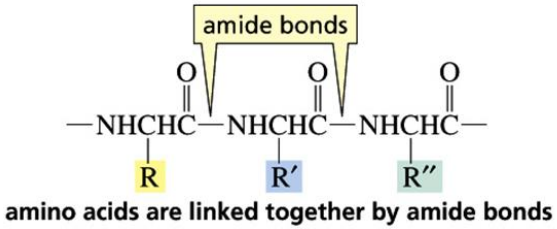
Notes on
Synthesis and structure of peptides and proteins

By

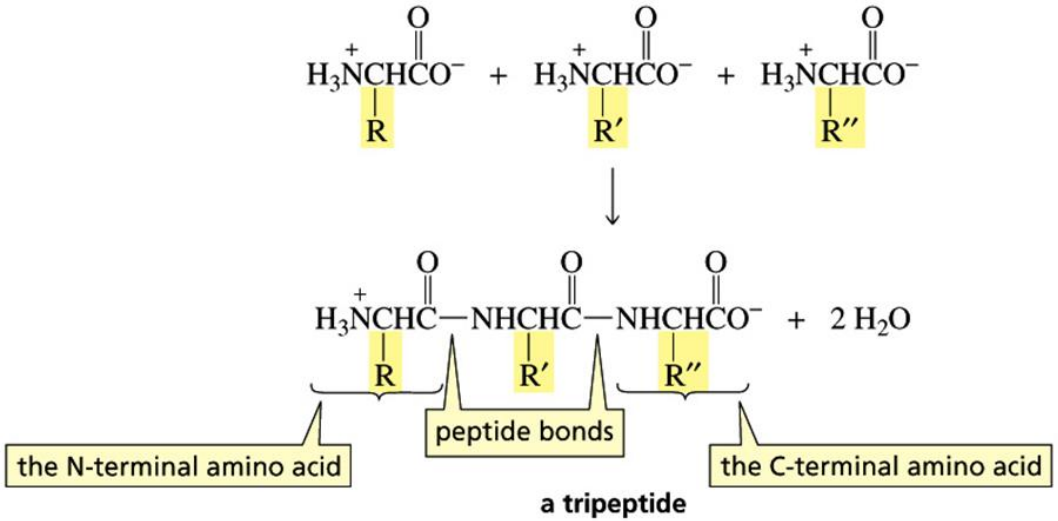
Dr. Gangutri Saikia

Peptides and proteins are polymers of amino acids linked together by amide bonds

Amino acid → **Peptides** → **Polypeptides** → **Proteins**



amino acids are linked together by amide bonds



a tripeptide

Structure of protein:

The primary structure of a protein is the sequence of amino acids in the chain. Also, the way in which the atoms of protein molecules are joined to one another by covalent bonds to form chains.

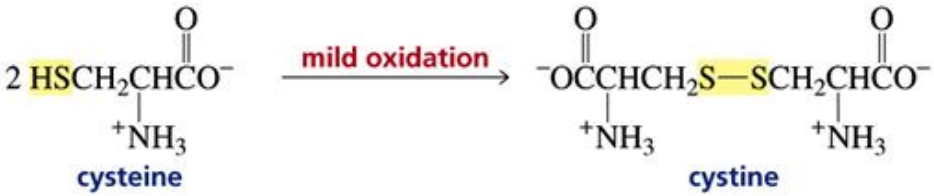
The secondary structure describes the regular conformation assumed by segments of protein's backbone.

The tertiary structure describes the three dimensional structure of the entire polypeptide.

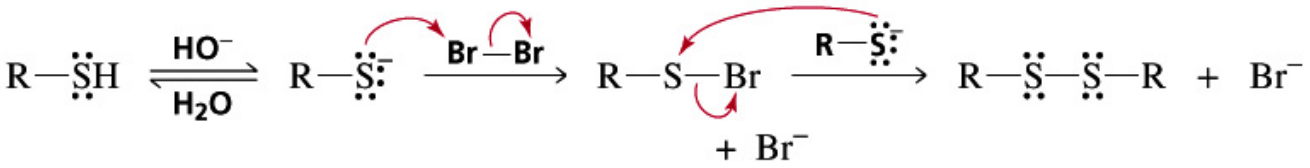
The quaternary structure of a protein is the way individual protein chains are arranged with respect to each other. If a protein has more than one polypeptide chain, it has quaternary structure.

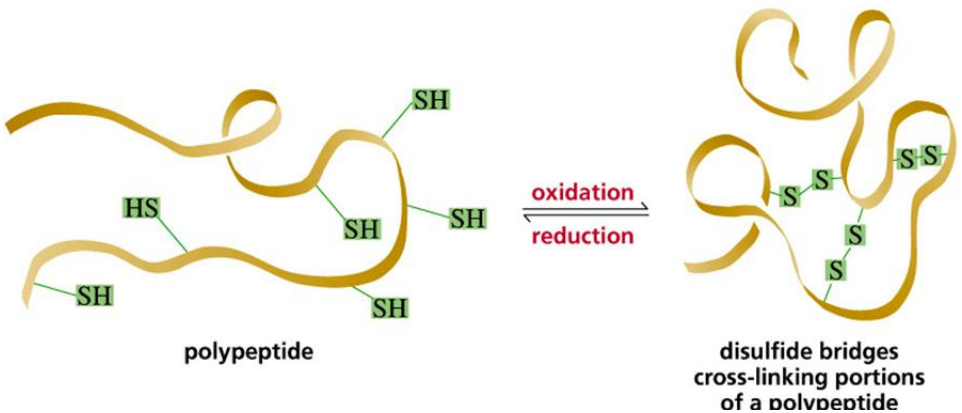
Various types of bonds and/or forces are responsible for the stabilisation of these protein structures other than the primary structure such as H-bond, electrostatic forces, inter and intramolecular chemical bonds, and hydrophobic bonds etc.

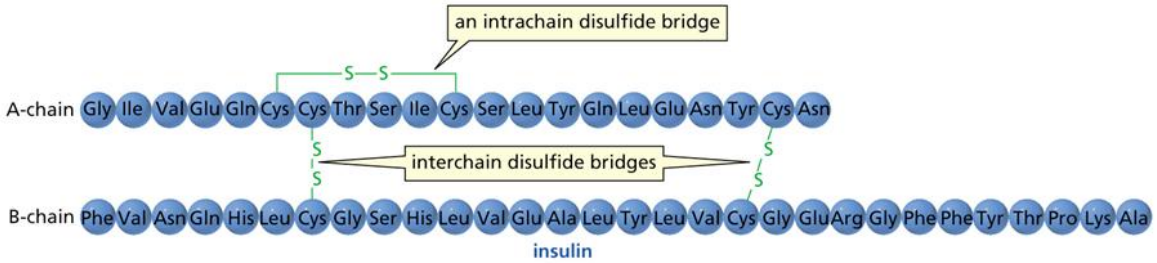
Disulfide bond



mechanism for oxidation of a thiol to a disulfide







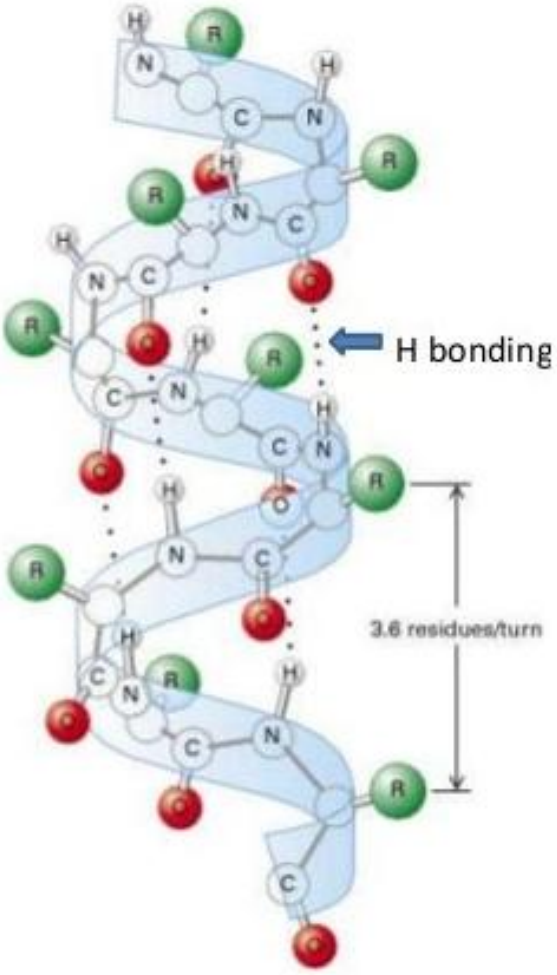
Secondary structure of protein:

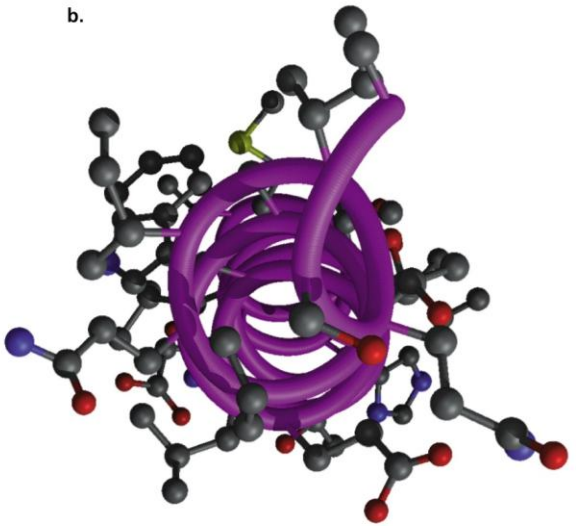
Three factors determine the choice of secondary structure:

- ❖ The regional planarity about each peptide bond, which limits the possible conformations of the peptide chain.
- ❖ Maximizing the number of peptide groups that engage in hydrogen bonding.
- ❖ Adequate separation between nearby to avoid steric hindrance and repulsion of like charges.

α-Helix:

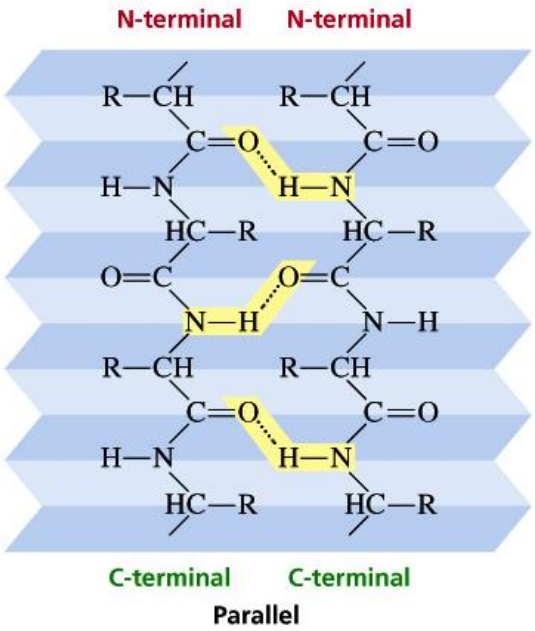
In α-helix, the backbone of the polypeptide coils around the long axis of the protein molecule. The helix is stabilized by hydrogen bonds- each hydrogen attached to an amide nitrogen is hydrogen bonded to a carbonyl oxygen of an amino acid four residues away. The substituents on the α-carbons of the amino acids protrude outward from the helix, thereby minimizing steric hindrance. Each turn of the helix contains 3.6 amino acid residues, and the repeat distance of the helix is 5.4 Å.

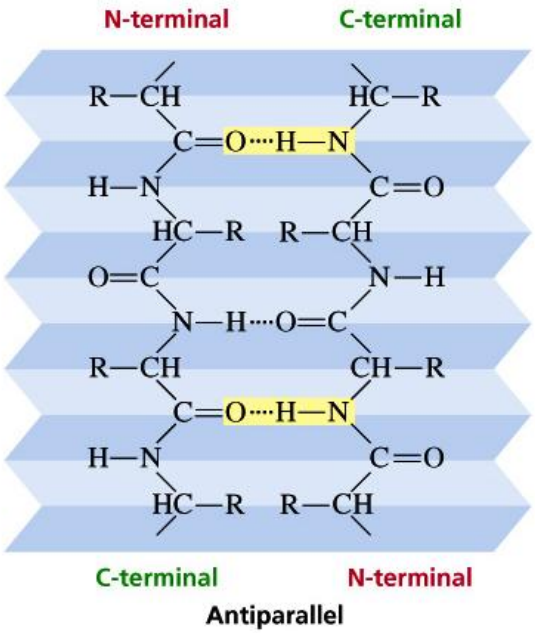




β-Pleated sheet:

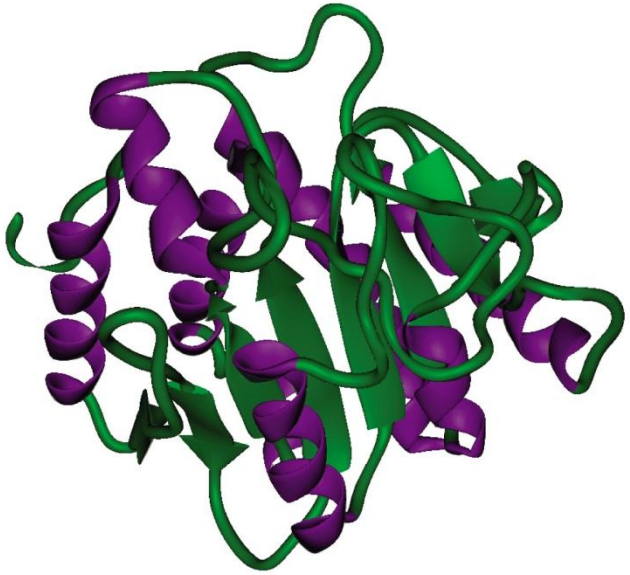
In β-pleated sheet, the polypeptide backbone is extended in a zigzag structure resembling a series of pleats. A β-pleated sheet is almost fully extended- the average two residue repeat distance is 7.0 Å. The hydrogen bonding in a β-pleated sheet occurs between neighbouring peptide chains. The adjacent hydrogen bonded peptide chains can run in the same or opposite direction.



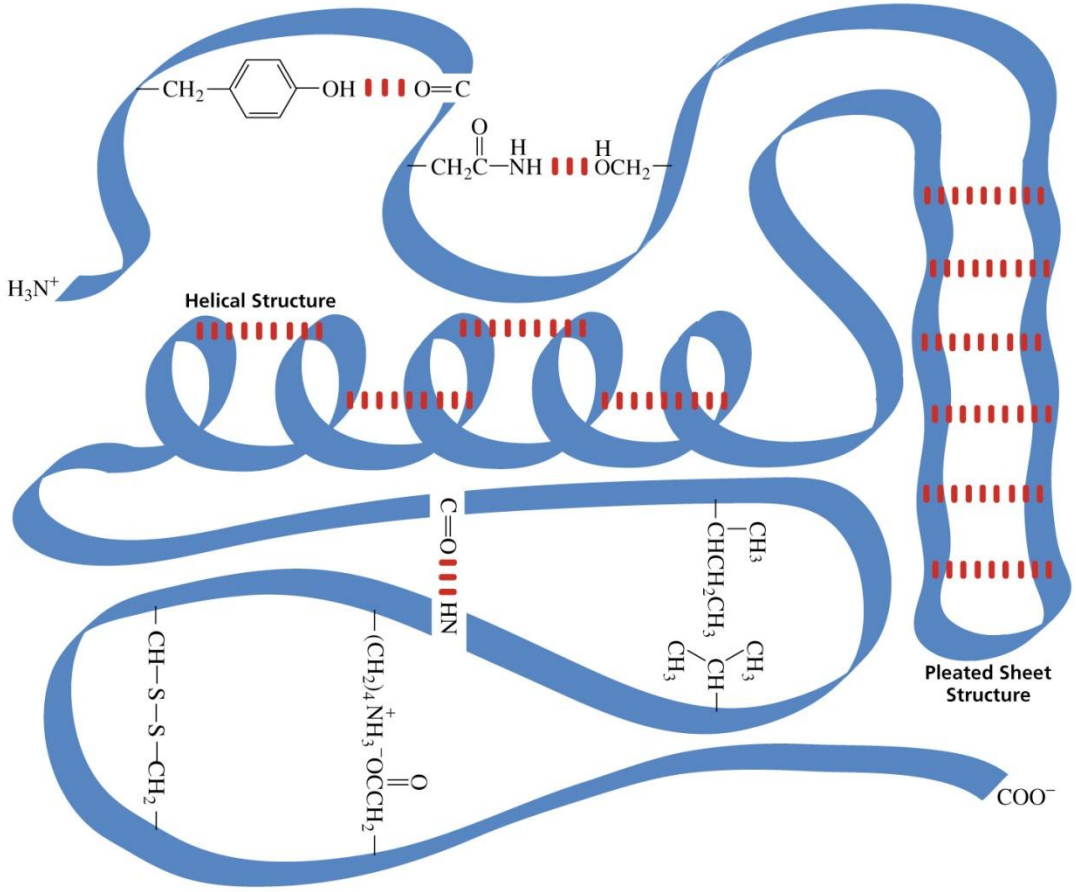


Coil conformation:

The coil conformation is unlike α-helix, very flexible and change from one form into another can be effected by change in pH or temperature.



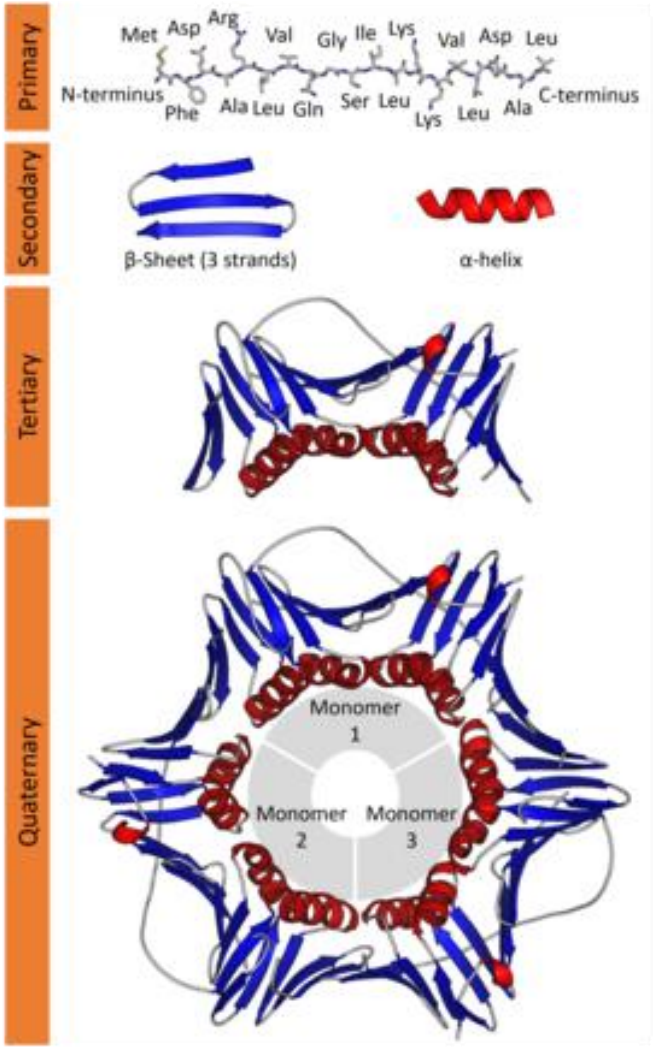
Tertiary structure of protein:



Proteins fold spontaneously in solution in order to maximize their stability. The stabilizing interaction that occur in folding are covalent bonds, hydrogen bonds, electrostatic interactions and hydrophobic interactions. The interactions can occur between peptide groups, between side-chain groups, and between peptide and side-chain groups. Disulfide bonds are the only covalent bonds that can form when protein folds. Most proteins exist in aqueous environment. Therefore, they tend to fold in a way that exposes the maximum number of polar groups to the aqueous environment and that buries the nonpolar groups in the interior of the protein, away from water.

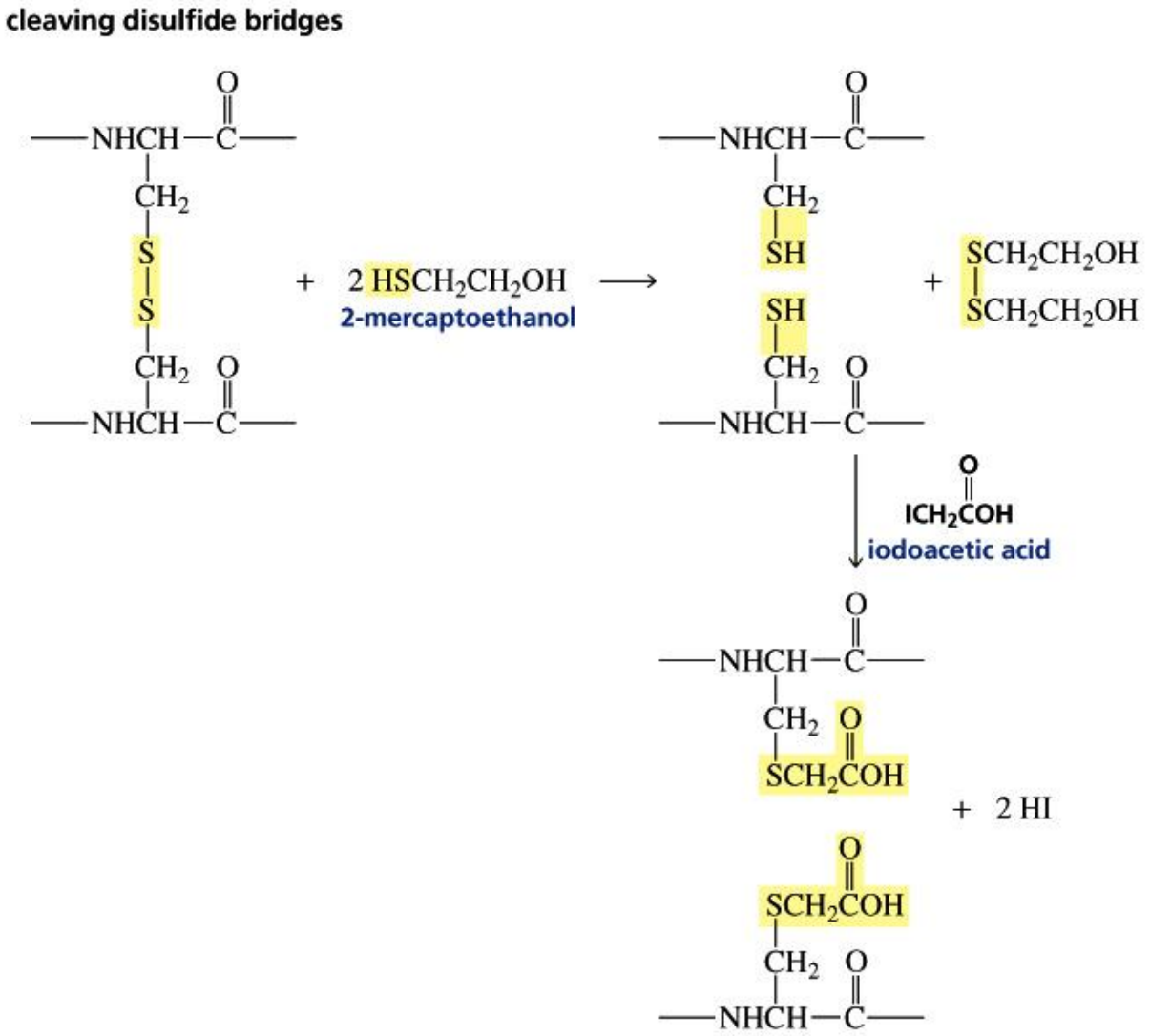
Quaternary structure of protein:

Proteins that have more than one peptide chain are called oligomers. The individual chains are called subunits. The subunits are held together by the same kinds of interactions that hold the individual protein chains in a particular three dimensional conformation- hydrophobic interactions, hydrogen bonding and electrostatic attractions. The quaternary structure of a protein describes the way the subunits are arranged in space.

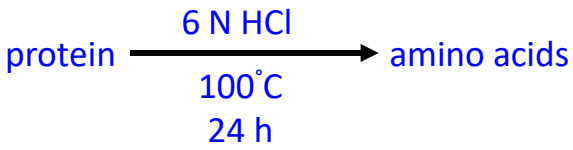


Determination of the primary structure of protein:

The first step in determining the sequence of amino acids in a peptide or a protein is to reduce any disulfide bridges in a protein. A commonly used reducing agent for this purpose is 2-mercaptoethanol, which is oxidized to a disulfide. To prevent the disulfide bridges from reforming, the thiol group of the protein reacts with iodoacetic acid.



The next step is to determine the number and kinds of amino acids in the peptide or protein



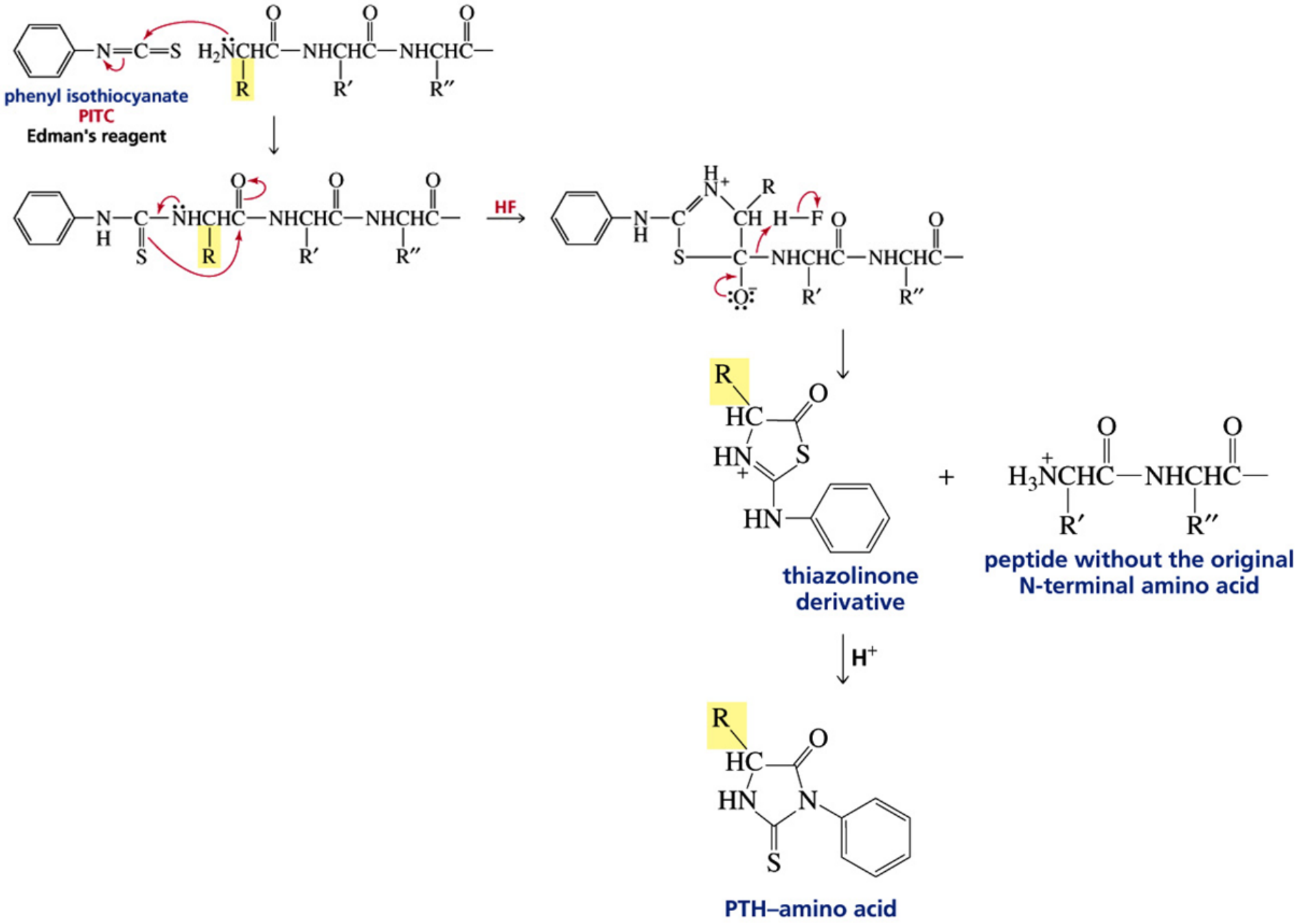
The mixture of amino acids is then passed through an amino acid analyzer to determine the number and kind of each amino acid in the peptide or protein.

This is not a general method for peptide bond hydrolysis because strongly acidic or alkaline conditions destroy several amino acid residues. There are several ways to identify N-terminal or C-terminal amino acid of peptide or a protein.

N-terminal amino acid determination. The Edman method:

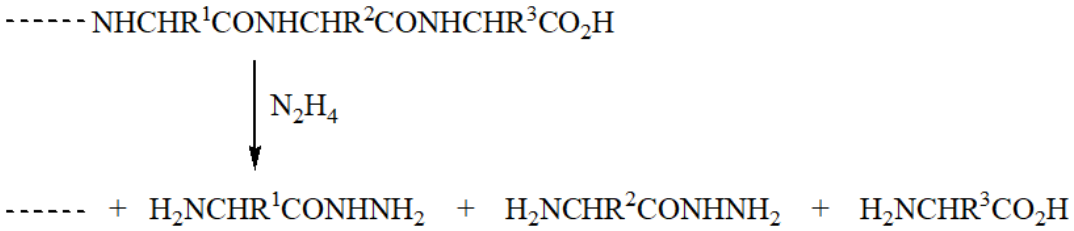
The basis of this method is the reaction between phenyl isothiocyanate (PITC) and the peptide or protein. Phenyl isothiocyanate reagent is more commonly known as Edman's reagent. It reacts with N-terminal amino group and the resulting thiazolinone derivative is cleaved from the protein under mildly acidic conditions. The thiazolinone derivative is extracted into an organic solvent and in presence of acid, rearranges to a more stable phenythiohydantoin (PTH).

Because each amino acid has a different substituent (R), each amino acid forms a different PTH. The particular PTH can be identified by chromatography using known standards. Several successive Edman degradations can be carried out on a protein.


phenyl isothiocyanate
PITC
Edman's reagent
HF
thiazolinone
derivative
peptide without the original
N-terminal amino acid
H+
PTH–amino acid

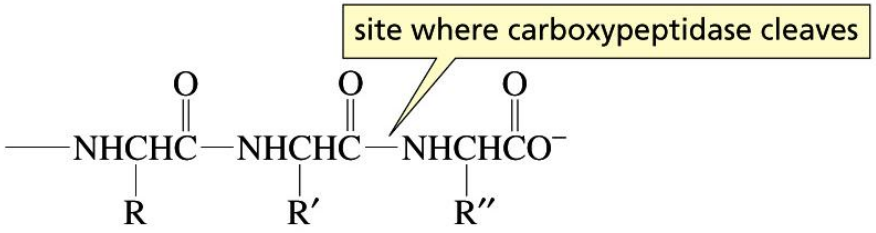
C-terminal amino acid determination:

The most widely used method is hydrazinolysis of protein. The peptide is heated with anhydrous hydrazine at 100 ºC. This converts all amino acid residues except the C-terminal one into amino acid hydrazides.



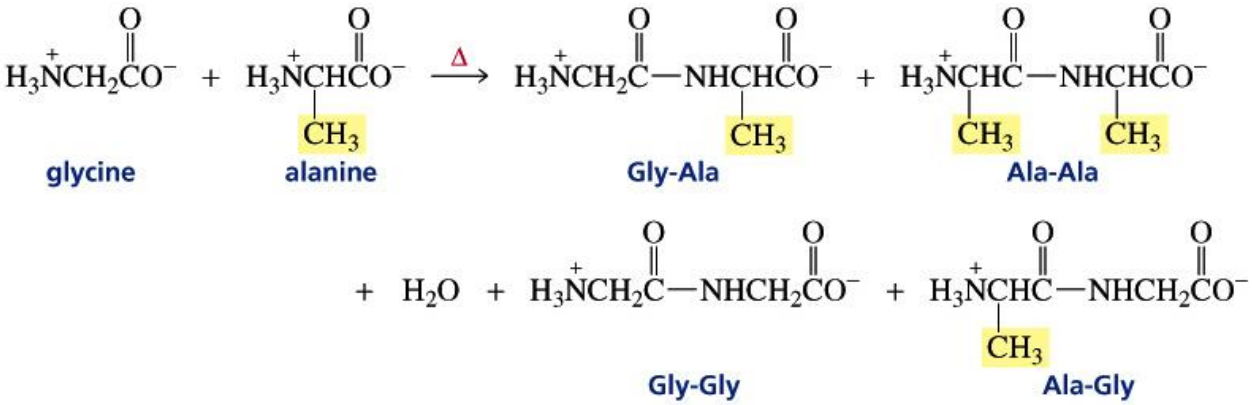
The mixture of products is subjected to chromatography on a column of a strong cation exchange resin. On elution, the strongly basic hydrazides are retained, but the free amino acid is eluted and can be identified.

Another widely used method that makes use of the enzyme carboxypeptidase. Carboxypeptidases are exopeptidases that catalyzes the hydrolysis of a peptide bond at the end of a peptide chain. This enzyme attacks proteins only at the end which contains free α-carboxyl group. When this terminal amino acid residue is liberated, the new terminal free carboxyl group is attacked by the enzyme. Thus, after a given time of hydrolysis, a number of successive terminal amino acids will have been liberated. Thus, by identification and quantitative determination of the amino acids, their sequence can be established.



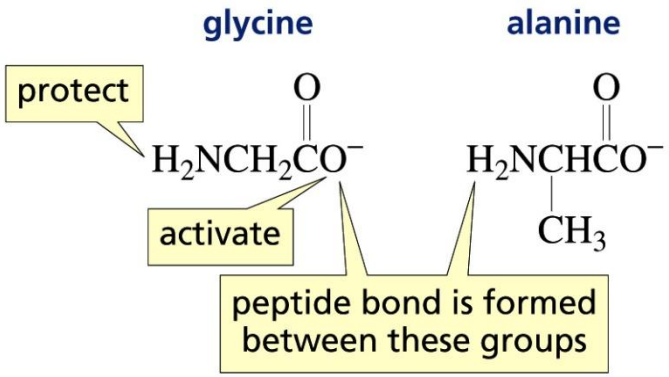
Synthesis of simple peptides:

Because amino acids have two functional groups, a problem arises when one attempts to make a particular peptide.

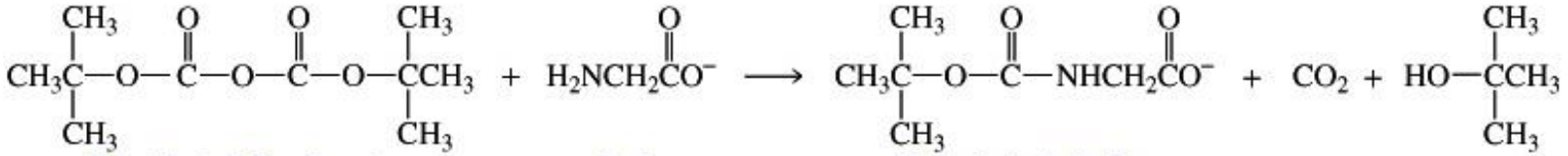


If the amino group of the amino acid that is to be on the N-terminal end is protected, it will not be available to form a peptide bond. If the carboxyl group of the same amino acid is activated before the second amino acid is added, the amino group of the added amino acid will react with the activated carboxyl group of glycine in preference to reacting with a nonactivated carboxyl group of another alanine molecule.

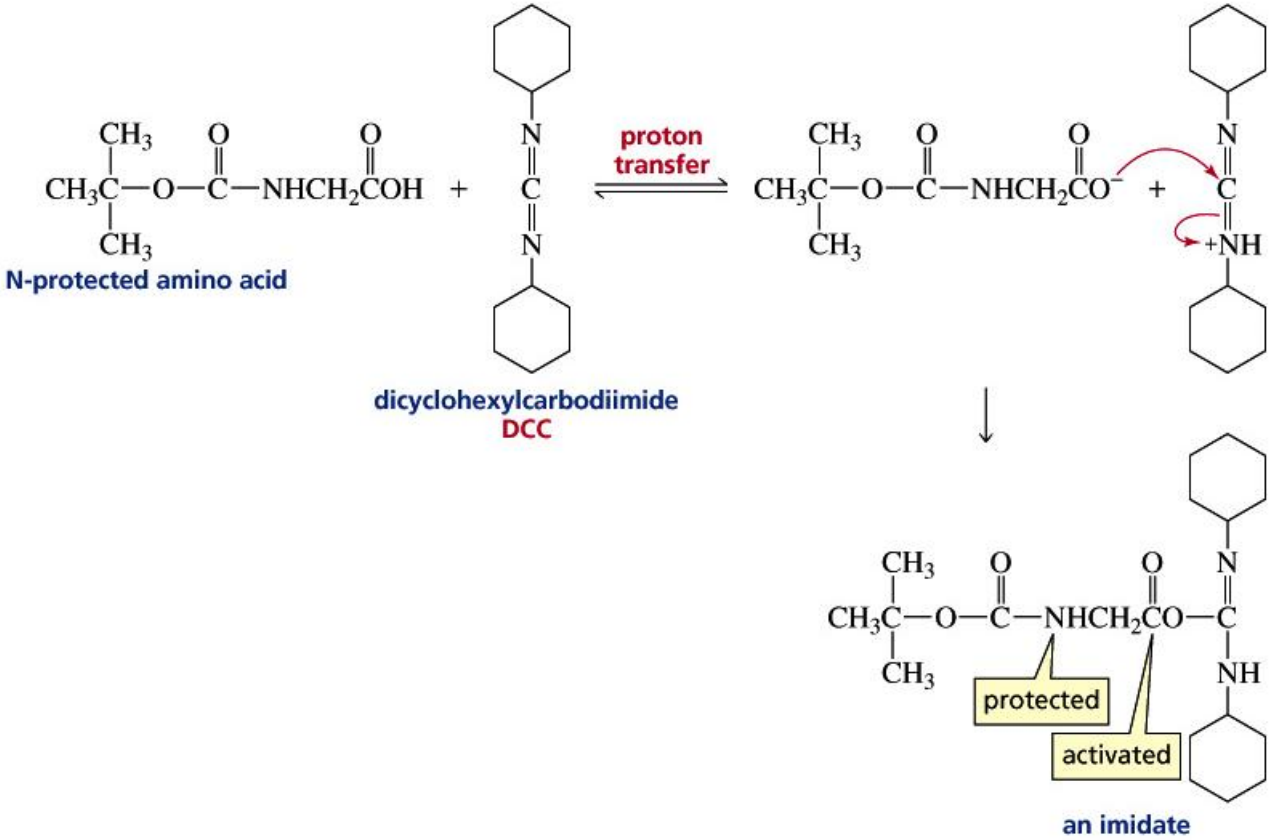
Strategy for Making a Specific Peptide Bond



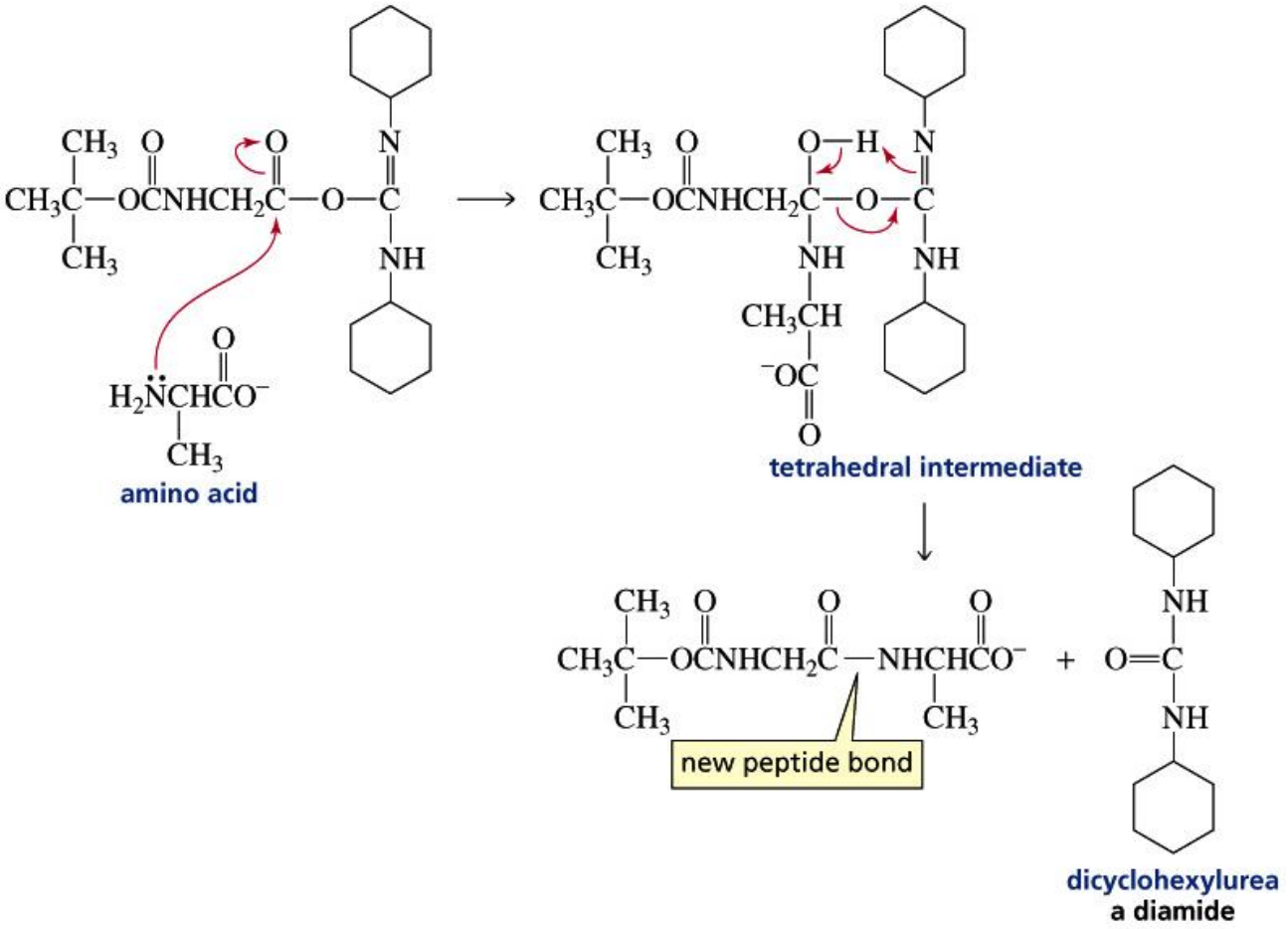
The reagent most often used to protect the amino group of an amino acid is di-tertbutyl dicarbonate (t-BOC). Its popularity is due to the ease with which the protecting group can be removed when the need for protection is over.



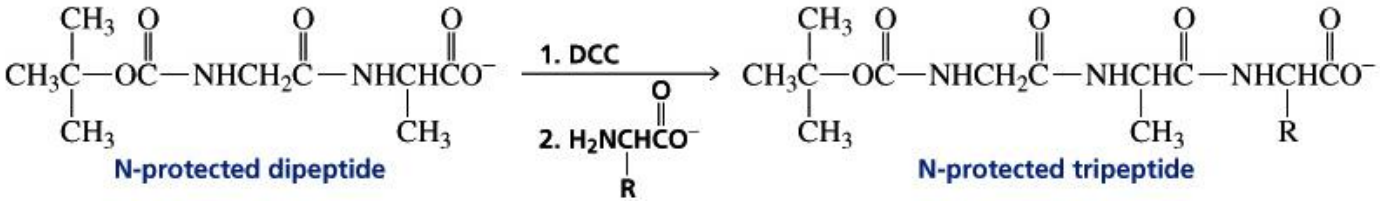
The preferred method for activating the carboxyl group of an N-protected amino acid is to convert it into an imidate by using dicyclohexylcarbodiimide (DCC). DCC activates a carboxyl group by putting a good leaving group on the carbonyl carbon.



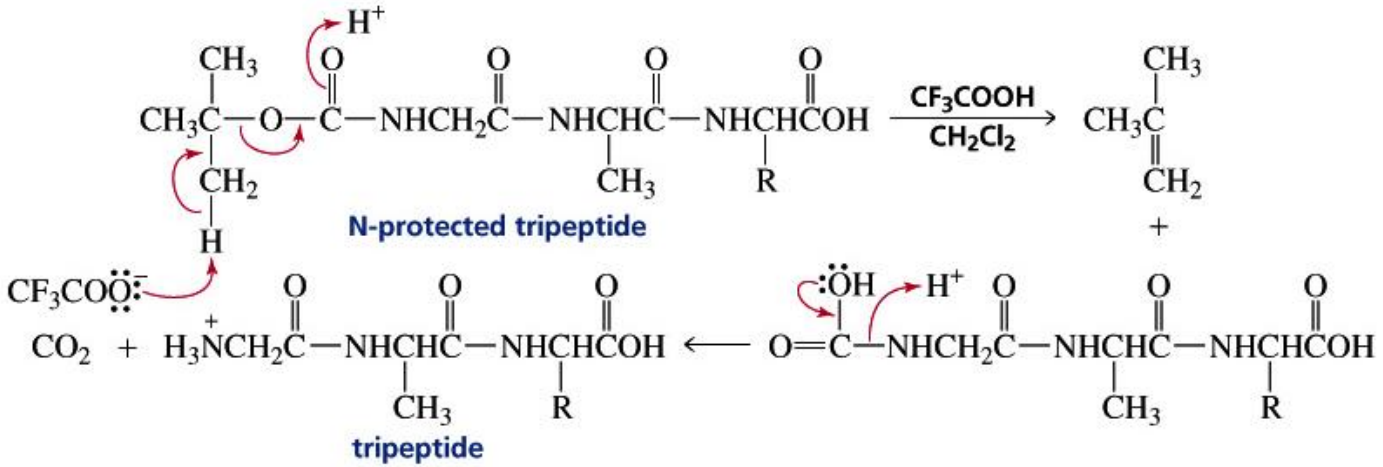
After the amino acid has its N-terminal group protected and C-terminal group activated, the second amino acid is added to form the new peptide bond. The C-O bond of the tetrahedral intermediate is easily broken, and dicylcohexylurea is formed as a stable diamide.



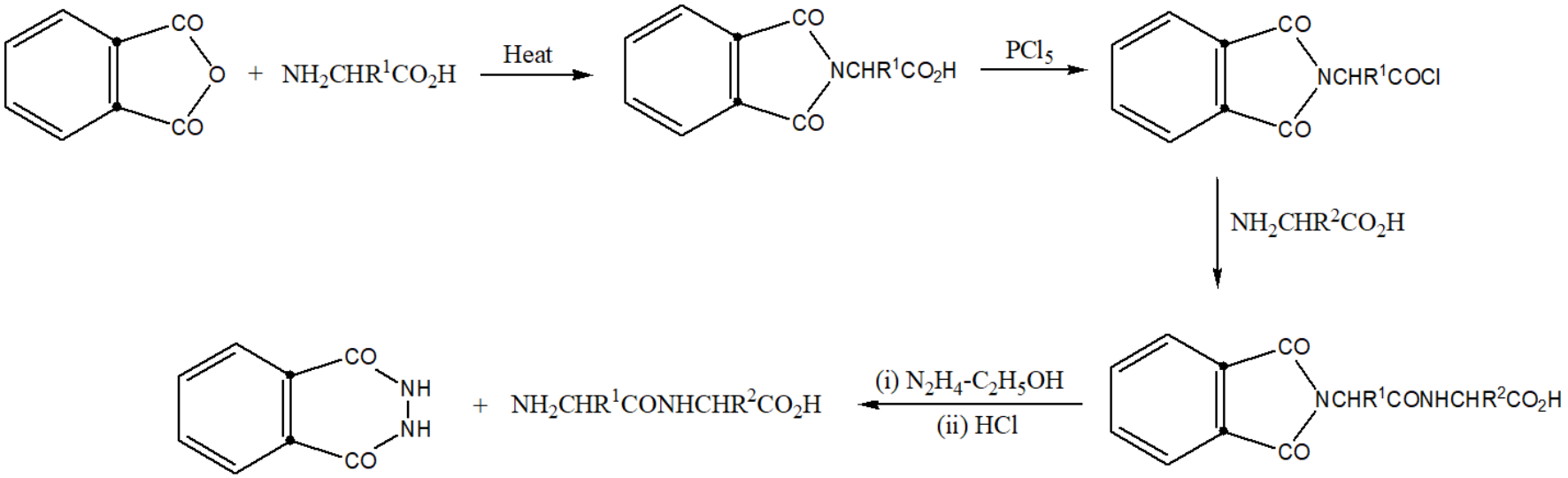
Amino acids can be added to the growing C-terminal end by repeating these two steps.



When the desired number of amino acids has been added to the chain, the protecting group on the N-terminal amino acid is removed. t-BOC is an ideal protecting group because it can be removed by washing with trifluoroacetic acid and methylene chloride. The protecting group is removed by an elimination reaction, forming isobutylene and carbon dioxide. Beause, these products are gases, they escape, driving the reaction to completion.



Sheehan et al. have used phthaloyl group as a means of protecting an end amino group.

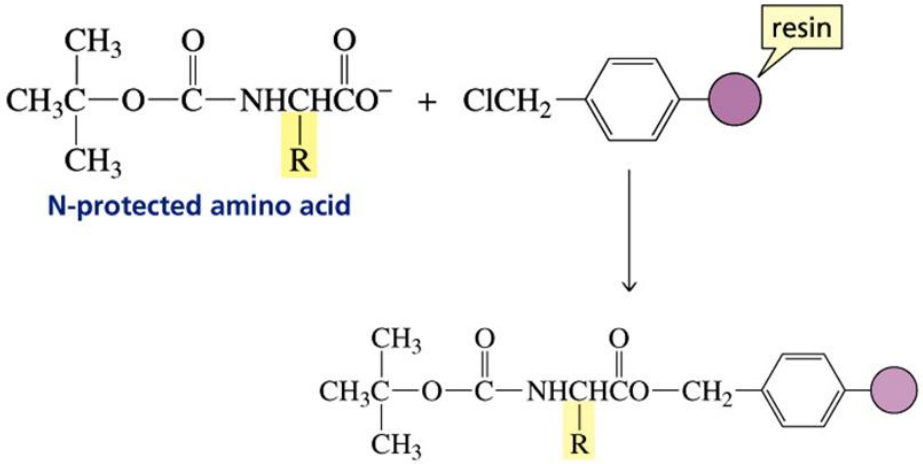


Solid phase peptide synthesis :

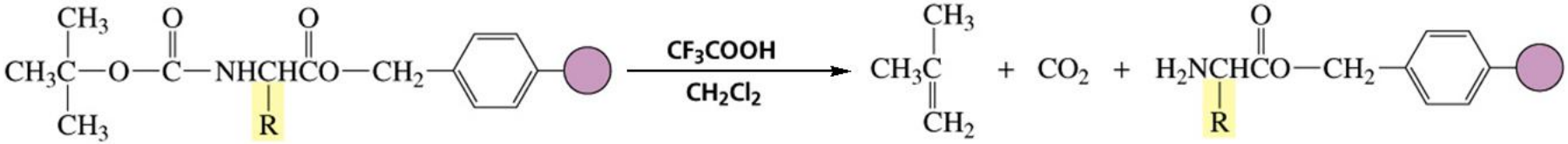
In 1969, Bruce Merrifield described a method that revolutionized the synthesis of peptides because it provides a much faster way to produce peptides in much higher yields. In this method, an amino acid or peptide is bound chemically to an insoluble synthetic resin and then the chain is built up, one amino acid residue at a time, at the free end. When the desired peptide has been synthesized, it is liberated from the solid support. The method has been automated, *i.e.*, each addition of the appropriate amino acid is carried out automatically at a predetermined time. Therefore, this method of protein synthesis is called automated solid phase peptide synthesis.

Some outstanding advantages of this solid phase method are: (i) because of the use of insoluble solid support, purification of products is not necessary, excess of reagents being removed by thorough washing with suitable solvents; (ii) high yields; (iii) shorter amount of time.

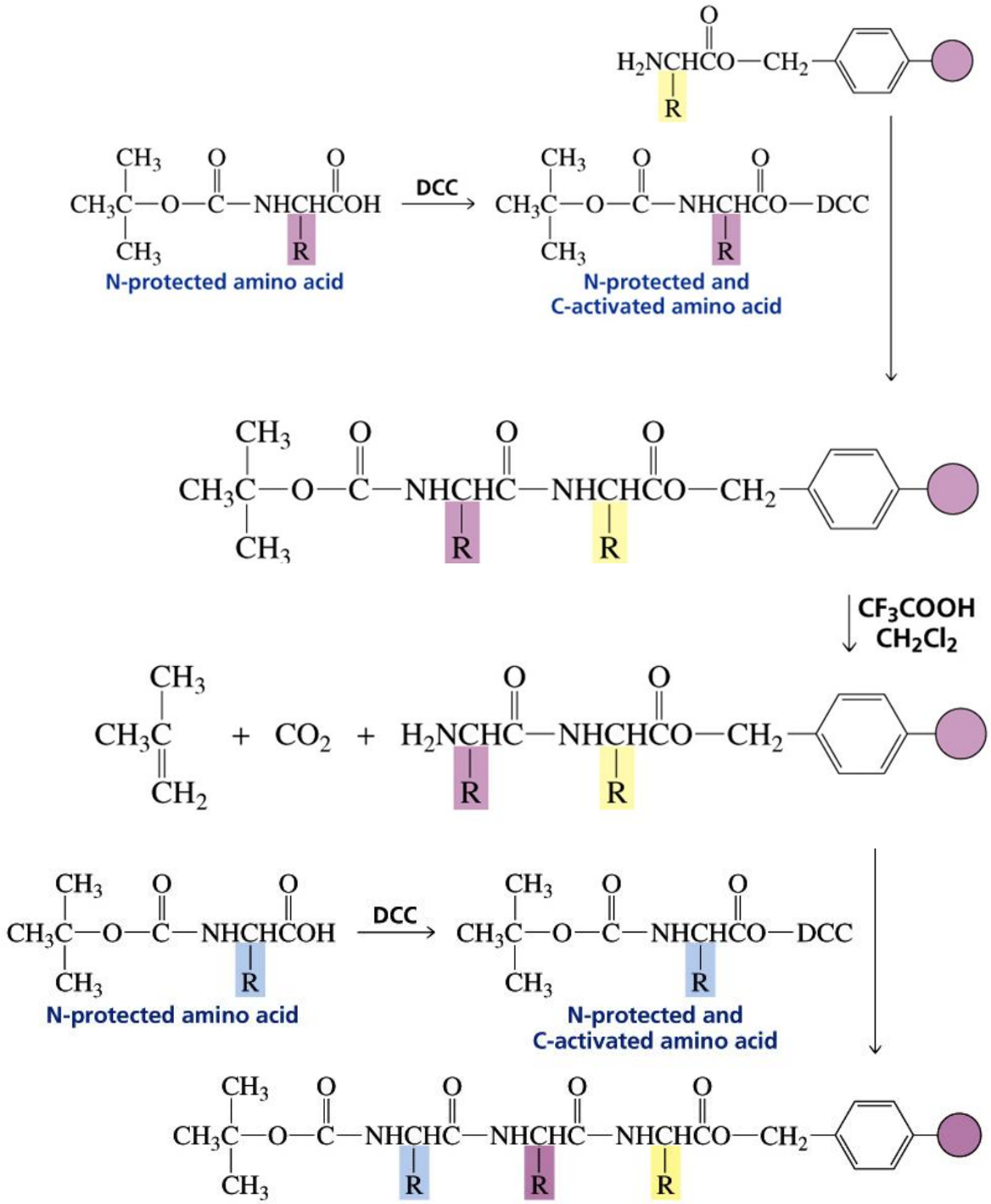
The solid support is a copolymer of styrene and divinylbenzene and is chloromethylated. This results in the formation of benzyl chloride groups through which the first amino acid becomes attached as benzyl ester. This first amino acid which is the C-terminal end of the peptide, is protected at its N-terminal end by t-butyloxycarbonyl group (t-BOC) to prevent the amino group from reacting with the resin. The C-terminal amino acid is attached to the resin by means of an SN2 reaction- its carboxyl group attacks a benzyl carbon of the resin, displacing a chloride ion.

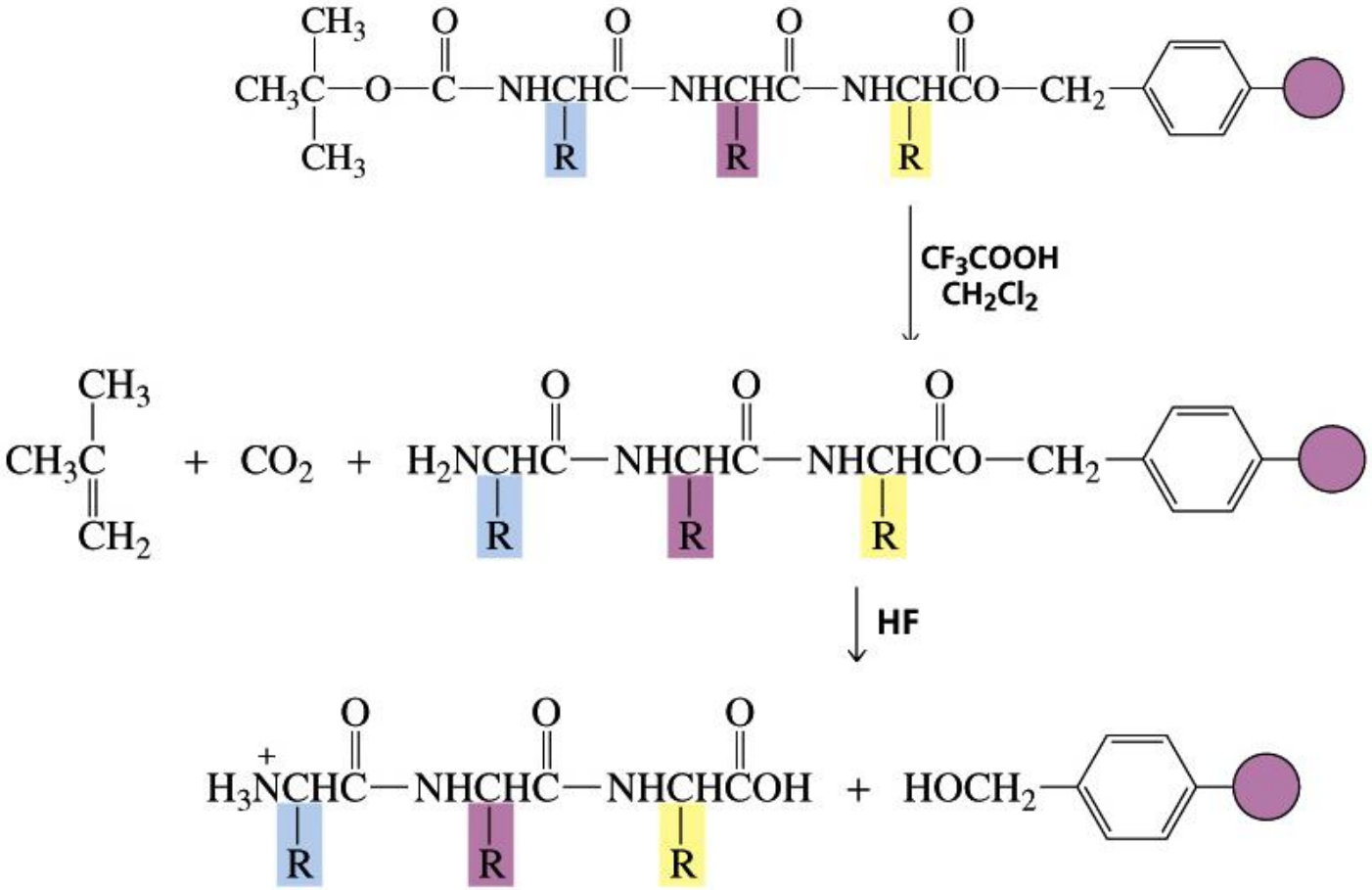


After C-terminal amino acid is attached to the resin, t-BOC protecting group is removed by washing with trifluroacetic acid and methylene chloride.



The next amino acid with its amino group protected with t-BOC and its carboxyl group activated with DCC, is added to the column. The cycle is then repeated with the N-protected third amino acid, and so on. When the desired peptide has been synthesized, the peptide can be removed from the resin by treatment with HF under mild conditions.


DCC
N-protected amino acid
N-protected and
C-activated amino acid
CF3COOH
CH2Cl2
DCC
N-protected amino acid
N-protected and
C-activated amino acid


CF3COOH
CH2Cl2
HF