CHEMISTRY OF PROTEINS

BY DR. VIJAY KUMAR
DEPT. OF CHEMISTRY
UP COLLEGE, VNS.

AMINO ACIDS :

Introduction : "Amino acids are compounds which contains at least one amino group and one carboxylic group"

* The amino group could be linked to carbon just next to carboxylic group or with gap of one or two carbons, they are called α, β or γ amino acids respectively.

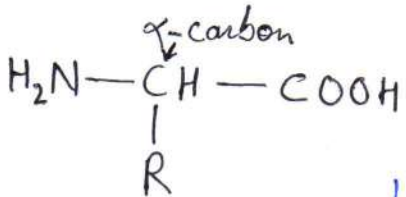
H_2N-CH_2-COOH ⇒ 2-Aminoethanoic acid OR α-Amino acid.

$H_2N-CH_2-CH_2-COOH$ ⇒ 3-Aminopropanoic acid OR β-amino acid

$H_2N-CH_2-CH_2-CH_2-COOH$ ⇒ 4-Aminobutanoic acid OR γ-amino acid

{* It is α-amino acids which have maximum importance because they are the constituent units of proteins.}

* In general, an α-amino acid can be represented as



R = H ⇒ Glycine
= CH_3 ⇒ Alanine
= CH_2-OH ⇒ Serine
= CH_2-SH ⇒ Cysteine

Where, R is an alkyl OR aryl group, it could also represent highly branched OR unsaturated carbon chain OR heterocyclic ring.

CLASSIFICATION OF AMINO ACIDS

Amino acids can be classified in different ways:-

A: As essential or non-essential amino acid

Amino acids which are very important or essential for growth of humans and animals are called essential amino acids. These amino acids can not be synthesised by the body and must be supplied in the diet as such. Lack of these amino acids in the diet may cause the disease Kwashiorkor

Amino acids which can be synthesised by our body are called non-essential amino acids.

B. As Polar or Non-polar [OR Acidic, basic or neutral] amino acids.

(i) Neutral Amino acids: Amino acids having one amino and one carboxylic groups are called neutral amino acids.

Example: Glycine,
alanine,
Valine etc.

(ii) Acidic Amino Acids: Amino acids having one additional (one - NH_2 and two -COOH) carboxylic group are called acidic amino acids.

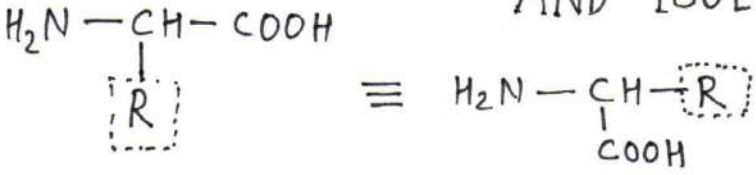
Examples: ① Aspartic acids ② Glutamic acids

(iii) Basic Amino acids:

Those amino acids which have one additional - NH_2 group (2 - NH_2 and 1 - COOH groups) are called basic aminoacids.

Examples: ① Lysine ② Arginine

AMINO ACIDS : CLASSIFICATION, STRUCTURE, NATURE AND ISOELECTRIC POINT



TYPE OR NATURE		S. No.	Name	Structure, R =	Isoelectric Points, pI
NEUTRAL	NON-POLAR	1	Glycine (Gly)	$-H$	5.97
		2	Alanine (Ala)	$-CH_3$	6.02
		3	Phenylalanine (Phe)*	$-CH_2-C_6H_5$ [$-CH_2-$⬡]	5.84
		4	Valine (Val) *	$-CH(CH_3)(CH_3)$	5.97
		5	Leucine (Leu)*	$-CH_2-CH(CH_3)(CH_3)$	5.98
		6	Isoleucine (Ile)*	$-CH(CH_3)-CH_2-CH_3$	6.02
		7	Proline (Pro)#	$-CH_2-CH_2-CH_2-$[HN—CH—COOH, $(CH_2)_2$, CH_2, #]	6.30
		8	Methionine (Met)*	$-CH_2-CH_2-S-CH_3$	5.74
	POLAR	9	Cysteine (Cys)	$-CH_2-SH$	5.02
		10	Serine (Ser)	$-CH_2-OH$	5.68
		11	Tyrosine (Tyr)	$-CH_2-C_6H_4-OH$ [$\equiv CH_2-$⬡$-OH$]	5.67
		12	Threonine (Thr)*	$-CH(OH)-CH_3$	5.60
		13	Asparagine (Asn)	$-CH_2-C(=O)-NH_2$	5.41
		14	Glutamine (Gln)	$-CH_2-CH_2-C(=O)-NH_2$	5.70
		15	Histidine (His) *	$-CH_2-C=CH$ (HN, $NH^{\oplus}$, CH)	7.59
		16	Tryptophan (Try)*	$-CH_2-C-C$ (NH)	5.88
Acidic		17	Aspartic Acid (Asp)	$-CH_2-COOH$	2.98
		18	Glutamic Acid (Glu)	$-CH_2-CH_2-COOH$	3.22
Basic		19	Lysine (Lys)*	$-CH_2-CH_2-CH_2-CH_2-NH_2$	9.74
		20	Arginine (Arg)*	$-CH_2-CH_2-CH_2-NH-C(=NH_2^{\oplus})-NH_2$	10.76

Note : * : Essential Amino Acids

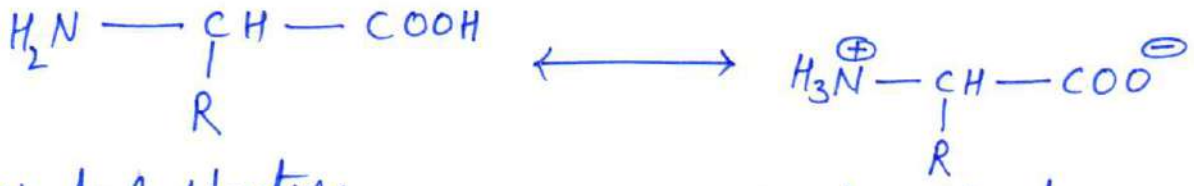
\# Complete structure of proline

Isoelectric Point, Zwitter Ion Structure

And Effect of pH on Amino Acids

Dipolar Nature of Amino Acids (Zwitter ion Structure)

It has been found that an amino acid molecule appears as dipole [means one end of it carrying ⊕ve charge and the other end carrying ⊖ve charge.



Neutral structure of amino acid

Dipolar structure of amino acid [Zwitter ion structure]

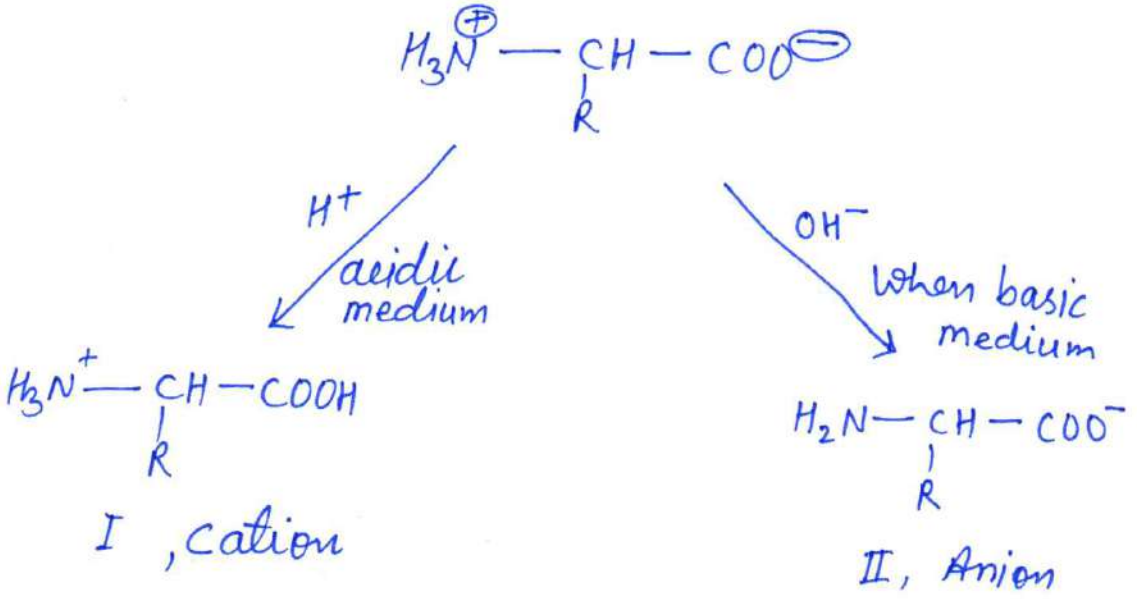
⇓

There is no free amino, $-NH_2$ or carboxylic, $-COOH$ group present in the molecule.

Evidence in support of Zwitter ion structure

1. Spectroscopic studies of amino acids do not show bands characteristics of $-NH_2$ & $-COOH$ groups.
2. Amino acids are insoluble in non-polar solvent and soluble in polar solvents.
3. Amino acids are non-volatile crystalline solids, which melts at high temperature.
4. Amino acids have high dipole moments
5. Dissociation constant K_a and K_b gives an idea about acid & base strengths. Amino acids have very low values of K_a & K_b, which indicates these molecules do not possess these groups ($-NH_2$ & $-COOH$).

Effect of pH on the structure of Amino Acids



ISOELECTRIC POINT :-

As amino acids are polar in nature, they show electrical properties. On applying electrical field to the solution of an amino acid, they migrate to one or the other end/ electrode depending on the following factors :-

(a) If the solution of amino acid is acidic, amino acid behave as cation, I, so it migrates towards cathod [⊖vely charged electrode].

Amino acid $\xrightarrow{pH < 7.0}$ ⊕ve ⇒ moves towards cathod, ⊖ve pole.

(b) If solution is basic → Amino acid behave as anion
⊖ve ⇒ It will move towards Anode ⊕ve pole.

At certain pH values (pH of solution of amino acid) the cationic and anionic structures will be in equal concentration and on passing electricity we shall observe that there is no movement of amino acid.

"The pH at which a particular Amino acid does not migrate under the influence of the electrical field is called isoelectric point"

Application of Isoelectric point : ELECTROPHORESIS

pI

⇓

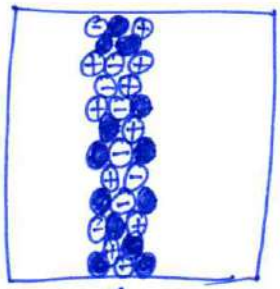
Separation of amino acids.

Every amino acid has a characteristic isoelectric point for example, glycine has 5·97, alanine has 6·02.

⇒ Electrophoresis is a process of separation and purification of compounds on the basis of movement of charged particle (amino acid) in an electric field.

Separation of mixture of amino acids can be brought about by electrophoresis

Isoelectric Point, $pI = \frac{Ka + Kb}{2}$

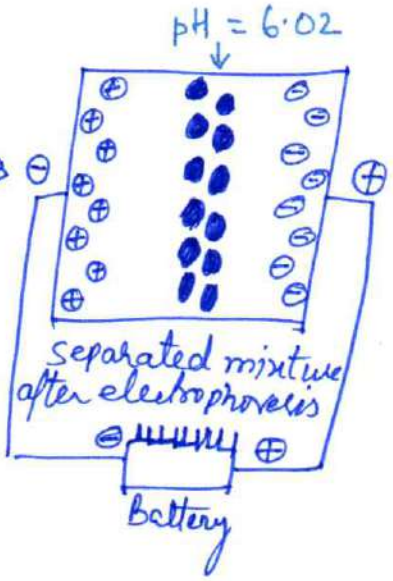


Mixture of 3 amino acid, say Asp, Ala & Lys

mixture before separation/electrophoresis

⟹ The desired pH (means pI of amino acid which are going to be separated) is attained or maintained

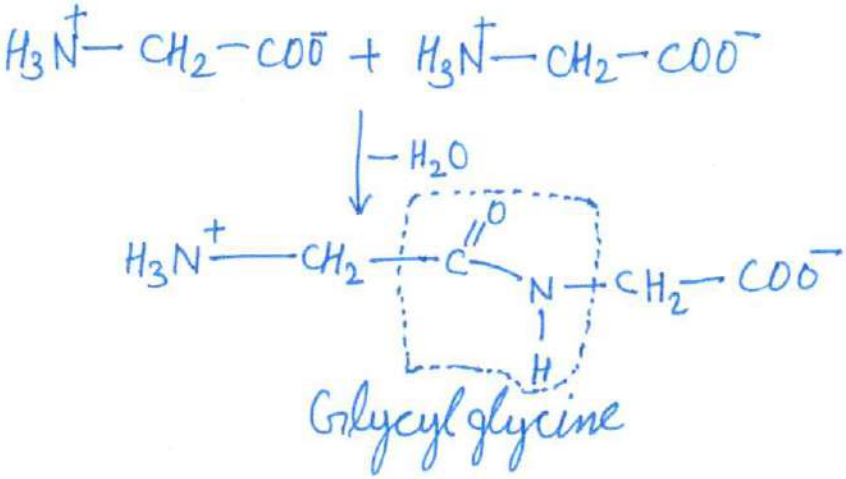
[pI Ala = 6·02]



PEPTIDES

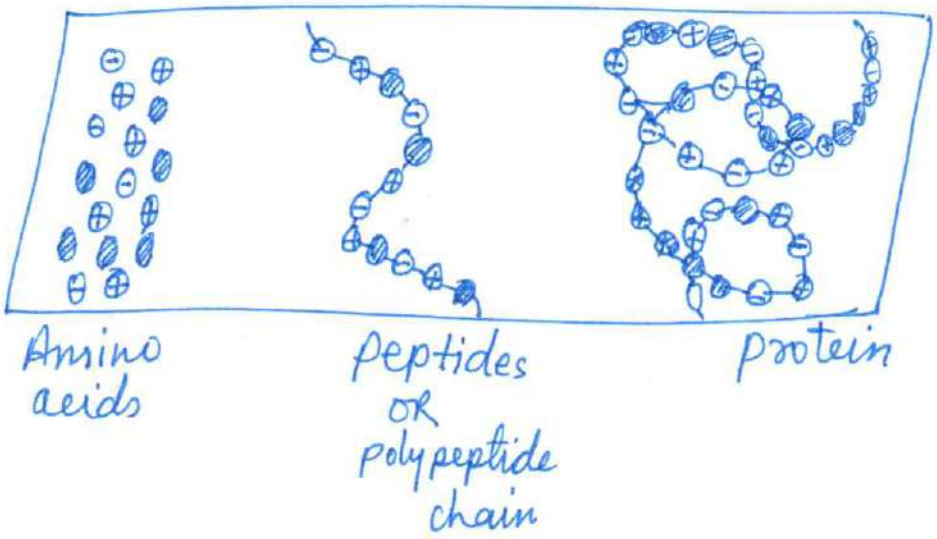
Peptides are amides obtained by interaction between the amino acids (amino and carboxylic groups) of two or more numbers.

For Examples; Two molecules of glycine combine to form amides substance known as glycyl glycine



Glycyl glycine

The group "—CO—NH—" in the peptides is called peptide linkage.



Classification: Dipeptide: Two amino acids joining together forming one peptide bond

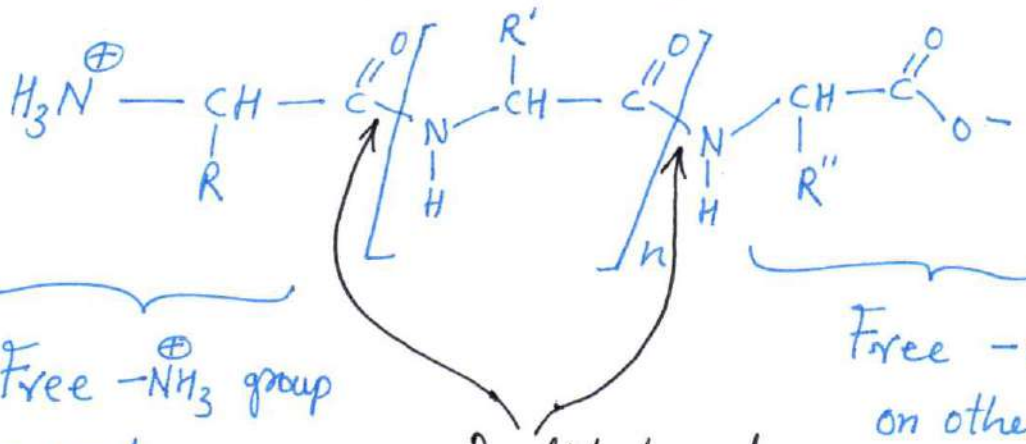
Tripeptide: Three amino acid joined [Two peptide bonds.]

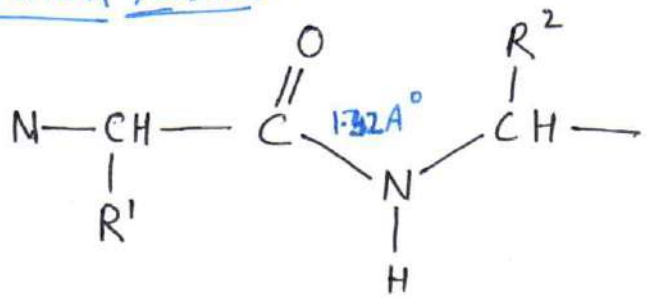
Tetra: Four amino acids joined [Three peptide bonds.]

Polypeptide: More than 4 amino acids joind together.

Polypeptides :

A polypeptide may be represented as





* Normal C — N bond length is 1.47A°, where as in peptide bond, it is 1.32A°, which indicates partial double bond character.

\# Peptides of molecular weight upto 10,000 are known as polypeptides, where as peptides of higher molecular weight are known as proteins.

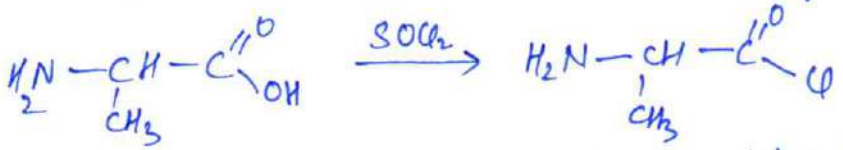
\# Conventionally, N-terminal amino acids residues is written at the left end while the C-terminal amino acid residue is written at the right end.

SYNTHESIS OF PEPTIDES FROM AMINO ACIDS

I Classical Peptide Synthesis :

The condensation reaction between two amino acids to form peptide linkage is an endothermic reaction and hence does not takes place easily. Therefore, in actual practice the carboxylic group of one amino acid molecule is activated by converting it in acid chloride & then this is reacted with amino group of another amino acid molecules.

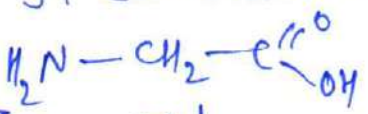
But there is another problem caused by the presence of both the aminoacid groups in the same molecule, hence both can react with activated/reactive molecule OR for example consider the formation of a simple dipeptide, alanylglycine, Ala-Gly, we first activate the carboxyl group of Ala by converting it in to acidchloride and then allow it to react with glycine. In practice alanyl chloride not only react with glycine but also with another alanylchloride (with $-NH_2$ gp).



Alanyl chloride

⇓

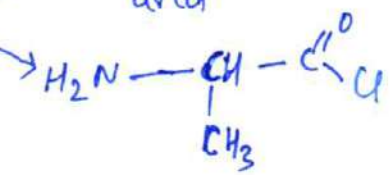
It can react with



⇒ Glycine

and

Reacting site →



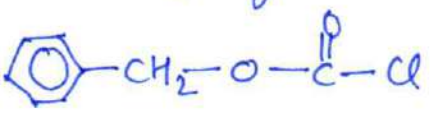
To overcome above problem, we protect the H_2N- group before activating the $-COOH$ group of particular amino acid.

Now we have different protecting group/reagents for $-NH_2$ and $-COOH$ group seperately.

Protection of $-NH_2$ Group :- A suitable protecting group would be that which can be introduced in the starting amino acid and can be removed after the reaction without disturbing the peptide bond formed.

① One such reagent is

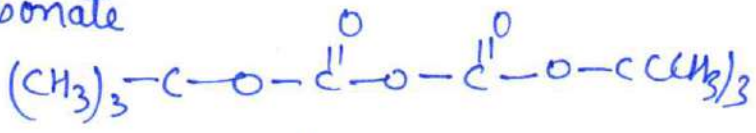
Benzyloxycarbonyl chloride



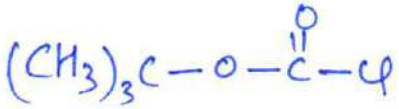
OR

carbonybenzoxy chloride

② di tertiarybutyl dicarbonate



OR



tertiary butoxycarbonyl

(BOC)

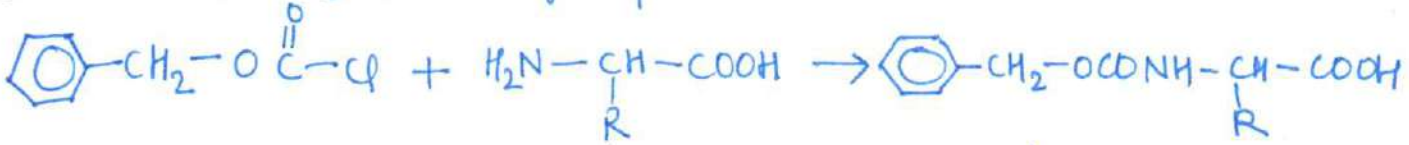
At the end, these group can be removed/detached/ eliminated by treatment with HBr, after the peptide has been formed.

* Synthesis of a peptide from amino acids involves the following steps :—

① Protection of H_2N- group
② Conversion of the $-COOH$ group into $-COCl$
③ Formation of peptide linkage
④ Repeating of steps ② & ③ upto desired number of peptides formed
⑤ Elemination of protecting group, by reacting HBr.

The step 4 can be skipped when one wish to synthesized dipeptide.

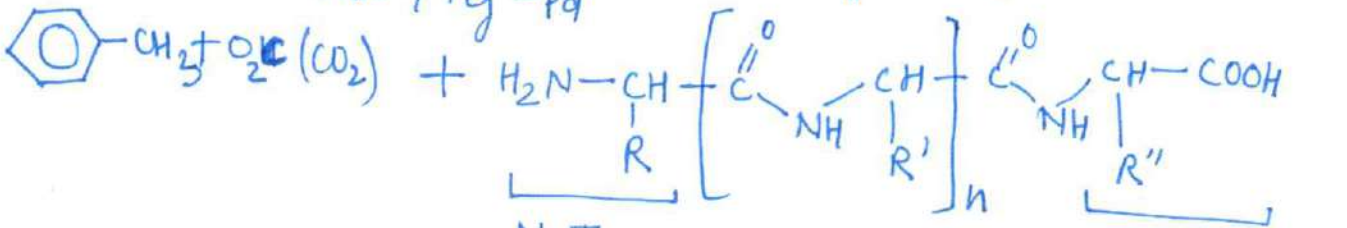
① Protection of amino group



④ Steps ② & ③ may be repeated According to number of peptide bonds

⑤ Elimination of Protecting group

$HBr / Hg-Pd$

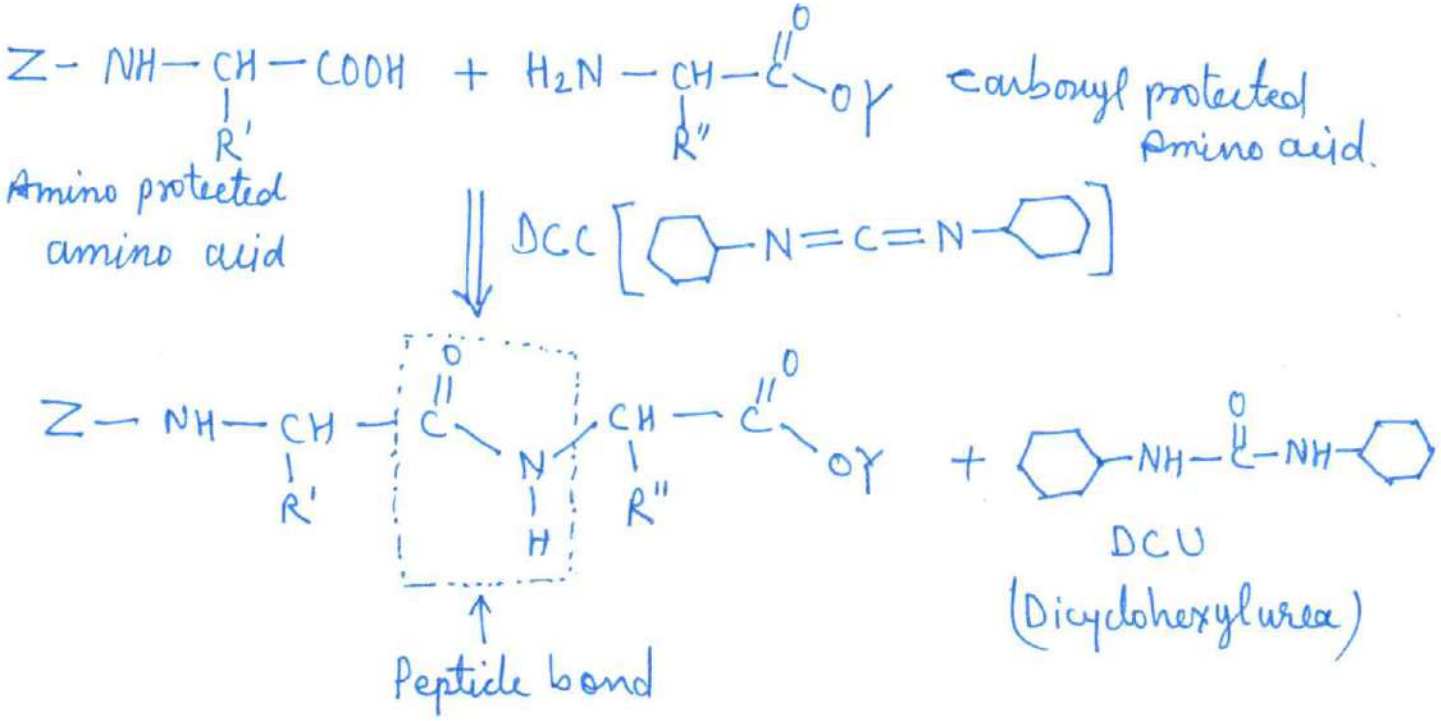


-COOH protecting Groups: Carboxyl group is generally protected by conversion to methyl, ethyl or benzylesters. After the peptide formation protecting groups are removed by hydrolysis

Use of DCC (Dicyclohexylcarbodiimide) as reagent in peptide bond formation.

The methods discussed above have certain drawbacks (racemisation of product). To over come, this problem, the use of DCC has been found with great advantage.

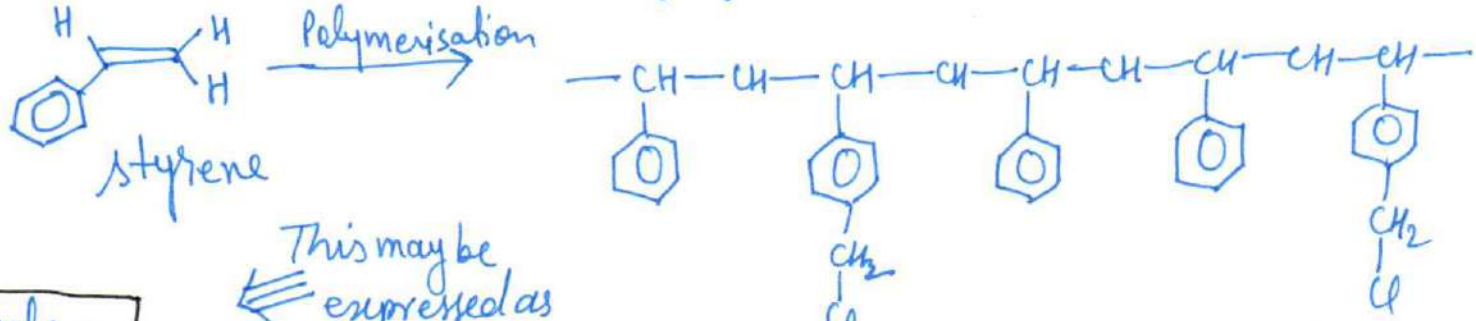
It was found that treatment of a solution containing an $-NH_2$ protected amino acid and $-COOH$ protected amino acid in presence of DCC, leads directly to peptide bond formation.



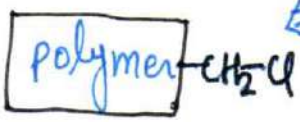
II Merrifield Solid Phase Peptide Synthesis

By using solid support, R.B. Merryfield develop a unique method of peptide synthesis, known as solid-phase peptide synthesis. Solid support used is p-chloromethylated polystyrene polymer

p-chloromethylation is done only on 2-5% of phenyl ring in polystyre



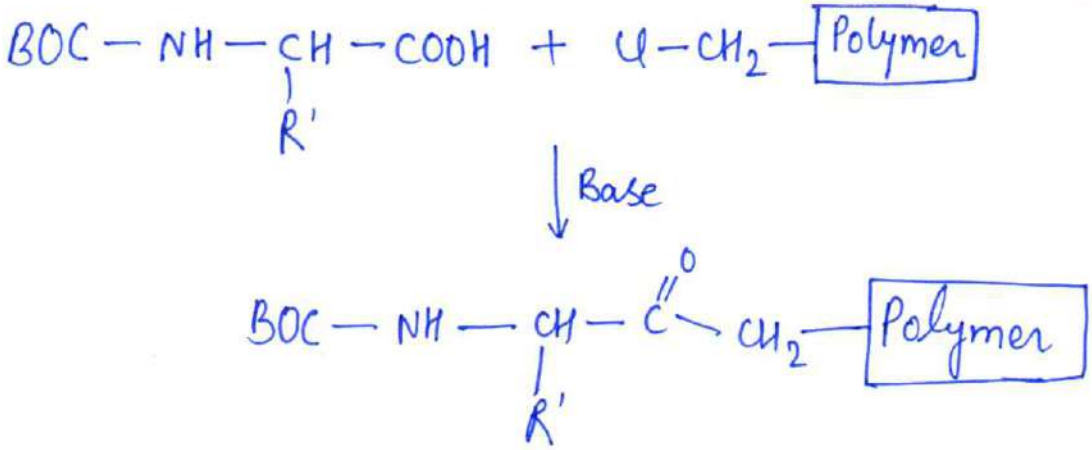
This may be expressed as



2-5%, p-chloromethylated derivated of polystyrene

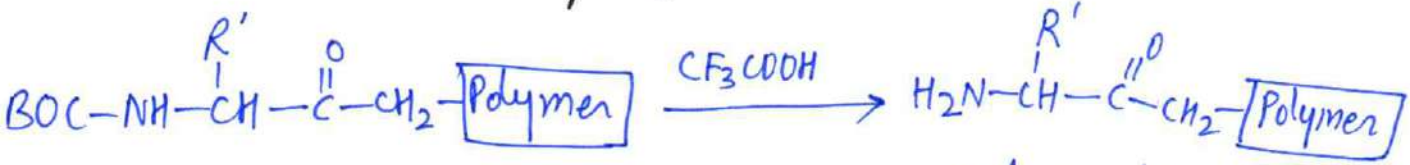
Merrifield solid phase peptide synthesis involves following 5 steps.

(i) **Protection of $-NH_2$ group by BOC [tert-butoxycarbonyl]**



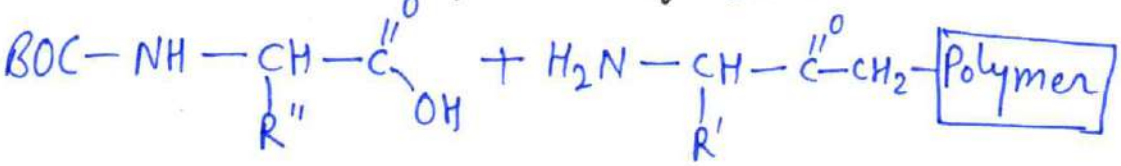
The excess reagent is removed by washing with suitable solvent.

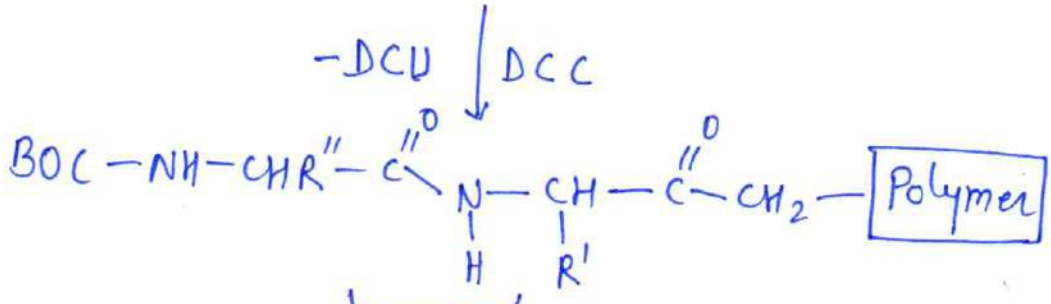
(ii) Removal of protecting group, BOC



polymer bound amino acid

(iii) Reaction of BOC protected amino acid with polymer linked amino acid in presence of DCC

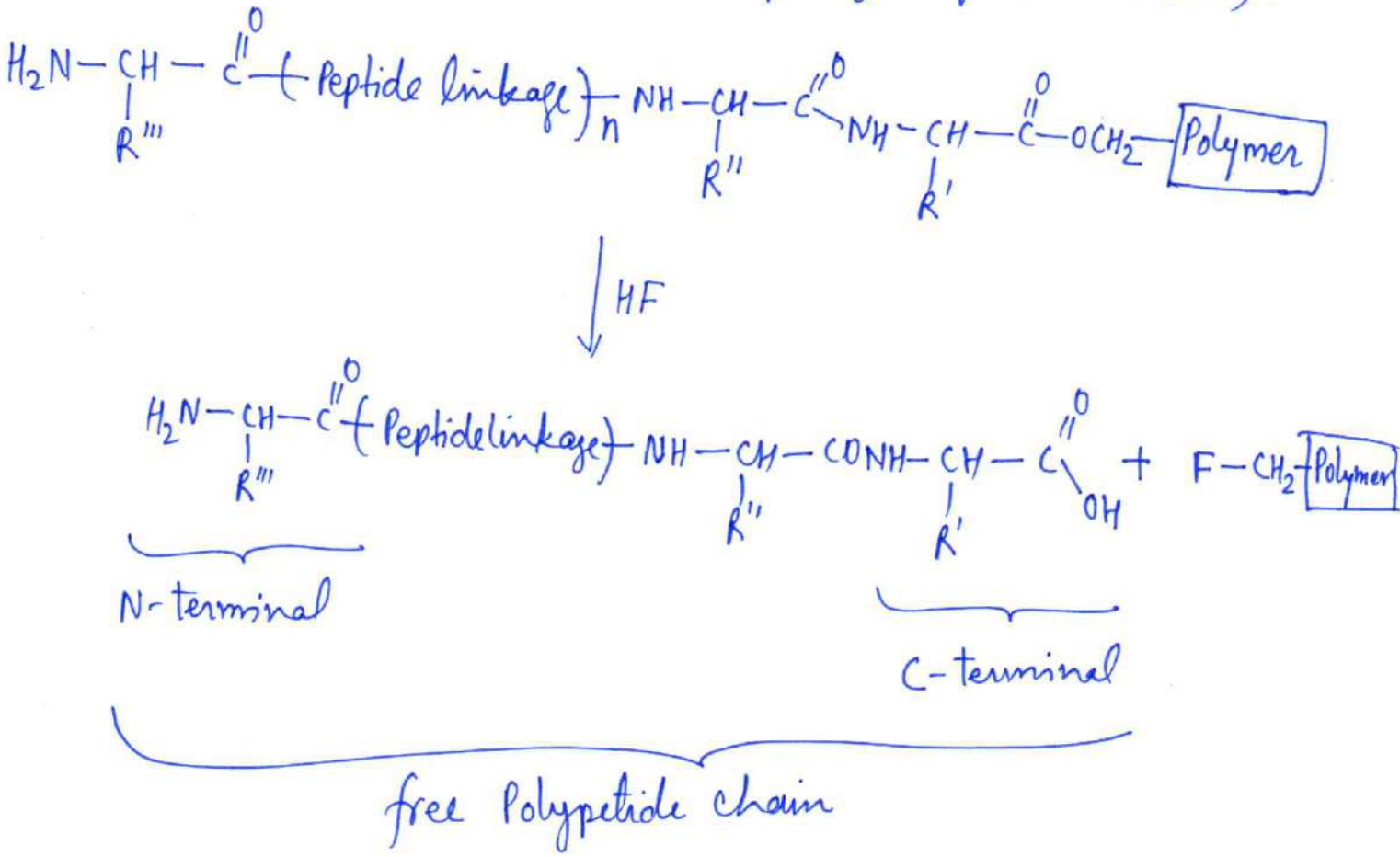




peptide bond formed.

(iv) Steps ② & ③ (ii & iii) are repeated to introduced as many amino acid units required to build the polypeptide chain

(v) The completed polypeptide is removed from the polymer by treatment with anhydrous hydrogen fluoride (HF).



Ribonucleage (an enzyme) which contains a ~~12~~ sequence of 124 amino acids and involves 369 chemical reactions and 11931 steps, has been successfully synthesised by this process.