| Course Content Developed by | Rakesh Singh<br>Associate Professor in Chemistry , Govt. Degree College, Kathua |
|---|---|
| Department | Department of Higher Education UT of Jammu and Kashmir |
| Paper | Organic Chemistry |
| Class | Semester 3rd (UG-CBCS, General) |
| University/College | Jammu University, Cluster University, Autonomous College- GCW Parade |
| Syllabus | *Determination of Primary structure of Peptides by degradation Edmann degradation (N-terminal) and C–terminal (thiohydantoin and with carboxypeptidase enzyme). Synthesis of simple peptides (upto dipeptides) by N-protection (t-butyloxycarbonyl and phthaloyl) & C-activating groups and Merrifield solid-phase synthesis.* ***Note :*** *Overview of Primary, Secondary, Tertiary and Quaternary Structures of Proteins is in syllabus of Cluster University and GCW Parade.* |

# Unit : Peptides and Proteins

## Topic 1 : Determination of Primary structure of Peptides

## Topic 2 : Methods of synthesis of Peptides

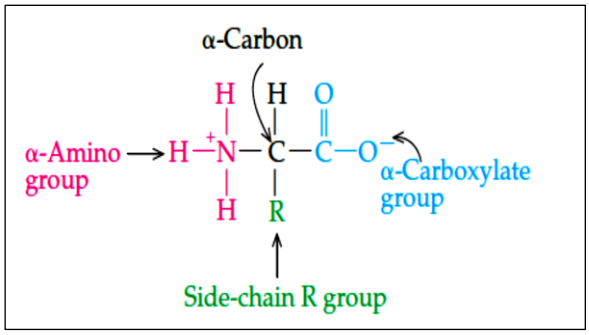
### Learning Objectives

### To Understand

- what are peptides
- Brief Classification of Peptides
- Nomenclature of Peptides
- Determination of Primary Structutre of Protein
- Methods of synthesis of Peptides
- Problems for Practice

## 1.1 Peptides and Proteins

Peptides and Proteins are fundamental components of cells which perform important biological functions. Structurally, proteins and peptides are quite similar, as both are made up of chains of α-amino acids that are held together by peptide bonds (also called amide bonds). The basic distinguishing factors are size and structure. Peptides are smaller than proteins. Traditionally, peptides are defined as molecules that consist of between 2 and 50 amino acids, whereas proteins are made up of 50 or more amino acids. In addition, peptides tend to be less well defined in structure than proteins, which can adopt complex conformations known as secondary, tertiary, and quaternary structures. All proteins are polypeptides but not all polypeptides are proteins.




Structure of α- amino acid

**1.2 Peptides:**

Peptides are defined as condensation products of two or more amino acids formed by reaction between amino group of one amino acid molecule with carboxylic group of other amino acid molecule. In peptide, amino acids are linked together by amide bonds. The amide bond between the amino group of one amino acid and the carboxyl of another is called a peptide bond. For example , when two molecules of glycine combine , it results in formation of dipeptide along with release of water as condensation by-product.

| | | | | | |
|---|---|---|---|---|---|
| $NH_2CH_2COO^-$ | + | $NH_2CH_2COO^-$ | ⟶ | $H_3N^+CH_2CONH\ CH_2COO^-$ | + $H_2O$ |
| Glycine | | Glycine | | Glycylglycine(Dipeptide) | + Water |

The amide linkage $-\overset{\overset{O}{\|}}{C}-NH-$ is called peptide bond or peptide linkage.

**1.3 Classification of Peptides:**

Depending upon number of amino acids involved in peptide formation, peptides have been classified into following types :

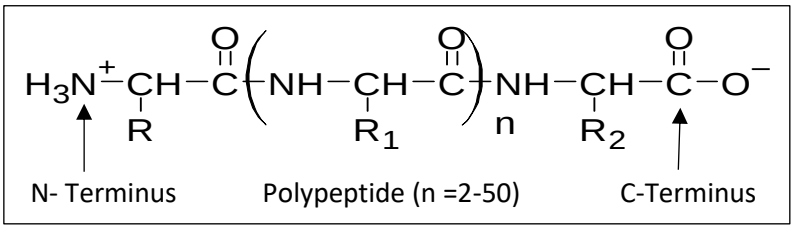
**a) Dipeptides :** A peptide formed by condensation of two amino acids is called dipeptide.

**b) Tripeptide :** A peptide formed by condensation of three amino acids is called tripeptide.

**c) Tetrapeptide :** A peptide formed by condensation of four amino acids is called tetrapeptide.

**d) Polypeptide :** A peptide formed by condensation of more than four amino acids is called polypeptide.

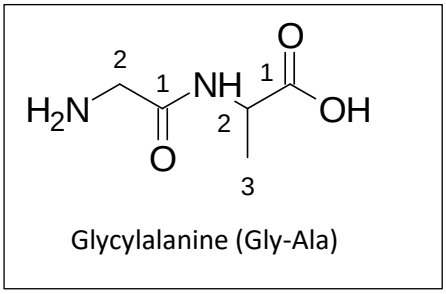
For example :



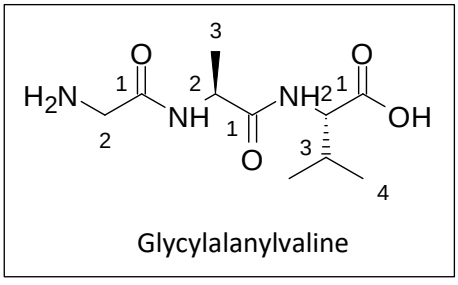

- Peptides possess a free $H_3N^+$-- group on one end and a free $-COO^-$ group at the other end.
- The amino acid having free $H_3N^+$-- group is called N-terminal amino acid residue
- The amino acid having free $-COO^-$ group is called C-terminal amino acid residue.
- While writing formula of peptide, we proceed from left with the N-terminal amino acid residue to the right with C-terminal amino acid residue. So it is a convention to write N-terminal amino acid on left hand side and C-terminal amino acid of right hand side.

**1.4 Nomenclature of Peptides:**

The names of amino acids which constitute a peptide are written from N-terminal i.e. starting from L.H.S to C-terminal i.e. R.H.S. While writing the names , the suffix "ine" of all amino acids, except the C-terminal amino acids is replaced by "yl" . Sometimes three letter abbreviations of amino acids are also used for writing name of polypeptide.



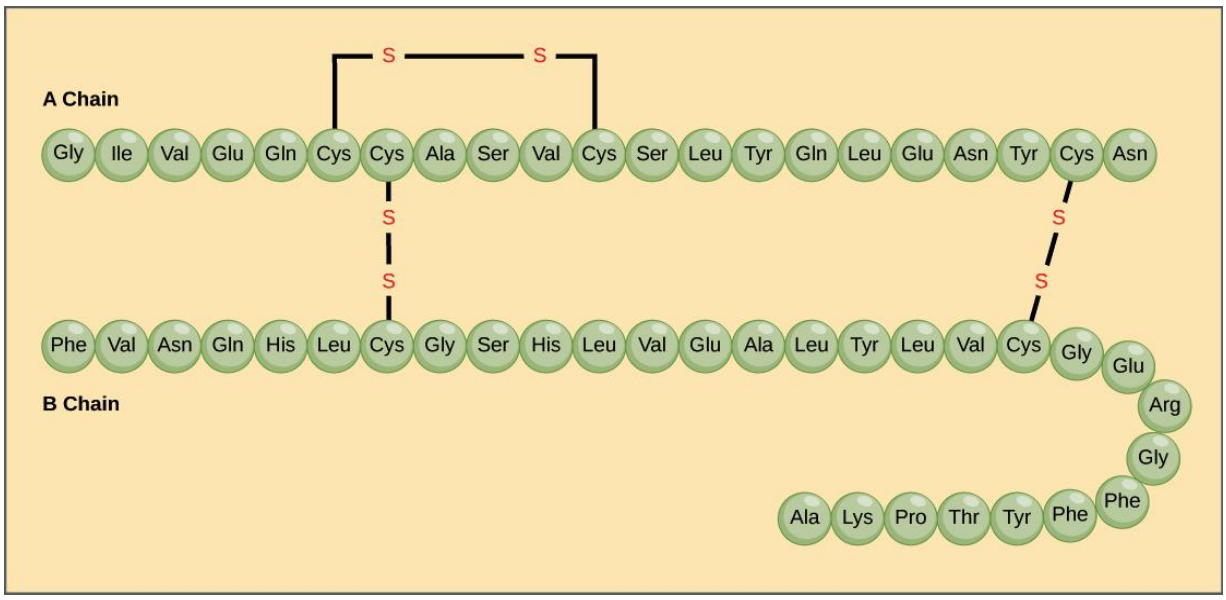
Glycylalanine (Gly-Ala)



Glycylalanylvaline

## 1.5 Overview of Primary, Secondary, Tertiary and Quaternary Structures of Proteins

**1.5.1 Primary Structure :** It refers to the sequence in which various amino acids are present in a polypeptide or protein are linked to each other. For example, the hormone insulin has two polypeptide chains, A and B, shown in diagram below (image credit :*OpenStax Biology*)




**Importance of Primary Structure:**

It must be noted that sequence of amino acid is very important because change of just one amino acid in a protein's sequence can affect the protein's overall structure and biological functions. e.g. change of just one amino acid in polypeptide sequence of haemoglobin causes a disease called sickle cell anaemia. In sickle cell anemia, one of the polypeptide chains that make up hemoglobin, the protein that carries oxygen in the blood, has a slight sequence change. The glutamic acid that is normally the sixth amino acid of the hemoglobin β chain (one of two types of protein chains that make up hemoglobin) is replaced by a valine. What is most remarkable to consider is that a hemoglobin molecule is made up of two α chains and two β chains, each consisting of about 150 amino acids, for a total of about 600 amino acids in the whole protein. The difference between a normal hemoglobin molecule and a sickle cell molecule is just 2 amino acids out of the approximately 600.

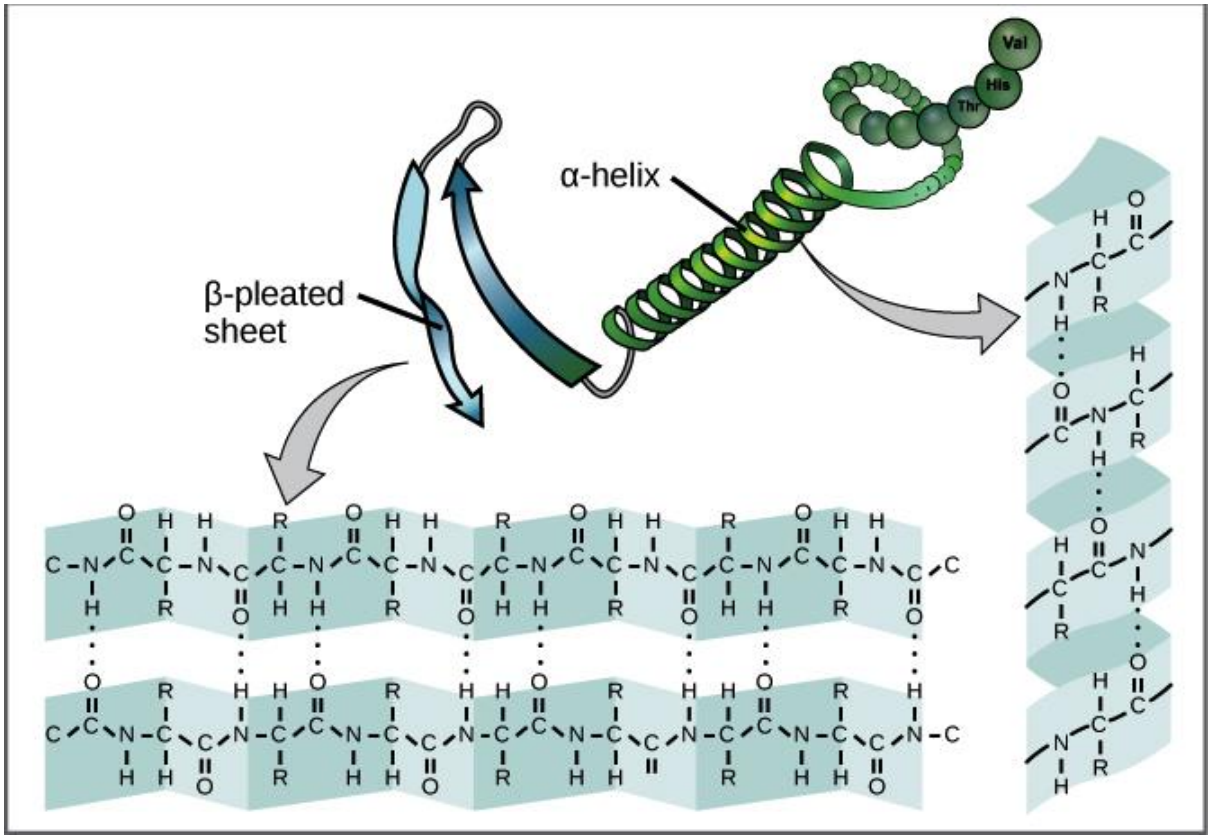
**1.5.2 Secondary structure:**

Secondary structure refers to the conformation, which a polypeptide chain assumes due to hydrogen bonding. Depending upon size of side chain R- attached to alpha carbon, the proteins show two types of secondary structure as :

**i) α-helix (by Linus Pauling in 1951):** If size of R-group is quite large , then proteins possess α-helix structure at secondary level. Alpha helix structure arises due to intramolecular H-bonding between , the carbonyl (C=O) of one amino acid residue to the N-H of forth amino acid residue. This causes the polypeptide chain to coil up into a spiral structure (helical structure) that resembles a curled ribbon, with each turn of the helix containing 3.6 amino acids. The R groups of the amino acids point outward from the α-helix, where they are free to interact. The α-helix may be right handed or left handed. However, Moffitt in 1956 , theoretically proved that for L-amino acid , right handed helix is more stable. Hence, proteins always possess right handed helical structure. i.e. polypeptide chain turns clockwise to form α-helix.

**ii) β-pleated sheet structure (Linus Pauling in 1951):** If size of R-groups is small or of moderate size, then In a β-pleated sheet structure, two or more segments of a polypeptide chain line up side by side, forming a sheet-like structure held together by hydrogen bonds, between carbonyl groups of chain and amino groups of other chain, while the R groups extend above and below the plane of the sheet. The strands of a β pleated sheet may be parallel, pointing in the same direction (meaning that their N- and C-termini match up), or antiparallel, pointing in opposite directions (meaning that the N-terminus of one strand is positioned next to the C-terminus of the other).

**Many proteins contain both α helices and β pleated sheets, though some contain just one type of secondary structure.**




*(Image credit: OpenStax Biology)*

**1.5.3 Tertiary structure :**

The overall three-dimensional structure of a polypeptide is called its tertiary structure. At normal $p^H$ and temperature, each protein molecule acquires tertiary structure which is most stable. For example , fibrous proteins which have almost same secondary structure through out their length possess rod like or rope like as its tertiary structure. (e.g. α-keratin, the major protein of hair and wool). On the other hand, globular proteins, don't have same secondary structure through out their length. Some parts may have α-helix, some other parts may have β-pleated sheet, whereas some other parts may have random coils. In such cases, the entire protein molecule may fold up to give spherical shape to protein molecule. So, tertiary structure of globular protein is spherical.

The tertiary structure is primarily due to various interactions like hydrogen bonding, ionic bonding, dipole-dipole interactions, and London dispersion forces, due to which protein molecule acquires a particular three dimensional shape. Also important to tertiary structure are hydrophobic interactions, between nonpolar, hydrophobic R groups. In addition the disulfide bond, i.e. covalent linkages between the sulfur-containing side chains of cysteines, are much stronger than the other types of bonds that contribute to tertiary structure. They act like molecular "safety pins," keeping parts of the polypeptide firmly attached to one another**.**

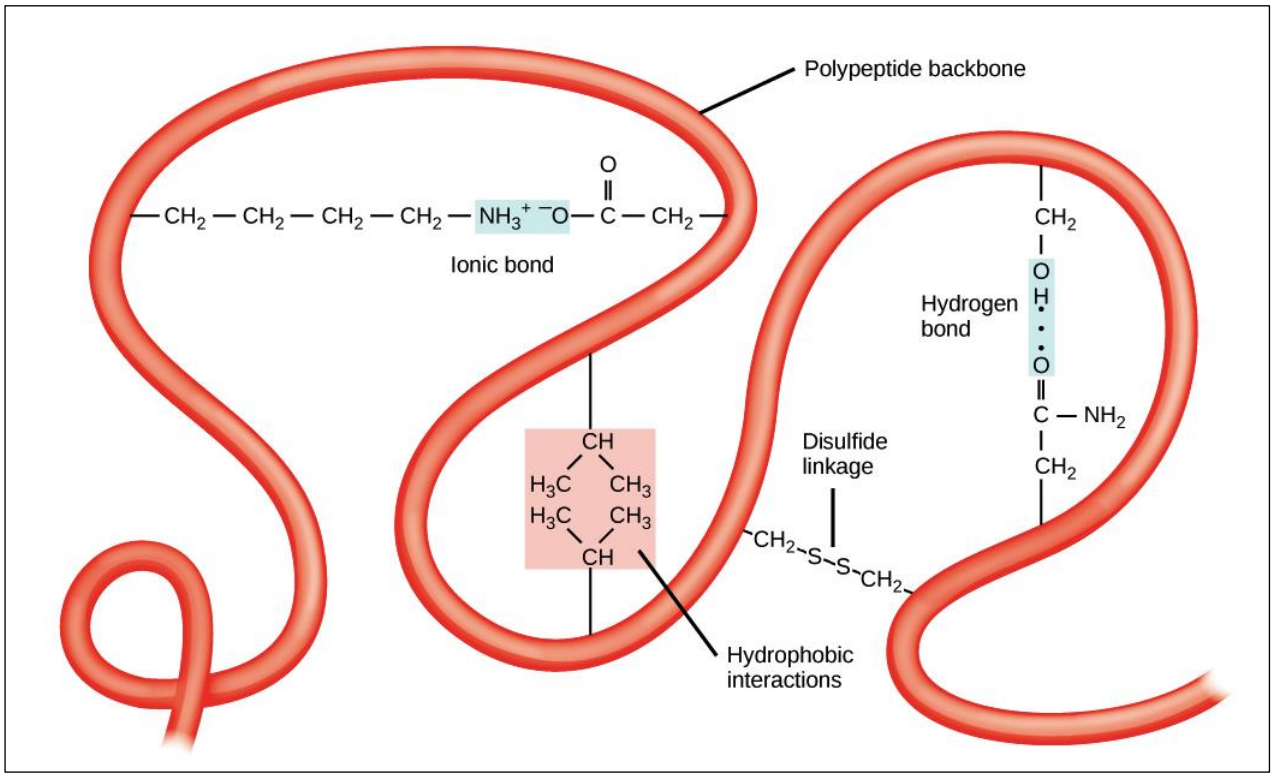
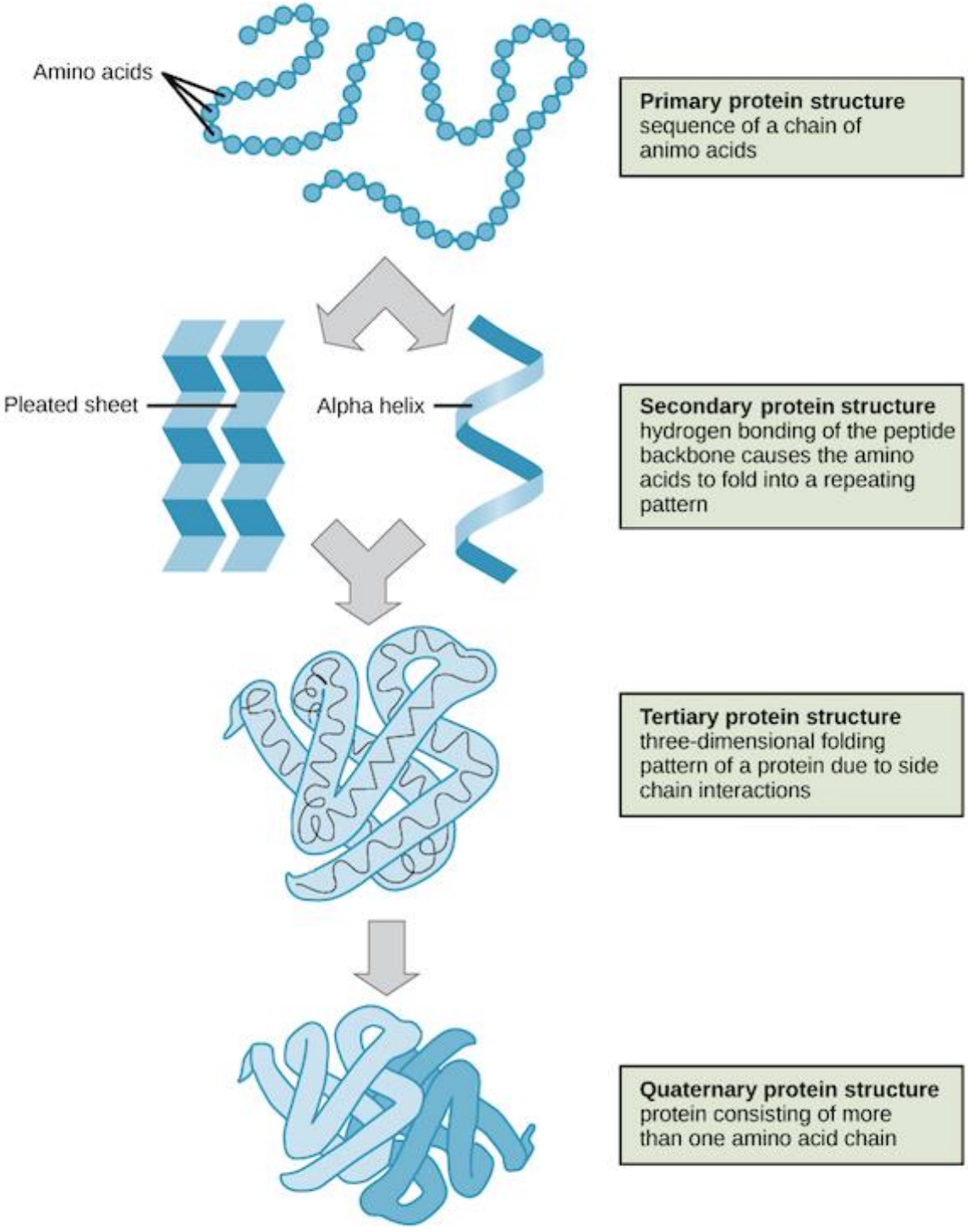



**Image credit OpenStax Biology**

### 1.5.4 Quaternary structure:

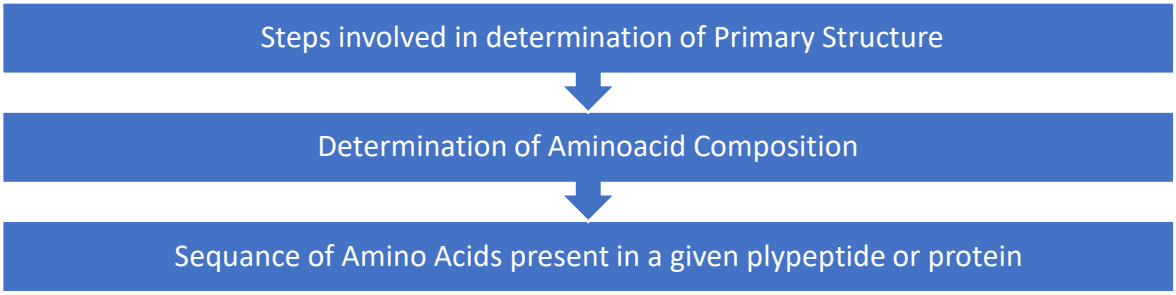
Many proteins are made up of a single polypeptide chain and have only three levels of structure i.e. Primary, secondary and tertiary. However, some proteins are made up of multiple polypeptide chains, also known as subunits. When these subunits come together, they give the protein its quaternary structure. For example, haemoglobin, which carries oxygen in the blood and is made up of four polypeptide chains or subunits [two each of the α (each chain with 141 amino acids )and β (each chain with 146 amino acid residues ) types]. Another example is DNA polymerase, an enzyme that synthesizes new strands of DNA and is composed of ten subunits -

In general, the same types of interactions that contribute to tertiary structure (mostly weak interactions, such as hydrogen bonding and London dispersion forces) also hold the subunits together to give quaternary structure. So quaternary structure of protein refers to determination of number of subunits and their overall arrangement in aggregate protein molecule as a whole. For example in haemoglobin , the four subunits lie at the vertices of regular tetrahedron.




***Credit :*** *Image modified from OpenStax Biology's modification of work by the National Human Genome Research Institute.*

## 1.6 Determination of Primary Structure of Peptide




**1.6.1 Step 1: Determination of Amino acid Composition :**

- The given polypeptide/protein is completely hydrolysed to its constituent amino acids.
- Hydrolysis is preferably done by 6N HCl at 373-393K or by enzymes.
- However, alkaline hydrolysis is not preferred because it causes racemization.
- The mixture of amino acids thus obtained is separated and individually identified by either ion-exchange or gas chromatography.
- The weights of each of the amino acid is noted .
- From their weights, the number of moles of each of the amino acid is determined.
- Hence number of each type of amino acid present in given polypeptide or protein is then calculated.

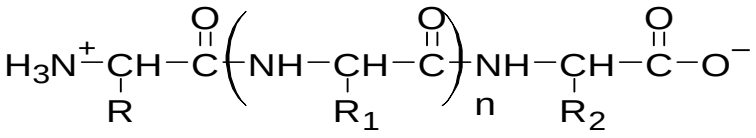
However, now a days the whole process is automated and is done with the help of an instrument called amino acid analyser.

It requires very small amount of peptide ($10^{-5}$ to $10^{-7}$ g)

**1.6.2 Step 2 : Sequencing of amino acids present in given polypeptide or protein :**

After determination of amino acids composition (i.e. number of each type of amino acid present in polypeptide), the next and the most important step is to determine sequence of amino acids in given polypeptide/protein. This is done by terminal residue analysis and partial hydrolysis.

**1.6.3 Terminal Residues analysis :** The terminal amino acid residue written on extreme left of polypeptide chain possesses free amino group and is called N-terminal amino acid, where as the amino acid residue present on extreme right of polypeptide chain possesses free -COOH group and is called C-terminal amino acid.



N- Terminus Polypeptide (n =2-50) C-Terminus

The whole process of terminal residue analysis involves treating a given polypeptide with a suitable reagent, which selectively reacts with either N-terminal amino acid or C-terminal amino acid. As a result the N-terminal or C-terminal amino acid gets labelled which is then selectively removed by partial hydrolysis of polypeptide and hence identified. The resulting degraded peptide is then again subjected to same treatment and one by one , the whole sequence of amino acids in the given polypeptide is determined.

The sequencing of amino acids can be done by following two methods:

**a) N-Terminal residue analysis b) C-terminal residue analysis**

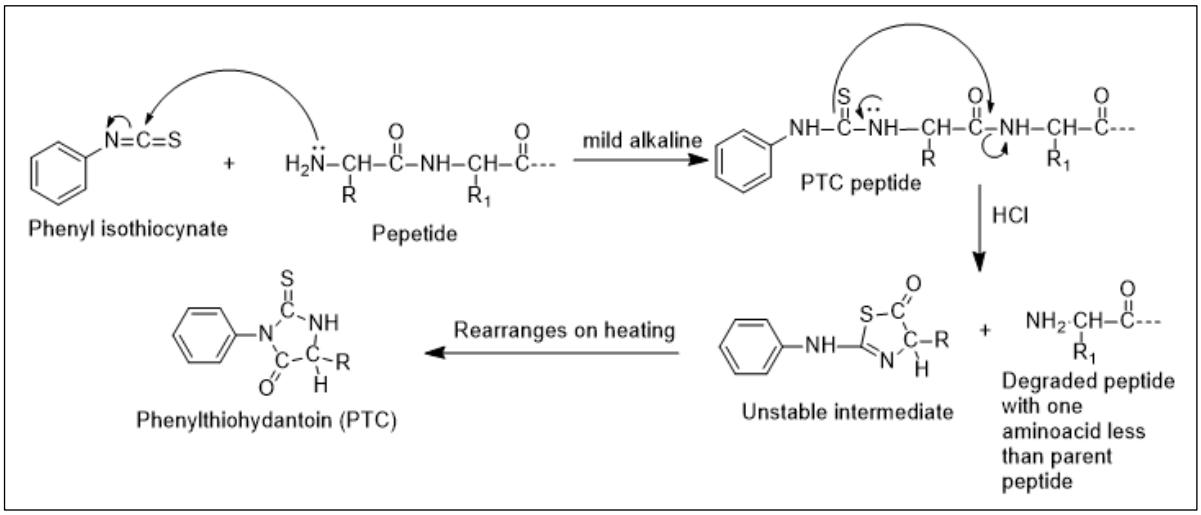
**a) N-terminal residue analysis :**

It is done by following methods:

**(i) Edman's method (Edman's degradation)**

This method was developed by Pehr Edman. The Edman degradation reaction was automated in 1967 by Edman and Beggs. In this process

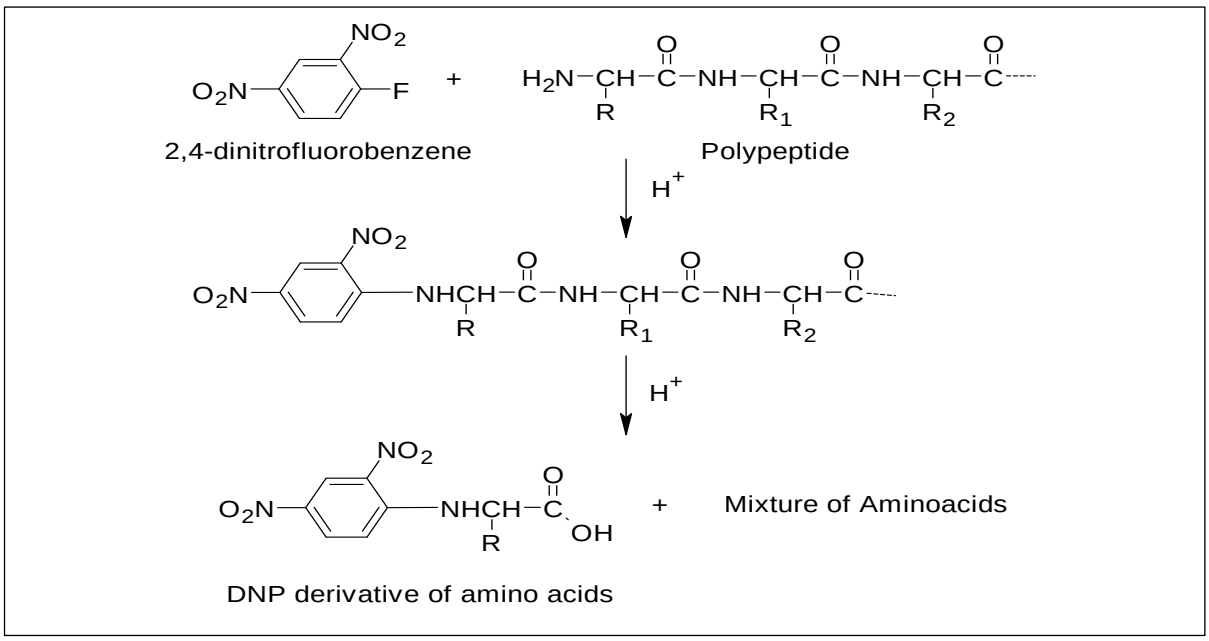
- The given peptide is reacted with phenyl isothiocyanate (PITC) under mildly alkaline conditions when the , $-NH_2$ group of N-terminus of polypeptide reacts to give a phenyl thiocarbamoyl derivative (PTC-peptide).
- The PTC-peptide upon mild hydrolysis with HCl, selectively removes, the N-terminal amino acid as phenyl thiohydantoin (PTH), while rest of polypeptide chain remains intact.
- The PTH so obtained is then separated from reaction mixture and then identified chromatographically. This is done by comparing with PTC obtained from known amino acids. Thus, the N-terminal amino acid get identified.
- The degraded peptide is then subjected to same process again and in this way the whole sequence of amino acids is determined.




**Note :** Theoretically this process can be can be repeated over and over again till whole sequence is determined, but practically , it is possible only up to 20 amino acids. However, nowadays, the automated Edman degradation (the protein sequenator) is used widely, and it can sequence peptides up to 50 amino acids.

**(ii) Sanger's method or DNFB method :** In this method

- The given peptide is treated with 2,4-dinitrofluorobenzene (DNFB), commonly known as Sanger's reagent.
- DNFB is very reactive in nucleophilic displacements with amines but not amides.
- The reagent reacts with free amino group of the N-terminal amino acid of peptide to form N-2,4-dinitrophenyl (DNP) derivative.
- The DNP derivative of polypeptide upon hydrolysis with dilute HCl gives 2,4-dinitrophenyl (DNP) derivative of N-terminal amino acid along with mixture of amino acids.
- The 2,4-dinitrophenyl (DNP) derivative of N-terminal amino acid is isolated and identified chromatographically.



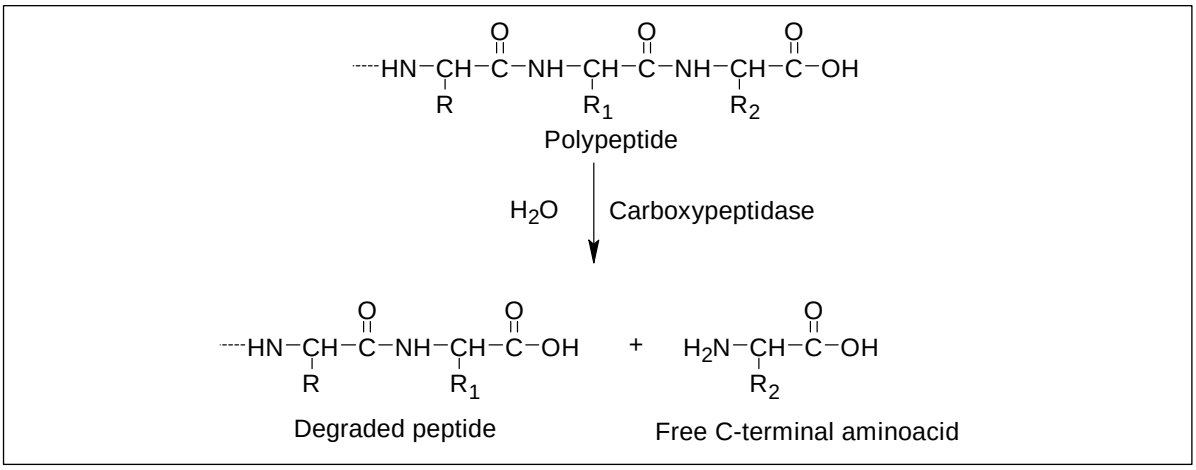

**Demerits :** By this method, only terminal amino acids can be identified.

**b) C-terminal residue analysis :**

The C-terminal residue is determined by the use of either a chemical reagent or the enzyme carboxypeptidase. The two commonly used methods are as :

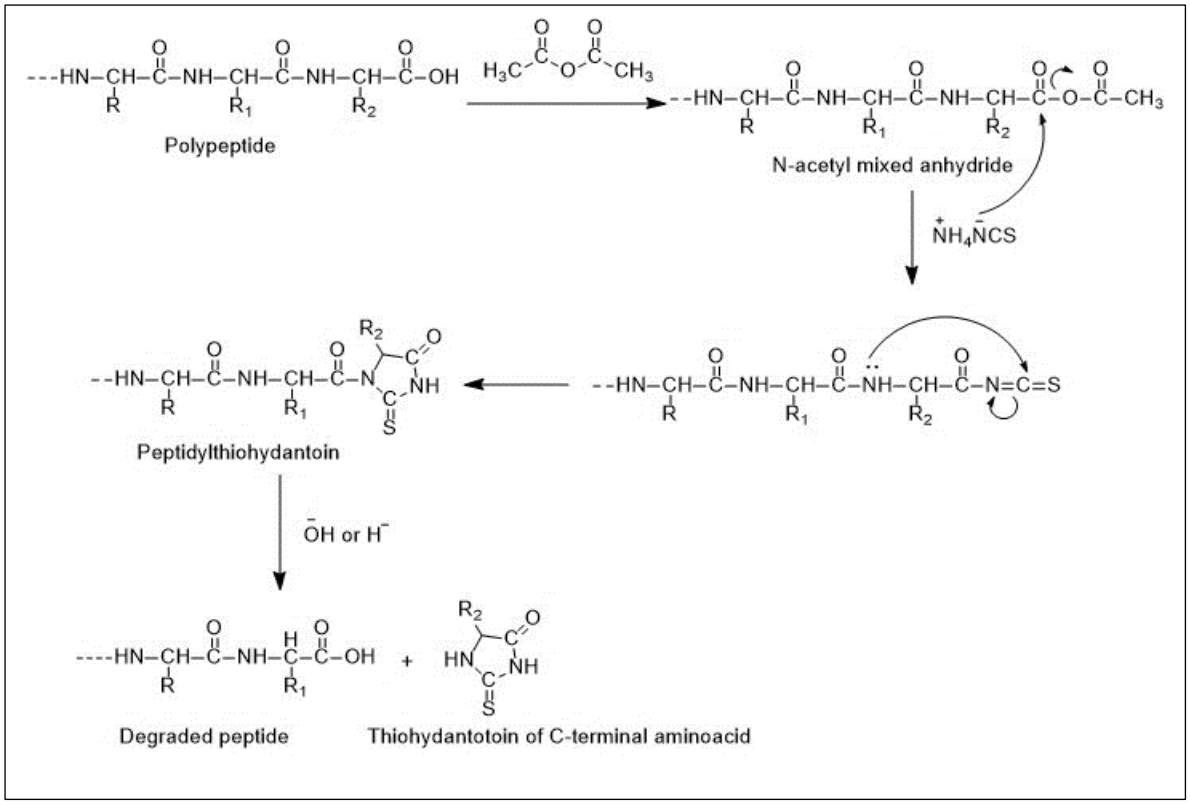
**i) Enzymatic Method ( with carboxypeptidase enzyme) :**

- The carboxypeptidase enzyme (obtained from pancreas) specifically hydrolyses the peptide bond adjacent to free carboxyl group.
- Hence when given polypeptide is treated with carboxypeptidase enzyme, it selectively separates the C-terminal amino acid along with formation of degraded peptide.
- The C-terminal amino acid so set free is identified. The process is repeated on degraded peptide and in this way the whole sequence of amino acids is determined.



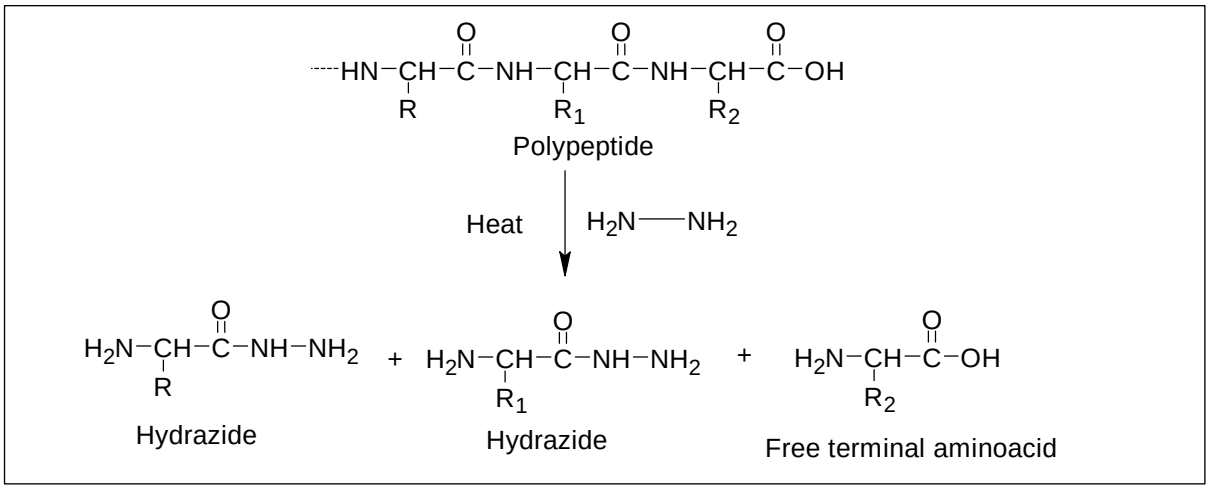
ii) **Thiohydantoin method : In this method,**

- First, side-chains of polypeptide having carboxyl groups and hydroxyl groups are protected as amides or esters.
- Then, the C-terminal carboxyl group is activated as an anhydride and reacted with thiocyanate.
- The C-terminal acyl thiocyanate peptide product automatically rearranges to a thiohydantoin incorporating the penultimate C-terminal unit.
- The peptidyl thiohydantoin so formed upon hydrolysis with acid gives thiohydantoin of C-terminal amino acid and degraded peptide. The of C-terminal amino acid is then identified.
- Thus, repetitive analyses is conducted on the degraded peptide and one by one whole sequence of peptide is determined.




**iii) Hydrazinolysis Method :**

- In this method, the given peptide is treated with anhydrous hydrazine at 373 K.
- The hydrazine forms aminoacyl hydrazide with every residue *except* the C-terminal amino acid.
- The C-terminus is thus readily identified by chromatographic procedures.
- The disadvantage of hydrazinolysis is that the entire sample is used to determine just one residue.




**Note 1** : The first ever primary structure of protein was determined by Fredric Sanger, a British Chemist and he was awarded Nobel Prize in 1958 for the same.

**Note 2 :** In practice it is not feasible to determine the sequence of all the residues in a long peptide chain by the stepwise removal of terminal residues. Instead, the chain is subjected to partial hydrolysis (acidic or enzymatic), and the fragments formed dipeptides, tripeptides, and so on-are identified, with the aid of terminal residue analysis. When enough of these smaller fragments have been identified. it is possible to work out the sequence of residues in the entire chain.

## 1.7: Synthesis of Peptides

Synthesis of polypeptides is of great significance. In recent times scientists have been successful in synthesizing some important polypeptides in laboratory. E.g. the insulin used for the treatment of diabetes was earlier obtained by extraction from the pancreas glands of cows and pigs. However, since the early 1980s, the "synthetic" insulin prepared by recombinant DNA technology has replaced "natural" insulin. Synthetic insulin is not only identical to human insulin, it is safer as well as less expensive than insulin obtained from animals.

Theoretically , it seems easy to join amino acids one by one starting from one terminal, but in actual practice , it is not so simple as, there are certain difficulties involved . The two important methods which are used for synthesis of polypeptides are as:

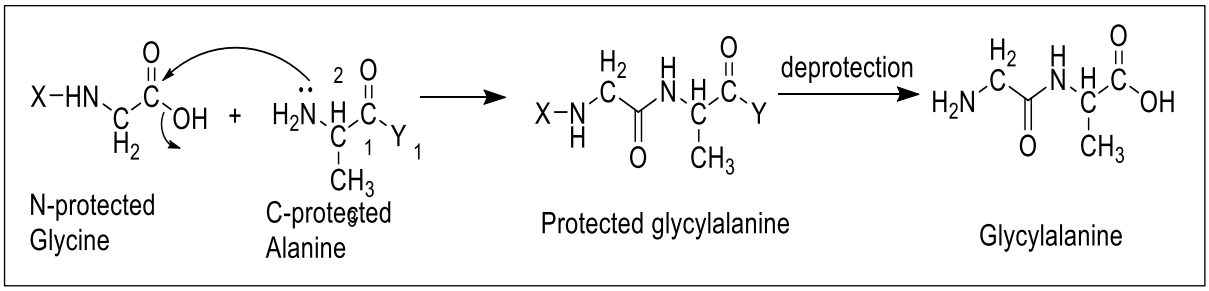
**i) Classical peptide synthesis ii) Merrifield Solid phase peptide synthesis**

## 1.7.1 : Classical peptide synthesis

It is also known as **solution phase synthesis of peptides**. The objective of peptide synthesis is to join amino acids in a specific sequence by peptide bond . Scientists have designed very effective methods and reagents for peptide bond formation, so that the joining of amino acids by amide linkages is not difficult. The real difficulty is in ensuring, the desired sequence. For example, If we want to synthesize a dipeptide, say Glycylalanine, by treating, glycine with alanine in presence of some dehydrating agent, there is equal chance that amino group of either of amino acid can react with -COOH group of other amino acid as well as of its own acid , thus forming four products instead of one.

Glycine + Alanine $\longrightarrow$ Gly-Gly + Gly-Ala + Ala-Ala + Ala-Gly + $H_2O$

To overcome this difficulty, the amino group of glycine and the carboxyl group of alanine must be protected so that they cannot react under the conditions of peptide bond formation. The formation of peptide bond with formation of desired dipeptide can be represented with the help of following equation. (where X and Y are $-NH_2$ and -COOH protecting groups, respectively)




Thus, dipeptide synthesis requires following sequence of four steps :

**Step 1:** Protection of the $-NH_2$ group of the N-terminal amino acid and the -COOH group of the C-terminal amino acid.

**Step 2 :** Activation of -COOH group of N-protected amino acid.

**Step 3 :** Reaction between two protected amino acids, so as to join them via amide bond or peptide bond formation.

**Step 4 :** Deprotection of the $-NH_2$ group of the N-terminus and the -COOH group of the C-terminus.

The same procedure can be further extended for synthesizing the higher peptides by using same logic as outlined for the synthesis of dipeptides.

The protecting group used for protecting amino group of N-terminal amino acid and carboxylic acid group of C-terminal amino acid must possess following characteristics:

(a) It should be easy to introduce.

(b) It should stable under experimental conditions.

(c) It should be easily removable leaving the desired dipeptide.

**Let us discuss, the synthesis of a dipeptide say Glycylalanine (Gly-Ala) :**

**Step 1a : Protection of amino group of N-terminal amino acid i.e. Glycine**

The two most widely groups which are used for protection of amino group are

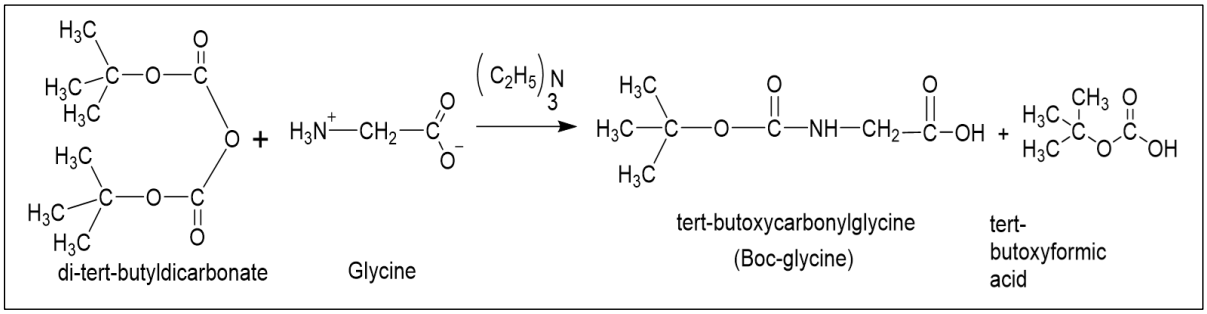
i) tert-butoxycarbonyl (Boc) group

ii) benzyloxycarbonyl (Z) group

The four reagents which are quite often used for protection of amino group are benzyl chloroformate, di-*tert*-butyl dicarbonate (sometimes abbreviated Boc2O, where Boc stands for *tert*-butyloxycarbonyl), N-ethoxycarbonylphthalimide and 9-fluorenylmethyl chloroformate:

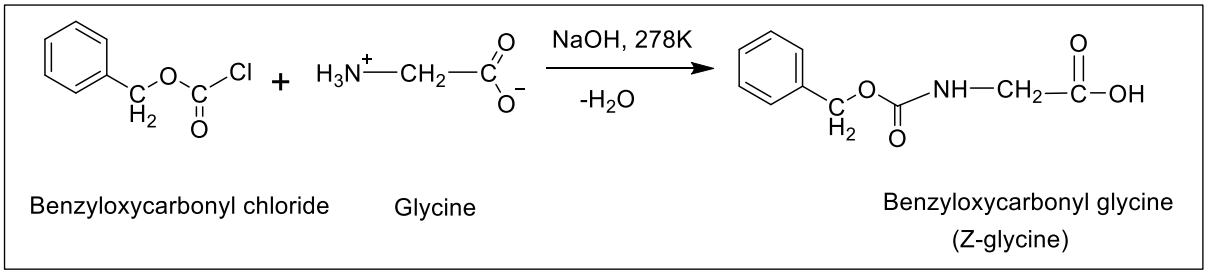
**i) Protection by tert-butoxycarbonyl (Boc) group :**

tert-butoxycarbonyl group { $(CH_3)_3C\text{-}OCO\text{-}$} is abbreviated as Boc. Protection by Boc is done by treating the amino acid with di-tert-butyldicarbonate i.p.o. triethylamine as:



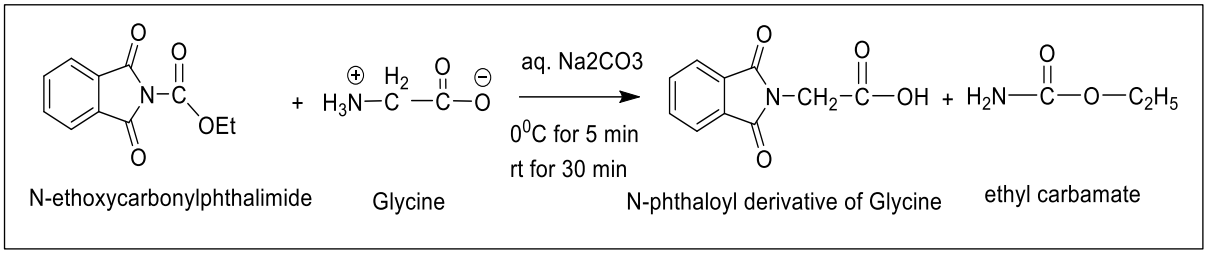

**ii) Protection by benzyloxycarbonyl (Z) group :**

Benzyloxycarbonyl group has been given symbol Z as per IUPAC system. Protection by Z is done by treating the amino acid with bezyloxycarbonyl chloride or Benzyl chloroformate ( earlier known by name carbobenzyloxy chloride) i.p.o. NaOH at $5^0$C.

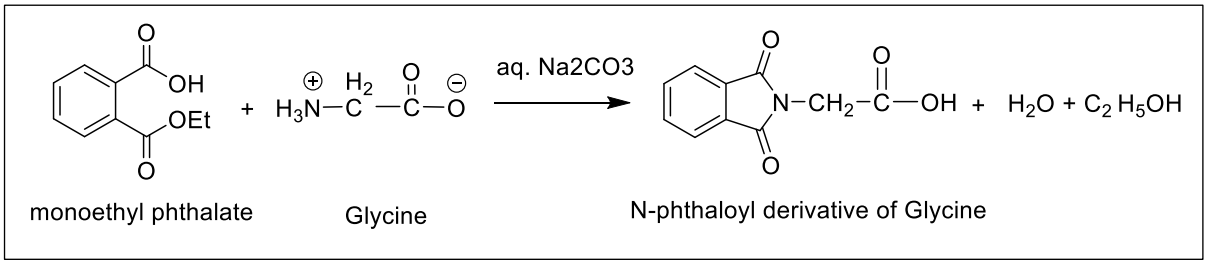


**iii) Protection by phthaloyl group :**

Protection by Phthaloyl group is done by treating the amino acid with N-ethoxycarbonylphthalimide i.p.o. aq $Na_2CO_3$ as depicted below.

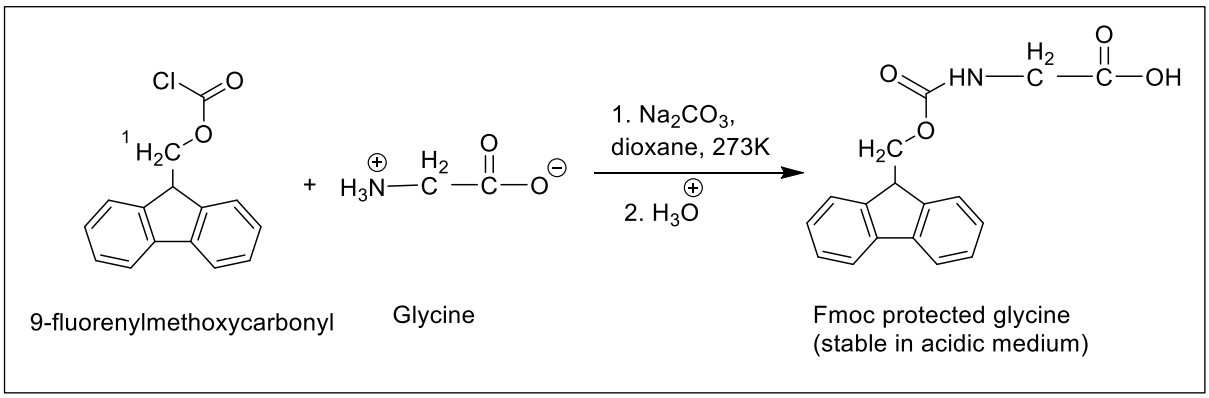



Protection by Phthaloyl group can also be done by treating the amino acid with monoethyl phthalate as:



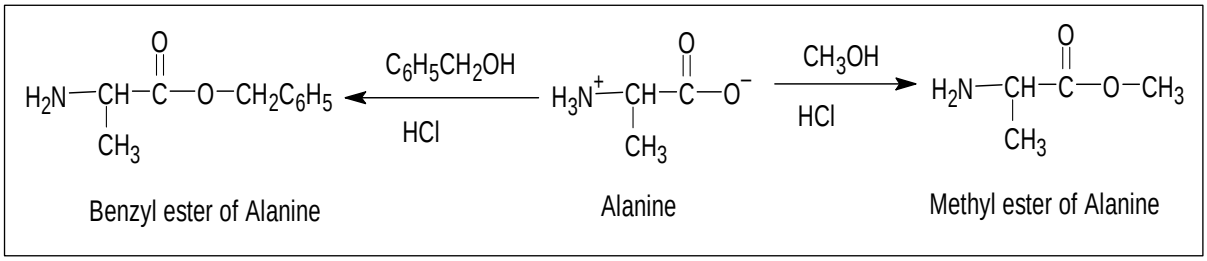

**iv) Protection by 9-fluorenylmethoxycarbonyl (Fmoc) group :**

Protection by **9-fluorenylmethoxycarbonyl** is done by treating the amino acid with N-9-fluorenylmethyl chloroformate i.p.o. sodium carbonate and dioxane at 273K . E.g. protection of amino group with Fmoc group is done as depicted below.



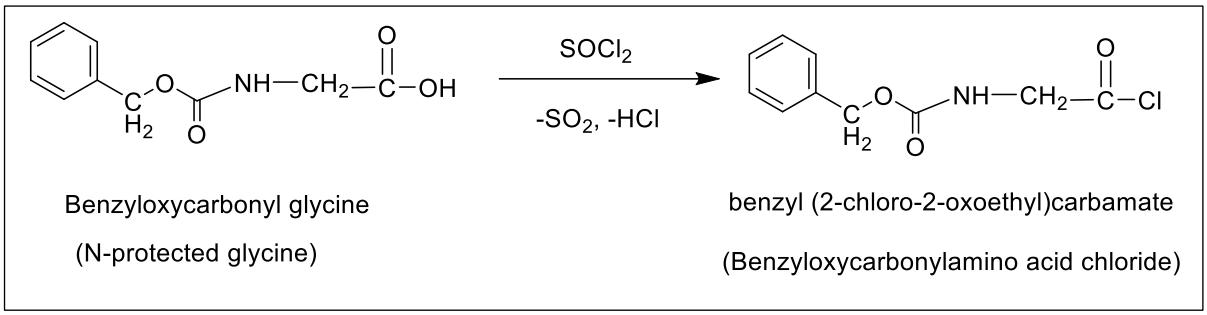

**Step 1b : Protection of -COOH group of C-terminal amino acid i.e. Alanine:**

This is done by converting the -COOH group into its methyl, ethyl, or benzyl ester. e.g In our example , of synthesis Glycylalanine, Alanine is C-terminal amino acid and hence, its -COOH group is protected as shown below :

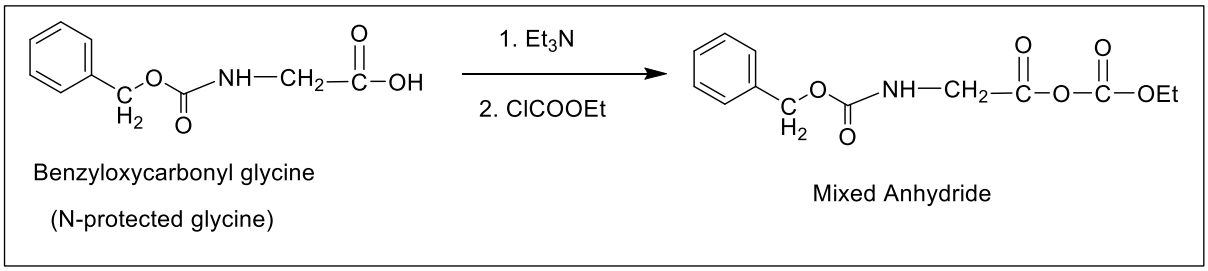



**Step 2 : Activation of the carboxyl Group of N-protected amino acid : It is done by following methods:**

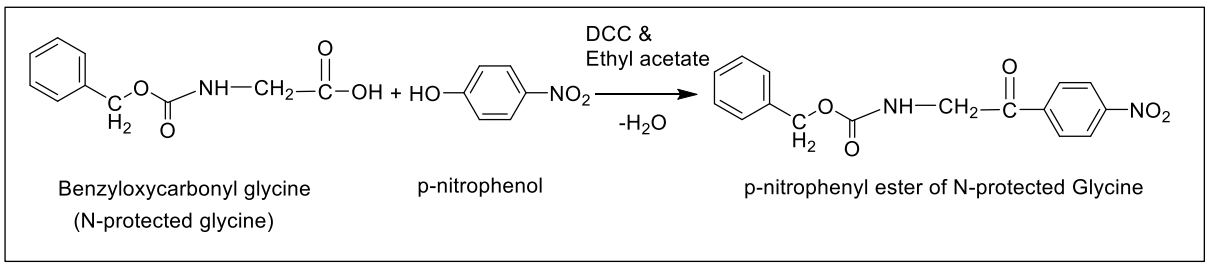
**i) By converting -COOH group into acid chloride :** In earlier methods, carboxyl group of N-protected amino acid was activated by converting it to an acyl chloride This is done by treating the N-protected amino acid with $SOCl_2$ or $PCl_5$. .




ii) **By converting into mixed anhydride :** The acid chlorides as formed in above method are actually more reactive than necessary. Due to which their use leads to complicating side reactions Hence, a much better method is to convert the carboxyl group of the N-protected amino acid to a mixed anhydride using ethyl chloroformate as described below.

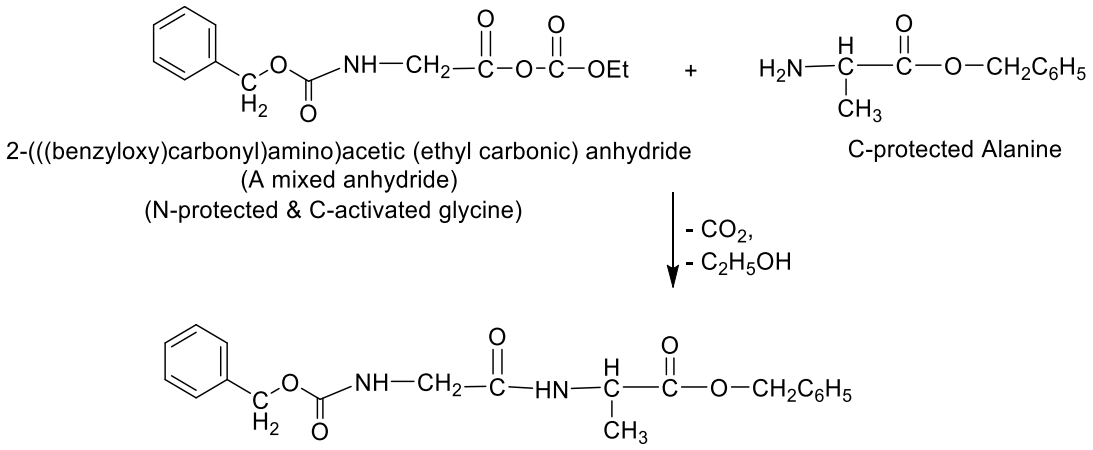



iii) **By converting -COOH group into a reactive p-nitrophenyl ester**: This is done by treating N-protected amino acid (Boc-glycine in our example) with p-nitrophenol i.p.o. N,N`-dicyclohexylcarbodiimide (DCC) under mild conditions as:



**Step 3 : Formation of Peptide Bond :**

The N- protected and C- activated amino acid is then treated with C-protected amino acid having free amino group, when we get desired N- and C-terminal protected dipeptide. For example, for synthesising Glycylalanine, a dipeptide, the N-protected and C-activated glycine is treated with C-protected Alanine, when it results in formation of peptide bond as:

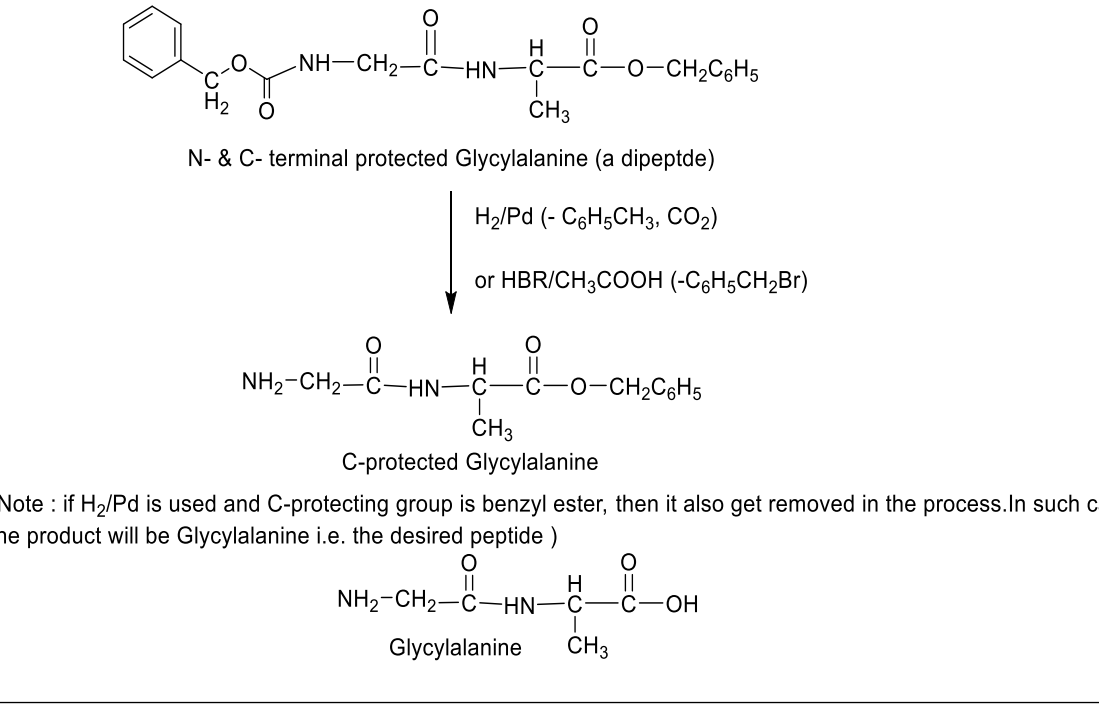



**Step 4 : Removal of Protecting Groups :**

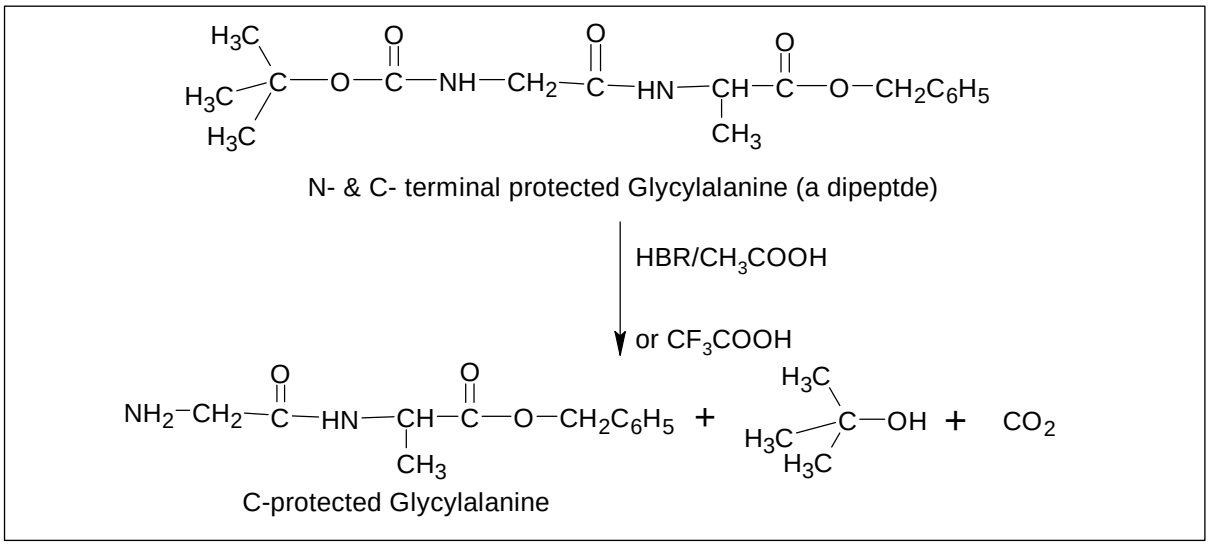
Finally, the protecting groups are removed, to obtain the desired dipeptide. Depending upon the type of protecting groups, following reagents and methods are commonly used as described below:

a) **Removal of N-protecting groups :** N- protecting groups are removed by following methods as ;

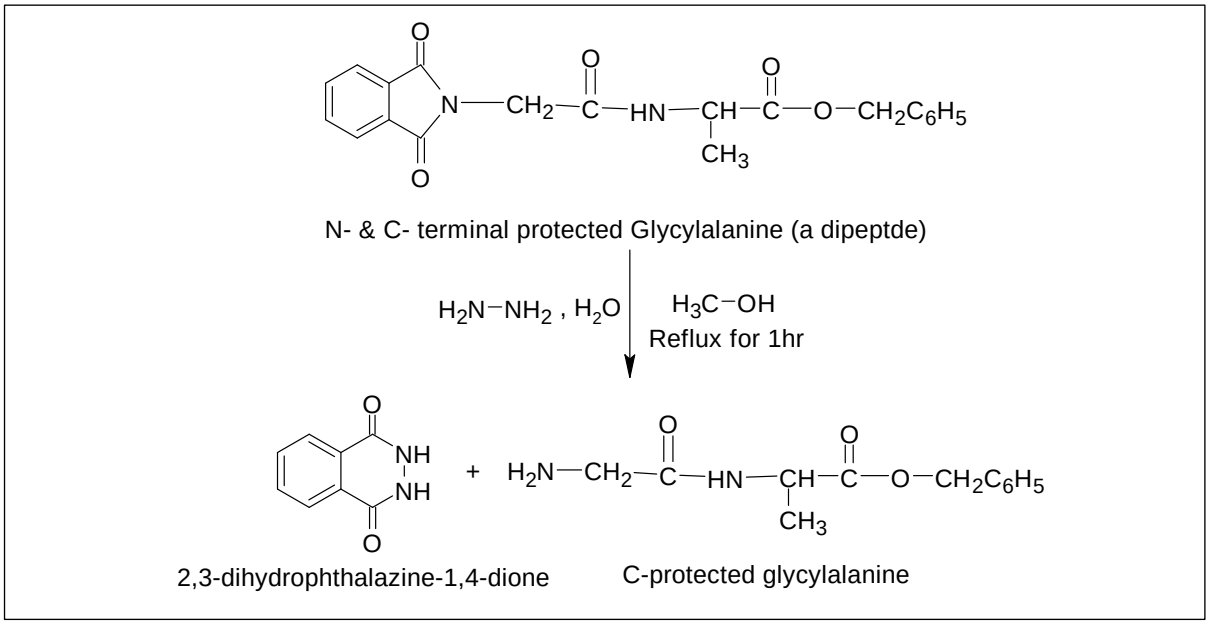
**i) Removal of benzyloxycarbonyl group** : Benzyloxycarbonyl (Boc) group is removed either by catalytic hydrogenation with $H_2$/Pd or by hydrolysis with HBr in cold $CH_3COOH$.



ii) **Removal of t-Butoxycarbonyl group :** It can be removed by treating the N-protected dipeptide with $HBr/CH_3COOH$ or by treatment with $CF_3COOH$

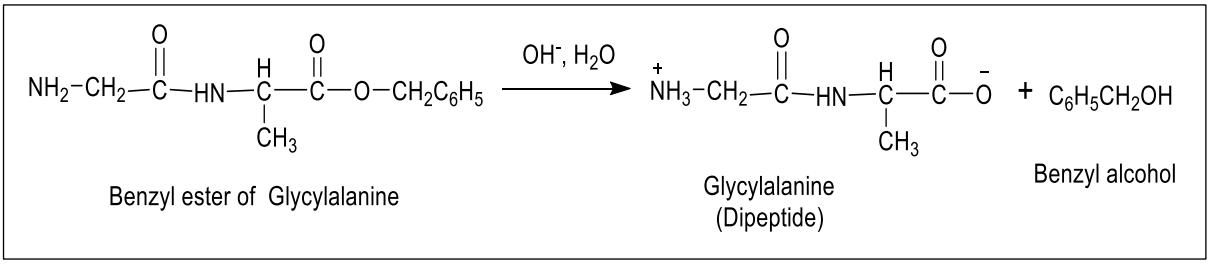



iii) **Removal of phthaloyl group :** The Phthaloyl group is removed by means of hydrazinolysis by treating the N-protected dipeptide with hydrazine hydrate in refluxing MeOH as :

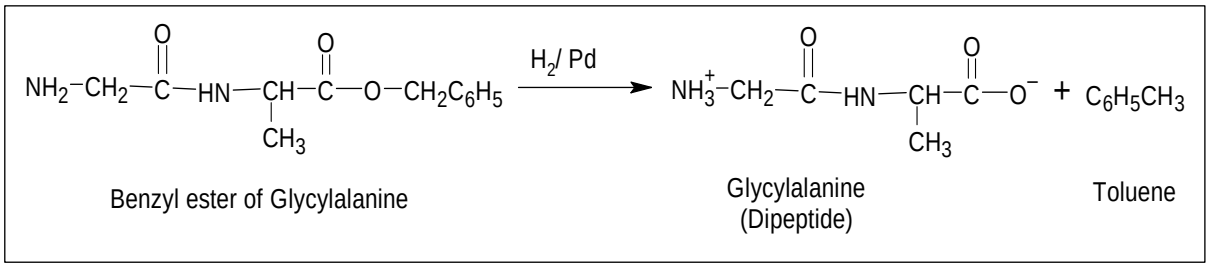



**b) Removal of Carboxyl or C-protecting groups :**

The C-protecting groups are removed by hydrolysis with an aqueous base or acid under mild conditions. Under mild conditions, only the ester get hydrolysed, where as amide bond remains intact. This is because esters are more reactive than amides.




Benzyl esters can also be removed by catalytic hydrogenation with $H_2$ /Pd i.e. by hydrogenolysis.




**1.7.2: Limitations of Classical Peptide Synthesis :**

The classical peptide synthesis involves condensation of amino acid molecules in a step wise manner. Hence, it can be presumed that this method cab used for thesis of polypeptides of any length. However, in actual practice, it has been found that this method can be used to synthesize polypeptides containing up to 10 amino acid residues. However, for the synthesis of polypeptides with larger number of amino acids, this method can't be used because of following limitations :

i) **Time consuming:** The classical method of peptide synthesis requires several steps for each new peptide bond formed, like protection, activation, condensation and deprotection . Moreover, the new peptide synthesized at each step needs to be purified before repeating the procedure . Hence, it requires lot amount of time, which makes this method quite time consuming, even for a small peptide.

**ii) Low yield.** Since , classical peptide synthesis involves number of steps, and each step needs purification of main product before proceeding to next step, the overall yield goes on decreasing after each step even though each individual step has an excellent yield of 80%. For example, if the yield of a dipeptide formed is 80% , then the yield of the tripeptide will be 80% of 80 i.e 64%. Similarly, the yield of tetrapeptide would be 80% of 64% i.e. 51.2%. If a dodecapeptide is to be synthesized, the overall yield would be around 8.576%. Hence, it is quite evident that the classical method is not useful for the synthesis of higher peptides.

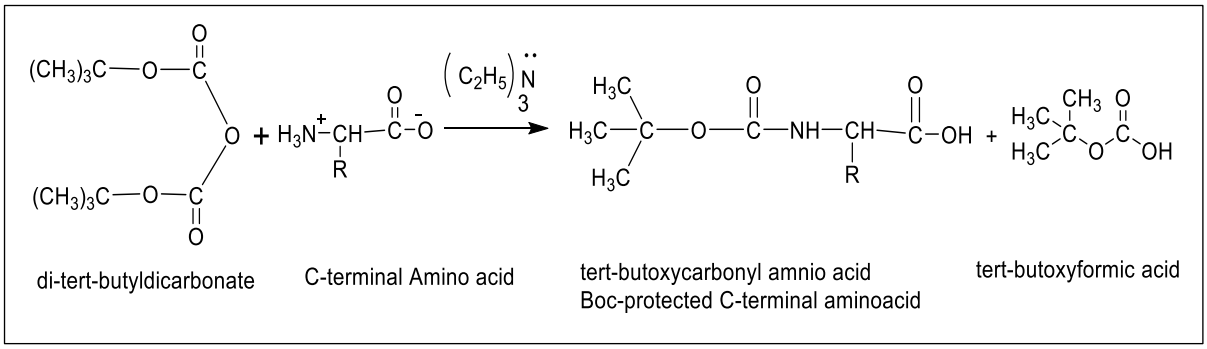
## 1.7.3: Merrifield Solid-Phase Peptide Synthesis

The classical method of peptide synthesis has been used to synthesize a number of polypeptides, including ones as large as insulin. However, the difficulties encountered in the classical peptide synthesis make the method less useful as at every stage, the peptides need to be purified. Furthermore, with each isolation and purification stage, significant loss of peptide may occur, thus decreasing substantially the yield of final polypeptide. The difficulties encountered in the classical peptide synthesis were overcome by R.B. Merrifield (Rockefeller University, USA) . He devised a brilliant method for peptide synthesis called the solid phase peptide synthesis in 1964. He was awarded Nobel Prize in Chemistry in 1984 for developing solid-phase peptide synthesis method.

**Principle :** The principle involved in this is method is to synthesize the peptide residue by residue while one end of the peptide remains attached to an insoluble solid support. The insoluble support is a polymer of styrene crossed linked with about 2% p-divinylbenzene. In addition, 5 % of the benzene rings carry chloromethyl group. The chloromethylated polystyrene polymer is used in the form of small beads, which are porous. It is important to note that protecting groups and other reagents are still necessary, but because the peptide being synthesized is attached to a solid support, hence by-products, excess reagents, and solvents can simply be rinsed off. When the desired polypeptide get synthesized, it is cleaved from the polymer support and finally purified by HPLC.

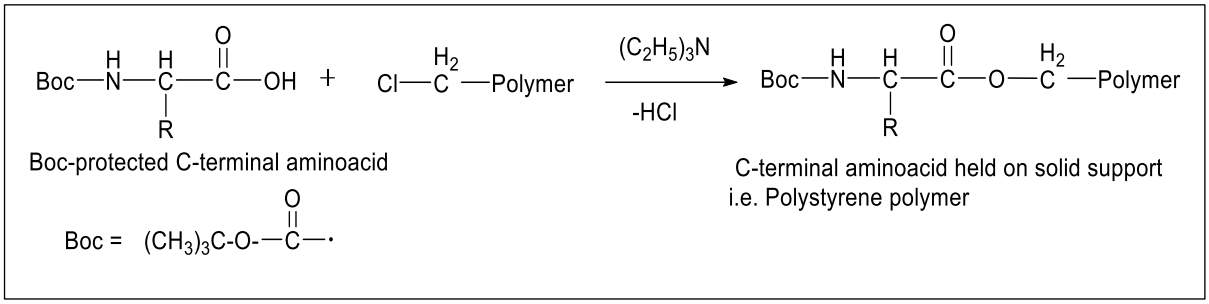
Let us explain the main steps involved in Solid Phase Peptide synthesis

**Step 1 : Protection of Amino group of C-terminal amino acid :** The amino group of C-terminal acid of desired peptide is protected by suitable protecting group. E.g. tert-butoxycarbonyl chloride (Boc) can be used as :

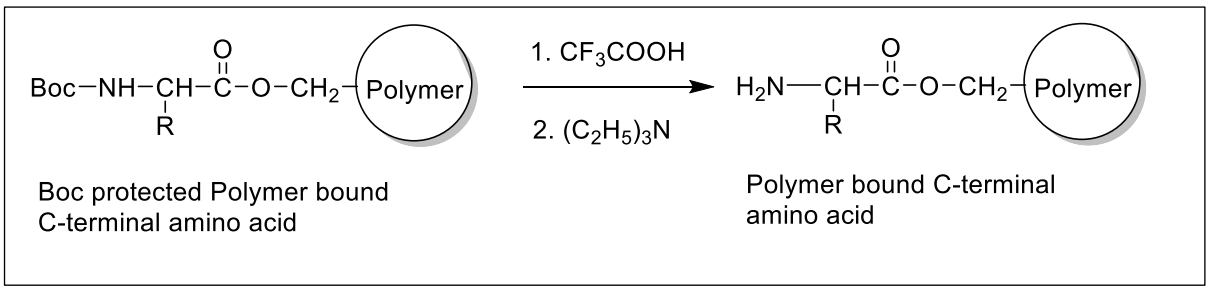



**Step 2 : Attachment of N-protected C-terminal amino acid to Polymer :** The Boc protected C- terminal amino acid is then attached to the polystyrene polymer via benzyl ester linkage by

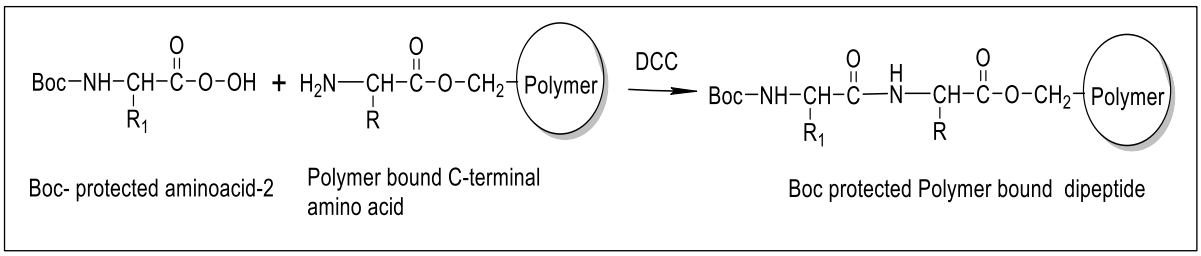
heating it in presence of triethyl amine in a suitable solvent . The excess reagents are removed by washing with suitable solvent.



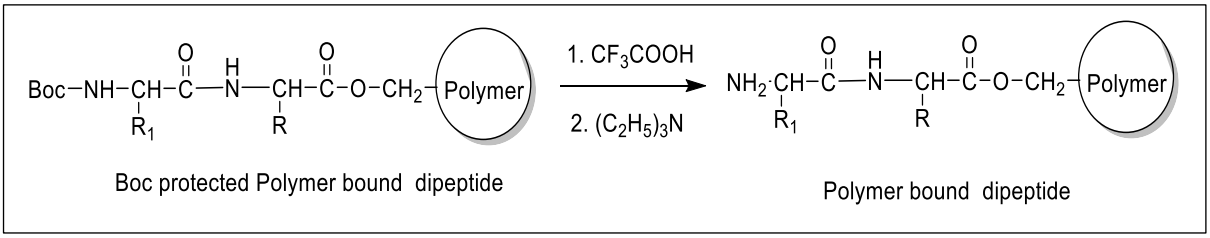

**Step 3 : Removal of Boc protecting group :** The Boc protecting Polymer bound amino acid is then removed by treating with trifluoro acetic acid and the salt of polymer bound amin acid formed is converted into free form by treating with excess of triethylamine. Finally, the Polymer bound amino acid is purified by washing.




**Step 4 : Addition of next Boc-protected amino acid** : The next Boc protected amino acid is then linked to polymer bound amino acid by condensation reaction with the help of dehydrating agent diisopropylcarbodiimide (DCC). The excess reagents are removed by washing.

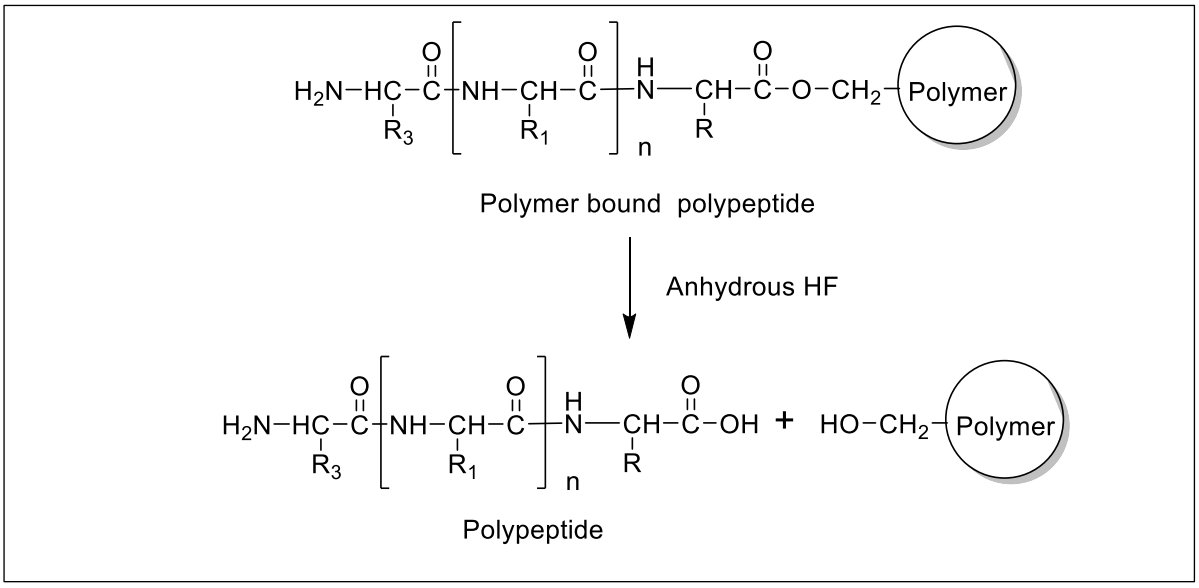



**Step 5 : Removal of Boc-protecting group :** The Boc-protecting group is removed by treating with trifluoro acetic acid as :



**Step 6 : Repetition of steps 4 & 5 :** The steps 4 and 5 are repeated to add desired amino acids . By repeating these two steps, we can add as many amino acids as we want and hence can synthesize the desired polypeptide.

**Step 7 : Separation of polypeptide chain from Polymer :** When the desired polypeptide has been synthesised, it is then separated from polymer by treating the Polymer bound polypeptide with anhydrous HF as :




**Note :** Fmoc is preferably used instead of Boc as protecting group because with each new addition of amino acid residue , the repetitive application of the acidic conditions to remove Boc-protecting groups from each new residue slowly interfere with the synthesis by prematurely cleaving some peptide molecules from the solid support and deprotecting some of the side chains. The basic conditions for Fmoc removal eliminates the probability of these problematic side reactions.

**Advantages of Solid Phase Peptide Synthesis(SPPS)** :

- The great advantage of solid-phase peptide synthesis is that purification of the peptide at each stage involves simply rinsing the plymer beads (solid support) so to wash away excess reagent, by-products, and solvents.
- Furthermore, the peptide being attached to a concrete solid support during the synthesis, facilitates the all the steps involved in the synthesis of desired peptide to be carried out by a machine in repeated cycles. Hence polypeptide synthesis can be carried out Automated peptide synthesizers that can complete one cycle in 40 min and can perform unattended operation of 45 cycles. So this saves lot of time. Though, it is not as efficient

as protein synthesis in the body, where enzymes directed by DNA can synthesize a protein with 150 amino acids in about 1 min.

- An example automated peptide synthesis was the synthesis of ribonuclease, a protein with 124 amino acid residues. The synthesis involved 369 chemical reactions and 11,930 automated steps—all carried out without isolating an intermediate. The synthetic ribonuclease not only had the same physical characteristics as the natural enzyme, it possessed the identical biological activity as well. The overall yield was 17%, which means that the average yield of each individual step was greater than 99%.

**Problems for Practice :**

**Problem 1:** What is meant by protection of group ? Name the conditions which a group must satisfy to act as protecting group.

**Problem 2**: Illustrate various methods used for protecting N-terminal of Amino acid.

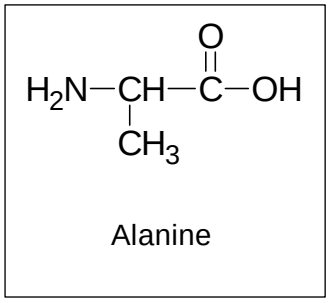
**Problem 3:** Illustrate various methods used for protecting C-terminal of Amino acid.

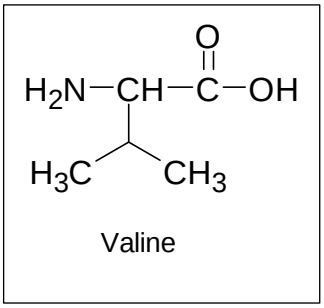
**Problem 4:** Write down various steps involved in Classical Peptide synthesis.

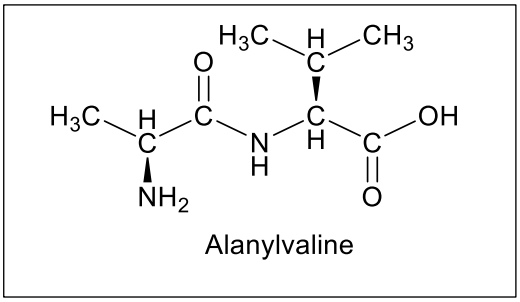
**Problem 5:** How solid phase peptide synthesis is superior to classical peptide synthesis.

**Problem 6:** Using Fmoc group as N-protecting group, describe synthesis of Alanylvaline by classical peptide synthesis method.

**(Hint :** In this dipeptide, Alanine is N-terminal amino acid and Valine is C-terminal amino acid)

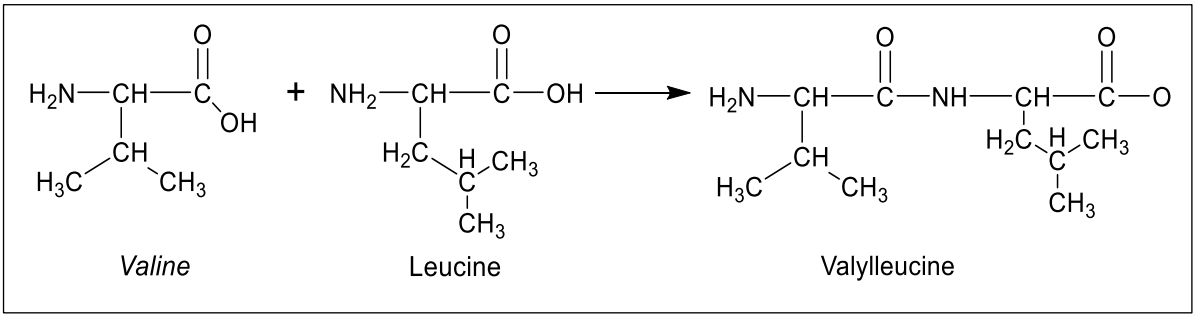



**Problem 7**: Using Boc group as N-protecting group, describe synthesis of Valylleucine by Solid Phase peptide synthesis method.

**(Hint: In this case Valine is N-terminal amino acid and Leucine is C- terminal amino acid. General reaction given. Write all the steps.)**



**Reference Books :**


- Kotz, J.C., Treichel, P.M. & Townsend, J.R. General Chemistry, Cengage Learning India Pvt. Ltd.: New Delhi (2009).
- Mahan, B.H. University Chemistry, 3rd Ed. Narosa (1998).
- Petrucci, R.H. General Chemistry, 5th Ed., Macmillan Publishing Co.: New York (1985).
- Morrison, R. T. & Boyd, R. N. Organic Chemistry, Dorling Kindersley (India) Pvt. Ltd. (Pearson Education).
- Finar, I. L. Organic Chemistry (Volume 1), Dorling Kindersley (India) Pvt. Ltd. (Pearson Education).
- Finar, I. L. Organic Chemistry (Volume 2), Dorling Kindersley (India) Pvt. Ltd. (Pearson Education). ·
- Nelson, D. L. & Cox, M. M. Lehninger's Principles of Biochemistry 7 th Ed., W. H. Freeman.
- Berg, J.M., Tymoczko, J.L. & Stryer, L. Biochemistry, W.H. Freeman, 2002.
- Graham Solomons, T. W., & Fryhle, C. B. (2011). Organic chemistry 10th edition.


**Some Important Links where students can have more e-resources:**

1. **Link for Vidya-mitra, Integrated E-Content Portal** : http://vidyamitra.inflibnet.ac.in/
2. **Link for e-PG Pathshala – Inflibnet** : https://epgp.inflibnet.ac.in/
3. **Link for** INFLIBNET Centre Gandhinagar : https://inflibnet.ac.in/