# 5-Biochemistry of Proteins

# Introduction

Proteins are the most abundant organic molecules of the living system. They occur in every part of the cell and constitute about 50 % of the cellular dry weight. Proteins form the fundamental basis of structure and function of life.

The term protein is derived from a Greek word proteios, meaning holding the first place.

Berzelius (Swedish chemist) suggested the name proteins to the group of organic compounds that are utmost important to life.

Mulder (Dutch chemist) in 1838 used the term **proteins** for the high molecular weight nitrogen- rich and most abundant substances present in animals and plants.

Proteins are predominantly constituted by five major elements in the following proportion- Carbon: 50-55 %, Hydrogen: 6-7.3 %, Oxygen: 19-24 %, Nitrogen: 13-19 % and Sulfur: 0-4 %.

The content of nitrogen, an essential component of proteins, on an average is 16 %. Estimation of nitrogen in the laboratory (mostly by Kjeldahl's method) is also used to find out the amount of protein in biological fluids and foods.

Proteins on complete hydrolysis (with concentrated HCl for several hours) yield L-$\alpha$-amino acids. This is a common property of all the proteins.

Proteins are the polymers of L-$\alpha$-amino acids.

Only 20-known as standard amino acids (Up to 300 amino acids occur in nature) are repeatedly found in the structure of proteins, isolated from different forms of life-animal, plant and microbial. This is because of the universal nature of the genetic code available for the incorporation of only 20 amino acids when the proteins are synthesized in the cells.

The process in turn is controlled by DNA, the genetic material of the cell. After the synthesis of proteins, some of the incorporated amino acids undergo modifications to form their derivatives.

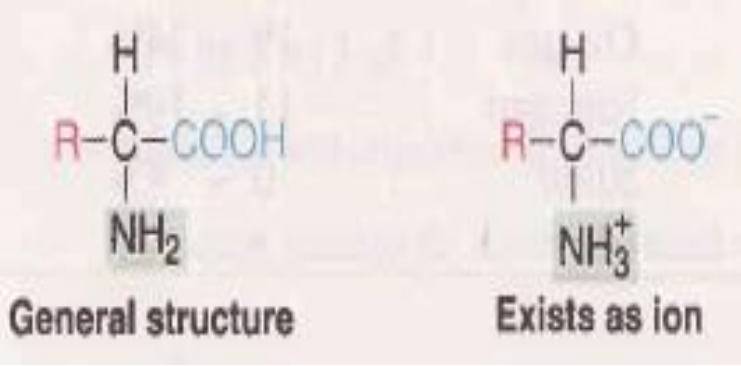
# Amino acids

Amino acids are a group of organic compounds containing two functional groups amino and carboxyl. The amino group (-NH2) is basic while the carboxyl group (-COOH) is acidic in nature.

**General structure of amino acids & optical isomers-**

The amino acids are termed as $\alpha$-amino acids, if both the carboxyl and amino groups are attached to the same carbon atom.

The $\alpha$-carbon atom binds to a side chain represented by R which is different for each of the 20 amino acids found in proteins. The amino acids mostly exist in the ionized form in the biological system. The amino acids possess four distinct groups (R, H, COO-, NH;) held by $\alpha$-carbon. All the amino acids (except glycine where R=H) have optical isomers.



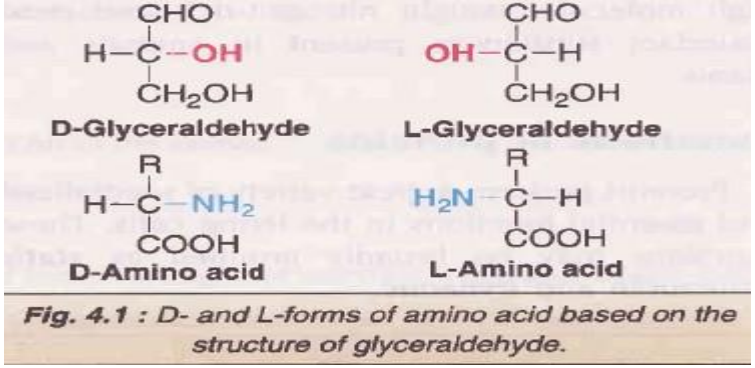




*Fig. 4.1 : D- and L-forms of amino acid based on the structure of glyceraldehyde.*

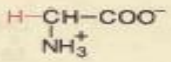
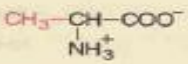
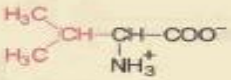
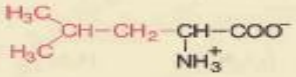
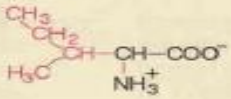
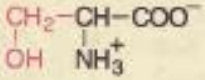
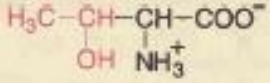
# Classification of amino acids

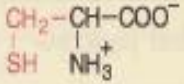
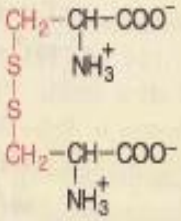
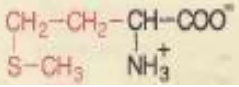
There are different ways of classifying the amino acids based on the structure and chemical nature, polarity, nutritional requirement, metabolic fate.

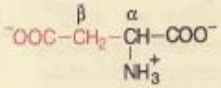
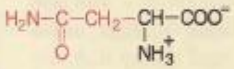
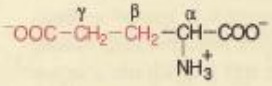
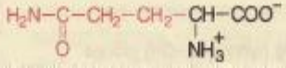
(A). Amino acid classification based on the structure:-

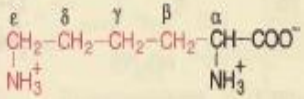
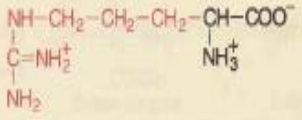
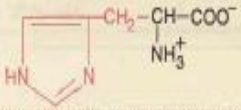
A comprehensive classification of amino acids is based on their structure and chemical nature. Each amino acid is assigned a 3 letter or 1 letter symbol. These symbols are commonly used to represent the amino acids in protein structure. The 20 amino acids found in proteins are divided into seven distinct groups.

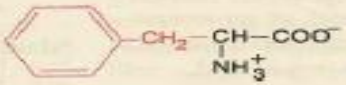
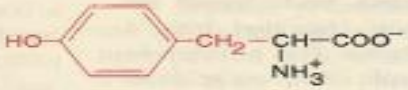
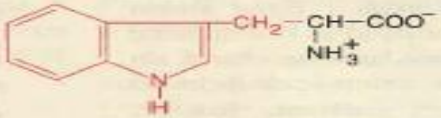
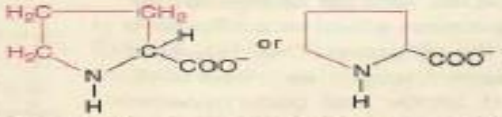
**Table 4.1 Structural classification of L-α-amino acids found in proteins**

| Name | Symbol | | Structure | Special group present |
|---|---|---|---|---|
| | 3 letters | 1 letter | | |
| **I. Amino acids with aliphatic side chains** | | | | |
| 1. Glycine | Gly | G |  | |
| 2. Alanine | Ala | A |  | |
| 3. Valine | Val | V |  | Branched chain |
| 4. Leucine | Leu | L |  | Branched chain |
| 5. Isoleucine | Ile | I |  | Branched chain |
| **II. Amino acids containing hydroxyl (—OH) groups** | | | | |
| 6. Serine | Ser | S |  | Hydroxyl |
| 7. Threonine | Thr | T |  | Hydroxyl |
| Tyrosine | Tyr | Y | See under aromatic | Hydroxyl |

| Name | Symbol | | Structure | Special group present |
|---|---|---|---|---|
| | 3 letters | 1 letter | | |
| **III. Sulfur containing amino acids** | | | | |
| 8. Cysteine | Cys | C |  | Sulfhydryl |
| Cystine | — | — |  | Disulfide |
| 9. Methionine | Met | M |  | Thioether |

| | | | | |
|---|---|---|---|---|
| **IV. Acidic amino acids and their amides** | | | | |
| 10. Aspartic acid | Asp | D |  | β-Carboxyl |
| 11. Asparagine | Asn | N |  | Amide |
| 12. Glutamic acid | Glu | E |  | γ-Carboxyl |
| 13. Glutamine | Gln | Q |  | Amide |

| | | | | |
|---|---|---|---|---|
| **V. Basic amino acids** | | | | |
| 14. Lysine | Lys | K |  | ε-Amino |
| 15. Arginine | Arg | R |  | Guanidino |
| 16. Histidine | His | H |  | Imidazole |

| Name | Symbol | | Structure | Special group present |
|---|---|---|---|---|
| | 3 letters | 1 letter | | |
| **VI. Aromatic amino acids** | | | | |
| 17. Phenylalanine | Phe | F |  | Benzene or phenyl |
| 18. Tyrosine | Tyr | Y |  | Phenol |
| 19. Tryptophan | Trp | W |  | Indole |
| **VII. Imino acid** | | | | |
| 20. Proline | Pro | P |  | Pyrrolidine |

*(Note : R group is shown in red)*

1.Amino acids with aliphatic side chains : These are monoamino monocarboxylic acids. This group consists of the most simple amino acids-glycine, alanine, valine, leucine and isoleucine. The three amino acids (Leu, lle, Val) contain branched aliphatic side chains, hence they are referred to as **branched chain amino acids.**

2.Hydroxyl group containing amino acids : Serine, threonine and tyrosine are hydroxyl group containing amino acids. Tyrosine-being aromatic in nature-is usually considered under aromatic amino acids.

3.Sulfur containing amino acids:- Cysteine with sulfhydryl group and methionine with thioether group are the two amino acids incorporated during the course of protein synthesis.

Cystine, another important sulfur containing amino acid, is formed by condensation of two molecule of cysteine.

4.Acidic amino acids and their amides:- Aspartic acid and glutamic acids are dicarboxylic monoamino acids while asparagine and glutamine are their respective amide derivatives.

5.Basic amino acids : The three amino acids lysine, arginine (with guanidino group) and histidine (with imidazole ring) are dibasic monocarboxylic acids. They are highly basic in character.

6.Aromatic amino acids:- Phenylalanine, tyrosine and tryptophan (with indole ring) are aromatic amino acids. Histidine may also be considered under this category.

7.lmino acids:- Proline containing pyrrolidine ring is a unique amino acid. lt has an imino group (=NH), instead of an amino group (-NH2) found in other amino acids. Proline is an $\alpha$-imino acid.

## (B). Classification of amino acids based on polarity:-

Amino acids are classified into 4 groups based on their polarity. The polarity in turn reflectst he functionalr ole of amino acids in protein structure.

**1.Non-polar amino acids :** These amino acids are also referred to as hydrophobic (water hating). They have no charge on the 'R' group. The amino acids included in this group are - alanine, leucine, isoleucine, valine, methionine, phenylalanine, tryptophan and proline.

**2.Polar amino acids with no charge on 'R‘ group :** These amino acids, as such, carry no charge on the 'R'group. They however possess groups such as hydroxyl, sulfhydryl and amide and participate in hydrogen bonding of protein structure.

The simple amino acid glycine (where R = H) is also considered in this category. The amino acids in this group are glycine, serine, threonine, cysteine, glutamine, asparagine and tyrosine.

**3. Polar amino acids with positive 'R' group :**

The three amino acids lysine, arginine and histidine are included in this group.

**4.Polar amino acids with negative 'R'group :** The dicarboxylic monoamino acids aspartic acid and glutamic acid are considered in this group.

## (C). Nutritional classification of amino acids:

Only 20 amino acids are required for the synthesis of variety of proteins, besides other biological functions. However, all these 20 amino acids need not be taken in the diet. Based on the nutritional requirements, amino acids are grouped into two class esssential and non-essential.

**1.Essential or indispensable amino acids:** The amino acids which cannot be synthesized by the body and, therefore, need to be supplied through the diet are called essential amino acids. They are required for proper growth and maintenance of the individual. The ten amino acids listed below are essential humans (and also rats)- Arginine, Valine, Histidine, Isoleucine, Leucine, Lysine, Methionine, Phenylalanine, Threonine, Tryptophan.

The two amino acids namely arginine and histidine can be synthesized by adults and not by growing children, hence these are considered as semi-essential amino acids. Thus, 8 amino acids are absolutely essential while 2 are semi-essential.

**2.Non-essential or dispensable amino acids:-** The body can synthesize about 10 amino acids to meet the biological needs, hence they need not be consumed in the diet. These are-glycine, alanine, serine, cysteine, aspartate, asparagnie, glutamate, glutamine, tyrosine and proline.

## (D). Amino acid classification based on their metabolic fate:-

The carbon skeleton of amino acids can serve as a precursor for the synthesis of glucose ( glycogenic) or fat (ketogenic) or both. From metabolic view point, amino acids are divided into three groups-

**1. Glycogenic amino acids:** These amino acids can serve as precursors for the formation of glucose or glycogen. e.g. alanine, aspartate, glycine, methionine etc.

**2.Ketogenic amino acids:** Fat can be synthesized from these amino acids. Two amino acids leucine and lysine are exclusively ketogenic.

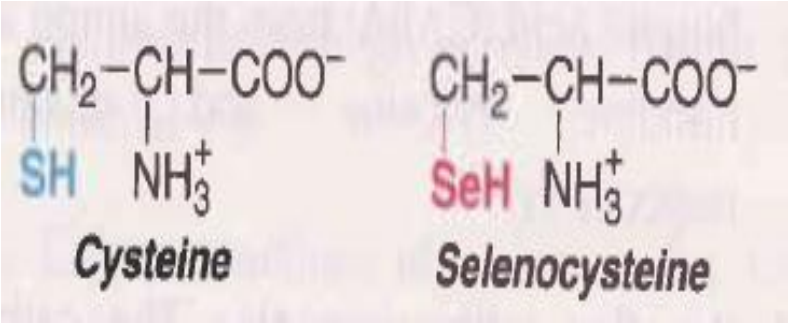
**3.Glycogenic and ketogenic amino acids:** The four amino acids isoleucine, phenylalanine, tryptophan, tyrosine are precursors for synthesis of glucose as well as fat.

# Selenocysteine - the 21st amino acid-

In recent years, a 21st amino acid namely selenocysteine has been added. It is found at the active sites of certain enzymes/proteins (selenoproteins). e.g. glutathione peroxidase, glycine reductase, 5'-deiodinase, thioredoxin reductase. Selenocysteine is an unusual amino acid containing the trace element selenium in place of the sulfur atom of cysteine.

Incorporation of selenocysteine into the proteins during translation is carried out by the codon namely **UGA.**

Another amino acid namely pyrrolysine as the 22nd amino acid present in protein. The stop codon **UAG** can code for pyrrolysine.



# Properties of amino acids

**A. Physical properties-**

1.Solubility:- Most of the amino acids are usually soluble in water and insoluble in organic solvents.

2.Melting points:- Amino acids generally melt at higher temperatures, often above 200°C.

3.Taste:- Amino acids may be sweet (Gly, Ala, Val), tasteless (Leu) or bitter (Arg, lle). Monosodium glutamate (MSG; ajinomoto) is used as a flavoring agent in food industry and Chinese foods to increase taste and flavor. In some individuals intolerant to **MSG, Chinese restaurant syndrome** (brief and reversible flu like symptoms) is observed.

4.Optical properties:- All the amino acids except glycine possess optical isomers due to the presence of asymmetric carbon atom. Some amino acids also have a second asymmetric carbon e.g. isoleucine, threonine.

5.Amino acids as ampholytes:- Amino acids contain both acidic (-COOH) and basic (-NH2) groups. They can donate a proton or accept a proton, hence amino acids are regarded as ampholytes.

Zwitterion or dipolar ion:- The name zwitter is derived from the German word which means hybrid. Zwitter ion (or dipolar ion) is a hybrid molecule containing positive and negative ionic groups.

The amino acids rarely exist in a neutral form with free carboxylic (-COOH) and free amino (-NH2) groups. In strongly acidic pH (low pH), the amino acid is positively charged (cation) while in strongly alkaline pH (high pH), it is negatively charged (anion). Each amino acid has a characteristic pH (e.g. leucine, pH 6.0) at which it carries both positive and negative charges and exists as zwitterion.

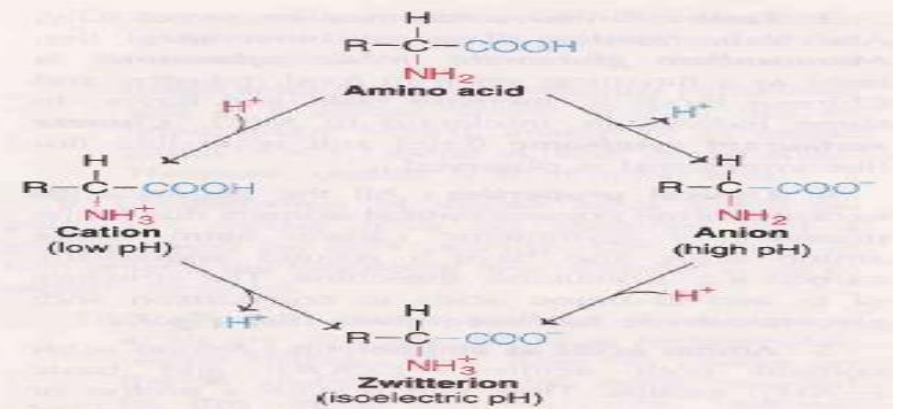
Leucine exists as cation at pH below 6 and anion at pH above 6. At the isoelectric pH (pI = 6.0), leucine is found as zwitterion.

**Isoelectric pH (symbol pI) -** is defined as the pH at which a molecule exists as a zwitterion or dipolar ion and carries no net charge. Thus, the molecule is electrically neutral.

$$pH = \frac{pK_1(COOH) + pK_2(NH_3^+)}{2}$$

$$pI = \frac{2.4 + 9.6}{2} = 6.0$$

The pI value can be calculated by taking the average pKa values corresponding to the ionizable groups.



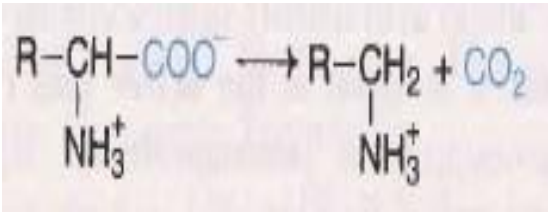
## B. Chemical Properties:-

### Reactions due to -COOH group-

1.Amino acids form salts (-COONa) with bases and esters (-COOR') with alcohols.

2.Decarboxylation:- Amino acids undergo decarboxylation to produce corresponding amines.

This reaction assumes significance in the living cells due to the formation of many biologically important amines. These include histamine, tyramine and gama-amino butyric acid (GABA) from the amino acids histidine, tyrosine and glutamate, respectively.



3. Reaction with ammonia:- The carboxyl group of dicarboxylic amino acids reacts with NH3 to form amide.

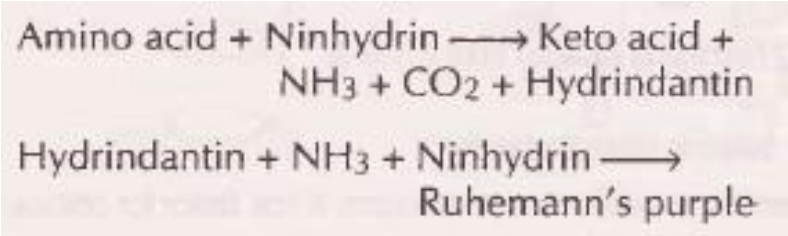
Aspartic acid + NH3 -----> Asparagine

Glutamic acid + NH3 ------> Glutamine

Reactions due to -NH2 group:-

4.The amino groups behave as bases and combine with acids (e.g. HCl) to form salts (-NH3+Cl-).

5.Reaction with ninhydrin:- The $\alpha$-amino acids react with ninhydrin (a powerful oxidizing agent, causes oxidative decarboxylation of α - amino acids) to form a purple, blue or pink colour complex (Ruhemann's purple). Ninhydrin reaction is effectively used for the quantitative determination of amino acids and proteins.



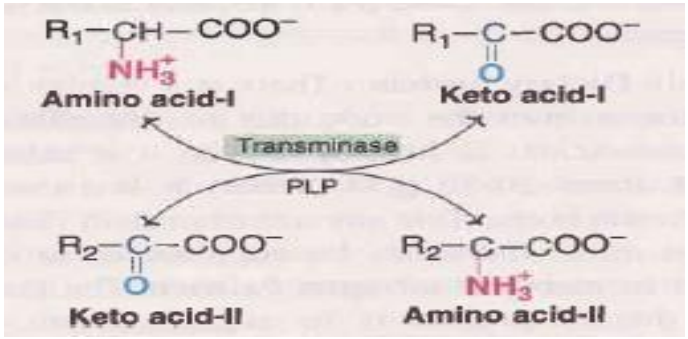

6. Others reactions are Sanger's reaction & Edman's reaction.

6.Colour reactions of amino acids:- Amino acids can be identified by specific colour reactions.

7.Transamination:- Transfer of an amino group from an amino acid to a keto acid to form a new amino acid is a very important reaction in amino acid metabolism.

8.Oxidative deamination:- Oxidative deamination is the liberation of free ammonia from the amino group of amino acids coupled with oxidation.

This takes place mostly in liver and kidney. The purpose of oxidative deamination is to provide NH3 for urea synthesis and $\alpha$-keto acids.



# NON-STANDARD AMINO ACIDS

These include the amino acid derivatives found in proteins, non-protein amino acids per forming specialized functions and the D-amino acids.

**(A). Amino acid derivatives in proteins:-** The 20 standard amino acids can be incorporated into proteins due to the presence of universal genetic code. Some of these amino acids undergo specific modification after the protein synthesis occurs. These derivatives of amino acids are very important for protein structure and functions. Examples -

Collagen- The most abundant protein in mammals contains 4-hydroxyproline and 5-hydroxylysine.

Histones- The proteins found in association with DNA.

gama-Carboxyglutamic acid - is found in certain plasma proteins involved in blood clotting.

Cystine is formed by combination of two cysteines. Cystine is also considered as derived amino acid.

**(B). D-Amino acids:-** The vast majority of amino acids isolated from animals and plants are of L-category. Certain D-amino acids are also found in the antibiotics (actinomycin-D, valinomycin, gramicidin-S). D-serine and D-aspartate are found in brain tissue.

**(C). Non-protein amino acids:-** These amino acids, although never found in proteins, perform several biologically important functions. They may be either $\alpha$-or non-$\alpha$-amino acids.

**Table 4.2 A selected list of important non-protein amino acids along with their functions**

| Amino acids | Function(s) |
|---|---|
| **I. α-Amino acids** | |
| Ornithine | Intermediates in the biosynthesis of urea. |
| Citrulline | |
| Arginosuccinic acid | |
| Thyroxine | Thyroid hormones derived from tyrosine. |
| Triiodothyronine | |
| S-Adenosylmethionine | Methyl donor in biological system. |
| Homocysteine | Intermediate in methionine metabolism. A risk factor for coronary heart diseases |
| Homoserine | Intermediate in threonine, aspartate and methionine metabolisms. |
| 3, 4-Dihydroxy phenylalanine (DOPA) | A neurotransmitter, serves as a precursor for melanin pigment. |
| Creatinine | Derived from muscle and excreted in urine |
| Ovothiol | Sulfur containing amino acid found in fertilized eggs, and acts as an antioxidant |
| Azaserine | An antibiotic |
| **II. Non-α-amino acids** | |
| β-Alanine | Component of vitamin pantothenic acid and coenzyme A |
| β-Aminoisobutyric acid | End product of pyrimidine metabolism. |
| γ-Aminobutyric acid (GABA) | A neurotransmitter produced from glutamic acid |
| δ-Aminolevulinic acid (ALA) | Intermediate in the synthesis of porphyrin (finally heme) |
| Taurine | Found in association with bile acids. |

Certain non-standard amino acids that are used as drugs-

D-Penicillamine (D-dimethylglycine)- a metabolite of penicillin, is employed in the chelation therapy of Wilson's disease.

N-Acetylcysteine- is used in cystic fibrosis, and chronic renal insuffiiciency, as it can function as an antioxidant.

Gabapentin- is used as an anticonvulsant.

# PROTEINS AND THEIR BIOLOGICAL SIGNIFICANCE

Proteins are synthesized from amino acids, which are joined together by peptide bond to form a linear chain. Functions of proteins depend on the amino acid sequence.

Most of the enzymes involved in the biochemical reactions in the body are protein in nature.

Many hormones are proteins or peptides in nature. Eg: Insulin

Proteins are seen in association with DNA molecules, where it controls the gene transcription and translation.

Proteins are involved in the transport processes. eg- Hemoglobin in erythrocytes involved in the transport of $O_2$. Some transport proteins bind with steroid hormones and transport them to the other parts of the body for action. Lipoproteins transport lipids.

Proteins have protective role in the body. Immunoglobulins and interferons are proteins that protect humans against bacterial and viral infections.

Structural proteins like collagen and elastin provide structural strength and elasticity to organs and the vascular systems.

Some proteins are used as nutrients eg: ovalbumin of egg white and casein of milk are used as nutrients.

Proteins are involved in the maintenance of osmotic pressure of plasma.

Some proteins like actin and myosin are involved in the contraction of skeletal muscles.

# CLASSIFICATION OF PROTEINS

Based on the solubility and physical properties, proteins are classified into 3 major groups, namely,

Simple proteins

Conjugated proteins

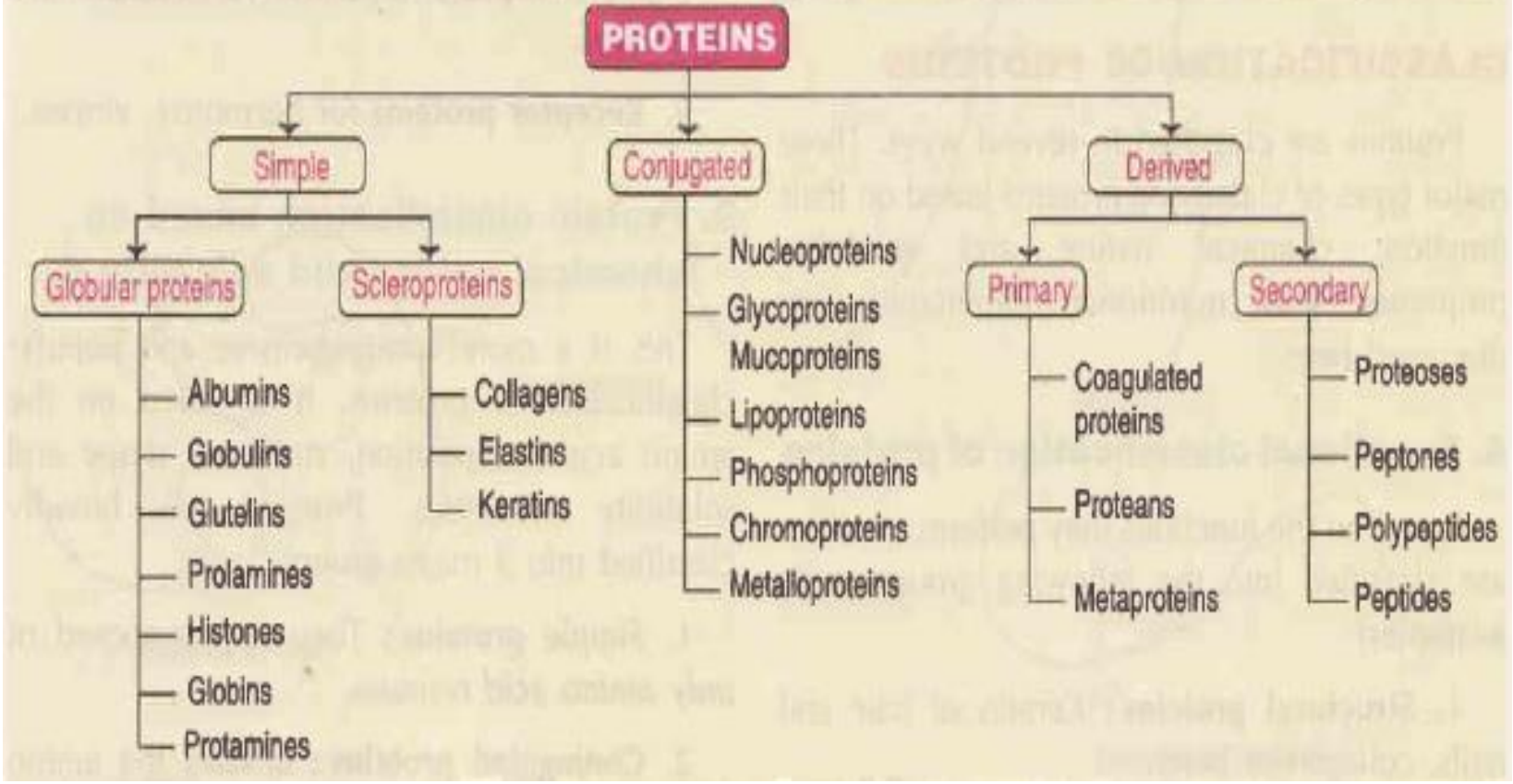
Derived proteins

Proteins can also be classified, according to their shapes

Globular protein

Fibrous protein

## Table 4.4 Summary of classification of proteins



# SIMPLE PROTEINS

These are proteins, which on complete hydrolysis yield only amino acids as an end product.

They are further subdivided, into

Albumins: These proteins are soluble in water coagulated by heat and precipitated by saturated salt solution like ammonium sulphate. Eg: serum albumin and lactalbumin.

Globulins: These proteins are insoluble in pure water, but soluble in dilute salt solutions. They may be coagulated by heat. Eg: Serum globulin, ovoglobulin.

Glutelins: These are soluble in dilute acids and alkalis and insoluble in water and neutral solvents. Eg: glutenin from wheat.

Prolamine: Soluble in 70-80 % alcohol, insoluble in water, absolute alcohol and other neutral solvents. Eg: zein of corn and glyadin of wheat.

Histones: These proteins are soluble in water and very dilute acids, insoluble in dilute ammoniun hydroxide. These proteins are not coagulated by heat. They are strongly basic in nature due to the presence of excess amounts of arginine and lysine. Eg: Histones associated with nucleic acids.

Protamines: Basic polypeptide, soluble in water or ammonium hydroxide, not coagulated by heat. Basic amino acids will predominate in their structure.Eg: protamines of sperm cells.

Albuminoids (Sclero proteins): Insoluble in all neutral solvents and in dilute acids and alkalis. These are the proteins of supportive tissue. Eg: Keratins and collagen.

# CONJUGATED PROTEINS

These are simple proteins conjugated to non-protein substances known as prosthetic group**.**
Classification is based on the nature of the prosthetic group, attached to the simple proteins.

| | |
|---|---|
| **Nucleoproteins** | **Simple proteins associated with nucleic acids (DNA and RNA). Eg: chromatin of cell.** |
| **Glycoproteins** | These are proteins having carbohydrates as prosthetic group.<br><br>Glycoproteins contain less than 4% of carbohydrates whereas mucoproteins contain more than 4% of carbohydrates. Eg: mucin in saliva. |
| **Lipoproteins** | These are proteins associated with cholesterol, phospholipids and fatty acids. |
| **Phosphoproteins** | Phosphoric acid is the prosthetic group. Eg: casein in milk, vitelline in egg yolk. |
| **Metalloproteins** | These proteins are attached to various metal ions such as, copper, cobalt, iron, manganese and zinc. E.g. ceruloplasmin containing copper, carbonic anhydrase containing zinc. |
| **Chromoproteins** | Proteins are attached to colored pigments. Eg: hemoglobin, cytochromes and flavoproteins. |

# DERIVED PROTEINS

As the name implies that these proteins are formed from simple and conjugated proteins, from the action of heat, enzymes or chemicals.

They are sub divided into,

- Primary derived proteins
- Secondary derived proteins

**Primary derived proteins**

The structure of these protein derivatives are slightly changed from original proteins. These are also called as denatured proteins. E.g. coagulated proteins. They are produced by action of alcohol and heat.

Metaproteins: they are formed by the action of acid and alkali on proteins.

**Secondary derived proteins**

These are smaller molecules produced by the hydrolysis of proteins. They are generally water-soluble and not coagulated by heat. Eg: proteases, peptones and peptides are formed by the hydrolytic cleavage of proteins.

1. Complete proteins: These proteins have all the ten essential amino acids in the required proportion by the human body to promote good growth. e.g. egg albumin, milk casein.

2.Partiatly incomplete proteins: These proteins are partially lacking one or more essential amino acids and hence can promote moderate growth. e.g. wheat and rice proteins (limiting Lys, Thr).

3.Incomplete proteins: These proteins completely lack one or more essential amino acids. Hence they do not promote growth at all. e.g. gelatin( lacks Trp), zein (lacks Trp, Lys).

## GLOBULAR AND FIBROUS PROTEINS:-

### Globular proteins-

These are proteins, in which the polypeptide chain or chains are tightly coiled in three dimensions to form globular molecules. E.g., enzymes and plasma proteins.

They are soluble in water.

### Fibrous proteins-

Fibrous proteins are those in which the polypeptide chains are either extended or coiled to form linear fibers.

They are insoluble in water.

They provide mechanical support to the cells or organism. E.g., keratin (the major component of hair and nail), collagen (component of skin, bones, teeth, blood vessel and connective tissues) and elastin (structural component of skin and blood vessels).

# Functional classification of proteins-

Based on the functions they perform, proteins are classified into the following groups (with examples)-

1.Structural proteins : Keratin of hair and nails, collagen of bone.

2. Enzymes or catalytic proteins: Hexokinase, pepsin

3. Transport proteins: Hemoglobin, serum albumin.

4. Hormonal proteins: Insulin, growth hormone.

5. Contractile proteins : Actin, myosin.

6. Storage proteins: Ovalbumin, glutelin.

7. Genetic proteins : Nucleoproteins.

8.Defense proteins: Snake venoms, immunoglobulins

9. Receptor proteins for hormones, viruses

# Structure of Proteins

Proteins are the polymers of L-$\alpha$- amino acids.

The structure of proteins is rather complex which can be divided into 4 levels of organization-

1.Primary structure:- The linear sequence of amino acids forming the backbone of proteins (polypeptides).

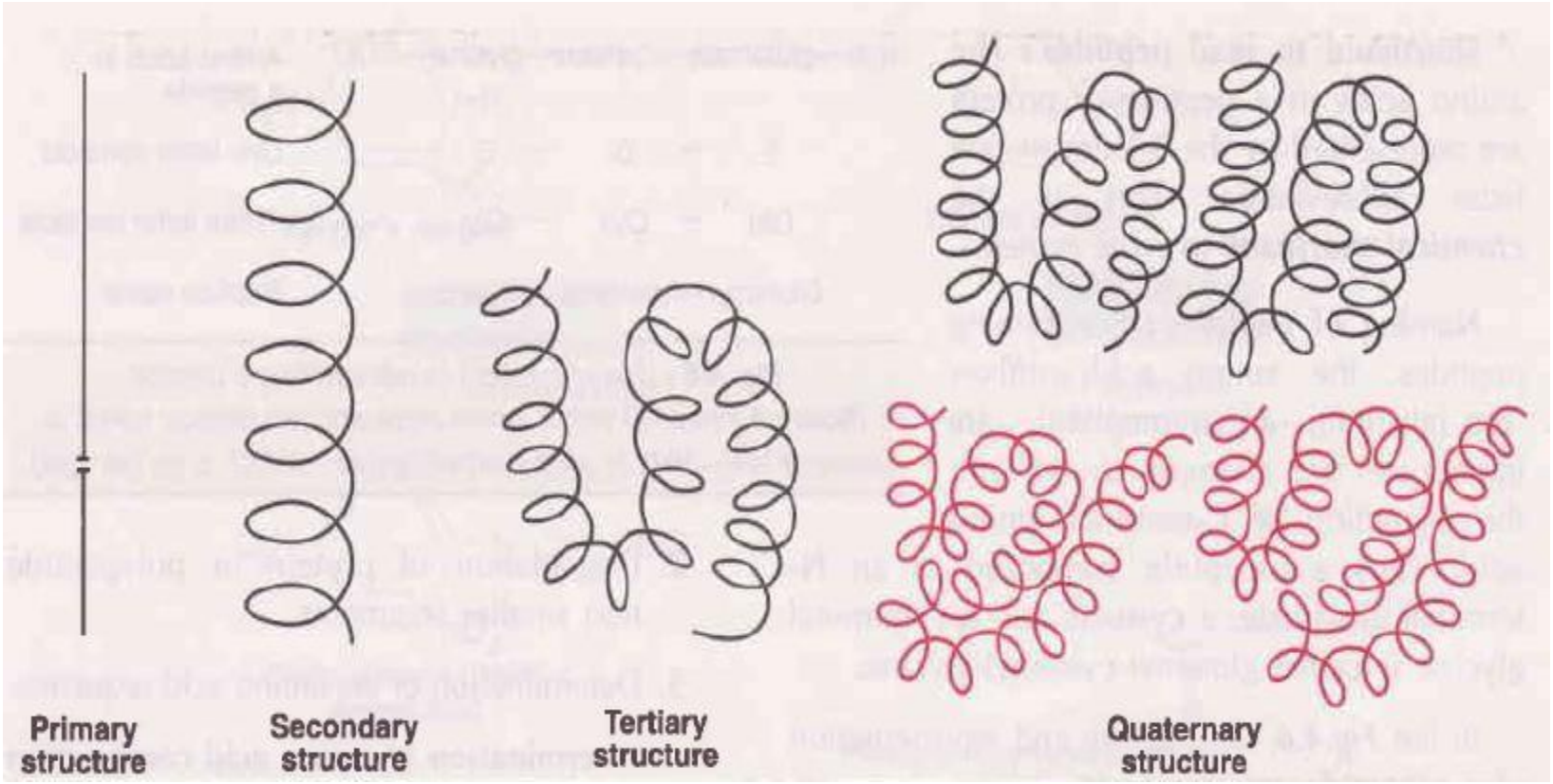
2.Secondary structure:- The spatial arrangement of protein by twisting of the polypeptide chain.

3. Tertiary structure:- The three dimensional structure of a functional protein.

4.Quaternary structure:- Some of the proteins are composed of two or more polypeptide chains referred to as subunits. The spatial arrangement of these subunits is known as quaternary structure.

The term protein is generally used for a polypeptide containing more than 50 amino acids. But some authors have been using 'polypeptide‘ even if the number of amino acids is a few hundreds. They prefer to use protein to an assembly of polypeptide chains with quaternary structure.

# Diagrammatic representation of protein structure



## PRIMARY STRUCTURE OF PROTEIN:-

Each protein has a unique sequence of amino acids which is determined by the genes contained in DNA. The primary structure of a protein is largely responsible for its function.

A vast majority of genetic diseases are due to abnormalities in the amino acid sequences of proteins i.e. changes associated with **primary structure of protein.**

The amino acid composition of a protein determines its physical and chemical properties.

**Peptide bond-**

The amino acids are held together in a protein by covalent peptide bonds or linkages. These bonds are rather strong and serve as the cementing material between the individual amino acids.

# Formation of a peptide bond

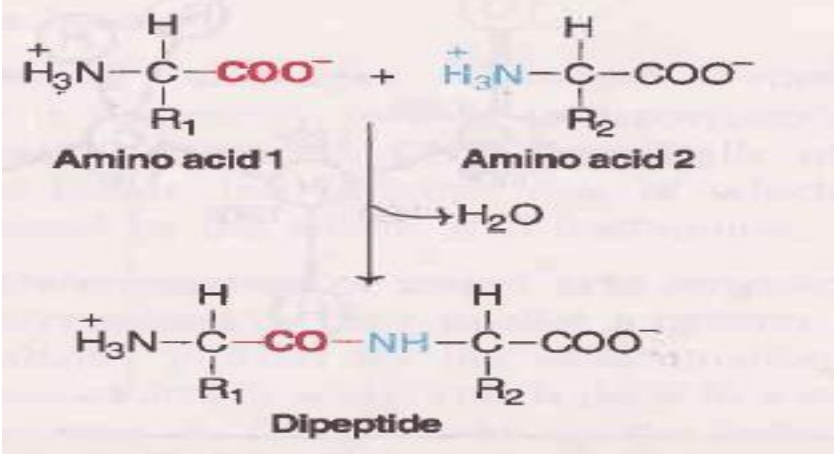
When the amino group of an amino acid combines with the carboxyl group of another amino acid, a peptide bond is formed.

Note that a dipeptide will have two amino acids and one peptide (not two) bond. Peptides containing more than 10 amino acids (decapeptide) are referred to as polypeptides.

## Characteristics of peptide bonds:

The peptide bond is rigid and planar with partial double bond in character.

lt generally exists in trans configuration. Both -**C=O and –NH groups** of peptide bonds are polar and are involved in hydrogen bond formation.




**Fomation of a peptide bond**

## Writing of peptide structures:-

Conventionally, the peptide chains are written with the free amino end (N-terminal residue) at the left, and the free carboxyl end (C-terminal residue) at the right. The amino acid sequenceis read from N-terminal end to C-terminal end.

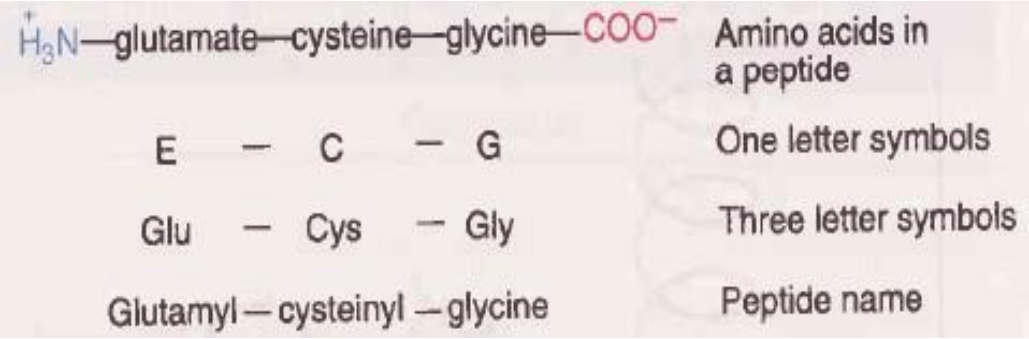
The protein biosynthesis also starts from the N-terminal amino acid.

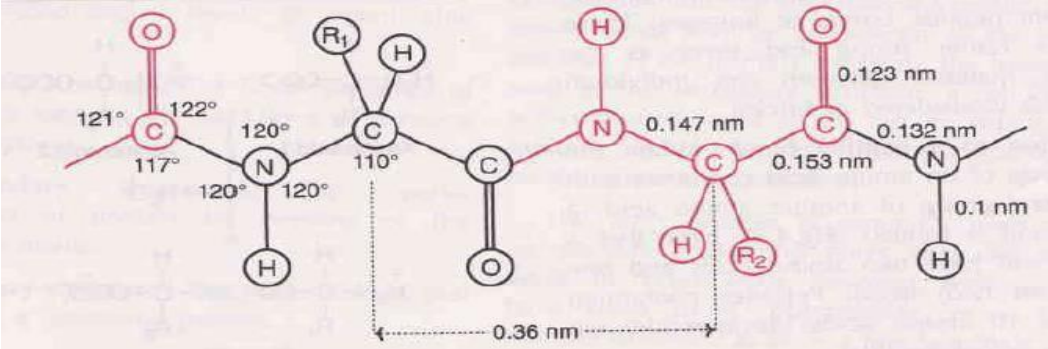
### Shorthand naming to read peptides:-

The amino acids in a peptide or protein are represented by the 3-letter or one letter abbreviation. This is the chemical shorthand to write proteins.

# Naming of peptides:

For naming peptides, the amino acid suffixes -ine (glycine), -an (tryptophan), -ate Glutamate) are changed to -yl with the exception of C terminal amino acid. Thus a tripeptide composed of an N terminal glutamate, a cysteine and a C-terminal glycine is called glutamyl-cysteinyl-glycine .



# Determnination of primary structure-

The primary structure comprises the identification of constituent amino acids with regard to their quality, quantity and sequence in a protein structure. A pure sample of a protein or a polypeptide is essential for the determination of primary structure which involves 3 stages:-

1. Determination of amino acid composition

2.Degradation of protein or polypeptide into smaller fragments.

3. Determination of the amino acid sequence

Pronase is a mixture of non-specific proteolytic enzymes that causes complete hydrolysis of proteins.

Cyanogen bromide (CNBr) is commonly used to split polypeptides into smaller fragments.

**Sanger's reagent:-**

Sanger used 1-fluoro 2 ,4-dinitrobenzene (FDNB) to determine insulin structure. FDNB specifically binds with N-terminal amino acid to form a dinitrophenyl (DNP) derivative of peptide. This on hydrolysis yields DNP amino acid (N-terminal) and free amino acids from the rest of the peptide chain.

**Edman's reagent:-**

Phenyl isothiocyanate is the Edman's reagent. lt reacts with the N-terminal amino acid of peptide to form a phenyl thiocarbamyl derivative. On treatment with mild acid, phenyl thiohydantoin (PTH)-amino acid, a cyclic compounds is liberated.

**Sequenator:-**

This is an automatic machine to determine the amino acid sequence in a polypeptide (with around 100 residues). It is based on the principle of Edman's degradation.

**Overlapping peptides-**

In the determination of primary structure of protein, several methods (enzymatic or chemical) are simultaneously applied. This results in the formation of overlapping peptides. This is due to the specific action of different agents on different sites in the polypeptide. Overlapping peptides are very useful in determining the amino acid sequence.

## SECONDARY STRUCTURE OF PROTEIN-

The conformation of polypeptide chain by twisting or folding is referred to as secondary structure. The amino acids are located close to each other in their sequence. Two types of secondary structures, $\alpha$-helix and β-sheet, are mainly identified.

Indian scientist **Ramachandran** made a significant contribution in understanding the spatial arrangement of polypeptide chains.

# $\alpha$-Helix-

$\alpha$-Helix is the most common spiral structure of protein. It has a rigid arrangement of polypeptide chain. $\alpha$-Helical structure was proposed by Pauling and Corey (1951).

1.The $\alpha$-helix is a tightly packed coiled structure with amino acid side chains extending outward from the central axis.

2.The $\alpha$-helix is stabilized by extensive hydrogen bonding. It is formed between H atom attached to peptide N, and O atom attached to peptide C. The hydrogen bonds are individually weak but collectively, they are strong enough to stabilize the helix.

3.All the peptide bonds, except the first and last in a polypeptide chain, participate in hydrogen bonding.

4.Each turn of a-helix contains 3.6 amino acids and travels a distance of 0.54 nm. The spacing of each amino acid is 0.15 nm.

5.$\alpha$-helix is a stable conformation formed spontaneously with the lowest energy.

6. The right handed $\alpha$-helix is more stable than left handed helix.

7.Certain amino acids( particularly proline) disrupt the $\alpha$-helix. Large number of acidic (Asp, Glu) or basic (Lys, Arg, His) amino acids also interfere with $\alpha$-helix structure.

**β-Pleated sheet-**

This is the second type of structure proposed by Pauling and Corey. **β**-Pleated sheets (Simply **β**-sheets) are composed of two or more segments of fully extended peptide chains. ln the **β**-sheets, the hydrogen bonds are formed between the neighbouring segments of polypeptide chain(s).

**Parallel and anti-parallel β-sheets-**

The polypeptide chains in the **β**-sheets may be arranged either in parallel (the same direction) or anti-parallel (opposite direction).

**β**-Pleated sheet may be formed either by separate polypeptide chains (H-bonds are interchain) or a single polypeptide chain folding back on to itself.

Many proteins contain **β**-pleated sheets. As such, the $\alpha$-helix and **β**-sheet are commonly found in the same protein structure.

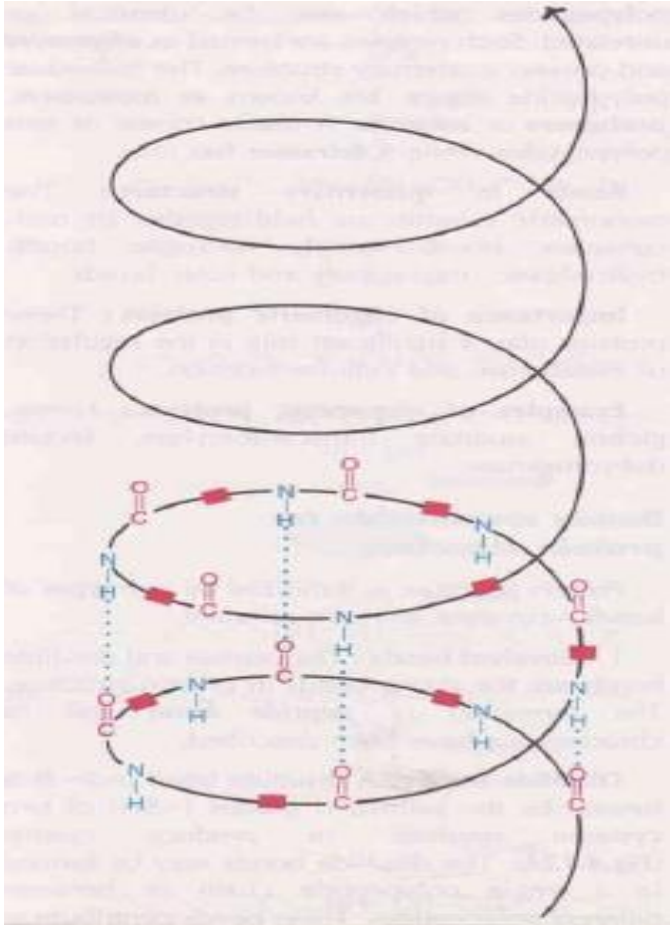


*Fig. 4.9 : Diagrammatic representation of secondary structure of protein—a right handed α-helix*

*(■ -indicate* $\begin{matrix} H \\ | \\ -C- \\ | \\ R \end{matrix}$ *groups of amino acids; dotted blue lines are hydrogen bonds;*

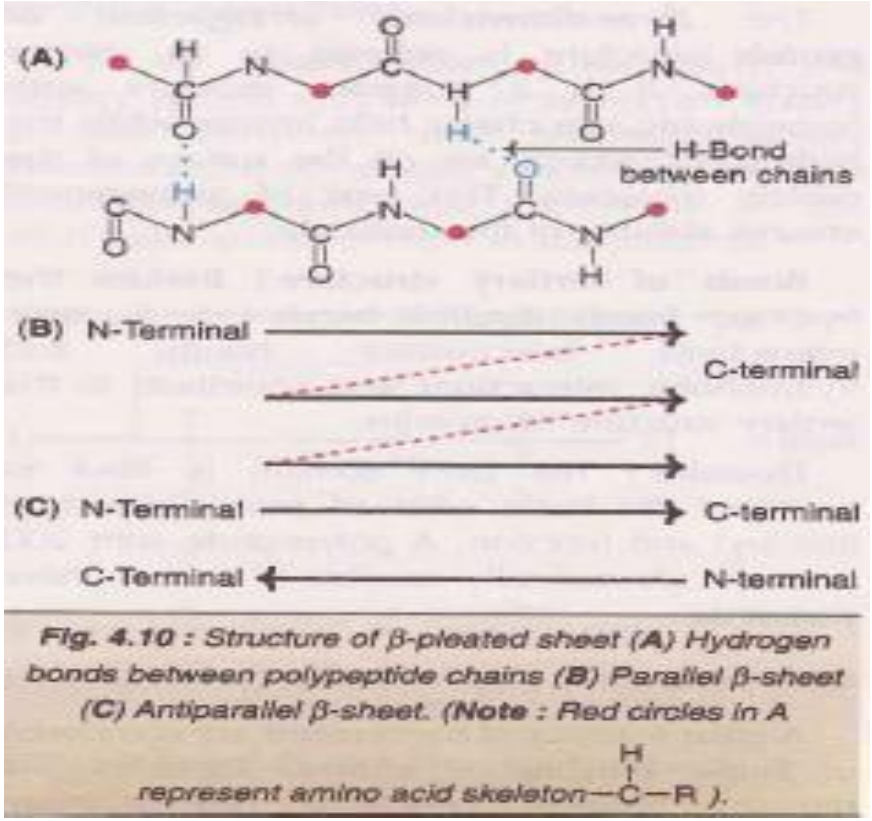



*Fig. 4.10 : Structure of β-pleated sheet (A) Hydrogen bonds between polypeptide chains (B) Parallel β-sheet (C) Antiparallel β-sheet. (**Note** : Red circles in A represent amino acid skeleton* $\begin{matrix} H \\ | \\ -C- \\ | \\ R \end{matrix}$ *).*

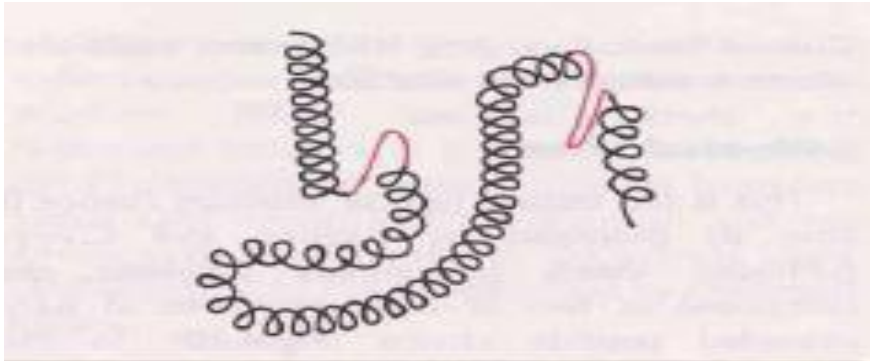


*Fig. 4.11 : Diagrammatic representation of a protein containing α-helix and β-pleated sheet (blue).*

# TERTIARY STRUCTURE OF PROTEIN-

The three-dimensional arrangement of protein structure is referred to as tertiary structure. lt is a compact structure with hydrophobic side chains held interior while the hydrophilic groups are on the surface of the protein molecule. This type of arrangement checks stability of the molecule.

## Bonds of tertiary structure:

Besides the hydrogen bonds, disulfide bonds ( -S-S), ionic interactions (electrostatic bonds) and hydrophobic interactions also contribute to the tertiary structure of proteins.

## Domains:

The term domain is used to represent the basic units of protein structure (tertiary) and function. A polypeptide with 200 amino acids normally consists of two or more domains.

## QUATERNARY STRUCTURE OF PROTEIN:-

A great majority of the proteins are composed of single polypeptide chains and some of the proteins consist of two or more polypeptides which may be identical or unrelated.

Such proteins are termed as oligomers and possess quaternary structure. The individual polypeptide chains are known as monomers, protomers or subunits. A dimer consist of two polypeptides while a tetramer has four.

**Bonds in quaternary structure-**

The monomeric subunits are held together by non-convalent bonds namely hydrogen bonds, hydrophobic interactions and ionic bonds.

# Importance of oligomeric proteins-

These proteins play a significant role in the regulation of metabolism and cellular function.

Examples- Hemoglobin, aspartate transcarbomylase lactate dehydrogenase.

**Bonds responsible for protein structure-**

Proteins tructureis stabilized by two types of bonds-covalent and non-covalent.

1. Covalent bonds:-

The peptide and disulfide bonds are the strong bonds in protein structure.

Disulfide bonds- A disulfide bond ( **-S-S-**) is formed by the sulfhydryl groups (-SH) of two cysteine residues, to produce cystine.

The disulfide bonds may be formed in a single polypeptide chain or between different polypeptides. These bonds contribute to the structural conformation and stability of proteins.

## 2. Non-covalent bonds:-

There are, mainly, four types of non covalent bonds.

(a)Hydrogen bonds- The hydrogen bonds are formed by sharing of hydrogen atoms between the nitrogen and carbonyl oxygen of different peptide bonds.

Each hydrogen bond is weak but collectively they are strong. A large number of hydrogen bonds significantly contribute to the protein structure.

(b)Hydrophobic bonds- The non-polar side chains of neutral amino acids tend to be closely associated with each other in proteins. The occurrence of hydrophobic forces is observed in aqueous environment where in the molecules are forced to stay together.

(c)Electrostatic bonds- These bonds are formed by interactions between negatively charged groups (e.g. COO-) of acidic amino acids with positively charged groups (e.g. NH3+) of basic amino acids.

(d)Van der Waals forces- These are the non-covalent associations between electrically neutral molecules.

**Structure of human insulin:-** Insulin consists of two polypeptide chains, A and B. The A chain has glycine at the N-terminal end and asparagine at the C-terminal end. The B chain has phenylalanine and alanine at the N- and C –terminal ends, respectively. Originally, insulin is synthesized as a single polypeptide preproinsulin which undergoes proteolytic processing to give proinsulin and finally insulin.

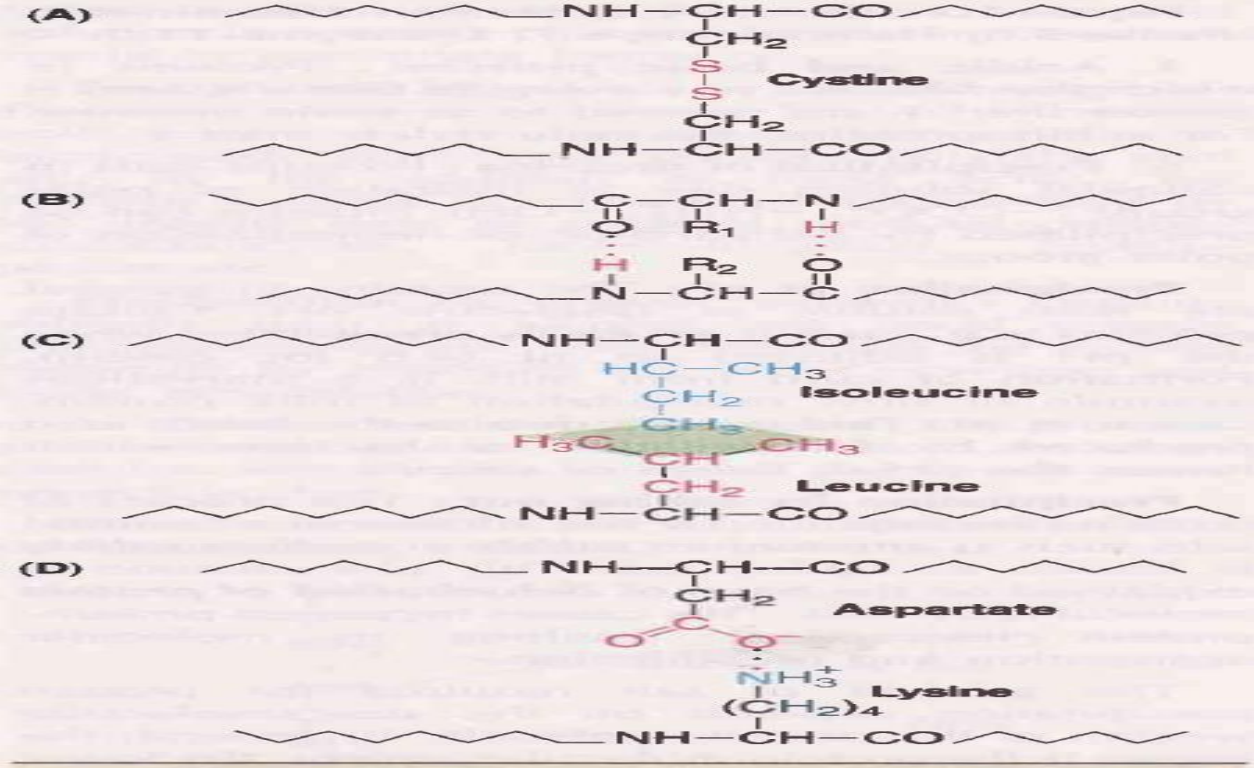



*Fig. 4.12 : Major bonds in protein structure (A) Disulfide bond (B) Hydrogen bonds (C) Hydrophic bonds (D) Electrostatic bond.*

**Representation of differents types of bonds**

# PROPERTIES OF PROTEINS

## Denaturation of proteins-

Secondary, tertiary and quaternary structure of proteins can be disrupted by chemicals like urea (6mol/L), 5M guanidinium hydrochloride, heat, high and low pH, detergents such as sodium dodecyl sulphate (1%) and sulfhydryl reagents such as mercaptoethanol.

High temperature disrupts a variety week interactions of proteins, hence, proteins lose the solubility and are then precipitated. An example is the denaturation of proteins caused by heating egg.

Sodium dodecyl sulphate alters the protein structures by interacting with the non-polar residues of proteins thereby interfering with the hydrophobic interactions.

High concentration of water-soluble organic substances such as aliphatic alcohols also interacts with hydrophobic forces.

Solute such as urea can also precipitates proteins by disrupting hydrophobic interactions.

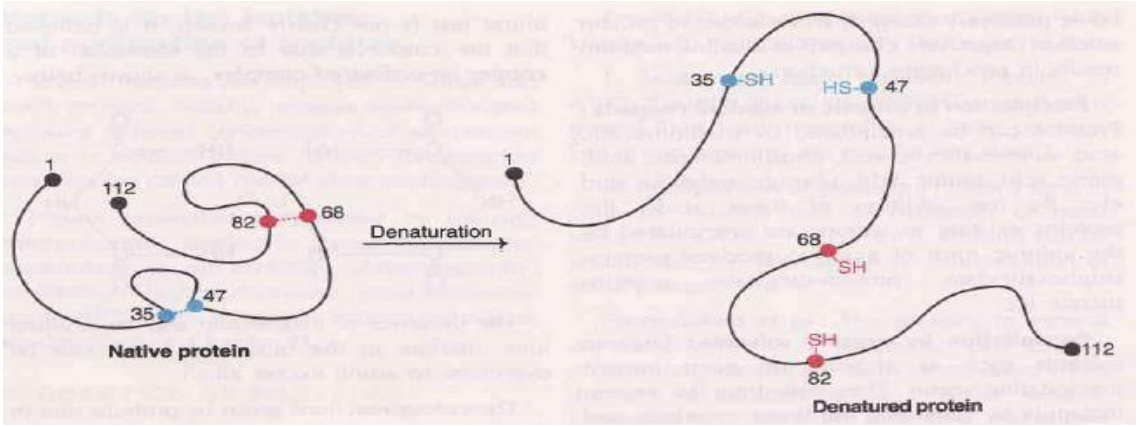
The above conditions overcome the weak forces on which polypeptide chains are folded and disrupt hydrogen bonds causing polypeptides to unfold. Such unfolding is called denaturation. It is accompanied by loss of the normal biological function (e.g. enzyme activity) of the protein.

Denatured proteins are usually not soluble in water, in part because denaturation exposes internal hydrophobic R groups. Denaturation does not break the primary structure of protein.

## Agents of denaturation-

Physical agents: Heat, violent shaking, X-ravs, UV radiation.

Chemical agents: Acids, alkalies, organic solvents (ether, alcohol), salts of heavy metals (Pb, Hg), urea, salicylate.



## Coagulation:-

The term 'coagulum' refers to a semi-solid viscous precipitate of protein. lrreversible denaturation results in coagulation.

Coagulation is optimum and requires lowest temperature at isoelectric pH. Albumins and globulins are coagulable proteins.

Heat coagulation test is commonly used to detect the presence of albumin in urine.

## Flocculation:-

lt is the process of protein precipitation at isoelectric pH. The precipitate is referred to as flocculum. Casein (milk protein) can be easily precipitated when adjusted to isoelectric pH (4.6) by dilute acetic acid.

Flocculation is reversible. On application of heat, flocculum can be converted into an irreversible mass, coagulum.

**Renaturation of proteins-**

If a denatured protein returns to its native state, after the removal of denaturing agent, the process is called as renaturation.

**Amphoteric nature-**

Due to the presence of –NH2 and –COOH group proteins are amphoteric.

Charges on proteins are mainly due to the presence of the side chains of the amino acid residues, the N – terminal amino group and the C – terminal carboxyl group. The other carboxyl and amino group of each amino acids are involved in the peptide bond formation. Therefore, pI of proteins depends on the pH of the solution and pKa of amino acids forming the proteins.

In solutions with pH values above the isoelectric point, the protein will have a net negative charge and at lesser pH values it will be positively charged.

**Precipitation of proteins-**

At isoelectric pH the proteins exist as zwitter ion (net charge is zero) at this pH the proteins are easily precipitated and it also shows no migration under an electric field.

Proteins can be precipitated by trichloroacetic acid, sulphosalicylic acid, phosphotungstic acid, picric acid, tannic acid, phosphomolybdic acid or by the addition of heavy metals like lead, silver and mercury.

**Salting out-**

Solubility of proteins is decreased by increasing the concentration of salt such as ammonium sulfate or sodium sulfate. The salts gradually dehydrate proteins by binding to water. The dehydrated proteins aggregate and precipitates.

**Salting in-**

Solubility of some proteins is increased by increasing the concentration of neutral salts. This process is known as salting in.

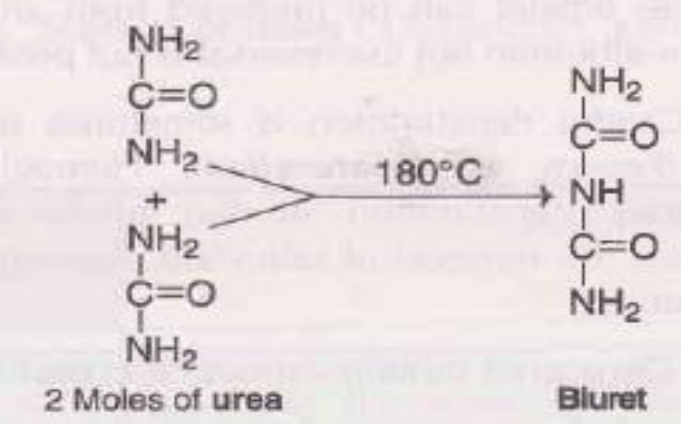
# Colour reactions of proteins:-

The proteins give several colour reactions which are often useful to identify the nature of the amino acids present in them.

Biuret reaction: Biuret is a compound formed by heating urea to 180°C.

When biuret is treated with dilute copper sulfate in alkaline medium, a purple colour is obtained. Single amino acids and dipeptides do not give the biuret reaction and the colour reaction is due to the formation of a copper co-ordinated complex.

This is the basis of biuret test widely used for identification of proteins and peptides.



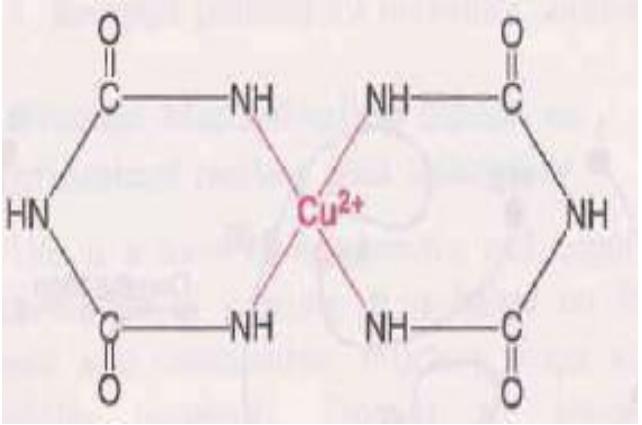
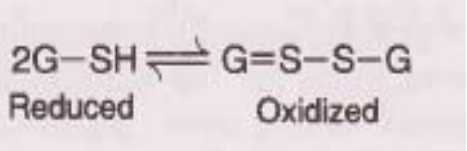



**TABLE 4.3 Colour reactions of proteins/amino acids**

| Reaction | Specific group or amino acid |
|---|---|
| 1. Biuret reaction | Two peptide linkages |
| 2. Ninhydrin reaction | α-Amino acids |
| 3. Xanthoproteic reaction | Benzene ring of aromatic amino acids (Phe, Tyr, Trp) |
| 4. Millions reaction | Phenolic group (Tyr) |
| 5. Hopkins-Cole reaction | Indole ring (Trp) |
| 6. Sakaguchi reaction | Guanidino group (Arg) |
| 7. Nitroprusside reaction | Sulfhydryl groups (Cys) |
| 8. Sulfur test | Sulfhydryl groups (Cys) |
| 9. Pauly's test | Imidazole ring (His) |
| 10. Folin–Coicalteau's test | Phenolic groups (Tyr) |

# BIOLOGICALLY IMPORTANT PEPTIDES

1. Glutathione:- lt is a tripeptide composed of 3- amino acids. Chemically, glutathione is gama-glutamyl-cysteinyl-glycine. It is widely distributed in nature and exists in reduced or oxidized states.




<u>Function-</u>

Glutathione serves as a coenzyme for certain enzymes e.g. prostaglandin PGE2 synthetase, glyoxylase.

lt prevents the oxidation of sulfhydryl (-SH) groups of several proteins to disulfide (-S-S-) groups. This is essential for the protein function, including that of enzymes.

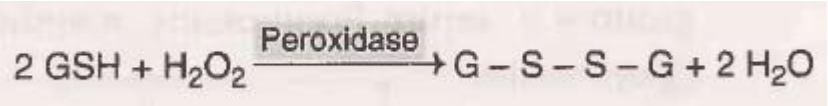
Glutathione (reduced) performs specialized functions in erythrocytes

(i) lt maintains RBC membrane structure and integrity.

(ii)lt protects hemoglobin from getting oxidized by agents such as H2O2.

Glutathione is involved in the detoxication process. The toxic substances (organophosphates, nitro compounds) are converted to mercapturic acids.

Toxic amounts of peroxides and free radicals produced in the cells are scavanged by glutathionep eroxidase ( a selenium containing enzyme).



2. Thyrotropin releasing hormone (TRH) :

It is a tripeptide secreted by hypothalamus. TRH stlmulates pituitary gland to release thyrotropic hormone.

3. Oxytocin:

lt is a hormone secreted by posterior pituitary gland and contains 9 amino acids (nonapeptide). Oxytocin causes contraction of uterus.

4. Vasopressin (antidiuretic hormone, ADH)-

ADH is a nonapeptide produced by posterior pituitary gland. It stimulates kidneys to retain water and thus increases the blood pressure.

5. Angiotensins :-

A ngiotensinI is a decapeptide (10 amino acids) which isconverted to angiotensin ll (8 amino acids). The later has more hypertensive effect. Angiotensin ll also stimulates the release of aldosterone from adrenal gland.

6. Methionine enkephalin-

lt is a pentapeptide found in the brain and has opiate like function. lt inhibits the sense of a pain.

7. Bradykinin and kallidin-

They are nonaand decapeptides, respectively. Both of them act as powerful vasodilators. They are produced from plasma proteins by snake venom enzymes.

8. Peptide antibiotics-

Antibiotics such as gramicidin, bacitracin, tyrocidin and actinomycin are peptide in nature.

9. Aspartame-

lt is a dipeptide (aspartyl-phenylalanine methyl ester), produced by a combination of aspartic acid and phenylalanine. Aspartame is about 200 times sweeter than sucrose, and is used as a low-calorie artificial sweetner in soft drink industry.

10. Gastrointestinal hormones-

Gastrin, secretin etc. are the gastrointestinal peptides which serve as hormones.

# THANKS