28.8 DETERMINATION OF SEQUENCE OF AMINO ACIDS IN A GIVEN POLYPEPTIDE

One of the most challenging works in the study of amino acids is to determine the constituent amino acids in a polypeptide and the sequence in which these are arranged. The determination of sequence of amino acids in polypeptide involves the following steps:

(1) A polypeptide (protein) is isolated and purified. Purification of a polypeptide may be carried out by different techniques, for example, dialysis, gel filtration chromatography, ion-exchange chromatography, electrophoresis, and so on. The electrophoresis is highly efficient and powerful technique for separation and purification.

(2) The constituent amino acids of a polypeptide are determined by hydrolysis of polypeptide chain. The entire chain is degraded by hydrolysis in the presence of an acid (6N HCl) and this results in the formation of a mixture of amino acids.

(3) The qualitative and quantitative estimation of amino acids is carried out with the help of an automated amino acid analyzer. The basic principle involved here is ion-exchange chromatography. In the process, the acidic amino acids are released first and the most basic amino acid are released in the end.

(4) The next step involves the identification of N- and C-terminals, that is, ***end group analysis***.

(5) Once the end groups are identified, stepwise degradation of peptide chain is carried out which helps in determining the order of amino acids (sequence) in a chain. The process involves the non-enzymatic hydrolysis of polypeptide into smaller fragments.

The last two steps, namely end group analysis and sequence determination of polypeptides are now discussed in detail.

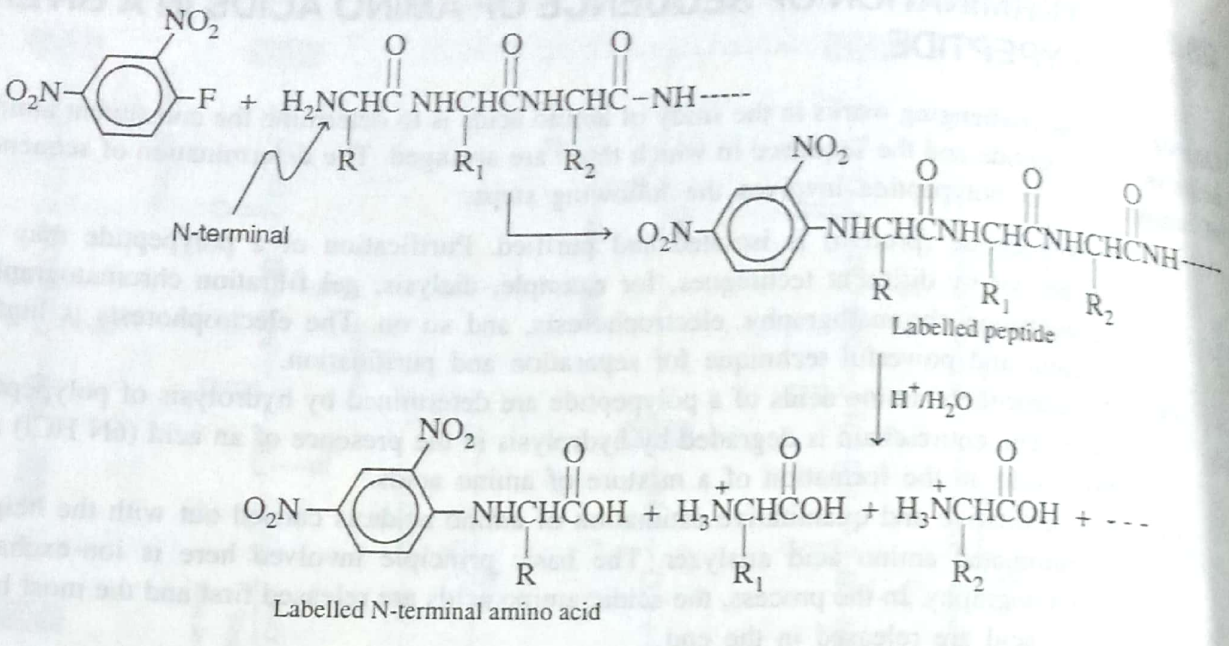
28.8.1 End Group Analysis

The N-terminal contains a free amino group and the C-terminal contains a free carboxylic group. Thus, amino acids present at two terminals are entirely different from each other as well as from other amino acids present in a chain.

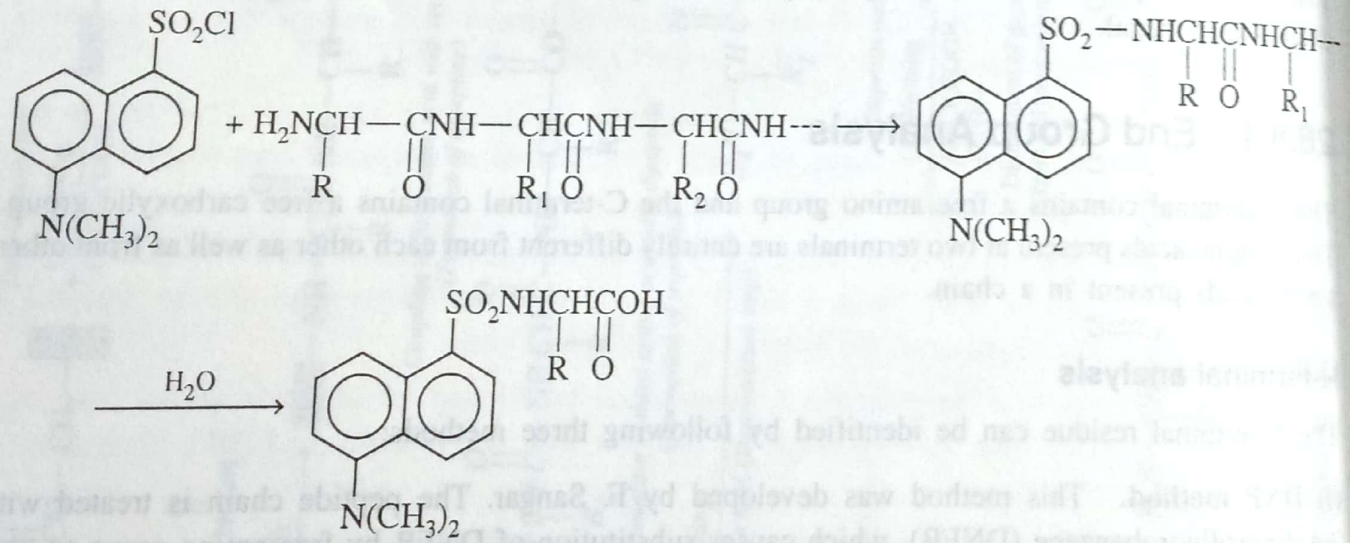
N-terminal analysis

The N-terminal residue can be identified by following three methods:

(i) DNP method. This method was developed by F. Sangar. The peptide chain is treated with 2,4-dinitrofluorobenzene (DNFB), which causes substitution of DNFB by free amino group to give a N-Dinitrophenyl (DNP) derivative. The hydrolysis of peptide chain results in the separation of labelled N-terminal amino acid (DNP-amino acid) and a mixture of another amino acids. Thus, DNP-amino acid so separated is identified to determine the N-terminal.



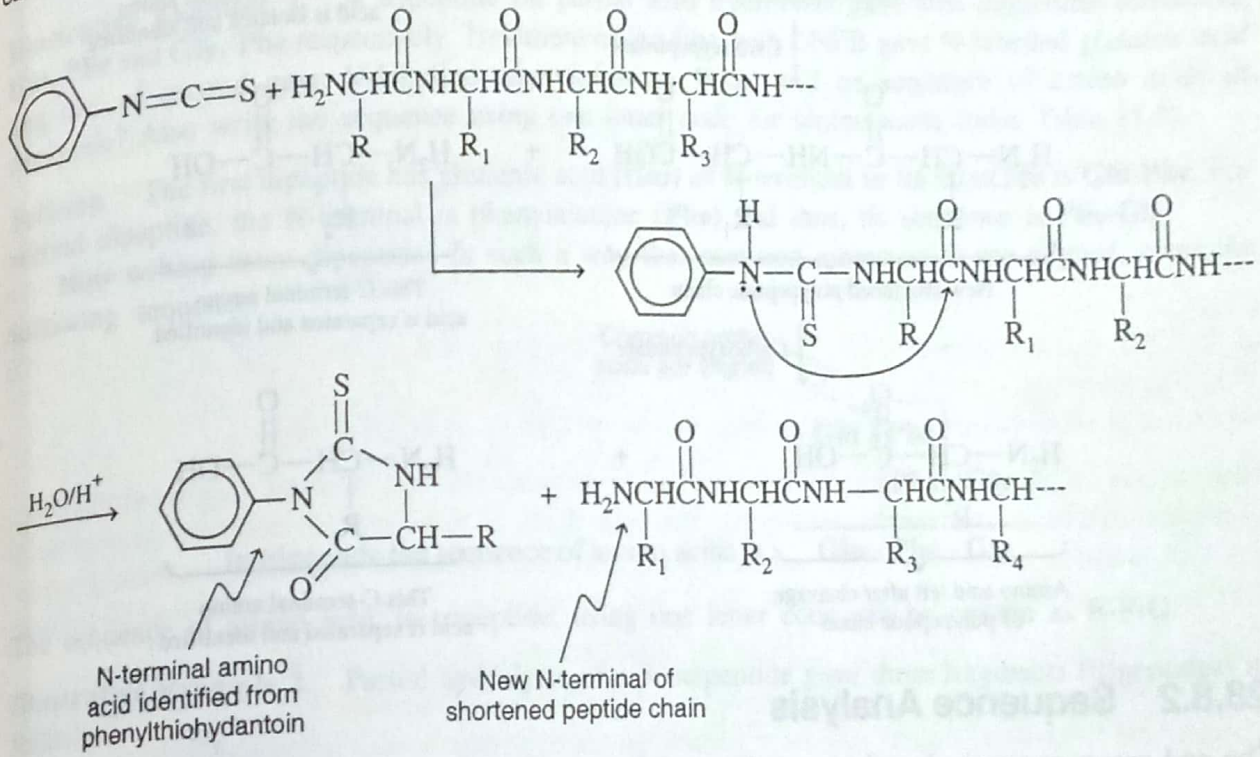
(ii) Dansyl method. In this method the reaction of peptide chain with 5-dimethylaminonaphthalene-1-sulfonylchloride (Dansyl chloride) occurs at the N-terminal (amino end). The hydrolysis of peptide chain results in the formation of Dansyl amino acid. This Densyl amino acid on irradiation with UV light shows fluorescence. Thus, the N-terminal is identified. This method is 100 times more sensitive than DNP method in detection of the N-terminal.



(iii) Edman degradation. One of the most important and widely used method which depends on the selective removal of N-terminal amino acids from polypeptide chain (proteins) is the *Edman degradation*. The method removes N-terminal amino acid and leaves the rest of the polypeptide chain intact. Thus, the method can be used repeatedly to identify the N-terminal of shortened peptide chain.

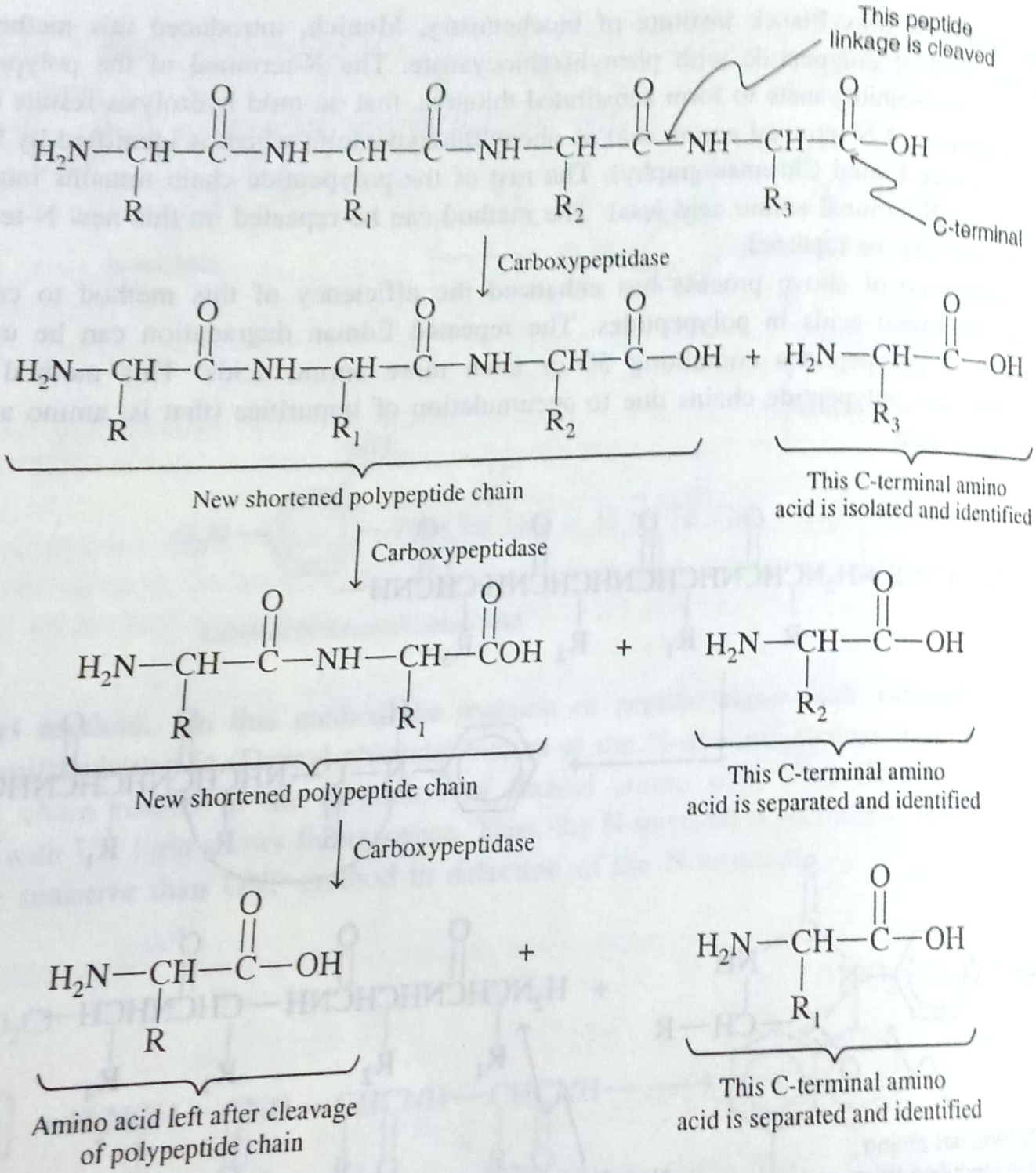
Pehr Edman of Max Planck institute of biochemistry, Munich, introduced this method. It involves treatment of polypeptide with phenylisothiocyanate. The N-terminal of the polypeptide reacts with phenylisothiocyanate to form substituted-thiourea, that on mild hydrolysis results in the selective separation of N-terminal amino acid as phenylthiohydantoin, which is identified by HPLC (High Performance Liquid Chromatography). The rest of the polypeptide chain remains intact (of course with one N-terminal amino acid less). The method can be repeated on this new N-terminal and the process may be repeated.

The automation of above process has enhanced the efficiency of this method to continue sequencing the amino acids in polypeptides. The repeated Edman degradation can be used for identification of polypeptides containing 50 or even more amino acids. This method is not applicable for high polypeptide chains due to accumulation of impurities (that is, amino acids) in each hydrolysis step.



C-terminal analysis

The C-terminal analysis is carried out even more efficiently by enzymatic methods rather than chemical methods. The enzyme *carboxypeptidase*, cleaves selectively the peptide linkage adjacent to C-terminal (free carboxylic group) in polypeptide chain. The removal of C-terminal residue results in the formation of shortened polypeptide chain, which can be treated with the *carboxypeptidase* enzyme to determine the new C-terminals.



.2 Sequence Analysis

nd group or terminal analysis can be used for selective removal of terminal groups. However, difficult process to identify a polypeptide chain by stepwise continuous removal of the al residues. The sequence of amino acids is identified by carrying out partial hydrolysis of ypeptide chain. Partial hydrolysis breaks the chain selectively into smaller fragments that can tified further. Specific enzymes are used which cleave the polypeptide at specific sites only. aller fragments can then be identified by Edman degradation.