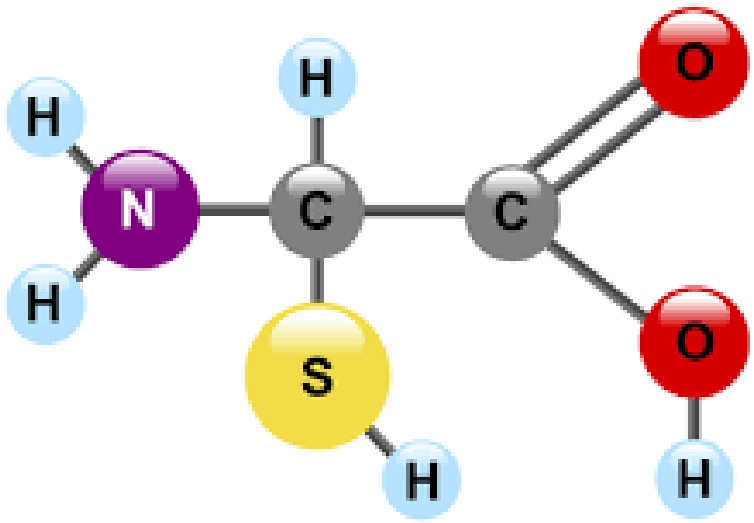
وزارة التعليم العالي والبحث العلمي
جامعة البصرة / كلية الصيدلة
فرع العلوم المختبرية والسريرية

Proteins

[Amino acids and peptides bonds]



By \ Dr : Basim Jasim ……University of Basra \ college of pharmacy

Chemistry of amino acids , peptides and proteins \

Overview

In blood stream , the Protein it is natural substance with high molecular weights (macromolecules , varying from 5000 to many millions Daltons. Besides , carbon hydrogen and oxygen , they also contain nitrogen and some times sulfur or phosphorous.

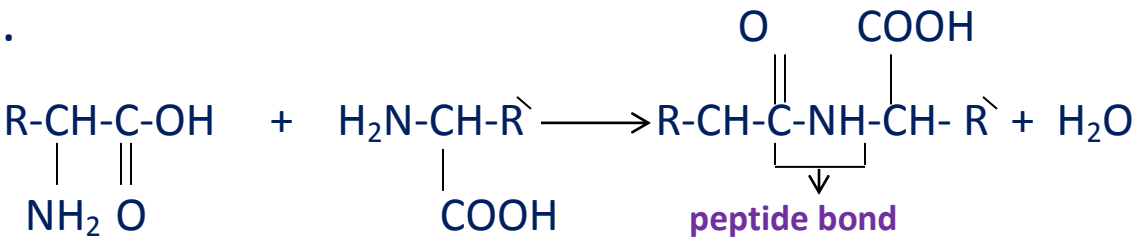
Proteins are most abundant and functionally a diverse molecules in the living system , virtually , every life process depends on this class of substances . e.x many of body proteins perform the most valuable function of catalyzing the innumerable chemical reactions constantly taking place inside the body , these protein are known as enzymes , which make up nearly 90 % of the total proteins in the cell .Another important example of proteins is polypeptide hormones , which play a critical role in the regulation of the metabolic pathways in the body. Also , the contractile proteins in the muscles permit the movement , whereas in the bones , the protein collagen forms a framework for the deposition of calcium phosphate crystals acting like the steel cables in reinforced concrete.

proteins , such as hemoglobin and plasma albumins , shuttle molecules essential to life . whereas immunoglobulin fight against the antigens such infectious bacteria or viruses.

Amino acids \

Proteins are large or huge molecules , but on hydrolysis , they produce a small molecules known as amino acids R-CH(NH_2)-COOH .These amino acids (principle structural units) are joined together through the combination of –NH group (amino acid) of one amino acid molecule with –COOH group(carboxyl group)

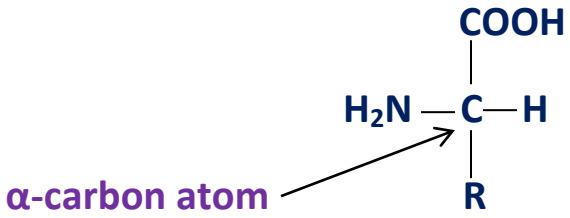
Of another one forming what is called peptide (amide) bond -CO-NH- .



Although about 300 types of amino acids are known to occur in the nature, only about 20 types are commonly found as constituents of mammalian proteins.

Each amino acid has a free carboxyl group (–COOH) ,free amino group (–NH_2 , except proline) , hydrogen atom (–H) , and distinctive side chain (R-group) , where all these groups are bonded to the α-carbon atom. This α-carbon atom asymmetric (chiral) atom in all amino acids except glycine (H_2N-CH_2-COOH) , in which the R-group is replaced H.

According to the presences of this asymmetrical carbon atom, every amino acid has two optical isomers; one isomer rotates the plane of polarized light to the right side (+) , and the another one to the left side (-).



Structure of amino acid

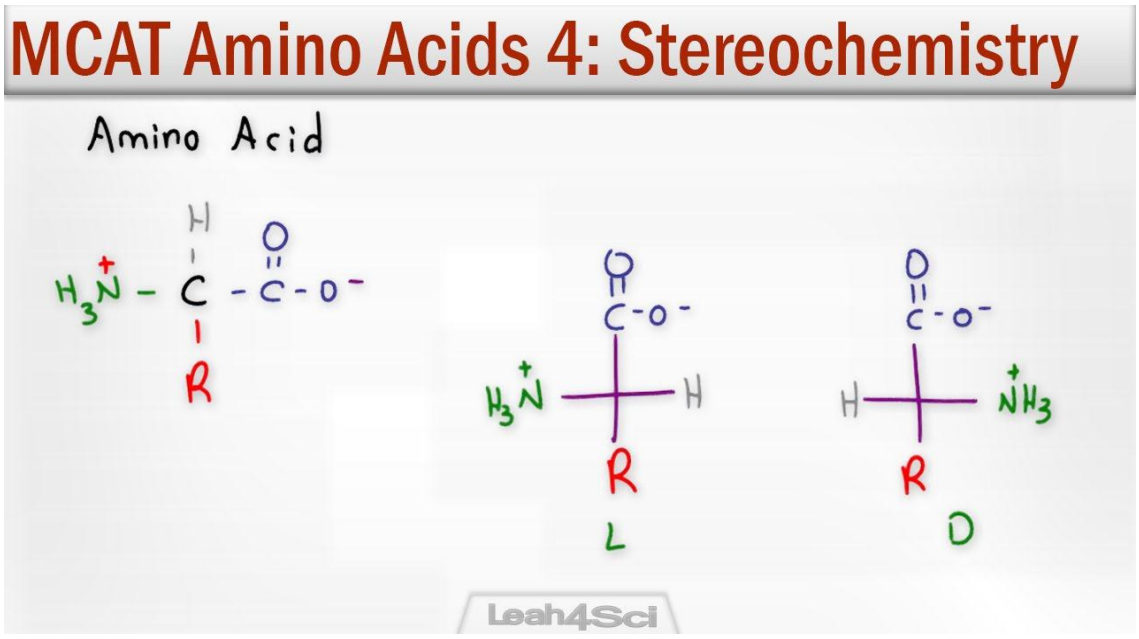
It may be recalled that the configurations of glyceraldehydes are also similar.

D (+) –glyceraldehydes L(-) -glyceraldehydes

Thus, as aldoses and ketoses, amino acid also have the D-series and the L-series of compounds , depending on the configurations of the α-carbon atom. The H-atom of α-carbon is on the left in the D-series , and is on the right in the L-series. The amino acids also do not show any relationship between these symbols (D and L) and the nature of the optical rotation.

Hence, it is necessary to write D or L to denote the configurations of the α-carbon and sign (+) or (-) to indicate the nature of rotation.

Although some amino acids of proteins are dextrorotatory and the other are levorotatory , all share the absolute configuration of the L–glyceraldehydes and thus , they frequently exist as L- α-amino acid.Rarely ,D-amino acids are found naturally , include free D-serine and D-aspartate in the brain tissue , and D-alanine and D-glutamate in the cell walls of gram-positive bacteria.



However, in amino acids , the specific rotation is governed by the PH of solution , temperature , type of solvent and presence of salt. Thus , there may be L(+) –alanine and also L(-) –alanine. The former may be prepared synthetically , and the latter is produced by protein hydrolysis, All amino acids obtained by enzyme hydrolysis of protein are L-isomer , but alkaline hydrolysis can produce DL-isomer (Racemic mixture).

- **There are many classification of amino acids , but the more suitable one is depend on the classification of R-group of amino acids according to aliphatic , aromatic , acidic and basic group or to certain atom or functional group which containing in the R-group side chain :-**

**1. Aliphatic amino acids **

a)Glycine (Gly.) ---- R= H .

b)Alanine (Ala.) ----- R=-CH_3 .

c)Valine (Val) ----- R= -CH(-CH_3)(-CH_3) .

d)Leucine (Leu) ----- R= -CH_2-CH(-CH_3)(-CH_3) .

e)Isoleucine (Ile) --- R= -CH(-CH_3)-CH_2-CH_3 .

**2. Amino acids with hydroxyl group (-OH) containing side chains **

a)Serine (Ser.) --- R=-CH_2OH.

b)Threonine (Thr.)--- R= -CH(-CH_3)-OH.

3. Amino acids with sulfur atom (-S) containing side chains \

a)Cysteine (Cys.) --R=- CH_2-SH.

b)Methionine(Meth.)—R= —(CH_2)— S —CH_3 .

4.Acidic amino acids \

a)Aspartic (Asp.) ---R=-CH_2COOH.

b) Asparagine(Asn) ---R=CH_2CONH_2.

c)Glutamic (Glu.) -----R=-$(CH_2)_2$-COOH.

d)Glutamine (Gln.)----R=-$(CH_2)_2$-$CONH_2$.

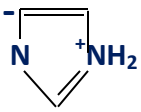
---

5. Basic amino acids \

a) Arginine (Arg.)--- R= —$(CH_2)_3$—NH—C($^+NH_2$)— NH_2.

b)Lysine (Lys.)--------R=—$(CH_2)_4$—$^+NH_3$.

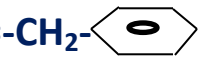
c) Histidine (His.) \ R= - CH-



---

6. Aromatic amino acids \

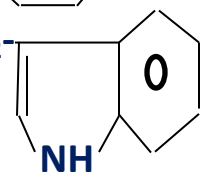
a)Phenylalanine (Phe.) \ R=-CH_2-



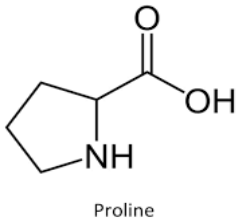
cb)Tyrosine (Tyr.) \ R=-CH_2-

- OH

c)Tryptophan (Trp)\ R=-CH_2-



7.Cyclic amino acids (imino acids) \



Proline

a)Proline (Pro.) \

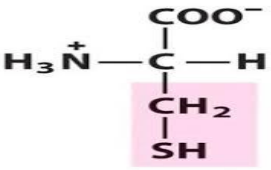
There are some acids produced from the principle Amino acid s (derivatives) :-

- 4-Hydroxyproline

Proline

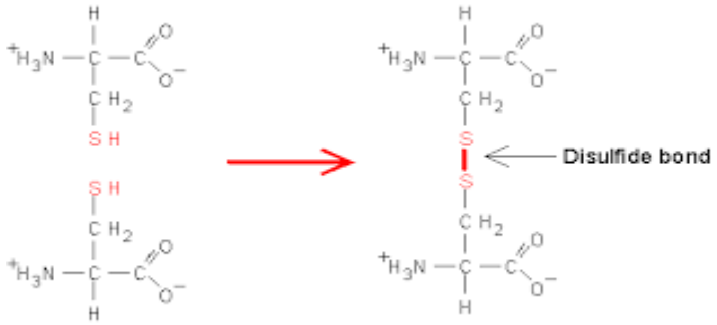
Proline **4-Hydroxyproline**

b)Cysteine \



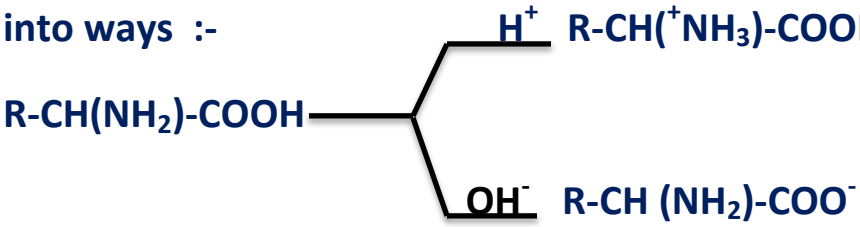
Cysteine

Two Cysteines bounded by oxidation



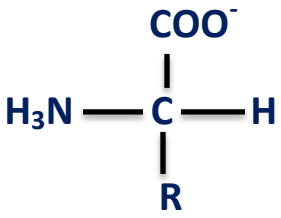
Physical and chemical properties of amino acids \

Amino acids (a.a) are colorless crystalline solids , which are generally soluble in water (except cysteine which is slight water insoluble). They are slightly soluble in alcohol and less in ether. Amino acids obtainable from protein hydrolysis and levorotatory. Every amino acid has a free carboxyl group and free amino group can be ionized into ways :-



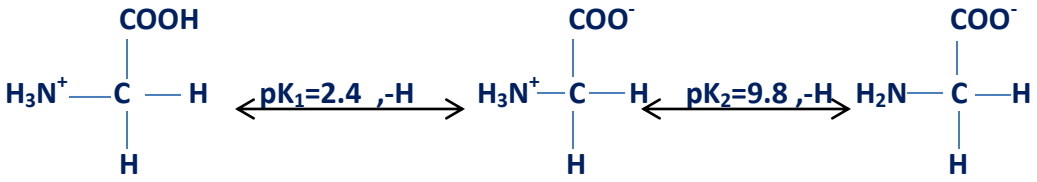
At physiological pH (pH= 7.4) , the carboxy group is dissociated forming negatively charged carboxylate ion ($-COO^-$) , Whereas the amino group is protonated as ($-^+NH_3$). IN protein , all the carboxyl and amino groups are combined peptide linkages , thus , The nature of side chain would be determined the role of amino acids in the proteins.

The molecules that contain equal number of ionizable groups opposite charges and therefore , the bear no net charge and could be termed as (Zwitter ion).

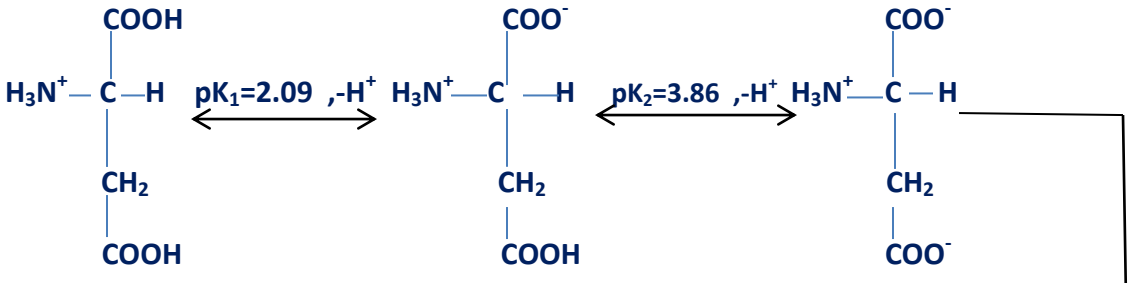


(Zwitter ion)

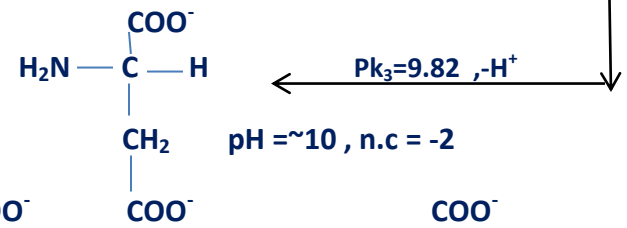
The protonic equilibriums of amino acids are shown below :-

- 

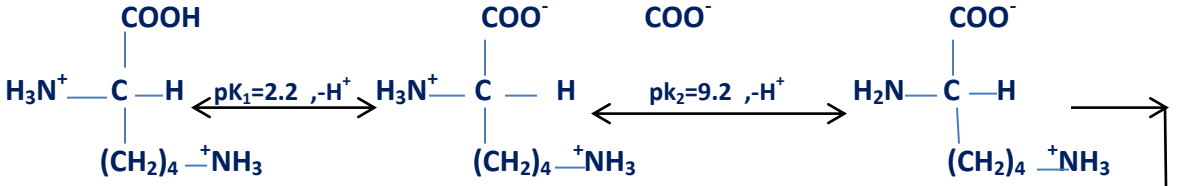
pH = ~1 , n.c=±1 pH =~3 ,n.c=0 pH=~8.5 , n.c=-1

- 

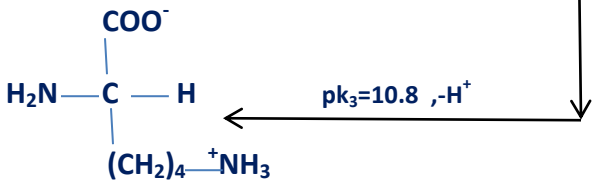
pH =~1 , n.c=+1 pH =~3, n.c=0 pH =~4 , n.c=-1



pH =~10 , n.c = -2

- 

pH =~1 , n.c=+2 pH =~3 , n.c=+1 pH =~7.5 , n.c=0



pH =~10 ,

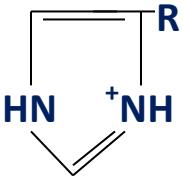
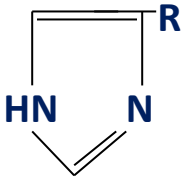
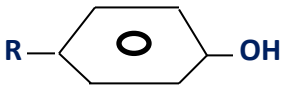
n.c=-1

The isoelectric species is the form of a molecule that has an equal number of positive and negative charges and thus, is electrically neutral. The isoelectric pH (pI) , is the midway between pKa values on either side of isoelectric species :-

(PI) of Glycine $= \frac{pK1 + pK2}{2} = \frac{2.4 + 9.8}{2} = 6.1$

(PI) of Aspartic $= \frac{pK1 + pK2}{2} = \frac{2.09 + 3.86}{2} =$

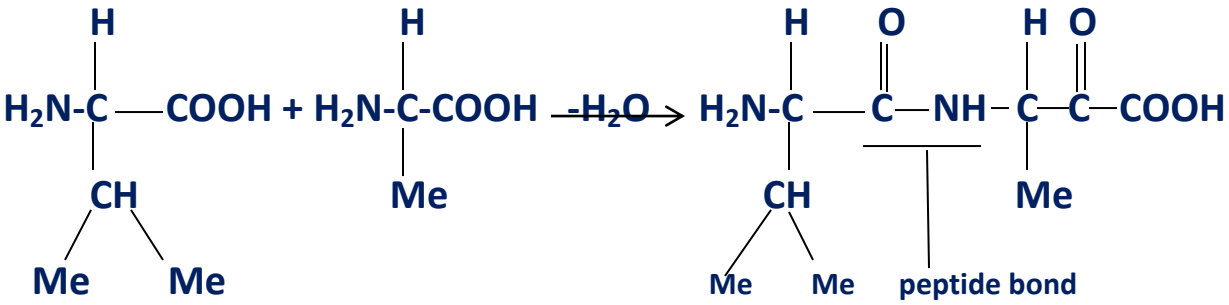
(PI) of Lysine $= \frac{pK2 + pK3}{2} = \frac{9.2 + 10.8}{2} =$

Group	Conjugate acid	Conjugate base	PKa
1) α - Carboxyl	R-COOH	$R\text{-}COO^-$	2.1 ± 0.5
2) non α – Carboxyl	R-COOH	$R\text{-}COO^-$	4.0 ± 0.3
3) Imidazole (Histidine)			0.6
4) α –amino	$R\text{-}^+NH_3$	$R\text{-}NH_2$	9.8 ± 1.0
5) E-amino (Lysine)	$R\text{-}^+NH_3$	$R\text{-}NH_2$	10.5
6) Phenolic OH (Tyr)			10.1
7) Guanido (Arg)	$R\text{-}NH\text{—}\underset{\underset{^+NH}{\|}}{C}\text{-}NH_2$	$R\text{-}NH\text{—}\underset{\underset{NH}{\|}}{C}\text{-}NH_2$	12.5
8)Sulfahydryl (Cys)	R-SH	$R\text{-}S^-$	8.3

Peptides \

They are a part of the proteins , which are joined covalently by peptide bonds (amide linkage). The peptide bonds are achieved between the α – amino group of another one to form a peptide or polypeptide (primary structure of protein), e.x \ Valine and alanine can form a dipeptide (valylalanine or Val-ala.) through the formation of peptide bond.

Peptides bonds are not broken at the same conditions that denature the proteins , such as heating or high concentrations of urea, whereas prolonged exposure to strong acids or bases at elevated temperatures is required to hydrolyzed these bonds non-enzymatically.

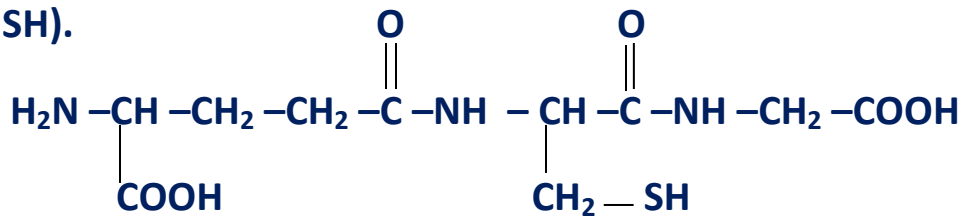


The sequences of amino acids in the peptides are achieved as following :-

The free amino end of the peptide chain (N-terminal) is written to the left and the free carboxy end (C-terminal) to the righ. Therefore , all amino acids sequences are read from the N-terminal to the C-terminal end to peptide chain. Frequently , the linkage of many amino acid through peptide bonds result in an unbranched chain which is called a polypeptide.

where each amino acids component in this polypeptide is celled (Residue) or Moiety. When the polypeptide is named , all amino acid residue that have the suffixes (-ine , -an , -ic , -ate) are converted to the suffix (-yl) with exception of amino acid at the C-terminal end e.x : a tripeptide composed of valine at N-terminal , glycine and Leucine at C-terminal is called as (Valyl-glycyl-leucine) or written (Val-Gly-Leu).

Animal ,plant and bacterial cells contain a variety of molecular weight polypeptides (3-100 amino acid residues) with profound physiological activities such as octapeptide vasopressin and oxytocin , ACTH ,and melanocyte-stimulating hormones. An example of one of the many known naturally occurring polypeptides is the tripeptide glutathione (GSH).



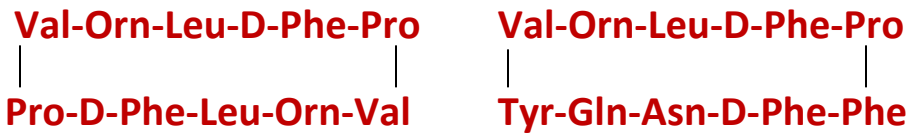
Glutathione (ɤ-glutamyl-cysteinyl-glycine)

Other important naturally occurring peptides include bradykinin and kallidin (Lysyl-bradykinin-), Which are vasodilators and hypotensive agents liberated from specific plasma proteins.

Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg

(structure of Bradykinin (9 a.a)

The cyclic peptides gramicidin S and tyrocidin are antibiotics and contain D-phenylalanine in their structures :



[Gramicidin S (10 a.a)] [Tyrosine (10 a.a)]

Proteins \

There are different systems of classification of proteins based on structure , solubility , function, chemical composition or other properties but non of them are universally satisfactory.

Proteins are classified into two groups, simple and conjugated proteins :-

(A)Simple protein \

Simple proteins are defined as those proteins that yield only amino acids when they are hydrolyzed.E.X \

1.Albumine :

They are soluble in water and diluted solvents , coagulated by heat. They may be precipitated from saturated ammonium sulfate solution.They are por ducts of both plants and animals , e.x\ egg albumin, serum albumin of blood and leucosin of wheat.

2.Globulins :

They are insoluble in water , but they are soluble in salts solutions of alkalis and acids, and the are heat coagulable. Also, they precipitated from saturated sodium chloride solution. The globulins constituent an important and widely distributed group of animal and plasma proteins ,e.x: ovoglobulin of egg yolk , blood serum globulin, myosin of muscles and seeds.

3.Glutelins :

The glutelins are soluble in very diluted acid and alkalis, but they are insoluble in neutral solutions. They are plant proteins(seeds) and they characterized a large content of arginine , proline and glutamic acid ,e.x: glutelin of wheat and rice.

4.Protamines:

They are strongly basic proteins and yield chiefly basic amino acid after hydrolysis, especially Arginine. They occur usually occur in tissues in salt combination with acids, Particularly with nucleic acids as nucleoproteins ,e.x: nucleoprotein of fish sperum.

5.Histones:

The histones are soluble in water and insoluble in diluted ammonia. They yield a large proportion of amino acid after hydrolysis. The histones being basic, usually occur in tissues in combinations with acidic substances such as nucleic acids e.x: globin of hemoglobin.

6.Collagens :

Collagens are essential proteins in the connective tissues and are insoluble in water, but they converted into gelatins that easily digestible ofter boiled in water or by affectivity of diluted acids or alkaline. They have a large proportion of hydroxy-proline and hydroxy-Lysine.e.x: skin ,hair , nails, and all connective tissues.

[B] Conjugated proteins \

They are composed of simple protein combined with some protein substance. The non-protein group is referred to as the prosthetic group.

1.Phosphoproteins :

Phosphoric acid is the prosthetic group of the phosphoproteins where the acid is usually attached to the serine or threonine by ester linkage. Casein of milk and vitellin of egg yolk are the the best-known phosphoproteins.

2.Mucoproein(Glycoprotein) :

They are composed of simple proteins combine with mucopolysaccharides. Their mucopolysaccharides are composed of hexosamine and hexo-sugars.

Water-soluble mucoproteins have been obtained from serum , egg white and human urine.

These proteins are not easily denaturation by heat or precipitated by agents such as picric and trichloro-acetic acids.The mucoproteins are important constituents of connective tissues and blood-group substances.

3.Chromoproteins :

The Chromoproteins are composed of simple proteins bonded with a colored prosthetic group. Many proteins of important biological functions belong to this group. Hemoglobins ,respiratory proteins in which the prosthetic group is porphyrin and iron. Cytochromes, cellular oxidation reduction proteins in which the prosthetic group is riboflavin.

4.Nucleoproteins :

They are composed of simple basic proteins (Protamine or histone) in salt combinations with nucleic acids as the prosthetic groups. They are the proteins of cell nucleic and are the chief constituents of chromatin (chromosome).

5.Lipoproteins:

They are formed by combination of protein with a lipid such as Lecithin , cephalin , fatty acid ,….etc. They are widely distributed in animal and plant material where they occur in milk ,blood , cell nuclei, egg yolk and chloroplasts of plant and also ,they are found bacterial antigens and viruses.

- **Some functions that proteins serve and examples of specific functional proteins are the following :-**

1. **Enzymatic proteins \ pepsin, Trypsin, Ribonuclase.**
2. **Structural proteins \ Collagen, Elastin, Keratins.**
3. **Contractile proteins \ Myosin, actin.**
4. **Transport proteins \ Hemoglobin, hemocyanin, myoglobin, serum protein.**
5. **Genetic protein \ nucleoprotein, histones.**
6. **Hormonal proteins \ Insulin, glucagon.**
7. **Storage proteins \ Casein , ovalbumin.**
8. **Immune proteins \ Gamma-globulins.**
9. **Venoms proteins \ Shake venom.**

Characterization and purification proteins \

Many techniques for separation and purification of proteins are used which involve the following ;-

1.Differential centrifugation \

The principle of this technique is based on the difference in molecular weight , where the centrifugation process divides the samples into two fractions : (a) sediment and supernatant fraction that can be easily separated ,e.x: separation of cell contents ,enzymes , and sperms.

2.Isoelectric precipitation \

Proteins tend to show a minimum solubility of their isoelectric pH, in which the number of positive charges would be equal to the negative charges.

3.Salt precipitation \

Protein also show a variation in solubility that depends on the concentration of salts in the solution. These frequently complex effects may involve specific interactions between charged side chain and solution ions.

The increasing in protein solubility after salts addition such as sodium chloride NaCl is often referred to as "salting in", Whereas the decreasing in protein solubility after sdlts addition such as $(NH_4)_2SO4$ or Na_2SO_4 is called as " salting out". Usually ,this process is achieved at saturated solution of ammonium sulfate salts, then the differential centrifugation is used to separate the precipitate proteins from other non-precipitated ones.

4.Organic solvents precipitation :

The precipitation of proteins can be achieved by using organic solvents such as acetone, methanol and ethanol, but using such these solvents leads to denaturation of proteins usually , therefore , this process must be achieved at low temperature.

5. Chromatographic techniques:

a)Ion-exchange chromatography.

b) Molecular sieve (Gel) filtration.

C) Electrophoresis.

Structures of proteins \

The twenty amino acids commonly found in proteins are joined together by peptide bonds. The linear sequence of the linked amino acids contain the necessary information to generate a protein molecule with unique three-dimensional shape. The complexity of protein structure is best analyzed by levels, namely ,primary ,secondary , tertiary and quaternary:-

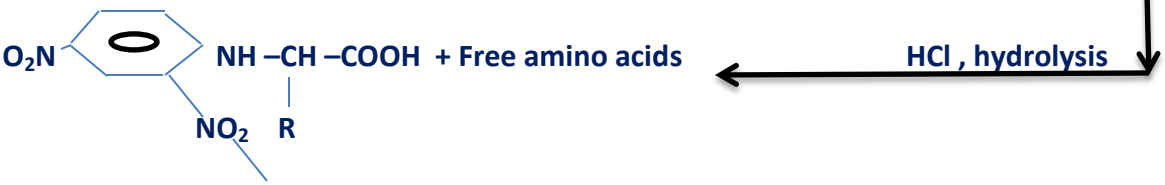
1. Primary structure of proteins \

The sequence of amino acid in a protein is called the primary structure. The first step in the determining the primary structure of polypeptide or protein is to identify and quantitate its content amino acids. A purified sample of the polypeptide to be analyzed is first hydrolyzed by strong acid at 110 C° for 24 hrs. The free amino acids can be separated by ion-exchange chromatography ,whereas the amount of each amino acid is determined by spectrophotometric methods.

Another step is the determination of N-terminal and C-terminal ends of polypeptide. The N-terminal can be determined by many ways such as :-

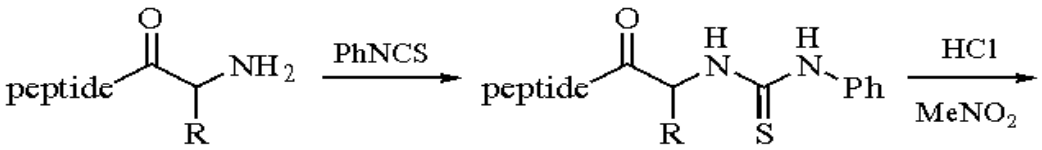
(a)Sanger reaction :

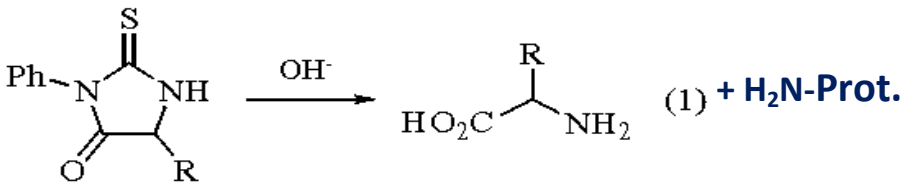
FDNB Protein DNB-protein



DNB –amino acid (yellow derivatives)

(b)Edman reaction : involve prepare phenyl-isothiocyanate peptide\





By using this reaction, we can determine the sequence of peptide amino acids reach up to 32 amino acid.

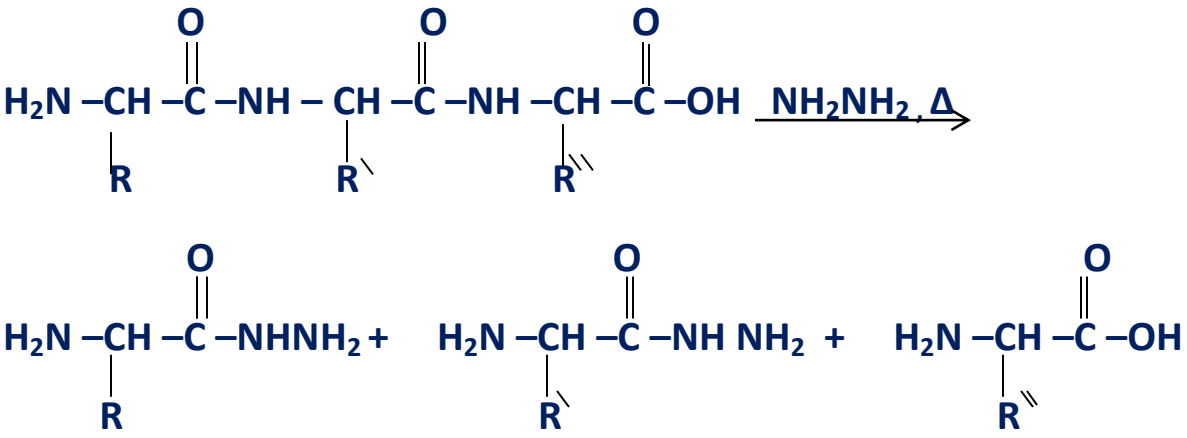
(c)Enzymes \

Enzymes also are used to determine the N-terminal end such as exopeptidase (leucine aminopeptidase). This enzyme needs free amino group to work , therefore, its liberates all amino acids in the N-terminal of peptide or protein except the proline.

- **C-terminal amino acid can be determined also by many ways such as **

1.Reaction with hydralazine(Hydralazinolysis) :-

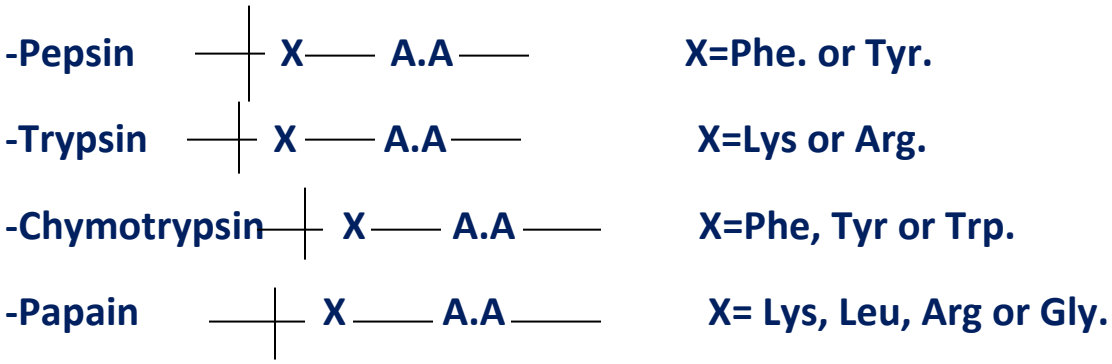
This method is very important to identify the C-terminal amino aci, where the polypeptide or protein is boiled with dry hydralazine at 100 C° for 6 hrs, Which lead to hydrolyze the peptide bond to produce the hydrazide derivatives of amino acids except the amino acid at C-terminal end.



2.Enzymes \

Enzymes also are used to determine the N-terminal end such as exopeptidase which called carboxypeptidase where the carboxypeptidase A which its liberates all amino acids in the N-terminal of peptide or protein except the "Arg , Lys ,and Pro.", whereas carboxypeptidase B which liberates "Arg , Lys", while carboxypeptide C will liberates "pro".

- **The next step in the determination of amino acids sequences are defragmentation or cutting the polypeptide chain or protein into small chains or fragments. This process is achieved by action of a kind of endopeptidase such as pepsin, trypsin, chymotrypsin and papain.**



---

2. Secondary structure of proteins \

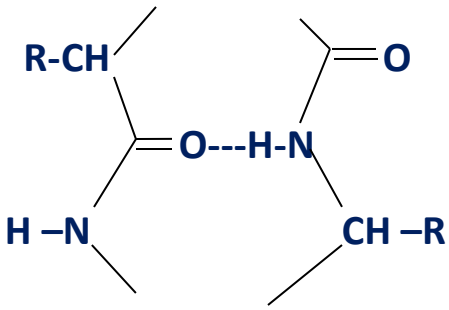
The polypeptide backbone does not assume a random three-dimensional structure, but instead generally forms regular arrangements of amino acids that are located near to each other in the linear sequence. These arrangement are termed the secondary structure of the polypeptide.

The α-helix and β-sheet are examples of secondary structures frequently encountered in proteins.

(a) α-Helix \

There are several different polypeptide helixes found in nature, but the α-helix is most common ,it is spiral structure, consisting of a tightly packed , coiled polypeptide backbone core, with side chains of the components amino acids extending outward from the central axis to avoid interfering sterically with each other many proteins contain α-helixes ,for examples ,the keratins are a family of closely related fibrous proteins which are a major component of tissues such as hair and skin.

1) **α-helix structure is stabilized by hydrogen bonding between the peptide bond carbonyl oxygen's amide hydrogen that are part of the poly peptide backbone. The hydrogen bonds extend up the spiral from the carbonyl oxygen of one peptide bond to –NH- group of a peptide linkage four residues a head in the polypeptide. Hydrogen bonds are weak, but they collecting serve to stabilize the helix.**



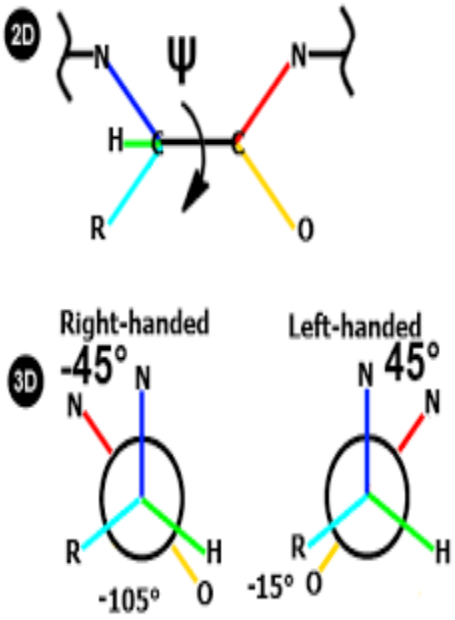
Hydrogen bonding

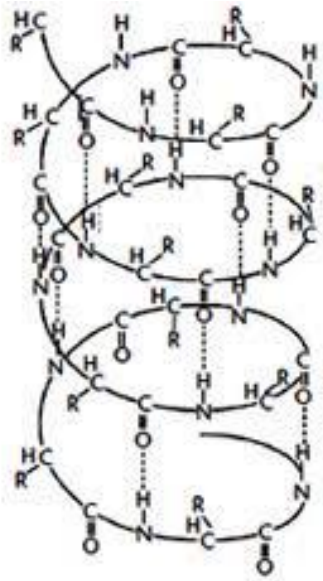
2.Each turn of α-helix contain 3.6 amino acids, thus ,amino acids residues spaced three or four a part in the primary sequence are especially close together when folded in the α-helix.

3.Proline destroy α-helix because it is imino acid group is not geometrically compatible with the right-handed spiral of the α-helix. Large numbers of charged amino acids (Glu, Asp, Lys, His, Arg) will destroy the α-helix by forming ionic bonds or by electrostatically repelling each other. Also, amino acids with bulky side chains , such as tryptophan or amino acids such as Valine or isoleucine that branched at the β-carbon can interfere with formation of the α-helix if they present in large number.

The α-helix may be right handed or left handed , but when the polypeptide have L-amino acids residues , the right-handed one is more stable.

The α-helix structures

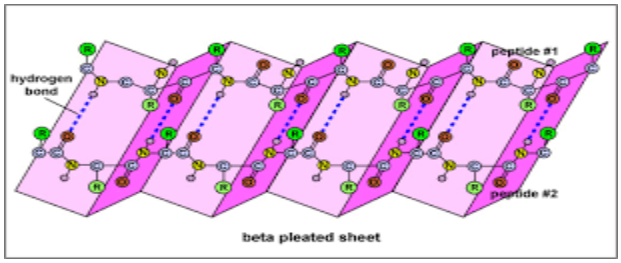


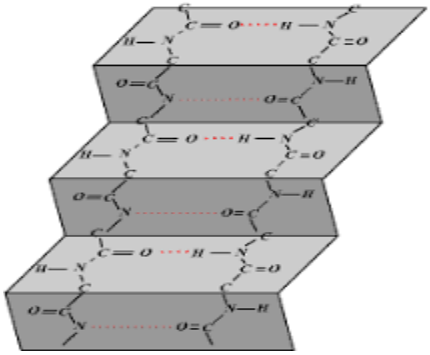


Right-handed α helix

(b) β-sheet \

The β-sheet is another from of secondary structure in which all of the peptide bond components are involved in hydrogen bonding. The surfaces of β-sheet appear "pleated" and these structures are,therefore. Often called "β-pleated sheet".

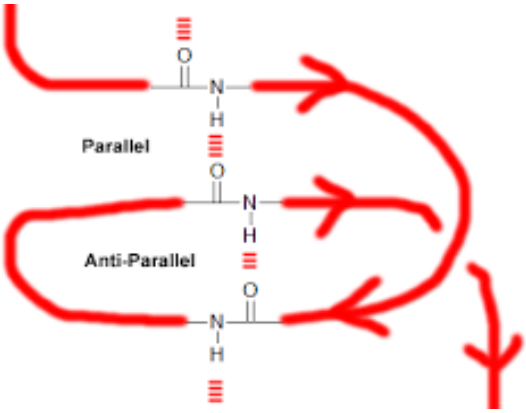


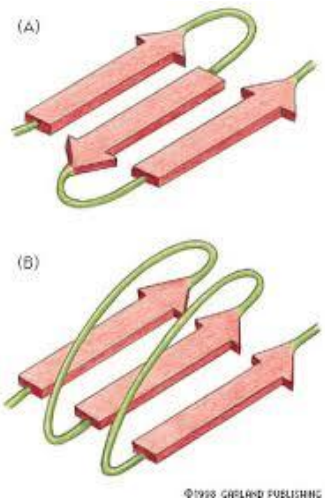


1.Unlike the α-helix , β-sheet are composed of two or more peptide chains (β-sheet) or segments of polypeptide chains, which are almost fully extended. The hydrogen bonds are perpendicular to the polypeptide backbone.

2. β-sheet can be formed from two or more separated polypeptide chains or segments of polypeptide chains that are arranged either antiparallel to each other or parallel, When the hydrogen bonds are formed between separated polypeptide chain. They termed (interaction bonds).

But when these bonds are formed in single polypeptide chain which folding back on itself, here are termed (intrachain bonds).





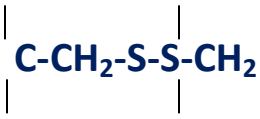
The bonds of give it a compact and globular shape hey usually found on the surface of protein molecules and often include charged residues.

3.Tertiary structure of proteins \

It interferes to the folding of domains (basic units of structure and function) ,e.x: the structure of globular protein in aqueous solution is compact, where the hydrophobic side chains would be located in the interior of molecule , whereas the hydrophilic side chains would be located in the surface of molecule.

These movements or charges in the locations of amino acids side chains lead to form the tertiary structures of proteins, also the happens lead to form new interactions inside protein molecule which collectively cooperate in the stabilization of this structure.

-Disulfide bonds \

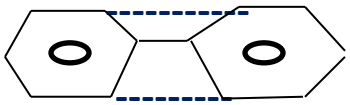


-Ionic interaction \

Asp Lys

-Hydrogen bonding. \

-Hydrophobic (Vander Waals) interaction\



or $-CH_3$-------CH_3