COMPILED AND CIRCULATED BY DR. POULAMI ADHIKARY MUKHERJEE, ASSISTANT PROFESSOR, DEPARTMENT OF ZOOLOGY, NARAJOLE RAJ COLLEGE

PROTEINS

BY

DR. POULAMI ADHIKARY MUKHERJEE
ASSISTANT PROFESSOR
DEPARTMENT OF ZOOLOGY
NARAJOLE RAJ COLLEGE

ZOOLOGY: SEM- III, PAPER- C7T: FUNDAMENTALS OF BIOCHEMISTRY, UNIT 3: PROTEINS

PROTEINS

John J. Berzelius (1838) first coined the term 'protein' (Gr. proteios — of the first rank) to stress the importance of this class of polymers.

- Proteins are nitrogenous organic compounds of high molecular weight, which play a vital role in living organisms.

- They are large biomolecules, or macromolecules, consisting of one or more long chains of amino acid residues.

- Proteins perform a vast array of functions within organisms, including catalysing metabolic reactions, DNA replication, responding to stimuli, providing structure to cells, and organisms, and transporting molecules from one location to another.

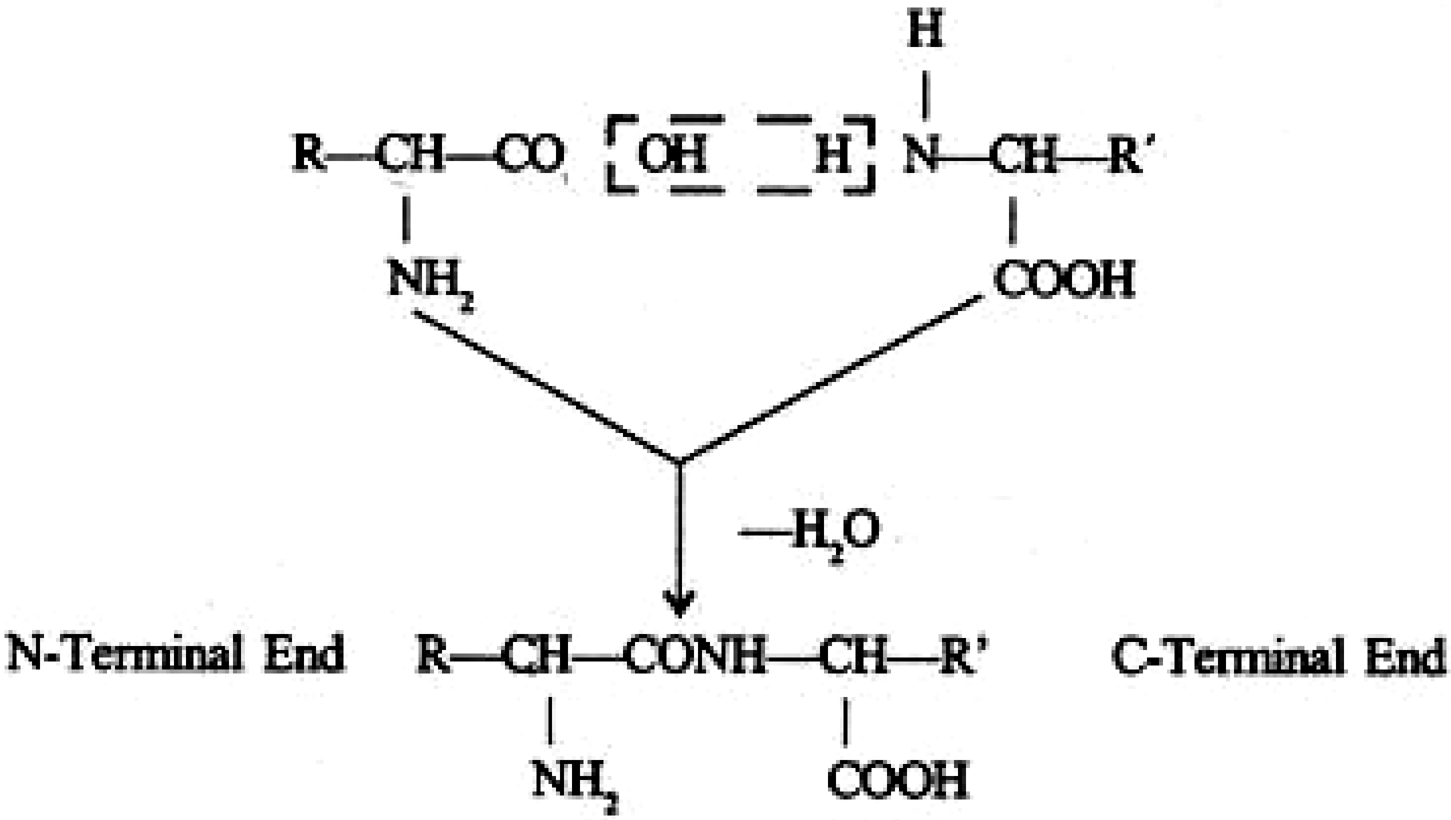
- Proteins differ from one another primarily in their sequence of amino acids, which is dictated by the nucleotide sequence of their genes, and which usually results in protein folding into a specific 3D structure that determines its activity.

- Proteins are organic nitrogenous compounds in which a large number of amino acids are joined together by **peptide linkages** to form **long polypeptide chains**.

- Peptide-linkage (—CONH—) is formed when amino group (—NH_2) of one amino acid con-denses with carboxylic group (—COOH) of another amino acid eliminating one molecule of water.

- That end of the polypeptide chain where the —COOH group of the amino acid is not involved in peptide linkage is called as C-terminal end.

- The other end of the polypeptide chain with amino acid having free —NH_2 group is called as N-terminal end.



- Although there may be hundreds of amino acids in a single polypeptide chain but fundamentally there are only about 20 different types of amino acids that constitute proteins in plants (there may be repetition of amino acids continuously or at intervals in the polypeptide chain).

- Because of their very large size the proteins are often called as gigantic molecules or macromolecules of the cells.

Their molecular weight may range from few thousands to over a million (10 lakhs)

- Proteins are the macromolecules composed of one or more polypeptide chains, each of which is a mixed polymer of L-α-amino acid residues joined end-to-end by peptide bonds.

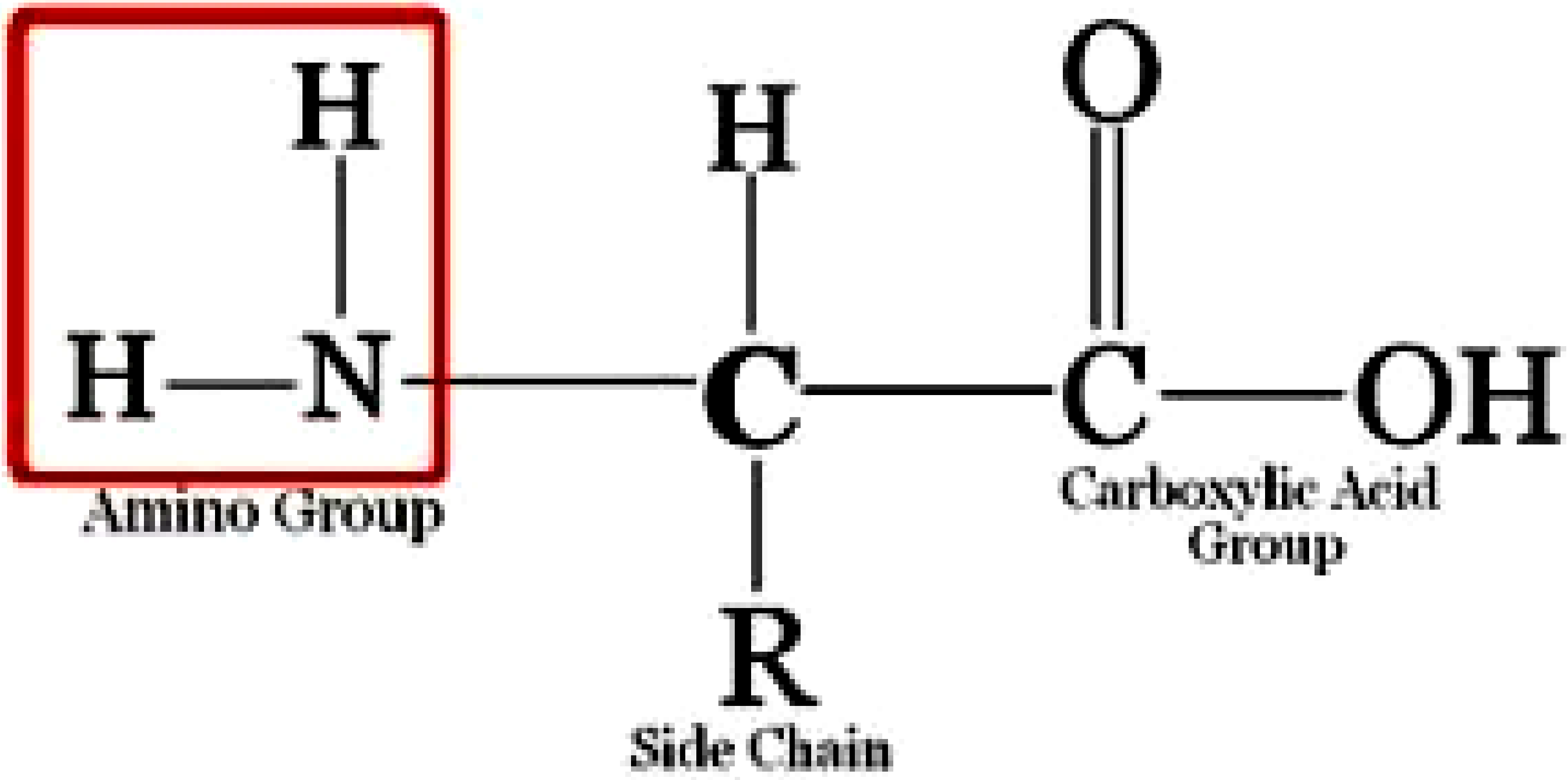
- Monomeric protein consists of single polypeptide chain, e.g., lysozyme, myoglobin. The oligomeric or multimeric protein consists of 2 or more polypeptide chains, each of which is called a **protomere** or subunit.

- Rubisco consists of 24 polypeptides, haemoglobin (Hb) is a tetrameric consists of two α-chains and two β- chains, immunoglobulins consists of 2 H-chains and 2H-chains etc.

Composition of Proteins:

Proteins are large molecules consisting of many amino-acids connected by **"peptide linkages".**

Peptide bond is produced when carboxyl radical $(-\overset{O}{\overset{\|}{C}}-OH)$ of one amino acid reacts with the amino ($-NH_2$) group of the other amino acid. The basic structural formula of amino acids is shown in below:


H
H
O
H—N
C
C
OH
Amino Group
Carboxylic Acid Group
R
Side Chain

It consists of one alpha (a) carbon atom that is associated with an amino group (-NH_2) with a potential (+) charge, a carboxyl group $(-\overset{O}{\overset{||}{C}}-OH)$ with a (-) charge, a hydrogen atom and a side chain "R" that varies in the different amino acids.

There are usually 20 amino acids found in proteins (for structural formulae of the amino acids, any book of biochemistry may be consulted). These twenty amino-acids are divided into 7 groups. All the 20 amino acids need not be present in a given protein.

The side chains (R) of amino acids are responsible for the different properties, such as, water solubility, interaction with other amino acids etc. of the amino acids. The amino acids possessing $-CH_3$ group are much less soluble in water and they are called **"hydrophobic"** amino acids, e.g., leucine, isoleucine, valine.

The amino acids that are water soluble are called "hydrophilic" amino-acids, e.g., lysine (+ charge) and aspartic acid (-charge).

The sulfhydryl group (-SH) of cysteine can interact with the -SH group of other cysteine in the protein chain to make a disulfide linkage (S-S). The H atoms of hydroxyl group (-OH) or carboxyl group $(-\overset{\overset{O}{||}}{C}-OH)$ of the "R ' chain can make hydrogen bonding with other amino acids in the protein chain. The bonds are required for stabilizing the structure of protein molecules.

Properties of Proteins:

1. Denaturation: Partial or complete unfolding of the native (natural) conformation of the polypeptide chain is known as denaturation. This is caused by heat, acids, alkalies, alcohol, acetone, urea, beta- mercaptoethanol.

2. Coagulation: When proteins are denatured by heat, they form insoluble aggregates known as coagulum. All the proteins are not heat coagulable, only a few like the albumins, globulins are heat coagulable.

3. Isoelectric pH (pH^{1}):

The pH at which a protein has equal number of positive and negative charges is known as isoelectric pH. When subjected to an electric field the proteins do not move either towards anode or cathode, hence this property is used to isolate proteins. The proteins become least soluble at pH^{I} and get precipitated. The pH^I of casein is 4.5 and at this pH the casein in milk curdles producing the curd.

4. Molecular Weights of Proteins:

The average molecular weight of an amino acid is taken to be 110. The total number of amino acids in a protein multiplied by 110 gives the approximate molecular weight of that protein. Different proteins have different amino acid composition and hence their molecular weights differ. The molecular weights of proteins range from 5000 to 10^9 Daltons. Experimentally the molecular weight can be determined by methods like gel filtration, PAGE, ultra centrifugation or viscosity measurements.

Classification of Proteins:

Proteins are classified based upon:

(1) Their solubility and

(2) Their structural complexity.

A. Classification Based upon Solubility:

On the basis of their solubility in water, proteins are classified into:

1. Fibrous proteins:

These are insoluble in water. They include the structural proteins. They have supportive function (e.g., collagen) and/or protective function (e.g., hair keratin and fibrin).

2. Globular proteins:

They are soluble in water. They include the functional proteins, e.g., enzymes, haemoglobin, etc.

B. Classification Based upon Structural Complexity:

On the basis of their structural complexity they are further divided into:

(1) Simple

(2) Conjugated and

(3) Derived proteins.

1. Simple proteins: Proteins which are made up of amino acids only are known as simple proteins.

They are further sub-divided into:

(a) Albumins: They are water soluble, heat coagulable and are precipitated on full saturation with ammonium sulphate, e.g., serum albumin, lactalbumin and ovalbumin.

(b) Globulins: They are insoluble in water, but soluble in dilute salt solutions. They are heat coagulable and precipitate on half-saturation with ammonium sulphate, e.g., serum globulin and ovo-globulin.

(c) Glutelins: They are insoluble in water and neutral solvents. Soluble in dilute acids and alkalies. They are coagulated by heat, e.g., glutelin of wheat.

(d) Prolamines: Water insoluble but soluble in 70% alcohol, e.g., gliadin of wheat, proteins of corn, barley, etc.

(e) Histories: Water soluble, basic in nature due to the presence of arginine and lysine, found in nucleus. They help in

DNA packaging in the cell. They form the protein moiety of nucleoprotein.

(f) Protamine’s: Water soluble, basic in nature, not-heat coagulable. Found in sperm cells, hence component of sperm nucleoprotein.

(g) Globin’s: They are water soluble, non-heat coagulable. e.g., globin of haemoglobin.

(h) Albuminoids or scleroproteins: Insoluble in all neutral solvents, dilute acids or alkalies, e.g., keratin of hair and proteins of bone and cartilage.

2. Conjugated proteins:

Proteins which are made up of amino acids and a non-amino acid/protein substance called the prosthetic group are known as conjugated proteins.

The various types of conjugated proteins are:

(a) Chromo proteins: Here the non-protein part is a coloured compound in addition to the protein part. Ex. Haemoglobin has heme as the prosthetic group and cytochromes also have heme.

(b) Nucleoproteins: These proteins are bound to nucleic acids, e.g., chromatin (histones + nucleic acids).

(c) Glycoproteins: When a small amount of carbohydrate is attached to a protein it is known as glycoproteins, e.g., mucin of saliva. (Note: Glycoproteins have major amounts of protein and some amount of carbohydrates and proteoglycans contain major amounts of carbohydrates and little amount of proteins).

(d) Phosphoprotein: Phosphoric acid is present with the protein. Ex. Milk casein and egg yolk (vitellin).

(e) Lipoproteins: Proteins in combination with lipids, e.g., LDL, HDL.

(f) Metalloproteins: They contain metal ion in addition to the amino acids, e.g., hemoglobin (iron), ceruloplasmin (copper).

3. Derived proteins: They are the proteins of low molecular weight produced from large molecular weight proteins by the action of heat, enzymes or chemical agents.

Proteins → Proteans → Proteoses → Peptones → Peptides → Amino acids

Structure of Proteins

The structure of protein can be described in terms of mainly four levels of hierarchy in protein conformation: Primary, Secondary, Tertiary and Quaternary. However, recent studies revealed two additional levels of protein organization i.e. Super- secondary structures or motifs and domains.

A. PRIMARY STRUCTURE:

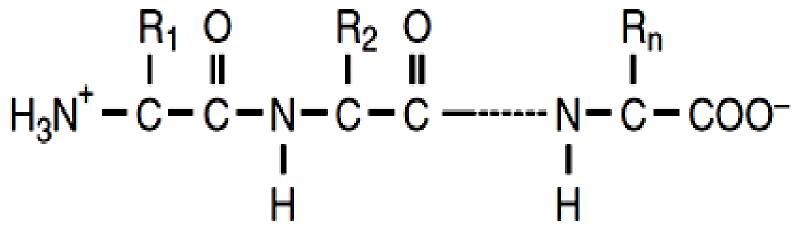
Primary Structure refers to the order of the amino acids in the peptide chain. Primary structure of a protein means the sequences amino acid residues of its polypeptide chain (s) which read in N-terminus → C-terminus direction. It is the 1st level of organization of protein determined by the codons of mRNA or cistron of DNA. The 1° structure is stabilized by the peptide bonds as well as and disulfide bonds between cysteine residues, if there are any.

Frederick Sanger (1953) first determined the 1° structure of bovine insulin. Now, the 1° structure of a polypeptide is determined by an automated device called spinning cup sequenator, developed by Pehr Edman and Geoffrey Begg.

1. The free α-amino group, written to the left, is called the amino-terminal or N-terminal end.

2. The free α-carboxyl group, written to the right, is called the carboxyl-terminal or C-terminal end.

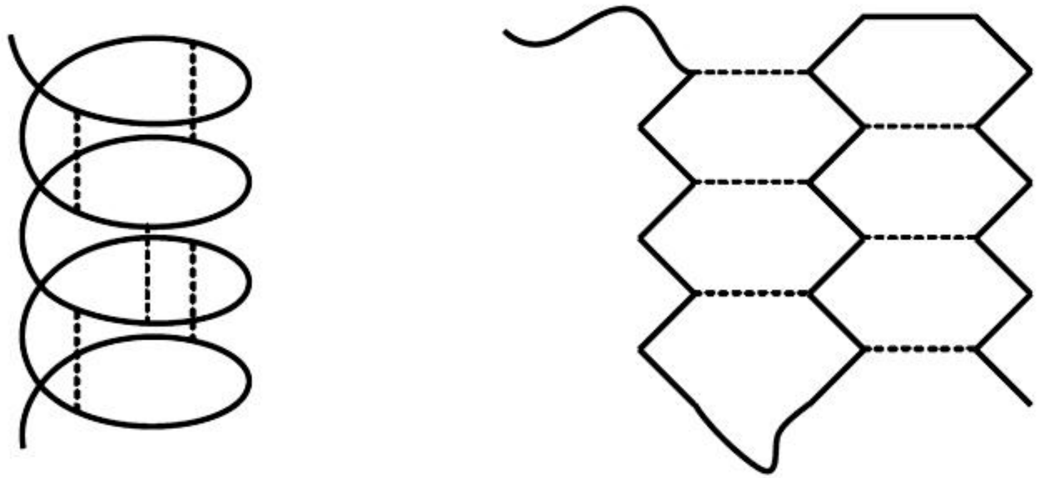
Primary structure (always written with the free amino group to the left):



B. SECONDARY STRUCTURE:

SECONDARY STRUCTURE is the arrangement of hydrogen bonds between the peptide nitrogens and the peptide carbonyl oxygens of different amino acid residues.

Secondary structures



α-Helix (intramolecular hydrogen bonds)

β-Pleated sheet (intramolecular hydrogen bonds)

1. In helical coils, the hydrogen-bonded nitrogens and oxygens are on nearby amino acid residues.

a. The most common helical coil is a right-handed α-helix.

b. α-keratin from hair and nails is an α-helical protein.

c. Myoglobin has several α-helical regions.

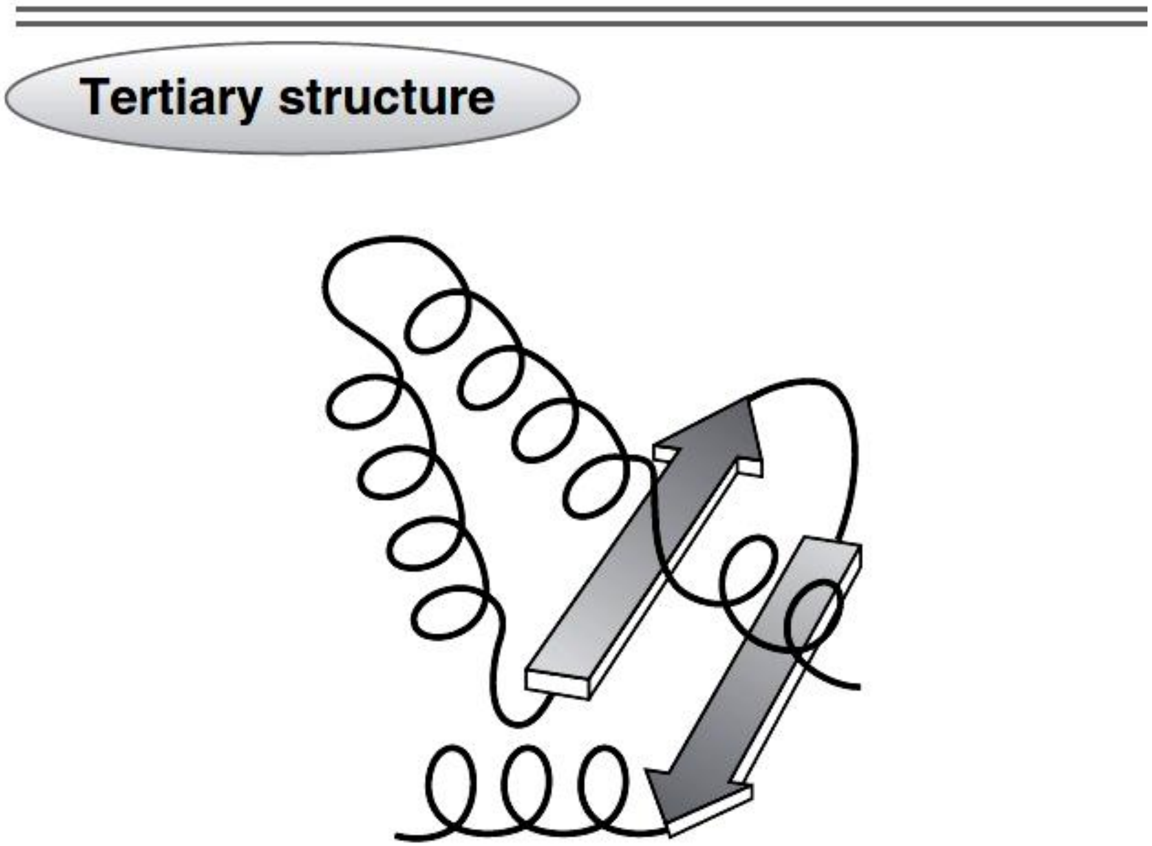
d. Proline, glycine, and asparagine are seldom found in α-helices; they are “helix breakers.”

2. In β-sheets (pleated sheets), the hydrogen bonds occur between residues on neighboring peptide chains.

a. The hydrogen bonds may be on different chains or distant regions of the same chain.
b. The strands may run parallel or antiparallel.
c. Fibroin in silk is a β-sheet protein.

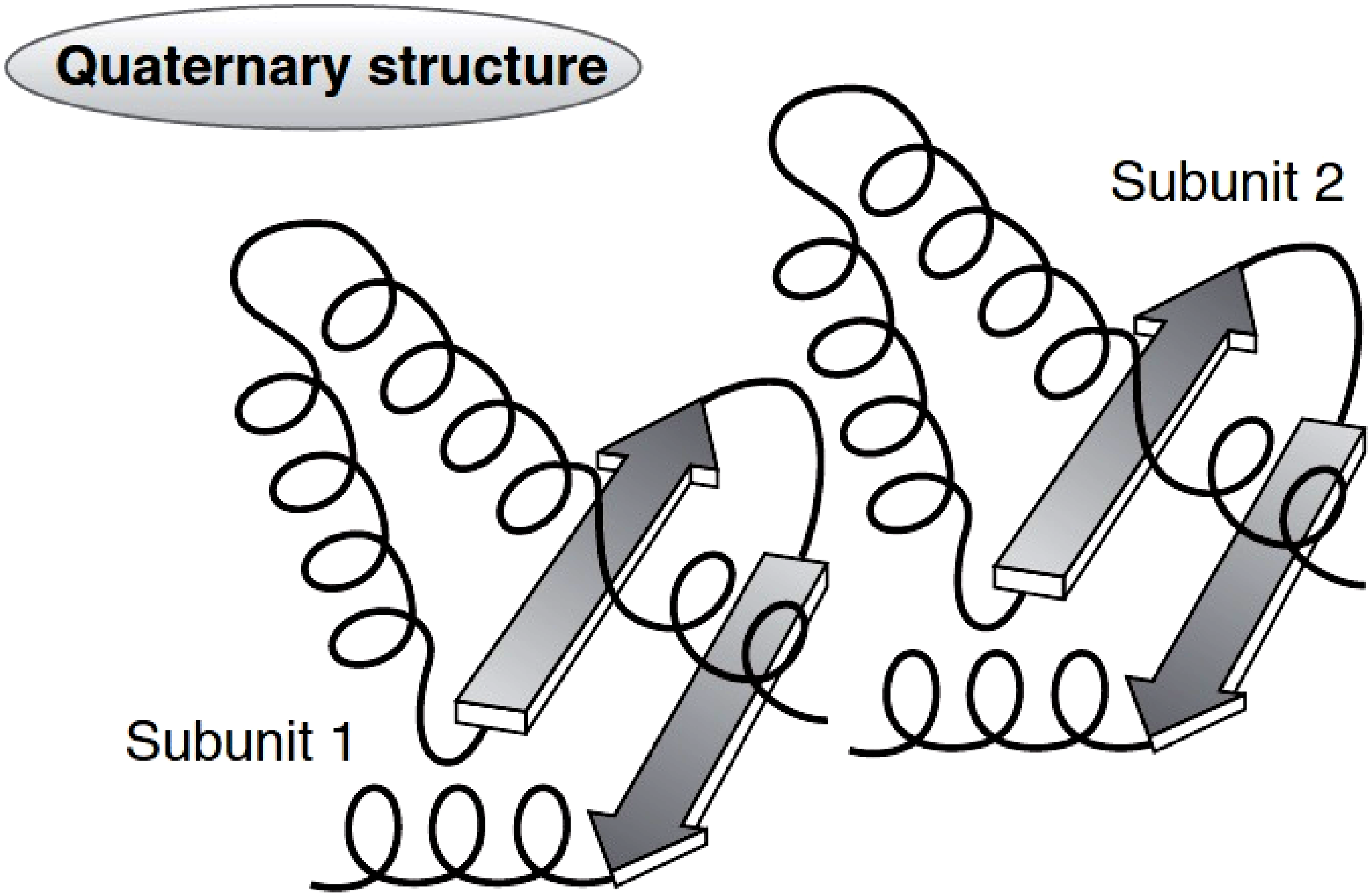
C. TERTIARY STRUCTURE:

TERTIARY STRUCTURE refers to the three-dimensional arrangement of a polypeptide chain that has assumed its secondary structure. Disulfide bonds between cysteine residues may stabilize tertiary structure.



D. QUATERNARY STRUCTURE:

QUATERNARY STRUCTURE is the arrangement of the subunits of a protein that has more than one polypeptide chain.


Quaternary structure
Subunit 2
Subunit 1

E. LEFT-HANDED HELICAL STRANDS:

LEFT-HANDED HELICAL STRANDS are wound into a supercoiled triple helix in collagen. The major structural protein in the body, collagen makes up 25% of all vertebrate protein.

a. The primary structure of collagen includes long stretches of the repeating sequence glycine-X-Y, where X and Y are frequently proline or lysine. The high proportion of proline residues leads to formation of the left-handed helical strands.

b. Only glycine has an R-group small enough to fit into the interior of the righthanded triple helix.

c. Collagen also contains hydroxyproline and hydroxylysine. The hydroxyl groups are added to proline and lysine residues by post-translational modification.

Functions of Proteins:

The main functions of proteins in human body are:

1. They serve as body building units, e.g., muscle proteins.

2. They provide support and protection to various tissues, e.g., collagen and keratin.

3. All chemical reactions in the body are catalysed by proteinaceous enzymes, e.g., trypsin.

4. They transport various molecules and ions from one organ to the other, e.g., hemoglobin, serum albumin.

5. They store and provide nutrients, e.g., milk casein, ovalbumin.

6. They defend the body from harmful foreign organisms, e.g., immunoglobulin’s, fibrinogen.

7. They help to regulate cellular or physiological activity, e.g., hormones, viz., insulin, GH.

THANK YOU