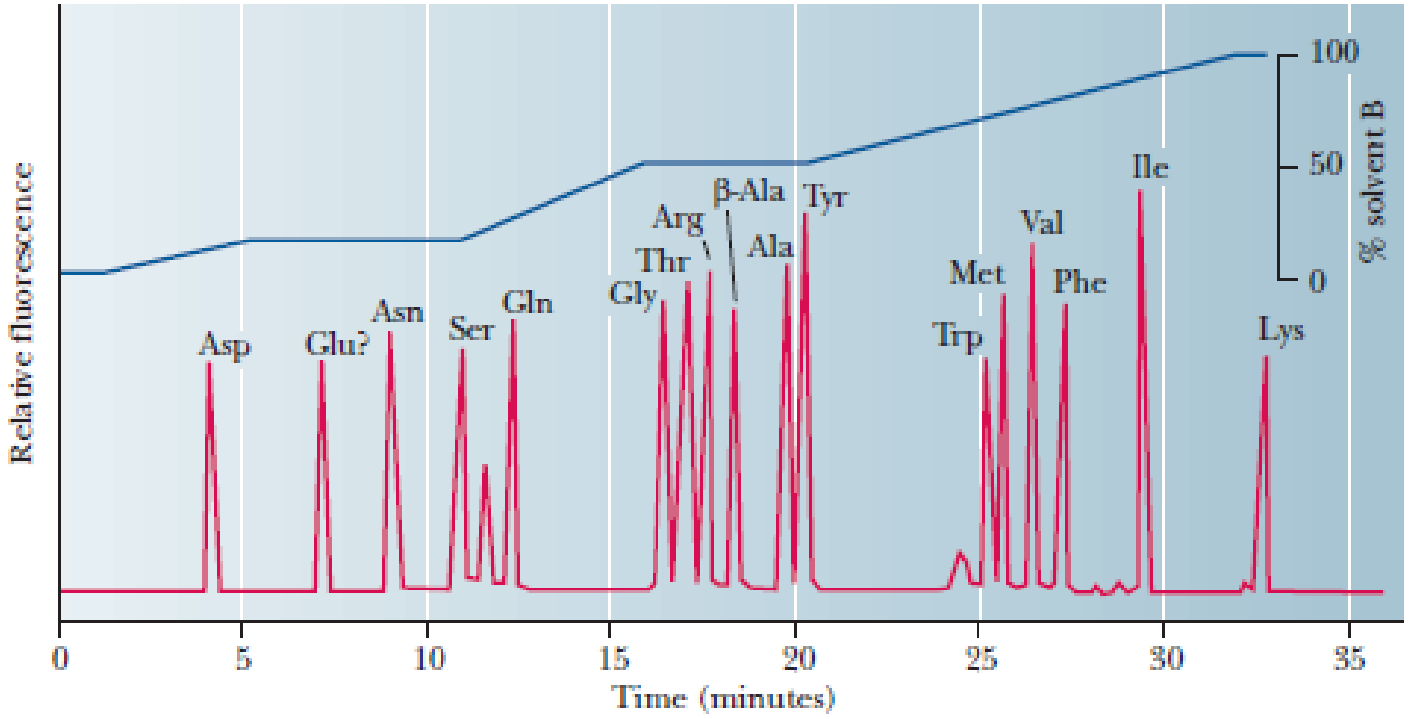
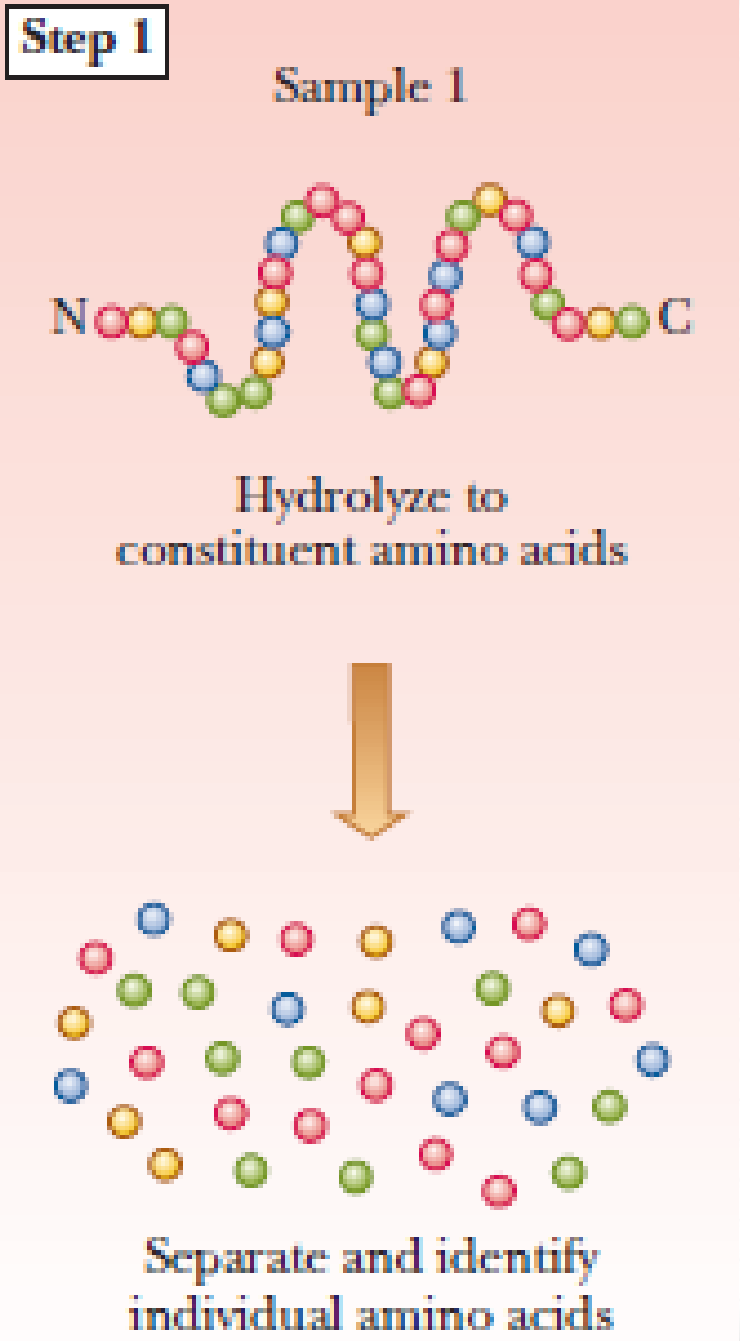
# Protein sequencing

- Protein sequencing is basically the process of knowing the amino sequence of a protein or a peptide.
- One technique is known as Edman Degradation.
- This procedure involves a step-by-step cleavage of the N-terminal residue of a peptide, allowing for the identification of each cleaved residue.

# Protein sequencing - Edman Method

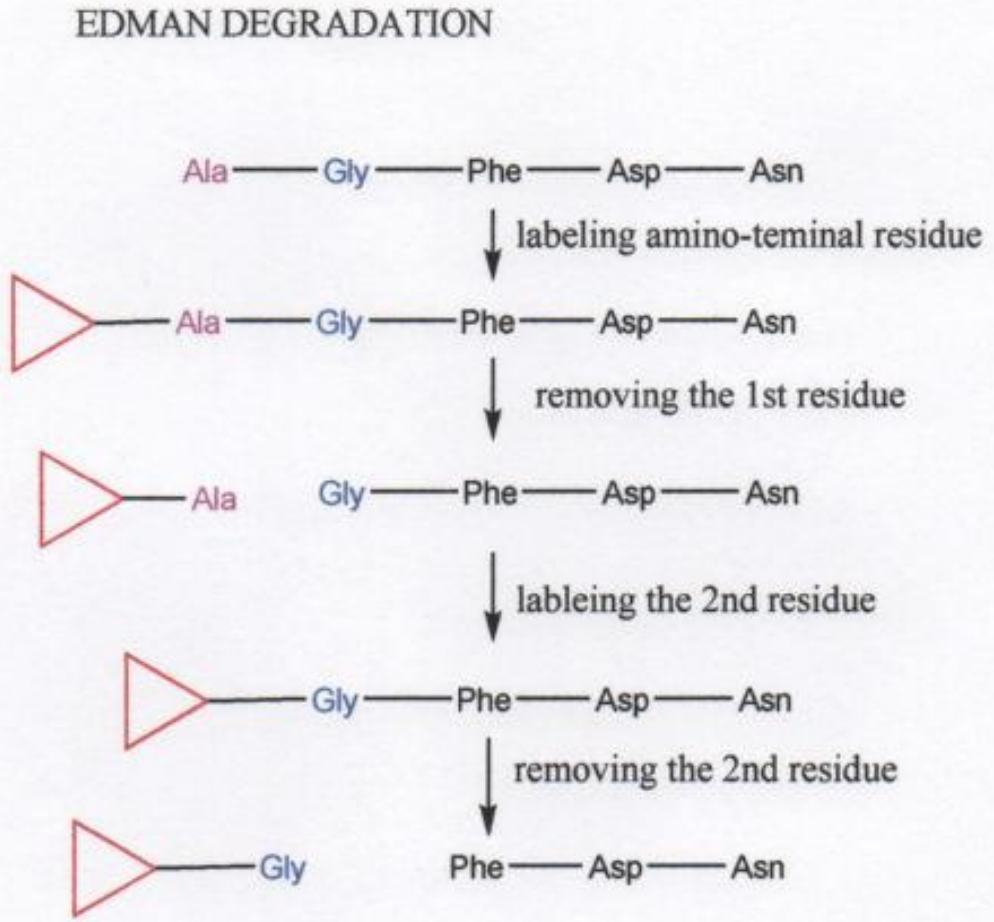
- how much and which amino acids are involved
- **Hydrolysis** (heating + HCl) **& Separation** (ion-exchange chromatography or by **high performance liquid chromatography, HPLC)**



# Procedure

- This method utilizes phenylisothiocyanate (PITC) to react with the N-terminal residue.
- The resultant amino acid is hydrolyzed, liberated from the peptide, and identified by chromatographic procedures.



# Advantage

- Since the remainder of the peptide is intact, the entire sequence of reactions can be repeated over and over to obtain the sequences of the peptide.

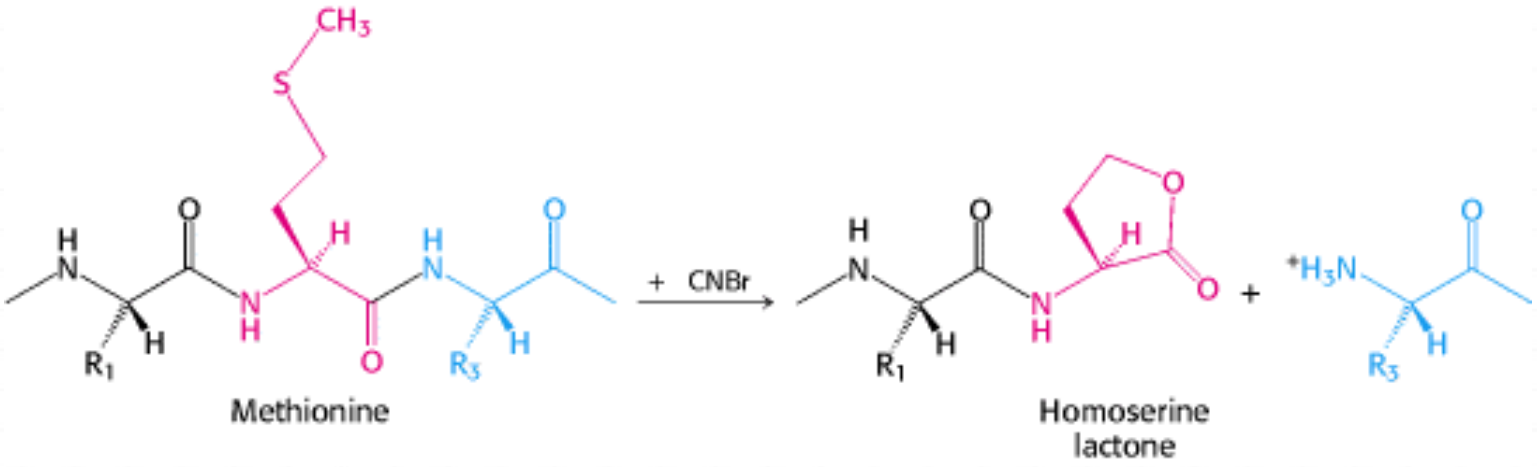
- The Edman degradation technique does not allow peptides more than 50 residues to be sequenced.

# Cleavage methods

- It is possible to sequence whole proteins by cleaving them into smaller peptides.
- This is facilitated by three methods:
  - Chemical digestion
  - Endopeptidases
  - Exopeptidases

# Chemical digestion

- The most commonly utilized chemical reagent that cleaves peptide bonds by recognition of specific amino acid residues is cyanogen bromide (CNBr).
- This reagent causes specific cleavage at the C-terminal side of methionine residues.
- A protein that has 10 methionine residues will usually yield 11 peptides on cleavage with CNBr.



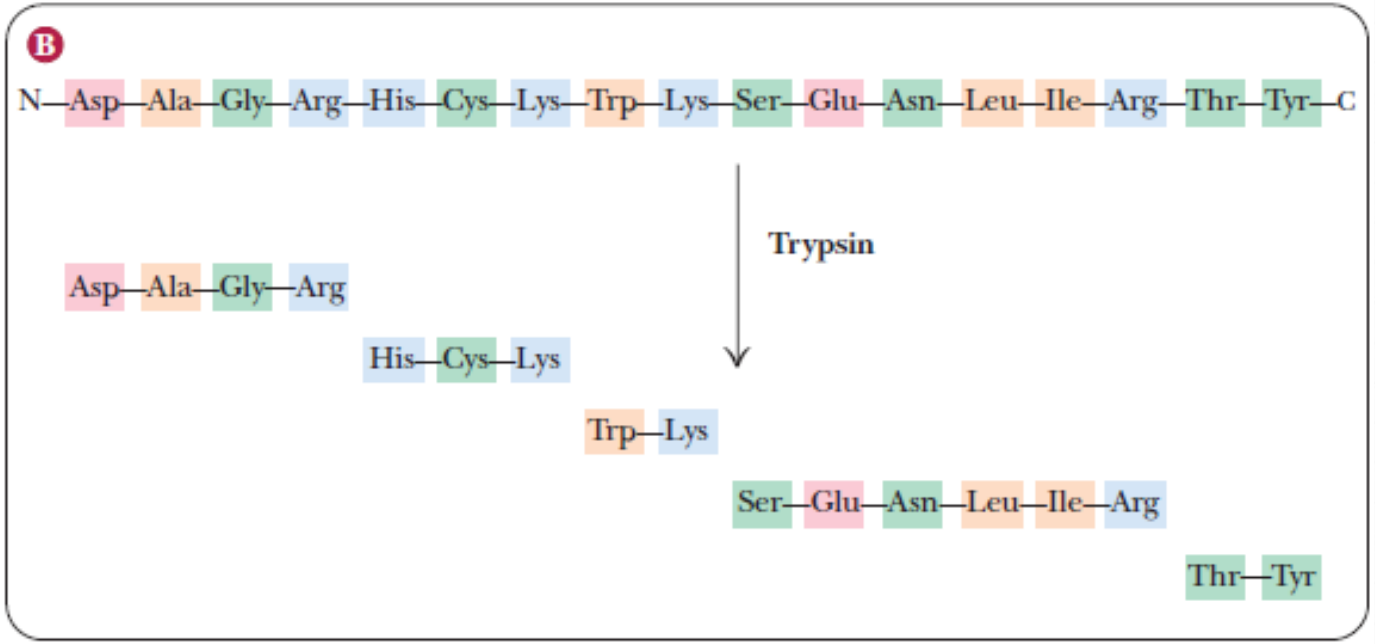
# Endopeptidases

- These are enzymes that cleave at specific sites within the primary sequence of proteins.
- The resultant smaller peptides can be chromatographically separated and subjected to Edman degradation sequencing reactions.

# Example

- Trypsin cleaves polypeptide chains on the carboxyl side of arginine and lysine residues.
- A protein that contains 9 lysine and 7 arginine residues will usually yield 17 peptides on digestion with trypsin.

**Another example**



# Other examples

| Enzyme | Specificity |
| --- | --- |
| Trypsin | peptide bond C-terminal to Arg, Lys, but not if followed by Pro (at C-terminal) |
| Chymotrypsin | peptide bond C-terminal to Phe, Tyr, Trp but not if followed by Pro (at C-terminal) |
| Elastase | peptide bond C-terminal to Ala, Gly, Ser, Val, but not if followed by Pro (at C-terminal) |
| Pepsin | peptide bond N-terminal to Leu, Phe, Trp, Tyr, but not when preceded by Pro (at N-terminal) |

# Exopeptidases

- These are enzymes that cleave amino acids starting at the end of the peptide.
- There are two types:
  - Aminopeptidases that cleave at the N-terminus
  - Carboxypeptidases that cleave at the C-terminus

# Protein sequencing – prediction from DNA & RNA

- If the sequence of the gene is known, this is very easy
- If the sequence of the gene is unknown (newly isolated proteins)? Sequence a short segment, complementary RNA, isolate mRNA, PCR, gene sequencing