AMINO ACIDS

Eric Mbindo Njunju

Bsc; Msc

Proteins

- Paramount importance in biological systems
- Play major structural and functional aspects of the body
- Polymers of amino acids linked by peptide bonds

- 300 amino acids in nature. Only 20 in human body
- Most are alpha (α). Proline is not
- Amino group attached to the same carbon atom to which the carboxyl group is attached

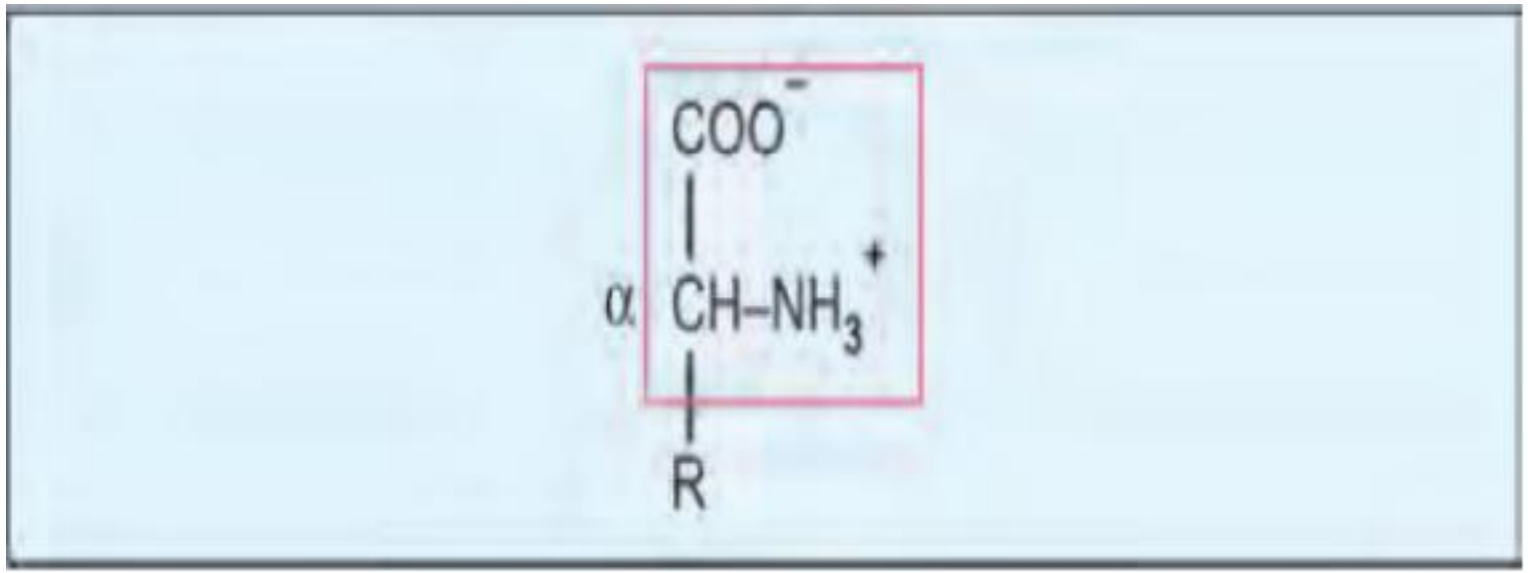


Fig: General structure

CLASSIFICATION OF AMINO ACIDS

Based on structure

A. Aliphatic amino acids

Monoamino monocarboxylic acids

Simple amino acids: Glycine, Alanine

Branched chain amino acids: Valine, Leucine, Isoleucine

Hydroxyamino acids: Serine, Threonine

Sulfur-containing amino acids: Cysteine, Methionine

Amino acids with amide group: Asparagine, Glutamine

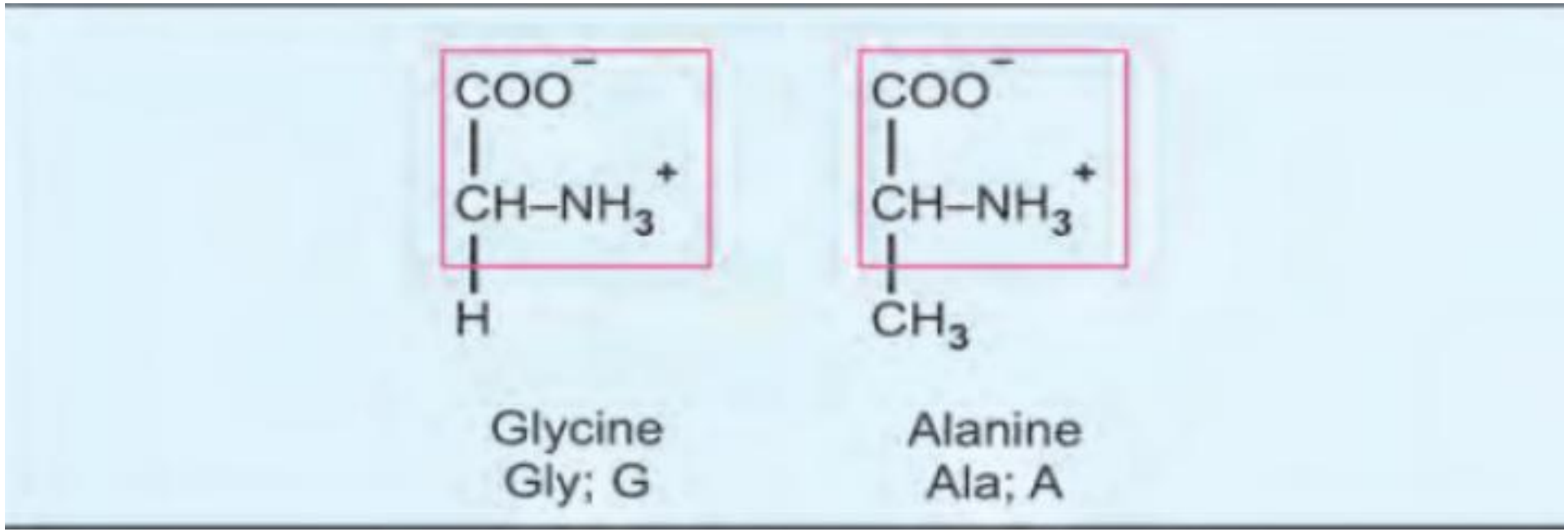


Fig: Simple amino acids

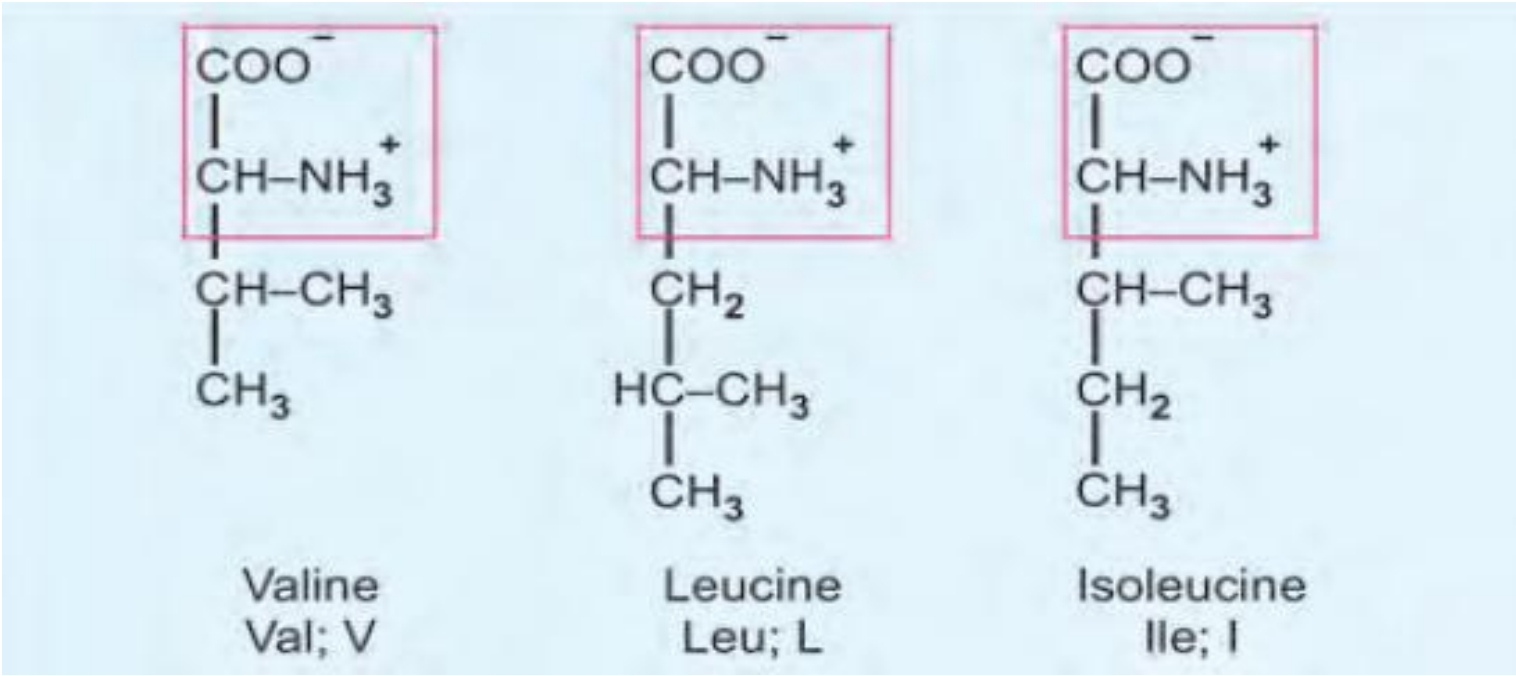


Fig: Branched chain amino acids

Fig: Hydroxyamino acids

Fig: Sulfur-containing amino acids

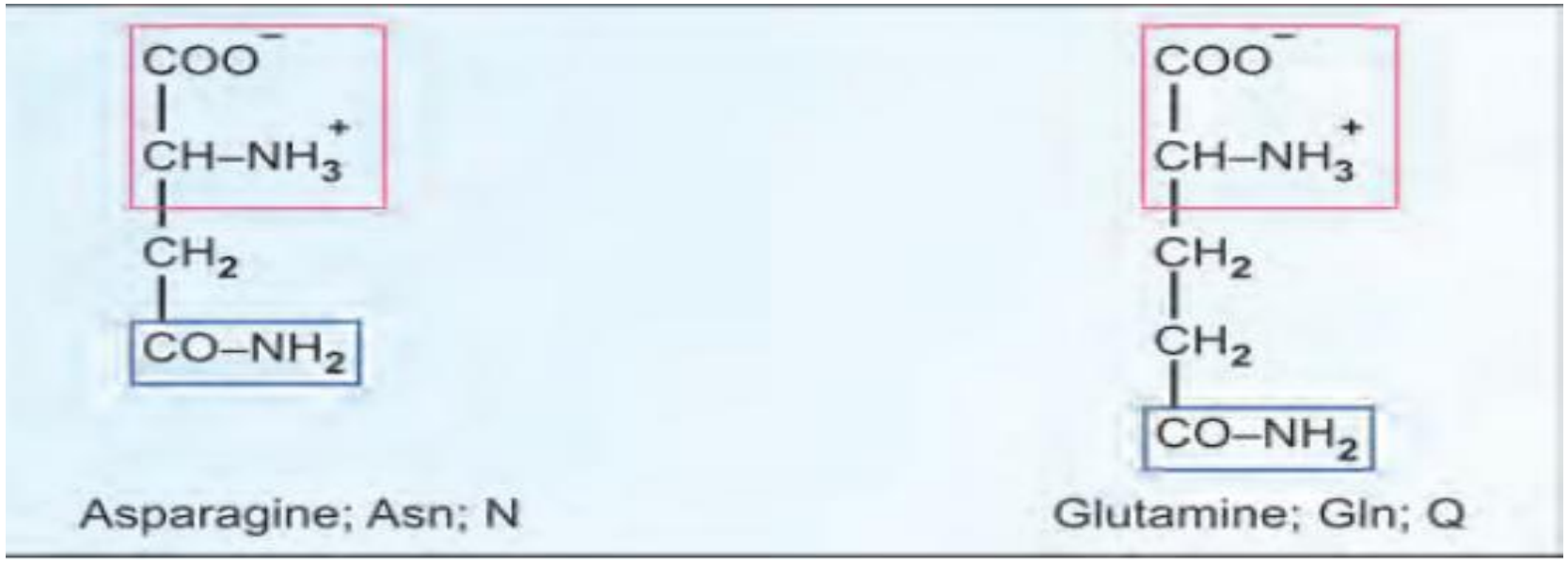


Fig: Amino acids with amide groups

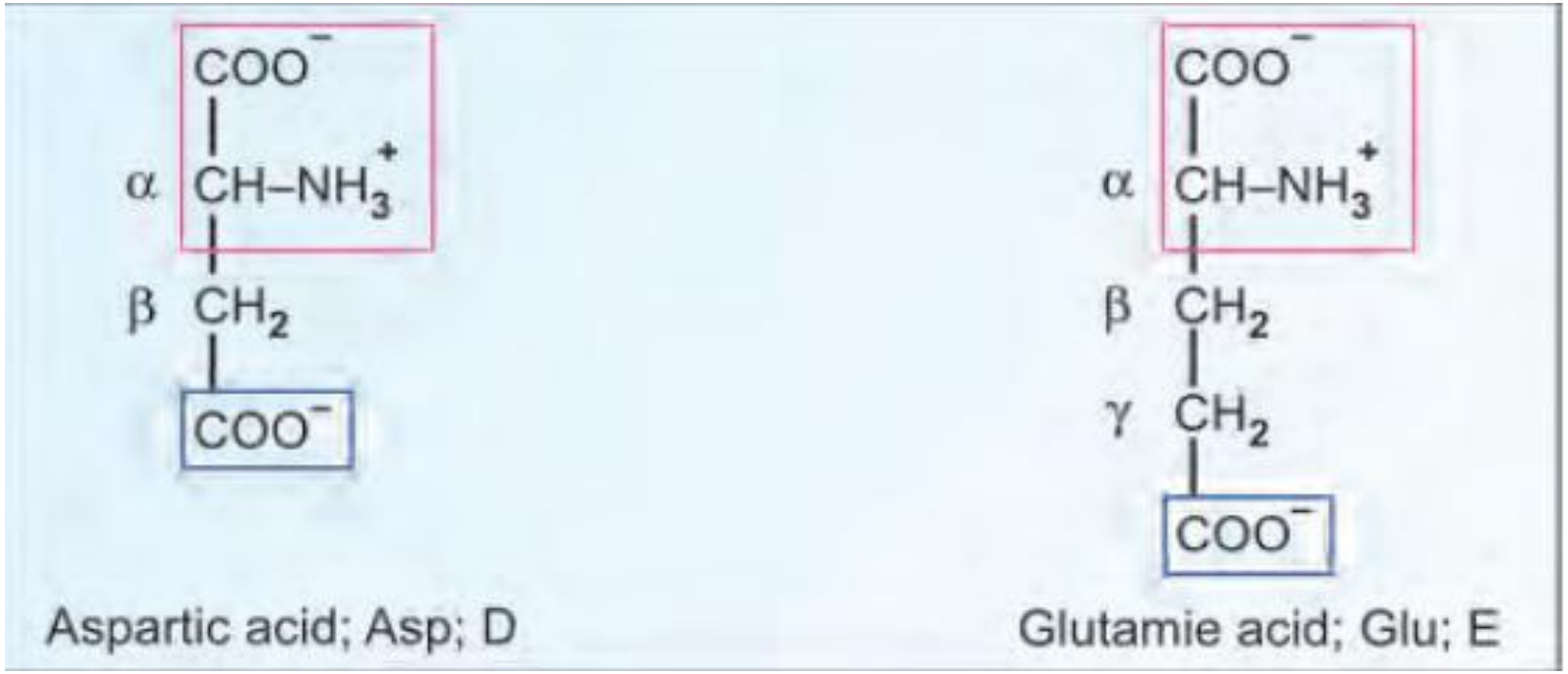
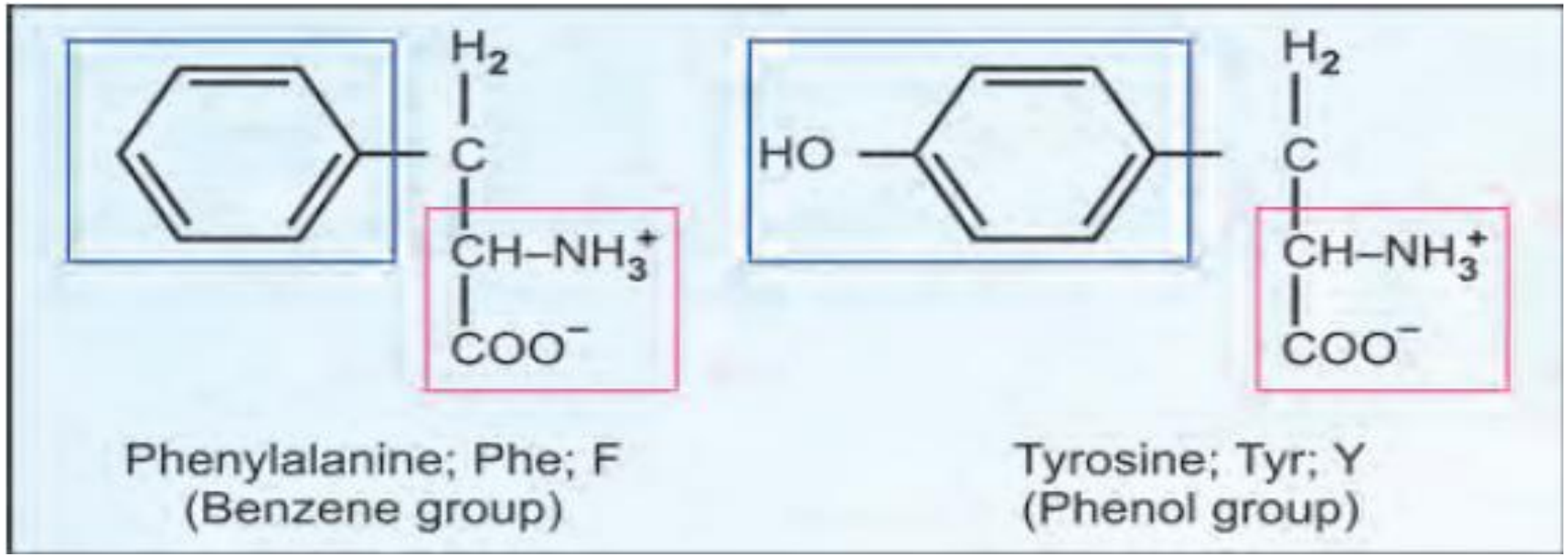
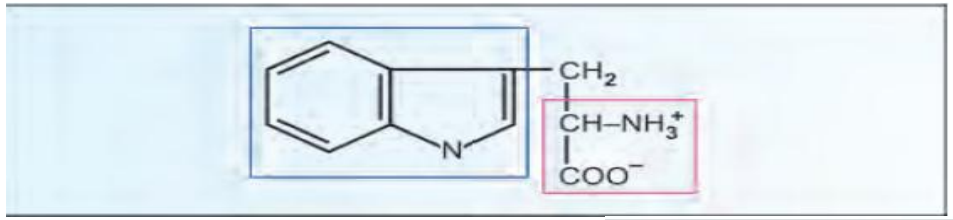


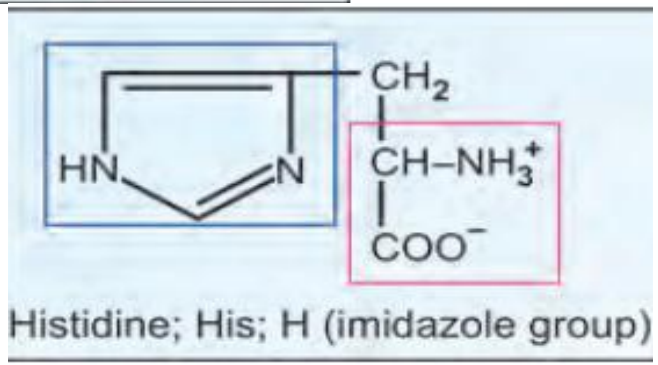
Fig: Mono amino dicarboxylic acids: Aspartic acid, Glutamic acid

Fig: Dibasic monocarboxylic acids: Lysine, Arginine

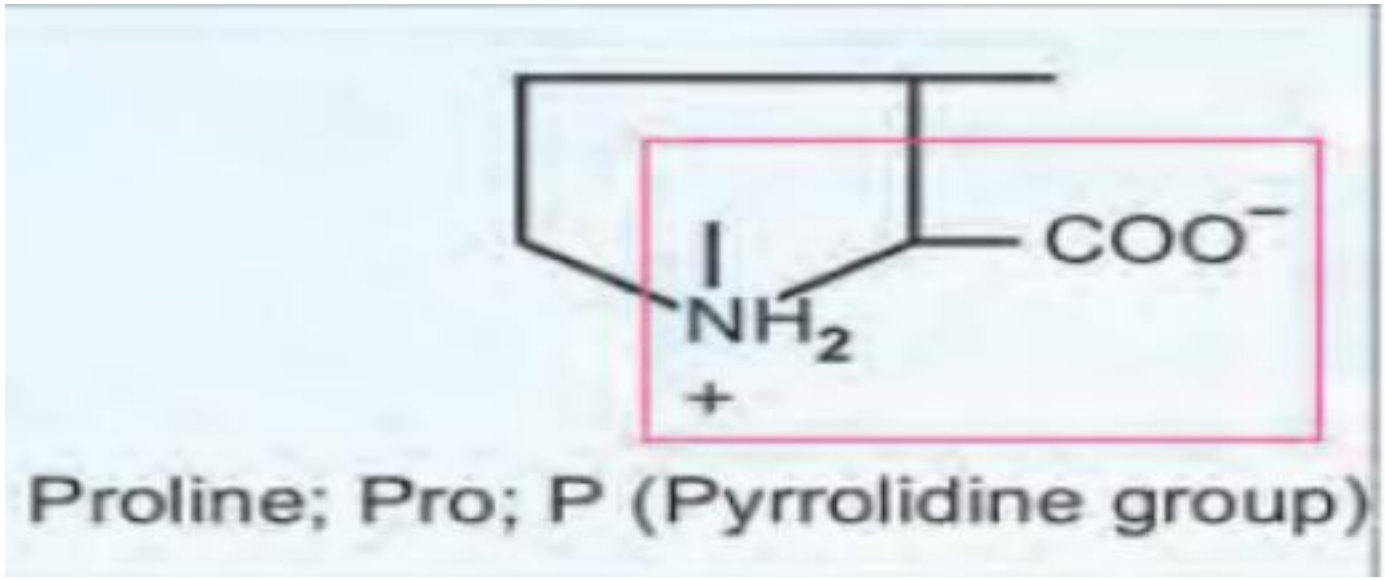


B. Aromatic amino acids: Phenylalanine, Tyrosine





Heterocyclic amino acids: Tyrptophan, Histidine



D. Imino acid: Proline

E. Derived amino acids:

I. Derived amino acids in proteins: Hydroxyproline and hydroxylysine

Important components of collagen

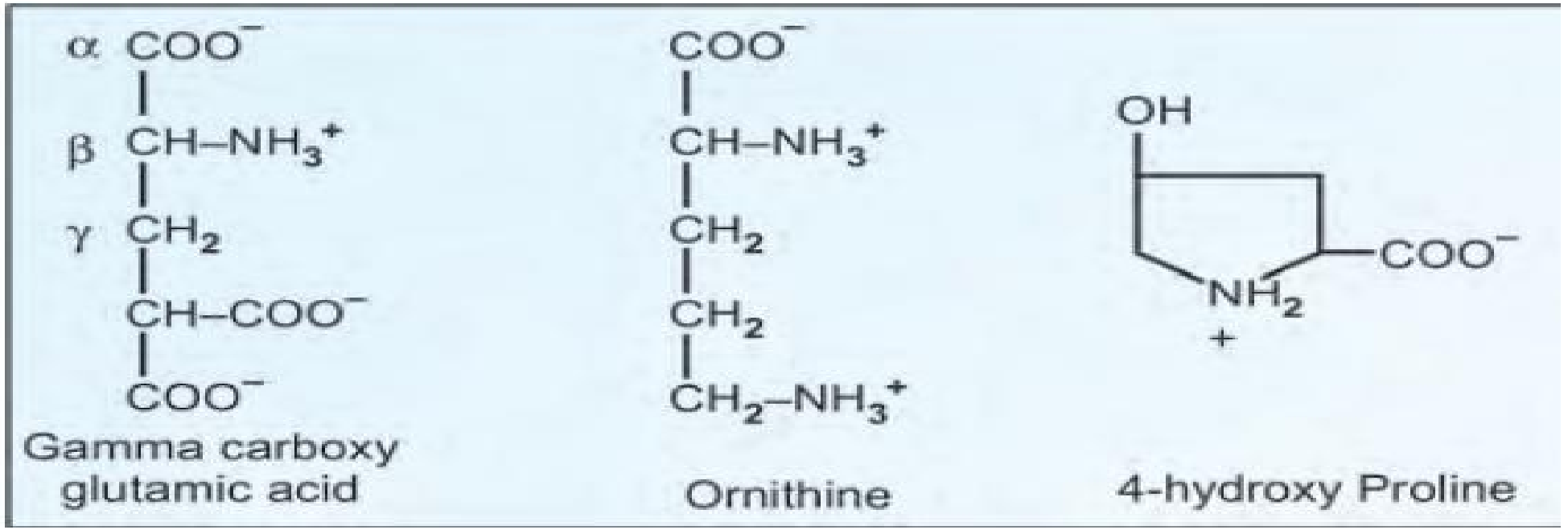
Gamma carboxylation of glutamic acid residues-blood clotting processes

In ribosomal proteins and in Histones: extensively methylated and acetylated

II. Derived amino acids not seen in proteins- free in cells: Ornithine, Citrulline, Homocysteine

- All produced during metabolism of amino acids

III. Non-alpha amino acids: Gamma amino butyric acid (GABA) derived from glutamic acid



Some derived amino acids

Name of amino acid	*Special group present*	*3-letter abbre-viation*	*1-letter abbre-viation*	*Mole-cular weight*
Glycine		Gly	G	77
Alanine		Ala	A	89
Valine		Val	V	117
Leucine		Leu	L	131
Isoleucine		Ile	I	131
Serine	Hydroxyl	Ser	S	105
Threonine	Hydroxyl	Thr	T	119
Cysteine	Sulfhydryl	Cys	C	121
Methionine	Thioether	Met	M	149
Asparagine	Amide	Asn	N	132
Glutamine	Amide	Gln	Q	146
Aspartic acid	β-carboxyl	Asp	D	133
Glutamic acid	γ-carboxyl	Glu	E	147
Lysine	ε-amino	Lys	K	146
Arginine	Guanidinium	Arg	R	174
Phenylalanine	Benzene	Phe	F	165
Tyrosine	Phenol	Tyr	Y	181
Tryptophan	Indole	Trp	W	204
Histidine	Imidazole	His	H	155
Proline (imino acid)	Pyrrolidine	Pro	P	115

TABLE: Common amino acids

Special groups:

- Arginine- guanidinium group
- Phenyl alanine-Benzene
- Tyrosine-Phenol
- Tryptophan-Indole
- Histidine-Imidazole
- Proline- Pyrrolidine

Based on Side chains

A. Non polar side chains: Alanine, Valine, Leucine, Isoleucine, Methionine, Proline, Phenylalanine and Tryptophan

Hydrophobic (water repellant) and lipophilic groups

B. Uncharged or non-ionic polar side chains Glycine, serine, Threonine, Cysteine, Tyrosine, Glutamine, Asparagine

Hydrophilic

C. Charged or ionic polar side chains:

Acidic: Aspartic acid and Glutamic acid

Basic: Lysine, Arginine and Histidine

Based on Metabolism

A. Purely ketogenic: Leucine- converted to ketone bodies
B. Ketogenic and glucogenic: Lysine, Isoleucine, Phenylalanine, Tryosine and Tryptophan
C. Purely Glucogenic: Remaining are purely

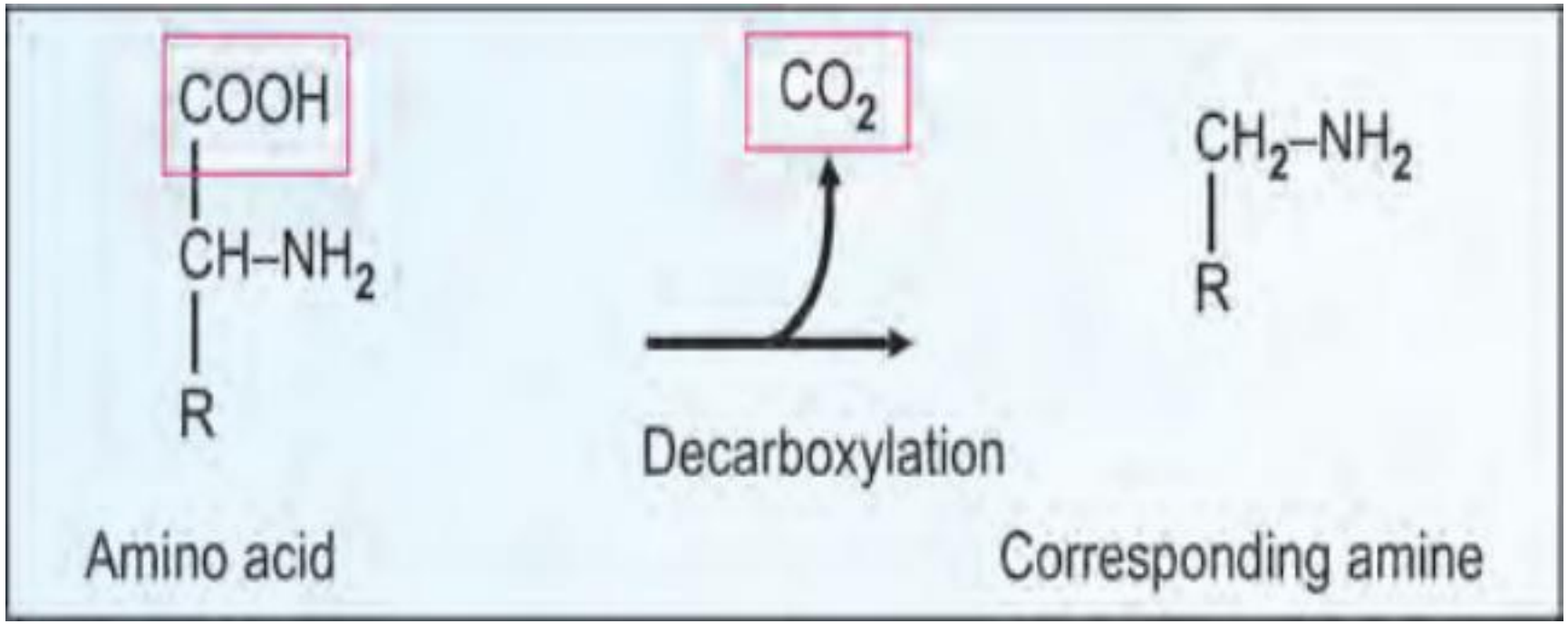
Based of Nutritional Requirements

A. Essential or Indispensable: Isoleucine, Threonine, Lysine, Methionine, Phenylalanine, Tryptophan and Valine
B. Partially essential (semi essential): Histidine and Arginine
C. Non essential: can be synthesized in the body
D. Conditionally essential: Arginine, Glycine, Cysteine, Tyrosine, Proline, Glutamine

General reactions

- Due to Carboxyl group
- Decarboxylation-Alpha decarboxylation to form the corresponding amine
- Histidine undergoes decarboxylation to give Histamine and carbon dioxide
- Glutamic acid gives Gamma amino butyric acid (GABA) and carbon dioxide

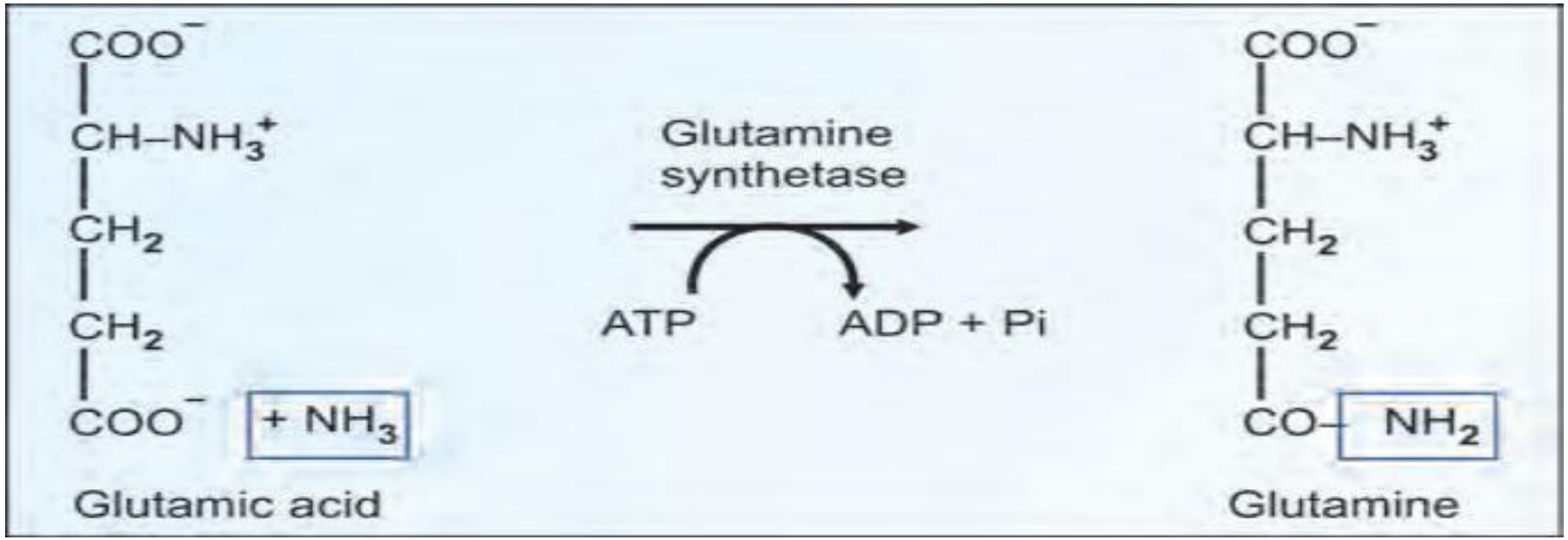
DECARBOXYLATION



Amide formation- COOH group of dicarboxylic amino acids (other than alpha carboxyl combines with ammonia to from an amide)

- Aspartic acid + NH_3 gives Asparagine
- Glutamic acid +NH_3 gives Glutamine
- Amides are also components of protein structure
- Amide group of glutamine- source of nitrogen for nucleic acid synthesis

AMIDE FORMATION

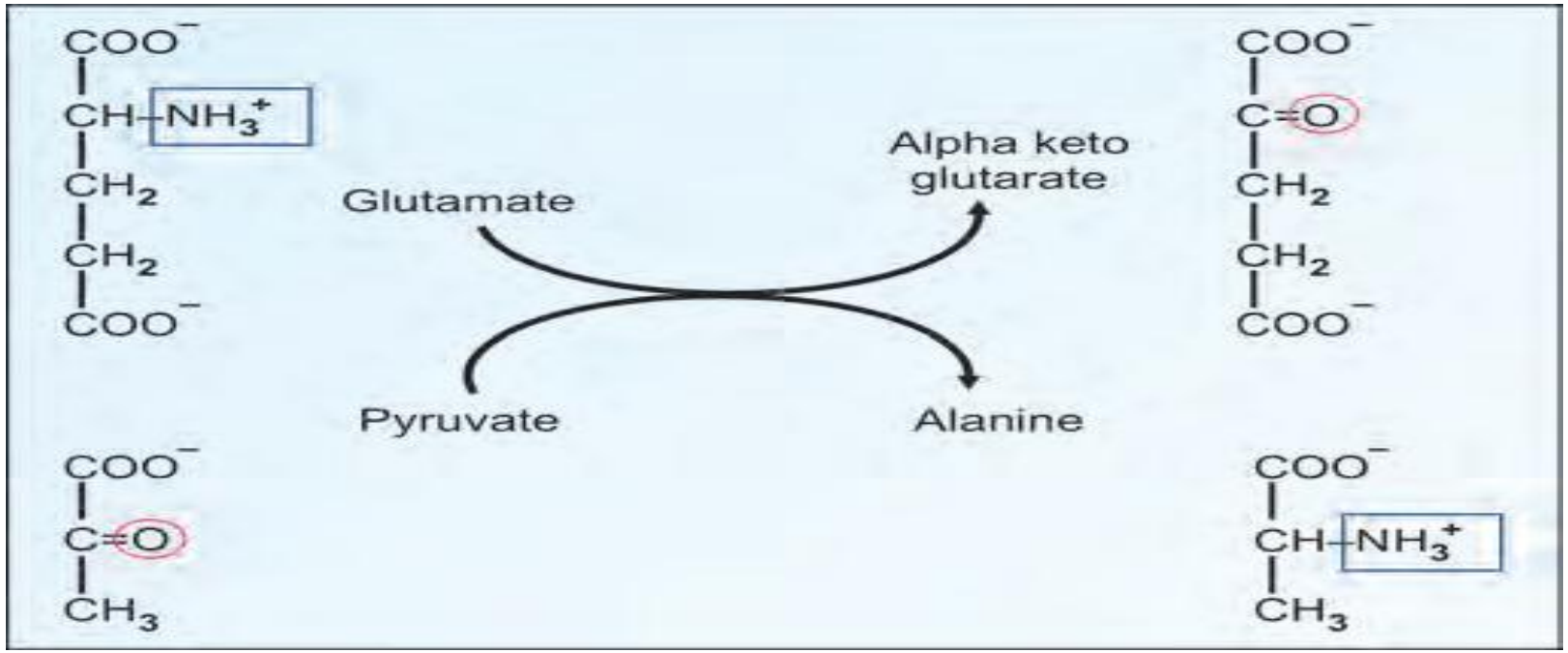


Due to Amino group

Transamination-Alpha amino group transferred to alpha keto acid to form the corresponding new amino acid and alpha keto acid

Important reaction- interconversion of amino acids and synthesis of non essential amino acids

TRANSAMINATION



Oxidative Deamination

Alpha amino group removed to form corresponding keto acid and ammonia.

Glutamic acid most common amino acid to under go this reaction

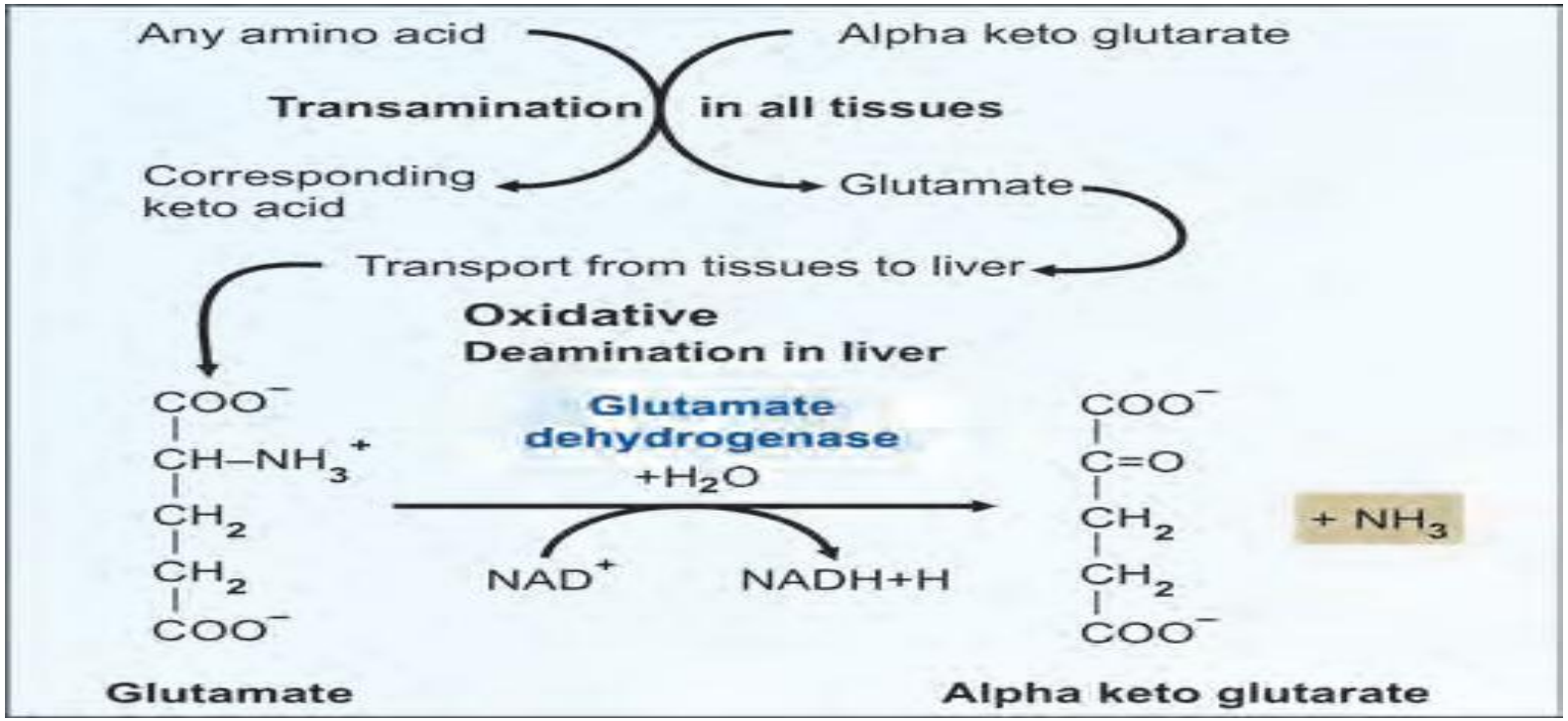
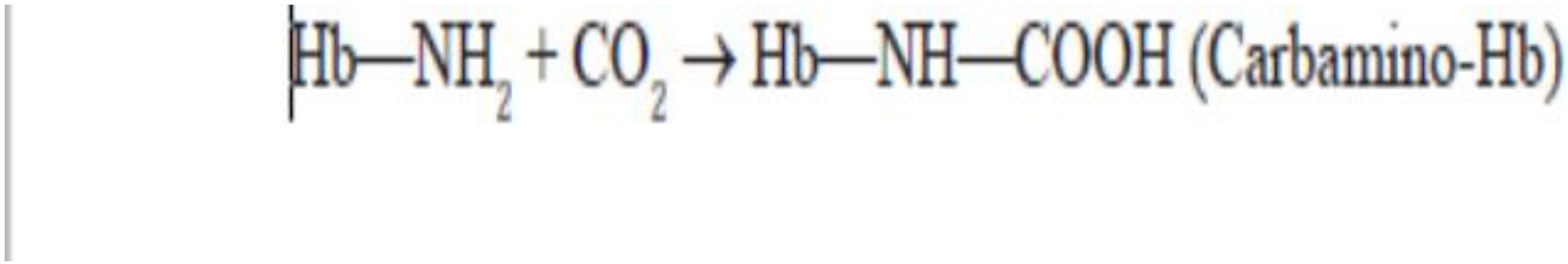


Fig: Transamination + deamination = transdeamination

Formation of Carbamino Compound

Carbon dioxide adds to the alpha amino group to form carbamino compounds

Mechanism for transport of carbon dioxide from tissues to the lungs



Due to side chains

Transmethylation

Methyl group of methionine, after activation transferred to an acceptor

Ester formation by the OH group

Hydroxy amino acids form esters with phosphoric acid to give phosphoproteins- O-glycosidic bonds

Similar to formation of glycoproteins

Reaction of amide groups of Glutamine and Asparagine- N-glycosidic bonds with carbohydrates

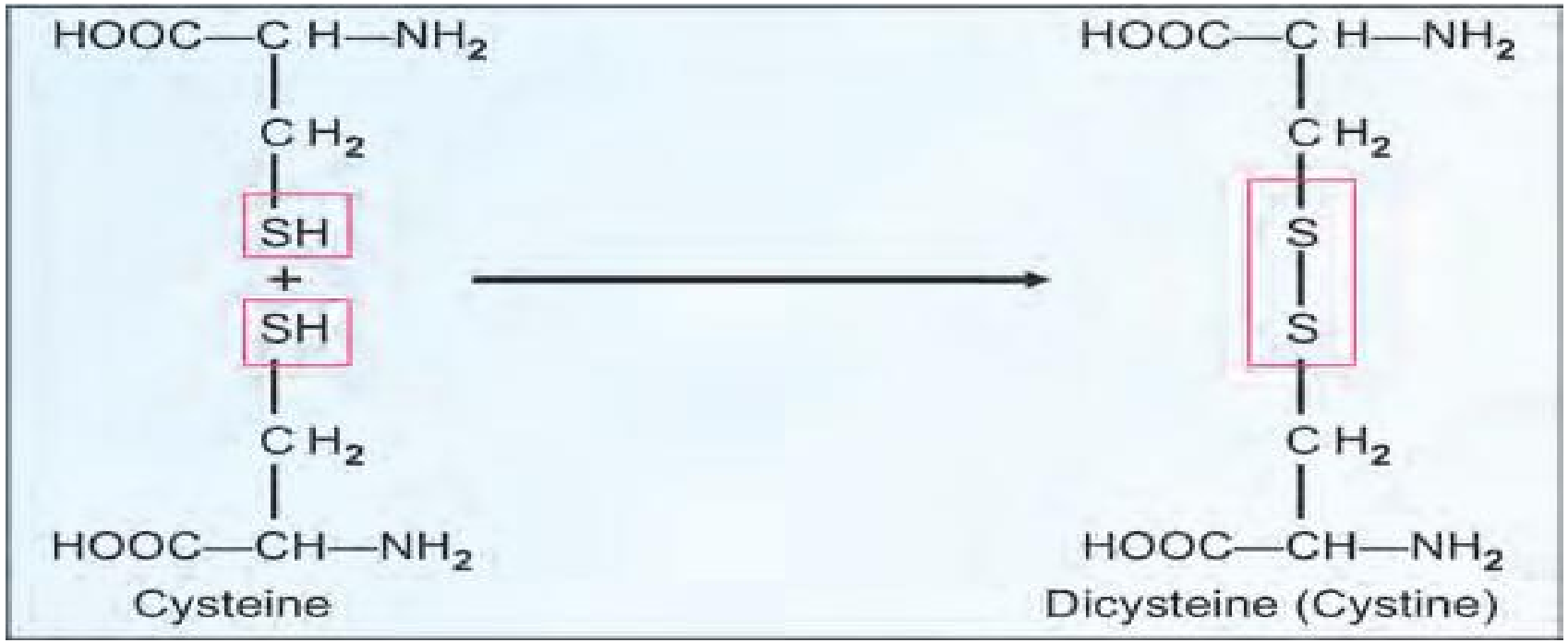
Reaction of SH group

Cysteine forming a disulfide bond with another cysteine residue.

Two cysteine residues connect two polypeptide chains-inter chain disulfide bonds or links

Dimer of two cysteine residues (Cystine or Dicysteine)

FORMATION OF DISULFIDE BRIDGES



Amino acid derivatives of Importance

- Gamma amino butyric acid (GABA)-derivative of glutamic acid and dopamine (derived from tyrosine) is a neurotransmitter
- Histamine synthesized from histidine- a mediator of allergic reactions
- Thyroxine (from tyrosine an important thyroid hormone
- Ornithine and citrulline are derivatives of arginine and are essential for urea synthesis

PEPTIDE BOND FORMATION

Alpha carboxyl group of one amino acid reacts with alpha amino group of another amino acid to form a peptide bond or CO-NH bridge

Proteins are made by polymerization of amino acids through peptide bonds.

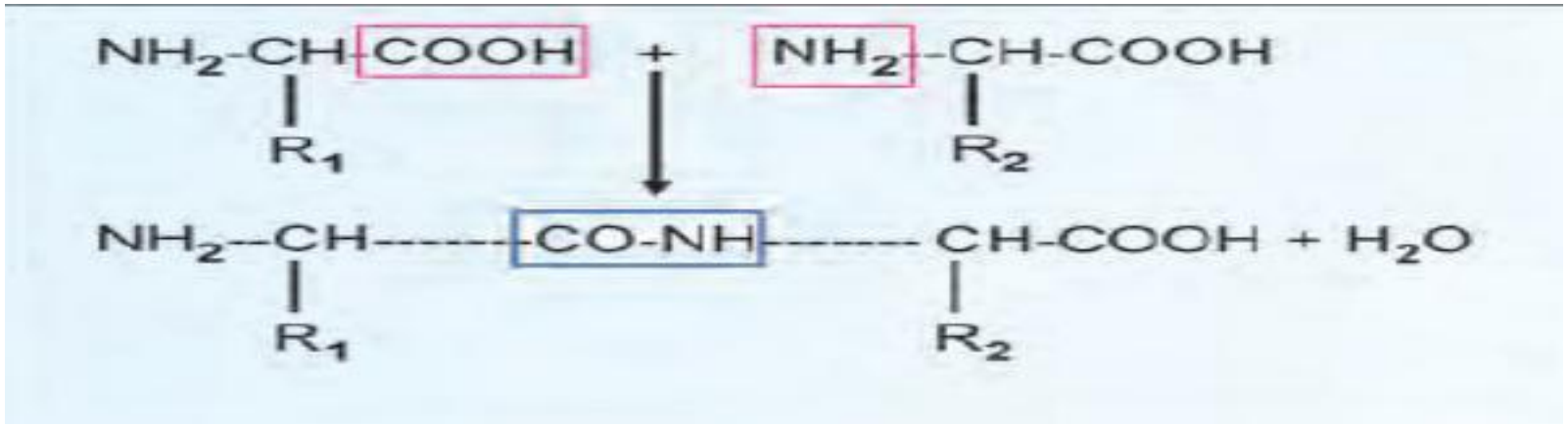


Fig: Peptide bond formation

Reaction	*Answered by specific group*
1. Ninhydrin	Alpha amino group
2. Biuret reaction	Peptide bonds
3. Xanthoproteic test	Benzene ring (Phe,Tyr, Trp)
4. Millon's test	Phenol (Tyrosine)
5. Aldehyde test	Indole (Tryptophan)
6. Sakaguchi's test	Guanidinium (Arginine)
7. Sulfur test	Sulfhydryl (Cysteine)
8. Nitroprusside test	Sulfhydryl (Cysteine)
9. Pauly's test	Imidazole (Histidine)

Table: Colour reactions of amino acids

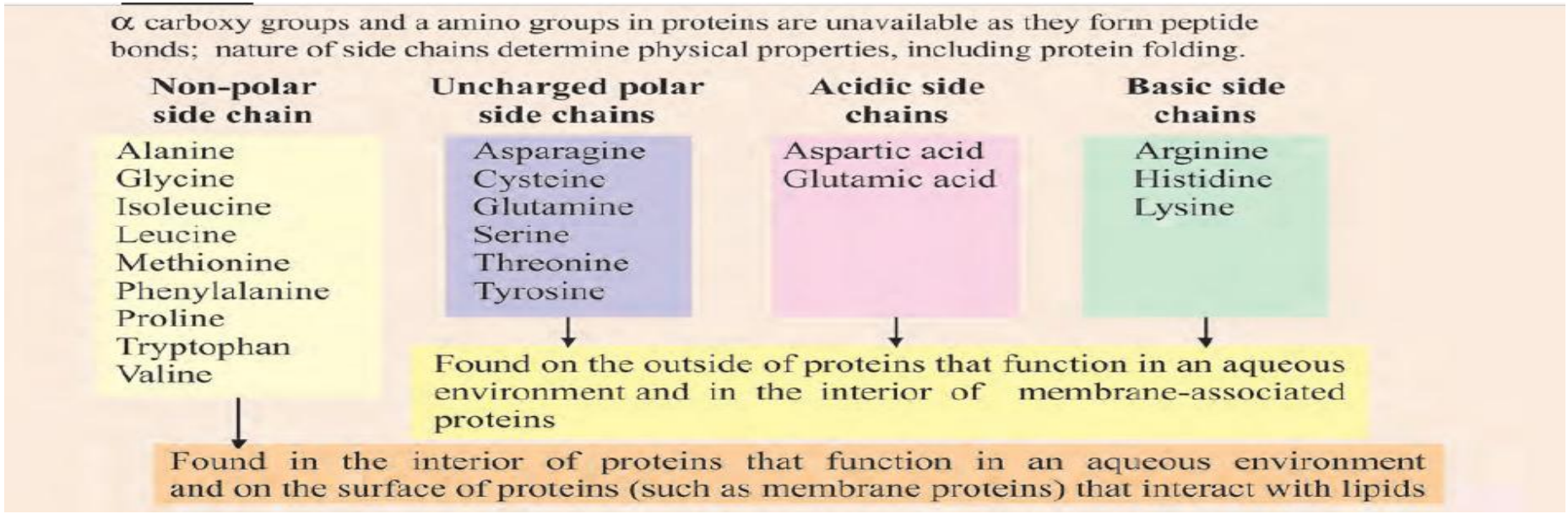


Table: Importance of side chains of amino acids

Selenocysteine-21st amino acid present in human proteins. An amino acid is given the individual status, when it is incorporated as such into proteins during protein biosynthesis, and having a separate codon. Selenocysteine is present in some enzymes. Instead of SH (sulfhydryl) group in cysteine, SeH (selenium) is present in selenocysteine. Abbreviated as SeCys or SeC.

About 25 proteins incorporate selenocysteine.

Pyrrolysine (Pyl)-the 22nd amino acid.

It is a lysine in an amide linkage to substituted-pyrroline-5-carboxylate. Present in methyl transferase enzymes of certain bacteria.

Both SeC and Pyl are encoded by codons that normally function as stop signals.

PROTEINS

Proteois (greek)- primary

- Paramount importance in biological systems
- 3/4 of the total body weight
- Body building (major structural and functional aspects of the body)
- Abnormality- lead to diseases

- Contain Carbon, Hydrogen, Oxygen and Nitrogen-Major components
- Sulphur, phosphorous-minor
- Nitrogen – 16% weight
- Polymer of amino acids

Reaction- Alpha carboxyl group reacts with alpha amino group of another amino acid- Peptide bond (CO-NH bridge)- polymerization of amino acids

- Dipeptide
- Tripeptide
- Tetrapeptide
- Oligopeptide
- 10 to 50-Oligopeptide
- Above 50 -proteins

Tripeptides

- 20^3- 8000 different permutations and combinations possible

- Ordinary proteins- 20^{100} possibilities.
- Changing sequence gives enormous number of markedly different proteins

STRUCTURE OF PROTEINS

Organisation

Different levels of structural organization

- Primary
- Secondary
- Tertiary
- Quaternary

- Primary structure- Number and sequence of amino acids in the protein
- Note: Higher levels of organization determined by the primary structure
- Covalent peptide bonds maintain the primary structure

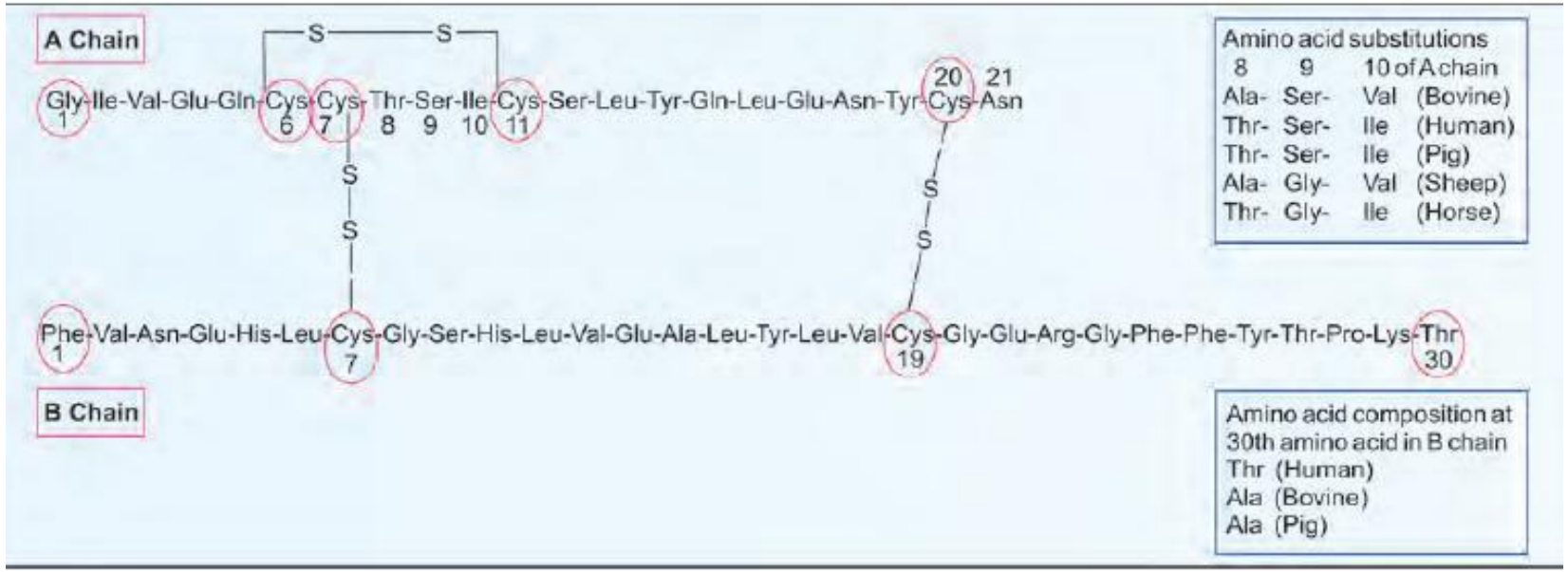


Fig: Primary structure of human insulin

Insulin has two polypeptide chains.

- The A chain (Glycine chain) has 21 aas and B (Phenylalanine) has 30 aas
- Held together by two interchain disulfide bonds. A chain 7^{th} cysteine and B chain 7^{th} cysteine are connected.
- Similarly, A chain 20^{th} cysteine and B chain 19^{th} cysteine are connected.
- There is another intrachain disulfide bond between 6^{th} and 11^{th} cysteine residues of A chain.

Species variation restricted to amino acids in position 8, 9 and 10 in A chain and in C-terminal of B chain.

Amino acid sequence conserved to a great extent during evolution

Pro-insulin

Beta cells of pancreas synthesize insulin as a prohormone.

- Proinsulin- single polypeptide chain with 86 aas.
- Biologically active insulin (2 chains) formed by removal of the central portion of the pro-insulin before release.
- The C-peptide (connecting peptide) is also released into the circulation (Fig.).

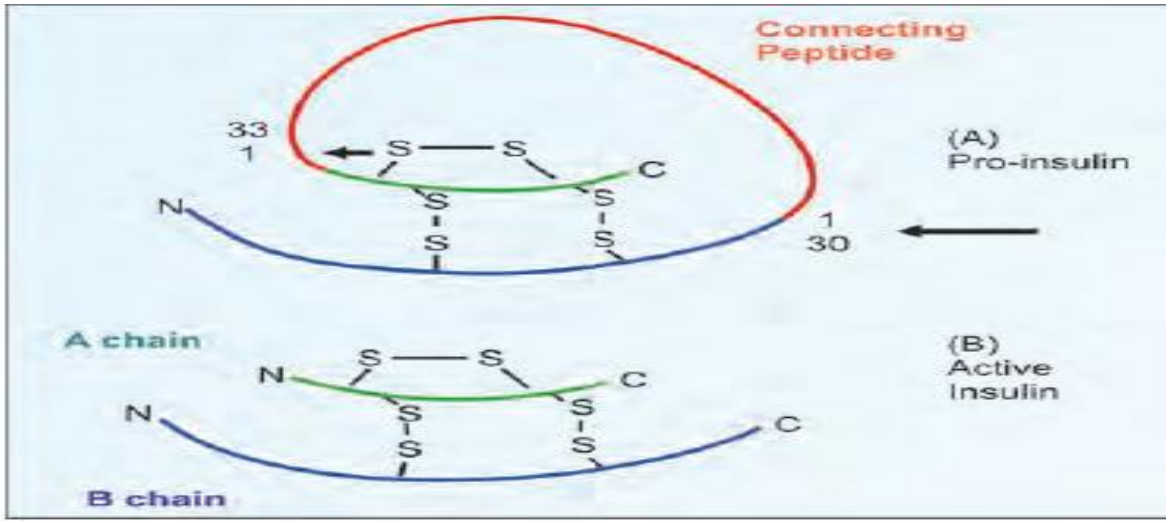


Fig: Conversion of pro-insulin to active insulin. Arrows = site of action of proteolytic enzymes

Primary Structure Determines Biological Activity

- A protein with a specific primary structure will automatically form its natural three dimensional shape.
- Higher levels of organization: Dependent on the primary structure.
- Even a single amino acid change (mutation) in the linear sequence may have profound biological effects on the function, e.g. in HbA (normal hemoglobin) the 6^{th} amino acid in the beta chain is glutamic acid; it is changed to valine in HbS (sickle cell anemia).

Secondary Structure of Proteins

Denotes configurational relationship between residues, which are about 3-4 amino acids apart in the linear sequence.

Secondary and tertiary levels of protein structure preserved by noncovalent forces or bonds like hydrogen bonds, electrostatic bonds, hydrophobic interactions and van der Waals forces.

A hydrogen bond-weak electrostatic attraction between one electronegative atom like O or N and a hydrogen atom covalently linked to a second electronegative atom.

Hydrogen atoms can be donated by -NH (imidazole, indole, peptide); -OH (serine, threonine) and $-NH_2$ (arginine, lysine).

Hydrogen accepting groups are COO^- (aspartic, glutamic), C =O (peptide); and S–S (disulfide).

Electrostatic bonds (ionic bonds): Positive charges donated by epsilon amino group of lysine, guanidinium group of arginine and imidazolium group of histidine.

Negative charges provided by beta and gamma carboxyl groups of aspartic and glutamic acids.

Hydrophobic bonds formed by interactions between nonpolar hydrophobic side chains by eliminating water molecules. This serves to hold lipophilic side chains together.

van der Waals forces-very weak, collectively contribute maximum towards the stability of protein structure.

Alpha helix

- i. Most common and stable conformation for a polypeptide chain. In proteins like hemoglobin and myoglobin, the alpha helix is abundant, whereas it is virtually absent in chymotrypsin.
- ii. The alpha helix is a spiral structure (Fig). The polypeptide bonds form the back-bone and the side chains of amino acids extend outward.
- iii. The structure is stabilized by hydrogen bonds between NH and C=O groups of the main chain.
- iv. Each turn is formed by 3.6 residues. The distance between each amino acid residue (translation) is 1.5 Å.
- v. The alpha helix is generally right handed. Left handed alpha helix is rare, because amino acids found in proteins are of L-variety, which exclude left handedness. Proline and hydroxy proline will not allow the formation of alpha helix.

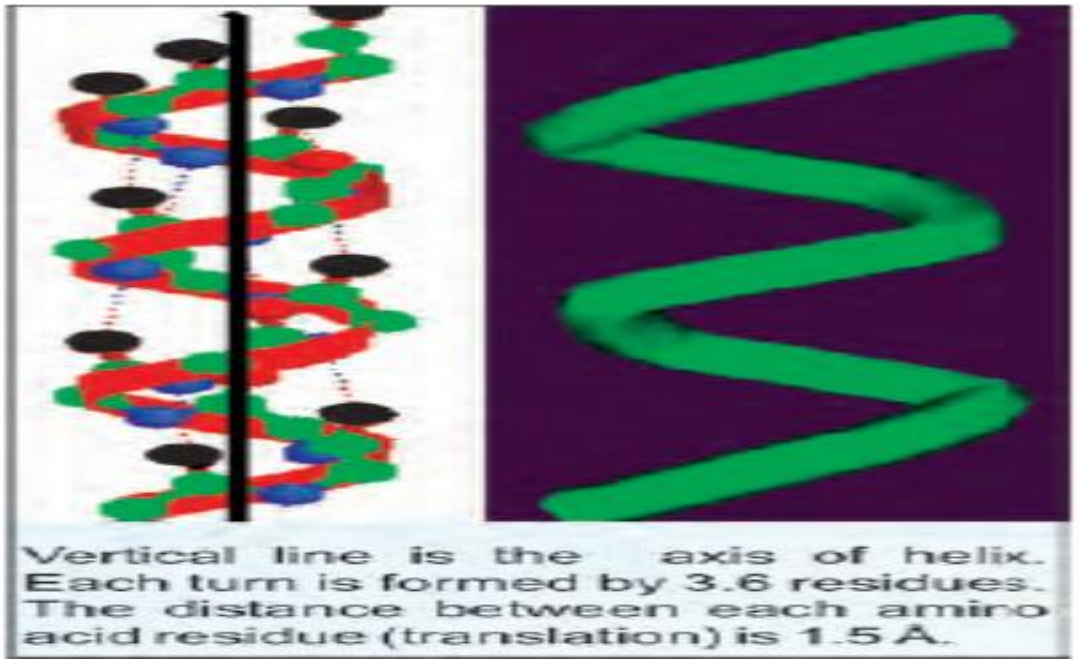


Fig: Structure of alpha helix

Beta-Pleated Sheet

The polypeptide chains in beta-pleated sheet is almost fully extended. Distance between adjacent amino acids is 3.5Å.

- Stabilized by hydrogen bonds between NH and C=O groups of neighboring polypeptide segments.
- Adjacent strands in a sheet can run in the same direction with regard to the amino and carboxy terminal ends of the polypeptide chain (parallel) or in opposite direction (anti-parallel beta sheet) (Fig.). Major structural motif in proteins like silk Fibroin (anti-parallel), Flavodoxin (parallel) and Carbonic anhydrase (both).
- Beta bends may be formed in many proteins by the abrupt U-turn folding of the chain. Intrachain disulfide bridges stabilize these bends.

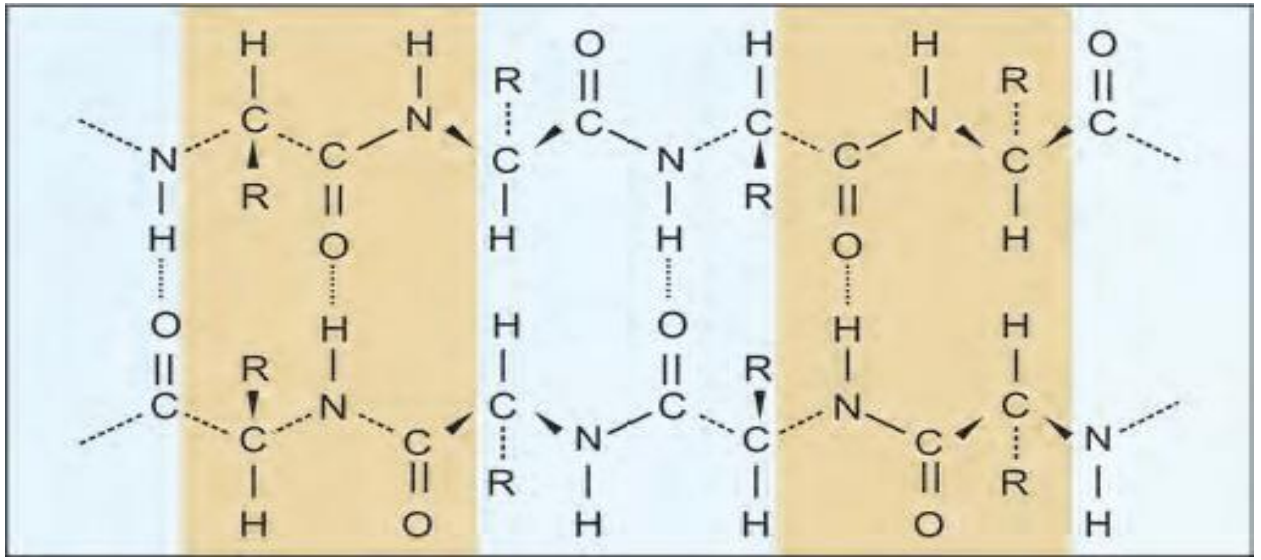


Fig: Structure of beta-pleated sheet

- *Collagen Helix*-triple helical structure found in collagen.

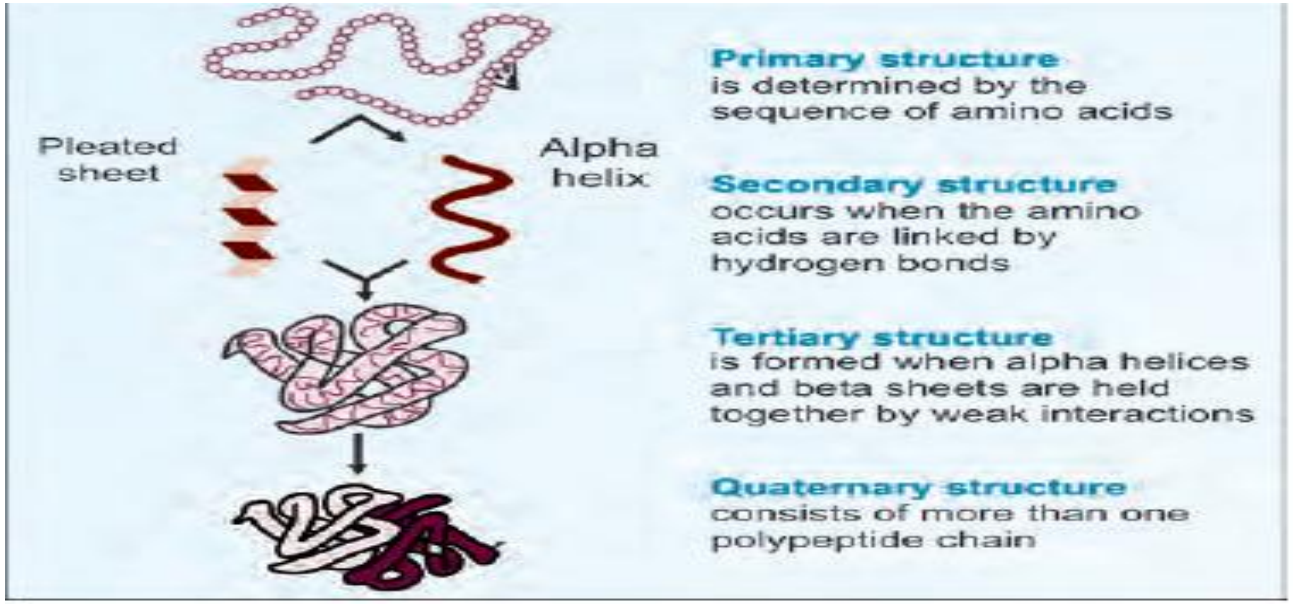


Fig: Levels of organization of proteins

Tertiary Structure

- Secondary structure denotes the configurational relationship between residues which are about 3-4 amino acids apart; or secondary level defines the organization at immediate vicinity of amino acids.
- The tertiary structure denotes three dimensional structure of the whole protein.
- Defines the steric relationship of amino acids which are far apart from each other in the linear sequence, but are close in the three-dimensional aspect.

- Maintained by **noncovalent** interactions such as hydrophobic bonds, electrostatic bonds and van der Waals forces.

Note: Tertiary structure acquired by native protein is always thermodynamically most stable.

- **Domain-**denotes a compact globular functional unit of a protein. Relatively independent region of the protein, and may represent a functional unit.
- Usually connected with relatively flexible areas of protein like immunoglobulins,
- Phenyl alanine hydroxylase enzyme contains 3 domains, one regulatory, one catalytic and one protein-protein interaction domains.

- The arrangements of the tertiary structure elements in a protein form a “fold”.
- A typical example: Calmodulin, the calcium binding regulatory protein which regulates intracellular calcium level.

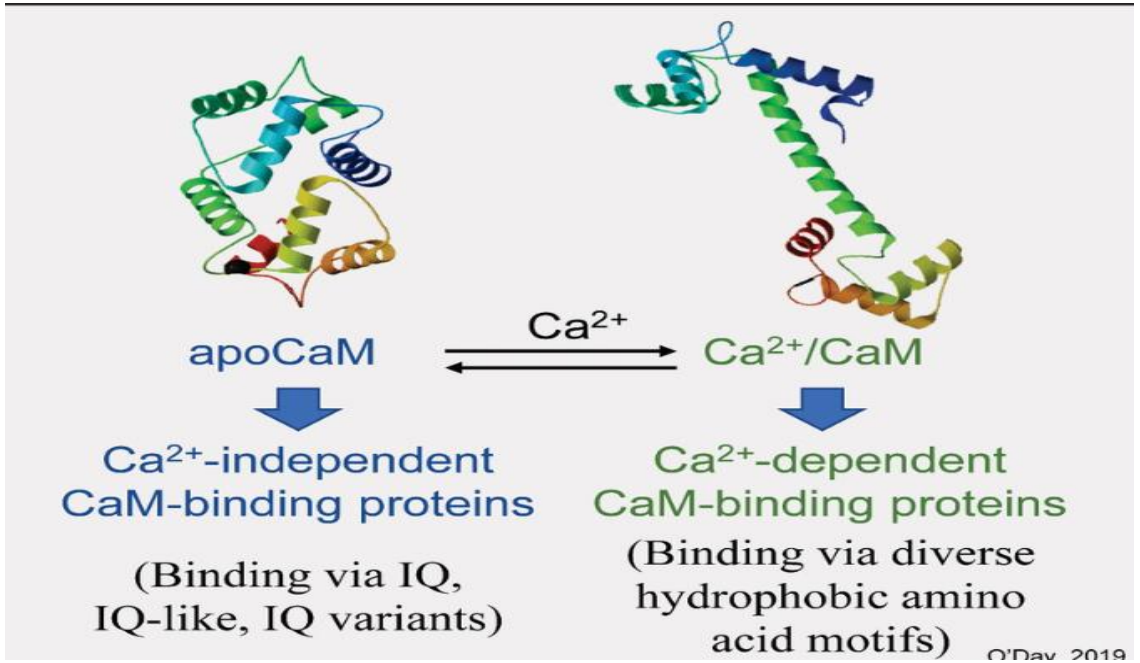


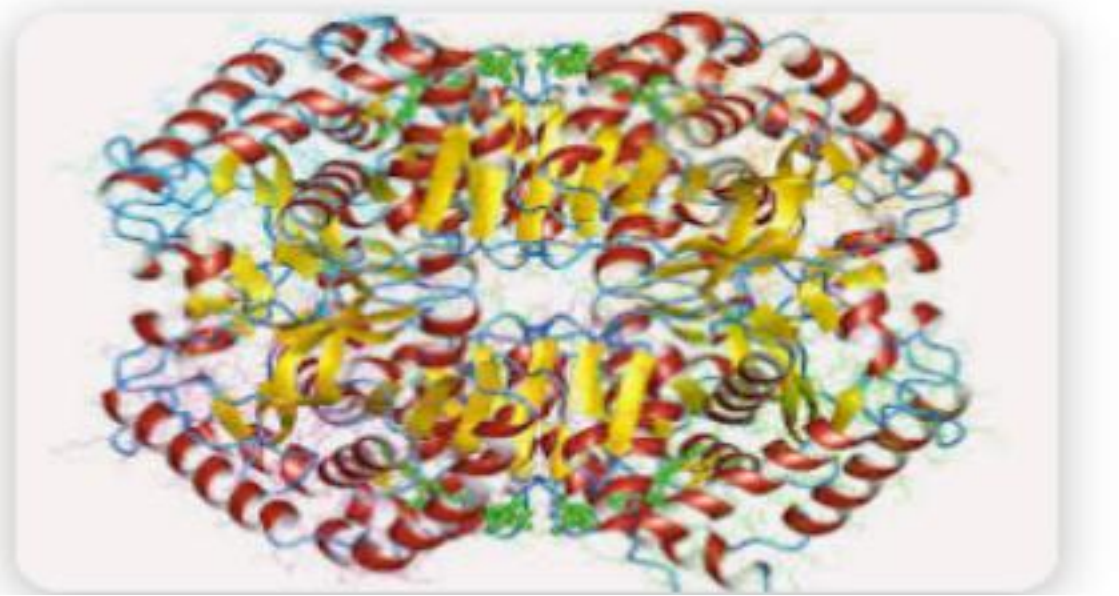
Fig: Structure of Calmodulin

Protein	*Structural motif present*
Myoglobin	Alpha helix and beta pleated sheet
Collagen	Triple helix
Keratin	Coiled coil
Elastin	No specifc motif
Superoxide dismutase	Antiparallel beta pleated sheet

TABLE: Specific structural motifs in common proteins

Quaternary Structure

- Certain polypeptides will aggregate to form one functional protein.
- The protein will lose its function when the subunits are dissociated.
- Forces that keep the quaternary structure are hydrogen bonds, electrostatic bonds, hydrophobic bonds and van der Waals forces.
- Depending on the number of polypeptide chain, the protein may be termed as monomer (1 chain), dimer (2 chains), tetramer (4 chains) and so on. Each polypeptide chain is termed as subunit or monomer.
- Homodimer contains two copies of the same polypeptide chain. Heterodimer contains two different types of polypeptides as a fuctional unit.
- v. For example, 2 alpha-chains and 2 beta-chains form the hemoglobin molecule. Similarly, 2 heavy chains and 2 light chains form one molecule of immunoglobulin G. Creatine kinase (CK) is a dimer.
- Lactate dehydrogenase (LDH) is a tetramer.



- Fig: Lactate dehydrogenase tetramer

Structure-function Relationship

Functions of proteins are maintained because of their ability to recognise and interact with a variety of molecules.

Three-dimensional structural conformation provides and maintains the functional characteristics.

Hydrophilic polar charged residues are seen on the outer surface. Non-polar hydrophobic residues inside, out of contact with water.

The three-dimensional structure, in turn, is dependent on the primary structure. So, any difference in the primary structure may produce a protein which cannot serve its function.

Enzymes

- The first step in enzymatic catalysis is the binding of the substrate to the enzyme. This, in turn, depends on the structural conformation of the active site of the enzyme, which is precisely oriented for substrate binding. **Carbonic anhydrase** catalyses the reversible hydration of carbon dioxide. Enzyme makes it possible for the precise positioning of the CO_2 molecule and the hydroxyl (OH^-) ion for the formation of bicarbonate ion. The zinc ion is located at a deep cleft coordinated to histidine residues. The CO_2 binding residues are very near to the zinc ion. Water binds to zinc ion, gets ionized to hydroxyl ion and it binds to the CO_2 which is proximally located. The substrates are brought in close proximity for the reaction to proceed.

Transport Proteins

Hemoglobin, the transporter of oxygen is a tetrameric protein (alpha 2, beta 2), with each monomer having a heme unit. Binding of oxygen to one heme facilitates oxygen binding by other subunits. Binding of H^+ and CO_2 promotes release of O_2 from hemoglobin.

- This allosteric interaction is physiologically important, and is termed as Bohr effect. Even a single amino acid substitution alters the structure and thereby the function, e.g. in sickle cell anemia (HbS), the 6^{th} amino acid in the beta chain is altered, leading to profound clinical manifestations.

Structural Proteins

Collagen-most abundant protein in mammals and is the main fibrous component of skin, bone, tendon, cartilage and teeth.

Forms a superhelical cable where the 3 polypeptide chains are wound around itself.

Every 3rd residue is a glycine. The only amino acid that can fit into the triple stranded helix is glycine. Replacement of the central glycine by mutations can lead to brittle bone disease.

- The triple helix of collagen is stabilized by the steric repulsion of the rings of hydroxyproline and also by the hydrogen bonds between them. In vitamin C deficiency failure of hydroxylation of proline/lysine leads to reduced hydrogen bonding and consequent weakness of collagen.
- The quarter staggered triple helical structure of collagen is responsible for its tensile strength. Different arrangements of collagen fibrils in tissues are seen. Parallel bundles in tendons and sheets layered at many angles in skin. Heat denatured collagen is gelatin.

STUDY OF PROTEIN STRUCTURE

- The first protein to be sequenced was insulin by Sanger in1955.
- Before studying the structure, first a pure sample of the protein has to be available.
- Proteins are extracted and purified by various chromatography techniques (ion exchange, adsorption, partition, size exclusion, affinity, HPLC). The purity of the protein thus isolated is studied by electrophoresis (agar, PAGE, iso electric focusing). Further, molecular weight is determined by mass spectroscopy.

- Mammals contain alpha keratin. It is classified into soft and hard keratin depending on the sulfur content. The cysteine residues are responsible for disulfide bridge formation which confers characteristic texture for each type of the protein. Soft keratin having low sulfur content is present in skin. Hard keratin is present in hair, horn and nails and has high sulfur content. The disulfide bridges resist the forces that try to deform them. Springiness of hair is due to the characteristic coiled coil structural motif. When stretched, the coiled coil will untwist and resume the original structure. Hair styling like curling and straightening is based on reduction of the existing disulfide bond and then reoxidation so that new bonds are formed. Stretching using moist heat, breaks disulfide bonds. Abnormalities in keratin structure will cause loss of skin integrity and results in diseases like Epidermolysis bullosa (fragile skin and blisters).

PHYSICAL PROPERTIES OF PROTEINS

- Protein solutions exhibit colloidal properties and therefore scatter light and exert osmotic pressure. Osmotic pressure of plasma proteins is clinically important
- Molecular weights of some of the proteins are: Insulin (5,700); Hemoglobin (68,000); Albumin (69,000); Immunoglobulins (1,50,000); Rabbit Papilloma Virus Protein (4,70,00,000).
- Shape of the proteins also vary. Thus, Insulin is globular, Albumin is oval in shape, while Fibrinogen molecule is elongated. Bigger and elongated molecules will increase the viscosity of the solution.
- 4. Isoelectric pH. Since proteins are made of amino acids, the pI of all the constituent amino acids will influence the pI of the protein.

PRECIPITATION REACTIONS OF PROTEINS

- Purification of enzymes and other proteins usually start with precipitating them from solution. The stability of proteins in solution will depend mainly on the charge and hydration. Polar groups of the proteins (-NH_2, COOH, OH groups) tend to attract water molecules around them to produce a shell of hydration. Any factor, which neutralizes the charge or removes water of hydration will therefore cause precipitation of proteins. The following procedures are used for protein precipitation:
- Salting Out
- When a neutral salt, such as ammonium sulfate or sodium sulfate is added to protein solution, the shell of hydration is removed and the protein is precipitated. This is called salting out. As a general rule, higher the molecular weight of a protein, the salt required for precipitation is lesser.
- Thus globulins are precipitated at half saturation of ammonium sulfate; but albumin will need full saturation of ammonium sulfate.
- Isoelectric Precipitation Proteins are least soluble at their isoelectric pH. Some proteins are precipitated immediately when adjusted to their isoelectric pH.

- The best example is **Casein** which forms a flocculent precipitate at pH 4.6 and re-dissolves in highly acidic or alkaline solutions. When milk is curdled, the casein forms the white curd, because lactic acid produced by the fermentation process lowers the pH to the isoelectric point of casein

Precipitation by Organic Solvents

When an organic solvent is added to the protein solution, water molecules available for proteins are reduced, and precipitation occurs. Organic solvents reduce the dielectric constant of the medium which also favors protein precipitation.

- Hence **alcohol** is a powerful protein precipitating agent. This may explain the disinfectant effect of alcohol.

Precipitation by Heavy Metal Ions

- In alkaline medium, proteins have net negative charge, or are anions. To such a solution, if salts of heavy metals are added, positively charged metal ions can complex with protein molecules and metal proteinates are precipitated.
- Salts of **Copper, Zinc, Lead, Cadmium and Mercury** are toxic, because they tend to precipitate normal proteins of the gastrointestinal wall. Based on this principle, **raw egg** is sometimes used as an **antidote** for mercury poisoning.

Precipitation by Alkaloidal Reagents

- Tungstic acid, phosphotungstic acid, trichloro acetic acid, picric acid, sulfosalicylic acid and tannic acid are powerful protein precipitating agents. These acids lower the pH of medium, when proteins carry net positive charges. These protein cations are complexed with negatively charged ions to form protein-tungstate, protein-picrate, etc., and thick flocculent precipitate is formed. In clinical laboratory phosphotungstic or trichloro acetic acid are usually used for precipitating proteins. Tanning in leather processing is based on the protein precipitating effect of tannic acid.
- Under certain conditions, proteins undergo denaturation, which is a mild form of precipitation reaction.
- Heat coagulation is an irreversible precipitation process.

CLASSIFICATION OF PROTEINS

Almost impossible to correctly classify all proteins.

- **Classification Based on Functions**

➢Catalytic proteins, e.g. enzymes
➢Structural proteins, e.g. collagen, elastin
➢Contractile proteins, e.g. myosin, actin.
➢Transport proteins, e.g. hemoglobin, myoglobin, albumin, transferrin
➢Regulatory proteins or hormones, e.g. ACTH, insulin, growth hormone
➢Genetic proteins, e.g. histones
➢Protective proteins, e.g. immunoglobulins, interferons, clotting factors

Classification based on Composition and Solubility

Simple Proteins

- According to definition, they contain only amino acids.

i. Albumins: Soluble in water and coagulated by heat. Human serum albumin has a molecular weight of 69,000, e.g. lactalbumin of milk and egg albumin.

ii. Globulins: Insoluble in pure water, but soluble in dilute salt solutions. Also coagulated by heat, e.g. egg globulin, serum globulins, legumin of peas.

iii. Protamines: Soluble in water, dilute acids. They are not coagulated by heating. They contain large number of arginine and lysine residues, and so are strongly basic. Hence they can combine with other acidic proteins. Protamine zinc insulinate is a common commercial preparation of insulin.

iv. Prolamines: Soluble in 70-80% alcohol, but insoluble in pure water. They are rich in proline but lack in lysine, e.g. zein from corn, gliadin of wheat, hordein of barley.

- v. **Lectins:** Lectins are precipitated by 30–60% **ammonium sulfate**.

Proteins having high affinity to sugar groups. Lectin from *Dolichos biflorus* will agglutinate human blood group A1 RBCs. Phytohemagglutinin (PHA), a lectin from *Phaseolus vulgaris* (red kidney bean) agglutinates all RBCs and WBCs. Concanavalin-A (ConA) from legumes will specifically attach to mannose and glucose. The attachment of lectins on normal and cancer cells will be different, this property is sometimes utilized to differentiate cancer cells.

- vi. **Scleroproteins:** Insoluble in water, salt solutions and organic solvents and soluble only in **hot strong acids**. Form supporting tissues, e.g. collagen of bone, cartilage and tendon; keratin of hair, horn, nail and hoof.

Conjugated Proteins

Combinations of protein with a non-protein part, called **prosthetic group**. Conjugated proteins may be classified as follows:

CONJUGATED PROTEINS

- i. **Glycoproteins:** These are proteins combined with carbohydrates. Hydroxyl groups of serine or threonine and amide groups of asparagine and glutamine form linkages with carbohydrate residues. **Blood group antigens** and many serum proteins are glycoproteins. When the carbohydrate content is more than 10% of the molecule, the viscosity is correspondingly increased; they are sometimes known as **mucoproteins** or proteoglycans.
- ii. **Lipoproteins:** These are proteins loosely combined with lipid components. They occur in blood and on cell membranes.
- iii. **Nucleoproteins:** These are proteins attached to nucleic acids, e.g. Histones. The DNA carries negative charges, which combines with positively charged proteins.

- iv. **Chromoproteins:** These are proteins with coloured prosthetic groups. Hemoglobin (Heme, red); Flavoproteins (Riboflavin, yellow), Visual purple (Vitamin A, purple) are some examples of chromoproteins.
- v. **Phosphoproteins:** These contain phosphorus. **Casein** of milk and **vitellin** of egg yolk are examples. The phosphoric acid is esterified to the hydroxyl groups of serine and threonine residues of proteins.
- vi. **Metalloproteins:** They contain metal ions. Examples are Hemoglobin (Iron), Cytochrome (Iron), Tyrosinase (Copper) and Carbonic anhydrase (Zinc).

Derived Proteins

- They are degradation products of native proteins. Progressive hydrolysis of protein results in smaller and smaller chains: Protein → peptones → peptides → amino acids.
- **Classification Based on Shape**
- *Globular Proteins*

Spherical or oval in shape, easily soluble, e.g. albumins, globulins and protamines

- *Fibrous Proteins*
- Molecules are elongated or needle shaped. Solubility is minimum, resist digestion. Collagen, elastin and keratins are examples.

Conjugated Protein	*Protein part*	*Prosthetic group*
Hemoglobin	Globin	Heme
Nucleoprotein	Histones	DNA
Rhodopsin	Opsin	11-cis-retinal
Succinate dehydrogenase	Protein	Riboflavin as FAD
Ferritin	Apoferritin	Iron

TABLE: Examples of conjugated proteins

Classification Based on Nutritional Value

Nutritionally Rich Proteins

Called **complete proteins or first class proteins**. Contain all the essential amino acids in the required proportions. On supplying these proteins in the diet, children will grow satisfactorily. A good example is **casein** of milk.

Incomplete Proteins

- They **lack one essential amino acid.** They cannot promote body growth in children; but may be able to sustain the body weight in adults. Proteins from **pulses are deficient in methionine**, while proteins of **cereals lack in lysine**. If both of them are combined in the diet, adequate growth maybe obtained.

- *Poor Proteins*

They **lack in many essential amino acids** and a diet based on these proteins will not even sustain the original body weight. Zein from corn lacks tryptophan and lysine.

- **Biologically Important Peptides**

When 10 or less number of amino acids are joined together, it is called an oligopeptide. Some of them are biologically active. A few examples are given below:

- i. **Thyrotropin releasing hormone** (TRH) is a tripeptide with the sequence of Glu-His-Pro; but the Glu and Pro are modified.
- ii. **Glutathione** is a tripeptide. It is gamma glutamyl cysteinyl glycine. It is involved in erythrocyte membrane integrity and is important in keeping enzymes in active state.
- iii. **Oxytocin and vasopressin** (ADH) are nanopeptides; with 9 amino acids. They are secreted by posteriorpituitary.

- iv. **Angiotensin I** has 10 amino acids and Angiotensin II has 8 amino acids. They are pressor agents; they elevate blood pressure.
- v. **Gramicidin S**, an antibiotic produced by *Bacillus brevis*, contains 10 amino acids. It is circular and contains D-phenyl alanine (usual proteins contain only L-amino acids).
- Polypeptide hormones (more than 10 amino acids)

QUANTITATIVE ESTIMATION OF PROTEINS

- **Biuret Method**

i. Cupric ions chelate with peptide bonds of proteins in alkaline medium produce a pink or violet color. The intensity of the color is proportional to the number of peptide bonds. The color is then compared with a standard protein solution treated with the biuret reagent, and estimated colorimetrically.

- ii. **Advantage:** The biuret method is simple one step process, and is the most widely used method for plasma protein estimations.
- iii. **Disadvantage:** The sensitivity of the method is less and is unsuitable for estimation of proteins in milligram or microgram quantities.

Proteomics

- Proteomics is the study of the entire galaxy of proteins produced by a cell under different conditions. At a particular time, a gene is “on” in a particular cell; but it will be “off” in another cell. Expression of proteins during growth and development will be different from the resting cell. Proteins produced by a gastro intestinal cell and a neuronal cell will be entirely different. Many proteins undergo post-translational modification, that too, at different levels at various organs. But genes are the same in

- all cells at all times. Therefore study of genes (genomics) will give only a partial picture of what is going on in nature. Even though DNA determines the basic genetic structure of an organism, it is the protein which actually carries out the body functions. Proteomics aims at studying the protein structure and function. A variety of techniques are available for studying proteins; these include mass spectrometry, NMR spectrometry, isoelectric focusing, etc. Human body contains hundreds of different cells, which express thousands of proteins, at different times and under the influence of different stimuli. Proteomics attempt to study this multifaceted picture in toto.