

Wednesday
20/9/2023

Credit - I

Protein Chemistry.

Page No.

Date

1) * Structure of Amino Acids :-

- proteins are polymer of amino acids.
- 23 amino acids are commonly found in proteins.

First amino acid discovered was 'Asparagine' found in Asparagus.

Glutamate was found in Wheat glutane gluten.

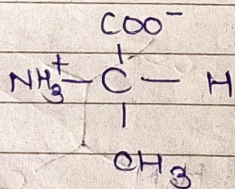
Tyrosine - (In Greek word 'Tyros' means chese hence the name is Tyrosine).

- Glycine (Glykas means sweet)

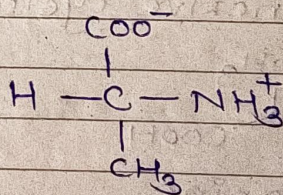
- Amino acids have carboxyl group & amino group bounded to same carbon atom (α carbon)

- All amino acids except 'Glycine' the α -carbon is bonded to four diff. groups [carboxyl $\rightarrow$ COO^- & amino group $\rightarrow$ NH_3^+ & 'R' group & 'H' atom]

eg.



L-Alanine



D-Alanine

- Stereoisomerism in α -amino acids are non superimposable mirror images of each other i.e. Enantiomers

(loss of water)

- Amino acid residues in protein molecules are exclusively L-stereoisomers when one amino acid joined to another.

* Classification of Amino acid :-

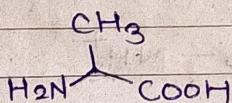
- Amino acids can be classified by 'R' group.
- polarity of 'R' group varies widely from non-polar & Hydrophobic to highly polar & Hydrophilic.

✓ 1) Non-polar, Aliphatic 'R' Group -
Amino acid in this group are non-polar & Hydrophobic (water insoluble)
eg.

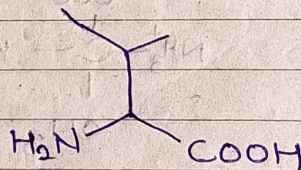
- Alanine, valine, leucine & isoleucine cluster together within the proteins.

Exception →

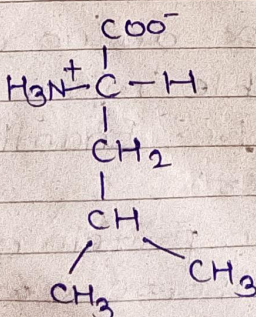
- Glycine is easily grouped with non-polar amino acid.
- Methionine having 2 sulphur containing amino acid.
- proline is aliphatic side chain with distinctive cyclic structure.



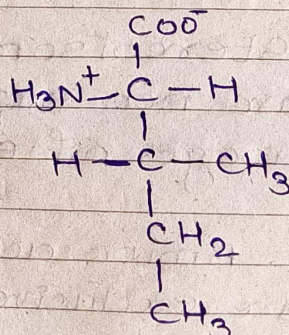
Alanine



valine



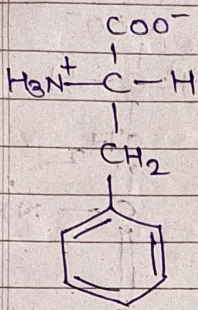
Leucine



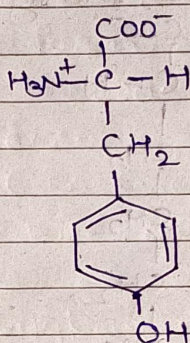
Isoleucine

2) Aromatic 'R' groups —

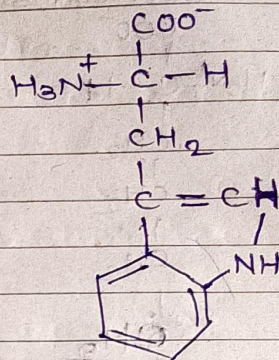
Phenyl Alanine, Tyrosine and Tryptophan with aromatic side chains & these amino acids are non-polar.



Phenylalanine



Tyrosine

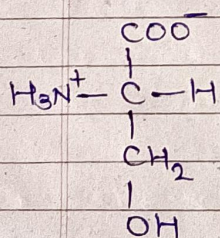


Tryptophan

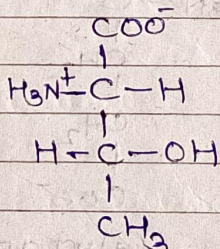
3) Polar and Uncharged 'R' Groups —

These amino acids are more soluble in water.

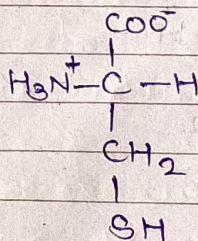
Serine, Threonine, Cysteine, Asparagine and Glutamine



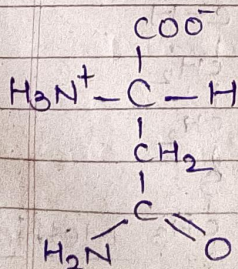
Serine



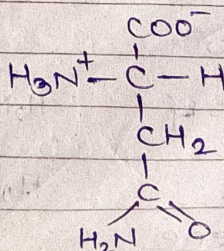
Threonine



Cysteine



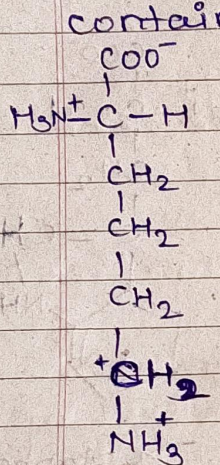
Asparagine



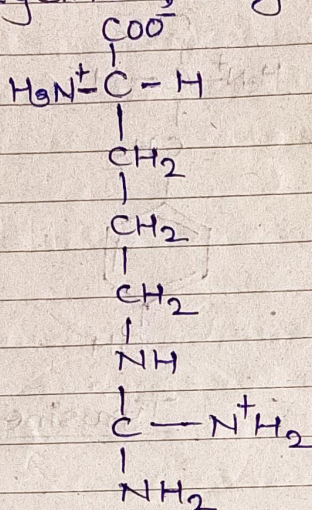
Glutamine

4) Positively charged (Basic) 'R' Groups —

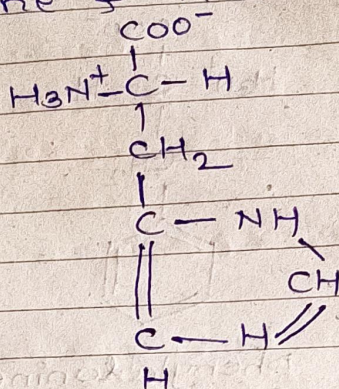
- Hydrophilic 'R' groups are either positively charged or Negatively charged
- 'R' group with positive charge at pH 7 contains Lysine, Arginine & Histidine.



Lysine



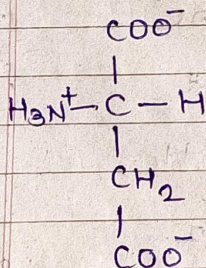
Arginine



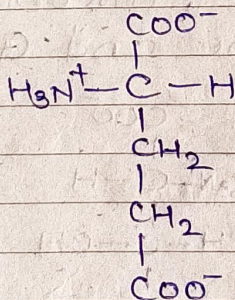
Histidine

5) Negatively charged (Acidic) 'R' Groups —

- 'R' group with negative charge at pH 7 are Aspartate & Glutamate.

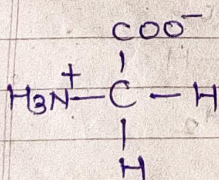


Aspartate

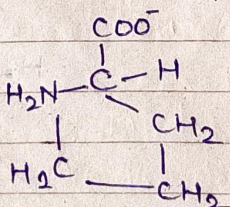


Glutamate

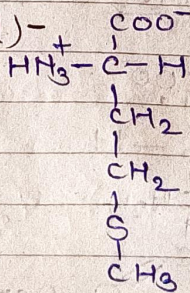
* Exceptional (not grouped) —



Glycine



Proline



Methionine

21/9/23
Thursday

-ve charges forms acidic sol?
+ve charge forms basic sol?

Page No.

Date

* Amino acids as buffers :—

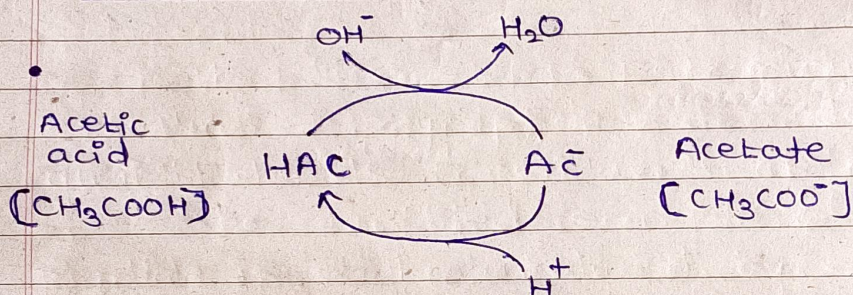
- Ionizable R groups of some amino acids along with amino (basic) and carboxyl (Acidic) group functions as weak acid & bases. (buffer)
- Amino acids lacks in ionisable R groups is dissolved in water at neutral pH. & exist in solution as dipolar or Zwitter ion which can act as either an acid or base.

✓ 2) * Henderson - Hasselbalch equation relates pH, pKa and buffer concentration :-

- The shape of Titration curve of any weak acid is described by 'Henderson - Hasselbalch equation'

- It is important for understanding buffer action & acid base balance in blood & tissue of vertebrates.

• $K_w = [H^+][OH^-]$ ($K_w \rightarrow$ water constant)



• Acetic acid - acetate pairs as buffer system.

$K_a = \frac{[H^+][A^-]}{[HAc]}$ $\left[\frac{\text{conc. of Reactant}}{\text{conc. of product}} \right]$
 (acetic acid constant)

- $\therefore HAC \rightarrow$ proton donor of acetic acid.
- Conjugate acid base pair (acetic acid - acetate) ion tends to resist a change in pH when small amount of acid or base is added.

• Buffering action is simply the consequence of two reversible reactions taking place simultaneously & reaching their points of equilibrium as their equilibrium constant K_w and K_a

• Now,

$K_a = \frac{[H^+][A^-]}{[HA]}$ $\Rightarrow$ $[H^+] = \frac{K_a[HA]}{[A^-]}$
 (acetate constant)

- Taking negative logarithm on both sides
(Due to pH definition i.e. pH is a negative log of H^+ ions)

$$\therefore -\log[H^+] = -\log K_a - \log \frac{[HA]}{[A^-]} \dots \textcircled{2}$$

- Now,
substitute pH for $-\log[H^+]$ and
pKa for $-\log K_a$ in eqⁿ (2)

$$\therefore \boxed{pH = pKa - \log \frac{[HA]}{[A^-]}}$$

- Now,
To obtain Henderson-Hasselbalch eqⁿ.
in word $-\log \frac{[HA]}{[A^-]}$ which involves changing
its sign.

$$\therefore \boxed{pH = pKa + \log \frac{[A^-]}{[HA]}}$$

- This is Henderson-Hasselbalch equation.
- This equation fits the titration curve of all weak acids & enables us to deduce some important quantitative relations.
- At midpoint of Titration $pKa = pH$
i.e. when $[A^-] = [HA]$ $\therefore pKa = pH$.

$$\therefore pH = pKa + \log \frac{[A^-]}{[HA]}$$

$$\therefore [A^-] = [HA]$$

$$\therefore pH = pKa + \log \frac{[HA]}{[HA]}$$

$$\therefore pH = pKa + \log 1$$

$$\boxed{pH = pKa} \dots (\log 1 \text{ is } 0)$$

amino acids → N terminal → $(-NH_2)$ (amino) group attached
→ C terminal → carboxyl group attached $(-COOH)$

* Uses of Henderson - Hasselbalch equation —

- Henderson - Hasselbalch equation also allows us :

- 1) Calculate pK_a , given pH & molar ratio of proton donor & Acceptor.
- 2) Calculate

6/10/23
Friday

Chapter 3
Lehninger

* Determination of primary structure of Polypeptide —

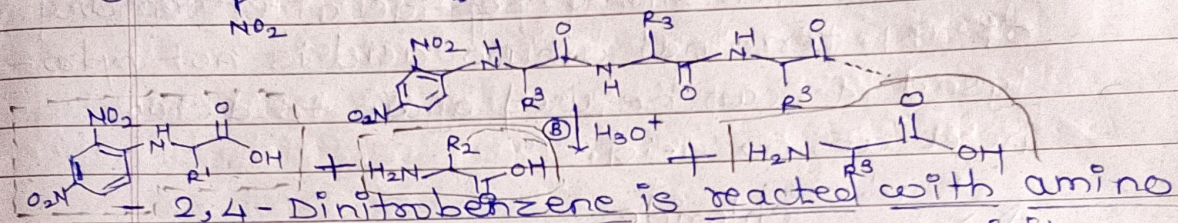
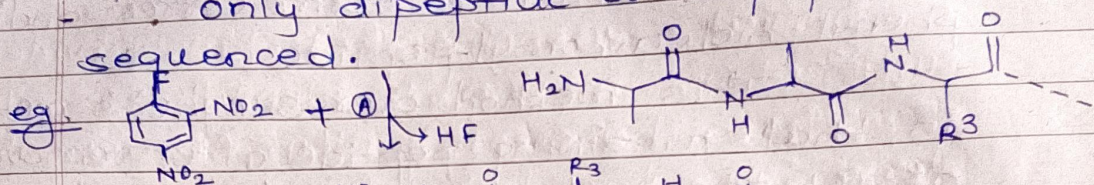
- The first step in Determining primary structure of polypeptide is to Identify & quantitate its constituents amino acids.
- The purified sample of polypeptide is to be analysed by strong acid (6M HCl) at $110^\circ C$ for 24 hrs.
- These cleaves peptide bonds and release amino acids which can be separated by cation exchange chromatography.
- The amount of amino acid is determined by spectrophotometrically by measuring amount of light absorbed by Ninhydrin derivative.
- Identification of N-terminal & C-terminal amino acids in polypeptide chain is called as End group analysis.

* Identification of N-terminal End : —

1) Sangher's Method —

- It is the first technique to determine sequence of proteins.
- Sangher's reagent → 2,4-Dinitrofluorobenzene (2,4-dinitrofluorobenzene)

- first protein to be sequenced is insulin by Fredrick Sanger
- only dipeptide or tripeptide can be sequenced.



2,4-Dinitrobenzene is reacted with amino group of peptide to form 2,4-Dinitro-phenyl derivative of N-terminal amino acid which is yellow in colour.

- Then is subjected to acid hydrolysis which cleaves all peptide bonds except the bond between 2,4-Dinitrofluorobenzene (2,4-DNFB) & NH_2 group which is resistant to acid hydrolysis.

- It is separated by chromatography.

✓ Edman's Degradation Method :- (Pehr Edman)

- Edman's reagent $\rightarrow$ Phenyl isothiocyanate

- It is used to identify specific amino acids at each position in peptide chain beginning at amino terminal end.

- Edman's technique can sequence many residues of single polypeptide sample.

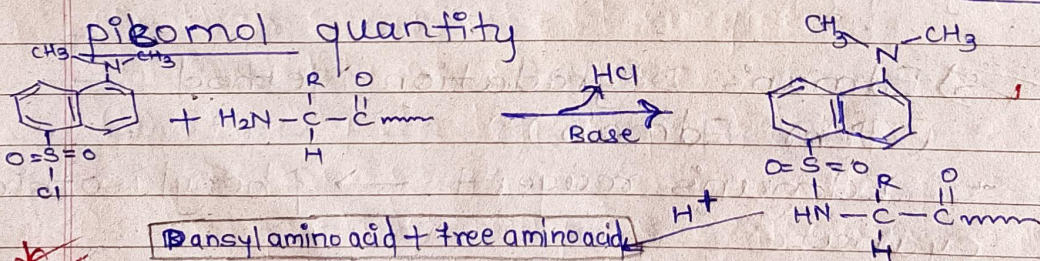
eg.

- Phenylisothiocyanate (PITC) react with amino group of polypeptide under mild alkaline condition to form corresponding phenylthiocarbonyl-peptide
- This method over sangher's method is that remaining peptide after removal of N-terminal amino acid is not hydrolysed & can be used again to detect next amino acid.

8) Dansyl chloride fluoride Method : —

- Dansyl Reagent $\rightarrow$ Dansyl chloride (5'-dimethyl-1-Naphthalene-sulphonyl chloride)
- NH_2 terminal is reacted with dansyl chloride which is fluorescent compound to form fluorescent dansyl amino acid.
- It is separated by chromatography
- It is sensitive method & can detect

picomol quantity



Dansyl amino acid + free amino acid $\xrightarrow{\text{H}^+}$

1) Akabori Method : —

- Treatment with hydrazine.
- In this method Identification of C-terminal amino acids involves heating linear peptide in presence of anhydrous hydrazine in a sealed tubes for several hours.

- The amino group of each peptide bond react with hydrazine to form corresponding amino acid hydrazide.
- It is identify by chromatography.
- The application of microwave irradiation to this reaction was investigated. It is was observed that microwave irradiation reduce the time required for completion of Akabori reaction. from hours to minutes.
- This approach not only about C-terminous but also N-terminous.
[approach $\rightarrow$ Microwave Irradiation (only)]

2) ~~By~~ Carboxy peptidase Method :-

- Carboxy peptidase cleaves a polypeptide ~~from~~ c-terminal removing 1-amino acid at a time by cleaving peptide bond.
- Removal of one amino acid leaves behind a new polypeptide which is the target of carboxypeptidase.
- Carboxypeptidases that have stronger preference for those amino acids containing aromatic or branched chain of hydrocarbon are called carboxy-peptidase - A
- It cleaves positively charged amino acids (arginine, lysine) are called carboxypeptidase - B.
- plant contains carboxypeptidase - C that liberates the amino acid proline as well as being able to release many of other protein amino acids.

* Methods of sequencing of peptides : —

- protein sequencing is used to identify the protein or characterize its post transcriptional modifications.
- partial sequencing of proteins provides sufficient information to identify.
- Data bases of protein sequences derive from translation of genes.
- Mass spectrometry are most widely used for protein sequencing & Identification.

1) The protein of interest is isolated by SDS Page or chromatography.

2) The isolated protein of interest is chemically modified - to stabilised amino acid cysteine residue.

3) The protein is digested with specific protease to generate peptides its advantages include :-

(i) frequency of Lysine & Arginine residue in protein.

(ii) High specificity of enzymes.

(iii) Stability of enzyme.

(iv) Stability of tryptic peptide for mass spectrometry.

4) The peptides may be De-isolated to remove ionisable contaminants & subjected to MAL-DE-ToF mass, spectrometry.

- Direct measurement of masses of peptides may provide sufficient information to identify protein (peptide mass finger printing) but further fragmentation inside the mass spectrometry is used to gain information about peptide sequencing.

5) Peptides may be separated by reverse phase HPLC & introduce to mass spectrometer.

* Similarities in sequences & functions are members of family of proteins that share common ~~chara~~ ancestor. This members of protein families are called homologous protein.

- The proteins in family are present in same species they are referred to as paralogous proteins.

- Homologous from different species are called as orthologous protein.

- Software packages usually generates a report showing the identity of each identified protein, its ~~mat~~ matching score & provide a measure of relative strength of matching where multiple proteins are identify.

• Denovo Sequencing —

- The pattern of fragmentation of peptide allow for direct determination of its sequence by De-novo sequencing.

- This sequence may be used to match databases of protein sequences or to investigate post-translational or chemical modifications.

- It may provide additional evidence for protein identifications performed as by mass spectrometry or chromatography or SDS page.

* Structural classification of proteins :-

• protein -

- protein structure are defined as polymer of amino acid joint by peptide bonds.

✓ Classification of protein -

1) fibrous protein (Linear) -

- When polypeptide chain run parallel & held together by hydrogen or disulfide bond & form fibre like structure are called fibrous protein.

- They are generally insoluble in water
eg. Keratin.

2) Globular protein -

- When chains of proteins coils around to give spherical structure / shape are called as globular protein.

- They are soluble in water.
eg. Insulin, albumin.

✓ Primary structure of protein -

- The exact sequence of protein is very important to determine the final fold & therefore the function of protein.

- The protein primary structure and an amino acid sequence in its polypeptide chain.

- Covalent, peptide bond which connect a acid together to maintain the primary structure of protein.

- All documented genetic disorder such as cystic fibrosis, sickle cell anemia are caused by mutation resulting in alteration in primary structure of proteins.

✓ Secondary structure of protein :—

- 2° structure of protein refers to local folded structures that form within polypeptide due to interaction between atoms of backbone.
- This polypeptide chains usually fold due to interaction between amine & carboxyl group of peptide link.
- They are found to exist in two diff. types of structures : α -helix & β -plated sheets.
- This structure arises due to regular folding of backbone to hydrogen bonding between 'COO⁻' group and 'NH' group of peptide bond.

✓ (a) α -Helix —

- α -helix is one of the common way in which a polypeptide chain form all possible hydrogen bond by twisting into right handed screw with 'NH' group of a acid residue hydrogen bonded to COO⁻ of adjacent turn of helix.

✓ (b) β -plated sheet —

- In this arrangement the polypeptide chains are stretched out beside one another & then bonded by inter-molecular hydrogen bond.
- In this structure all peptide chains are stretched out to nearly maximum extensives & then lead side by side which held together which.
- The structure resembles to plated force of drapery therefore is known as β -plated sheet

✓ • Tertiary structure of protein : —

- This structure arises from further folding of 2° structure.
- hydrogen bond, electrostatic forces, disulphide linkages, coonder-wal-force stabilizes this equation.
- It gives rise to two measure molecular shape called fibrous & globular.
- The main forces which stabilized the secondary & tertiary structure of protein are hydrogen bond disulphide & von-dex-wall force, electrostatic forces of attraction.

✓ • Quaternary Structure of protein : —

- The special arrangement of various tertiary structure gives rise to quaternary structure.
- The exact amino acid sequence of each proteins derives to fold it unit & biologically active three-dimensional fold also known as tertiary structure.
- protein consist of different combinations of 2° element some of each are simple or complex.

✓ The part of protein chain which has there own three fold & can be important attributed to some function are called domains

- Many proteins most of which are enzyme contain organic or elemental components needed for there activity & stability.

18/10/23
Wednesday

Page No.

Date

Polar - more soluble

Non polar - less Soluble

* Non-Covalent Interaction -

- Non covalent interactions are weak interaction between atoms or molecules where no chemical reaction takes place, because no formation of chemical bonds is induced,
- non-covalent interaction are often called as Non-bonded interaction.

- Three types -

(1) * Vander-Wall Interaction :-

- They are short range interaction & occurs when two atoms or molecules come close to each other hence it is also called as short range interaction.
- The interaction which become relevant at distances comparable with size of interacting atom.
- They are comparatively weak and therefore more susceptible to more disturbance.
- This are described as combination of London-dispersion forces between instantaneously induced dipole & dehye... forces between permanent dipoles and induced dipoles & Keesom force between permanent molecule dipole whose rotational orientations are dynamically average over time.
- The strength of vander-wall bonds increases when with higher polarizability of participating atoms.

(2) * Electrostatic Interaction :-

- In contrast to the other two types of non-covalent interaction the electrostatic interactions are long range.

- This are relevant beyond the limits of close neighbours.
- This makes the description somewhat complicated in which ions involves the interaction between opposite electric charges such as cation & anion.

(i) * Electrostatic Attraction :-

- When opposite charges interaction leads to formation of bond.

(ii) * Electrostatic Repulsion :-

- When same charges interact each other. they repulse & bond formation do not occur.

- Polymeric micelle formation of ionic interaction have great advantages of encapsulating the active agent, protein, peptides & nucleic acid delivery.

(3) * Hydrogen Bonds :-

- Hydrogen bonds are interaction which are at boundary between chemical bonds & non-covalent interactions.
- They takes place between parts of atoms only if one of them is proton donor & other one is proton acceptor.

21/10/2023
Saturday

Page No.

Date

* Confirmations of protein :-

- protein conformation may be defined as arrangement in space of its constituent atoms, that determine the overall shape of molecule
- A series of computational examples of protein conformational changes under different pH environment, temperature & ligand binding.
- Some useful methods such as constant pH molecular dynamics, molecular docking and molecular mechanics, poisson-boltzmann surface area / generalised bond & surface area calculations.
- In comparison with experimental results the methods mentioned above are reasonable to detect residue, residue & DNA, and residue and ligand.
- additionally some crucial interactions that cause protein conformational changes are discovered it give penetrating information to understand the pH, temperature and ligand induced conformational change mechanism.
- The conformation of protein arises from bonding arrangement within the structure.

* Techniques used for protein Determination :-

- If the protein contains only peptide bonds all proteins would be random in shape due to C α -C and C α -N bond

rotation, investigation on a large no. of proteins have indicated that proteins are highly organised structure with defined shapes.

1) NMR Although proton nuclear magnetic resonance (^1H -NMR) has been recently applied to the determination of protein conformation, the study of conformation of protein has been dominated by the techniques of x-ray crystallography which offers a view of relative position of atoms within a structure.

2) X-ray crystallography The x-ray diffraction studies of 1940 highlighted general features regarding the shape of the protein.

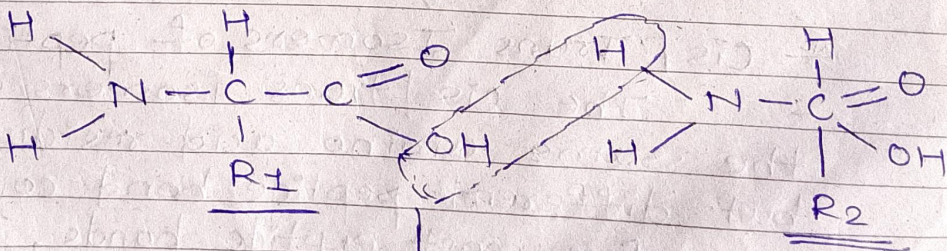
- The distance between major recurrent features is termed major periodicity & between minor recurrent features is the minor periodicity.
- X-ray crystallography creates the impression that each atom in the protein molecule is firmly fixed in position.
- This view of static structure is being replaced by an appreciation of dynamic nature of molecule in which the atoms of proteins are in state of constant motion.
- Therefore, the x-ray diffraction studies produce an average position around which each atom moves.
- From the knowledge of dimensions & characteristics the backbone bonds scale models were constructed in 1940.

to define geometric relationship of neighbouring amino acid residues.

31/10/2023

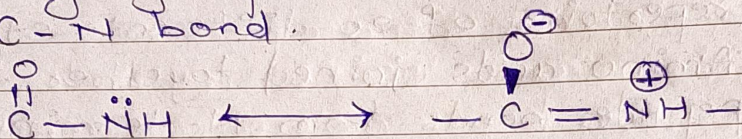
* Polypeptide Chain Geometry :-

- protein polypeptide chains are linear polymers that are assembled from repertoire of 20 different standard Amino acids joined together through peptide bonds from N-terminus & C-terminus.
- The backbone is built of α -C atoms & the side-chains are represented by pseudo C β atoms corresponding to 20 natural amino acids.
- Identity of each amino acid is determined by its side chain, known as an 'R' group.
- The 'R' group is important because it gives different properties to amino acids. Some amino acids have large R groups that create steric hindrance in polypeptide chain while others may have acidic or basic 'R' groups.

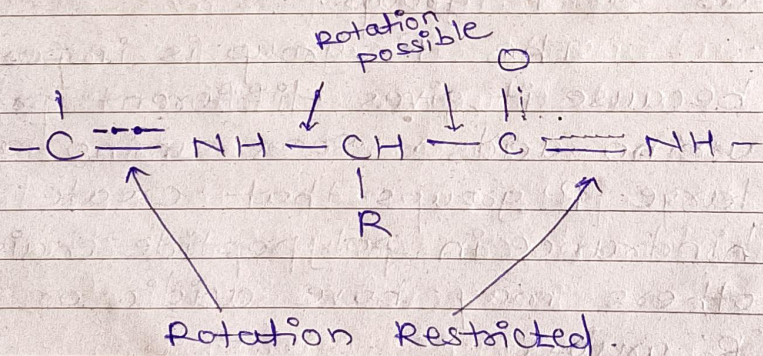


H₂O Dehydration
 Polypeptide formation.

- Resonance form of peptide Group :—
 - The General structure of peptide group is rigid & planar.
 - The stability of peptide bond is because of resonance of amide.
 - The resonance structure form due to the interaction between electrons of carbonyl groups double bond with those of C-N bond.



- because the bond between the carbonyl carbon & nitrogen has partial double bond character, rotation around the bond is restricted, thus the peptide unit is planar, rigid structure and rotation in peptide backbone is restricted to the bonds involving carbon.



- * Cis / Trans Isomers of peptide group :—
 - The cis / trans isomers are with the same amino acid residue content, but different peptide bond configuration.
 - for most peptide bonds, trans configuration is favoured about 1,000 times more than the cis configuration.

- In trans Configuration, the 2 alpha carbon atoms of connected amino acids are on opposite sides of peptide bond.
- In cis configuration they are on same side of peptide bond.
- In most cases, peptide bonds in proteins are trans.
- Trans Configuration is more stable because steric clash between side chains.

