INTRODUCTION

Biopolymers present in nature are Polysaccharides [Carbohydrates], Proteins, and Nucleic Acids

➢ ***Proteins are made up of hundreds of smaller units called amino acids that are attached to one another by peptide bonds, forming a long chain.***

- ***Amino acids are monomers of peptide and protein biopolymers***

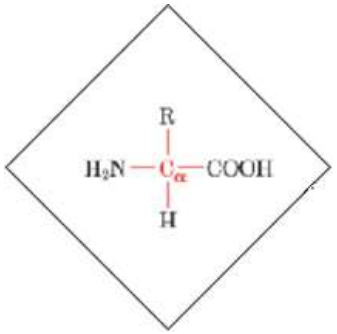
Amino Acids —*Peptide/ amide linkage*→ ***Peptides {Di-. Tri- , Tetra-, Oligo-}*** —*covalent/ H-bond/ van der Waal & 3D folding*→ ***Proteins***

- ***The sequence and quantity of amino acid arrangement determine the final protein structure***

- ***Side chain of an amino acid determines***
  - ***protein folding***
  - ***binding to specific ligand***
  - ***interaction with its environment***

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

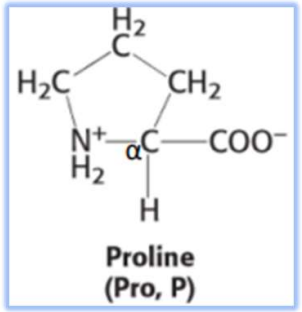
Amino Acids



***Amino acid**: is defined as a carboxylic acid having a protonated amino gp on the α-C atom. For all the common amino acids except glycine, the α-carbon is bonded to four different gps: a carboxyl group, an amino group, an R group, and a hydrogen atom. In case of glycine, the R group is replaced by another H-atom. They have a 1º amino group & a carboxylic acid substituent on the same carbon atom, with the exception of proline, (which has a 2º amino group).*

Proline classified as an imino acid, its α-amine is a 2º amine with its a nitrogen having two covalent bonds to carbon (to the α-carbon and side chain carbon), rather than 1º amine.

Incorporation of amino nitrogen into a five membered ring constrains rotational freedom around $-N_{\alpha}$ - C_{α}-bond in proline to specific rotational angle, reduces structural flexibility of polypeptide regions containing proline.

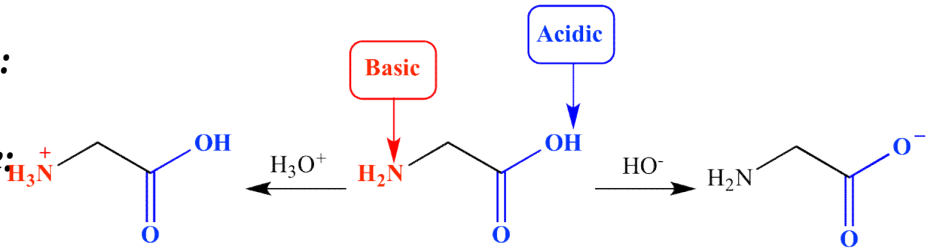


Proline
(Pro, P)

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Characteristic features of Amino Acids:

1. Amphoteric/ amphiprotic in nature:



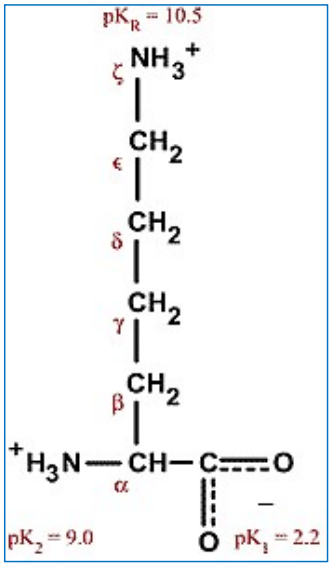
They have acidic and basic tendencies. The carboxyl group is able to lose a proton and the amine group is able to accept a proton. Amino acids are also ionic in character, and behave as ampholytes, meaning they move to their isoelectric points [pI] when placed in a pH gradient under an electric field.

2. Distinct R – gp determines the structure and thus unique biochemical functions of a protein:

As we shall study next, all naturally occurring sugars belong to the D series. It is interesting that nearly all known plant and animal proteins are composed entirely of L-amino acids. However, certain bacteria contain D-amino acids in their cell walls, and several antibiotics (e.g., actinomycin D and the gramicidins) contain varying amounts of D-leucine, D-phenylalanine, and D-valine.

The side chain or R group attached to the α-carbon of each amino acid has unique characteristics arising from the size, shape, solubility, charge and ionization properties of its R group

For all the common amino acids except glycine, the α-carbon is bonded to four different groups: an amino, a carboxyl, an R group, and a hydrogen atom. In case of glycine, the R group is replaced by another hydrogen atom. The additional carbons in an R group are commonly designated β , γ, δ, ε, and so forth, proceeding out from the α-carbon.

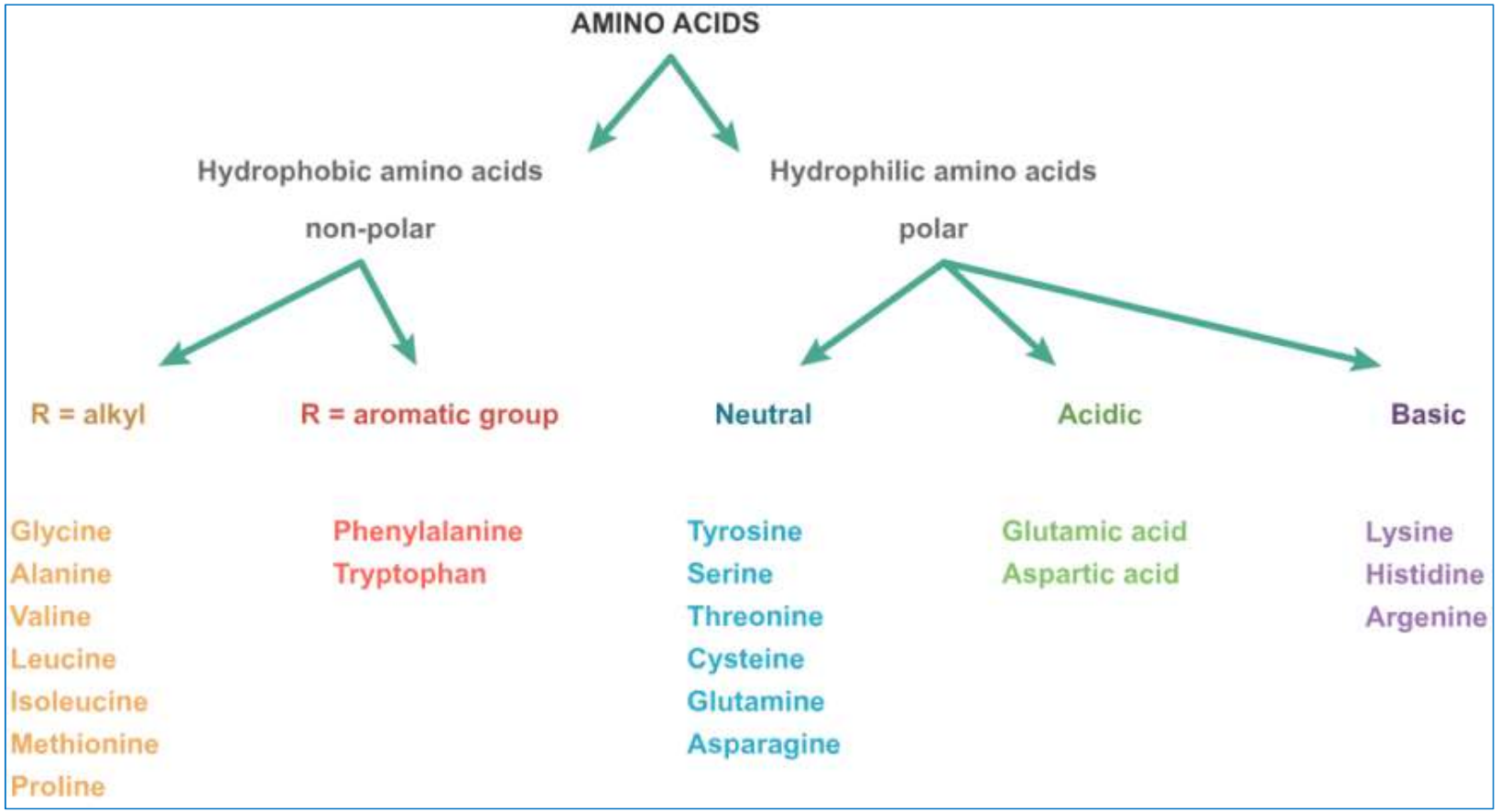


Solubility is decided by polar/ non-polar nature of R group [hydrophilic/ hydrophobic R]

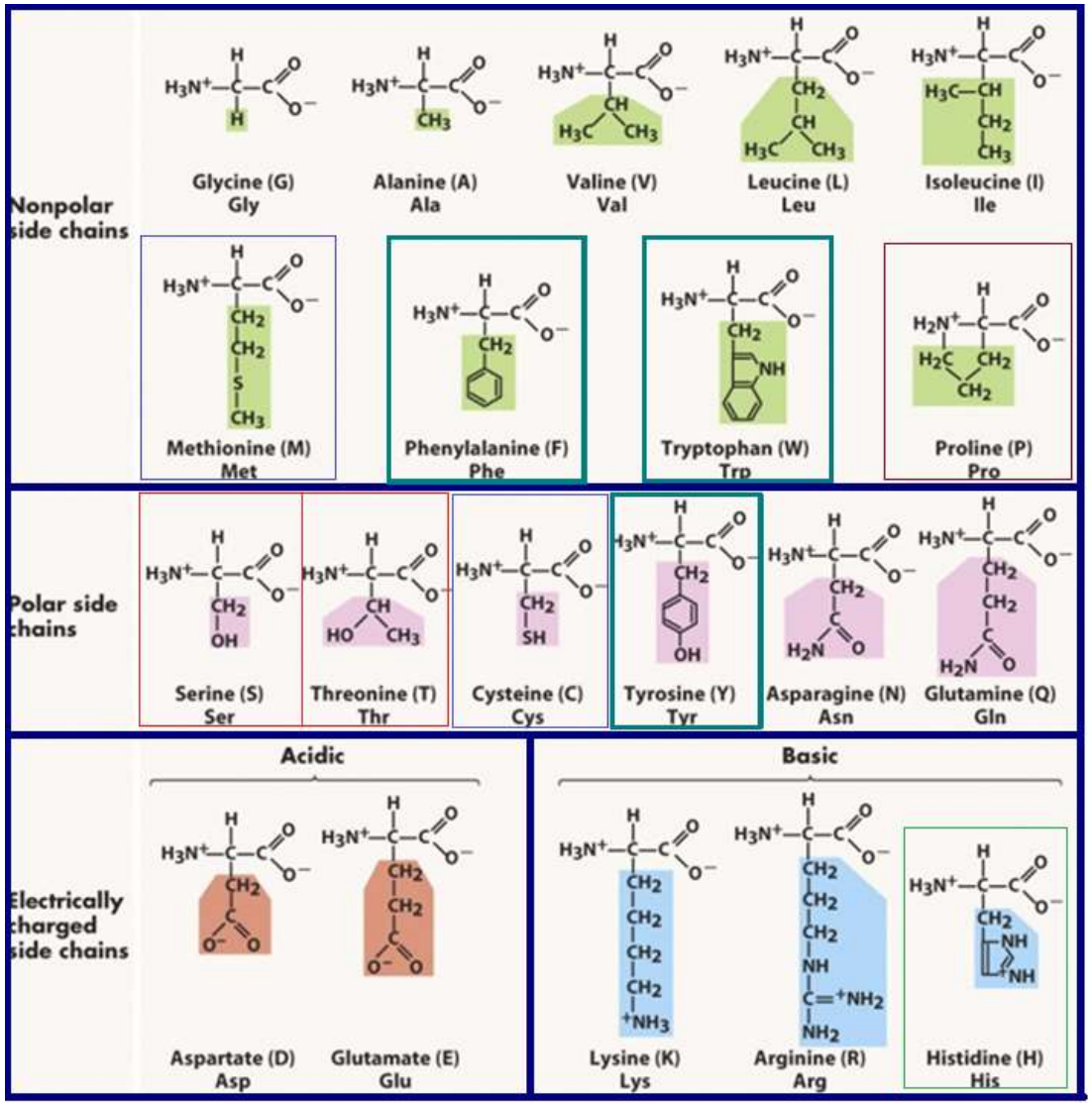
Classification according to the functional group on the side chain at neutral pH is

- *Nonpolar [hydrophobic]*
- *Polar but uncharged [hydrophilic]*
- *Polar negatively charged [hydrophilic]*
- *Polar positively charged [hydrophilic]*
- *Aromatic [hydrophilic/ hydrophobic]*

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati


AMINO ACIDS
Hydrophobic amino acids
non-polar
Hydrophilic amino acids
polar
R = alkyl
R = aromatic group
Neutral
Acidic
Basic
Glycine
Alanine
Valine
Leucine
Isoleucine
Methionine
Proline
Phenylalanine
Tryptophan
Tyrosine
Serine
Threonine
Cysteine
Glutamine
Asparagine
Glutamic acid
Aspartic acid
Lysine
Histidine
Argenine

Classification of Amino acids by their R group



Aromatic

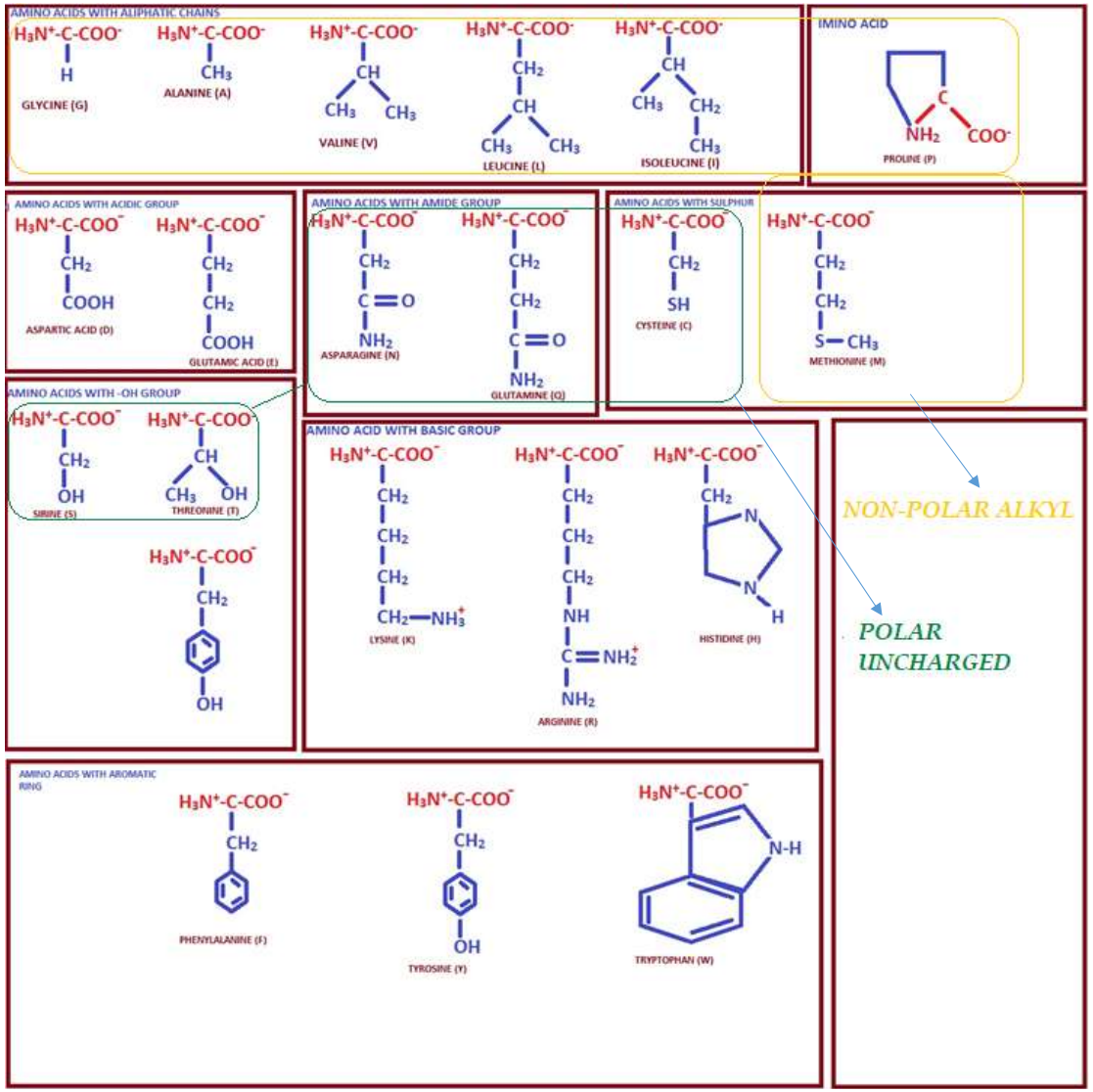
S- containing R:
Met, Cys

OH containing R:
Ser, Thr

Imino containing R:
Pro

Imidazole containing R:
His

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

The amino acids differ only in the substituent R attached to the α-carbon atom. This variation in R group gives proteins their structural and therefore functional diversity (as seen in enzymes/co-factors/ hormones).

Amino acids are classified based on:

- *Standard and Non-standard amino acids (aa)*
- *Essential and non-essential aa*
- *Side chain functional group*

1. Standard amino-acids: Those 20 amino acids are encoded by universal genetic code

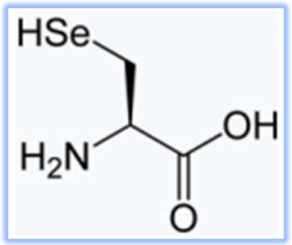
2. Non-Standard amino-acids: Two amino acids incorporated into proteins by unique synthetic mechanism

•*Selenocysteine: Incorporated when mRNA translated included SECIS (selenocysteine insertion seq) element, causes the UGA codon to encode selenocysteine instead of stop codon)*

•*Pyrrolysine: used by methanogenic archaea in enzyme that they use to produce methane. It is coded for UAG stop codon.*

Non-standard Amino Acids

- *Selenocysteine, 21st protein L-α amino acids*
- *Selenium atom replaces the sulfur of its elemental analog, cysteine*
- *Selenocysteine is not the product of a posttranslational modification, but is inserted directly into a growing polypeptide during translation.*
- *Selenocysteine is charged on a special tRNA called tRNASec specific for UGA (STOP) codon inserted into growing polypeptide during translation*



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

There are 20 common amino acids occurring in nature and most proteins comprise of these. Of these 20, 10 amino acids are ***essential*** *amino acids and must be obtained from our diets as the human body cannot synthesize them in adequate amounts. They are:*

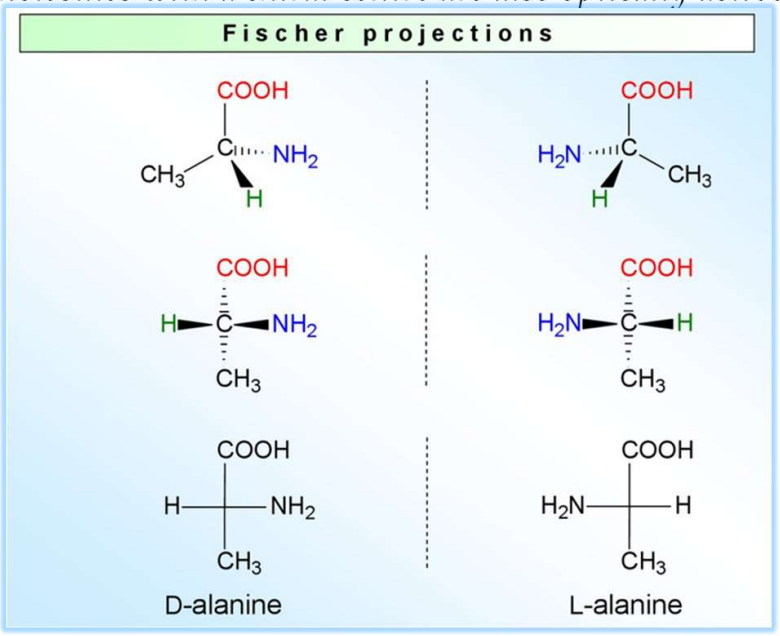
Essential	Non-essential
Arginine	Alanine
Histidine	Asparagine
Isoleucine	Aspartate
Leucine	Cysteine
Lysine	Glutamate
Methionine	Glutamine
Phenylalanine	Glycine
Threonine	Hydroxyproline
Tryptophan	Hydroxylysine
Valine	Proline
	Serine
	Tyrosine

derived

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Configuration of Amino Acids

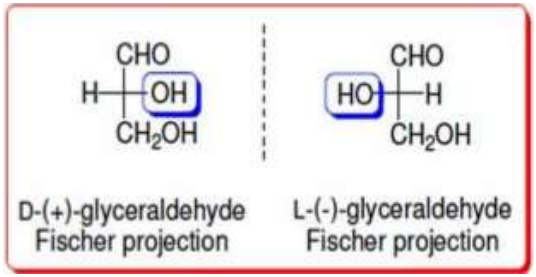
The α-carbon atom is thus a chiral centre. Because of the tetrahedral arrangement of the bonding orbitals around the α-carbon atom, the four different groups can occupy two unique spatial arrangements, and thus amino acids have two possible stereoisomers, non-superimposable mirror images of each other, called ***enantiomers****. All molecules with a chiral centre are also optically active-that is, they rotate the plane of plane polarized light.*

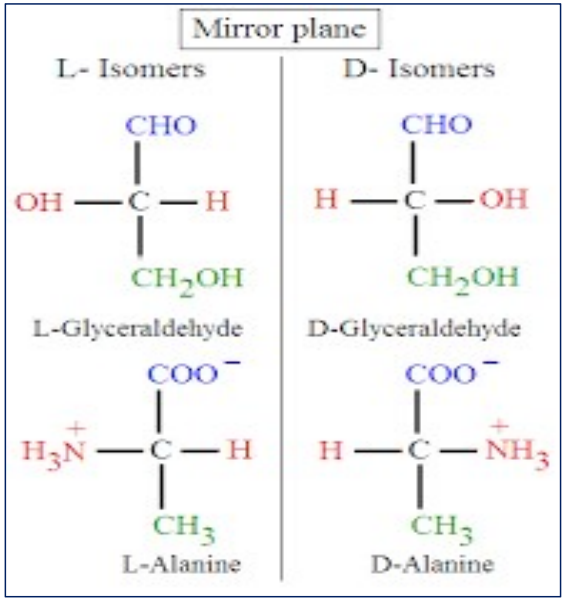


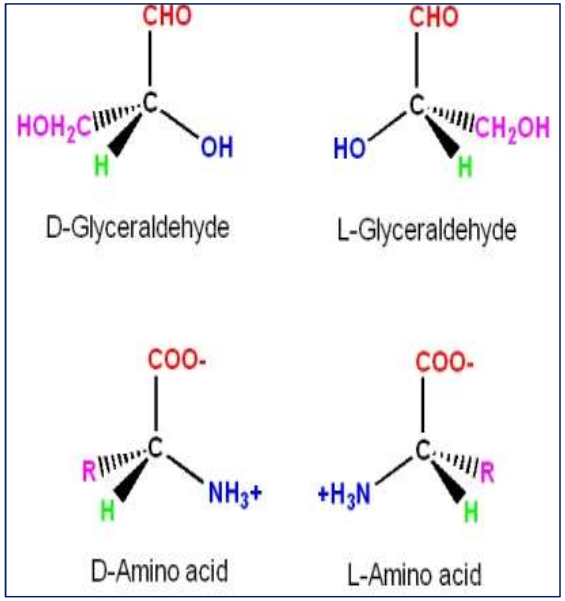
Only the L-stereoisomers with a configuration related to the absolute configuration of the reference molecule L –glyceraldehyde are found in proteins.

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

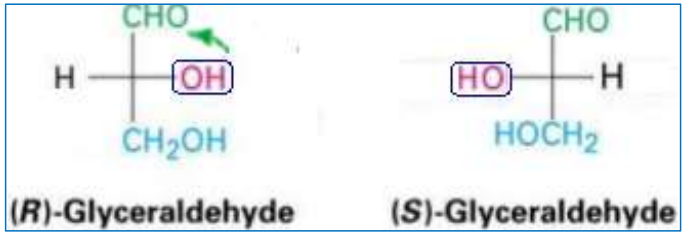
A D-amino acid is one where, in its Fischer projection formula, with carboxyl group on the top and R-group at the bottom of the vertical axis, amino group is on the right. Similarly, an L-amino acid has amino group on the left.





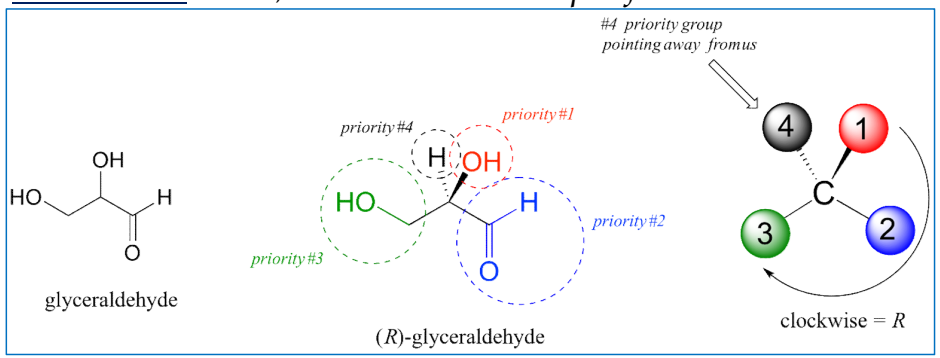


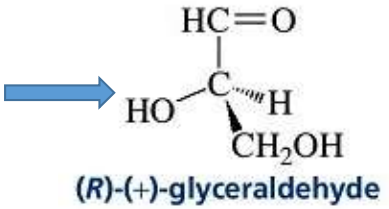
Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati



Counterclockwise is R (as H is on a horizontal bond)

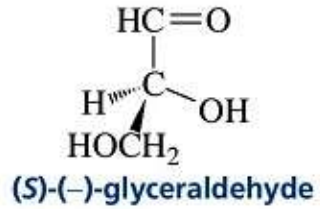
For the configuration of a Fischer projection formula, ***only if the lowest priority substituent is on a vertical bond****, will a clockwise arrow from the highest priority to the next highest priority substituent denote R-configuration & counterclockwise denote S. On the other hand,* ***if the lowest priority substituent is on a horizontal bond****, counterclockwise specifies R and clockwise specifies S configuration.*



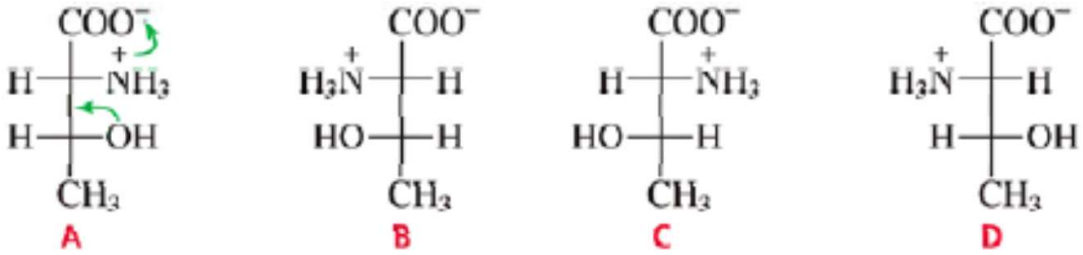


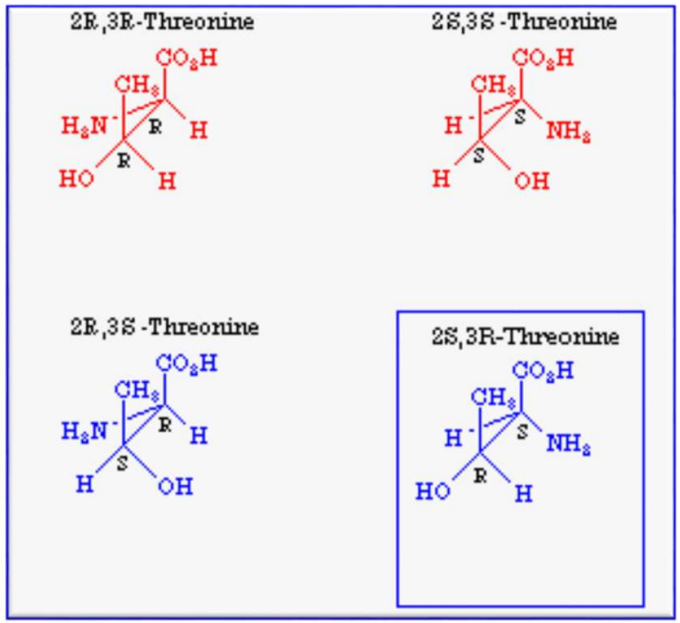
Clockwise is R

Anticlockwise is S



Threonine has two chiral centers: C-2, C-3; so it has $2^n = 2^2 = 4$ stereoisomers.





Naturally occurring L-threonine is (2S, 3R)-threonine. Thus, since H is on a horizontal bond, (2S, 3R) for C-2 & C-3 is denoted by stereoisomer D, where $^+NH_3$ group is on the left.

The common amino acids of proteins have been assigned three-letter abbreviations and one-letter symbols.

The three-letter code comprises of the first three letters of the amino acid name. e.g. Glycine: Gly

The one letter code indicates the composition and sequence of amino acids in proteins. e.g. Valine: V

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Abbreviations for Amino Acids

- Each amino acid has standard 3-letter and 1-letter abbreviations (shown in the table below)

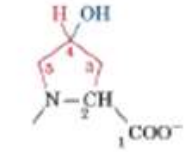
Name	3-Let	1-Let	Name	3-Let	1-Let
Glycine	Gly	G	Serine	Ser	S
Alanine	Ala	A	Threonine	Thr	T
Valine	Val	V	Asparagine	Asn	N
Leucine	Leu	L	Glutamine	Gln	Q
Isoleucine	Ile	I	Tyrosine	Tyr	Y
Phenylalanine	Phe	F	Aspartic Acid	Asp	D
Methionine	Met	M	Glutamic Acid	Glu	E
Proline	Pro	P	Lysine	Lys	K
Tryptophan	Trp	W	Arginine	Arg	R
Cysteine	Cys	C	Histidine	His	H

In addition to the 20 common amino acids, proteins may contain residues created by modification of common residues already incorporated into a polypeptide.

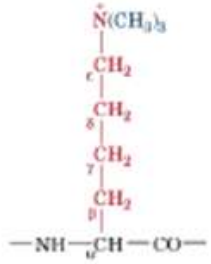
Among these uncommon amino acids are 4-hydroxyproline, a derivative of proline, and 5-hydroxylysine, derived from lysine. The former is found in plant cell wall proteins, and both are found in collagen, a fibrous protein of connective tissues.

6-N-Methyllysine is a constituent of myosin, a contractile protein of muscle. Another important uncommon amino acid is 1-carboxyglutamate, found in the blood-clotting protein prothrombin and in certain other proteins that bind Ca^{2+} as part of their biological function. More complex is desmosine, a derivative of four Lys residues, which is found in the fibrous protein elastin. Selenocysteine is a special case. This rare amino acid residue is introduced during protein synthesis rather than created through a postsynthetic modification. It contains selenium rather than the sulfur of cysteine. Actually derived from serine, selenocysteine is a constituent of just a few known proteins.

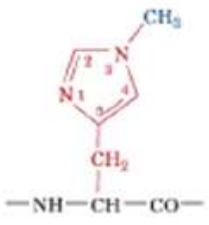
Uncommon amino acids that are components of certain proteins



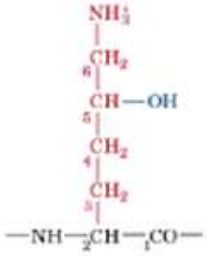
4-Hydroxyproline



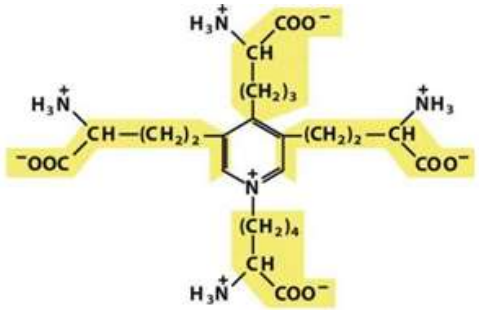
ε-*N*,*N*,*N*-Trimethyllysine



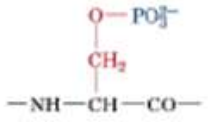
3-Methylhistidine



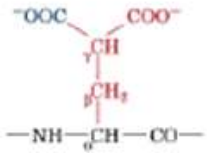
5-Hydroxylysine



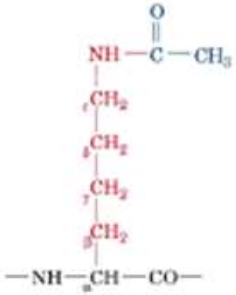
Desmosine



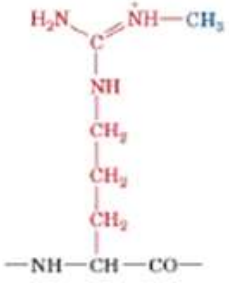
O-Phosphoserine



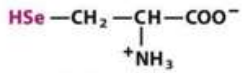
γ-Carboxyglutamate



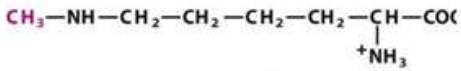
ε-*N*-Acetyllysine



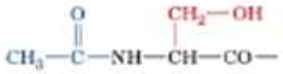
ω-*N*-Methylarginine



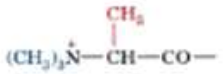
Selenocysteine



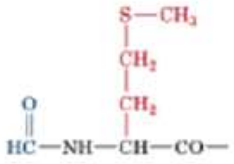
6-*N*-Methyllysine



N-Acetylserine



N,*N*,*N*-Trimethylalanine

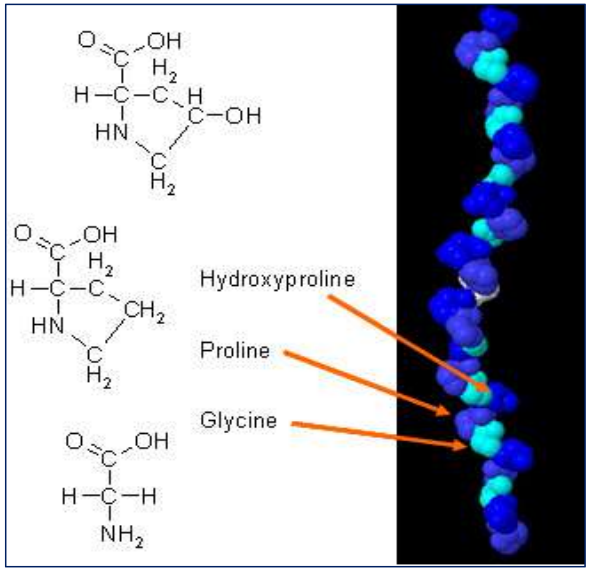


N-Formylmethionine

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Collagen is a tough protein found in tendons and connective tissue, in skin, bone, and teeth. It contains large amounts of glycine, proline, and hydroxyproline.

When the collagen molecule is assembled, it incorporates Pro [Proline] where Hyp [4-Hydroxyproline] is required. Once the protein is complete, some of the proline residues are oxidized to 4-hydroxyproline. This oxidation requires vitamin C to catalyze it. Thus vitamin C deficiency causes scurvy—symptoms of scurvy suffered by sailors (loose teeth, sores, and blisters) is caused by the inability to make collagen.



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

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

When an amino acid lacking an ionizable R group is dissolved in water at neutral pH, it exists in solution as the dipolar ion, or zwitterion ("hybrid ion"), which can act as either an acid or a base therefore amphoteric and are often called ampholytes.

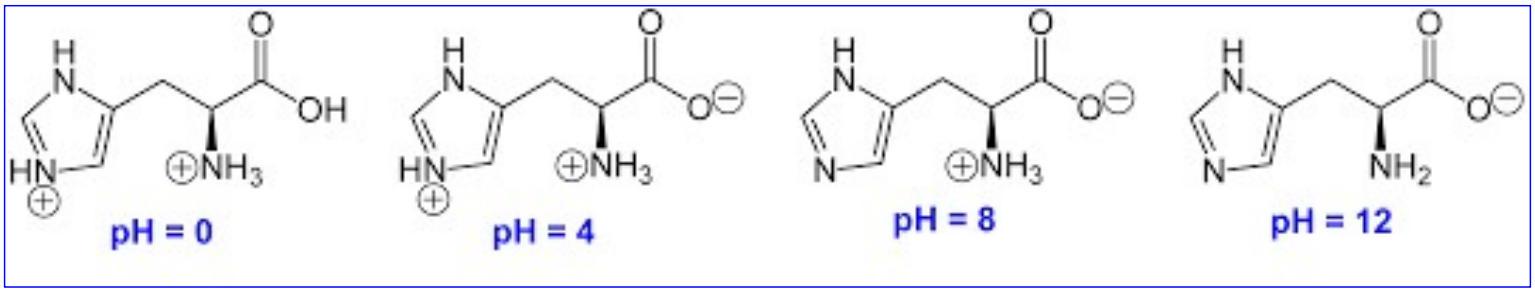
The pK_a is a measure of the tendency of a group to give up a proton, and it decreases ten times as the pK_a increases by one unit.

The pK_a of any functional group is greatly affected by its chemical environment, a phenomenon sometimes exploited in the active sites of enzymes to promote enzyme-catalyzed reaction mechanisms.

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

When an amino acid having side chain with ionizable hydrogens R group is dissolved in water it can exist in different forms or tautomers depending on the pH of the solution.

e.g. Histidine has 4 different forms



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

The characteristic pH at which the net electric charge is nil is called the ***isoelectric point*** *or* ***isoelectric pH****, designated as* ***pI****.*

pI = pH at which there is no net electric charge

For glycine, which has no ionizable group in its side chain, the isoelectric point is simply the arithmetic mean of the two pK_a values:

$$pI = \frac{2.34 + 9.69}{2} = 6.02$$

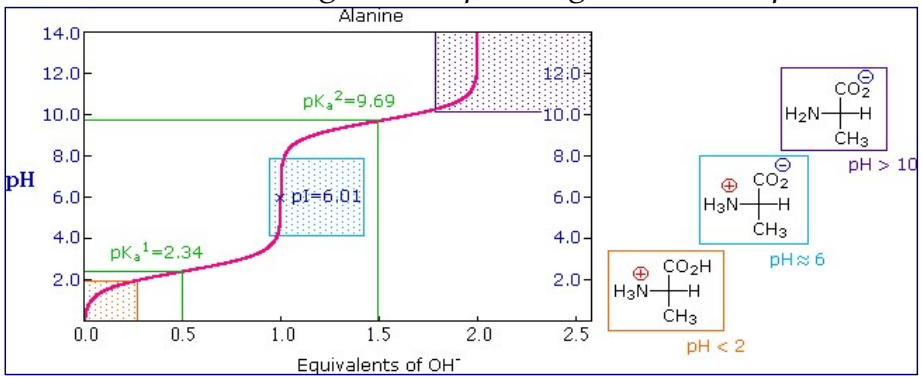
The pI of an amino acid whose side chain is not ionizable, is an average of its two pK_a values.

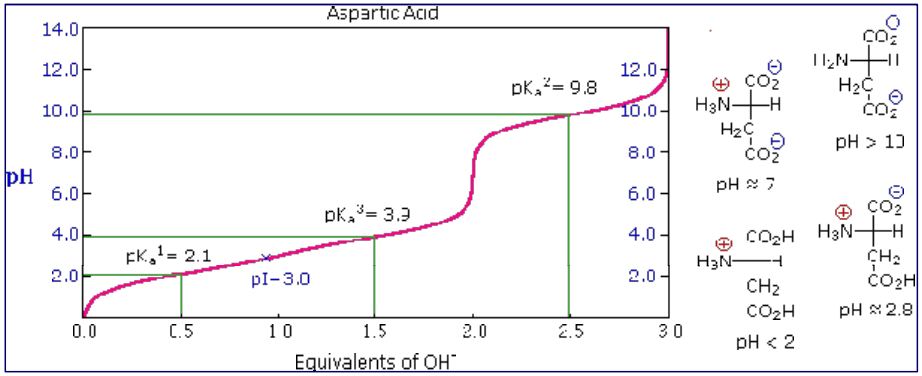
The pI of an amino acid which has an ionizable side chain, is the average of the pK_a values of similarly ionizing groups (either positively charged to uncharged groups or uncharged groups to negatively charged groups).

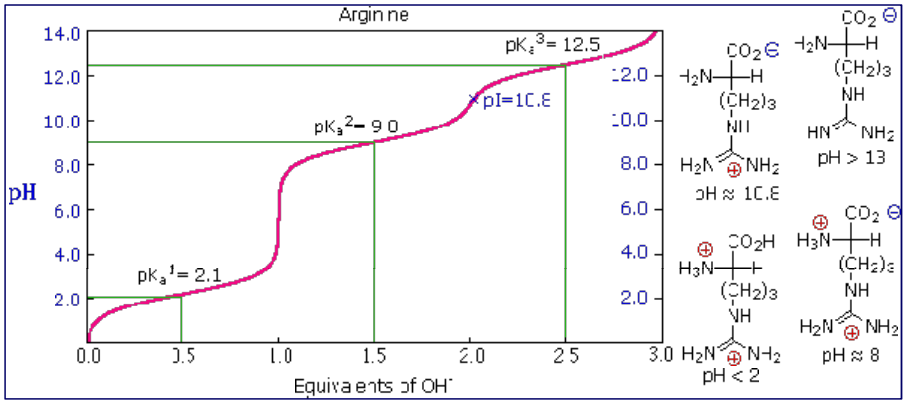
pK_a values of Amino acids

Amino acid	pK_a α-COOH	pK_a α–NH3	pK_a side chain
Alanine	2.34	9.69	-
Arginine	2.17	9.04	12.48
Asparagine	2.02	8.84	-
Aspartic acid	2.09	9.82	3.86
Cysteine	1.92	10.46	8.35
Glutamic acid	2.19	9.67	4.25
Glutamine	2.17	9.13	-
Glycine	2.34	9.60	-
Histidine	1.82	9.17	6.04
Isoleucine	2.36	9.68	-
Leucine	2.36	9.60	-
Lysine	2.18	8.95	10.79
Methionine	2.28	9.21	-
Phenylalanine	2.16	9.18	-
Proline	1.99	10.60	-
Serine	2.21	9.15	-
Threonine	2.63	9.10	-
Tryptophan	2.38	9.39	-
Tyrosine	2.20	9.11	10.07
Valine	2.32	9.62	-

Amino acids vary in their acid-base properties and have characteristic titration curves. All amino acids with a single α-amino group, a single α-carboxyl group, and an R group that does not ionize have titration curves resembling that of glycine/ alanine. However, amino acids with an ionizable R group have more complex titration curves, with three stages corresponding to the three possible ionization steps; thus they have three pK_a values.





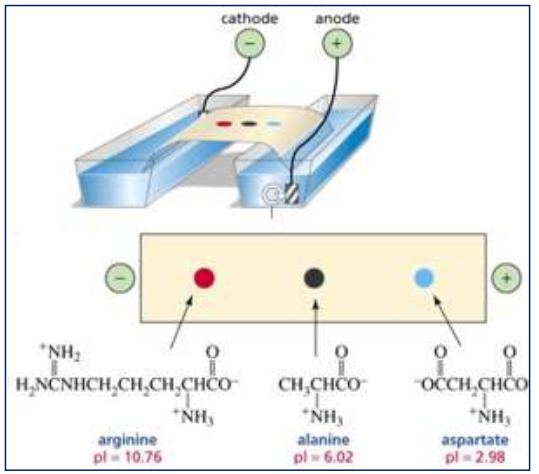


Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

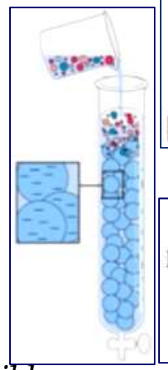
Separation of Amino Acids

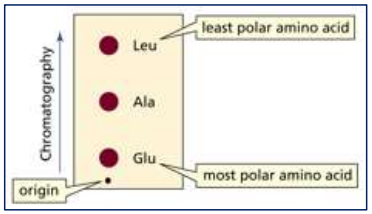
The various techniques by which the amino acids can be separated are:

Electrophoresis*: separates amino acids on basis of their* ***pI*** *values. Paper/gel containing amino acid mixture is placed in a buffered solution between two electrodes & an electric field is applied. Amino acid with* ***pI< pH*** *of buffer has overall negative charge & migrates towards anode. After separation, paper is dipped into ninhydrin solution & dried in an oven.*

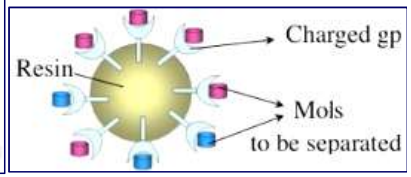


Paper/ Thin Layer Chromatography*: separates amino acids on basis of their* ***polarity****. Least polar amino acid travels farthest, most polar amino acid travels least. Then paper is developed with ninhydrin to reveal the spots.*

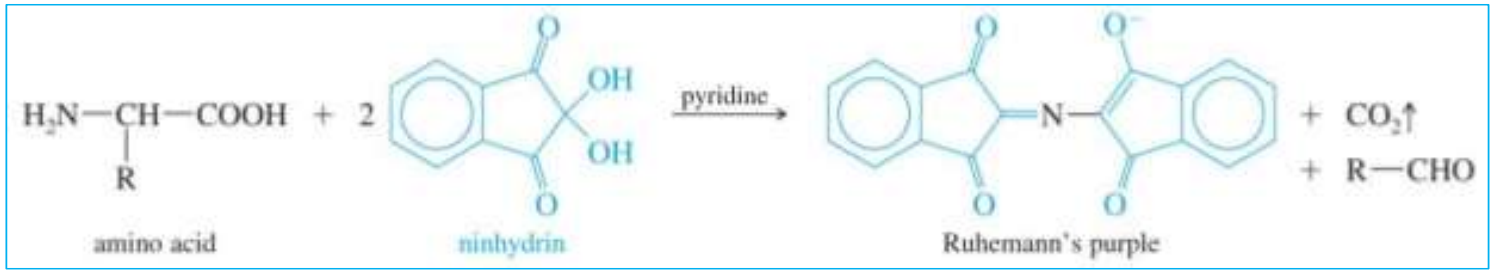




Ion-Exchange Chromatography*: can both identify & determine relative amounts of AAs in the mixture. A soln of mixture of amino acids is loaded onto the top of a column packed with an insoluble resin and move through the column at different rates separated differentially in fractions, ninhydrin added, relative amounts of amino acids found using visible spectroscopy. An instrument that automates ion-exchange chromatography is called an amino acid analyzer.*



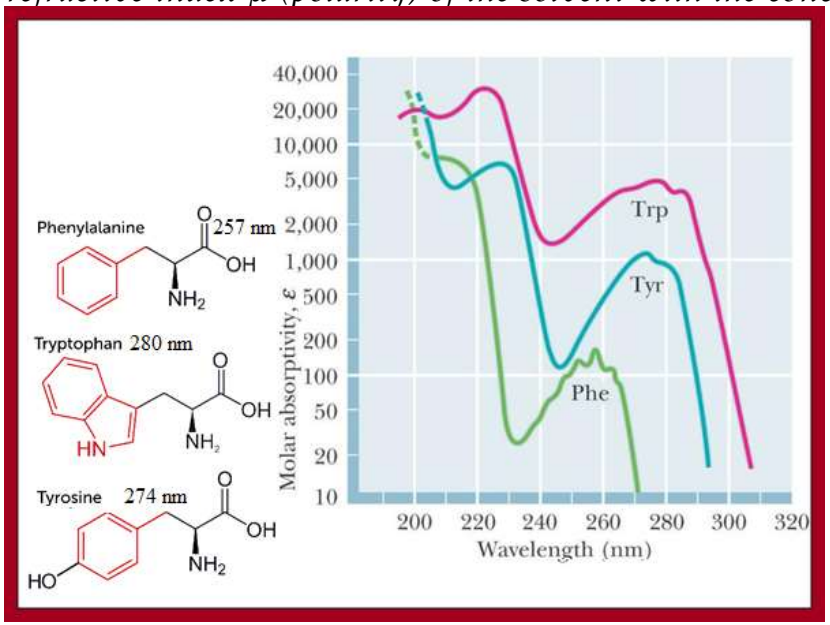
Electrophoresis followed by Ninhydrin spray for identification of amino acids:

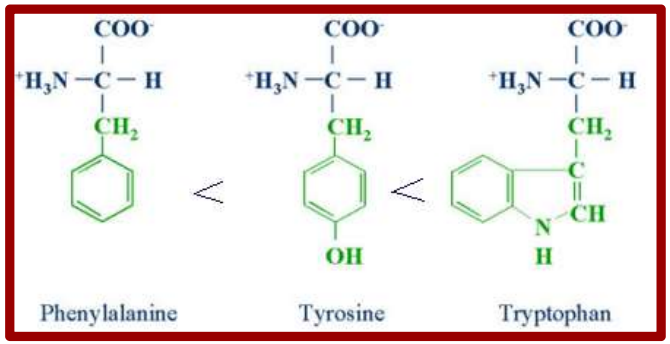


Which aldehyde is formed when valine is treated with ninhydrin?

UV absorbance properties of the Aromatic Amino Acids

The aromatic side-chains of Tyr, Trp and Phe absorb light in the 'far-UV' range (180–230 nm) and, also in the 240–300 nm region. Absorption spectroscopy is excellent for monitoring ligand-binding interactions, enzyme catalysis & conformational transitions in proteins: there will be a difference in absorbance between the 3D-folded protein and unfolded/denatured forms of a protein. Denaturation/ unfolding by heat or reagents [urea or guanidinium chloride] can be quantitatively monitored by locating the absorbance patterns at 287–292 nm [Tyrosine and Tryptophan at 287 nm and at 291 nm, respectively] as a function of temperature or denaturant concentration, increasing slightly, due to alterations in refractive index μ (polarity) of the solvent with the concentration of denaturants.





Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Absorption of Light by molecules: Beer- Lambert's Law

Measurement of light absorption by a spectrophotometer is used to detect and identify biomolecules like amino acids and to measure their concentration in solution. The incident light absorbed by a solution at a given wavelength is related to the thickness of the absorbing layer (path length) and to the concentration of the absorbing species.

These two relationships are combined into the Lambert-Beer law:

$$\boxed{Log\ I_o/I = \varepsilon cl}$$

where,

I_o *is the intensity of the incident light,*

I is the intensity of the transmitted light, the ratio I/I_o *(inverse of the ratio in the equation) is the transmittance,*

ε is the molar extinction coefficient (in units of moles per litre),

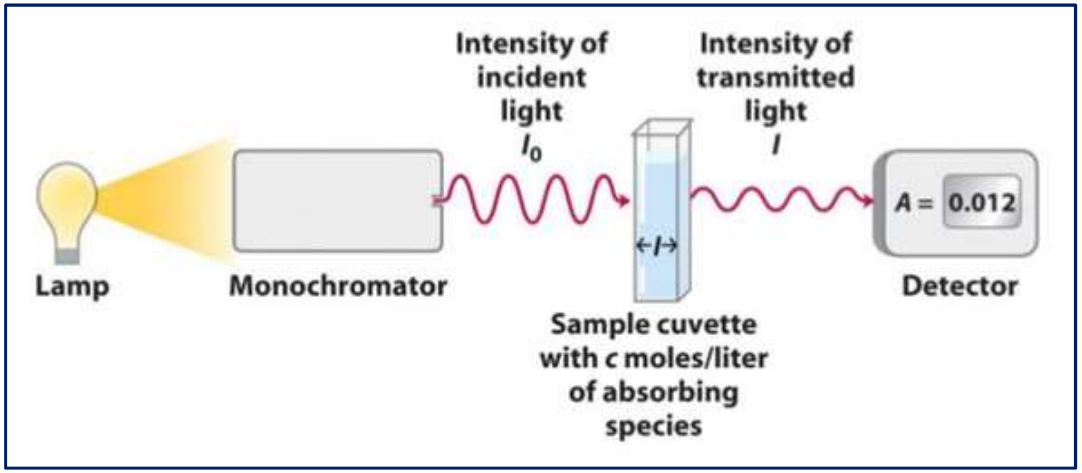
c is the concentration of the absorbing species (in moles per litre), and

l is the path length of the light-absorbing sample (in cm),

$Log(I_o/I)$ *is called absorbance A, and is directly proportional to the conc. of the absorbing solute.*

ε , the molar extinction coefficient, changes with the nature of the absorbing compound, the solvent, the wavelength & also with pH.

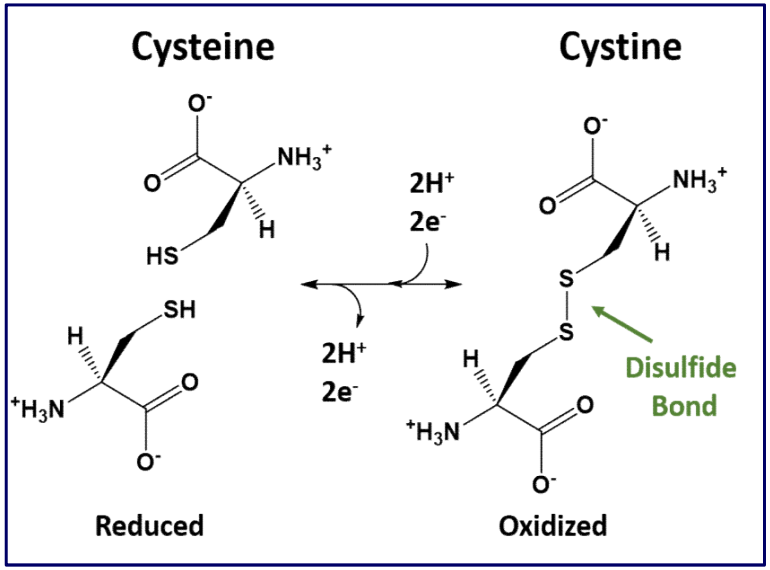
Measurement of light absorption by a Spectrophotometer



The Beer-Lambert law assumes that the incident light is parallel and monochromatic (of a single wavelength) and that the solvent and solute molecules are randomly oriented. The expression log (I_0/I) is called the absorbance, designated as A.

With an absorbing layer of fixed path length, the absorbance A, is directly proportional to the concentration of the absorbing solute.

The molar extinction coefficient varies with the nature of the absorbing compound, the solvent, and the wavelength, and also with pH if the light-absorbing species is in equilibrium with an ionization state that has different absorbance properties.



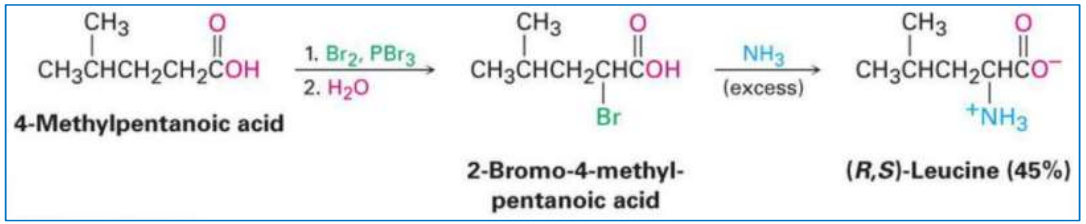
Reversible disulfide bond formation by the oxidation of two molecules of cysteine to form cystine. Disulfide linkages stabilize protein structure.

- *They influence the thermodynamics of protein folding: disulfide bonds stabilize the native conformation of a protein by destabilizing the unfolded form, as the entropy of the unfolded form is lower.*
- *Disulfide bonds restrict energetically favourable conformational changes.*
- *They also maintain protein integrity; oxidants and proteolytic enzymes in the extracellular environment may lead to protein inactivation.*

The Synthesis of Amino Acids

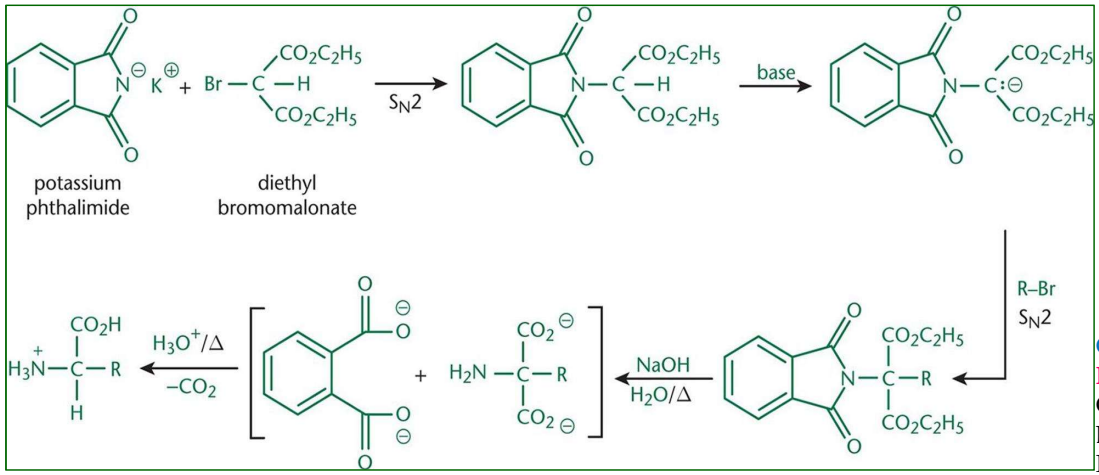
The various ways in which amino acids can be synthesized are:

1. *HVZ Reaction [Reaction with NH_3]:*



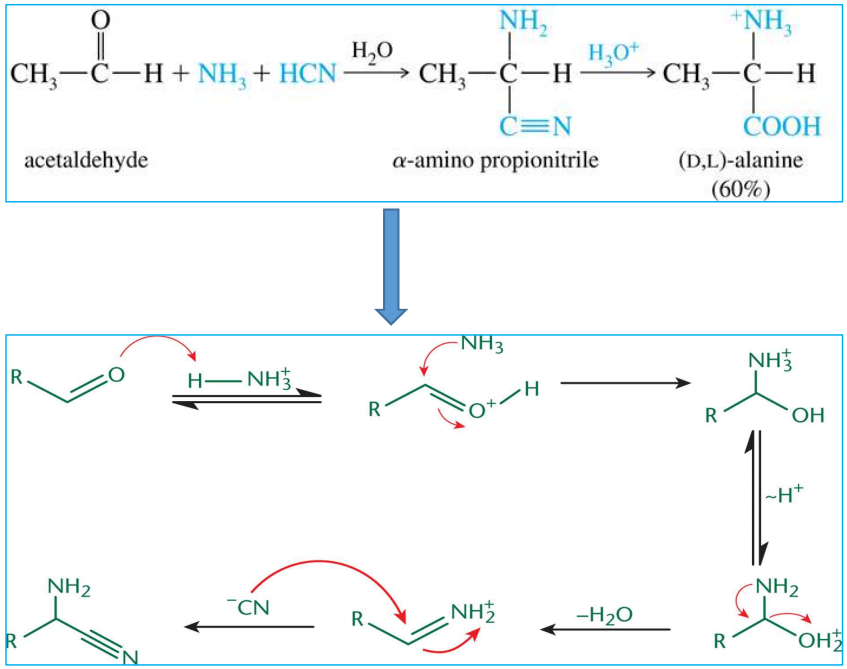
2. *Reductive Amination: of an α-keto acid:*

3. *Gabriel {N-Phthalimidomalonic Ester} Synthesis:*



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

4. Strecker Synthesis



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

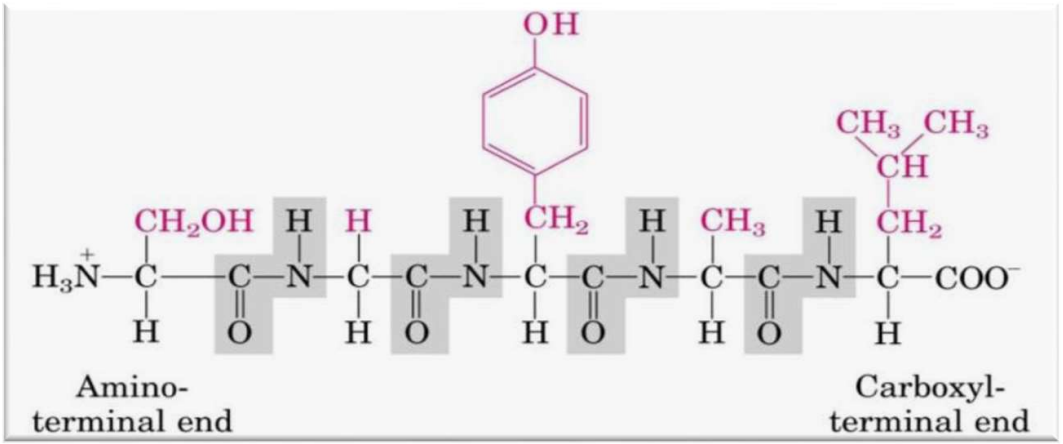
Synthesis of Peptides

Two amino acid molecules can be covalently joined through a substituted amide linkage, termed a peptide bond, to yield a dipeptide. Few (2-20) amino acids bond to form oligopeptides, and many to form polypeptides (molecular weights below 10,000). Proteins are higher polypeptides (40 to 4000 amino acids).

This is an example of a condensation reaction in which the water of dehydration that is removed comes from the α-carboxyl group of one amino acid and the α-amino group of another.

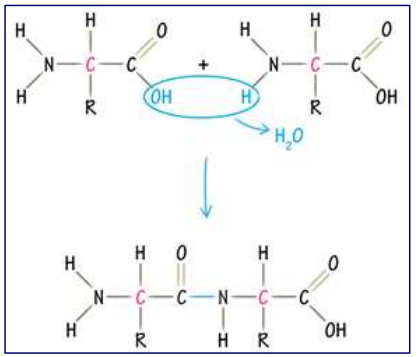
Peptides contain only one free α-amino group and one free α-carboxyl group, at opposite ends of the chain.

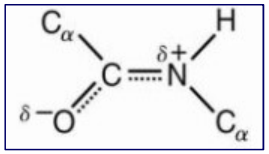
Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati



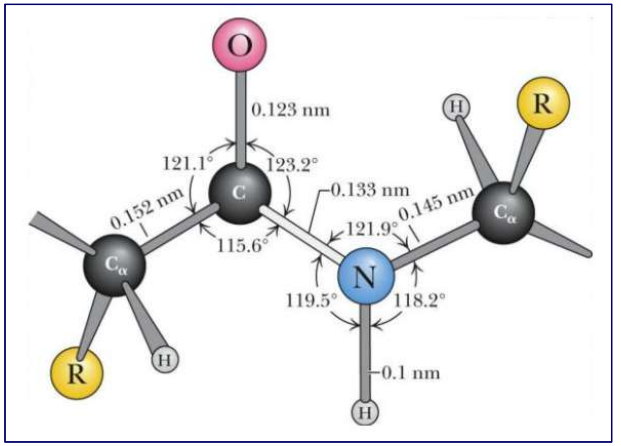
The pentapeptide serylglycyltyrosylalanylleucine, Ser-Gly-Tyr-Ala-Leu, or SGYAL. Peptides are named beginning with the amino-terminal residue, which by convention is placed at the left.

Peptides & Proteins





- *Peptides are functionally amides formed by interaction of carboxyl group (-COOH) of one amino acid and amino group (NH_2) of another amino acid by condensation/ dehydration.*
- *The peptide covalent bond exists in trans conformation*
- *It has partial [~40%] double bond character (N: δ+, O: δ-)[A pure double bond between C & O permits free rotation about the linked C – N bond*
- *The 6 atoms of peptide bond / amide gp are always planar.*
- *The bond is intermediate between a single & double bond [0.133 nm]*
- *The peptide backbone comprises of the repeating sequence N-C_α-C_O*



They are generally classified as:

Oligopeptides (2-20 amino acids)
- *Dipeptide (2 amino acids)*
- *Tripeptide (3 amino acids)*
- *Tetrapeptide (4 amino acids)*

Polypeptides (20-10,000 amino acids)

Proteins (>10,000 amino acids)

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

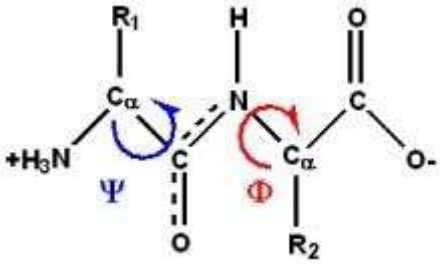
Rotation about two single bonds in a polypeptide chain determines its path of folding

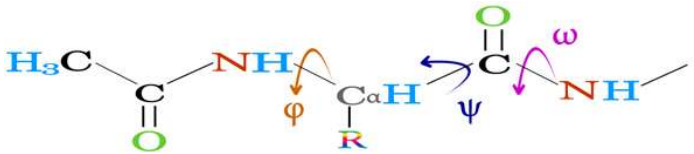
*The **alpha carbon (C_α)** in the center of each amino acid is held in the main chain by two rotatable bonds. The dihedral (torsion) angles of these bonds are called **Phi** and **Psi** (**φ** and **ψ**). The two angles of rotation Phi & Psi are responsible for the 3D polypeptide/ protein folding.*

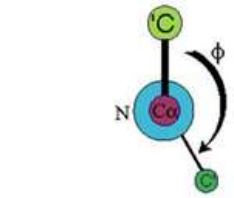
*Phi (**φ**) is the torsional angle of rotation about the N-C_α bond*

*Psi (**ψ**) is the torsional angle of rotation about the C_α-C_O bond*

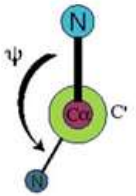
$-180 \leq \varphi / \psi \geq 180$







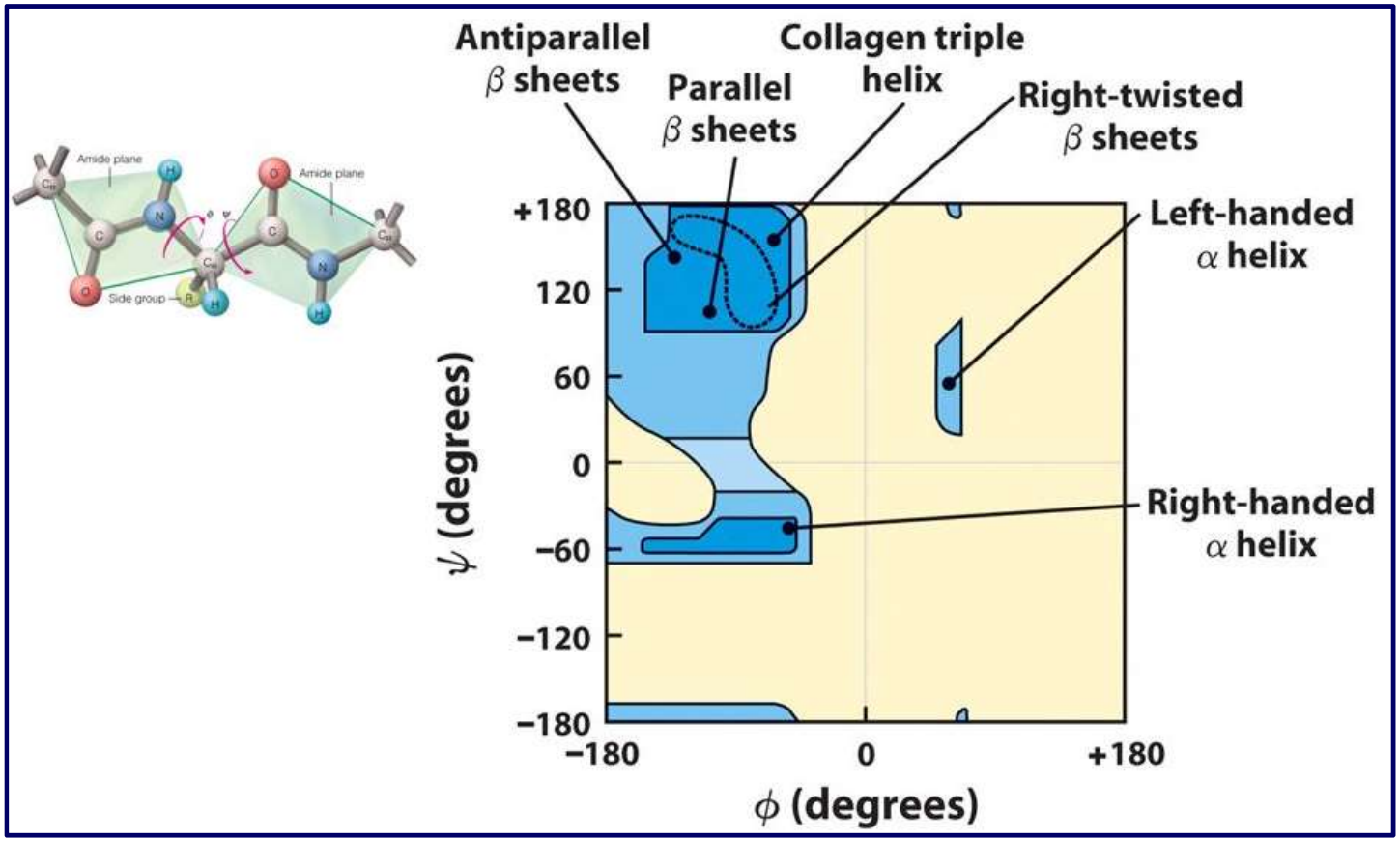
This torsion angle ϕ is right handed and is therefore positive



This torsion angle ψ is left-handed and is therefore negative

*In the early 1960's, **G N Ramachandran** computationally determined the phi and psi angles that avoid steric collisions, initially treating the atoms simply as rigid spheres. He showed that the physically allowed angle combinations (that avoid clashes) correspond largely to the secondary structures observed in proteins: α-helices, β-sheets and turns/loops.*

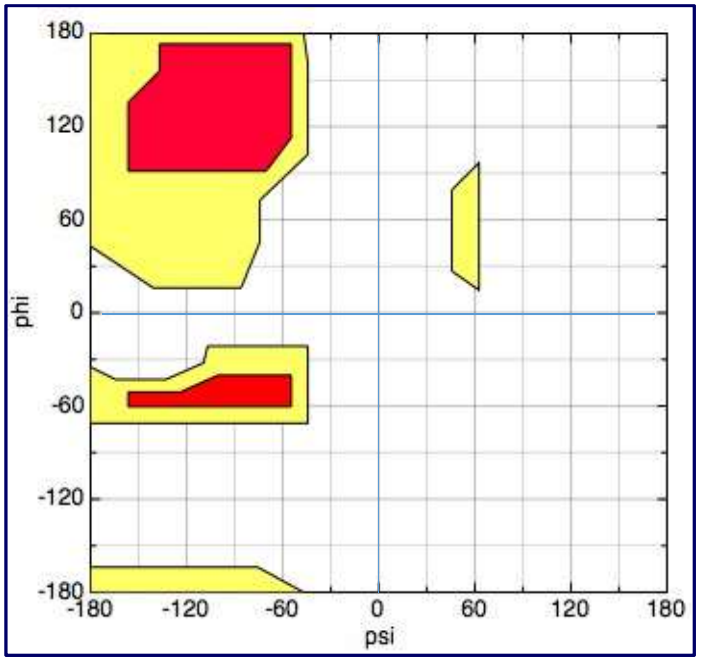
Ramachandran Plot



Glycine and proline are not included in Ramachandran plot

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

- *In the diagram above the white areas correspond to conformations where atoms in the polypeptide come closer than the sum of their van der Waals radii. These regions are sterically disallowed for all amino acids except glycine which is unique in that it lacks a side chain.*
- *The red regions correspond to conformations where there are no steric clashes, -the allowed regions namely the α-helical and β-sheet conformations.*
- *The yellow areas show the allowed regions if slightly shorter van der Waals radii are used in the calculation, i.e. the atoms are allowed to come a little closer together. This brings out an additional region which corresponds to the left-handed α-helix.*
- *Disallowed regions generally involve steric hindrance between the side chain C-β-methylene group and main chain atoms.*
- *Glycine has no side chain and therefore can adopt phi and psi angles in all four quadrants of the Ramachandran plot. Hence it frequently occurs in turn regions of proteins where any other residue would be sterically hindered.*



Glycine has a much wider low-energy area because it does not have a Ca atom

Glycine is formally nonpolar, its very small side chain makes no real contribution to hydrophobic interactions

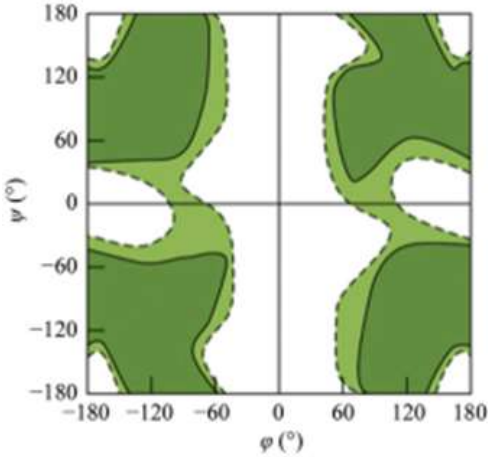
Proline has its side chain covalently bound to backbone amine; hence its phi angle is limited to the range of phi = -60° +/- 20°

Proline has an aliphatic side chain with a distinctive cyclic structure. The secondary amino (imino) group of proline residues is held in a rigid conformation that reduces the structural flexibility of polypeptide regions containing proline

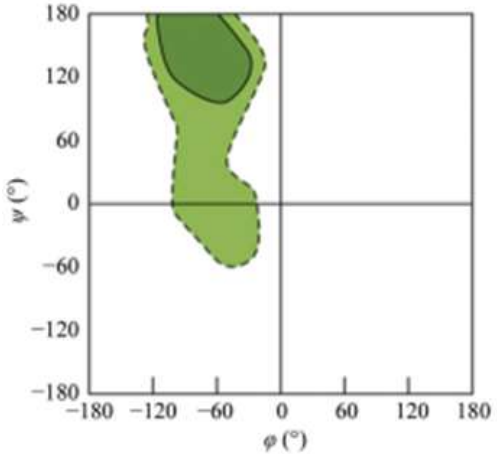
Proline mainly occurs in the first turn of an a-helix because it can not donate a hydrogen bond in the middle of a helix, and it creates sterical problems in a-helical conformation

Glycine is an exceptional amino acid because it has the flexibility to occur at phi-psi combinations that are not tolerated for other amino acids

Glycine



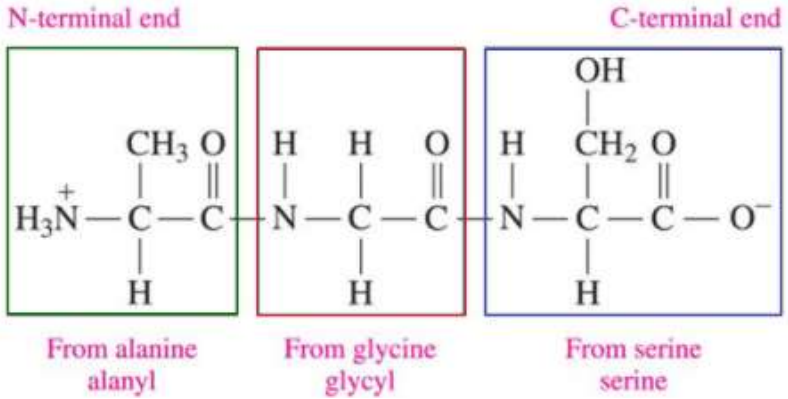
Proline

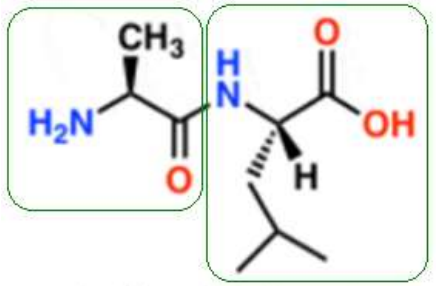


Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Nomenclature of Peptides

- *The C-terminal amino acid with –NH_2 end is considered as first {LHS}*
- *'-ine' replaced by '-yl'*
- *The C-terminal amino acid with carboxyl end is considered as parent, so name is unchanged {RHS}*

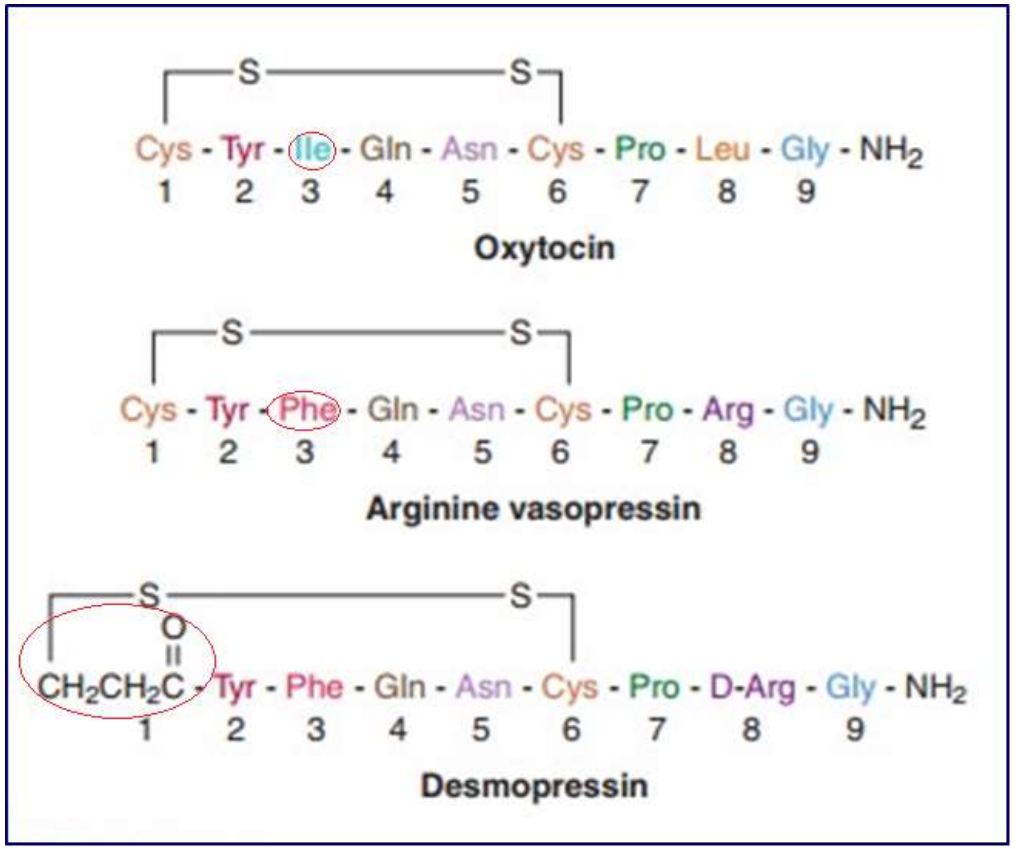




Alanylglycylserine

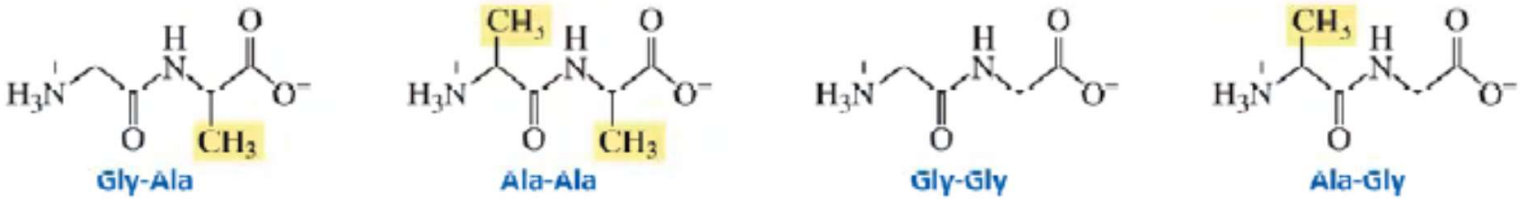
Ala-Gly-Ser, AGS

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

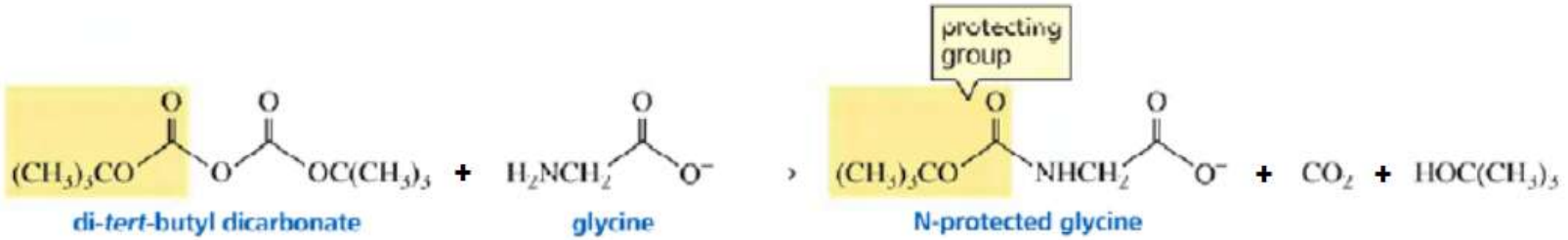


Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

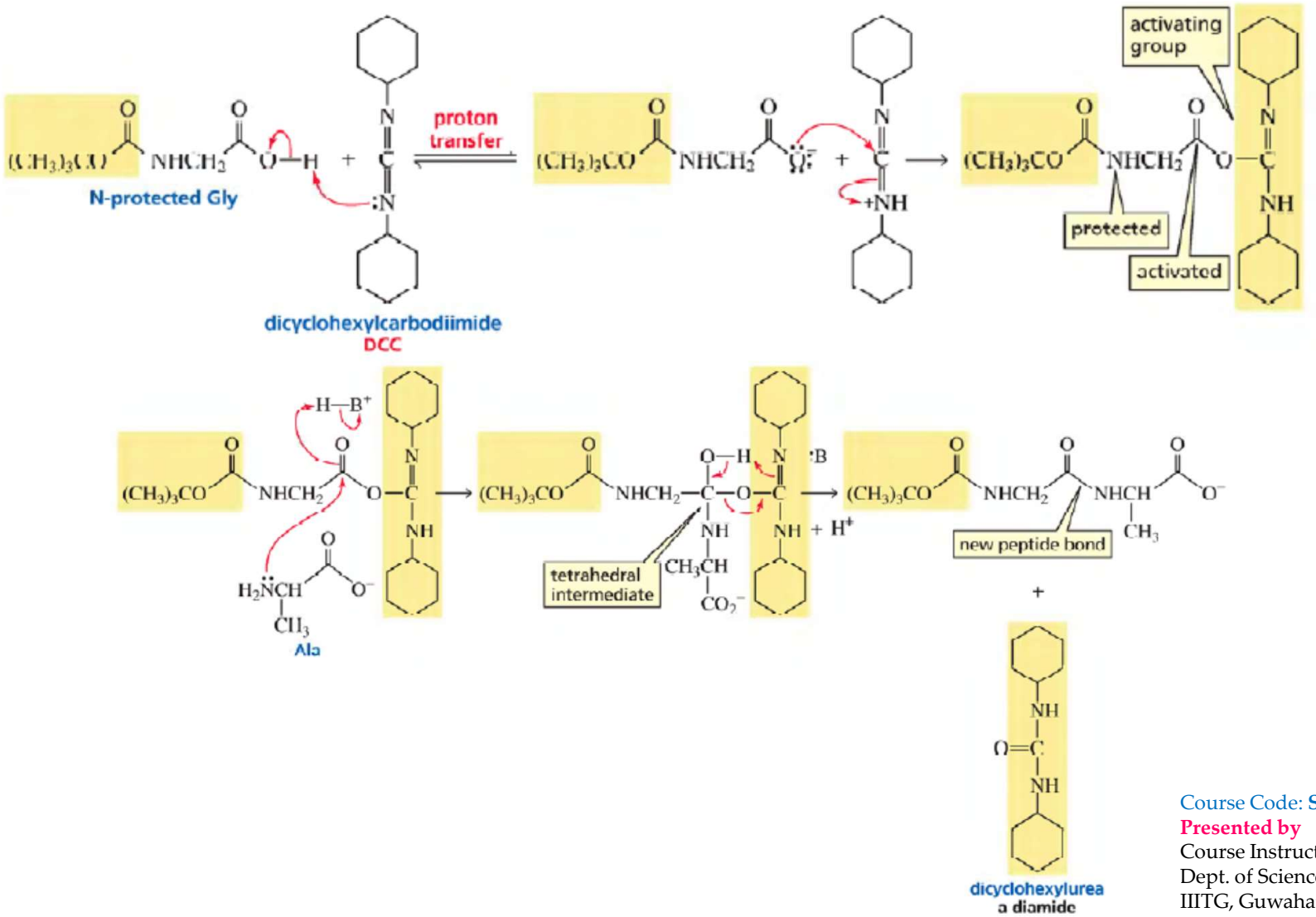
Strategy of Peptide bond synthesis: *N-protection & C-activation: For dipeptide Gly-Ala, Combination of Glycine and Alanine without N-protection & C-activation gives a mixture*



With N-protection & C-activation:



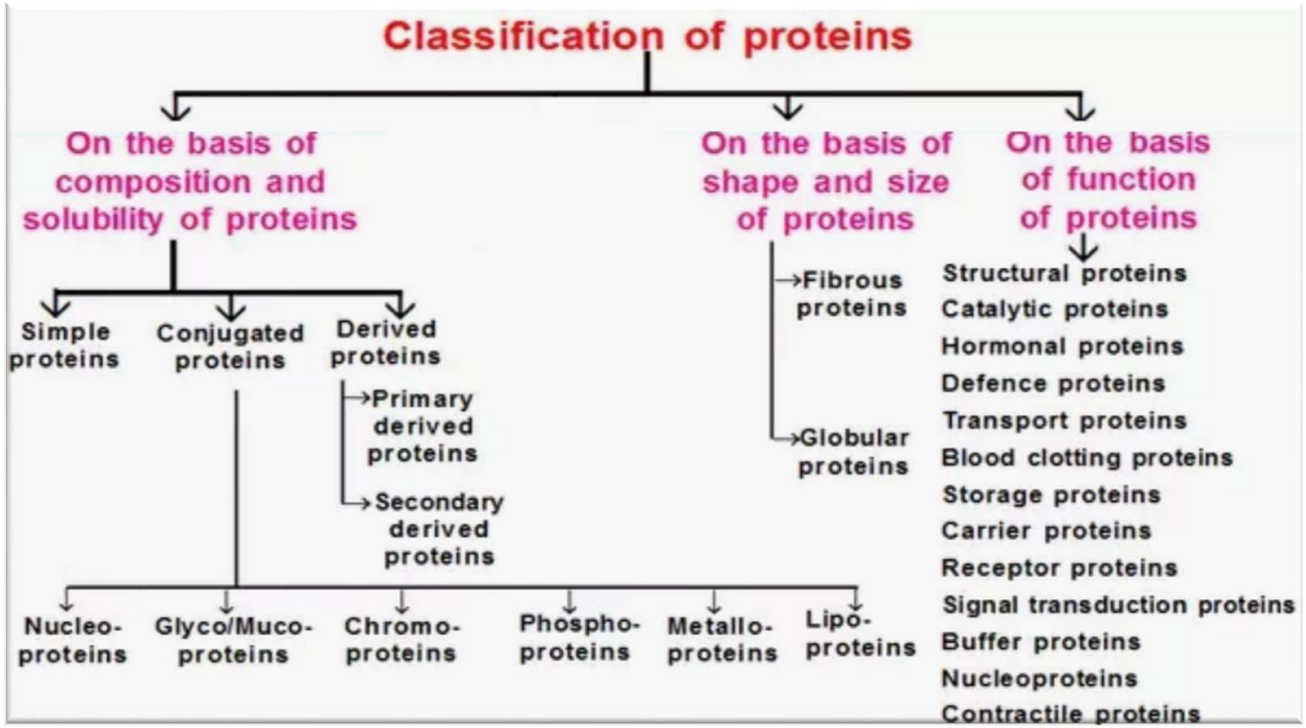
Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Proteins

Biologically occurring polypeptides (40-4000 amino acids) consisting of 40 to thousands of linked amino acid residues are called proteins.



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

*Proteins are classified on account of their **structure** as :*

***Fibrous**: These are insoluble in water, long, thread-like fibres e.g. keratin (skin, hair, nails), collagen (bones, teeth), elastin*

- Axial ratio more than 10.
- Long thread like molecule.
- Their helical strands mainly form fibers.
- These protein are insoluble in water.
- Form structure of the tissue
- Present where support is required.

***Globular**: These are soluble in water or aqueous solutions of acids, bases or salts and are of spheroidal shape. e.g. all enzymes & many hormones; hemoglobin (which transports oxygen from the lungs to the tissues), albumin, globulin*

- Axial ratio less than 10.
- Spheroid or ovoid in shape.
- Enzyme are mostly globular in shape.
- Subdivided into two type of protein…

1. Albumins: Water soluble.
2. Globulin: Soluble in dilute salt solution.

Simple & Conjugated Proteins

Based on their chemical composition, proteins are classified as:

***Simple proteins**: Many proteins, for example the enzymes ribonuclease A and chymotrypsin, contain only amino acid residues and no other chemical constituents; these are considered simple proteins.*

***Conjugated proteins**: some proteins contain permanently associated chemical components in addition to amino acids; these are called conjugated proteins. The non-amino acid part of a conjugated protein is usually called its prosthetic group. Conjugated proteins are classified on the basis of the chemical nature of their prosthetic groups; for example,*

Lipoproteins contain lipids,

Glycoproteins contain sugar groups,

Phosphoproteins contain phosphate groups,

Hemoproteins contain iron porphyrin,

Flavoproteins contain flavin nucleotides

Metalloproteins contain a specific metal

Types	Prosthetic Group	Properties	Example
Nucleoproteins	Nucleic acid (DNA,RNA)	Large, compact complexes	Chromatin, ribosomes
Mucoproteins*	Carbohydrate	More than 4% carbohydrate by weight	Human chorionic gonadotrophin, a hormone used to test for pregnancy
Glycoproteins * Lipoproteins	Carbohydrate Lipid	Less than 4% carbohydrate Water soluble	Antibodies Serum lipoproteins
Proteolipids	Lipid	Not very water soluble, soluble in nonpolar solvents	Cell membranes
Hemoproteins	Heme group	Characteristic color	Hemoglobin, cytochrome c
Metalloproteins	Metal ion (Fe^{3+}, Zn^{2+} Mg^{2+}, Mn^{2+})	Require a metal ion to function	Carbonic anhydrase

Some proteins contain more than one prosthetic group.

The primary structure of a protein dictates how it folds into a 3D structure, thus determining its function.

Based on their functions, proteins can be classified as:

Structural proteins: provide strength and protection e.g. collagen in bones

Protective proteins: protect from disease e.g. peptide antibiotics, blood-clotting proteins

Enzymes: catalyze cellular reactions

Hormones: regulate biological reactions

Proteins with physiological functions: transport and store oxygen in the body, muscles.

Storage proteins: that act as reserves of energy/minerals

Class of Protein	Function in the body	Examples
Structural	Provide structural components	Collagen Keratin
Contractile	Move muscles	Myosin Actin
Transport	Carry essential substances throughout the body	Hemoglobin Lipoprotein
Storage	Store nutrients	Casein Ferritin
Hormone	Regulate body metabolism and nervous system	Insulin Growth hormone
Enzyme	Catalyze biochemical reactions in the cells	Sucrase Trypsin
Protection	Recognize and destroy foreign substances	immunoglobulins

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Structure of Proteins

Primary structure *is the sequence of amino acid residues. [Assembly]*

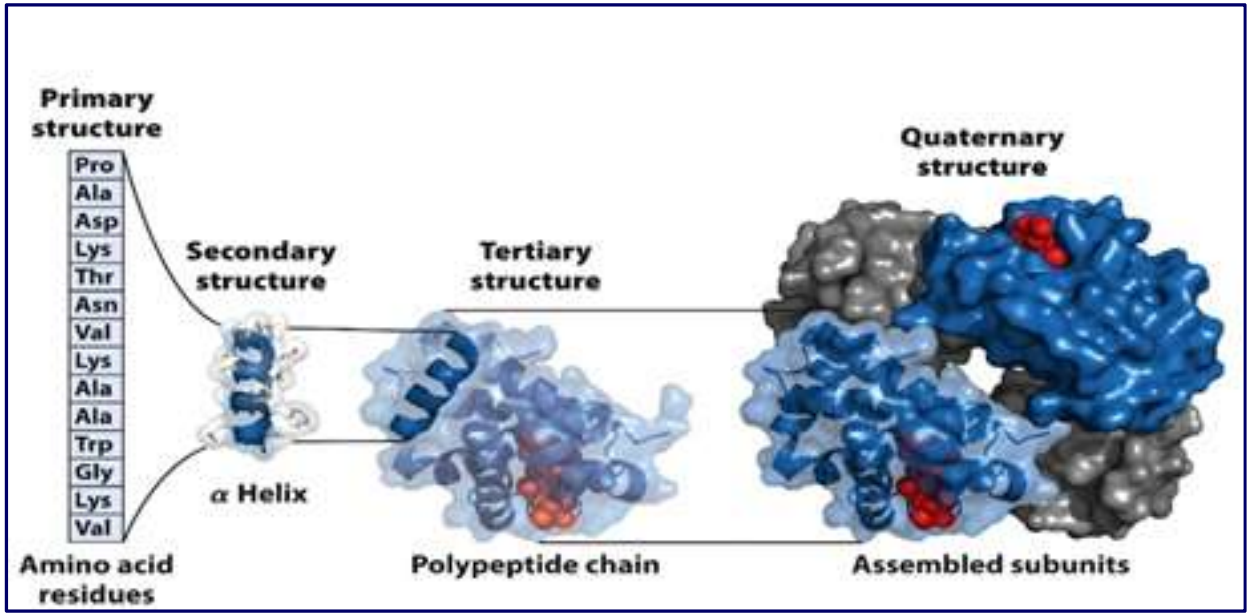
Secondary structure *refers to particularly stable arrangements of amino acid residues*

giving rise to recurring structural patterns. e.g. α-helix, β-sheet, loops or reverse turns etc. [Folding]

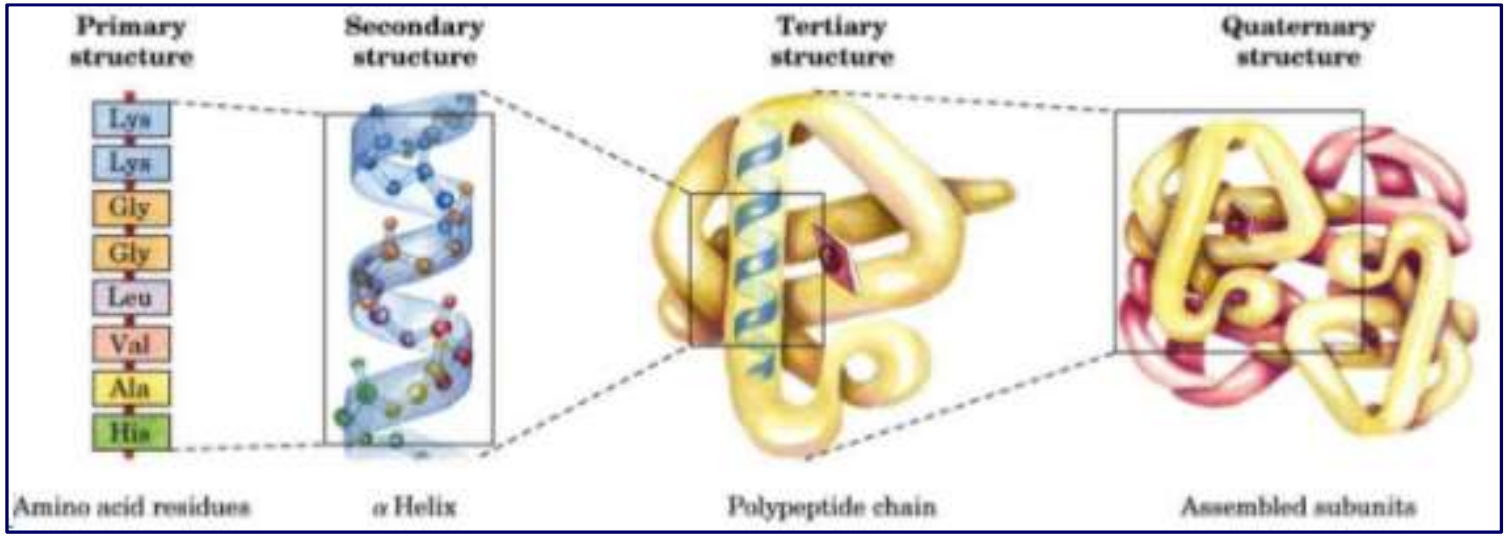
Tertiary structure *describes all aspects of the three-dimensional folding of a polypeptide. [Packing]*

When a protein has two or more polypeptide subunits, their arrangement in space is referred to as ***quaternary structure****.*

The 4˚structure describes the way the subunits are arranged with respect to each other. [Interaction]



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati



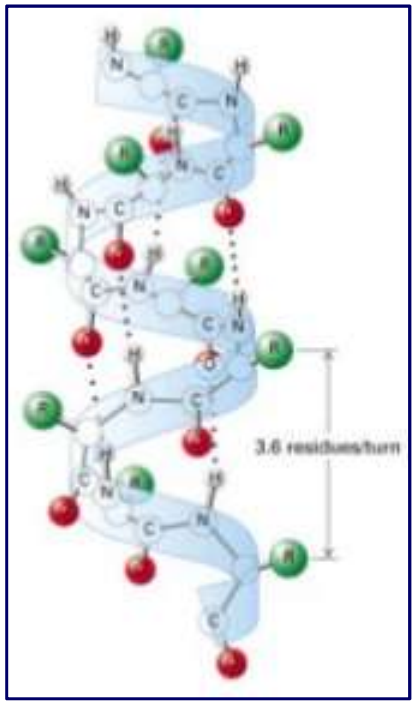
Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Secondary Structure

- The polar N-H and C=O peptide units in the interior of the protein are held by **H-bonds**
- Two types which are regular structures in protein
- **α-helix** and **β-pleated sheet**

α- Helix features

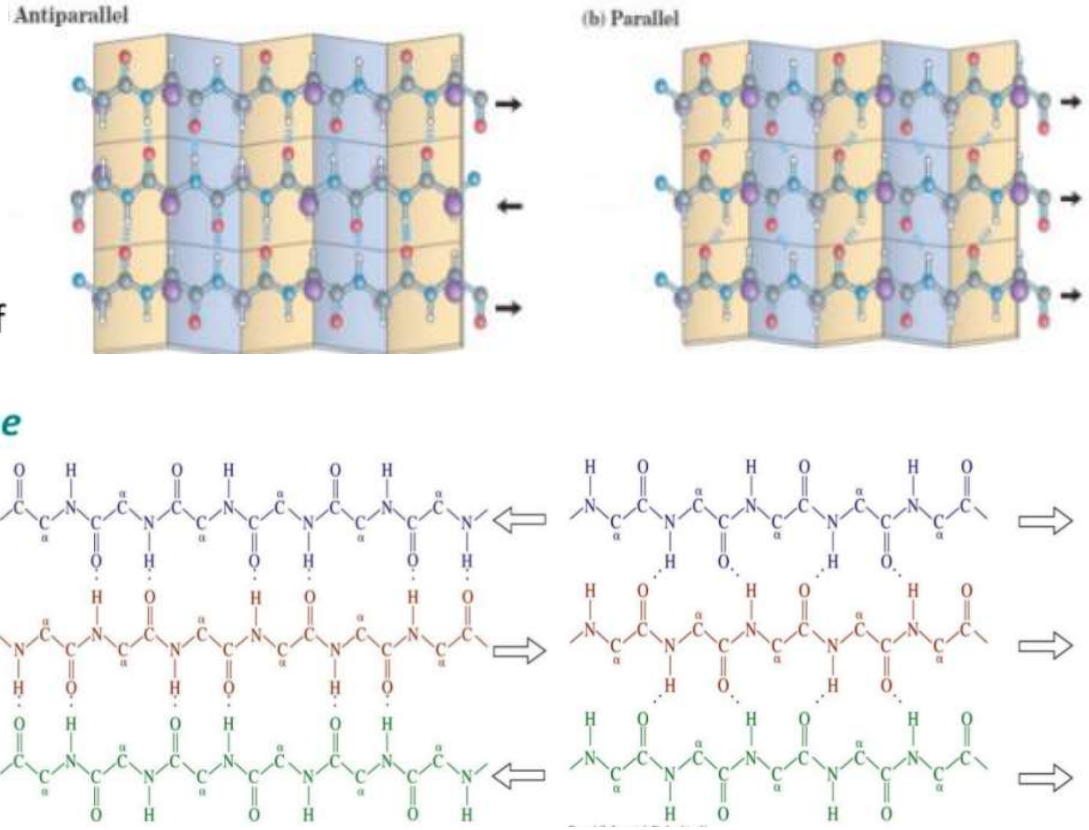
- Coil direction – **left** handed or **right** handed
- **L- amino acids** favor the **right hand coil**
- One coil has about 3.6 aa residues; there can be several coils with 650 aa residues(1000Å)
- Average length of helix in a globular protein is 15-20Å
- H-bonds occur between **1st O of backbone C=O** to **13th H atom of backbone NH**
- The presence of the ff amino acids *do not favor the helix* formation: **Pro,** adjacent basic or acidic amino acids, **Asn, Tyr, Ser, Thr, Ile and Cys**



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

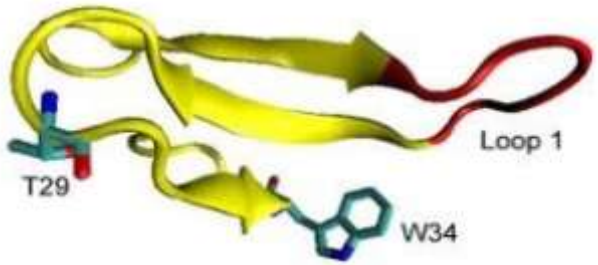
β-sheet features:

- Two adjacent peptides
- **Parallel** (both N→C or C→N)
- **Antiparallel** (N to C running in opposite directions)
- Antiparallel more common in the structure of proteins
- Peptides with this structure are ***rich in alanine and glycine*** (silk fiber and spider web)



Loop or coil conformation

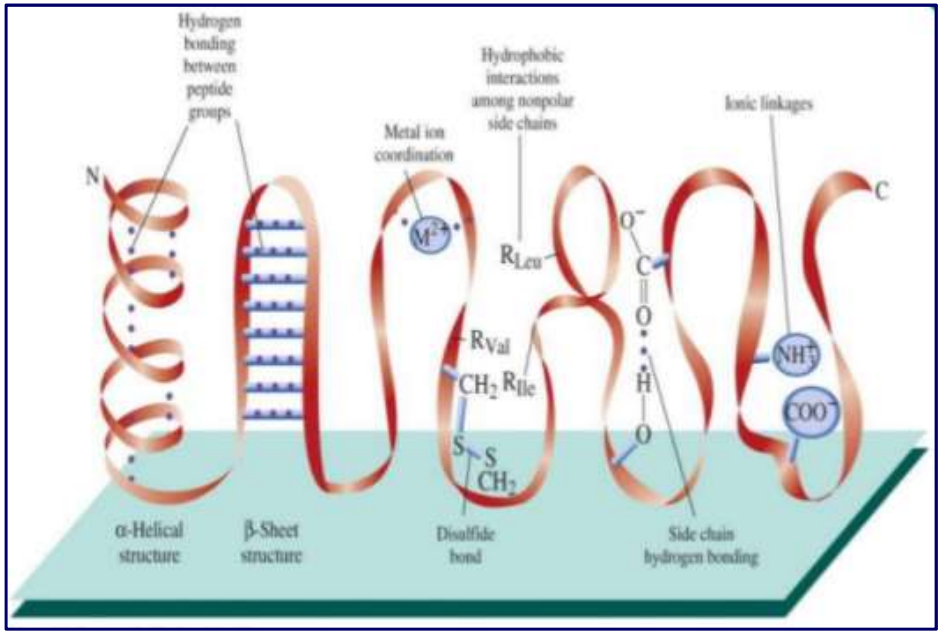
1. Present mainly in globular protein.
2. Connect two Alpha helix or Beta sheath.
3. Present in those area where bend is required.



Tertiary Structure

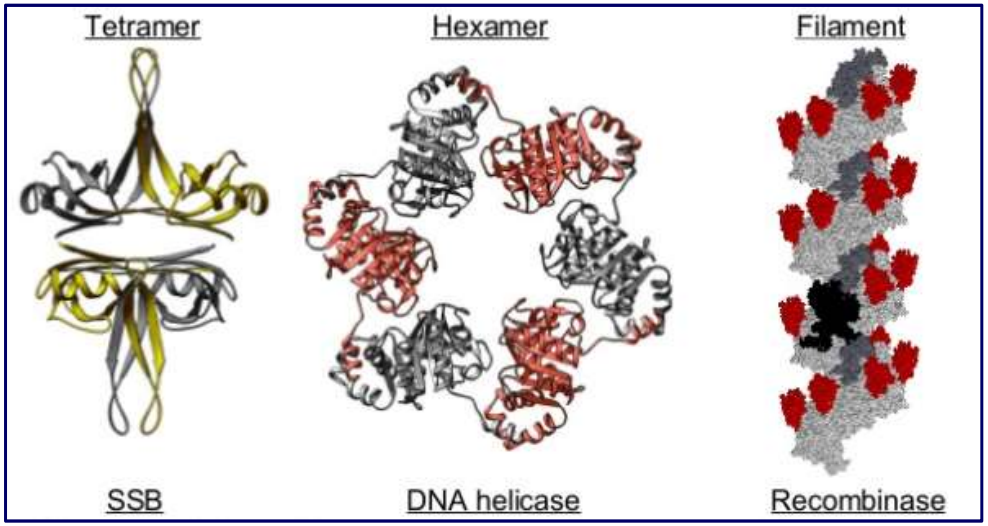
- **Combination of several motifs** of secondary structures into a compact arrangement

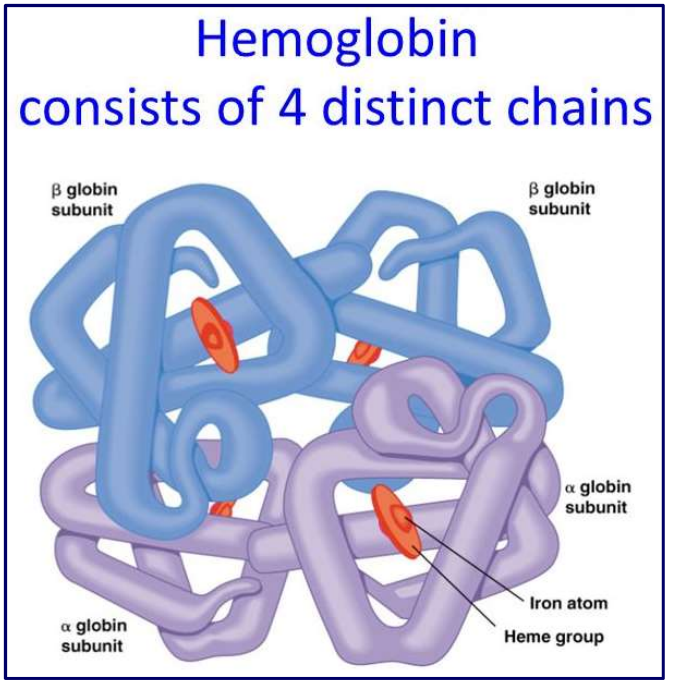
- **Noncovalent forces** bring about the interactions and stability;
  - H-bonds,
  - electrostatic,
  - hydrophobic,
  - Van Der Waal's,
  - pi-pi complexation between R-side chains
  - Disulfide bonds occur between Cys residues



Quaternary Structure

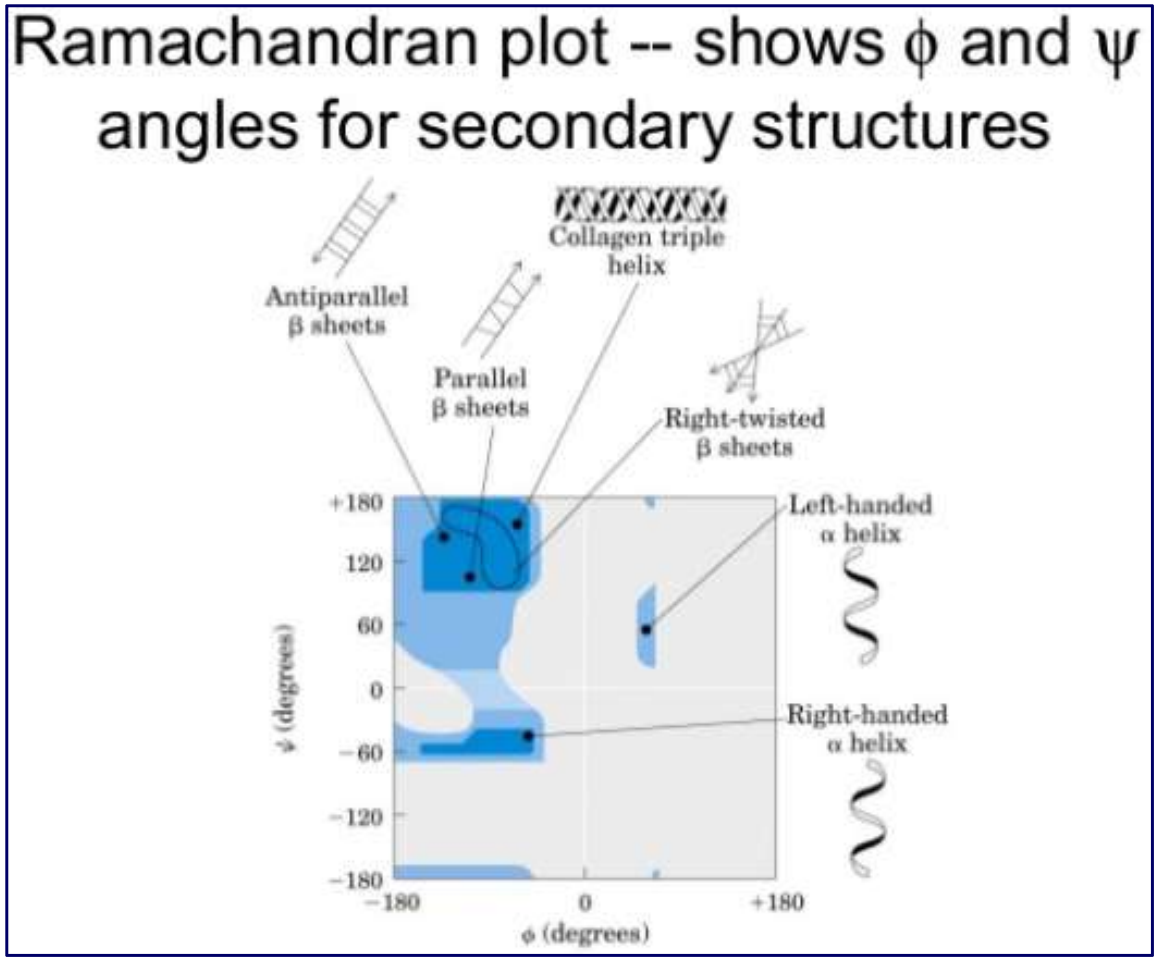
- **Oligomeric** –two or more polypeptide chains; subunits
- **Homotypic** – almost identical subunits
- **Heterotypic** – different subunits
- Defines the arrangement and position of each subunit in an intact protein



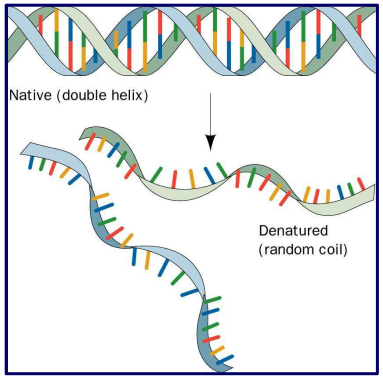


Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Ramachandran plot -- shows ϕ and ψ angles for secondary structures



Protein Denaturation



Denaturation *is a process in which proteins lose their highly organized quaternary and tertiary structure causing their 3-dimensional shape to be lost due to unfolding, by disruption/ destruction of both the 2^o and 3^o structures.*

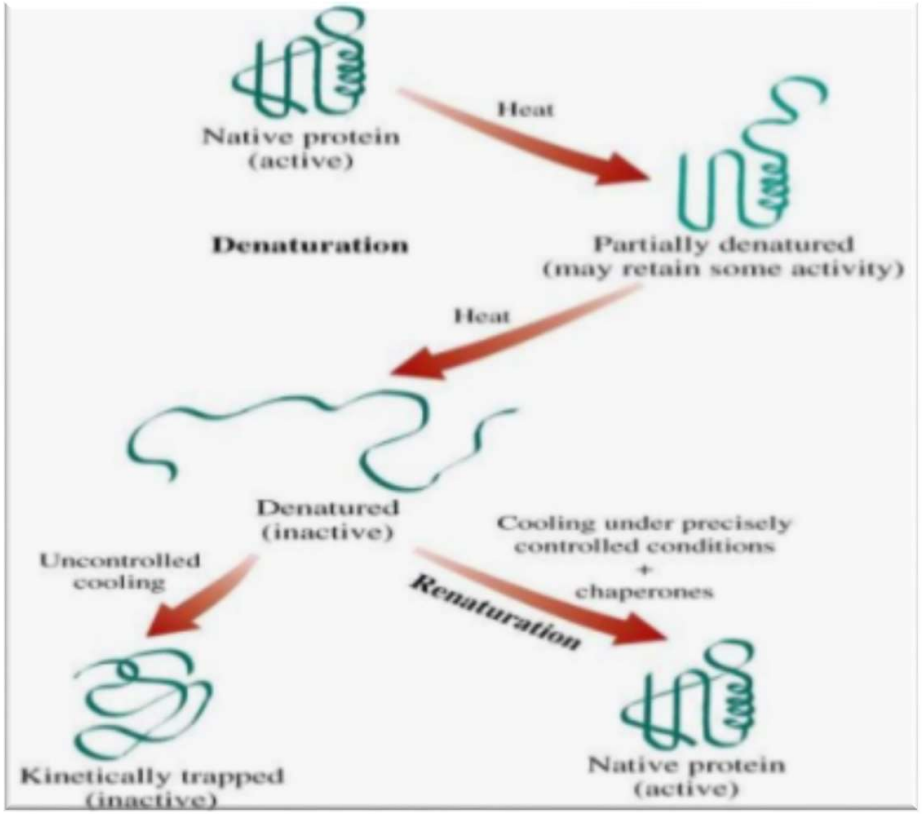
Denaturation is not strong enough to break the peptide bonds, thus 1^o structure or the sequence of amino acids is retained post-denaturation.

Denaturation unfolds the α-helices and β-sheets of a protein into a random shape.

Denaturants (pH, salt, concentration, temperature etc.) disrupt bonds and forces of attraction

Causes for denaturation [denaturants] are*:*

- *Change in pH [breaking of H-bonds & electrostatic attractions]*
- *Some reagents such as guanidine hydrochloride and urea form H-bonds to protein groups which are stronger than intramolecular H-bonds.*
- *Detergents like sodium dodecyl sulfate & organic solvents associate with nonpolar groups of the protein thus impairing hydrophobic interactions.*
- *Heating or vigorous shaking which disrupts attractive interactions.*



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

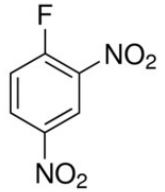
Determination of Amino acid sequence

Short Polypeptides are Sequenced with Automated procedures

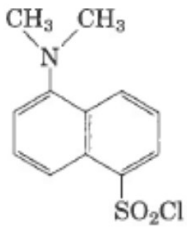
It involves the labelling and identifying of the amino-terminal amino acid residue.

Various reagents used are:

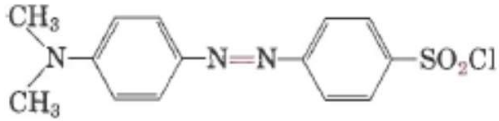
1. *1-fluoro-2,4-dinitrobenzene [FDNB]*
   [Sanger's reagent]



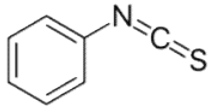
2. *Dansyl chloride*
   1-dimethylaminonaphthalene-5-sulfonyl chloride



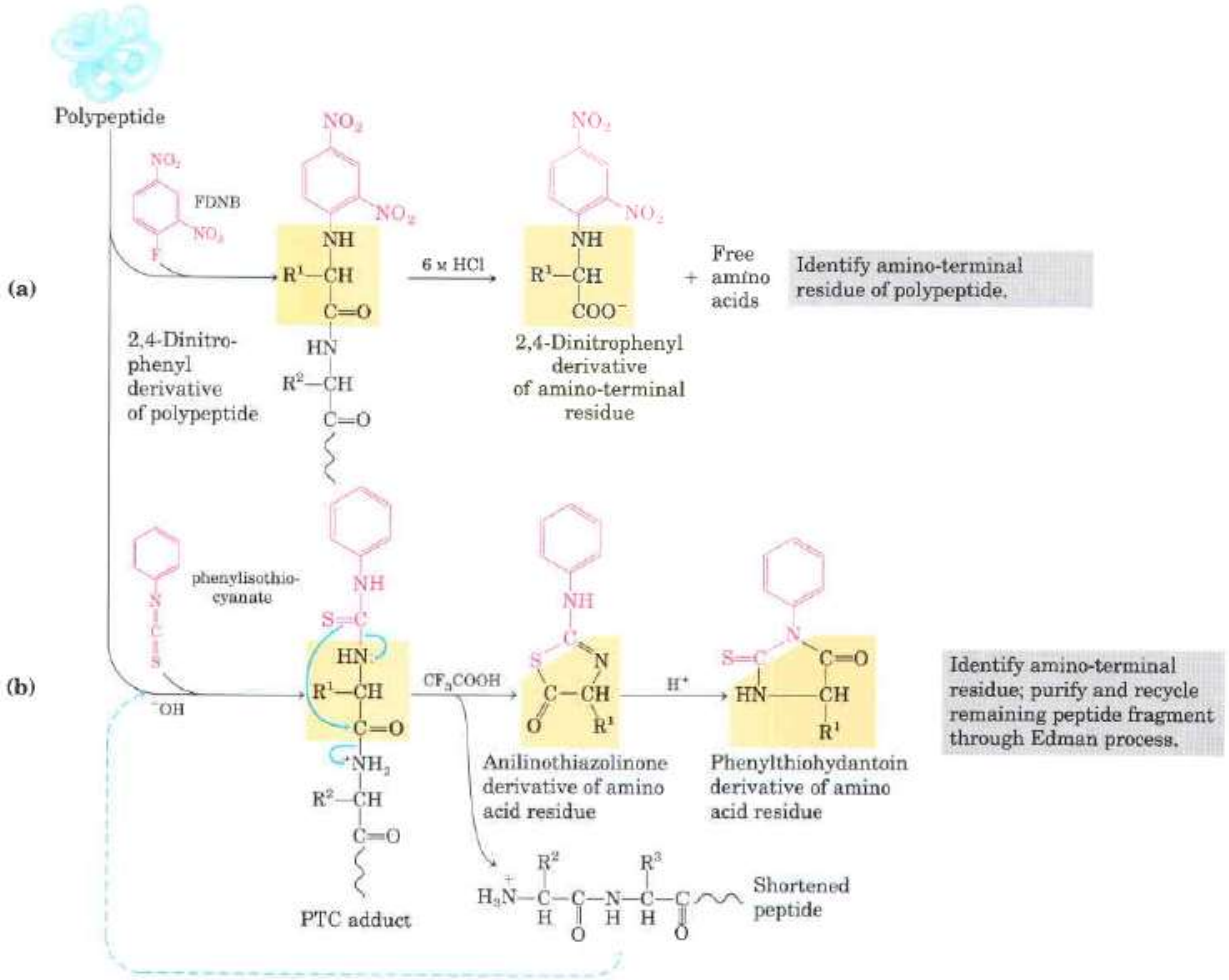
3. *Dabsyl chloride*
   (4-N,N-dimethylaminoazobenzene-4'-sulfonyl chloride)



4. *Phenylisothiocyannate [PTC]*
   [Edman's reagent]

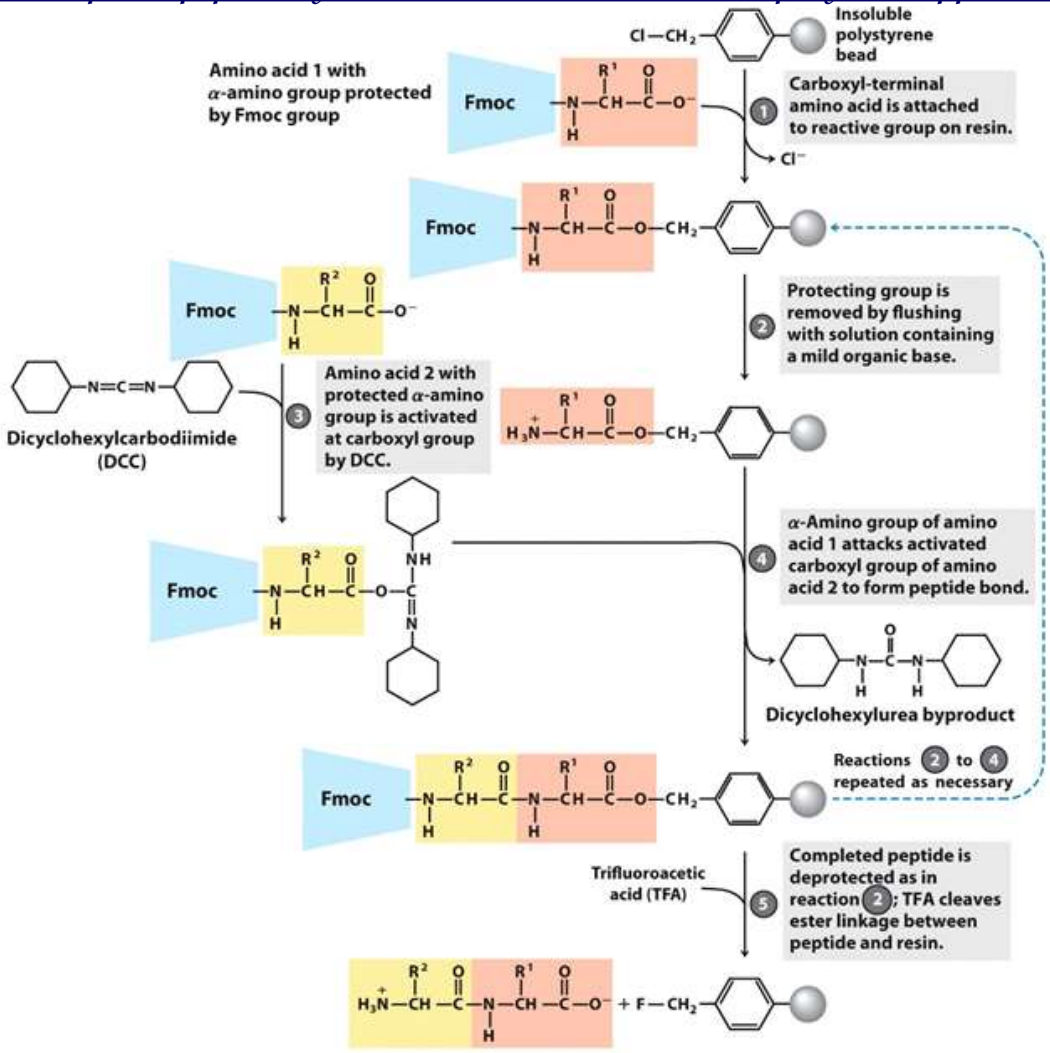


Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati



(a) Sanger's method of amino-terminal identification. (b) Edman degradation

Automated chemical solid phase peptide synthesis [SPPS]on an insoluble polymer support [Merrifield synthesizer]



Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science &Mathematics
IIITG, Guwahati

Reagents for Peptide or Protein cleavage

Reagent	***Specificity***
Chemical	
Edman's reagent	*removes N-terminal amino acid*
Cyanogen bromide	*hydrolyzes on C-side of Met*
Exopeptidases	
Carboxypeptidase A	*removes C-terminal amino acid (not Arg/Lys)*
Carboxypeptidase B	*removes C-terminal amino acid (only Arg/Lys)*
Endopeptidases	
Trypsin	*hydrolyzes on C-side of Arg and Lys*
Chymotrypsin	*hydrolyzes on C-side of aromatic 6-membd R-gps*
Elastase	*hydrolyzes on C-side of small amino acids (Gly, Ala, Val, Ser)*
Thermolysin	*hydrolyzes on C-side of Ile, Met, Phe, Trp, Tyr, Val*

However, in case of cysteine, if the sulfur in cysteine attacked cyanogen bromide, the bromide ion would deprotonate the cyanide adduct, leaving sulfur uncharged and the β-carbon of cysteine not electrophilic. The strongest electrophile would be the cyanide nitrogen, which, if attacked by water, would yield cyanic acid and the original cysteine.

Purification of proteins

Reverse phase (RP HPLC): high-pressure pumps that speed the movement of the protein molecules down the column

Column chromatography: takes advantage of differences in protein charge, size, binding affinity, and other properties (most powerful)

Affinity chromatography: based on binding affinity of protein to the ligand

Ion-exchange chromatography: exploits differences in the sign and magnitude of the net electric charge of proteins at a given pH

Size-exclusion chromatography (also called gel filtration separates proteins according to size),

Dialysis: is a procedure that separates proteins from small solutes by taking advantage of the proteins' larger size.

2D electrophoresis: is separation of proteins based on migration of charged proteins in an electric field.

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Enzymes

Enzymes*: are globular proteins that catalyze the biochemical reactions that occur in living systems. With the exception of a small group of catalytic RNA molecules, all enzymes are proteins. Enzymes, like other proteins, have molecular weights ranging from about 12,000 to more than 1 million. Enzymes enhance reaction rates by a factor of* 10^5 *to* 10^{17}.

Some enzymes require no chemical groups for activity other than their amino acid residues. Others require an additional chemical component called a cofactor-inorganic ions, such as Fe^{2+}, Mg^{2+}, *or a complex organic or metalloorganic molecule called a* ***coenzyme****. A coenzyme or metal ion that is very tightly or even covalently bound to the enzyme protein is called a prosthetic group.*

A complete, catalytically active enzyme together with its bound coenzyme and/or metal ions is called a ***holoenzyme****. The protein part of such an enzyme is called the* ***apoenzyme or apoprotein****. Enzymes are extremely specific in their catalytic actions.*

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

*The distinguishing feature of an enzyme-catalyzed reaction is that it takes place within the confines of a pocket on the enzyme called the **active site**. The molecule that is bound in the active site and acted upon by the enzyme is called the **substrate**.*

IUB [International Union of Biochemistry] enzyme classification:

a) Oxidoreductases

b) Transferases

c) Hydrolases

d) Lyases

e) Isomerases

f) Ligases.

Measurement and expression of enzyme activity:

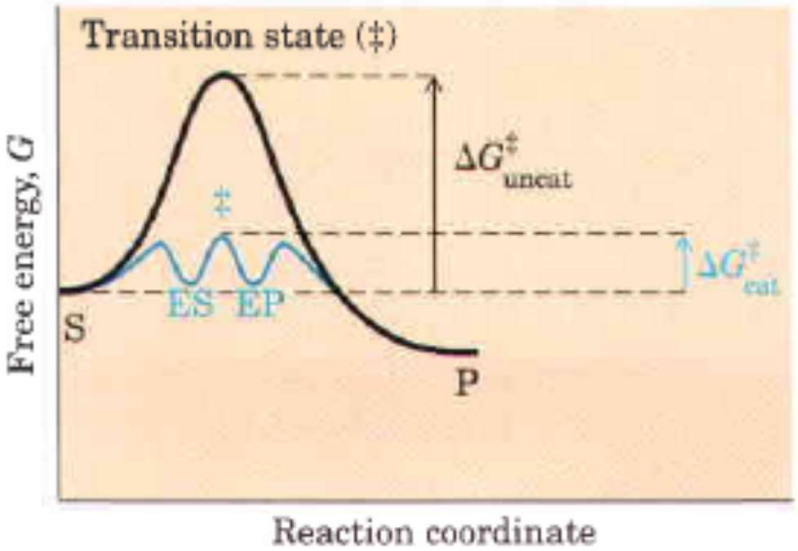
*Katal: One **katal** is the amount of **enzyme** that is needed to transform 1 mole of substrate per second.*

Enzymes are classified by the reactions they catalyze

Class	***Chemical Reaction Catalyzed***	***Sample Enzymes***
Oxidoreductase	*Oxidation-reduction in which oxygen and hydrogen are gained or lost*	*Cytochrome oxidase, lactate dehydrogenase*
Transferase	*Transfer of functional groups, such as an amino group, acetyl group, or phosphate group*	*Acetate kinase, alanine deaminase*
Hydrolase	*Hydrolysis (addition of water)*	*Lipase, sucrase*
Lyase	*Removal of groups of atoms without hydrolysis*	*Oxalate decarboxylase, isocitrate lyase*
Isomerase	*Rearrangement of atoms within a molecule*	*Glucose-phosphate isomerase, alanine racemase*
Ligase	*Joining of two molecules (using energy usually derived from the breakdown of ATP)*	*Acetyl-CoA synthetase, DNA ligase*

An enzyme does not change the equilibrium constant of the reaction; it only changes the rate at which the product is formed. An enzymatic reaction can be written as $E + S \leftrightarrow ES \leftrightarrow EP \leftrightarrow E + P$
E: Enzyme, S: Substrate, P: Product, ES: Transient complexes of the enzyme with substrate, EP: Transient complexes of the enzyme with product

Energy in biological systems is described in terms of ***free energy, G****. In the coordinate diagram, the free energy of the system is plotted against the progress of the reaction. The starting point any reaction is called its ground state. There is an energy barrier between S and P, comprising of energy required for alignment of reacting groups, formation of transient unstable charges, bond rearrangements and other transformations that are required for the reaction to proceed. To overcome this barrier, the molecules must be raised to a higher energy level, called the transition state. The difference between the energy levels of the ground state and the transition state is the* ***activation energy,*** $\Delta G^{\ddagger}$*. A higher activation energy corresponds to a slower reaction. Enzymes accelerate reaction rates by lowering activation energies of biochemical reactions.*



Comparison of enzyme-catalyzed and uncatalyzed reactions

The equilibrium constant, $K_{eq} = [P]/[S]$

The standard free-energy change for the reaction, $\Delta G'^{o}$ *is linked to the activation energy,* $\Delta G^{\ddagger}$

From thermodynamics, the relationship between K_{eq} *and* $\Delta G'^{o}$ *can be deduced as*

$\Delta G'^{o} = -RT \ln K_{eq}$

At transition state, $k = \frac{kT}{h} e^{-\Delta G^{\ddagger}/RT}$

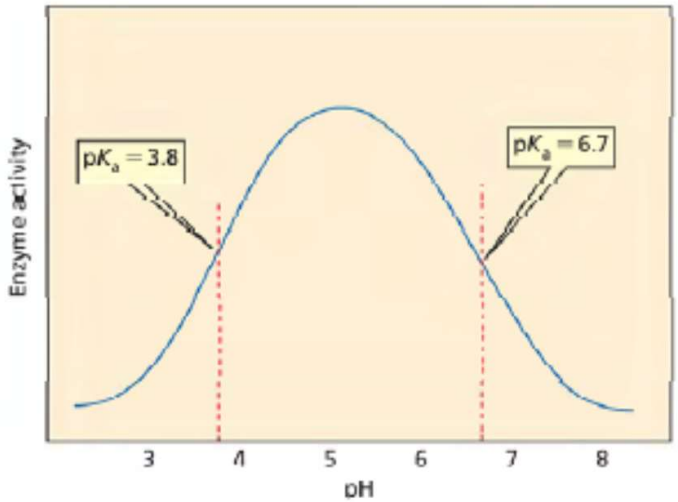
where k*: the Boltzmann constant and h: Planck's constant, T: absolute temperature, R: gas constant, 8.315 J /mol. K,* k*: rate constant*

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science & Mathematics
IIITG, Guwahati

Factors related to the catalytic activity of enzymes are:

1. *Reacting groups are brought together at active site/pocket of the enzyme in proper orientation*
2. *Amino acid side chains are well oriented relative to the substrate.*

Activity of an enzyme as a function of pH of the reaction mixture is called the pH-activity profile or pH-rate profile.

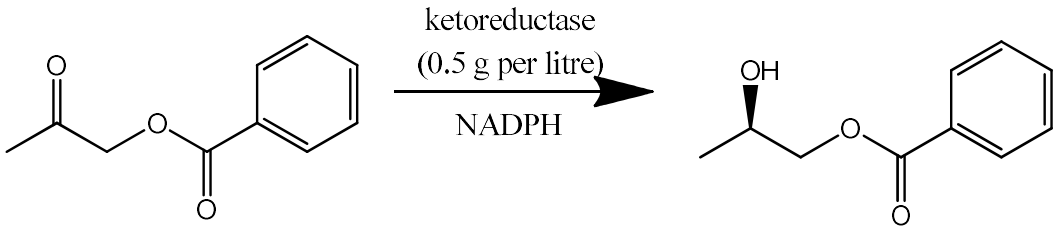


E.g. Activity of lysozyme is highest at pH = 5.3, 50% at pH = 3.8 & pH = 6.7.

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science &Mathematics
IIITG, Guwahati

Applications:

*Enzymes act as catalysts for asymmetric reactions, both **in vivo** as well as **in vitro**.*
In biocatalysis, enzymes are used to catalyze the synthesis of enantiomerically pure molecules.



91% yield; >98%ee

Course Code: **SC202**
Presented by
Course Instructor: **Dr Bhar Saha**
Dept. of Science &Mathematics
IIITG, Guwahati