SNS COLLEGE OF PHARMACY AND HEALTH SCIENCES
Sathy Main Road, SNS Kalvi Nagar,
Saravanampatti Post, Coimbatore - 641 035,
Tamil Nadu.

Chemistry of Proteins and Amino Acids

- **Proteins are the most abundant organic molecules of the living system.**
- **They occur in the 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.**
- **In 1839 Dutch chemist G.J.Mulder while investing the substances such as those found in milk, egg, found that they could be coagulated on heating and were nitrogenous compounds.**

- **The term protein is derived from a Greek word *proteios,* meaning first place.**
- ***Berzelius (Swedish chemist)* suggested the name proteins to the group of organic compounds that are utmost important to life.**
- **The proteins are nitrogenous macromolecules composed of many amino acids.**

Biomedical importance of proteins:

- **Proteins are the main structural components of the cytoskeleton. They are the sole source to replace nitrogen of the body.**
- **Bio chemical catalysts known as enzymes are proteins.**
- **Proteins known as immunoglobulins serve as the first line of defense against bacterial and viral infections.**

- **Several hormones are protein in nature.**
- **Structural proteins like actin and myosin are contractile proteins and help in the movement of muscle fibre.**

  Some proteins present in cell membrane, cytoplasm and nucleus of the cell act as receptors.
- **The transport proteins carry out the function of transporting specific substances either across the membrane or in the body fluids.**

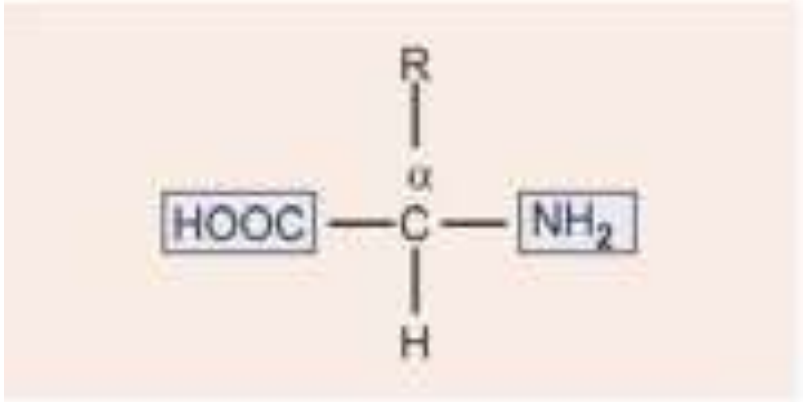
- **Storage proteins bind with specific substances and store them, e.g. iron is stored as ferritin.**
- **Few proteins are constituents of respiratory pigments and occur in electron transport chain, e.g. Cytochromes, hemoglobin, myoglobin**
- **Under certain conditions proteins can be catabolized to supply energy.**
- **Proteins by means of exerting osmotic pressure help in maintenance of electrolyte and water balance in the body.**

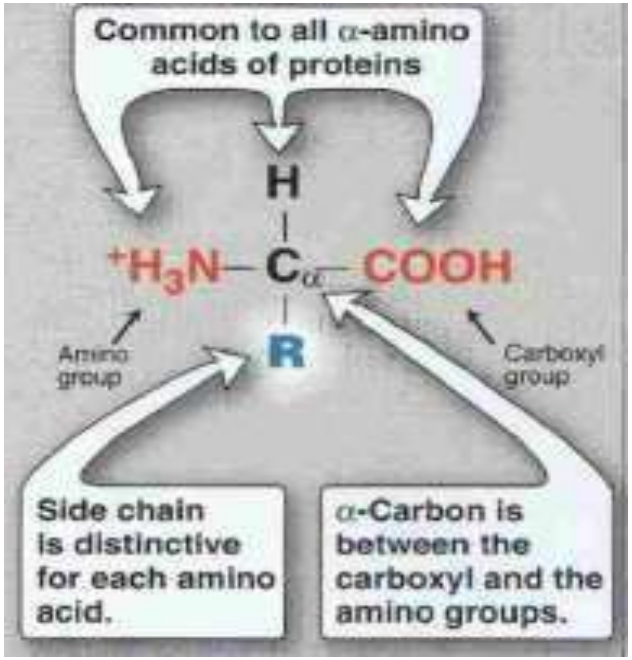
Amino acids

- **Amino acids are a group of organic compounds containing two functional groups – *amino and carboxyl.***
- **The amino group ($-NH_2$) is basic while the carboxyl group (-COOH) is acidic in nature**
- **There are about 300 amino acids occur in nature. Only 20 of them occur in proteins.**

Structure of amino acids:

Each amino acid has 4 different groups attached to α- carbon (which is C-atom next to COOH). These 4 groups are : amino group, COOH, Hydrogen atom and side Chain (R)

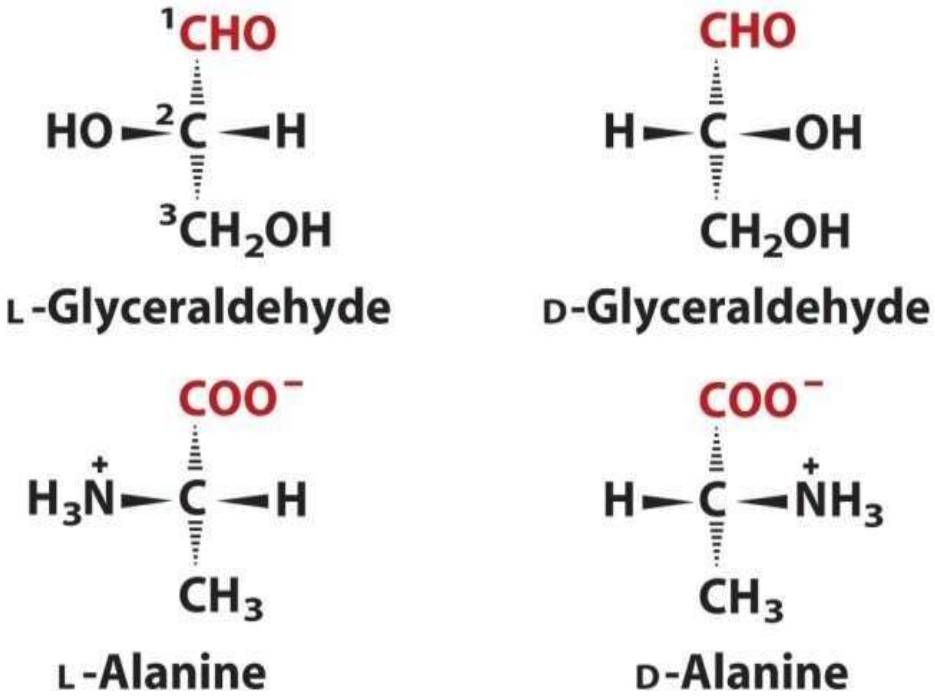




Isomerism

Stereoisomerism

In D-amino acids – NH_2 group is on the right hand while in L-amino acids it is oriented to the left.



Optical isomers of amino acids

- **If a carbon atom is attached to four different groups, it is asymmetric and therefore exhibits isomerism.**
- **The amino acids (except glycine) possess four distinct groups (R, H, COO-, NH_3+) held by a α-carbon.**
- **Thus all the amino acids have optical isomers.**
- **The proteins are composed of L- α-amino acids.**

Classification of amino acids

Amino acid classification based on the structure

The 20 standard amino acids found in protein structure are divided into seven distinct groups.

1.Aliphatic amino acids:

2.Hydroxyl group containing amino acids:

3.Sulfur containing amino acids:

4.Acidic amino acids and their amides:

5.Basic amino acids:

6.Aromatic amino acids:

7.Imino acids:

1. Aliphatic amino acids

These are monoamino monocarboxylic acids.

This group consists of most simple amino acids.

A) Glycine - Gly - G

B) Alanine - Ala - A

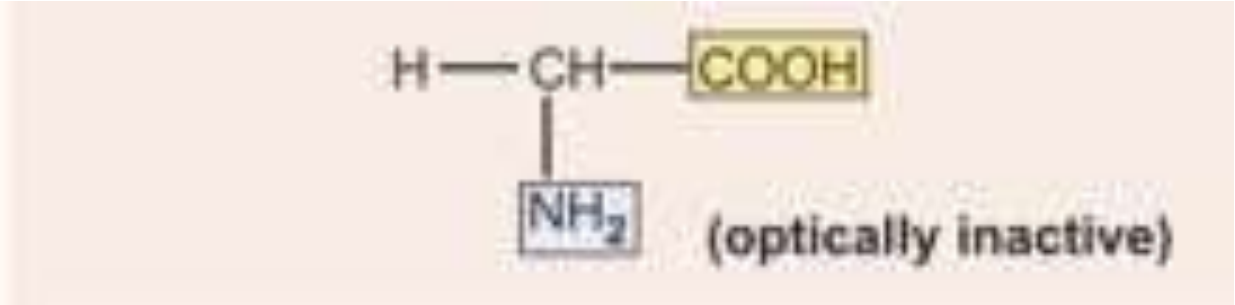
C) Valine - Val - V

D) Leucine - Leu - L

E) Isoleucine - Ile - I

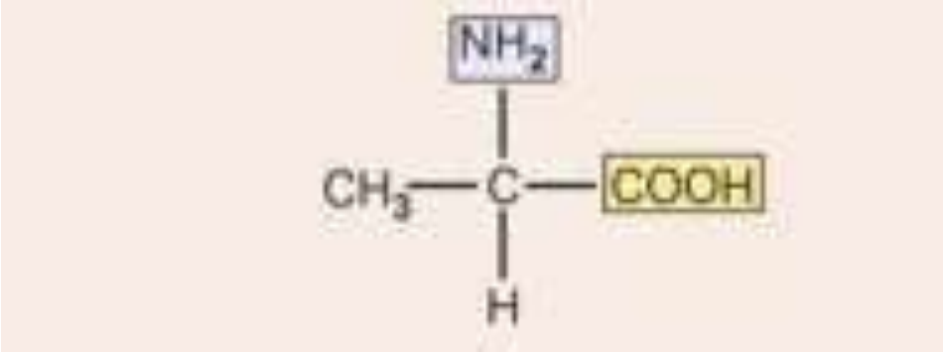
Glycine:

- **Small, simple amino acid. R – group is hydrogen**
- **It is a non essential amino acid.**
- **Glycine is allosteric inhibitor of glutathione synthetase.**



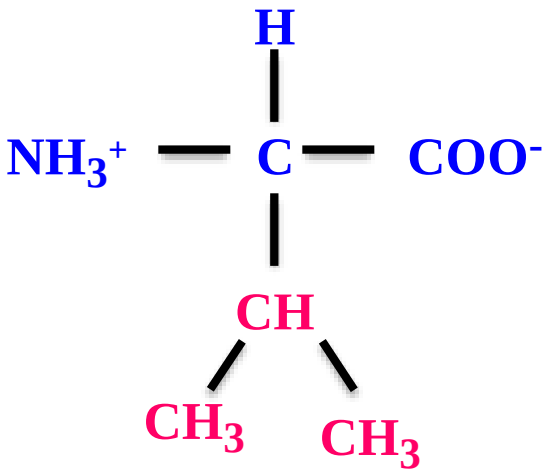
Alanine:

- **It is a non essential amino acid. Alanine is allosteric inhibitor of glutathione synthetase.**
- **D-Alanine: is a component of bacterial cell wall.**
- **B-Alanine is found in pantothenic acid.**



Valine:

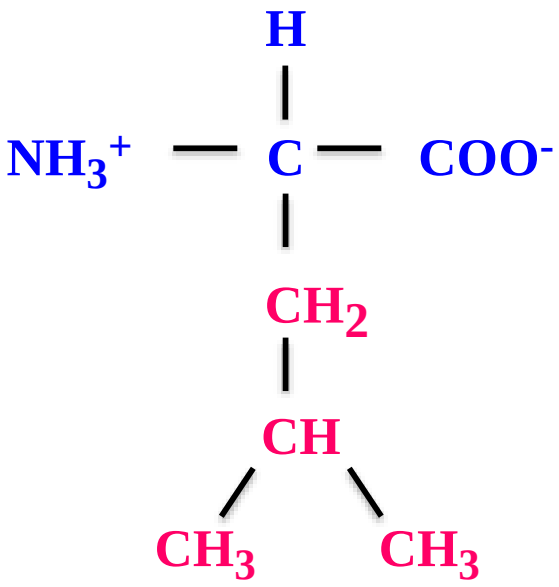
- **It is essential amino acid.**
- **It has an aliphatic hydrophobic isopropyl (3Carbon) side chain.**



Branched chain

Leucine:

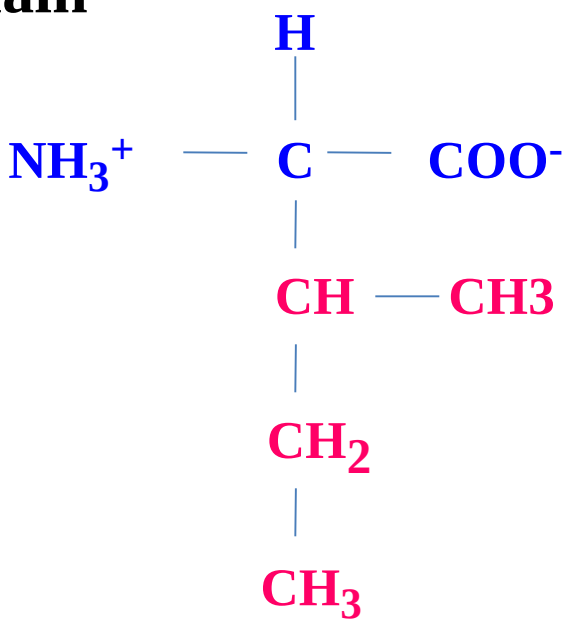
- **It is an essential amino acid.**
- **It has an aliphatic hydrophobic isopropyl (4Carbon) side chain.**



Branched chain

Isoleucine:

- **It is an essential amino acid.**
- **It has an aliphatic hydrophobic isopropyl (4Carbon) side chain**



Branched chain

2. Hydroxyl gr. containing amino acids (-OH)

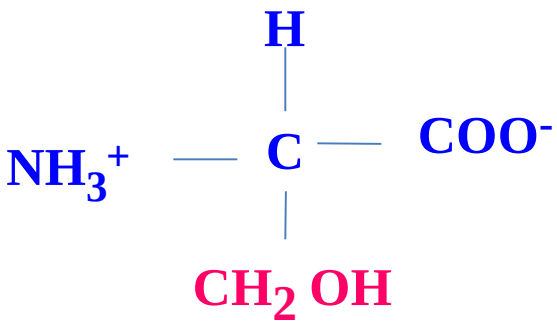
Serine - Ser - S

Threonine - Thr - T

Tyrosine - Tyr - Y

Serine:

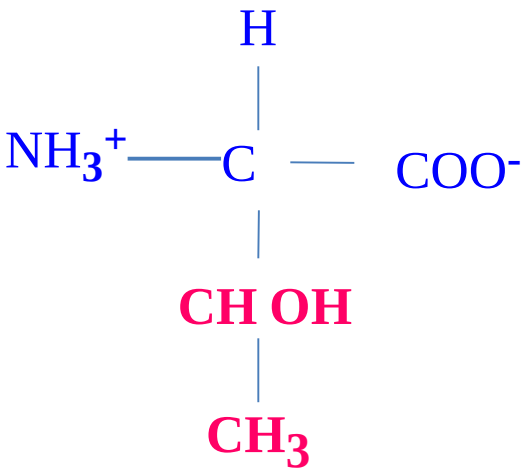
- **It is a non essential amino acid.**
- **It has a alcohol group, is a site for phosphorylation of many proteins.**



Hydroxyl

Threonine:

- **It is a essential amino acid.**
- **Threonine alcohol side group is a target for phosphorylation of proteins.**



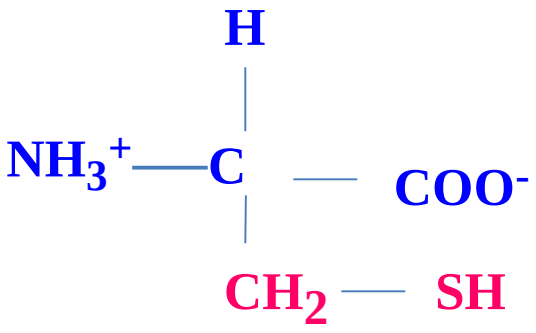
Hydroxyl

3. Sulfur containing amino acids:

- **Cysteine - Cys - C**
- **Methionine - Met - M**

Cysteine:

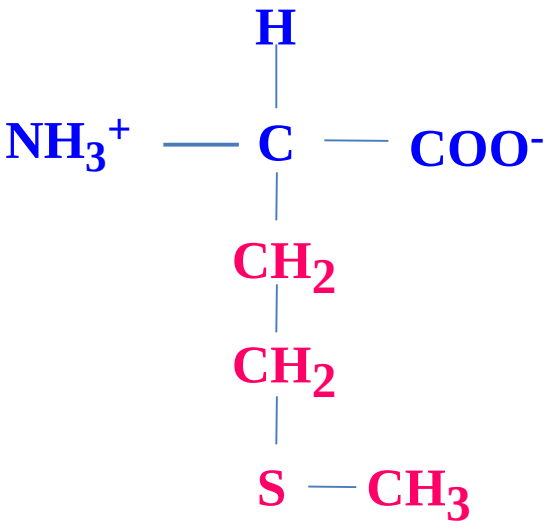
- **It is a non essential amino acid.**
- **Cystine, another important sulfur containing amino acid, is formed by the condensation of two molecules of cysteine.**



Sulfydryl

Methionine:

- **It is an essential amino acid.**
- **In genetic code methionine is coded by codon AUG. This codon is called start codon. Methionine is first amino acid used to build a protein chain.**



Thioether

4. Acidic amino acids and their amides:

Aspartic acid - Asp - D

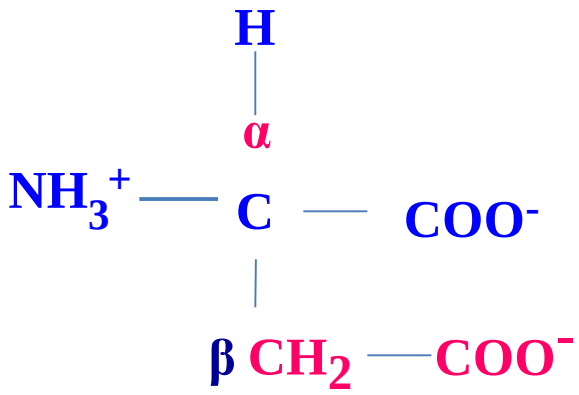
Asparagine - Asn - N

Glutamic acid - Glu - E

Glutamine - Gln - Q

Aspartic acid :

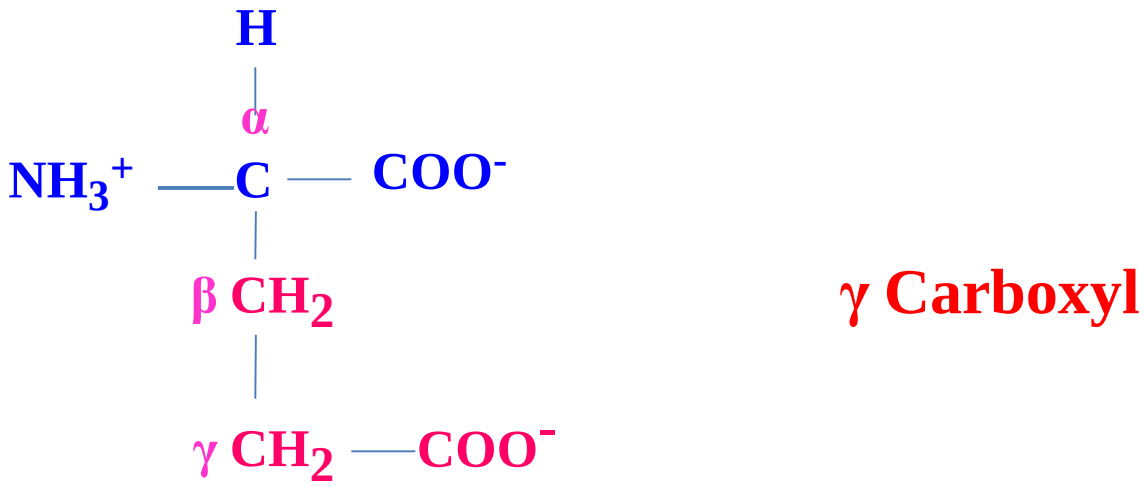
- **It is a non essential amino acid.**
- **Aspartic is capable of forming ionic bonds and involved in chemical reactions.**



β Carboxyl

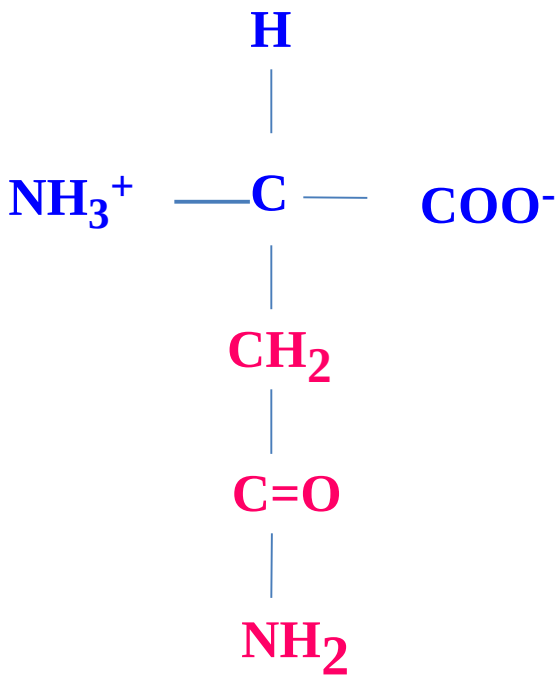
Glutamic acid :

- **It is a non essential amino acid.**
- **Vitamin K2 carboxylates glutamate residues in certain proteins to give carboxy glutamate. This modification allows protein to bind calcium an essential event in blood clotting cascade.**



Asparagine:

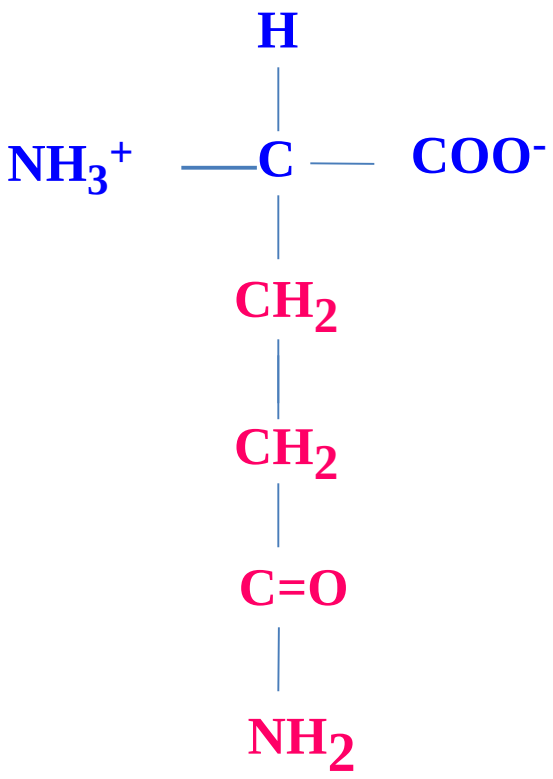
- **It is non essential amino acid.**



Amide

Glutamine :

- **It is a non essential amino acid.**
- **It is very important compound in transamination reactions.**



Amide

5.Basic amino acids:

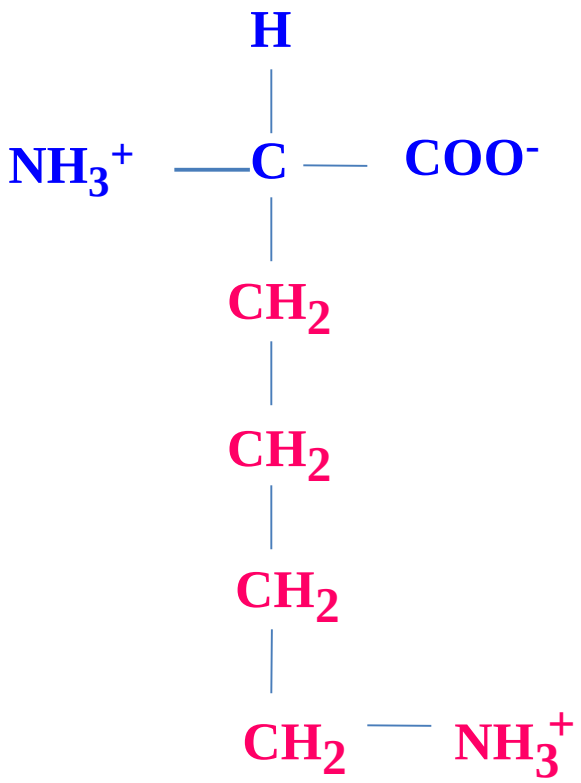
Lysine - Lys - K

Arginine - Arg - R

Histidine - His - H

Lysine :

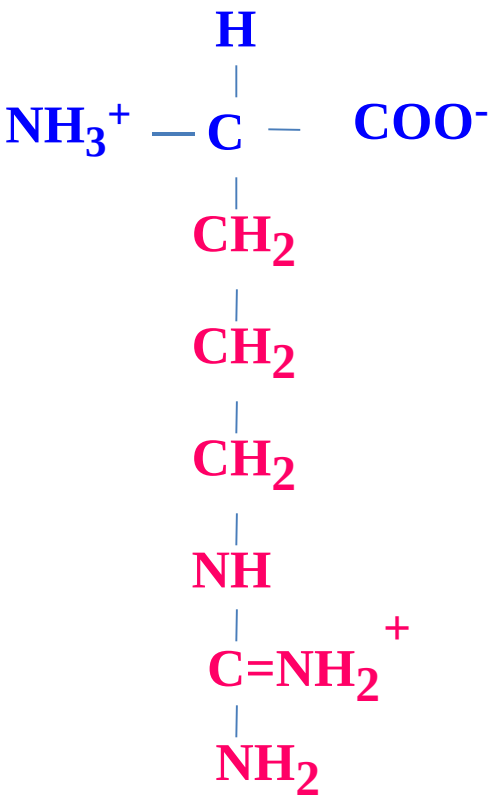
- **It is an essential amino acid. These are strongly polar.**



Amide

Arginine :

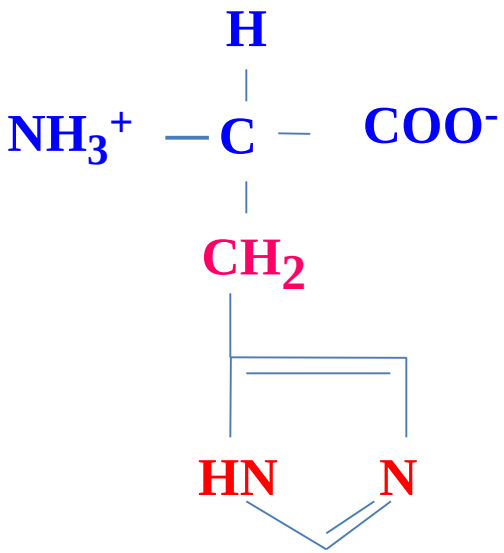
- **It is a semi essential amino acid.**
- **It contain guanidino group and is monocarboxylic acid.**
- **It is an intermediate in urea cycle and is precursor for nitric oxide.**



Guanidinium

Histidine:

- **It is an semi essential amino acid.**
- **It contains imidazole ring.**



Imidazole

6. Aromatic amino acids:

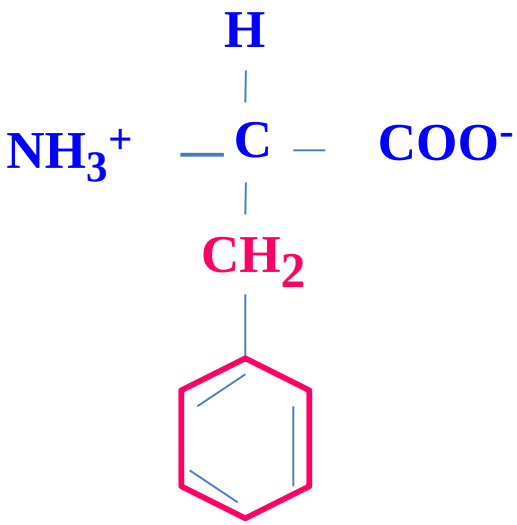
Phenyl alanine - Phe - F

Tyrosine - Tyr - Y

Tryptophan - Trp - W

Phenyl alanine :

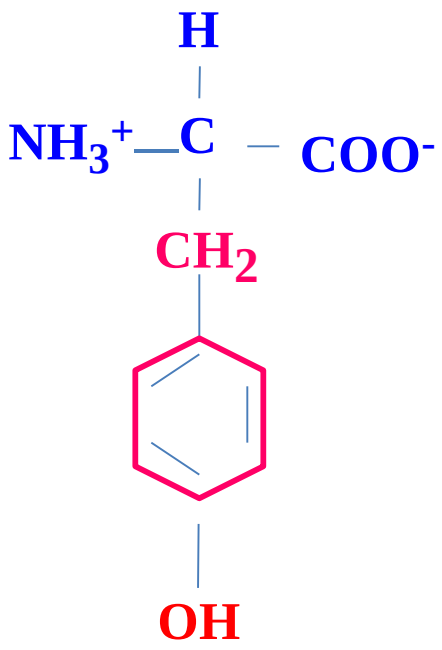
- It is an essential amino acid.
- It contains benzene ring.



Benzene Ring

Tyrosine :

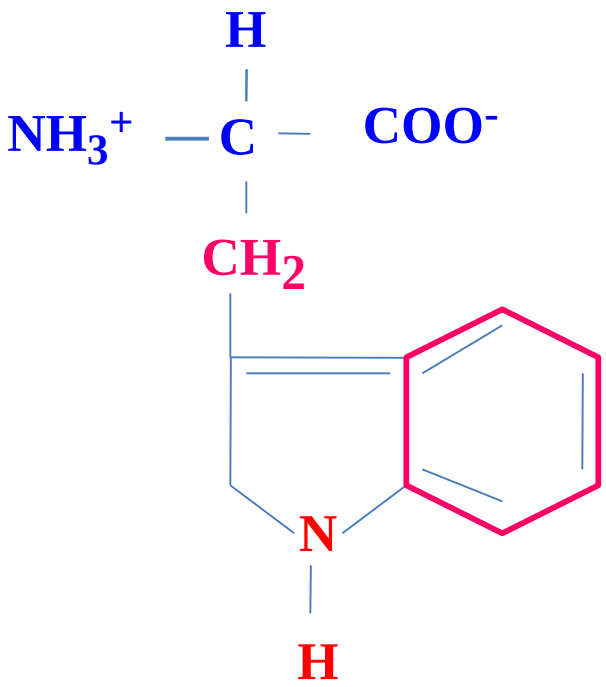
- **It is a non essential amino acid.**
- **The hydroxyl group of tyrosine imparts slight polarity to the side chain and is a site for phosphorylation in proteins.**



Phenol

Tryptophan :

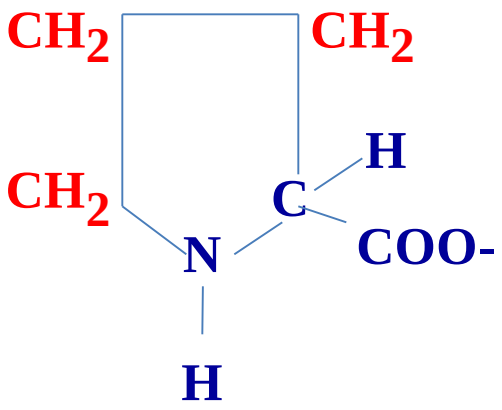
- **It is a essential amino acid.**
- **It contains indole ring.**



Indole

7.Imino acids :

- **Proline containing pyrrolidine ring is a unique amino acid.**
- **It has an imino group (=NH). Threrefore, proline is an α-imino acid.**



Pyrroline

Classification of amino acids based on polarity:

Amino acids are classified into 4 groups based on their polarity.

1. Non – polar amino acids:

- **These amino acids are also referred to as hydrophobic.**
- **They have no change on R-group.**

e.g. glycine, alanine, leucine, isoleucine, valine, methionine, phenyl alanine, 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.**

e.g., glycine, serine, threonine, cysteine, glutamine, asparagine and tyrosine.

3. Polar amino acids with positive R- group:

The three amino acids

- **lysine,**
- **arginine**
- **histidine**

4. Polar amino acids with negative R- group:

- **The dicarboxylic mono amino acids – aspartic acid and glutamic acid are considered in this group.**

3. Nutritional classification of amino acids:

A) Essential or indispensable amino acids:

- **The amino acids which cannot be synthesized by the body and need to be supplied through the diet are called essential amino acids.**
- **They are required for proper growth and maintenance of the individual**

The 10 essential amino acids are

1. Arginine -Semi essential
2. Valine
3. Histidine -Semi essential
4. Isoleucine
5. Leucine
6. Lysine
7. Methionine
8. Phenyl alanine
9. Threonine
10. Tryptophan

AVHILLMPTT

B) Non – essential 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, asparagine, glutamate, glutamine, tyrosine and proline.**

Based on their metabolic fate:

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

1. Glucogenic amino acids:

- **These amino acids can serve as precursor 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 ketogenic.

3.Glucogenic and ketogenic Amino acids

The four amino acids

- Phenyl alanine
- Isoleucine
- Tryptophan
- Tyrosine

Are precursors for the synthesis of glucose as well as fat

New amino acids:

- **In addition to 20 L – amino acids that take part in protein synthesis, recently two more new amino acids are described. They are**

1. **Selenocysteine – 21st amino acid**
2. **Pyrrolysine – 22 nd amino acid**

1. Selenocysteine

occurs at the active site of several enzymes.

e.g., Thioredoxin reductase

- **Glutathione peroxidase**
- **De – iodinase**
- **Glycine reductase**

- **Selenoprotein P, a glycoprotein containing 10 selenocysteine residues, found in mammalian blood. It has an antioxidant function and its concentration falls in selenium deficiency.**
- **The stop codon UGA can code for Selenocysteine**
- **Selenocysteine is enzymatically generated from serine directly on the t RNA and then incorporated into proteins.**

2. Pyrrolysine :

The STOP codon UAG can code for pyrrolysine.

Amino acids
(fully protonated)

are composed of

- **α-Carboxyl group (–COOH)** → *is* → deprotonated (COO^-) at physiologic pH
- **α-Amino group ($–NH_3^+$)** → *is* → protonated (NH_3^+) at physiologic pH
- **Side chains of 20 different types** → *composed of*

Nonpolar side chains	Uncharged polar side chains	Acidic side chains	Basic side chains
Alanine	Asparagine	Aspartic acid	Arginine
Glycine	Cysteine	Glutamic acid	Histidine
Isoleucine	Glutamine		Lysine
Leucine	Serine		
Methionine	Threonine		
Phenylalanine	Tyrosine		
Proline			
Tryptophan			
Valine			

found → in the interior of proteins that function in an aqueous environment and on the surface of proteins (such as membrane proteins) that interact with lipids

found → on the outside of proteins that function in an aqueous environment and in the interior of membrane-associated proteins

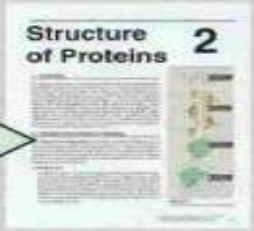
Acidic side chains *characterized by* → Side chain dissociates to $–COO^-$ at physiologic pH → *found* → on the outside of proteins that function in an aqueous environment and in the interior of membrane-associated proteins

Basic side chains *characterized by* → Side chain is protonated and generally has a positive charge at physiologic pH → *found* → on the outside of proteins that function in an aqueous environment and in the interior of membrane-associated proteins

can

- Release H^+
- *and act as*
- Weak acids
- *described by*
- Henderson-Hasselbalch equation: $pH = pK_a + \log \frac{A^-}{HA}$
- *predicts*
- Buffering capacity
- *predicts*
- Buffering occurs ± 1 pH unit of pK_a
- *predicts*
- Maximal buffer when $pH = pK_a$
- *predicts*
- $pH = pK_a$ when $[HA] = [A^-]$

In proteins, most α-COO^- and α-NH_3^+ of amino acids are combined in peptide bonds. → Therefore, these groups are not available for chemical reaction. → Thus, the chemical nature of the side chain determines the role that the amino acid plays in a protein, particularly . . . → . . . how the protein folds into its native conformation.



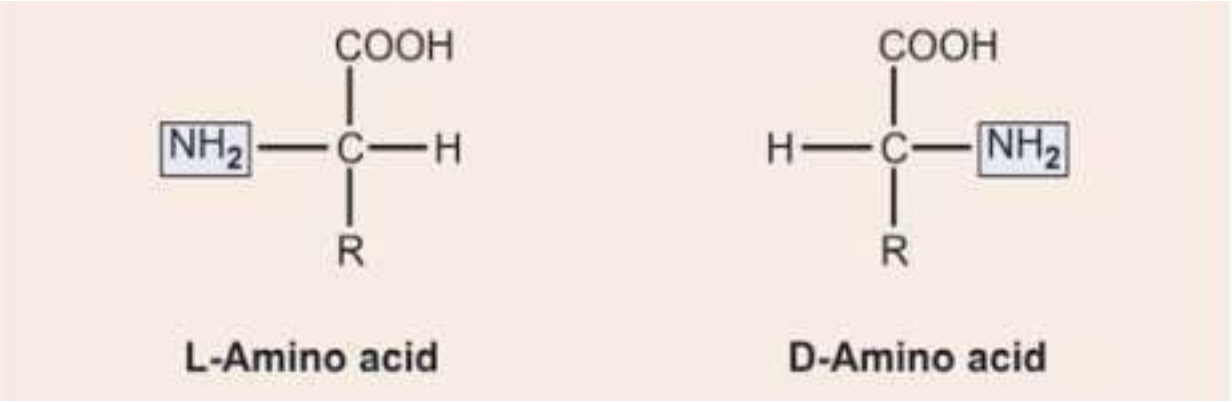
Properties of amino acids:

Physical properties:

- **Solubility: Most of the amino acids are soluble in water and insoluble in organic solvents.**
- **Melting points :Amino acids generally melt at higher temperatures, often above 200ºC.**

- **Taste : Amino acids may be sweet(Gly, Ala, Val), tasteless (Leu) or bitter (Arg, Ile).**
- **Monosodium glutamate is a salt of glutamic acid.**
- **It is employed as a flavoring agent in food industry to increase taste and flavor.**

Optical properties : **All amino acids except glycine possess optical isomers due to the presence of asymmetric carbon atom.**



Amino acids as ampholytes : **Amino acids contain both acidic (-COOH) and basic (NH_2) groups.**

- **They can donate a proton or accept a proton, hence amino acids are regarded as ampholytes.**

Zwitter ion or dipolar ion :

- **Zwitter ion is a hybrid molecule containing positive and negative ionic groups.**
- **The amino acids rarely exist in a neutral form with free carboxylic and free amino groups.**
- **In strongly acidic pH, the amino acid is positively charged (cation)**
- **In strongly alkaline pH, the amino acid is negatively charged (anion)**

- **Each amino acid has a characteristic pH at which it carries both positive and negative charges and exists as zwitter ion.**

- ***Isoelectric pH*** **is defined as the pH at which a molecule exists as a zwitter ion or dipolar ion and carries no net charge.**

Chemical properties:

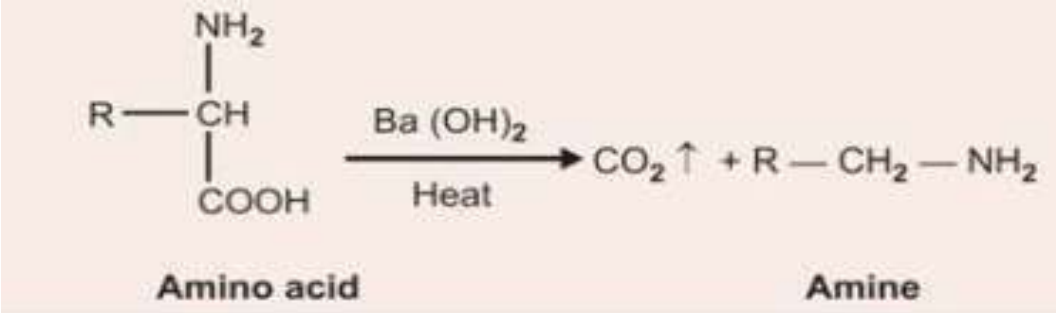
The general reactions of amino acids are mostly due to the presence of two functional groups namely carboxylic (-COOH) and amino (-NH_2) group.

- **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.**

- **e.g., histamine, tyramine, ɣ-amino butyric acid (GABA) from the amino acids histidine, tyrosine and glutamate respectively.**



3. Reaction with ammonia: **The carboxyl group of dicarboxylic amino acids react with NH_3 to form amide.**

- **Aspartic acid + NH_3 ⟶ Asparagine**
- **Glutamic acid + NH_3 ⟶ Glutamine**

Reactions due to $-NH_2$ group:

The amino groups behave as bases and combine with acids to form salts.

5.Reaction with ninhydrin:

- **The α-amino acids react with ninhydrin to form a purple, blue or pink colour complex(Ruhemanns`s purple).**
- **Amino acid + Ninhydrin** $\longrightarrow$ **Keto acid +** NH_3 **+** CO_2 **+ Hydrindantin**

  Hydrindantin + NH_3 **+ Ninhydrin** $\longrightarrow$ **Ruhemanns`s purple**

Colour reactions of Amino acids

Amino acids can be identified by specific Colour reactions

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 a.a.
4.Millons 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 group (Tyr)

7.Transamination:

- **Transfer of an amino group from an amino acid to a keto acid to form a new amino acid is known as transamination.**

8.Oxidative deamination:

- **The amino acids undergo oxidative deamination to liberate free ammonia.**

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 modifications after the protein synthesis occurs. These derivatives of amino acids are very important for protein structure and functions.**
- **Collagen-the most abundant protein in mammals-contains 4-hydroxyproline and 5-hydroxy lysine.**

- **Histones-the proteins found in association with DNA –contain many methylated, phosphorylated or acetylated amino acids.**
- **ɣ-Carboxyglutamic acid is found in certain plasma proteins involved in blood clotting.**
- **Cystine is formed by combination of two cysteines. It is also a derived amino acid.**

Non-Protein amino acids

α-Amino acids	Functions
1.Ornithine	Intermediate in the biosynthesis of urea
2.Citrulline	Intermediate in the biosynthesis of urea
3.Arginosuccinic acid	Intermediate in the biosynthesis of urea
4.Thyroxine	Thyroid hormone derived from tyrosine.
5.Triiodothyronine	Thyroid hormone derived from tyrosine
6.SAM	Methyl donor in biological system.
7.Homocysteine	Intermediate in methionine metabolism. A risk factor for coronary heart diseases.
8. 3,4-Dihydroxy phenyl alanine (DOPA)	A neurotransmitter, precursor for melanin
9.Creatinine	Derived from muscle and excreted in urine.

Non – α-amino acids:

Amino acids	Functions
1. β-Alanine	Component of pantothenic acid and coenzyme A.
2. β-Aminoisobutyric acid	End product of pyrimidine metabolism.
3. ɣ-Aminobutyric acid	A neurotransmitter produced from glutamic acid
4. δ-Amino levulinic acid	Intermediate in heme synthesis
5. Taurine	Found in association with bile acids

D-Amino acids:

- **Certain D-amino acids are also found in the antibiotics (Actinomycin-D, valinomycin, gramicidin-S).**
- **D-Glutamic acid and D-Alanine are present in bacterial cell wall.**
- **D-Serine and D-Aspartate are found in brain tissue.**

Peptides and Proteins

- 20 amino acids are commonly found in protein.
- These 20 amino acids are linked together through "peptide bond forming peptides and proteins.
- The chains containing less than 50 amino acids are called "peptides", while those containing greater than 50 amino acids are called proteins".

Peptide bond formation:

- α-carboxyl group of one amino acid (with side chain R1) forms a covalent peptide bond with α-amino group of another amino acid (with the side chain R2) by removal of a molecule of water.
- The result is : Dipeptide (i.e. Two amino acids linked by one peptide bond).

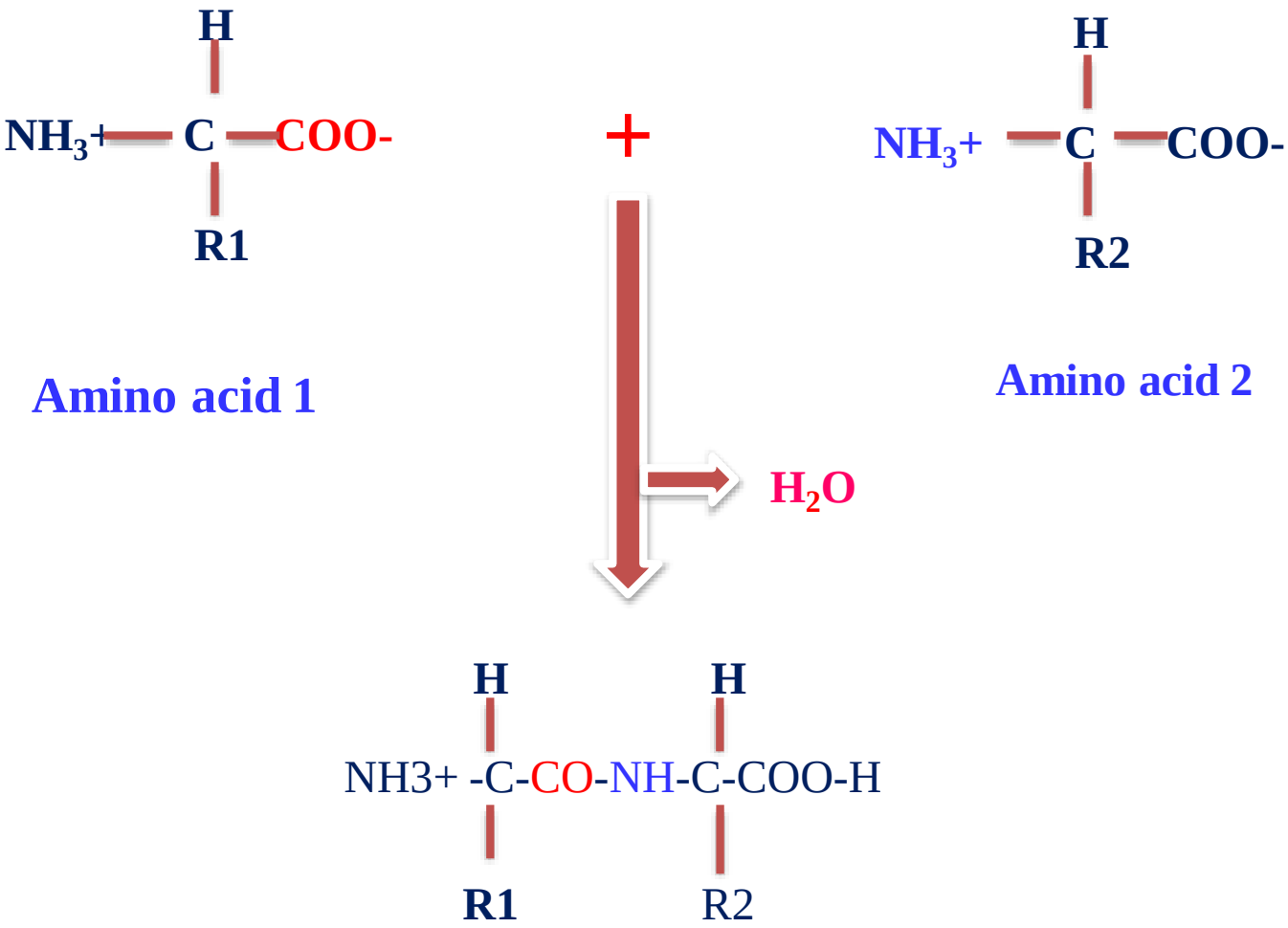
- The dipeptide can then forms a second peptide bond with a third amino acid (with side chain R3) to give Tripeptide.
- Repetition of this process generates a polypeptide or protein of specific amino acid sequence.

- Each polypeptide chain starts on the left side by free amino group of the first amino acid enter in chain formation.
- It is termed (N- terminus).
- Each polypeptide chain ends on the right side by free COOH group of the last amino acid and termed (C-terminus).

Peptide bond

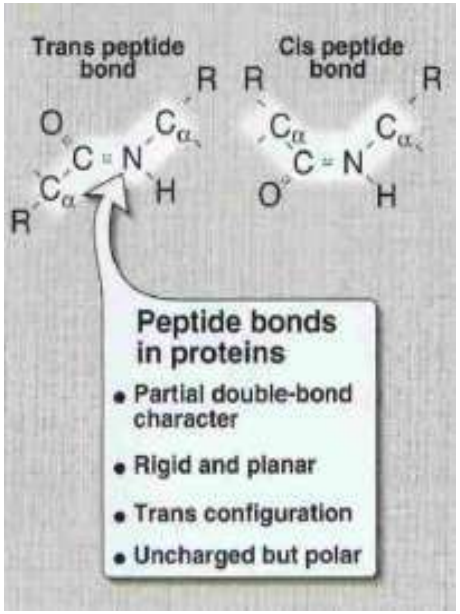
- The amino acids are held together in a protein by covalent peptide bonds or linkage cementing material between amino acids
- When amino group of an amino acid combines with the carboxyl group of another amino acid, a peptide bond is formed
- Peptides containing more than 10 amino acids are referred as polypeptide

Formation of peptide bond



Functions of peptide bond

- It usually found in trans configuration.
- The peptide bond is a partial double bond.
- N- partially positive, O- partially negative.
- Shorthand 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 peptide.



Naming of peptides

- The amino acid suffixes –ine (glycine), -an (tryptophan), -ate (glutamate) are changed to –yl with the exception of C-terminal amino acid

- **E.g, Glutamyl-cysteinyl-glycine**

$+H_3N$ –glutamate - cysteine - glycine - COO^- **Amino acids**

in a peptide

E - C - G **One letter symbols**

Glu - Cys - Gly **Three letter symbols**

Glutamyl - cysteinyl - glycine **Peptide name**

Examples of Peptides

- **Dipeptide**: Two amino acids joined by one peptide bond.

  E.g. Aspartame which acts as sweetening agent being used in replacement of cane sugar.

- It is composed of aspartic acid and phenyl alanine.
- **Tripeptides:** 3 amino acids linked by two peptide bonds.

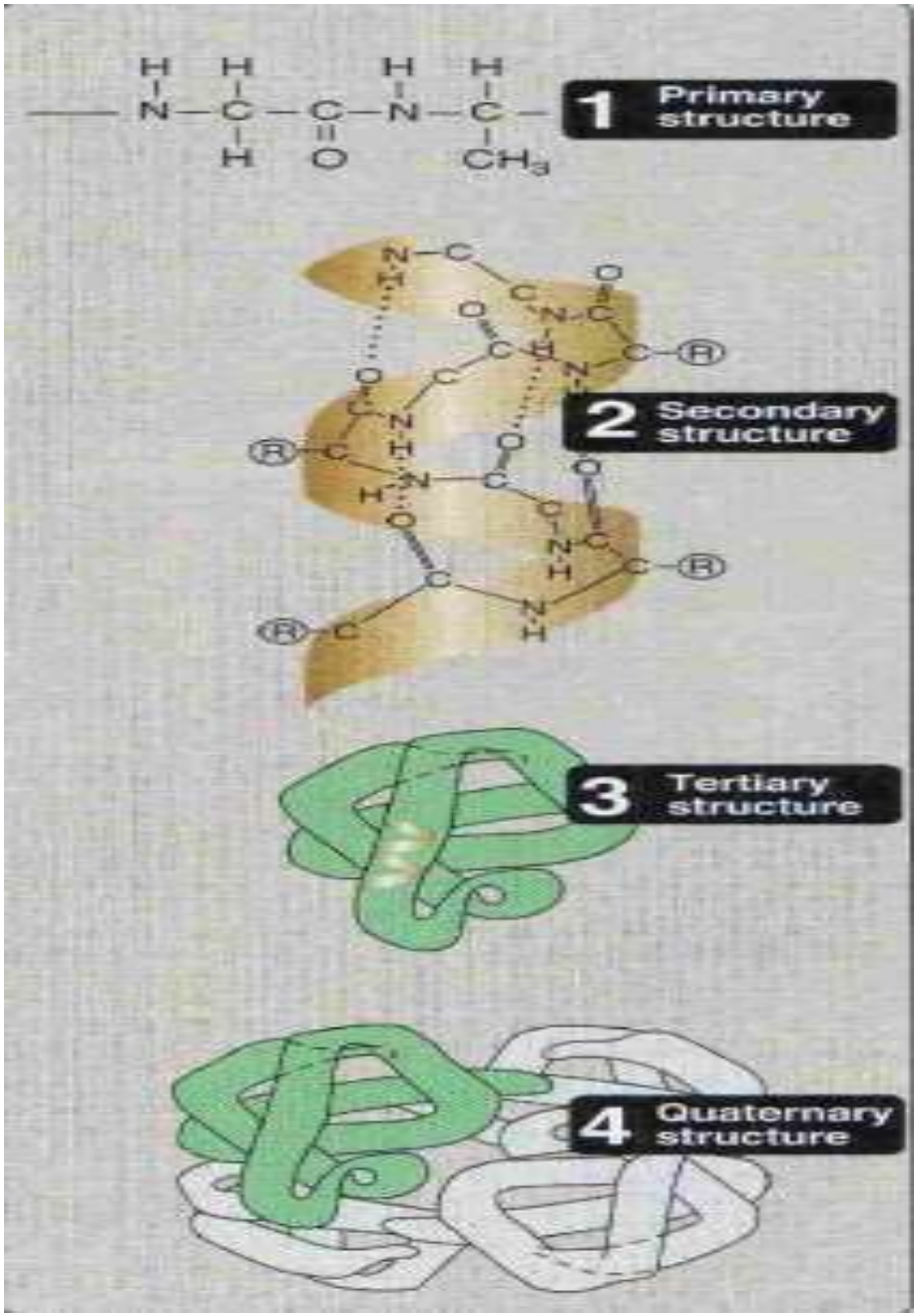
- e.g. GSH which is formed from 3 amino acids: glutamic acid, cysteine and glycine.
- It helps in absorption of amino acids, protects against hemolysis of RBC by breaking H_2O_2 which causes cell damage.
- **Polypeptides**: 10- 50 amino acids.
- E.g. Insulin, a polypeptide hormone.

Structure of proteins

- Proteins have different level of organization;
- Primary structure: The linear sequence of amino acids forming the backbone of proteins.
- Secondary structure: The spatial arrangement of protein by twisting of the polypeptide chain.
- Tertiary structure: The three dimensional structure of a functional protein.

- 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.

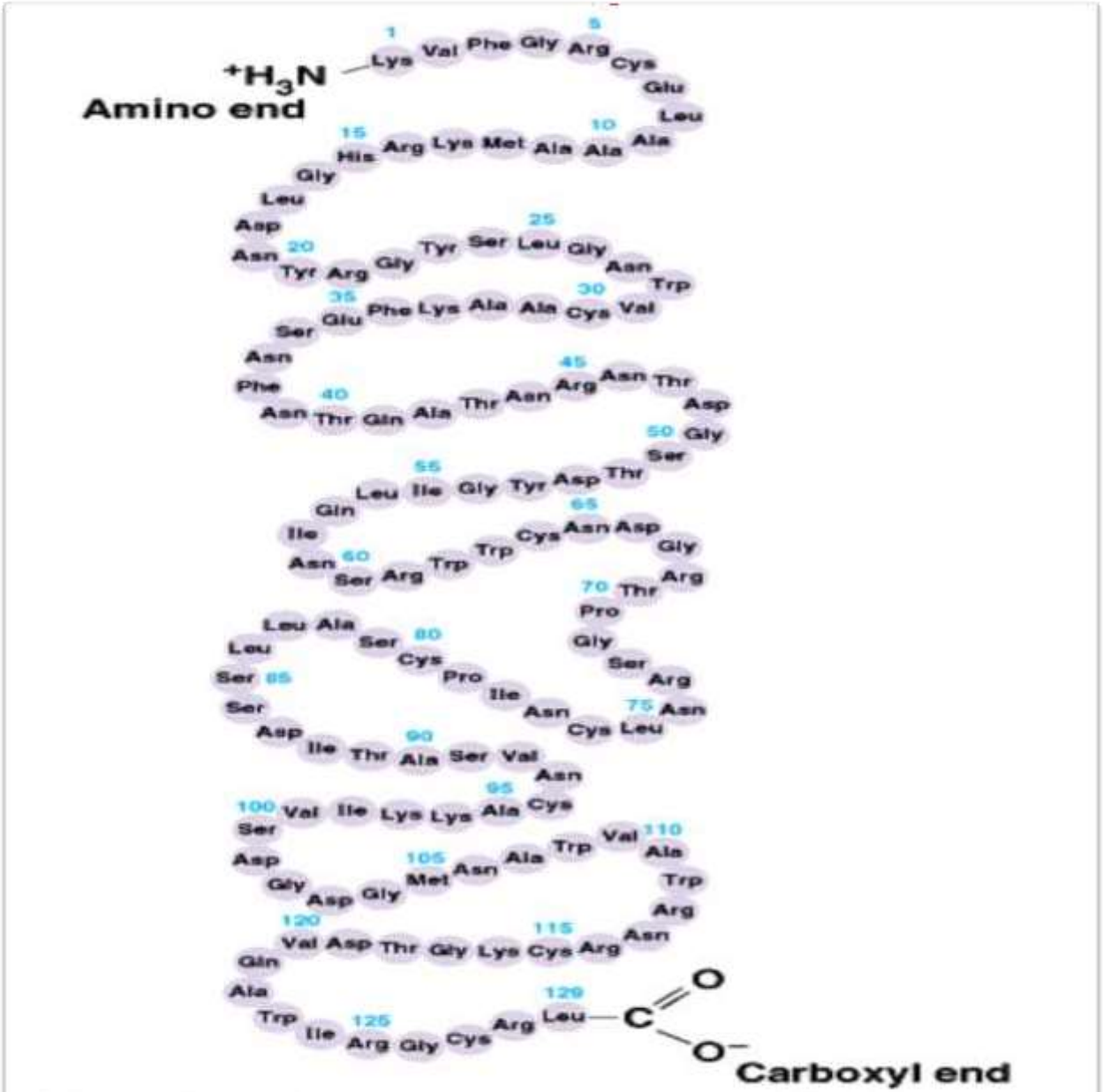


Structure of proteins

Primary structure: The linear sequence of amino acids held together by peptide bonds in its peptide chain.

- The peptide bonds form the back bone.
- The free $-NH_2$ group of the terminal amino acid is called as N-terminal end and the free –COOH end is called C-terminal end.

- It is a tradition to number the amino acids from N-terminal end as No.1 towards the C-terminal end.
- Presence of specific amino acids at a specific number is very significant for a particular function of a protein.



Determination of primary structure

- The primary structure comprises the identification of amino acids with regard to their quality, quantity and sequence in a protein structure.

Determination of primary structure involves 3 stages:

- Determination of amino acid composition.
- Degradation of protein or polypeptide into smaller fragments.
- Determination of the amino acid sequence

1. Determination of amino acid composition

- The protein or polypeptide is completely hydrolyzed to liberate the amino acids
- The hydrolysis may be carried out either by acid or alkali treatment or by enzyme hydrolysis.
- *Pronase* is a mixture of non-specific proteolytic enzymes that causes complete hydrolysis of proteins.

Separation and estimation of amino acids

- The mixture of amino acids liberated by protein hydrolysis can be determined by chromatographic techniques.

2. Degradation of protein or polypeptide into smaller fragments.

Liberation of polypeptides

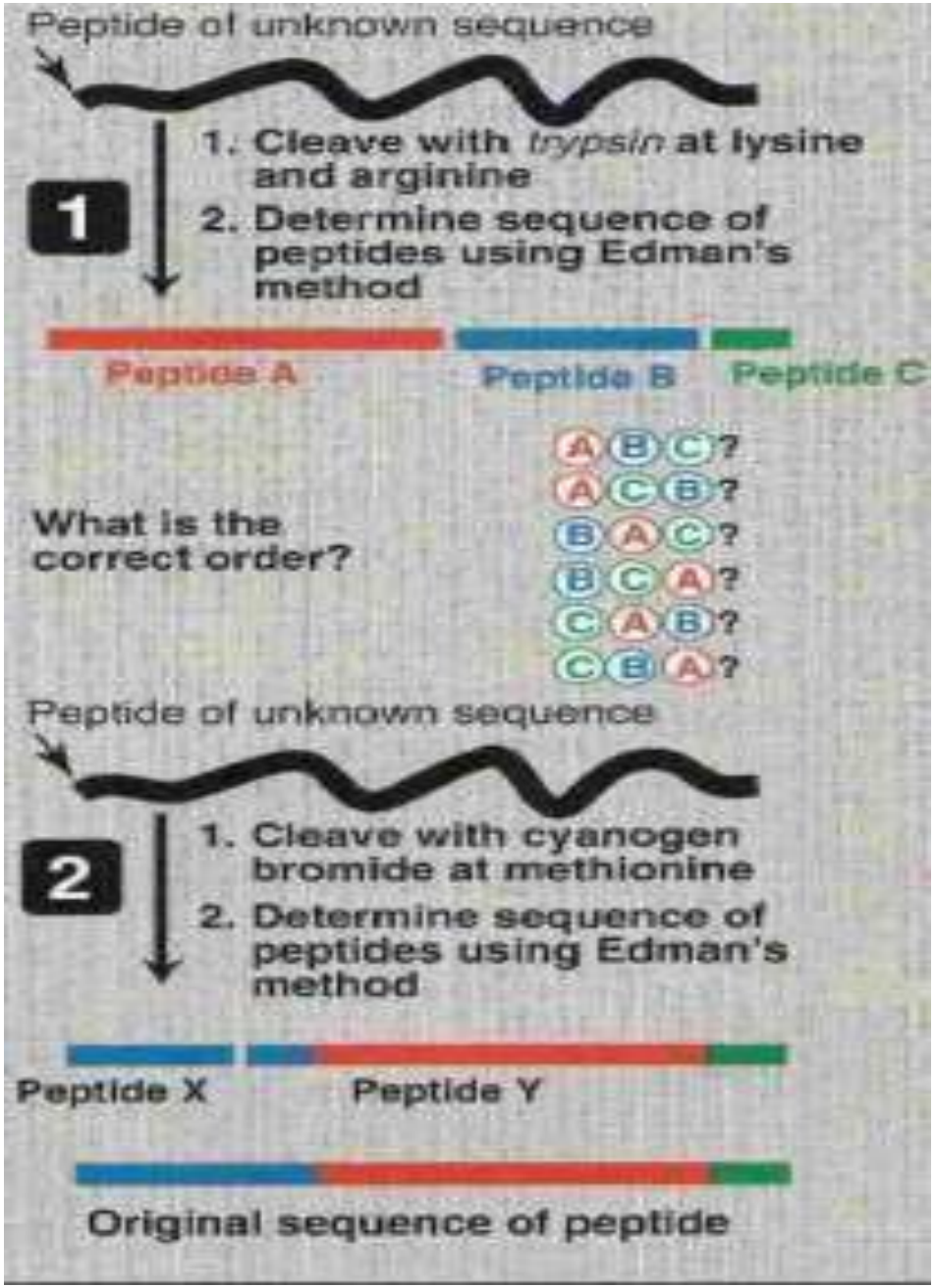
- Treatment with urea or guanidine hydrochloride disrupts the non-covalent bonds and dissociates the protein into polypeptide units.
- Cleaving the disulfide linkages between the polypeptide units, treatment with performic acid is necessary.

- Number of polypeptides: The number of polypeptide chains can be identified by treatment of protein with *dansyl chloride*.
- It specifically binds with N-terminal amino acids to form dansyl polypeptides which on hydrolysis yield N-terminal dansyl amino acid.
- The number of dansyl amino acids produced is equal to the number of polypeptide chains in a protein.

Breakdown of polypeptides into fragments

- Polypeptides are degraded into smaller peptides by enzymatic or chemical methods
- Enzymatic cleavage: The proteolytic enzymes such as *trypsin, chymotrypsin, pepsin and elastage* exhibit specificity in cleaving the peptide bonds.
- Among these enzymes, trypsin is commonly used.

- Trypsin hydrolyses the peptide bonds containing lysine or arginine on the carbonyl (-C=O) side of peptide linkage.
- Chemical cleavage: Cyanogen bromide (CNBr) is commonly used to split polypeptides into smaller fragments.
- CNBr specifically splits peptide bonds, the carbonyl side of which is contributed by the amino acid methionine.

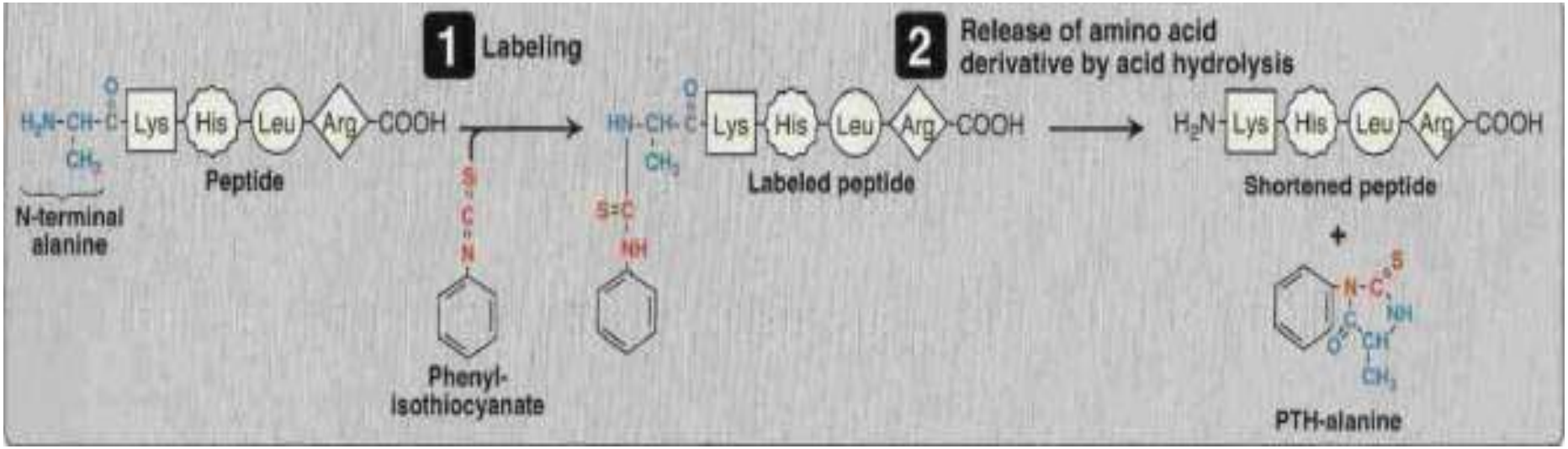


Determination of the amino acid sequence

- **Sanger's Reagent:** Sanger used 1-fluoro-2, 4-dinitrobenzene (FDNB) to determine insulin structure.
- FDNB specifically binds with N-terminal amino acids to form a dinitrophenyl (DNP) derivative of peptide.
- This, on hydrolysis yields DNP – amino acids (N-terminal) and free amino acids from the rest of the peptide chain.
- DNP amino acids can be identified by chromatography.

- Edmans reagent: Phenyl isothiocyanate is Edmans reagent.
- It reacts with the N-terminal amino acid of peptide to form a phenyl thiocarbamyl derivativve.
- On treatment with mild acid, phenyl thiohydrantoin (PTH) –amino acid, a cyclic compound is liberated.

- PTH amino acid can be identified by chromatography
- Edmans reagent has an advantage, a peptide can be sequentially degraded liberating N-terminal amino acids one after another which can be identified.
- This is due to the fact that the peptide as a whole is not hydrolyzed but only releases PTH- amino acids.



Sequenator

- This is an automatic machine to determine the amino acid sequence in a polypeptide
- It is based on the principle of Edman's degradation .
- Amino acids are determined sequentially from N-terminal end
- The PTH-amino acid liberated is identified by HPLC.
- Sequenator takes about 2 hours to determine each amino acid.

Overlapping peptides

- In the determination of primary structure of protein, several methods are simultaneously employed.
- This results in the formation of overlapping peptides
- Overlapping peptides are very useful in determining the amino acid sequence.

Reverse sequencing technique

- It is the genetic material (DNA) which ultimately determines the sequence of amino acids in a polypeptide chain
- By analyzing the nucleotide sequence of DNA that codes for protein, it is possible to translate the nucleotide sequence into amino acid sequence.
- This technique, however, fails to identify the disulfide bonds and changes that occur in the amino acids after the protein is synthesized.

Secondary structure

- The spatial arrangement of protein by twisting of the polypeptide chain
- The amino acids are located close to each other in their sequence.
- Two chief types of secondary structures, α-Helix and β-sheet

α-Helix

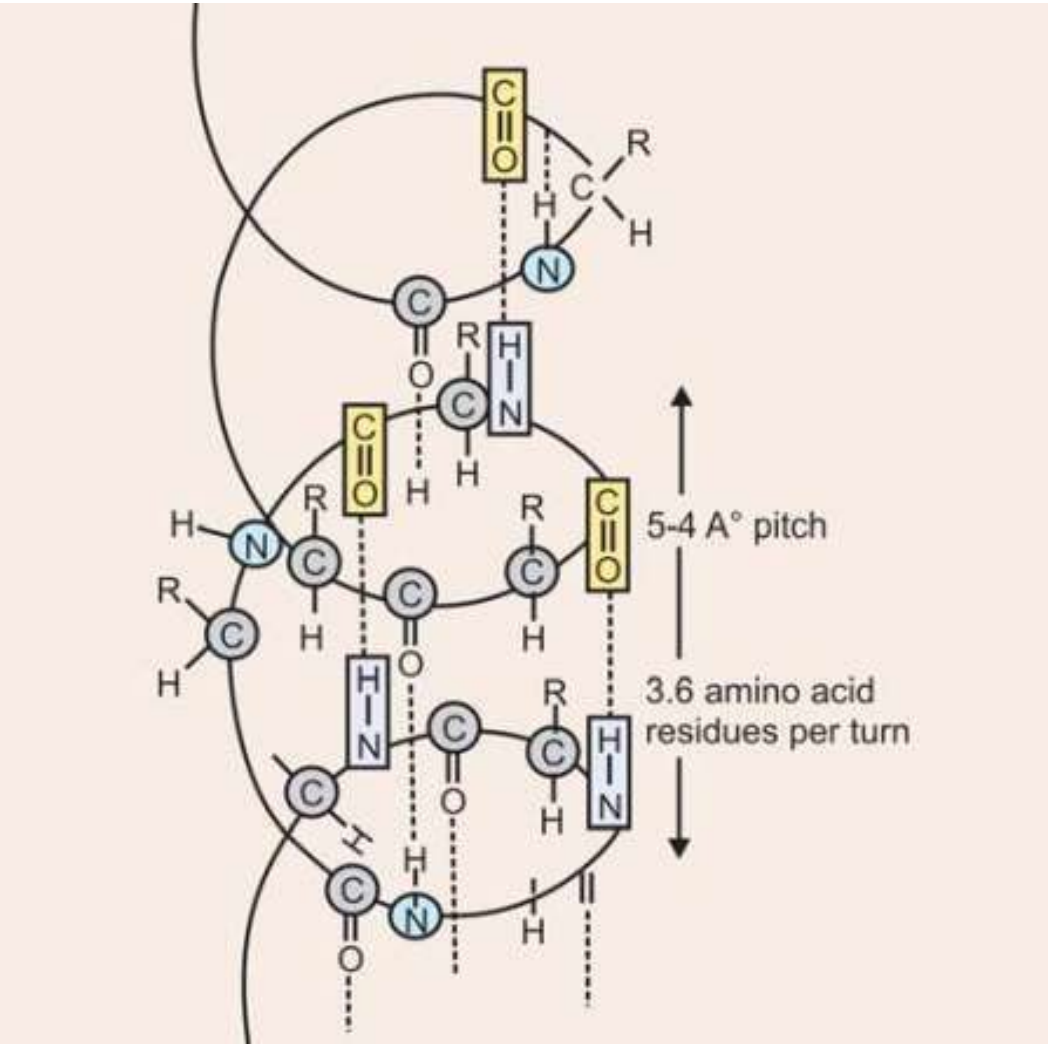
- α-Helix is the most common spiral structure of protein
- α-Helical structure was proposed by Pauling and Corey in 1965.

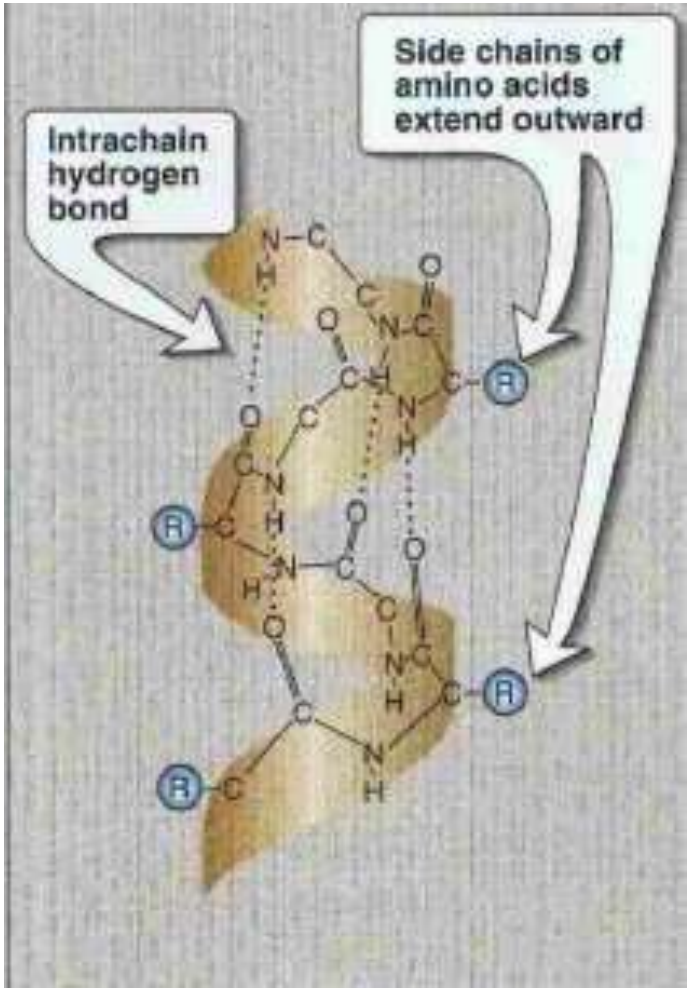
The salient feature:

- The α-Helix is a tightly packed coiled structure with amino acid side chains extending outward from the central axis.

- The α-Helix is stabilized by extensive hydrogen bonding
- Formed between H atom attached to peptide N, and O atom attached to peptide C.
- The H-bonds are individually weak but collectively, strong enough to stabilize the helix.
- All the peptide bonds, except the first and last in polypeptide chain, participate in hydrogen bonding.

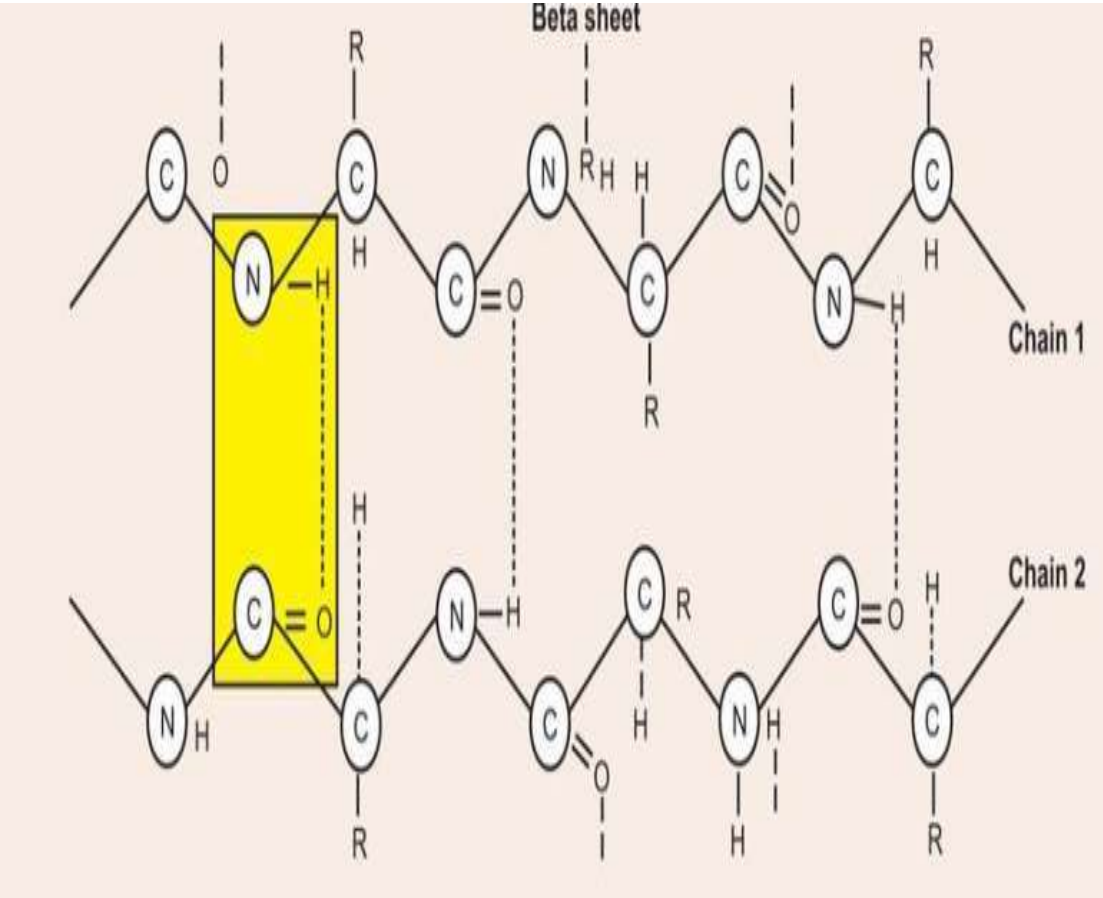
- Each turn of helix contains 3.6 amino acids and travels a distance of 0.54 nm.
- The spacing of each amino acid is 0.15 nm.
- α-Helix is a stable conformation formed spontaneously with the lowest energy.
- The right handed α-Helix is more stable than the left handed helix
- Certain amino acids (particularly proline) disrupt the α-Helix.
- Large number of acidic and basic amino acids are also interfere with α-Helix structure.

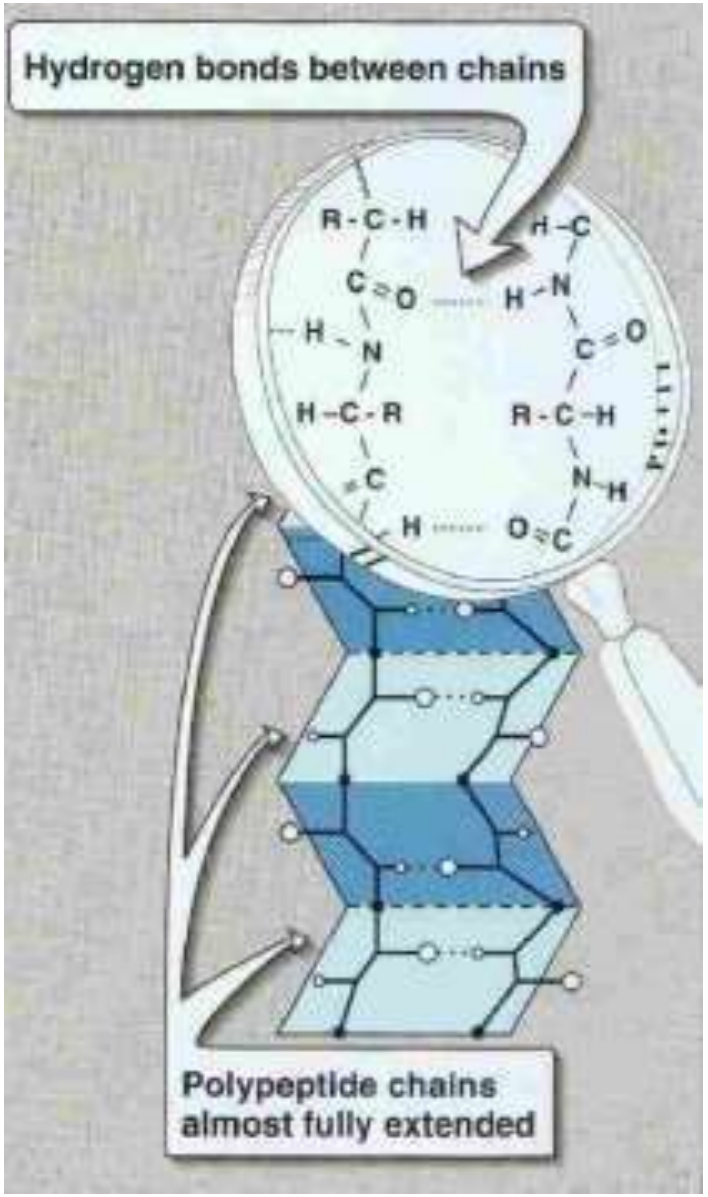

5-4 A° pitch
3.6 amino acid residues per turn


Intrachain hydrogen bond
Side chains of amino acids extend outward

β-pleated sheet

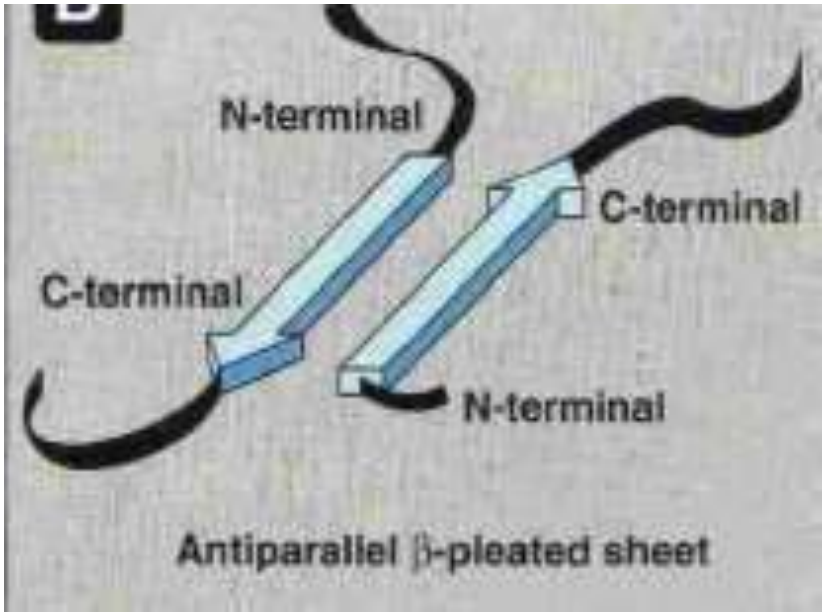
- An another form of secondary structure in which ***two or more polypeptides*** (or segments of the same peptide chain) are linked together by hydrogen bond between H- of NH- of one chain and carbonyl oxygen of adjacent chain (or segment).
- β-Pleated sheet may be formed either by separate polypeptide chains (H-bonds are inter chain) or a single polypeptide chain folding back on to itself (H-bonds are intrachain)

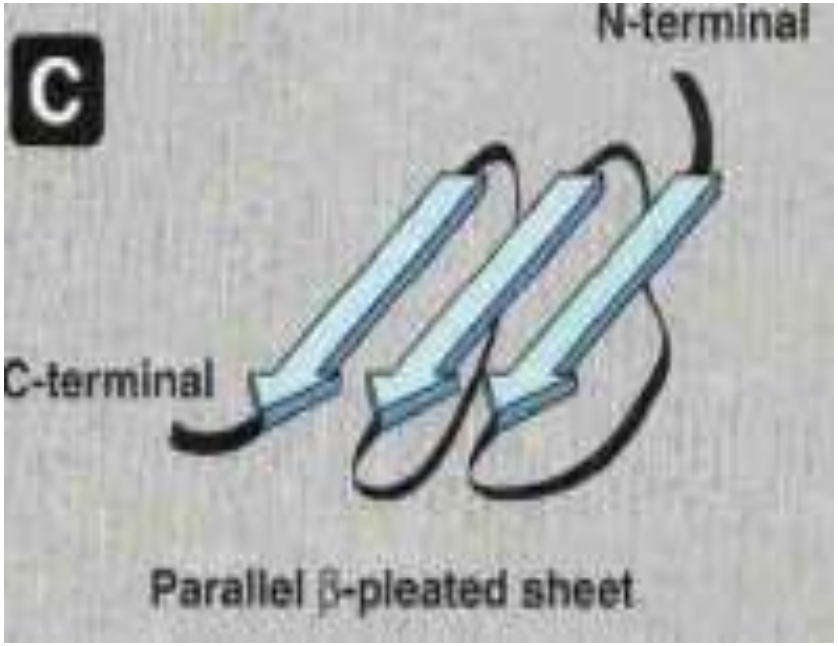

Beta sheet
Chain 1
Chain 2


Hydrogen bonds between chains
Polypeptide chains
almost fully extended

Parallel and anti-parallel β-sheets

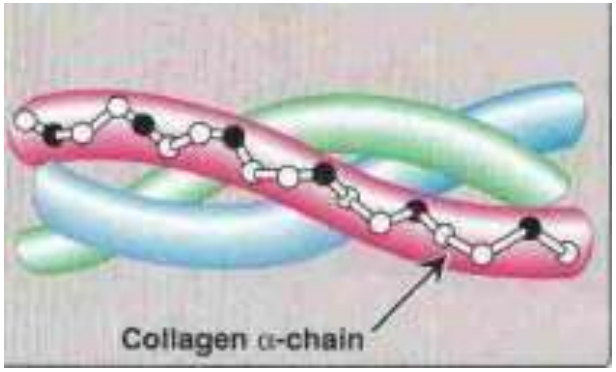
The polypeptide chains in the β-sheets may be arranged either in parallel or anti-parallel direction.





Triple helix

- Collagen is rich in proline and hydroxy proline and cannot form a α-helix or β-Pleated sheet.
- Collagen forms a triple helix.
- The triple helix is stabilized by both non covalent as well as covalent bonds.



Reverse Turns or β-bends

- Since the polypeptide chain of a globular protein changes direction two or more times when it folds, the conformation known as Reverse turns or β-bends.
- Reverse turns usually occur on the surfaces of globular proteins.

Super secondary structures (Motifs)

- Varies combinations of secondary structure, called super secondary structure, are commonly found in globular proteins.

These are

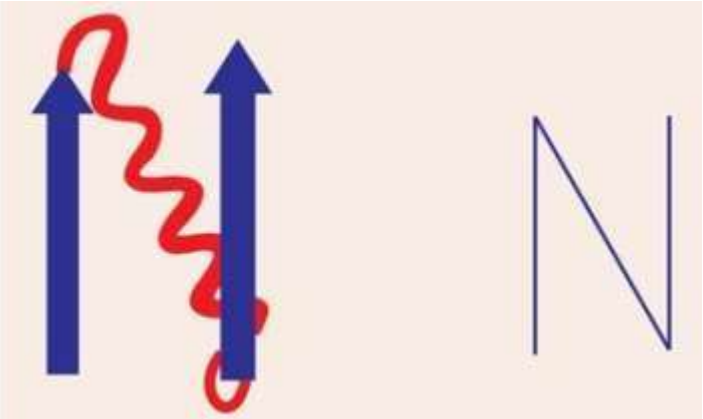
- β- α- β unit
- Greek key
- β-meander
- β- barrel

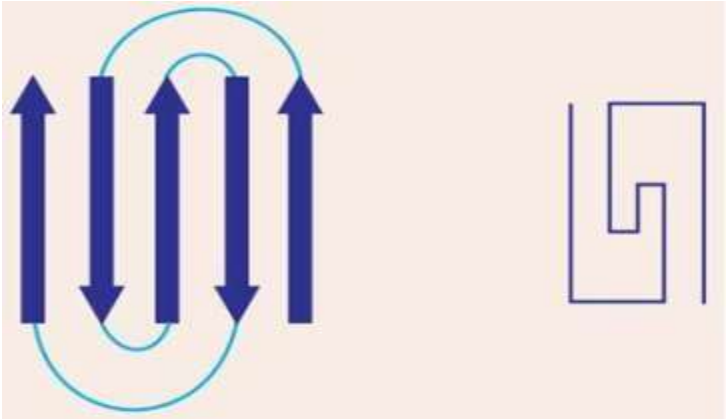
β- α- β unit :

- The β- α- β unit consists of two parallel β-pleated sheets connected by an intervening strand of α –helix.

Greek key :

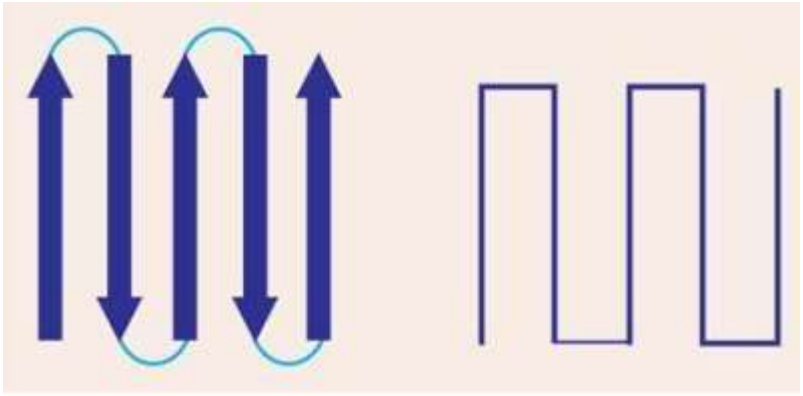
- A conformation that takes its name from a design often found on classical Greek pottery.

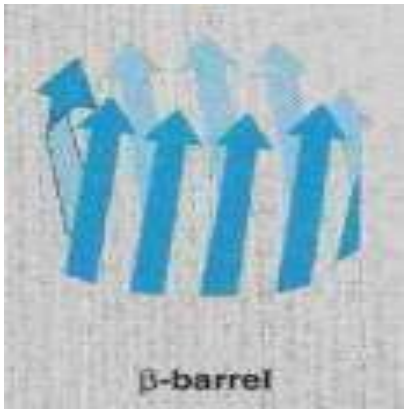




β-meander :

- The β-meander consists of five β-Pleated sheets connected by reverse turns.
- The β-meander contains nearly as many hydrogen bonds as an α-helix and its common occurrence probably reflects the stability conferred by this extensive hydrogen bonding.





Tertiary structure

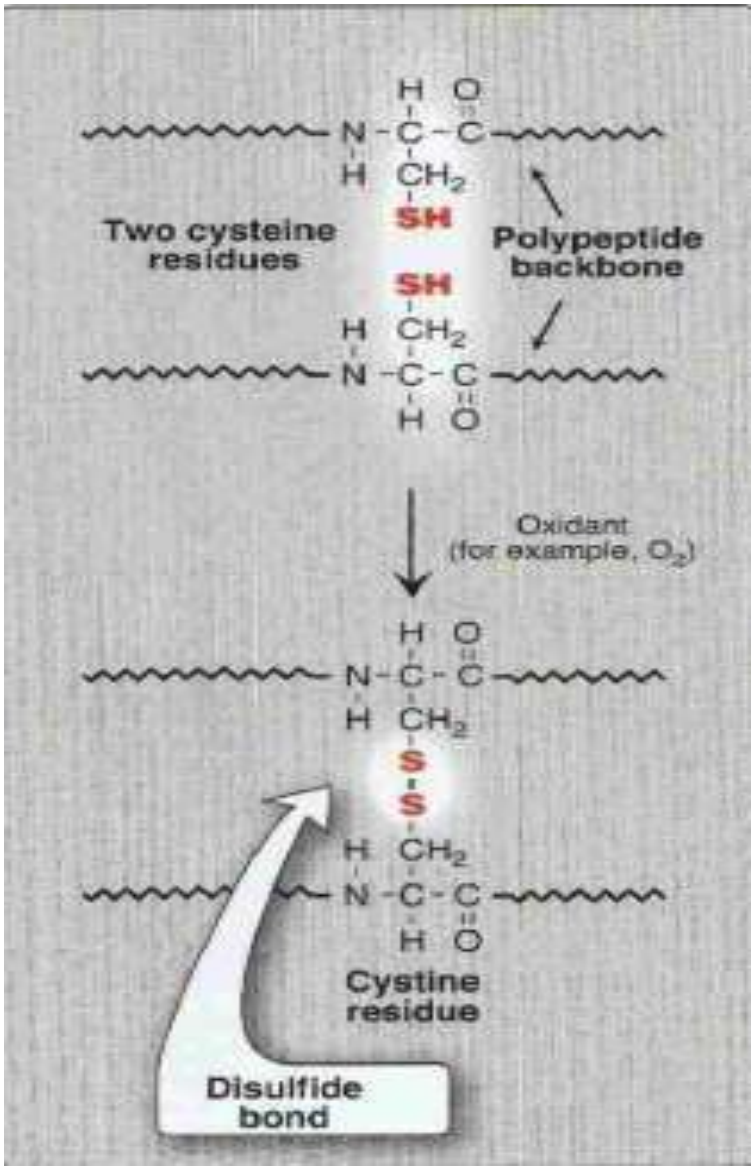
- The three dimensional structure of a protein.
- It is a compact structure with hydrophobic side chains held interior while the hydrophilic groups are on the surface of the protein.
- This type of arrangement ensures stability of the molecule.

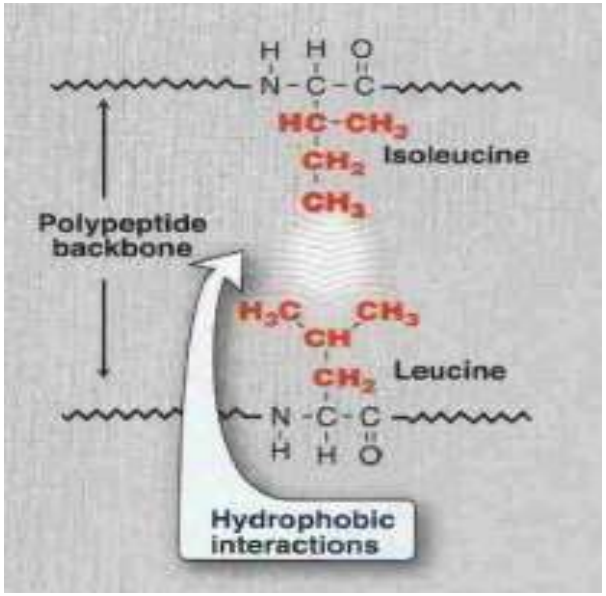
Bonds of Tertiary structure

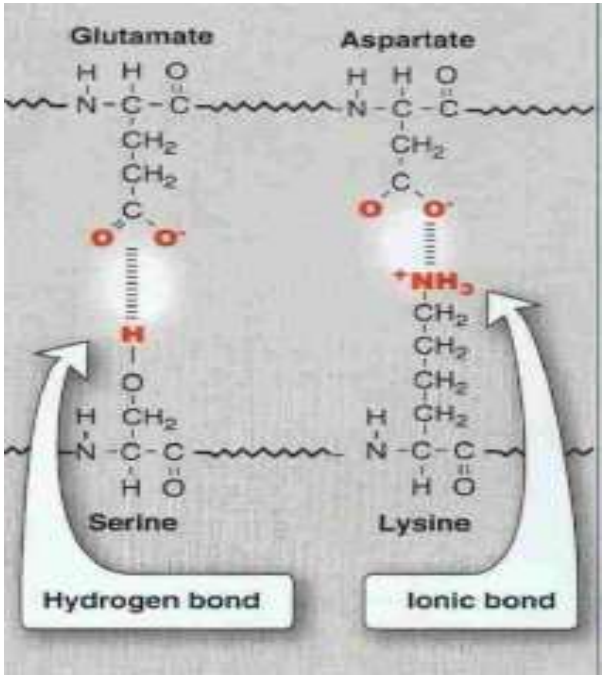
- The hydrogen bonds,
- Disulfide bonds(-S-S)
- Ionic interactions (electrostatic bonds) and
- Hydrophobic interactions also contribute to the tertiary structure of proteins.

Domains: Used to represent the basic unit of protein structure (tertiary) and function.

- A polypeptide with 200 amino acids normally consists of two or more domains.


Two cysteine residues
Polypeptide backbone
Oxidant (for example, O_2)
Cystine residue
Disulfide bond


Isoleucine
Polypeptide backbone
Leucine
Hydrophobic interactions


Glutamate
Aspartate
Serine
Lysine
Hydrogen bond
Ionic bond

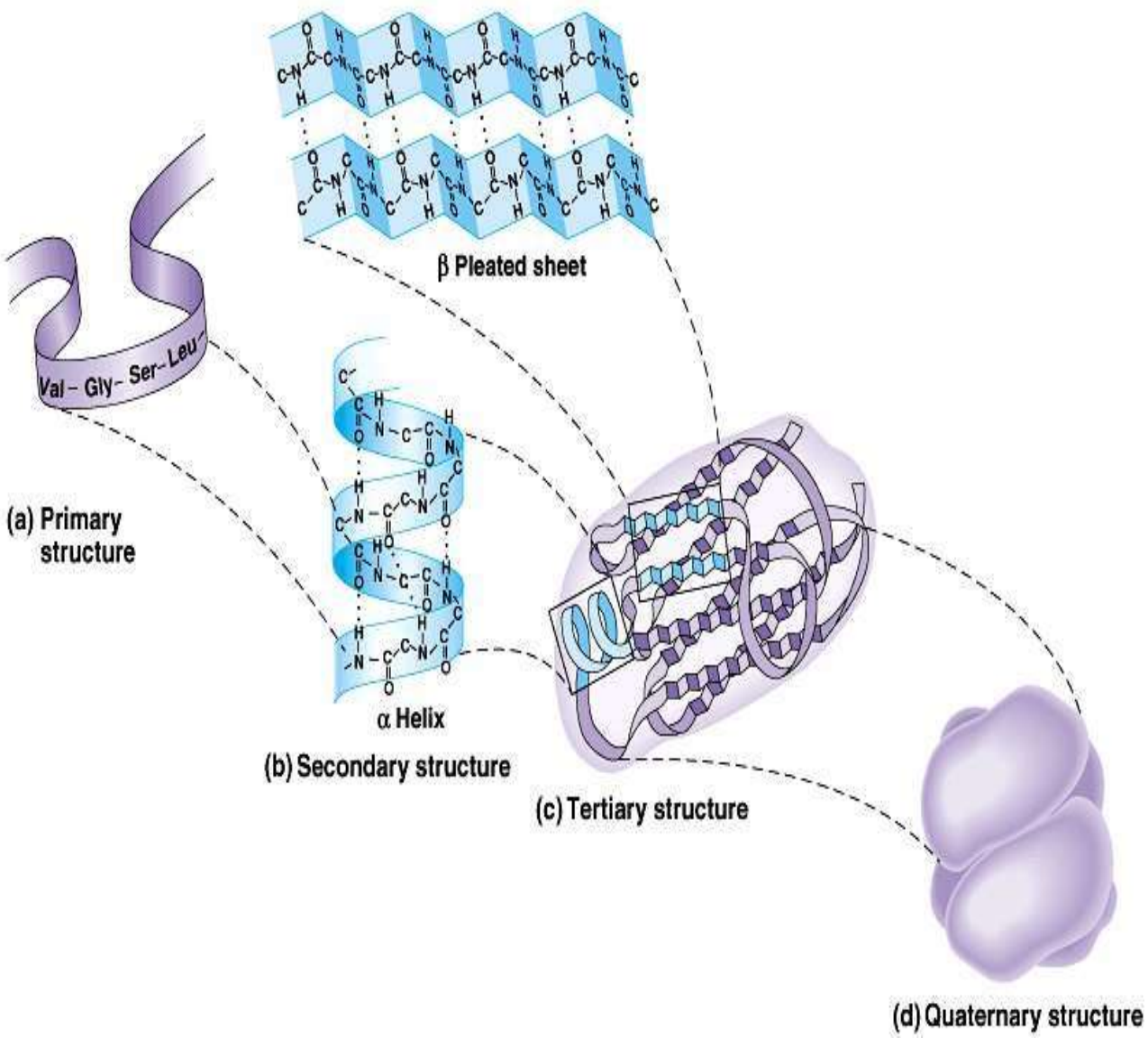
Quaternary structure

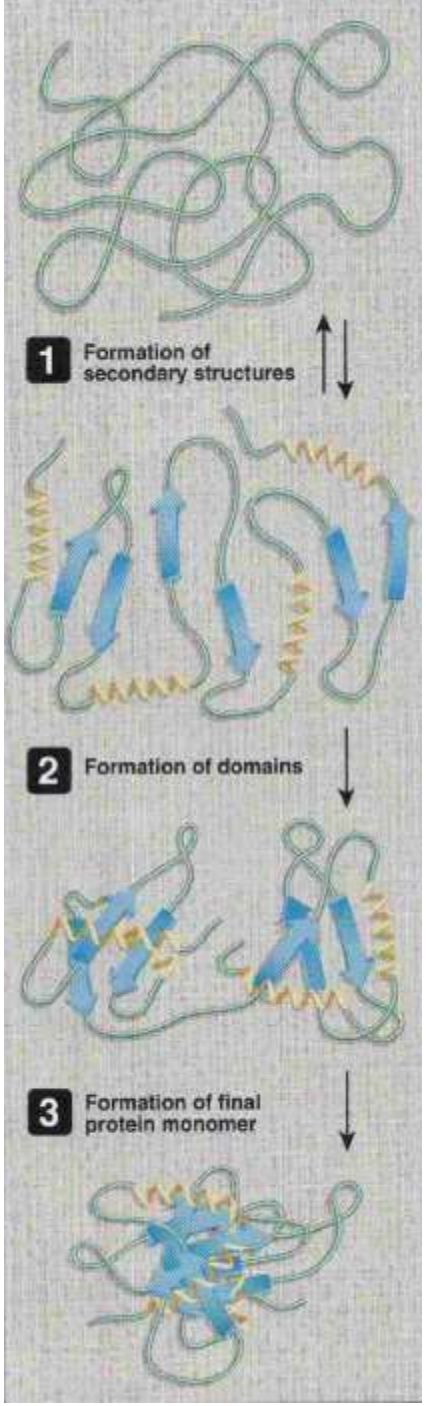
- Majority of the proteins are composed of single polypeptide chains.
- Some of the proteins, consists of two or more polypeptides which may be identical or unrelated.
- Such proteins are termed as oligomers and possess quaternary structure

- The individual polypeptide chains are known as monomers, protomers or subunits.
- A dimer consists of two polypeptides while a tetramer has four.
- Importance of oligomeric proteins:

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

e.g. hemoglobin, LDH.


β Pleated sheet
Val – Gly – Ser – Leu
(a) Primary structure
α Helix
(b) Secondary structure
(c) Tertiary structure
(d) Quaternary structure


1 Formation of secondary structures
2 Formation of domains
3 Formation of final protein monomer

Bonds responsible for protein structure

Two types of bonds

Covalent and non-covalent bonds

Covalent bonds:

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

Disulfide bond:

- A disulfide bond (-S-S-) is formed by the sulfhydryl groups (-SH) of two cystine 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 protein.

Non-covalent bonds

Hydrogen bonds (H-bonds):

- The H-bonds are formed by sharing of H-atoms between the Nitrogen and carbonyl oxygen of different peptide bonds.
- Each H-bond is weak but collectively they are strong.
- A large number of H-bonds significantly contribute to the protein structure.

Hydrophobic bonds

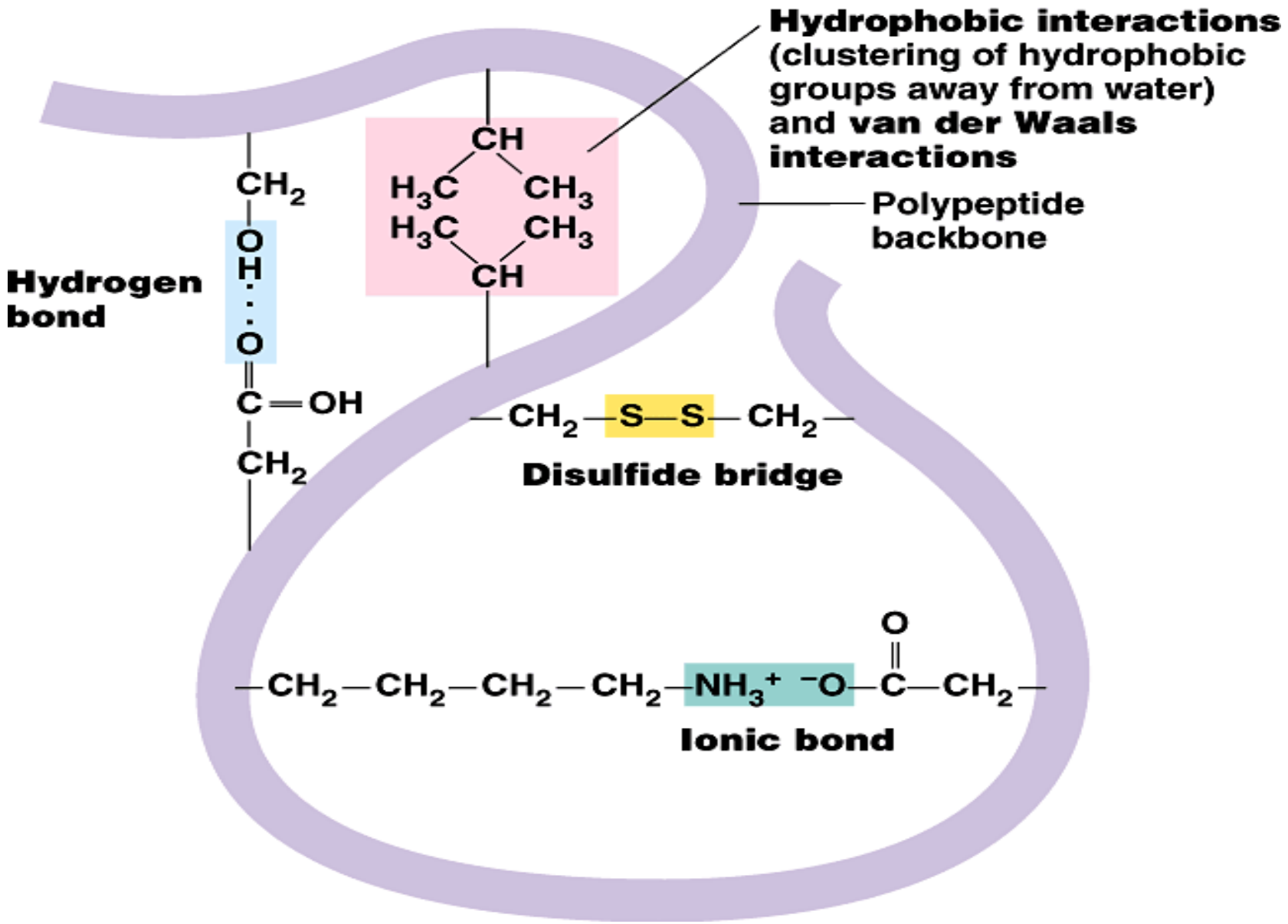
- The non-polar side chains of neutral amino acids tend to be closely associated with each other in proteins.
- These are not true bonds.
- The occurrence of hydrophobic forces is observed in aqueous environment wherein the molecules are forced to stay together.

Electrostatic bonds:

- These bonds are formed by interactions between negatively charged groups of acidic amino acids with positively charged groups of basic amino acids.

Van der Waals forces:

- **These are the non-covalent associations between electrically neutral molecules.**
- **They are formed by the electrostatic interactions due to permanent or induced dipoles**


Hydrophobic interactions
(clustering of hydrophobic groups away from water)
and van der Waals interactions
Polypeptide backbone
Hydrogen bond
Disulfide bridge
Ionic bond

Methods to determine protein structure

- Determination of secondary and tertiary protein structures, X-ray crystallography is most commonly used.
- Nuclear magnetic resonance (NMR) spectroscopy of proteins provides structural and functional information on the atoms and groups present in the proteins

Methods for the isolation and purification of proteins

- Proteins are fractionated by using different concentrations of ammonium sulfate or sodium sulfate.
- Protein fractionation may also be carried out by ultracentrifugation
- Protein separation is achieved by utilizing electrophoresis, isoelectric focussing, immunoelectrophoresis, ion-exchange chromatography, gel filtration, HPLC.

Properties of proteins

Solubility:

- Proteins form colloidal solutions instead of true solutions in water.
- This is due to huge size of protein molecule.

Molecular weight:

- The proteins vary in their molecular weights, dependent on the number of amino acid residues.
- Majority of proteins/polypeptides may be composed of 40 to 4,000 amino acids with a molecular weight ranging from 4,000 to 440,000.

Shape:

- There is a wide variation in the protein shape.
- It may be globular (insulin), Oval (albumin), fibrous or elongated (fibrinogen)

Isoelectric pH:

- At isoelectric pH, the proteins exist as zwitterion or dipolar ions.
- They are electrically neutral with minimum solubility, maximum precipitability and least buffering capacity.

Precipitation of proteins:

- Proteins exist in colloidal solution due to hydration of polar groups ($-COO^-$, $-NH_3^+$, -OH).
- Proteins can be precipitated by dehydration or neutralization of polar groups

Iso-electric precipitation:

- Proteins are least soluble at their iso-electric pH.
- Some of the proteins are precipitated immediately when adjusted to their iso-electric pH

e.g. Casein which forms a flocculent precipitate at pH 4.6 and redissolves in highly acidic or alkaline solutions.

- When the milk is curdled, the casein forms the white curd, because lactic acid produced by the fermentation process lowers the pH to the iso-electric point of casein.

Precipitation by salting out

- When a neutral salt such as ammonium sulphate or sodium sulphate is added to the protein solution , the shell of hydration is removed and the protein is removed and the protein is precipitated
- This is called as salting out
- Higher the molecular weight of a protein, the salt required for precipitation is lesser
- Globulins are precipitated by half saturation of ammonium sulphate; but albumin will need full saturation

Precipitation by Organic solvents:

- Organic solvents such as alcohol are good protein precipitating agents.
- They dehydrate the protein molecule by removing the water envelope and cause precipitation

Precipitation by salts of heavy metals:

- Heavy metal ions like Pb^{2+}, Hg^{2+}, Fe^{2+} , Zn^{2+} , Cd^{2+} cause precipitation of proteins.
- These metals being positively charged, when added to protein solution (negatively charged) in alkaline medium results in precipitate formation.

Precipitation by anionic or alkaloid reagents

- Proteins can be precipitated by trichloroacetic acid, sulphosalicylic acid, phosphotungstic acid, picric acid, tannic acid, phosphomolybdic acid etc.

  By the addition of these acids, the proteins existing as cations are precipitated by the anionic form of acids to produce protein-sulphosalicylate, protein-tungstate, protein-picrate etc.

Denaturation

- Defined as an disorganization of native protein structure
- Denaturation results in the loss of secondary, tertiary and quarternary structure of proteins due to cleavage of noncovalent bonds
- This involves a change in physical, chemical and biological properties of protein molecules

Note: Primary structure of protein molecule, i.e., peptide bond is not affected.

Denaturating agents

Physical agents

Heat, violent shaking, X-rays, UV radiation

Chemical agents:

Acids, alkalis, organic solvents (ether, alcohol), salts of heavy metals (Pb, Hg), urea, salicylate etc.

Characteristics of denaturation

- The native helical structure of protein is lost
- The primary structure of a protein with peptide linkages **remain intact. i.e. peptide bonds are not hydrolyzed**
- The protein loses its biological activity
- Denatured protein becomes insoluble in solvent in which it was originally soluble
- The viscosity of denatured protein (solution) increases while its surface tension decreases

- Denaturation is associated with increase in ionizable and sulfhydryl groups of protein.
- This is due to the loss of hydrogen and disulfide bonds
- Denatured protein is more easily digested.
- This is due to increased exposure of peptide bonds to enzymes
- Denaturation is irreversible.

e.g., omelet can be prepared from an egg but the reversal is not possible

- Careful denaturation is sometimes reversible (known as renaturation).

  Hemoglobin undergoes denaturation in the presence salicylate.
- By the removal of salicylate, hemoglobin is renaured
- Denatured protein cannot be crystallized

Coagulation

- Coagulation refers to a semi-solid viscous precipitate of protein.
- Irreversible denaturation results in coagulation

•Coagulation is optimum and requires lowest temperature at isoelectric pH

e.g., Albunins and globulins

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

Flocculation:

- It is the process of protein precipitation at isoelectric pH
- The precipitate is referred to as flocculum
- Flocculation is reversible.
- On the application of heat, flocculum can be converted into an irreversible mass, coagulum.

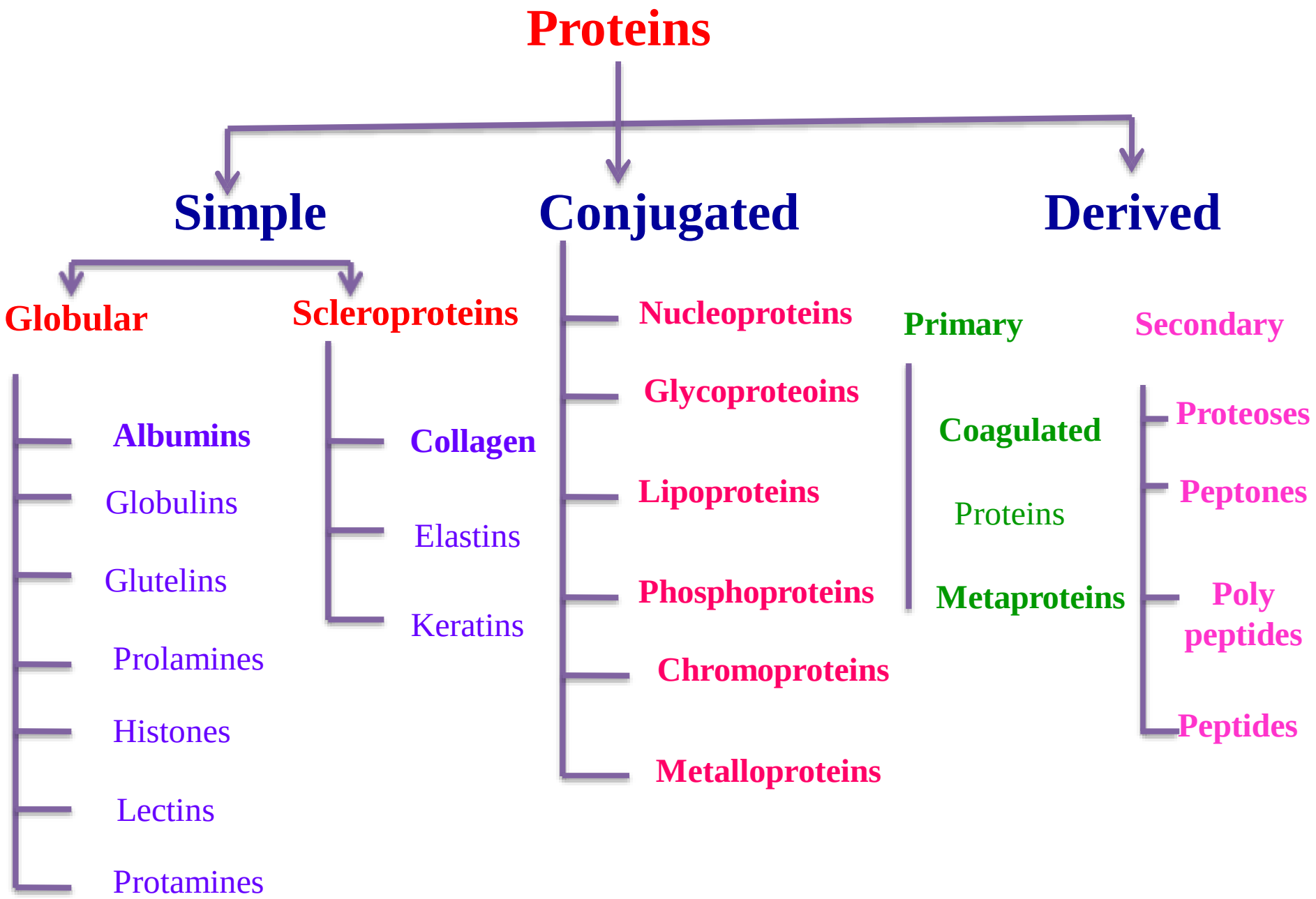
Classification of proteins

Three major types of classifying proteins based on their function, chemical nature and solubility and nutritional importance

- Functional classification of proteins
- Chemical nature and solubility
- Nutritional importance

Functional classification of proteins

- Structural proteins: Keratin of hair and nails, collagen of bone
- Contractile proteins: myosin, actin
- Enzymes or catalytic proteins: Hexokinase, pepsin
- Transport proteins: Hemoglobin, serum albumin
- Regulating proteins: Insulin, growth hormone


Proteins
Simple
Conjugated
Derived
Globular
Scleroproteins
Albumins
Globulins
Glutelins
Prolamines
Histones
Lectins
Protamines
Collagen
Elastins
Keratins
Nucleoproteins
Glycoproteoins
Lipoproteins
Phosphoproteins
Chromoproteins
Metalloproteins
Primary
Secondary
Coagulated
Proteins
Metaproteins
Proteoses
Peptones
Poly peptides
Peptides

Simple proteins

A) Globular proteins:

- These are spherical or oval in shape, soluble in water or other solvents and digestible

Albumins:

- They are soluble in water and coagulated by heat. e.g serum albumin, ovalbumin (egg), lactalbumin (milk)

Globulins:

- These are insoluble in pure water, but soluble in dilute salt solutions and coagulated by heat.

- They are precipitated by half saturation with ammonium sulphate or by full saturation with sodium chloride. e.g., serum globulins, egg globulins

Glutelins:

- These are plant proteins, insoluble in water or neutral salt solutions, soluble in dilute acids or alkalies
- They are rich in glutamic acid.
- They are large molecules and can be coagulated by heat. e.g., glutelin (wheat), oryzenin (rice).

Prolamines:

- They are soluble in 70-80% alcohol, but insoluble in pure water.
- They are rich in proline but lack in lysine. E.g., gliadin (wheat), zein (maize)

Histones:

- These are basic proteins, rich in arginine and histidine, with alkaline isoelectric pH.
- They are soluble in water, dilute acids and salt solutions but insoluble in ammonia

- They form conjugated proteins with nucleic acids (DNA) and porphyrins.

  e.g., nucleohistones, chromosomal nucleoproteins

Protamines:

- These are soluble in water, dilute acids and alkalies. They are not coagulated by heating..
- They contain large number of arginine and lysine residues, and are strongly basic. These are also found in association with nucleic acids

Scleroproteins or Albuminoids

- These are fibrous proteins with great stability and very low solubility and form supporting structures of animals

Collagens, are connective tissue proteins lacking tryptophan

Elastins, these proteins are found in elastic tissues such as tendons and arteries

Keratins, these are present in exoskeletal structures e.g. hair, nails, horns.

Human hair has a higher content of cysteine.

Conjugated proteins

They are combinations of proteins with a non-protein part, called prothetic group

These are

Glycoproteins:

- Glycoproteins are the proteins with carbohydrate moiety as the prosthetic group.
- The term mucoprotein is used if the carbohydrate content is more than 4%.
- Blood group antigens and many serum proteins are glycoproteins

Lipoproteins:

- These are the proteins loosely combined with lipid components.
- They occur in blood and on cell membranes
- Serum lipoproteins, membrane lipoproteins

Nucleoproteins:

- These are the proteins attached to nucleic acids, e.g. Histones.
- The DNA carries negative charges, which combines with positively charged proteins

Chromoproteins

•These are the proteins with colored prosthetic groups

Hemoproteins: All hemoproteins are Chromoproteins which carry heme as the prosthetic group

- Hemoglobin: Respiratory protein found in RBCs
- Cytochromes: These are the mitochondrial enzymes of the respiratory chain

Catalase:

- This enzyme decomposes H_2O_2 to water and O_2

Flavoproteins:

- Is a cellular oxidation-reduction protein which has riboflavin a constituent of B-complex vitamin as its prosthetic group.
- This is yellow in color

Visual purple:

- Is a protein of retina in which the prosthetic group is a corotenoid pigment which is purple in colour

Phosphoproteins:

- These contain phosphorous. E.g. casein and vitellin of egg yolk.
- The phosphoric acid is esterified to the hydroxyl groups of serine and threonine residues of proteins

Metalloproteins:

- They contain metal ion as their prothetic group.
- Several enzymes contain metallic elements such as Fe, Co, Mn, Zn, Cu, Mg, etc

  E.g., Ferritin (Fe), Carbonic anhydrase (Zn), Ceruloplasmin (Cu)

Derived proteins

•The derived proteins are of two types

Primary derived proteins:

- These are the denatured or coagulated or first hydrolysed products of proteins

Secondary derived proteins:

- These are the degraded (due to breakdown of peptide bonds) products of proteins

Primary derived proteins

Proteans:

- These are earliest products of protein hydrolysis by enzymes, dilute acids, alkalis etc.
- They are insoluble in water

•E.g., Myosan (from myosin), Elestan (from elastin)

Meta proteins:

- These are the second stage products of protein hydrolysis by treatment with slightly stronger acids and alkalis. e.g., acid and alkali metaproteins.

Coagulated proteins:

- These are the denatured proteins produced by agents such as heat, acids, alkalis etc, E.g., cooked proteins, coagulated albumin (egg white)

Secondary derived proteins:

- These are the degraded (due to breakdown of peptide bonds) products of proteins

Proteoses or albumoses:

- These are hydrolytic products of proteins which are soluble in water and coagulated by heat and precipitated by saturation with ammonium sulphate

Peptones:

- These are hydrolytic products of proteoses.
- They are soluble in water, not coagulated by heat and not precipitated by saturation with ammonium sulphate.
- They can be precipitated by phosphotungstic acid

Peptides:

- Peptides are composed of very small number of amino acids joined as peptide bonds.
- They are named according to the number of amino acids present in them

- Dipeptide: made up of two amino acids
- Tripeptides: made up of three amino acids
- Peptides are water soluble and not coagulated by heat
- Hydrolysis: The complete hydrolytic decomposition of a protein generally follows these stages
- Protein → Protean → Metaprotein → Proteose → Peptone → Peptides → Amino acids

Nutritional classification of proteins

Complete proteins or Nutritionally rich proteins: These proteins have all the essential amino acids in the required proportion.

- Also called as first class proteins. e.g., egg albumin, casein of milk

Partially incomplete proteins:

- These proteins are partially lacking one or more essential amino acids and can promote moderate growth. e.g., wheat and rice proteins. Limiting amino acids Lysine and Threonine
- Pulses lack Methionine

Incomplete or poor proteins:

- They lack in many essential amino acids

  Hence they do not promote growth at all

- e.g., gelatin (lacks Trp), zein (from corn, lacks Trp, Lys)

Biologically Important Peptides

- When 10 or less number of amino acids are joined together, it is called as an oligopeptide
- Some of them are biologically active

Glutathione:

- It is tripeptide composed of 3 amino acids.
- Chemically, glutathione is γ-glutamyl-cystinyl-glycine.
- It exists in reduced or oxidized states.

Functions:

- Glutathione serves as a coenzyme for certain enzymes e.g., prostaglandin PGE2 synthetase, glyoxylase
- It prevents the oxidation of sulfhydryl (-SH) groups of several proteins to disulfide (-S-S-) groups.
- This is essential for protein function
- Glutathione in association with glutathione reductase participate in the formation of correct disulfide bonds in several proteins

- Glutathione maintains RBC membrane structure and integrity
- Glutathione protects hemoglobin from getting oxidized by agents such as H_2O_2
- Glutathione is involved in the transport of amino acids in the intestine and kidney tubules via γ-glutamyl cycle or Meister cycle
- Glutathione involved in the detoxification process

- Toxic amounts of peroxides and free radicals produced in the cells are scavanged by glutathione peroxidase (a selenium containing enzyme)

Thyroprophin releasing hormone(TRH):

- It is tripeptide secreted by hypothalamus
- TRH is stimulated by pituitary gland to release thyrotrophic hormone

Oxytocin:

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

Vasopressin (antidiuretic hormone, ADH):

- ADH is also a nonapeptide secreted by posterior pituitary gland
- It stimulates kidney to retain water and thus increases blood pressure.

Angiotensins:

- Angiotensin is a decapeptide (10 amino acids) which is converted to angiotensin II (8 amino acids)
- Angiotensin II stimulates the release of aldosterone from adrenal gland

Methionine enkephalin:

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

Bradykinin:

- They are nona and decapeptides.
- Act as powerful vasodilators
- Produced from plasma proteins by snake venom enzymes

Peptide antibiotics:

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

GIT hormones:

- Gastrin, secretin etc. are serves as hormones