, and consequently are essential to cellular contraction and coordinated motion.

Whatever their structural or functional role, all proteins are polymers composed of the same building blocks, the **amino acids**. The physical and chemical properties of a protein are determined by its constituent amino acids. The individual amino acid subunits are joined by amide linkages called **peptide bonds**. The property that distinguishes peptide bonds from other ligands is not easily break up. Proteins in this count are strong and stable. They differ only in the number, the nature, and the sequential order of their constituent amino acids.

**Chemical Composition and Structure of Proteins**

Proteins are constituted from the same 20 different amino acids. Each amino acid comprises an amino group, a carboxyl group, a hydrogen atom, and a specific R group bonded to a carbon atom called the a-carbon. The R group is referred to as the ‘side-chain’ and varies in size, charge, shape and chemical composition from one amino acid to the other.

 

The carboxyl and amino groups of amino acids are ionized in solution at neutral physiological pH, with the carboxyl group bearing a negative charge and the amino group a positive charge. The carboxyl group loses a proton, giving a carboxylate ion, and the amino group is protonated to an ammonium ion. This structure is called a dipolar ion or a **zwitterion**. The dipolar nature of amino acids gives them some unusual properties. Amino acids have high melting points, generally over 200 °C. Amino acids are more soluble in water than they are in ether, dichloromethane, and other common organic solvents.

Because amino acids contain both acidic and basic groups, they are **amphoteric** (having both acidic and basic properties). The predominant form of the amino acid depends on the pH of the solution. In an acidic solution, the group is protonated to a free group, and the molecule has an overall positive charge. As the pH is raised, they loses its proton, the molecule has an overall negative charge. There is an intermediate pH where the amino acid is evenly balanced between the two forms, as the dipolar zwitterion with a net charge of zero. This pH is called the **isoelectric pH** or the **isoelectric point**. Electrophoresis uses differences in isoelectric points to separate mixtures of amino acids.

In general, the ionization properties of the constituent amino acids, including those of their side-chains, greatly influence the solubility, stability and structural organization of proteins. Similarly, the hydrophobicity/hydrophilicity of the side-chains plays an important role in the physicochemical behavior of polypeptide chains and their folding into three-dimensional structures.

Structurally, proteins and peptides are very similar, being made up of chains of amino acids that are held together by peptide bonds. So, what distinguishes a peptide from a protein? The basic distinguishing factors are size and structure. Peptides are smaller than proteins. Peptides, however, may be subdivided into oligopeptides, which have few amino acids (e.g., 2 to 20), and polypeptides, which have many amino acids (20 to 50). Proteins are formed from one or more polypeptides joined together. Hence, proteins essentially are very large peptides. In addition, proteins can adopt complex conformations known as secondary, tertiary, and quaternary structures.

*The primary structure* simply refers to the linear sequence of amino acids joined to each other through peptide bonds. The sequence of amino acids determines the basic structure of the protein. All the properties of the protein are determined, directly or indirectly, by the primary structure. Any folding, hydrogen bonding, or catalytic activity depends on the proper primary structure.

*The secondary structure* results from the formation of hydrogen bonds between contiguous parts of the polypeptide chain with particular amino acid sequences. Several folding patterns occur repeatedly in parts of protein molecules. There are essentially two types of secondary structure, alpha helix and beta sheet. When a peptide chain winds into a helical coil, each carbonyl oxygen can hydrogen-bond with a hydrogen on the next turn of the coil. Many proteins wind into a helix with the side chains positioned on the outside of the helix. Segments of peptides can also form orderly arrangements of hydrogen bonds by lining up side-by-side. In this arrangement, each carbonyl group on one chain forms a hydrogen bond with a hydrogen on an adjacent chain. This arrangement may involve many peptide molecules lined up side-by-side, resulting in a two-dimensional sheet. A protein may or may not have the same secondary structure throughout its length. Some parts may be curled into a helix, while other parts are lined up in a pleated sheet.

*The tertiary structure*, also called “native structure”, is the three-dimensional structure of the proteins. The tertiary structure includes all the secondary structure and all the kinks and folds in between. In tertiary structures, in addition to hydrogen bonding, amino acid side chains of the various secondary structures start interacting with each other in a number of ways. These interactions include hydrophobic interactions, ionic interactions, and disulfide bonds.

Many proteins are made up from two or more polypeptide chains, called subunits or monomers, which may have identical or different amino acid sequences. *The quaternary structure* refers to the association of two or more peptide chains in the protein. Therefore, it refers to the spatial arrangement of the individual chains and the nature of the forces that bind them together. Not all proteins have quaternary structure. The ones that do are those that associate together in their active form. For example, hemoglobin, the oxygen carrier in mammalian blood, consists of four peptide chains fitted together to form a globular protein.

**Classification of Proteins**

Proteins may be classified according to their composition, structure, or function. On the basis of their chemical composition, proteins may be divided into two classes: simple and complex. Simple proteins, are made up of only amino acids. Also, known as homoproteins. Examples are plasma albumin, collagen, and keratin. Complex or conjugated proteins are bonded to a non-protein prosthetic group. The prosthetic group may be either a metal or a compound. On decomposition with acids, these liberate the constituent amino acids as well as the prosthetic group. Their further classification is based on the nature of the prosthetic group present. The various divisions are chromoproteins, glycoproteins, phosphoproteins, lipoproteins and nucleoproteins. *Chromoprotein*, are proteins coupled with a coloured pigment. Typical examples are; hemoglobin and myoglobin, which bind heme groups; chlorophylls, which bind a porphyrin ring with a magnesium atom at its center. *Glycoproteins*, are the proteins containing carbohydrate as prosthetic group. Ovomucoid from egg white, mucin from saliva are some examples for this group. *Phosphoproteins*, are proteins linked with phosphoric acid; mainly acidic, e.g. casein from milk and ovovitellin from egg yolk. Proteins forming complexes with lipids (cephalin, lecithin, cholesterol) are called *lipoproteins*; soluble in water but insoluble in organic solvents. Examples are lipovitellin and lipovitellenin from egg yolk; lipoproteins of blood. *Nucleoproteins*, are compounds containing nucleic acid and protein, esp., protamines and histones. These are usually the salt-like compounds of proteins since the two components have opposite charges and are bound to each other by electrostatic forces. They are present in nuclear substances as well as in the cytoplasm. These may be considered as the sites for the synthesis of proteins and enzymes.

On the basis of their shape, proteins may be divided into two classes: fibrous and globular. Fibrous proteins, have primarily mechanical and structural functions, providing support to the cells as well as the whole organism. Their polypeptide chains form long filaments or sheets, where in most cases only one type of secondary structure, that repeats itself, is found. These are mainly of animal origin and are insoluble in water as they contain, both internally and on their surface, many hydrophobic amino acids.

The fibrous proteins are extremely strong and possess two important properties which are characteristic of the elastomers; they can stretch and later recoil to their original length and they have a tendency to creep. They are not digestible. It is a heterogeneous group and includes the proteins of connective tissues, bones, blood vessels, skin, hair, nails, horns, hoofs, wool and silk. *Collagens*, are of mesenchymal origin and form the major proteins of white connective tissues (tendons, cartilage) and of bone. The gelatin used in food preparation is a derivative of collagen by boiling in water, dilute acids or alkalines. *Elastins*, also of mesenchymal origin; form the major constituents of yellow elastic tissues (ligaments, blood vessels); differ from collagens in not being converted to soluble gelatins. *Keratins*, are of ectodermal origin; form the major constituents of epithelial tissues (skin, hair, feathers, horns, hoofs, nails); usually contain large amounts of sulfur in the form of cystine. *Fibrion*, is the principal constituent of the fibres of silk; composed mainly of glycine, alanine and serine units. It is produced by spiders and insects.

Globular proteins, possess a relatively spherical or ovoid shape. In this regard, tertiary and quaternary structures are found, in addition to the secondary structures. They are usually soluble in water or in aqueous media containing acids, bases, salts or alcohol, and diffuse readily. As a class, globular proteins are more complex in conformation than fibrous proteins, have a far greater variety of biological functions and are dynamic rather than static in their activities. At the intestinal level, most of the globular proteins of animal origin are hydrolyzed almost entirely to amino acids. Most of the proteins belong to this class. These are further classified mainly on their solubility basis as follows; albumins, protamins and histones, globulins, glutelins, prolamines. Albumins, are widely distributed in nature but more abundant in seeds; soluble in water,acids and bases and salts also coagulated by heat, e.g., leucosine in cereals, legumeline in legumes, ovalbumin from white of egg, serum albumin from blood plasma, myosin of muscles and lactalbumin of milk. *Globulins*, generally insoluble in water but soluble in natural sources, coagulated by heat. e.g. serum globulin from blood plasma, ovoglobulin from egg white; soybeans glycinine. *Glutelins*, have been isolated only from plant seeds; insoluble in water, dilute salt solutions and alcohol solutions but soluble in dilute acids and alkalies; coagulated by heat. Examples for this group are glutenin from wheat, glutelin from corn, oryzenin from rice. *Prolamines*, have also been isolated only from plant seeds; insoluble in water and dilute salt solutions but soluble in dilute acids and alkalies and also in 60 – 80% alcohol solutions; not coagulated by heat e.g., gliadin from wheat, zein from corn, hordein from oat. *Protamines and histones*, are basic proteins and occur almost entirely in animals, mainly in sperm cells; possess simplest structure and lowest molecular weight (approximately 5,000); soluble in water; unlike most other proteins, not coagulated by heat; strongly basic in character owing to high content of basic amino acids (lysine, arginine); form salts with mineral acids and nucleic proteins. Protamines are virtually devoid of sulfur and aromatic amino acids. Histones are somewhat weaker bases and are, therefore, insoluble in NH4OH solution, whereas the protamines are soluble. Examples of protamines; clupeine from herring sperm, salmine from salmon sperm, sturine from sturgeon and cyprinine from carp. Histones, nucleohistones of nuclei; globin of hemoglobin.

The multitude of functions that proteins perform is the consequence of both the folding of the polypeptide chain, therefore of their three-dimensional structure, and the presence of many different functional groups in the amino acid side chains, such as thiols, alcohols, thioethers, carboxamides, carboxylic acids and different basic groups. From the functional point of view, they may be divided into several groups like shown in the table.

|  |  |  |
| --- | --- | --- |
| **Class of protein** | **Function** | **Examples** |
| Enzymic proteins  | Biological catalysts Urease, | Amylase, Catalase,Cytochrome C, Alcoholdehydrogenase |
| Structural proteins  | Strengthening or protecting biological structures | Collagen, Elastin, Keratin, Fibroin |
| Transport or carrier proteins | Transport of ions or molecules in the body  | Myoglobin, Hemoglobin, Ceruloplasmin, Lipoproteins |
| Nutrient and storage proteins | Provide nutrition to growing embryos and store ions | Ovalbumin, Casein, Ferritin |
| Contractile or motile proteins | Function in the contractile system | Actin, Myosin, Tubulin |
| Defense proteins | Defend against other organisms  | Antibodies, Fibrinogen, Thrombin |
| Regulatory proteins | Regulate cellular or metabolic activities | Insulin, G proteins, Growth hormone |
| Toxic proteins | Hydrolyze (or degrade) enzymes | Snake venom, Ricin |