

Structure and Properties of Proteins

Proteins are macromolecules made up of linear polymers of amino acids. Proteins are most abundant molecules of living system and have high molecular weights. Proteins play very important functions in the cell such as enzyme catalysis, in immune response, transport and storage of molecules, in transmission of nerve impulse, control growth and differentiation, provide mechanical support, etc. Proteins are made up of carbon, hydrogen, oxygen and nitrogen and in some cases sulfur and phosphorus.

Amino acids

Proteins are linear polymers made up of small units of amino acids called monomers. Amino acids are thus building blocks of proteins. The general formula of an α -amino acid is shown in Figure 1.

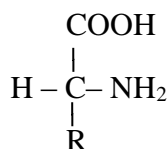


Figure 1: General formula of an amino acid

The central carbon atom is called α -carbon which is linked to four different groups; $-\text{COOH}$ group, NH_2 group, $-\text{H}$, and $-\text{R}$ group called side chain. Because of binding with four different groups α -carbon of amino acids is chiral or asymmetric. Because of this asymmetric α -carbon amino acids are present in two optically active forms or mirror image forms; L isomers and the D isomers (Figure 2). R can be $-\text{H}$ as in the case of glycine or it can be a $-\text{CH}_3$ as in alanine.



Figure 2: D and L amino acids.

In L isomers $-\text{NH}_2$ group is present to the left and in D-isomers $-\text{NH}_2$ group is present to the right of α -carbon. There are 20 standard amino acids present in proteins of biological system. These amino acids are L-isomers.

Electrochemical properties of amino acids

The α - COOH group and α - NH_2 groups of amino acids are ionized in solutions at physiological pH (pH=7). At this pH $-\text{COOH}$ group is present as deprotonated form ($-\text{COO}^-$) bearing negative charge and the amino group the ($-\text{NH}_3^+$) bearing a positive charge. This dipolar state of amino acids is called zwitterion (Figure 3).

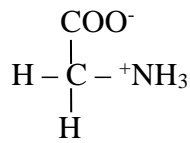


Figure 3: Zwitterion of glycine

Classification of amino acids

Classification of amino acids according to the polarities of R group is the most common way of classification. According to this classification there are five types of amino acids.

1. Non polar aliphatic amino acids

This group contains seven amino acids. Four amino acids glycine, alanine, valine, leucine and isoleucine have R groups of aliphatic hydrocarbon (Figure 4). Methionine, one of the two sulfur containing amino acids has a slightly non polar thiol ether side chain. Proline has a cyclic secondary amino (imino) group.

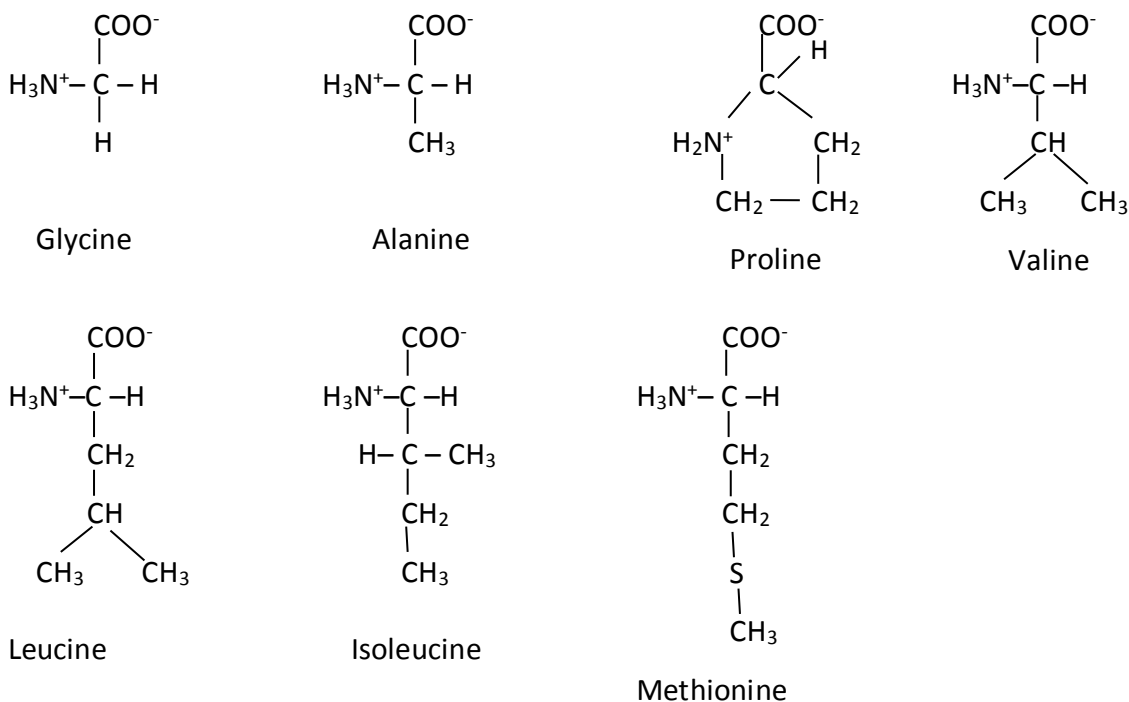


Figure 4: Structure of non-polar aliphatic amino acids.

2. Amino acids with aromatic R groups

Three amino acids namely phenylalanine, tyrosine and tryptophan have aromatic side chain that makes them slightly non polar (hydrophobic) and hence participate in hydrophobic interactions (Figure 5). Tyrosine and tryptophan are relatively more polar than phenylalanine due to presence of -OH group in tyrosine and the nitrogen atom in the indole ring of tryptophan. The

aromatic ring of these three amino acids absorb ultraviolet light which is the basis for the characteristic strong absorbance of light by proteins at a wavelength of 280 nm.

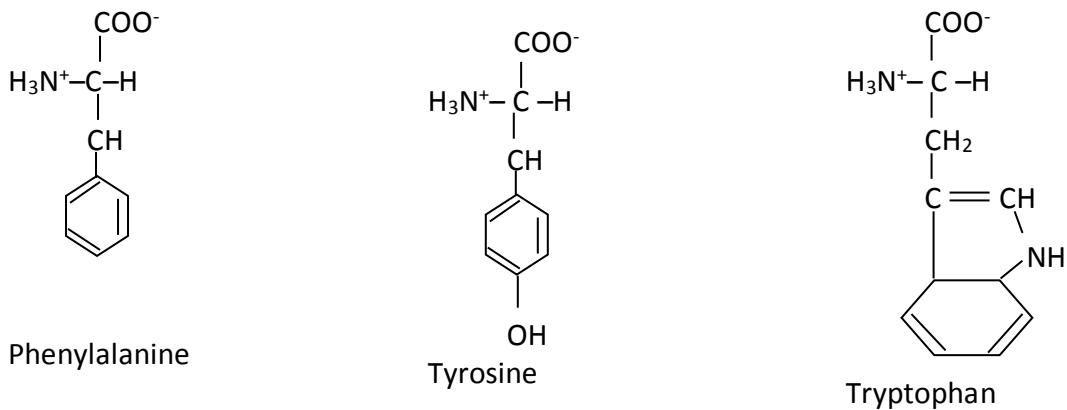


Figure 5: Structure of aromatic amino acids.

3. Amino acids with polar, uncharged R groups

The amino acids that belong to this class are serine, threonine, cysteine, asparagine and glutamine (Figure 6). The R groups of these amino acids are polar (hydrophilic) and therefore they are more soluble in water. Polarity in these amino acids is because they contain functional groups like -OH (serine and threonine), -CONH₂ (glutamine and asparagine), -SH (cysteine).

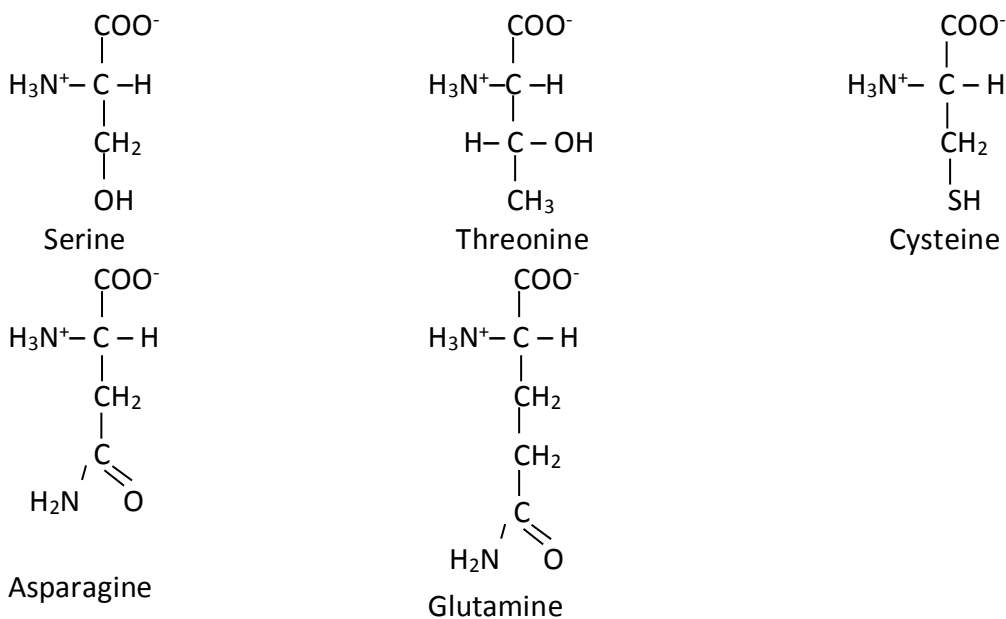


Figure 6: Structure of polar uncharged amino acids.

4. Amino acids with positively charged (basic) R groups

In this class of amino acids R groups have positive charge at physiological pH (Figure 7). Lysine has a second primary amino group at the ϵ position on its aliphatic chain. Arginine has a positively charged guanidium group and histidine has an aromatic imidazole group.

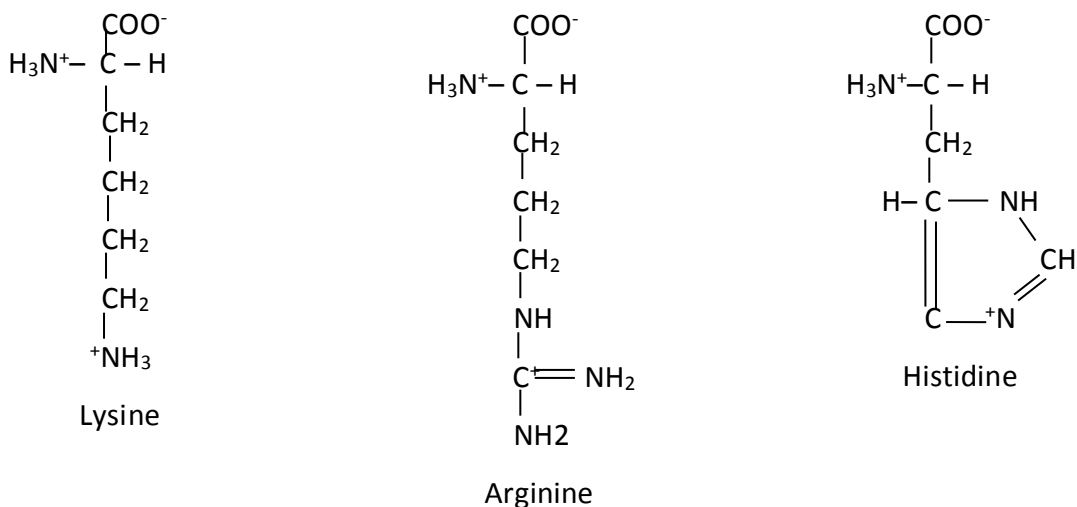


Figure 7: Structure of polar uncharged amino acids.

5. Amino acids with negatively charged (acidic) R groups

There are two amino acids having a negative charge at pH 7.0 (Figure 8). These are aspartate and glutamate; both contain a second -COOH group.



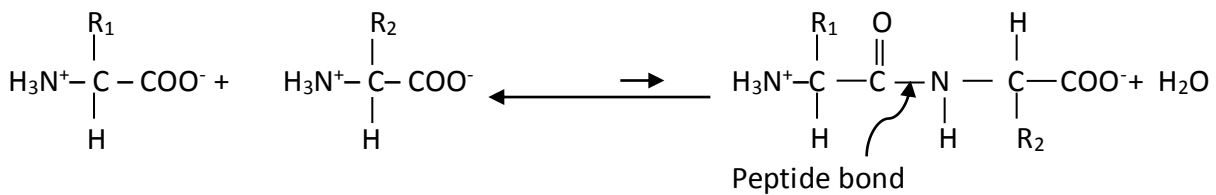
Figure 8: Structure of polar uncharged amino acids.

Primary structure of proteins

The 20 standard amino acids join through peptide bonds to form proteins. Protein has four levels of structure namely primary, secondary, tertiary and quaternary. The primary structure of proteins is the amino acid sequences that are linked by peptide bonds.

Peptide bond

The α -COOH of one amino acid links to α -NH₂ group of another amino acid through a peptide bond (also called amide bond) to form linear polymer. Formation of a peptide bond from two amino acids is accompanied by the loss of water molecule. The equilibrium of this reaction lies on the side of the hydrolysis rather than synthesis. Hence the biosynthesis of peptide bond requires an input of free energy.



Depending on the number of amino acids composing a chain, the peptides may be termed as a dipeptide (containing two amino acid units), a tripeptide (containing three amino acid units). Peptides containing less than ten amino acid residues are called oligopeptide. Peptides containing more than ten residues are termed polypeptide. Each amino acid unit in polypeptide is called a residue.

Conformation of peptide bond

The peptide bond which links amino acid residues in polypeptide chain is planar. For a pair of amino acids linked by a peptide bond, six atoms lie in the same plane. These atoms are α -carbon atom and C=O group from the first amino acid and NH group and α -carbon from the second amino acid. Planar structure is result of resonance interactions that gives the peptide bond an ~40% double bond character. Due to this double bond character the peptide bond can not rotate freely. A planar peptide bond has two possible conformations; trans and cis. In the trans conformation, the two α -carbon atoms are on opposite sides of the peptide bond. In cis conformation, these groups are on the same side of the peptide. Almost all peptide bonds in proteins are trans. Cis conformation is rare and less stable than trans conformation. However, cis conformation occurs in peptide bonds which are followed by proline residue. In these bonds nitrogen of proline is bonded to two tetrahedral carbon atoms, thus reducing the steric differences between the trans and cis conformations.

Flexibility of polypeptide backbone

In contrast to peptide bond, the bonds between the $-\text{NH}_2$ group and α -carbon and between α -carbon and $-\text{COOH}$ group are pure single bonds. The two adjacent rigid peptide units may rotate about these bonds taking on various orientations. This freedom of rotation about two bonds of each amino acids allows proteins to fold in many different ways. The rotation angles of these bonds can be specified by dihedral angles (torsion angles). The rotation angle about the $\text{C}_\alpha\text{-N}$ bond is called phi (ϕ). The angle of rotation between $\text{C}_\alpha\text{-C}$ bond is called psi (ψ) (Figure 9). The ϕ and ψ angles determine the path of the polypeptide chain. By convention both ϕ and ψ are defined 0° in the conformation in which the two peptide bonds connected to a single carbon in the same plane. In principle, ϕ and ψ can have any value between -180° and $+180^\circ$, but many values of ϕ and ψ are prohibited by steric interference between atoms in the

polypeptide backbone and the amino acid side chains. The conformation in which ϕ and ψ are both 0° is prohibited for this reason.

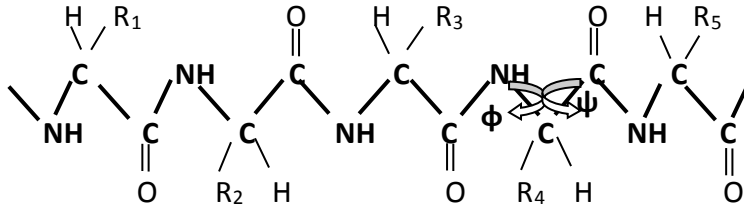


Figure 9: Rotation about bonds in polypeptide. ϕ is rotation angle between C_α-N bond and ψ is rotation angle between C_α-C bond.

Secondary structure of proteins

Folding of polypeptide chain is possible because of the presence of hydrogen bond. A regular secondary structure occurs when each dihedral angle ϕ and ψ remains the same or nearly same throughout the segment. In 1951 Linus Pauling and Robert Corey proposed two types of periodic structures called α helix and β pleated sheet.

α Helix

Pauling and Corey observed that a polypeptide chain with planar peptide bonds would form a right handed helical structure by simple twists about the C_α-N and the C_α-C bonds. They called this helical structure as α helix. An α helix is a rod like structure. The inner part of helix is made up of the tightly coiled backbone and the side chains extend outward in the helix. The protruding side chains determine the interaction of α helix both with other parts of the folded protein chain and with other protein molecules. The helix is stabilized by hydrogen bonds between NH and CO groups of the main chain. The α helix contains 3.6 amino acids per turn of the helix and has a pitch of 5.4 Å (0.54 nm), thereby giving a rise per turn residue of $5.4/3.6=1.5$ Å, which is the identity period of the helix. The value of rotational angles ϕ and ψ is -60° and -45° to -50° , respectively. The α helix can be right handed (clockwise) or left handed (counterclockwise). All known polypeptides contain right handed α helix. The occurrence of α helical content in proteins ranges widely. For example in ferritin, that helps storage of iron, has 75% of its amino acid residues form α helix.

β pleated sheet

The second type of periodic structure is β pleated sheet. In contrast to α helix β pleated sheet involve hydrogen bonds between groups from residues distant from each other in the linear sequence. In β sheets two or more strands widely separated in the protein sequence are arranged side by side, with hydrogen bonds between the strands. Based on the orientation of the strands β sheets are of two types. If strands run in the same direction they are called *parallel* β sheets; if strands run in opposite direction they are called *antiparallel* β sheets. In parallel arrangement the NH group is hydrogen bonded to the CO group of the one amino acid on the adjacent strand, whereas the CO group is hydrogen bonded to the NH group on the amino acids two residues farther along the chain.

Tertiary structure of protein

The overall three dimensional arrangement of all atoms in a protein is called as tertiary structure of protein. The amino acids that are far apart in polypeptide chain are in different

types of secondary structures may interact with each other to form completely folded structure of a protein. In tertiary structure there is an involvement of some additional bonds like disulfide, hydrogen, hydrophobic and ionic (Figure 10). This makes the protein globular in shape. The enzymes, transport protein, some peptide hormones and immunoglobulins are all globular in shape which is their tertiary structure. In these globular proteins the head group are located on the outer surface because of their hydrophilicity (water loving property) and non polar R groups are located interior where their interaction create a hydrophobic (water hating property) environment. The tertiary structure thus involves the folding of the helices of globular proteins.

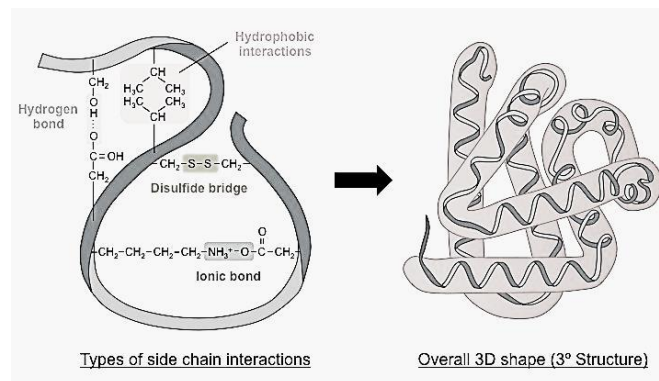


Figure 10: Tertiary structure of proteins.

Quaternary structure of protein

Some globular proteins consist of more than one polypeptide chain. Each polypeptide chain in such a protein is called subunit. These subunits may be identical or different in their primary structure. These subunits may associate specifically to each other to form large sized complex molecule known as quaternary structure (Figure 11.). Quaternary structure is the spatial arrangement of subunits and the nature of their contact. The same forces i.e. disulfide, hydrogen, hydrophobic and ionic bonds are involved in tertiary structure formation are also involved in quaternary structure to link various polypeptide chains. The simplest type of quaternary structure is a dimer consisting of two subunits. If the two subunits or polypeptides are identical the quaternary protein is termed homodimer. Example of homodimeric protein is Cro protein found in bacterial virus called λ (lambda) phage. If two different type of subunit are present then quaternary protein is called heterodimer. Example of heterodimeric protein is bovine insulin which regulates the metabolism of glucose. It consists of one subunit of one type (designated α) and one subunit of another type (designated β). Thus bovine insulin exists as $\alpha\beta$ heterodimer. Other examples are hemoglobin, which is the oxygen carrying protein in blood; consist of two subunits of one type (α) and two subunits of another type (β). Thus hemoglobin exists as $\alpha_2\beta_2$ heterotetramer.

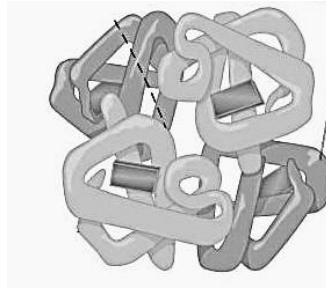


Figure 11: Quaternary structure of protein consisting of heterotetramer.