A NOTE ON
PROTEIN SEQUENCING

Submitted BY

Dr. Namrata Kahar

(Guest Faculty)

Department of Biotechnology

Govt. Digvijay Autonomous College Rajnandgaon

(C.G.) 491441, INDIA

SYNOPSIS –

- **Introduction**
- **History**
- **Determination of amino acid composition-**

  Hydrolysis

  Separation

  Quantitative analysis
- **Mechanism of protein sequencing by–**

  Chemical method

  Physical method
- **Application of protein sequencing**
- **Summary**
- **Conclusion**
- **References**

INTRODUCTION –

- Protein sequencing is a technique to determine the amino acid sequence of a protein.
- It is a method to understand the structure and function of proteins in living organism.
- Amino acid sequence determines the eventual three dimensional structure of the protein.
- All proteins are polymers of amino acid.

HISTORY –

Pehr Edman (1947) -

He found the method to decode the amino acid sequence of a protein using chemicals.

Fredrick Sanger (1955) -

He was able to present the complete sequence of insulin.

DITERMINATION OF AMINO ACID COMPOSITION-

- Amino acid composition and purity must be known before starting sequencing.
- The polypeptide chains of multimeric proteins should be separated and molecular weight of each chain should be measured.

- The determination of amino acid is done by

  Hydrolysis

  Separation

  Quantitative analysis

- **Hydrolysis**

  Peptide bonds are readily hydrolyzed by heating with either acid and base. The peptide is hydrolyzed into its constituent amino acids by heating it with 6N HCL at 110º c for 10 to 24 hours. An evacuated sealed tube is the usual procedure for complete hydrolysis.

- **Separation**

  Separation of protein is done by chromatography, dialysis, and fractionation etc.

SEPARATION BY GEL CHROMATOGRAPHY-

- Gel filtration (chromatography), is also known as molecular sieve chromatography.
- Gel filtration chromatography separates molecules according to their size and shape.
- The stationary phase consists of beads containing pores that span a relatively narrow size range.
- Smaller molecules spend more time inside the beads than larger molecules and therefore elute later (after a larger volume of mobile phase has passed through the column).

- It’s the best method for separation of molecules differing in molecular weight because:
- It doesn’t depend on temperature, pH, ionic strength and buffer composition. So separation can be carried out under any conditions.
- There is very little adsorption.
- There is less zonal spreading than in other techniques.
- The elution volume is related to the molecular weight.

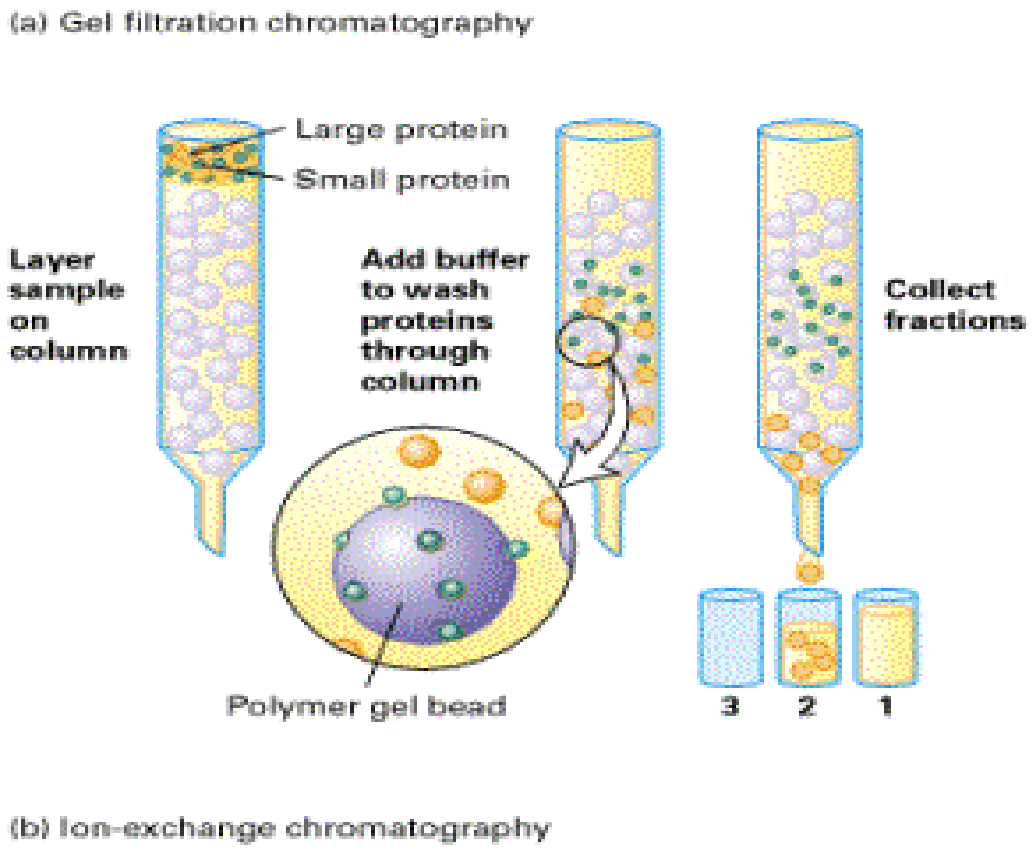

Fig . No.- 1. Gel chromatography

- **Quantitative Analysis**

The amino acid residue of peptides reacts quantitatively with ninhydrin. On heating, a α-amino acid reacts with two molecules of ninhydrin to yield an intensely coloured product. Purple colour is given in this test by all amino acids and peptide having a free α -amino group where as proline gives yellow colour.

Amino acid+ 2 ninhydrin $\longrightarrow$ **CO_2 + aldehyde + final Complex (yellow/purple) + $3H_2O$**

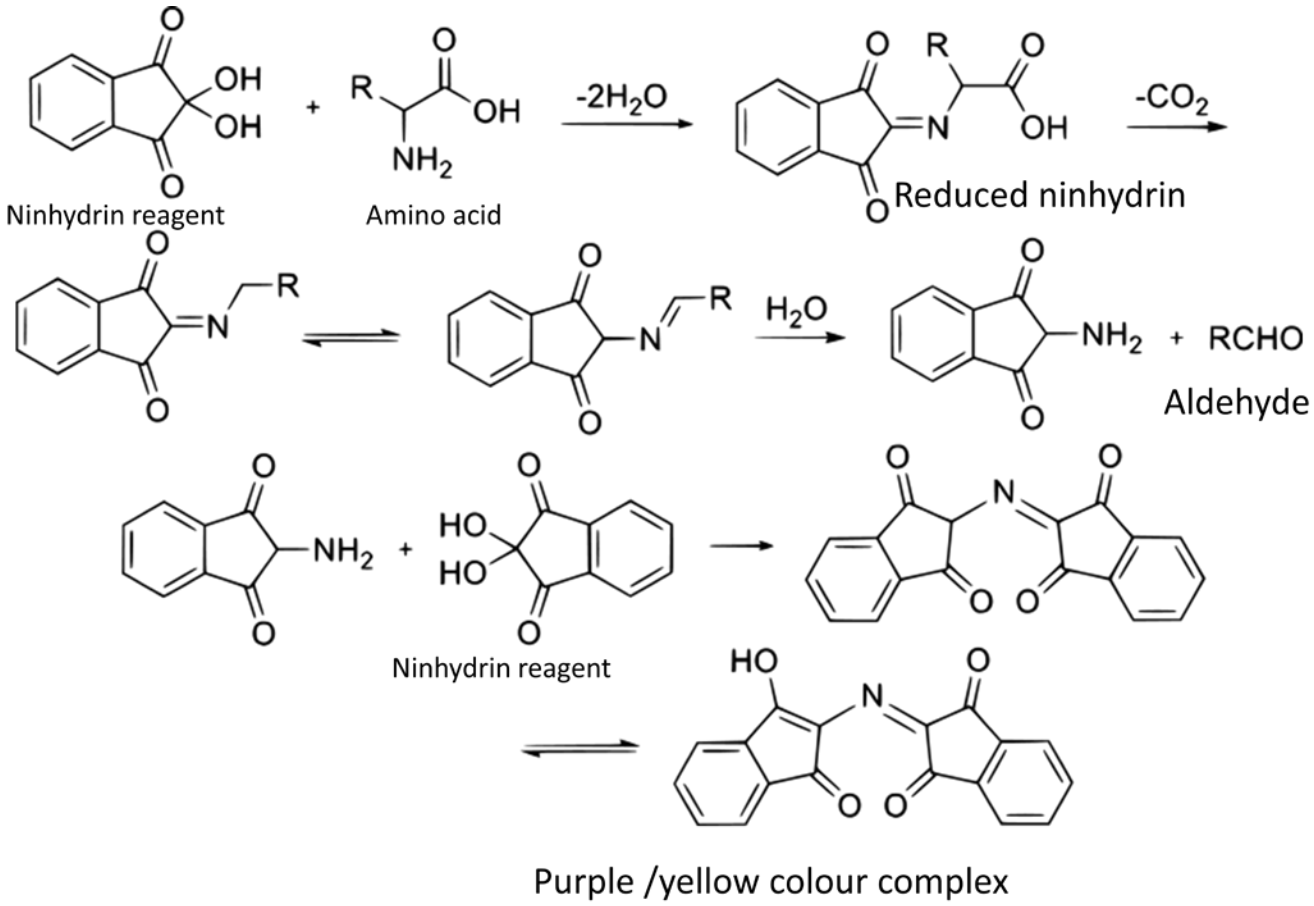

Fig . No.- 2. Ninhydrin reaction

MECHANISM OF PROTIEN SEQUENCING -

- **Chemical method-**

  Sanger's method

  Edman's method

- **Physical method –**Mass spectrometry

<u>Sanger's method</u>-

- Fredrick Sanger was a Nobel Laureate.
- He showed for the first time that amino acids are covalently together by α-amino and α carboxyl group.
- He suggested stepwise release and identification of amino acid starting from N-terminal.
- He used reagent fluoro dinitro benzene (FDB) which is commonly called Sanger's reagent.
- The FDB reacts with free NH_2 group of N-terminus.
- Upon hydrolysis a yellow coloured dinitrophenol (DNP) derivative of N-terminal amino acid is produced.
- The DNP amino acid is identified comparing it with a known standard DNP-amino acid by using gel chromatography.

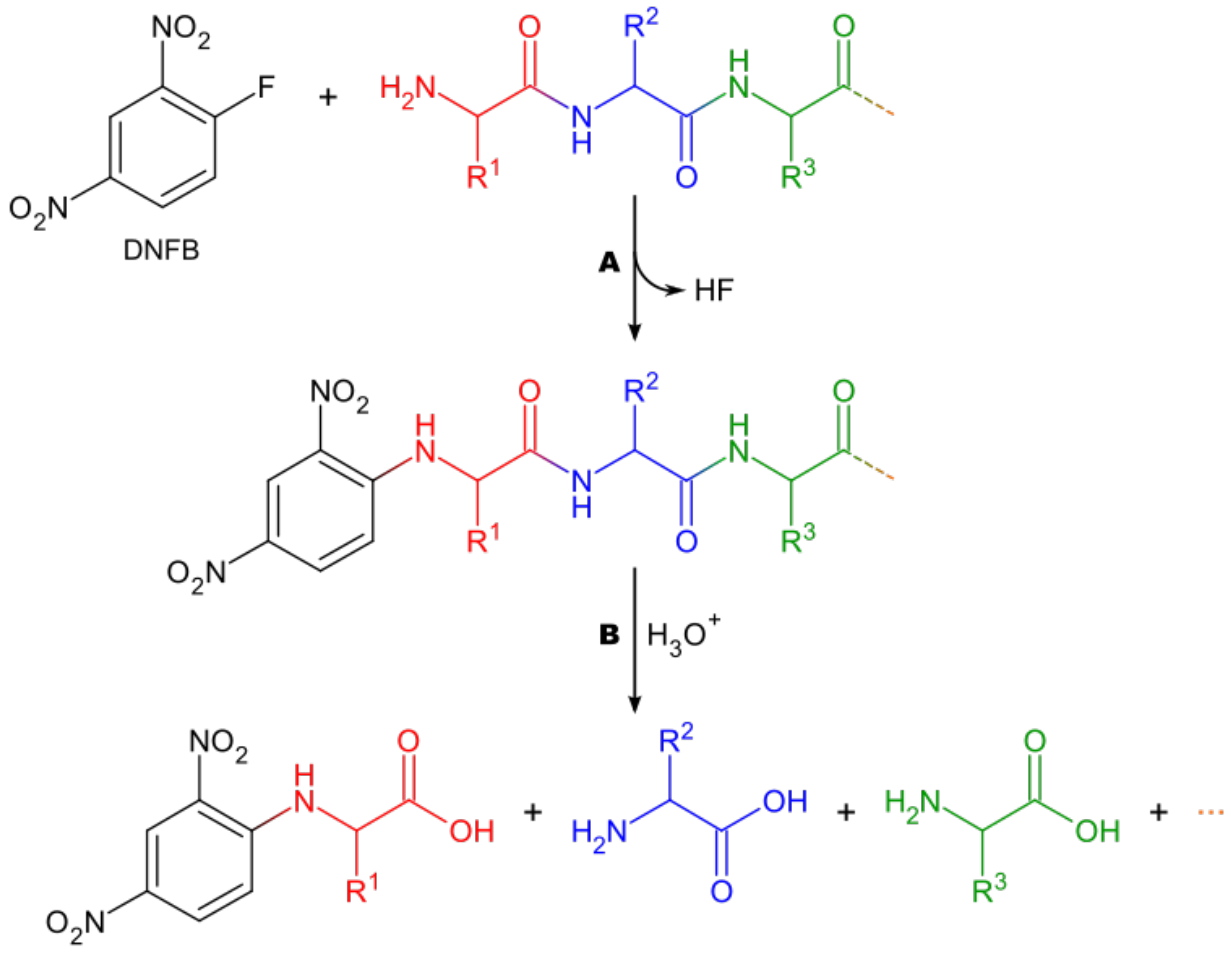

Fig. No. 3. Sanger's method

Dansyl Chloride -

- The dansyl chloride (dimethyl amino naphthalene 5 sulfonyl chloride) is commonly used because it forms intensively coloured derivatives.
- That can be detected with high sensitivity that the dinitro phenyl compound.
- It reacts with an uncharged C and N terminal to form a sulfonamide derivatives that is the stable under condition that by hydrolyser peptide bonds.
- Although the dansyl method for determining the amino terminal residue is sensitive and powerful.

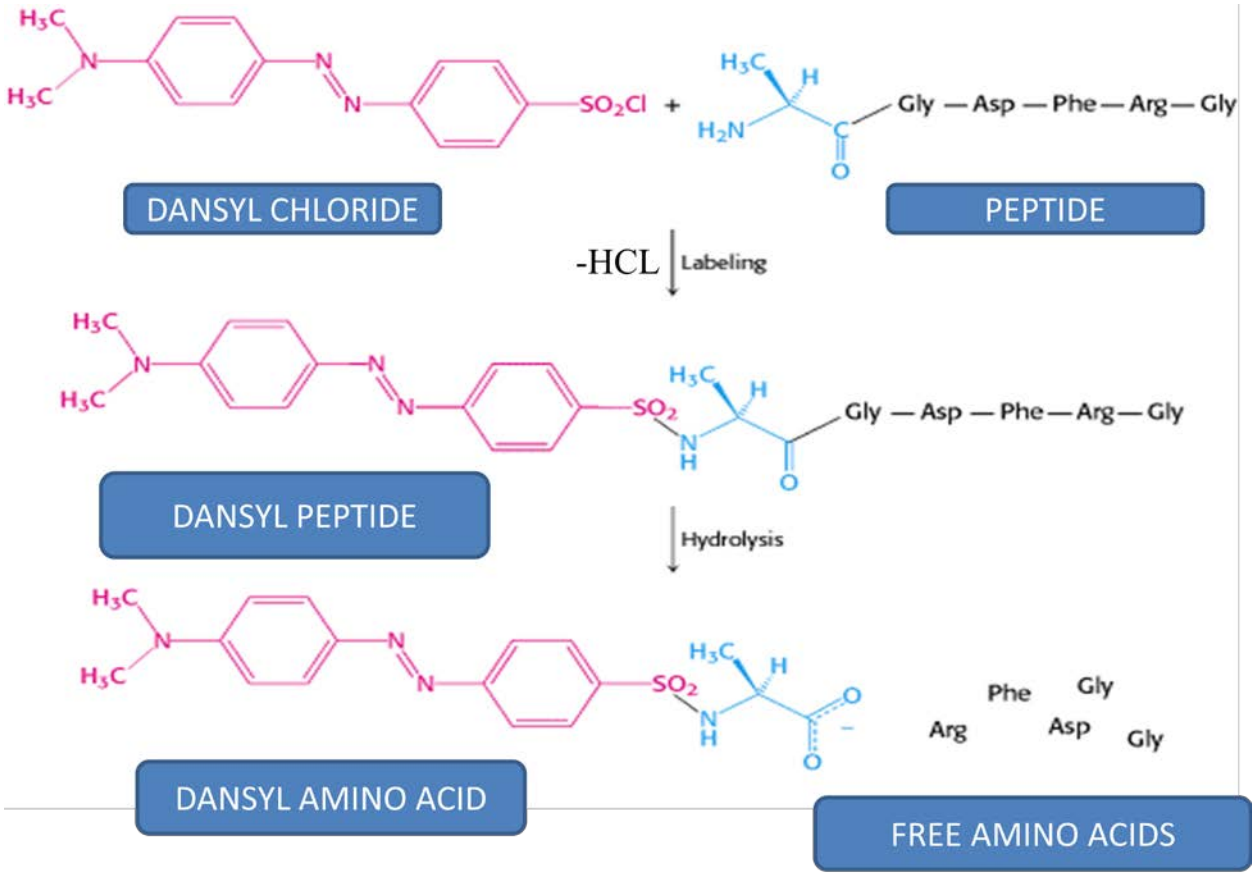

FIG. NO.- 4. DANSYL CHLORIDE METHOD

Disadvantages -

- ❖ It cannot be used repeatedly on the same peptide because the peptide is totally degraded in the acid hydrolysis step.
- ❖ The Sanger's method is more sensitive and efficient procedure.
- ❖ The dansyl chloride is 100 times more sensitive than the Sanger's method.
- ❖ The Sanger's method is only useful for identification of N-terminal residue of a peptide.

Edman degradation -

- ➢ It involves sequential identification of amino acids from N to C termini.

- Phenyl isothiocynate (PTC) reagent is used for the Edman degradation.
- The amino terminal (N-terminal residue of a protein can be identifies by reaction the protein with the PTC that forms a stable covalent link with the free α amino group prior to hydrolysis with 6M HCl
- Phenyl isothiocynate reacts with the uncharged terminal amino group of the peptide to form a phenyl thiocarbonyl derivation.
- The labeled N-terminal amino acid can be identified by comparison of its chromatographic properties with stanrdard fluorodinitro benzene and dansyl chloride.
- Then, under mild acidic the cyclic (PTH) Phenylthio hydantoin of the terminal amino acid is librated which leaves an intact peptide shorted by one amino acid.
- The released PTH amino acid is identified by high performance liquid chromatography.
- The sequencing technique has been automated and refined so that upwards of 50 residues from the N-termines of a protein can be sequenced from picomolequantities of material.

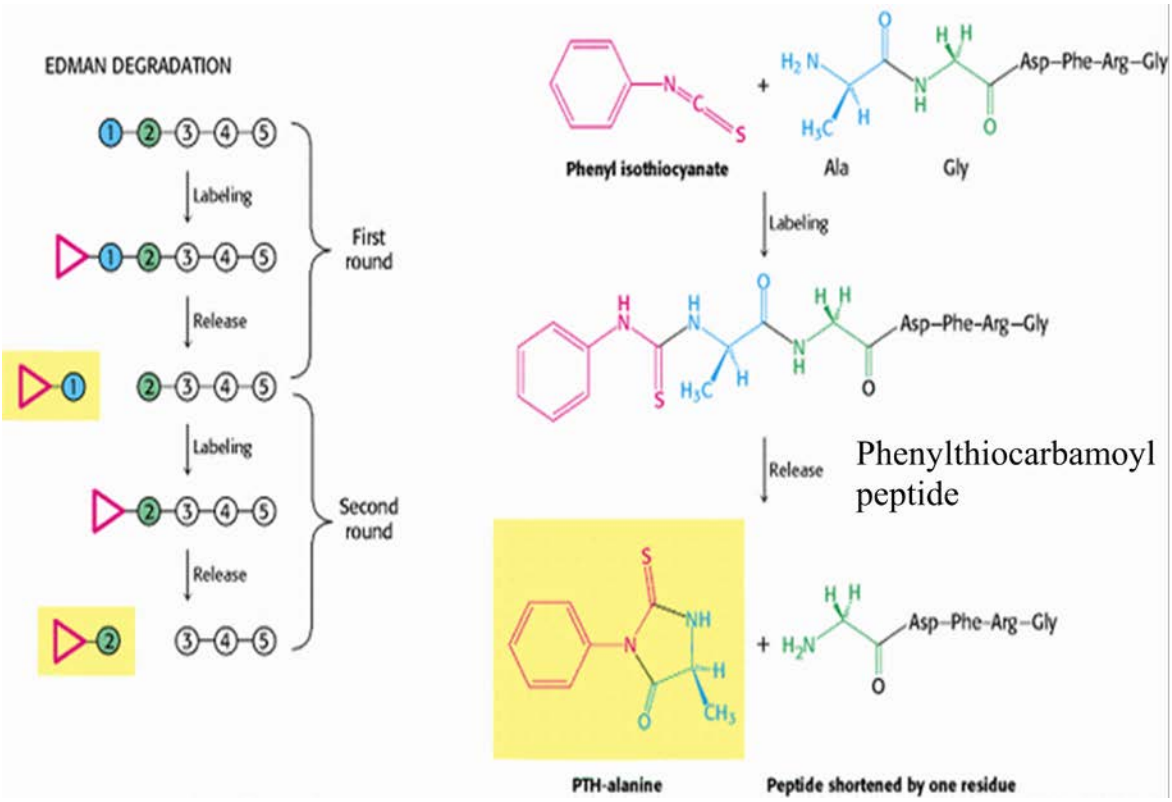

FIG. NO. 5. EDMAN DEGRADATION

Advantages –

The great advantage of the Edman method is that the rest of the peptide chain after removal of the N-terminal amino acid is left intact for further cycles of this procedure, thus the Edman method can be used in a sequential fashion to identify several or many consecutive amino acid residues starting from the N-terminal end.

Disadvantages -

- The Edman degradation proceeds from the N-terminus of the protein it will not work it the N-terminal amino acid has been chemically modified.
- It also required the use of either guess work or a separate- procedure to determine the position of disulfide bridges.

Protein sequencing by Mass Spectrometry -

- Mass spectrometry is an important emerging method for the characterization of proteins.
- The two primary methods for ionization of whole proteins are electro spray ionization (ESI) and matrix assisted laser desorption/ionization (MALDI).
- Mass spectrometer has three main parts for characterizing the protein –

1. An ionization source
2. A mass analyzer
3. An ion detector

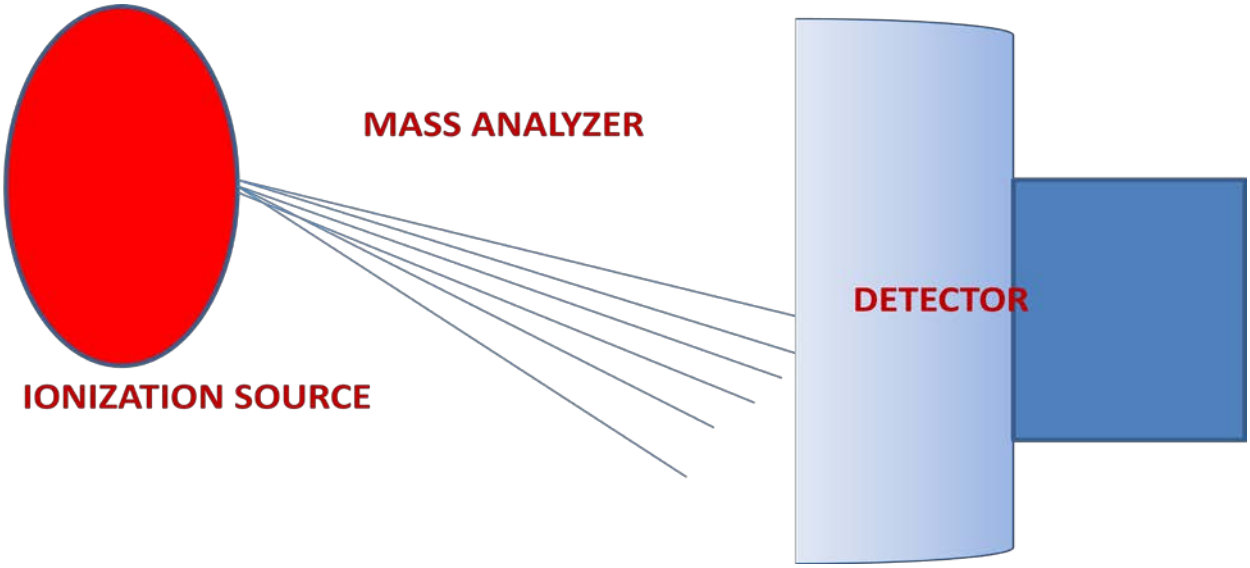

Fig no.6. Parts of mass spectrometer

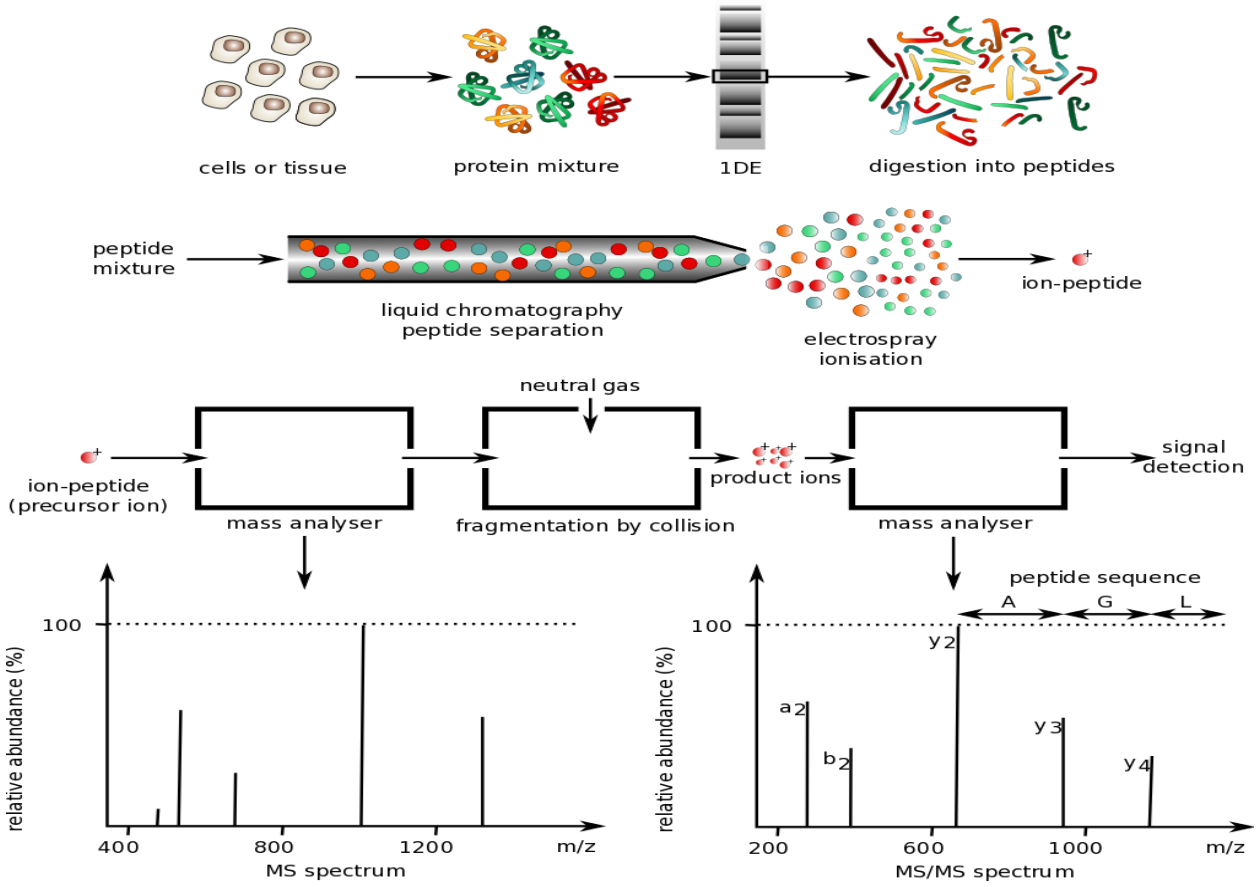

Fig no7. Mass Spectrometry

- Digested the peptide bonds (by protease enzyme).
- Ionization by MALDI (matrix assisted laser desorption/ionization) or electro spray.
- Then ionized molecules are introduced into mass- analyzer.
- Using the mass spectrometer the masses of each of these ionized peptides is calculated.
- Using the masses it is now possible to know the amino acid sequence of this protein.

ADVANTAGES –

- Stable isotopic labeling of proteins may be used to discriminate between contaminants and original partners using MS (mass spectrometry) techniques.
- MS (mass spectrometry) has also made advantages in the analysis of membrane proteins.

DIS-ADVANTAGES-

- Miscalibration is one of the main errors of MS spectrometry.
- MALDI doesn't favor the identification of hydrophobic peptides.

APPLICATIONS OF PROTEIN SEQUENCING

- mRNA /protein analysis and coding SNP (single nucleotide protein) scoring tools.
- Knowledge of the sequence of amino acids in a protein can offer insights into its three dimensional structure and its function in cellular location and evolution.
- Certain amino acid sequences serve as signals that determine the cellular location chemical modification on a half life of a protein.
- Protein sequences can elucidate the history of life on Earth.

SUMMARY-

- The protein is a polypeptide chain which is made up of amino acid monomers polymerization.

- Protein sequencing is the sequencing of monomer of peptide chain which is amino acid.

- For sequencing of protein firstly we determined the amino acid sequences which are done by hydrolysis, separation and quantitative analysis.

- The protein sequencing is done by two methods which are chemical and physical method.

CONCLUSION-

- Protein sequencing also gives information regarding which conformation the protein adopts.
- Discovering the structures and functions of proteins in living organisms is an important tool for understanding cellular processes.
- Not all proteins contain all the amino acid, nor do the amino acids occur with equal frequency.
- The amino acids have N-terminal end and C-terminal end.

REFERENCES –

Cox M. Michael and

Nelson L. David 2008 Principles of Biochemistry, 5th Edition

Lehninger L. Albert 2009 Biochemistry, 2nd Edition

Stryer Lambert 2009 Principle of Biochemistry, 2nd Edition