

# CHAPTER 4

## Chemistry of Proteins

- Introduction
- General Nature of Amino Acids
- Classification of Amino Acids
- Modified or Nonstandard Amino Acids
- Properties of Amino Acids
- Biologically Important Peptides

- Definition, Classification and Functions of Proteins
- Structure of Proteins
- Properties of Proteins
- Denaturation of Proteins
- Summary
- Exercise

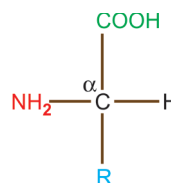
### INTRODUCTION

Proteins are the most abundant macromolecules in living cells. The term 'protein' was first used by Berzelius in 1838 and was derived from the Greek word "**protos**" which means **primary** or **holding first place**. As the name indicates, protein is the most important of cell constituents. They are responsible for almost every function that occurs in the body.

Proteins are linear chains of amino acids that are linked together by covalent, **peptide bonds**. Each protein has specific and unique sequence of amino acids that defines both its three-dimensional structure and its biologic function.

### GENERAL NATURE OF AMINO ACIDS

- There are approximately 300 amino acids present in various animal, plant and microbial systems, but only **20 amino acids are involved in the formation of proteins**.
- All the 20 amino acids found in proteins (**Table 4.1**) have a carboxyl group (-COOH) and an amino acid group (-NH<sub>2</sub>) bound to the same carbon atom called the **α-carbon** (**Figure 4.1**).
- Amino acids differ from each other in their **side chains** or **R-groups**, attached to the α-carbon.
- The 20 amino acids of proteins are often referred to as the **standard** or **primary** or **normal amino acids**.



**Figure 4.1:** General structure of α-amino acid found in protein

- The standard amino acids have been assigned three letters abbreviations and one letter symbol, e.g. amino acid **glycine** has abbreviated name **Gly** and symbol letter **G**.
- All the amino acids found in proteins are exclusively of the **L-configuration**.

### CLASSIFICATION OF AMINO ACIDS

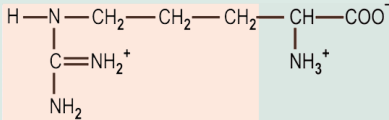
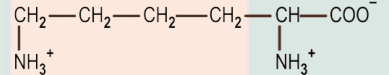
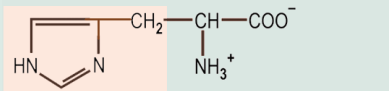
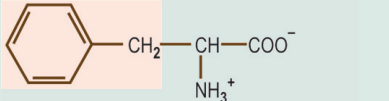
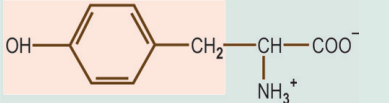
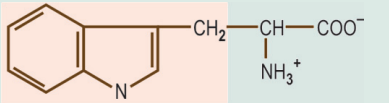
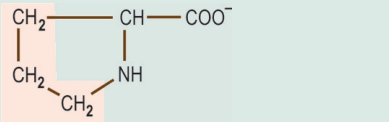
There are five ways of classifying amino acids depending on the:

1. Chemical nature of the amino acid in the solution
2. Structure of the side chain of the amino acids
3. Nutritional requirement of amino acids
4. Metabolic product of amino acids
5. Nature or polarity of the side chain of the amino acids.

Table 4.1: The 20, L- $\alpha$ -amino acids (standard amino acids) found in proteins

Name	Symbol	Structural formula
<b>Aliphatic side chain</b>		
Glycine	Gly (G)	$\begin{array}{c} \text{H}-\text{CH}-\text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$
Alanine	Ala (A)	$\begin{array}{c} \text{CH}_3-\text{CH}-\text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$
Valine	Val (V)	$\begin{array}{c} \text{CH}_3 \\   \\ \text{CH}-\text{CH}-\text{COO}^- \\   \\ \text{CH}_3 \quad \text{NH}_3^+ \end{array}$
Leucine	Leu (L)	$\begin{array}{c} \text{H}_3\text{C} \\   \\ \text{CH}-\text{CH}_2-\text{CH}-\text{COO}^- \\   \\ \text{H}_3\text{C} \quad \text{NH}_3^+ \end{array}$
Isoleucine	Ile (I)	$\begin{array}{c} \text{CH}_3 \\   \\ \text{CH}_2-\text{CH}-\text{CH}-\text{COO}^- \\   \quad   \\ \text{CH}_3 \quad \text{NH}_3^+ \end{array}$
<b>Hydroxylic (OH) group containing side chains</b>		
Serine	Ser (S)	$\begin{array}{c} \text{CH}_2-\text{CH}-\text{COO}^- \\   \quad   \\ \text{OH} \quad \text{NH}_3^+ \end{array}$
Threonine	Thr (T)	$\begin{array}{c} \text{CH}_3-\text{CH}-\text{CH}-\text{COO}^- \\   \quad   \\ \text{OH} \quad \text{NH}_3^+ \end{array}$
Tyrosine	Tyr (Y)	See aromatic group containing side chain amino acids
<b>Sulfur containing side chains</b>		
Cysteine	Cys (C)	$\begin{array}{c} \text{CH}_2-\text{CH}-\text{COO}^- \\   \quad   \\ \text{SH} \quad \text{NH}_3^+ \end{array}$
Methionine	Met (M)	$\begin{array}{c} \text{CH}_2-\text{CH}_2-\text{CH}-\text{COO}^- \\   \quad   \\ \text{S}-\text{CH}_3 \quad \text{NH}_3^+ \end{array}$
<b>Side chains containing acidic groups (–COOH) and their amides</b>		
Aspartic acid	Asp (D)	$\begin{array}{c} \text{COO}^- - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$
Asparagine	Asn (N)	$\begin{array}{c} \text{H}_2\text{N} - \text{C} - \text{CH}_2 - \text{CH} - \text{COO}^- \\    \quad   \\ \text{O} \quad \text{NH}_3^+ \end{array}$
Glutamic acid	Glu (E)	$\begin{array}{c} \text{OOC}^- - \text{CH}_2 - \text{CH}_2 - \text{CH} - \text{COO}^- \\   \\ \text{NH}_3^+ \end{array}$
Glutamine	Gln (Q)	$\begin{array}{c} \text{NH}_2 - \text{C} - \text{CH}_2 - \text{CH}_2 - \text{CH} - \text{COO}^- \\    \quad   \\ \text{O} \quad \text{NH}_3^+ \end{array}$

Contd...

Table 4:1 (contd...)			
Name	Symbol	Structural formula	
Basic groups containing side chains			
Arginine	Arg (R)		
Lysine	Lys (K)		
Histidine	His (H)		
Aromatic group containing side chains			
Histidine	His (H)	See above	
Phenylalanine	Phe (F)		
Tyrosine	Tyr (Y)		
Tryptophan	Trp (W)		
Imino acids			
Proline	Pro (P)		

### Classification Based on Chemical Nature of the Amino Acid in Solution

According to this type of classification, amino acids are classified as follows:

- Neutral amino acids
- Acidic amino acids
- Basic amino acids.

#### Neutral amino acids

The amino acids which are neutral in solution and are monoamino-monocarboxylic acids (i.e. having one amino group and one carboxylic group), e.g.

Glycine	Serine	Phenylalanine
Alanine	Threonine	Tyrosine
Valine	Cysteine	Tryptophan
Leucine	Methionine	Asparagine
Isoleucine	Proline	Glutamine

#### Acidic amino acid

These are acidic in solution and are monoamino dicarboxylic acids, e.g.

- Aspartic acid
- Glutamic acid.

**Basic amino acid**

These are basic in solution and are diamino-monocarboxylic acids, e.g.

- Lysine
- Arginine
- Histidine.

**Classification Based on Chemical Structure of Side Chain of the Amino Acid**

According to this type of classification, amino acids are classified as:

1. Aliphatic amino acids
2. Hydroxy amino acids
3. Sulfur containing amino acids
4. Dicarboxylic acid and their amides
5. Diamino acids
6. Aromatic amino acids
7. Imino acids or heterocyclic amino acids.

**Aliphatic amino acids**

Amino acids having aliphatic side chain, e.g.

- Glycine
- Alanine
- Valine
- Leucine
- Isoleucine.

**Hydroxy amino acids**

Amino acids having hydroxy group in the side chain, e.g.

- Threonine
- Serine
- Tyrosine.

**Sulfur containing amino acids**

Amino acids having sulfur in the side chain, e.g.

- Cysteine
- Methionine.

**Dicarboxylic acid and their amides**

Amino acids having carboxylic group in their side chain, e.g.

- Glutamic acid
- Glutamine (amide of glutamic acid)
- Aspartic acid
- Asparagine (amide of aspartic acid).

**Diamino acids**

Amino acids having amino group ( $-\text{NH}_2$ ) in the side chain, e.g.

- Lysine
- Arginine
- Histidine.

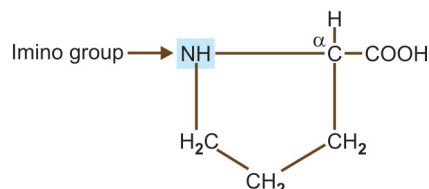
**Aromatic amino acids**

Amino acids containing aromatic ring in the side chain, e.g.

- Phenylalanine
- Tyrosine
- Tryptophan.

**Imino acids or heterocyclic amino acids**

- One of the 20 amino acids, **proline** is an **imino** ( $-\text{NH}$ ) acid not an amino ( $-\text{NH}_2$ ) acid as are other 19. The side chains of proline and its  $\alpha$ -amino group form a ring structure and thus proline differs from other amino acids, in that it contains an imino group, rather than an amino group.



Structure of proline

**Nutritional Classification of Amino Acids**

On the basis of nutritional requirement, amino acids are classified into two groups:

- i. Essential or indispensable amino acids
- ii. Nonessential or dispensable amino acids.

**Essential amino acids (also refer Chapter 11)**

Essential amino acids cannot be synthesized by the body and must, therefore, be essentially supplied through the diet. Ten amino acids, essential for humans include:

- |                 |              |
|-----------------|--------------|
| • Phenylalanine | • Methionine |
| • Valine        | • Histidine  |
| • Threonine     | • Arginine   |
| • Tryptophan    | • Lysine     |
| • Isoleucine    | • Leucine.   |

The mnemonics often used by students are **PVT. TIM. HALL** or **L.VITTHAL (MP)**.

Among the ten essential amino acids; **arginine** and **histidine** are known as **semi-essential** amino acids since these amino acids are synthesized partially in human body. Arginine and histidine become essential in diet during periods of rapid growth as in childhood and pregnancy.

A deficiency of an essential amino acid impairs protein synthesis and leads to a negative nitrogen balance (nitrogen excretion exceeds nitrogen intake).

**Nonessential amino acids**

Nonessential amino acids can be synthesized in human body and are not required in diet, e.g.

- Glycine
- Proline
- Serine
- Glutamic acid
- Glutamine
- Alanine
- Tyrosine
- Cysteine
- Aspartic acid
- Asparagine.

### Metabolic Classification of Amino Acids

On the basis of their catabolic end products, the twenty standard amino acids are divided in three groups (Table 4.2).

- Glucogenic amino acids:** Those which can be converted into glucose. Fourteen out of the twenty standard amino acids are glucogenic amino acids (Table 4.2).
- Ketogenic amino acids:** Those which can be converted to ketone bodies. Two amino acids **leucine** and **lysine** are exclusively ketogenic.
- Both glucogenic and ketogenic:** Those which can be converted to both glucose and ketone bodies. Four amino acids **isoleucine**, **phenylalanine**, **tryptophan** and **tyrosine** are glucogenic and ketogenic.

Table 4.2 : Metabolic classification of amino acids

Glucogenic	Ketogenic	Both ketogenic and glucogenic
Glycine, alanine, serine, cysteine, aspartic acid, asparagine, glutamic acid, glutamine, proline, histidine, arginine, methionine, threonine, valine	Leucine Lysine	Isoleucine Phenylalanine Tyrosine Tryptophan

### Classification Based on Nature or Polarity of Side Chain of Amino Acid

According to this type of classification, amino acids are classified into two major classes (Figure 4.2):

- Hydrophilic or polar amino acids
- Hydrophobic or nonpolar amino acids.

### Importance of Amino Acids

- Formation of proteins:** Amino acids are joined to each other by peptide bonds to form proteins and peptides.
- Formation of glucose:** Glucogenic amino acids are converted to glucose in the body.
- Enzyme activity:** The thiol (-SH) group of cysteine has an important role in certain enzyme activity.
- Transport and storage form of ammonia:** Amino acid glutamine plays an important role in transport and storage of amino nitrogen in the form of ammonia.
- As a buffer:** Both free amino acids and some amino acids present in protein can potentially act as buffer, e.g. histidine can serve as the best buffer at physiological pH.
- Detoxification reactions:** Glycine, cysteine and methionine are involved in the detoxification of toxic substances.
- Formation of biologically important compounds:** Specific amino acids can give rise to specific biologically important compounds in the body (Table 4.3).

### MODIFIED OR NONSTANDARD AMINO ACIDS

In addition to the standard amino acids, a small number of modified amino acids are found in proteins. These amino acids are formed by the modification of one of the

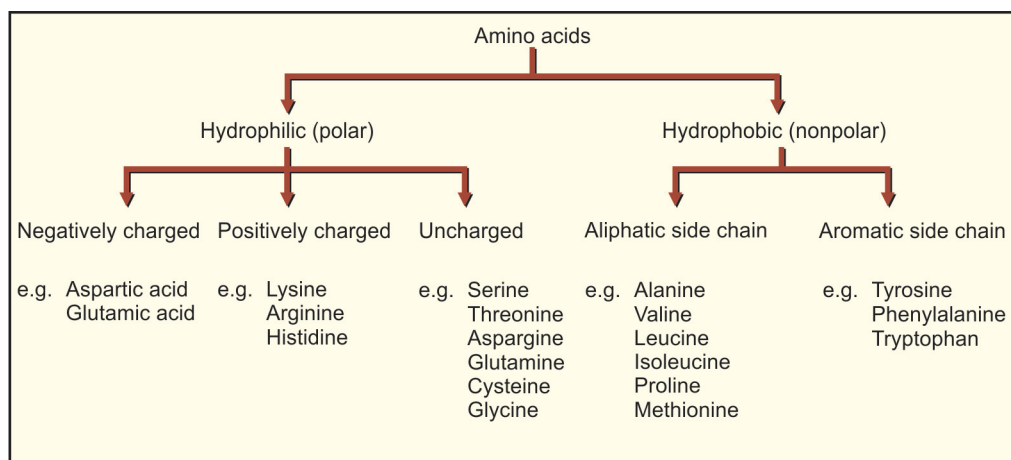


Figure 4.2: Classification of amino acids based on polarity

<i>Amino acid</i>	<i>Biologically important compound</i>
Tyrosine	Hormone, e.g. adrenaline and thyroxine, skin pigment, e.g. melanin
Tryptophan	Vitamin, e.g. niacin
Glycine, arginine and methionine	Creatine
Glycine and cysteine	Bile salts
Glycine	Heme
Aspartic acid and glutamic acid	Pyrimidine bases
Glycine, aspartic acid and glutamine	Purine bases
β-alanine	Coenzyme-A

## Cystine

$$\begin{array}{c} \text{NH}_2 \\ | \\ \text{CH}_2-\text{CH}-\text{COOH} \\ | \\ \text{SH} \\ + \\ \text{SH} \\ | \\ \text{CH}_2-\text{CH}-\text{COOH} \\ | \\ \text{NH}_2 \\ \text{Cysteine} \end{array} \longrightarrow \begin{array}{c} \text{NH}_2 \\ | \\ \text{CH}_2-\text{CH}-\text{COOH} \\ | \\ \text{S} \\ | \\ \text{S} \\ | \\ \text{CH}_2-\text{CH}-\text{COOH} \\ | \\ \text{NH}_2 \\ \text{Cystine} \end{array} + 2\text{H}^+ + 2\text{e}^-$$

### Hydroxyproline and Hydroxylysine

### *Desmosine and Isodesmosine*

### Gamma Carboxyglutamate

## Selenocysteine

- $$\begin{array}{c} \text{H} \\ | \\ \text{NH}_2 - \text{C} - \text{CH}_2 - \text{SeH} \\ | \\ \text{COOH} \end{array}$$
- Selenocysteine

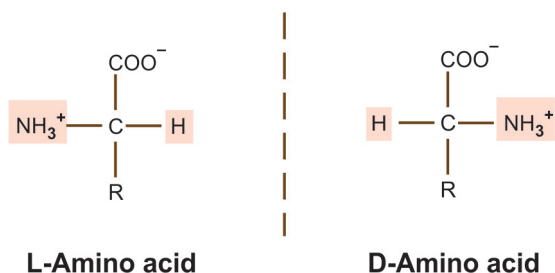
## Physical Properties

## Optical Properties

## ***Ionization of Amino Acids***

- Acid base behavior
- Amphoteric properties (zwitter ion formation)
- Buffering activity.

- The acid base properties of amino acids depend on the **amino** and **carboxyl** groups attached to the  $\alpha$ -carbon.
- The **carboxyl** ( $-\text{COOH}$ ) group of an amino acid can donate proton ( $\text{H}^+$ ) and behave as an acid, forming a negatively charged **anion**.
- Amino group ( $-\text{NH}_2$ ) of an amino acid can accept the proton ( $\text{H}^+$ ) which behave as a base, forming positively charged **cation**.



**Figure 4.4:** D and L forms of amino acids

Thus, amino acids in aqueous solution are ionized and act as acids and bases (**Figure 4.5**).

#### Amphoteric properties of amino acids and formation of zwitter ion at isoelectric pH

Substances having a two-way property are called *amphoteric* or *ampholytes* (Greek word *ampho* means *both*). As amino acids have both acidic and basic groups, they can donate a proton or accept a proton, hence amino acids are regarded as ampholytes.

#### Zwitter ion (Dipolar molecule)

- Monoamino monocarboxylic acids exist in aqueous solution as *dipolar molecule* or *zwitter ions*, which means that they have both positive and negative charges on the same amino acids.

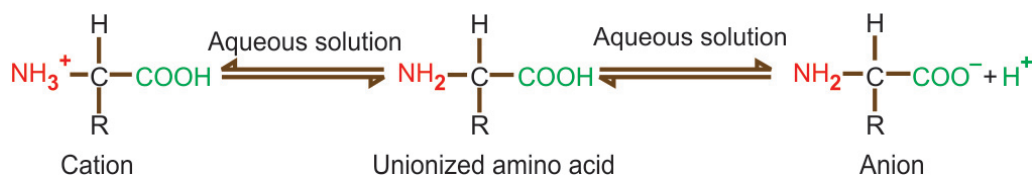
- The  $\alpha\text{-COOH}$  group is ionized and becomes negatively charged anion ( $\text{COO}^-$ ) and
- The  $\alpha\text{-NH}_2$  group is protonated to form a positively charged cation ( $\text{NH}_3^+$ ).
- Thus, the overall molecule is electrically neutral. Thus, the molecular species which contain an equal number of ionizable group of opposite charge and as a result bear no net charge, are called *zwitter ions*.
- The net charge of an amino acid depends upon the pH of the medium.
  - At acidic pH, amino acid is positively charged because ionized  $\text{COO}^-$  group accepts a proton and becomes uncharged ( $\text{COOH}$ ), so that the overall charge on the molecule is positive.
  - While at alkaline pH it is negatively charged as the  $\text{NH}_3^+$  group loses its proton and becomes uncharged; thus the overall charge on the molecule is negative (**Figure 4.6**).

#### Isoelectric pH (PI)

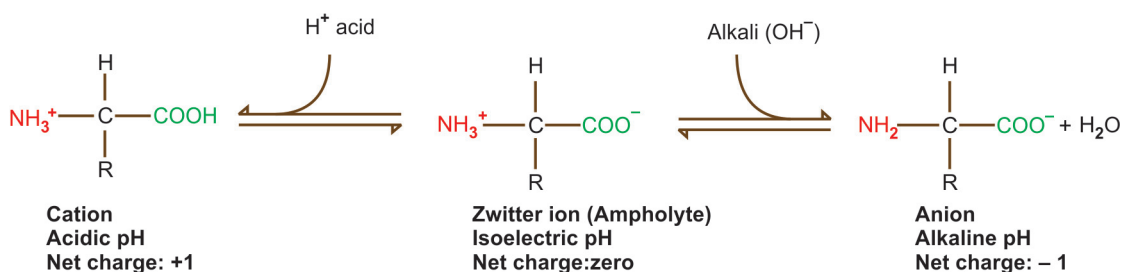
The pH at which amino acid bears no net charge (zwitter ion) and therefore does not move in an electric field is called *isoelectric pH (PI)*.

#### Buffering action of amino acid

- Amino acid can act as weak acid or weak base. In addition, each of the acidic and basic amino acids



**Figure 4.5:** Ionization of amino acid



**Figure 4.6:** Ionic forms of amino acid in acidic, basic and isoelectric pH



contains an ionizable group in its side chain. Thus, both free amino acids and some amino acids present in proteins can potentially act as buffers.

- Among the 20 standard amino acids, histidine serves as the best buffer at physiological pH. Because the side chain of histidine has a pKa of 6.0 and can serve as the best buffer at physiological pH. All other amino acids have pKa value, too far away from pH 7 to be an effective physiological buffer. *Maximum buffer capacity occurs at pH equal to the pKa.*

### BIOLOGICALLY IMPORTANT PEPTIDES

- Peptides are chains (polymer) of amino acids. Two amino acid molecules can be covalently joined through a *peptide bond*, to yield *dipeptide*. Peptide linkage is formed by the removal of a molecule of water from the  $\alpha$ -carboxyl group of one amino acid and the  $\alpha$ -amino group of another (Figure 4.7).
- When many amino acids are joined, the product is called *polypeptide*. Proteins are polypeptides with thousands of amino acids.
- There are many naturally occurring small polypeptides some of which have important biological activities and are called **biologically important peptides**. A few of them are given below.

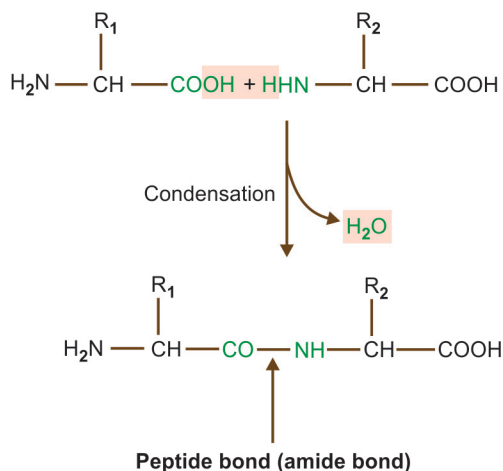


Figure 4.7: Formation of peptide bond

### Glutathione (GSH)

- GSH is a tripeptide ( $\gamma$  glutamyl – cysteinyl – glycine) containing glutamate, cysteine and glycine.
- Glutathione is found in all mammalian cells *except the neurons*.

- Glutathione may exist as the reduced (GSH) or oxidized (G-S-S-G) form (Figure 4.8) and can thus play a role in some oxidation–reduction reactions.
- In oxidized form, two molecules of glutathione are linked by disulfide bond.
- The sulfhydryl (-SH-) is the functional group primarily responsible for the properties of glutathione.

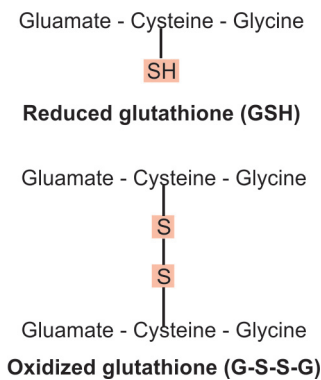
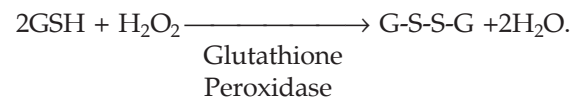


Figure 4.8: Reduced and oxidized glutathione

### Functions of glutathione

- The reduced form of glutathione with a free sulfhydryl (-SH-) group serves as a **redox buffer** regulating the redox state of the cell.
- It helps in keeping the enzymes in an active state by preventing the oxidation of sulfhydryl (-SH-) group of enzyme to disulfide (-S-S-) group.
- Reduced glutathione is essential for maintaining the normal structure of red blood cells and for keeping hemoglobin in the ferrous state. Cells with lowered level of reduced glutathione are more susceptible to hemolysis.
- Glutathione plays a key role in detoxification by reducing  $\text{H}_2\text{O}_2$ —the harmful byproduct of metabolism.



- Glutathione is involved in transport of amino acids across the cell membrane of the kidney and intestine.

### Thyrotropin Releasing Hormone (TRH)

TRH is a hypothalamic hormone of three amino acid residues. It stimulates the release of hormone thyrotropin, from the anterior pituitary gland.



**Oxytocin**

This is a 9-amino acid residue hormone secreted by posterior pituitary and stimulates uterine contractions.

**Vasopressin**

This is a 9-amino acid residue hormone secreted by posterior pituitary and it increases blood pressure and has an antidiuretic action.

**Gastrin**

This is a local hormone produced by stomach. It stimulates the production of gastric juice.

**Angiotensin**

Angiotensin II is a vasoconstrictor and elevates the arterial pressure and also promotes the synthesis of a steroid hormone called aldosterone that promotes sodium retention.

**Bradykinin**

It contains, 9-amino acid residue. It is a powerful vasodilator and causes contraction of smooth muscle and is mainly responsible for causing intense peripheral and visceral pain by stimulating the pain receptors.

**Insulin**

A pancreatic hormone contains two polypeptide chains: one having 30-amino acid residues and the other 21. Insulin regulates the glucose concentration in blood.

**Glucagon**

This is a pancreatic hormone of 29-residues that opposes the action of insulin.

### DEFINITION, CLASSIFICATION AND FUNCTIONS OF PROTEINS

**Definition**

Proteins are linear chains of amino acids that are linked together by covalent, *peptide bonds*. Each protein has specific and unique sequence of amino acids that defines both its three-dimensional structure and its biologic function.

**Classification of Proteins**

Proteins have been classified in several ways. They are most conveniently classified on the basis of their:

1. Function
2. Physical and chemical properties.

**Classification of Proteins Based on Functions**

In a functional classification, they are grouped according to their biological role. Some functions that proteins serve and examples of specific functional proteins are as follows:

**Catalytic Proteins or Enzymes**

These proteins act as enzymes, e.g.

- Glucokinase
- Dehydrogenases
- Transaminases
- Hydrolytic enzymes, pepsin, trypsin, etc.

**Transport Proteins**

These proteins are involved in the process of transportation, e.g.

- Hemoglobin transports oxygen
- Transferrin transports iron
- Albumin carries fatty acids and bilirubin.

**Storage Proteins**

Many proteins serve as storage form, e.g.

- Apoferritin stores iron in the form of ferritin
- Myoglobin stores oxygen in muscles.

**Contractile Proteins**

Some proteins have the ability to contract and function in the contractile system of skeletal muscle, e.g.

- Actin
- Myosin.

**Structural Proteins**

Many proteins serve as supporting framework of cells to give biological structure, strength or protection, e.g.

- Collagen in bone
- Cartilage, elastin of ligaments
- Keratin of hair, nail.

**Defence Proteins**

Many proteins involved in defence mechanism against invasion of foreign substances such as viruses, bacteria and cells. Examples of defence proteins are:

- Immunoglobulins or antibodies
- Fibrinogen and thrombin are blood clotting proteins that prevent loss of blood when the vascular system is injured.

**Regulatory Proteins**

Some proteins regulate cellular or physiological activity, e.g.

- Many hormones, e.g. *insulin*, regulate sugar metabolism; *growth hormone* of pituitary gland regulates growth of the cells.

### Classification of Proteins Based on Physical and Chemical Properties of Protein

According to the joint committee of the *American Society of Biological Chemists and American Physiological Society*, proteins are classified into three main groups as follows:

1. Simple proteins
2. Conjugated proteins
3. Derived proteins.

#### Simple Proteins

Simple proteins are defined as those proteins that upon hydrolysis, yield only amino acids or their derivatives. They are subclassified according to their solubility and heat coagulability as follows:

##### Albumins

The albumins are *soluble in water, coagulated by heat*. It is deficient in glycine, e.g.

- Egg albumin
- Serum albumin
- Lactalbumin of milk.

##### Globulins

The globulins are insoluble in water, but they are *soluble in dilute neutral salt solution and are heat coagulable*, e.g.

- Ovoglobulin of egg yolk
- Serum globulin
- Myosin of muscle.

##### Glutelins

Theutelins are *soluble in dilute acids and alkalies* but they are *insoluble in neutral solvents*. They are plant proteins, e.g.

- Glutelin of wheat
- Oryzenin of rice.

##### Prolamins or alcohol soluble proteins

The prolamins are soluble in **70 to 80% alcohol**, but they are *insoluble in water, neutral solvent or absolute alcohol*. The prolamins are rich in proline but are *deficient in lysine*. They are plant proteins, e.g.

- Zein of corn
- Gliadin of wheat.

##### Histones

The histones are *soluble in water, but are not coagulated by heat*. Histones are basic proteins as they are rich in

basic amino acids. The histones, being basic, usually occur in tissues in salt combinations with acidic substances, such as nucleic acids (RNA and DNA), e.g.

- Nucleoprotein.

##### Protamines

They are strongly basic and rich in basic amino acid arginine. The protamines are *soluble in water but are not heat coagulable*. Like histones, they occur in tissues with nucleic acids, e.g. nucleoproteins.

##### Scleroproteins (fibrous proteins)

Fibrous proteins are also called **sclero proteins**. They are insoluble, (in all common solvents like water, neutral salt solution, organic solvents, dilute acid and alkali) high molecular weight fibers. Examples of sclero proteins are:

- Collagen found in cartilage and tendons
- Elastin found in tendon and arteries
- Keratin of hair, skin and nail.

#### Conjugated Proteins

Conjugated proteins are composed of simple protein combined with some non-protein substance. The non-protein group is referred to as the **prosthetic** (additional) group. Following are the examples of conjugated proteins:

##### Nucleoproteins

The nucleoproteins are composed of simple basic proteins (histones or protamines) with **nucleic acids (RNA and DNA) as the prosthetic groups**. They are proteins of cell nuclei, e.g.

- Nucleohistone
- Nucleoprotamine.

##### Glycoproteins and proteoglycans or mucoproteins

These consist of simple protein and **carbohydrate as a prosthetic group**.

- When carbohydrate content is less than 4% of protein it is called **glycoprotein**, e.g.
  - Mucin of saliva
  - Immunoglobulins
- When carbohydrate content is more than 4%, it is called **mucoprotein** or **proteoglycans**, e.g. glycosaminoglycans.

##### Chromoproteins

Chromoproteins are composed of simple proteins with a **colored prosthetic group**, e.g.

- Hemoglobin
- Cytochromes
- Catalase
- Peroxidase.

In all these chromo proteins, prosthetic group is heme.

### Phosphoproteins

The phosphoproteins are formed by a combination of protein with *prosthetic group phosphoric acid*, e.g.

- Casein of milk
- Vitellin of egg yolk.

### Lipoproteins

The lipoproteins are formed by a combination of protein with a *prosthetic group lipid*, e.g.

- Serum lipoproteins like:
  - Chylomicrons
  - Very low density lipoprotein (VLDL)
  - Low density lipoprotein (LDL) and
  - High density lipoproteins (HDL).

### Metaloproteins

The *prosthetic group is metallic elements* such as: Fe, Co, Mn, Zn, Cu, Mg, etc., for example;

- Ceruloplasmin is a copper containing protein
- Carbonic anhydrase, carboxypeptidase and DNA polymerase are zinc containing proteins.

### Derived Proteins

This class of proteins as the name implies, includes those substances formed from simple and conjugated proteins. Derived proteins are subdivided into:

- Primary derived proteins (denatured proteins)
- Secondary derived proteins.

#### Primary derived proteins (denatured proteins)

These protein derivatives are formed by agents, such as heat, acids, alkalis, etc. which cause only slight changes in the protein molecule and its properties without hydrolytic cleavage of peptide bond. These are synonymous with denatured proteins (**Figure 4.9**), e.g.

- Proteans
- Metaproteins.

#### Proteans

These are the earliest product of protein hydrolysis by action of dilute acids or enzymes, e.g.

- Myosan from myosin
- Fibrin from fibrinogen.

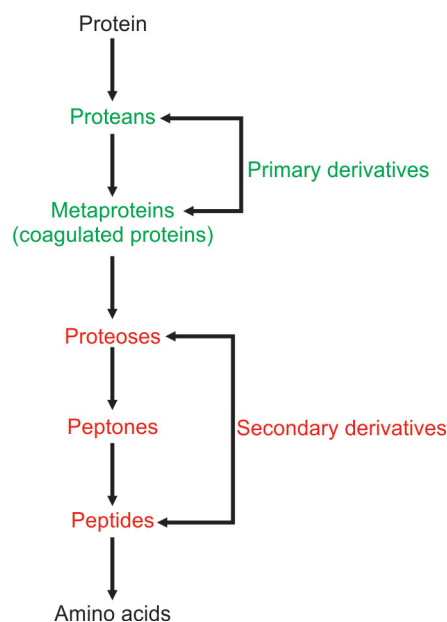
#### Metaproteins

The metaproteins are formed by further action of acids and alkalis on proteans, e.g.

- Acid and alkali albuminates.

#### Secondary derived proteins

These substances are formed in the progressive hydrolytic



**Figure 4.9:** Formation of primary and secondary derived proteins

cleavage of the peptide bonds of metaproteins (coagulated proteins) into progressive smaller molecules, e.g.

- Proteoses
- Peptones
- Peptides (**Figure 4.9**).

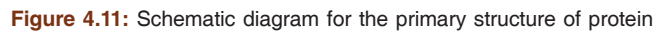
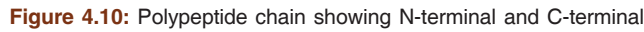
### STRUCTURE OF PROTEINS

Every protein in its native state has a unique three-dimensional structure which is referred to as its *conformation* and made up of only 20 different amino acids. The number and sequence of these amino acids are different in different proteins. Protein structure can be classified into four levels of organization:

1. Primary structure
2. Secondary structure
3. Tertiary structure
4. Quaternary structure.

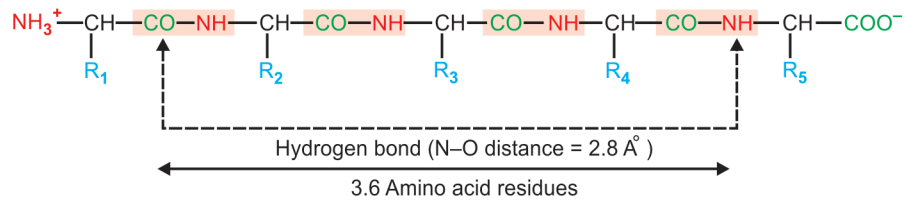
#### Primary Structure of Proteins

- The sequence of amino acids forming the backbone of proteins and location of any disulfide bond in a protein is called, the primary structure of the protein (**Figures 4.10 and 4.11**).
- In proteins, amino acids are joined covalently by *peptide bonds*, which are formed between  $\alpha$ -carboxyl group of one amino acid and  $\alpha$ -amino group of another with the elimination of a water molecule (**Figure 4.7**).

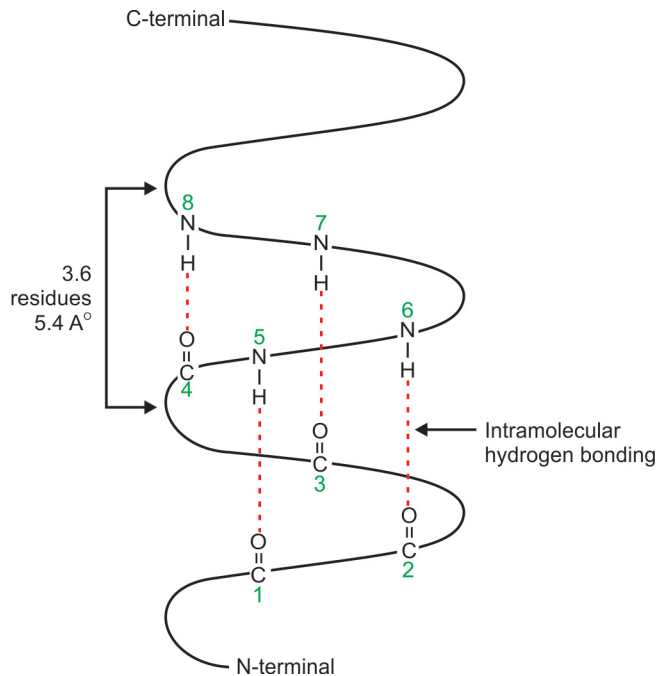


- The helix is stabilized by hydrogen bonds between the NH and CO groups of the same chain.
- These hydrogen bonds have an essentially optimal nitrogen to oxygen (N-O) distance of 2.8 Å. Thus, CO group of each amino acid is hydrogen bonded to the -NH of the amino acid that is situated four residues ahead in the linear sequence (**Figure 4.12**).
- The axial distance between adjacent amino acids is 1.5 Å and gives 3.6 amino acid residues per turn of helix (**Figure 4.13**).

It is called  $\alpha$  because it was the first structure elucidated by Pauling and Corey. If a backbone of polypeptide chain



**Figure 4.12:** Formation of hydrogen bond in  $\alpha$ -helix



**Figure 4.13:** Schematic diagram of  $\alpha$ -helical structure of protein. C=O group of each amino acid is hydrogen bonded to the NH of the amino acid that is situated four residues ahead (Number indicates the amino acid residues)

- In  $\beta$ -sheet, the hydrogen bonds are perpendicular to the polypeptide backbone rather than parallel as in the  $\alpha$ -helices.

The arrangement of polypeptide chains in  $\beta$ -pleated sheet conformation can occur in two ways:

1. Parallel pleated sheet
2. Anti-parallel pleated sheet.

#### Parallel pleated sheet

In parallel pleated sheet:

- The polypeptide chains lie side-by-side and in the same direction (with respect to N- and C-terminal), so that their N-terminal residues are at the same end

(N-terminal faces to N-terminal) and stabilized by hydrogen bonding.

- Here the hydrogen bonds are (interchain) formed between NH of a polypeptide in one chain and carbonyl (C=O) of a neighboring chain (**Figure 4.14 B**).

#### Anti-parallel pleated sheet

In the anti-parallel pleated sheet:

- The polypeptide chains lie in opposite directions, i.e. N-terminal end of one is next to the C-terminal of the other. (N-terminal faces to C-terminal) (**Figure 4.14A**).
- It is stabilized by interchain hydrogen bonding.

#### Other types of secondary structures

Besides the  $\alpha$ - and  $\beta$ -structures described above, the  $\beta$ -bends, loop regions and disordered regions, are also found in proteins.

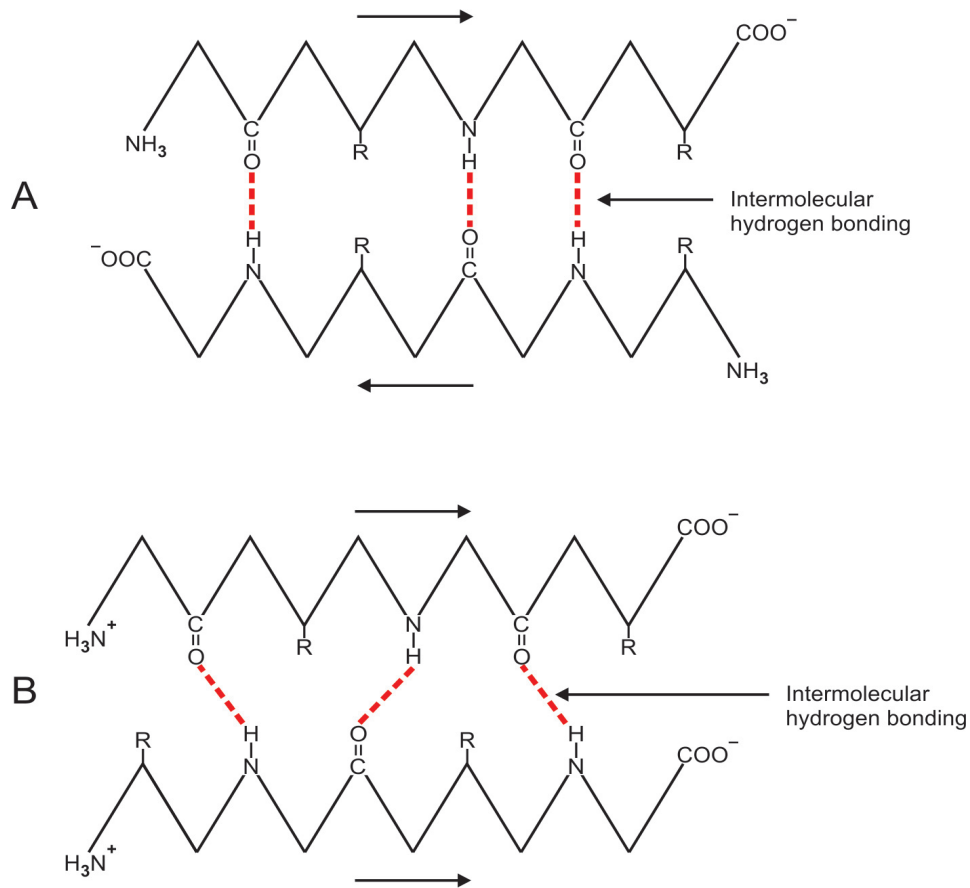
### Tertiary Structure

- The peptide chain, with its secondary structure, may be further folded and twisted about itself forming three-dimensional arrangement of the polypeptide chain (**Figure 4.15**).
- Amino acid residues which are very distant from one another in the sequence can be brought very near due to the folding and thus form regions essential for the functioning of the protein, like **active site** or catalytic site of enzymes.
- Thus, the three-dimensional folded compact and biologically active conformation of a protein is referred to as its tertiary structure, e.g. myoglobin.

#### Tertiary Structure Stabilizing Forces

The three-dimensional tertiary structure of a protein is stabilized by:

- Hydrogen bonds
- Hydrophobic interactions
- Van der Waals forces
- Disulfide bond
- Ionic (electrostatic) bonds or salt bridges.



**Figures 4.14 A and B:**  $\beta$ -pleated sheet structure. (A) Anti-parallel; (B) Parallel



**Figure 4.15:** Tertiary structure of protein

### Quaternary Structure of Protein

- Only those proteins that have **more than one polypeptide chain (polymeric)** have a quaternary structure. Not all proteins are polymeric. Many proteins consist of a **single polypeptide chain and are called monomeric** proteins, e.g. myoglobin.
- The arrangement of these polymeric polypeptide subunits in three-dimensional complexes is called the quaternary structure of the protein (**Figure 4.16**).

- Examples of proteins having quaternary structure are:
  - Lactate dehydrogenase
  - Pyruvate dehydrogenase
  - Hemoglobin.

### Quaternary Structure Stabilizing Forces

The subunits of polymeric protein are held together by noncovalent interactions or forces such as:

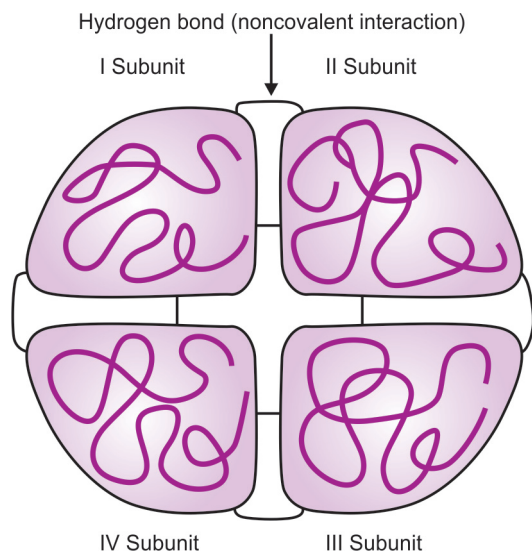
- Hydrophobic interactions
- Hydrogen bond
- Ionic bonds.

### Bonds Responsible for Protein Structure

Protein structure is stabilized by two types of bonds

1. Covalent bond, e.g.
  - Peptide bonds
  - Disulfide bond





**Figure 4.16:** Schematic representation of quaternary structure of polymeric protein

2. Noncovalent bond, e.g.
  - Hydrogen bond
  - Hydrophobic bond or interaction
  - Electrostatic or ionic bond or salt bond or salt bridge
  - Van der Waals interactions.

### Covalent Bond

#### Peptide bonds ( $-\text{CO}-\text{NH}-$ )

- A peptide bond is formed by the condensation of the amino group of one amino acid with the carboxyl group of another amino acid with a removal of a water molecule (See Figure 4.7).

#### Disulfide bond ( $-\text{S}-\text{S}-$ )

- A covalent bond formed between the sulfhydryl group ( $-\text{SH}$ ) of side chain of cysteine residues in the same or different peptide chains.
- These disulfide bonds help to stabilize against denaturation and confer additional stability.

### Noncovalent Bonds

#### Hydrogen bond

- Bond formed between  $-\text{NH}$  and  $-\text{CO}$  groups of peptide bond by sharing single hydrogen.
- Hydrogen bond may occur within the same polypeptide chain (intrachain) or between different polypeptide chains (interchain).
- Side chains of 11 out of the 20 standard amino acids can also participate in hydrogen bonding.

#### Hydrophobic bond or interaction

These are formed by interaction between nonpolar hydrophobic R groups (side chain) of amino acids like alanine, valine, leucine, isoleucine, methionine, phenylalanine and tryptophan.

#### Electrostatic or ionic bond or salt bond or salt bridge

These are formed between oppositely charged groups when they are close, such as amino ( $\text{NH}_3^+$ ) terminal and carboxyl ( $\text{COO}^-$ ) terminal groups of the peptide and the oppositely charged R-groups of polar amino acid residues.

#### Van der Waals interactions

Van der Waals forces are extremely weak and act only on extremely short distances and include both an attractive and a repulsive forces (between both polar and nonpolar side chain of amino acid residues).

## PROPERTIES OF PROTEINS

### Colloidal Nature

Protein molecules exist in the form of colloidal particle, 5-100 m $\mu$  in dimension, are heavier than water and sink.

### Colloidal Osmotic Pressure

- Colloidal protein molecules exert **osmotic pressure**. The osmotic pressure generated by plasma proteins is often called the **colloidal osmotic pressure** or **oncotic pressure** of plasma.
- The osmotic pressure of protein is proportional to its concentration, but inversely proportional to its molecular weight.
- In blood plasma, albumin contributes 75-80% of osmotic pressure (although it represents no more than half the plasma proteins), because its molecular weight is lower.
- Oncotic pressure exerted by protein is clinically important in maintaining blood volume.

### Molecular Weight

Proteins are macromolecules and have very high molecular weights with wide variations, e.g. molecular weight of:

- Albumin = 69,000
- $\gamma$ -Globulin = 1,60,000.

### Solubility

In general, globular proteins, such as, albumin have higher solubility than elongated fibrous proteins. Moreover, smaller molecules are more soluble than larger molecules.



### Shape of the Protein

There is a wide variation in the protein shape.

- Scleroproteins like keratin, collagen are in the form of fibers.
- While soluble proteins tend to be of rounded shape and are called globular proteins.

### Hydration of Proteins

- Proteins, when brought into contact with water, absorb water and swell up.
- The polar groups like COOH, NH<sub>2</sub>, OH of protein bind to the molecules of water by hydrogen bonds to hold a considerable amount of water. Thus, a relatively immobile shell-like layer of water, called the “*solvation layer*” or *water envelope* is held around each protein particle in an aqueous medium.

### Amphoteric Nature and Isoelectric pH of the Proteins

The isoelectric pH of amino acid has been described previously. One end of the protein molecule has free amino group, while the other end has free COOH group.

- In acid solution, the NH<sub>2</sub> groups accept H<sup>+</sup> ion and present as NH<sub>3</sub><sup>+</sup> (cation). Therefore protein in acid solution will be positively charged.
- In alkaline pH, the COOH groups donate H<sup>+</sup> ion and are present as COO<sup>-</sup> (anion). Hence, proteins in alkaline solution are negatively charged.

So, proteins are **ampholytes** acting both as donors and acceptors of H<sup>+</sup> ion.

### Isoelectric pH of the Protein

- For every protein in solution, there is a particular pH at which the number of anions formed is exactly equal to the number of cations, and the solution is electrically neutral. That pH is called the *isoelectric pH (PI)* of that protein and the protein exists as *zwitter ion*.
- At isoelectric pH:
  - Protein molecules do not migrate in an electric field.
  - Solubility, buffering capacity and viscosity will be minimum and precipitability will be maximum.
- PI of some proteins are given below:
  - Pepsin: 1.1
  - Casein: 4.6
  - Albumin: 4.7
  - Globulin: 6.4.
- At pH values above or below the isoelectric pH, they carry a net negative or positive charge and migrate to anode (+vely charged electrode) or cathode (-vely charged electrode).

### Precipitation of Proteins

The stability of protein in solution depends on the **charge** and **hydration** of the protein molecule.

- The factors which neutralize the charge or remove water of hydration will cause precipitation of proteins. The factors used for precipitation of proteins are:
  - Salting out
  - Isoelectric pH
  - Heavy positive or negative ions
  - Organic solvents.

#### Salting out method

When neutral salts such as ammonium sulfate or sodium sulfate are added to a protein solution, the addition may precipitate a protein from its solution.

- Mineral ions attract water molecules and consequently remove the shell of hydration (solvation layer) from around protein molecules.
- Since water layer around protein particles is removed, the protein is precipitated. This is called *salting out*.
- Higher molecular weight protein requires less salt to precipitate than low molecular weight protein. Thus, globulins are precipitated at half saturation of ammonium sulfate or 22% Na-sulfate, but albumin requires full saturation of ammonium sulfate or 28% Na-sulfate.

**Application:** These methods are therefore useful in separating albumin and globulin from serum proteins.

#### Precipitation at isoelectric pH

All proteins are least soluble at their isoelectric pH (PI) and can then be precipitated.

#### Precipitation by heavy positive or negative ions

- On the acidic side of its PI, a protein remains as cation (+vely charged) and may then be precipitated by neutralizing the charge on protein by adding anion (-vely charged) or alkaloidal reagents like *tungstate*, *trichloroacetate* or *picrate*.
- On the alkaline side of its PI, a protein exists as an anion and may then be precipitated as metal proteinates by heavy metal ions like Zn<sup>2+</sup>, lead, mercury, etc.

#### Precipitation by organic solvents

Organic solvent like alcohol, dehydrates and precipitates the proteins.

## DENATURATION OF PROTEINS

The three-dimensional conformation, the *primary*, *secondary*, *tertiary* and even in some cases *quaternary* structure is characteristic of a native protein. **Hydrogen**

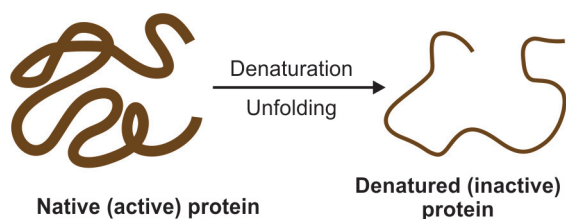


Figure 4.17: Denaturation of protein

*bond*, *ionic bond* and *hydrophobic bond* stabilize the structure to maintain its conformation in space. This conformation can upset and disorganized *without breakage of any peptide linkage*, only by the rupture of ionic bond, hydrogen bonds and hydrophobic bond which stabilize the structure. This is called **denaturation** (Figure 4.17). Denaturation of proteins leads to:

- Unfolding of natural coils of native protein.
- Decrease in solubility and increase in precipitability.
- Loss of biological activities, (e.g. enzyme activity) and antigenic properties.
- Increased digestibility.

#### Denaturing Agents

- Denaturation is brought about by certain:
  - Physical agents
  - Chemical agents
  - Mechanical means.
- **Physical agents** : Heat, Ultraviolet rays and ionizing radiations can denature proteins.
- **Chemical agents** : Acids, alkalies and certain acid solutions of heavy metals, e.g. mercury, lead, detergents; organic solvents like alcohol, acetone, etc. denature proteins.
- **Mechanical means** : Vigorous shaking or grinding leads to denaturation of the protein.

#### Examples of Denatured Protein

Cooked meat or boiled egg, milk paneer, etc.

#### Significance of Denaturation

- Digestibility of native protein is increased on denaturation by gastric HCl or by heat on cooking. Denaturation causes unfolding of native polypeptide coil so that hidden peptide bonds are exposed to the action of proteolytic enzyme in the gut. It also increases reactivity of certain groups.
- Denaturation property of a protein is used in blood analysis to eliminate the proteins of the blood (deproteinization of blood).

#### Coagulation

Denaturation may, in rare cases be reversible, in which case the protein refolds into its original native structure, when the denaturing agent is removed. However, most proteins, once denatured, remain permanently disordered and are called *irreversible denaturations* or *coagulation*, e.g. coagulated egg white of boiled egg.

#### SUMMARY

- The basic structural units of proteins are amino acids.
- All proteins in all species from bacteria to humans are constructed from the same set of twenty amino acids called standard amino acids.
- Each amino acid contains an  $\alpha$ -carboxyl group, an  $\alpha$ -amino group and distinctive R (side chain) group attached to the  $\alpha$ -carbon atom.
- The  $\alpha$ -carbon atom of the amino acids except glycine is asymmetric and thus can exist in stereoisomeric forms (D and L), only L- $\alpha$ -amino acid (L-stereoisomer) are found in protein.
- Amino acids can be classified on the basis of nature of amino acids in the solution or structure of the side chain or nutritional requirement or metabolic products of amino acids.
- PI is the pH at which an amino acid bears no net charge and does not move in an electrical field.
- Among the 20 standard amino acids, histidine serves as the best buffer at physiological pH.
- Proteins are classified on the basis of their function, and physical and chemical properties.
- Every protein has a unique three-dimensional structure that reflects its function.
- There are four levels of protein structure: Primary, secondary, tertiary and quaternary (only for oligomeric proteins).

#### EXERCISE

#### Multiple Choice Questions (MCQs)

1. All amino acids found in proteins are optically active, *except*:
 

a) Serine	b) Glycine
c) Threonine	d) Tyrosine
2. All of the following forces may play a role in the formation of tertiary structure, *except*:
 

a) Hydrogen bond
b) Disulfide bridges
c) Hydrophobic interaction
d) Peptide bonds

3. The greatest buffering capacity at physiological pH would be provided by a protein, rich in which of the following amino acids:
  - a) Glycine                      b) Lysine
  - c) Histidine                  d) Valine
4. In proteins, triple helix is an example of:
  - a) Primary structure    b) Secondary structure
  - c) Tertiary structure    d) Quaternary structure
5. Which of the following amino acid is exclusively ketogenic?
  - a) Leucine                      b) Phenylalanine
  - c) Threonine                  d) Isoleucine
6. In humans all of the following amino acids are essential, *except*:
  - a) Valine                        b) Isoleucine
  - c) Glycine                      d) Phenylalanine
7. During denaturation of protein the following bonds are disrupted, *except*:
  - a) Hydrogen                  b) Hydrophobic
  - c) Peptide                      d) Sulfide
8. The imino acid present in protein is:
  - a) Phenylalanine            b) Valine
  - c) Leucine                      d) Proline
9. Glycine is used for synthesis of the following, *except*:
  - a) Heme                        b) Serotonin
  - c) Purine                        d) Creatine
10. Which class of amino acids contain only non-essential amino acids?
  - a) Aromatic                  b) Acidic
  - c) Branched chain          d) Basic
11. Which of the following is a non-protein amino acid?
  - a) Proline
  - b) Histidine
  - c) Ornithine
  - d) Asparagine
12. Which of the following is a kind of secondary structure?
  - a)  $\alpha$ -helix
  - b)  $\beta$ -bend
  - c) Triple helix
  - d) All of the above
13. Glutathione is found in all mammalian cells *except*:
  - a) RBC                            b) Neurons
  - c) Skeletal muscle            d) Argentaffin cells
14. Albumin is deficient in which of the following amino acid?
  - a) Glycine                      b) Tryptophan
  - c) Cysteine                      d) Methionine
15. Which of the following acts as a redox buffer?
  - a) Insulin                        b) Glucagon
  - c) Glutathione                  d) Angiotensin

#### Correct Answers for MCQs

1-b	2-d	3-c	4-b
5-a	6-c	7-c	8-d
9-b	10-b	11-c	12-d
13-b	14-a	15-c	