INCORRECT PROTEIN FOLDING (MISFOLDING) AND NEURODEGENERATIVE DISEASES

- The correct three-dimensional structure is essential to function, although some parts of functional proteins may remain unfolded. Failure to fold into native structure (misfolding) produces inactive proteins that are usually toxic (infectious). Several neurodegenerative and other diseases are believed to result from the accumulation of amyloid fibrills formed by *misfolded* proteins.
- Aggregated proteins are associated with prion-related illnesses such as Creutzfeldt-Jakob disease (CJD), bovine spongiform encephalopathy (mad cow disease) and amyloid-related illnesses such as Alzheimer's disease, as well as diseases such as Huntington's and Parkinson's disease. These age onset degenerative diseases are associated with the multimerization of misfolded proteins into insoluble, extracellular aggregates and/or intracellular inclusions including cross-beta sheet amyloid fibrils.

Misfolding of some proteins may cause several rare diseases

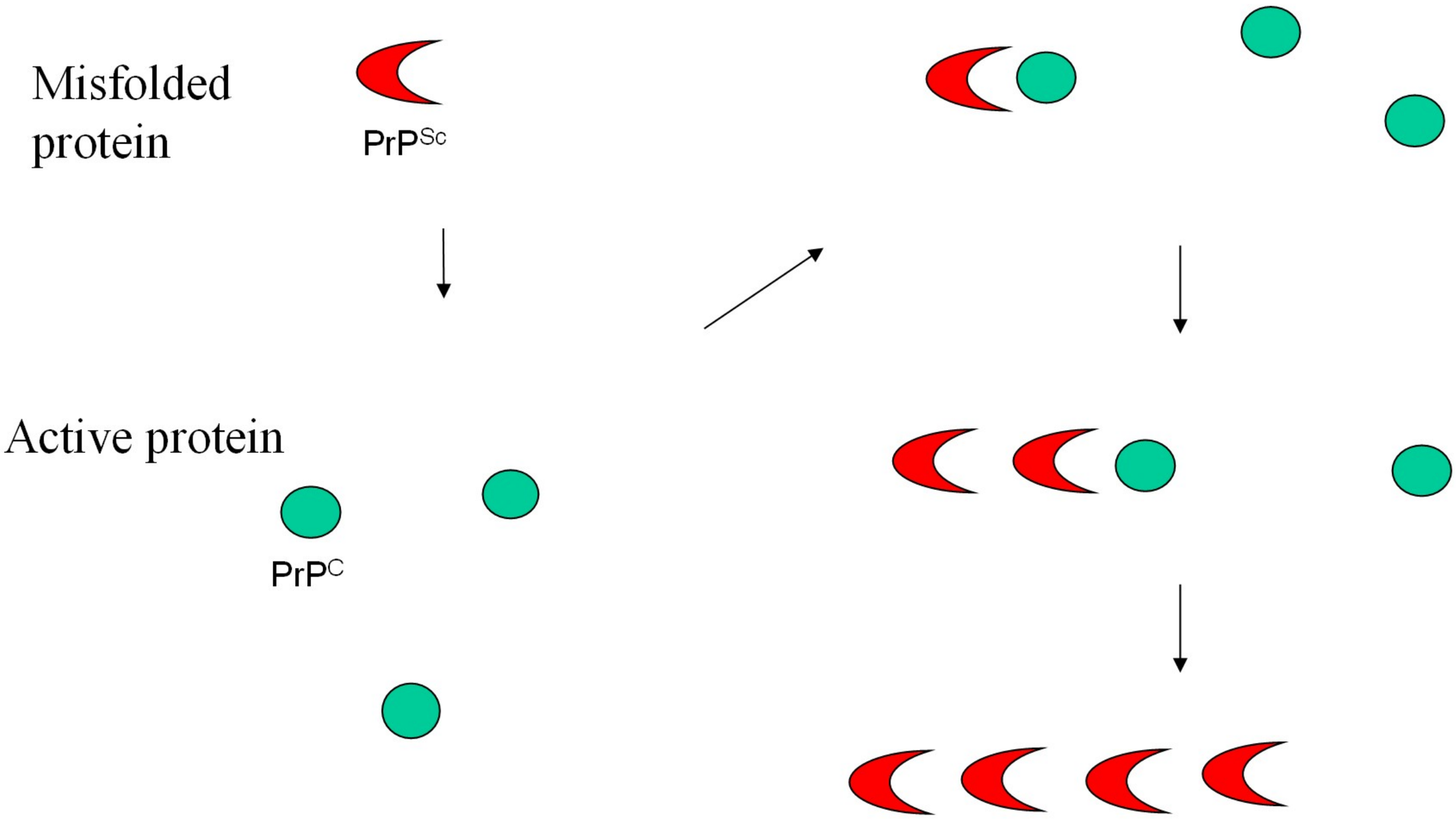
e.g. Prion diseases

An example of Prion diseases is mad cow disease, or bovine spongiform encephalopathy (BSE), which is a fatal brain disorder that occurs in cattle. Abnormal protein folding (misfolding) is considered crucial to the onset of the disease.

Evidence indicates that the infectious agent in transmissible spongiform encephalopathy is a protein (Prion protein). Stanley Prusiner pioneered the study of these proteins and received the Nobel Prize in 1997. He has named them prion proteins (referred to as PrP) or simply prions.

The normal protein is called PrP^{c} (for cellular) is a normal constituent of brain tissue in most mammals (function not well known) of a m.wt. of 28,000 Da. Its secondary structure is dominated by alpha helices. The abnormal, disease producing protein called PrP^{Sc} (for Scrapie), has the same primary structure as the normal protein, but its secondary structure is dominated by beta conformations. Interaction of misfolded form (PrP^{Sc}) with the normal PrP^{c}converts them very rapidly (domino effect) to PrP^{Sc}

Misfolded proteins can be infectious (Mad Cow's Disease, Prion proteins)

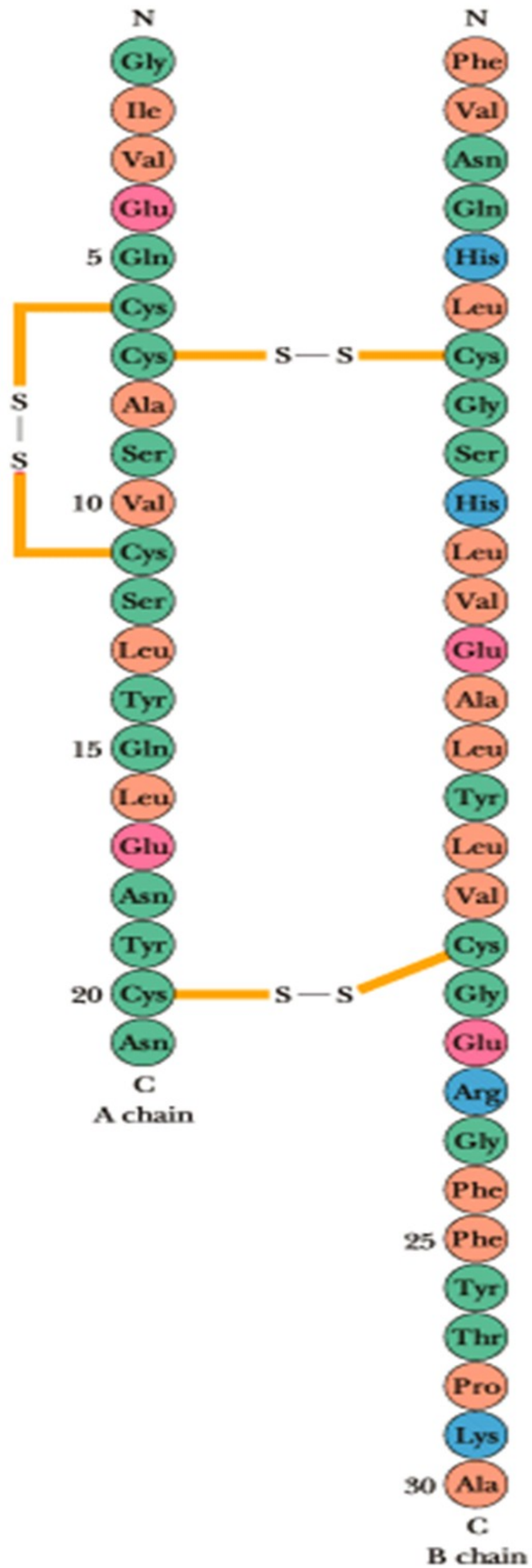


Stanely Prusiner: 1997 Nobel Prize in Medicine

PROTEIN SEQUENCING

First Sequence

- The first protein sequencing was achieved by Frederic Sanger in 1953.
- He determined the amino acid sequence of bovine insulin
- Sanger was awarded the Nobel Prize in 1958

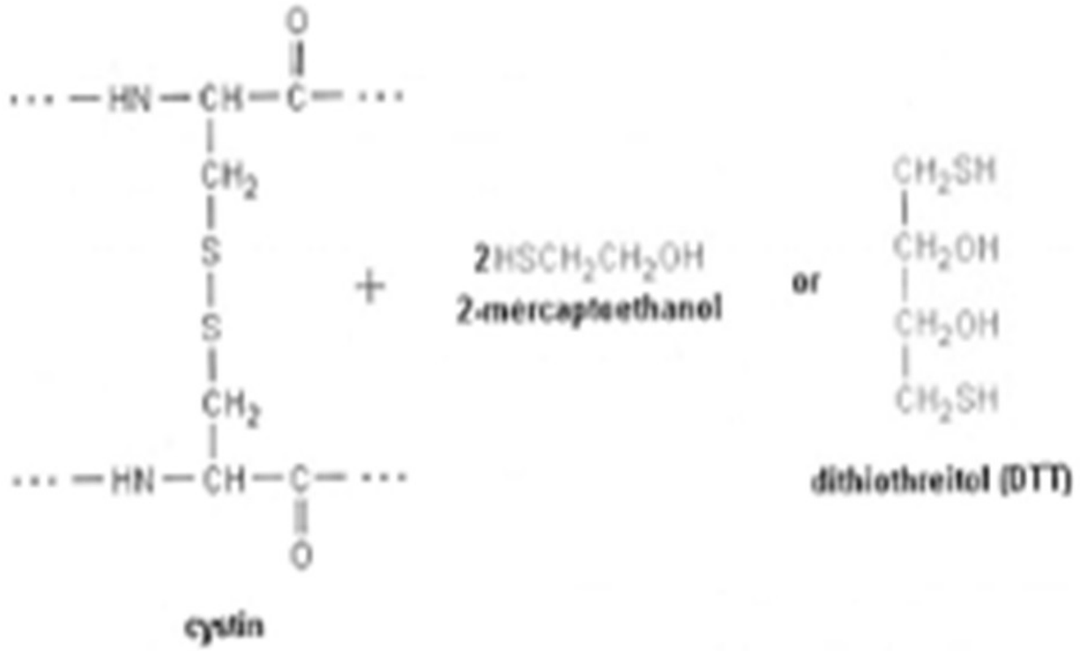


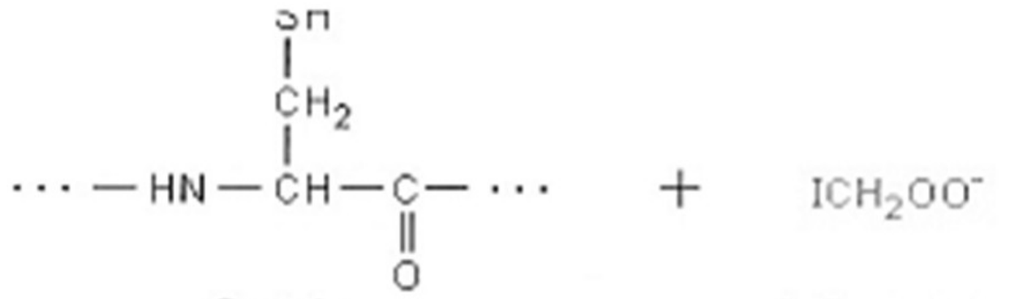
How To Sequence A Protein

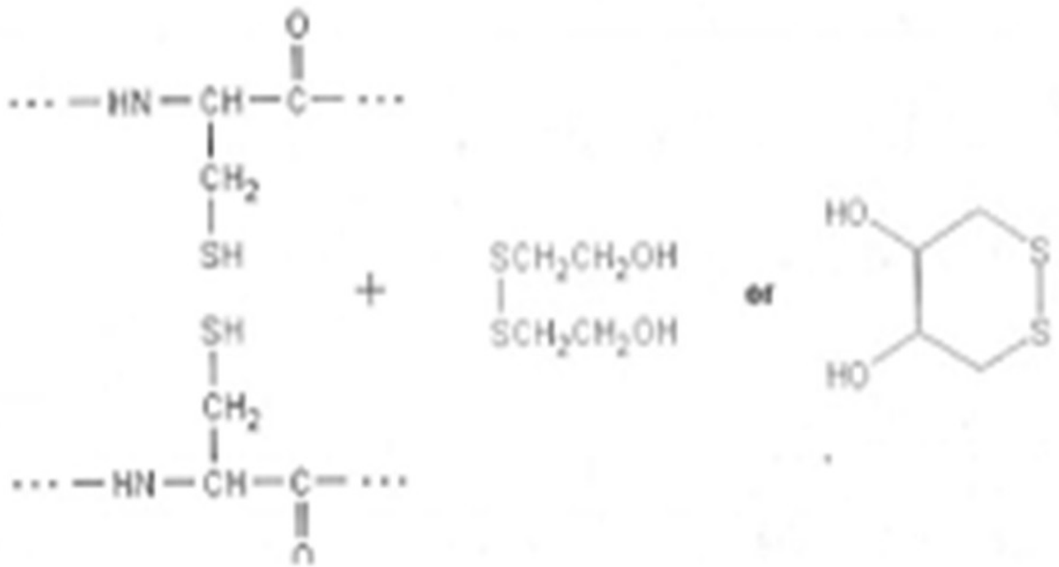
Preliminary Steps

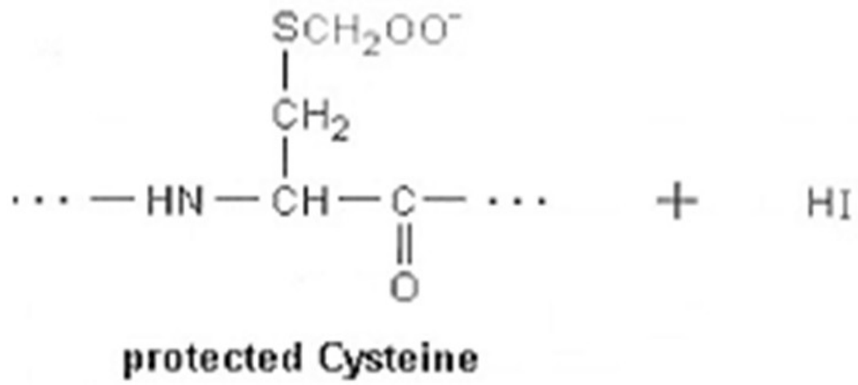
- For multisubunit proteins, the individual protein chains must first be separated
- Break interchain disulfide bonds, if necessary
- Three reagents are commonly used:
  - Performic acid
  - 2-Mercaptoethanol
  - Dithiotreitol (DTT)

Reduction of S-S bond using 2-Mercaptoethanol or DTT, followed by alkylation using Iodoacetate

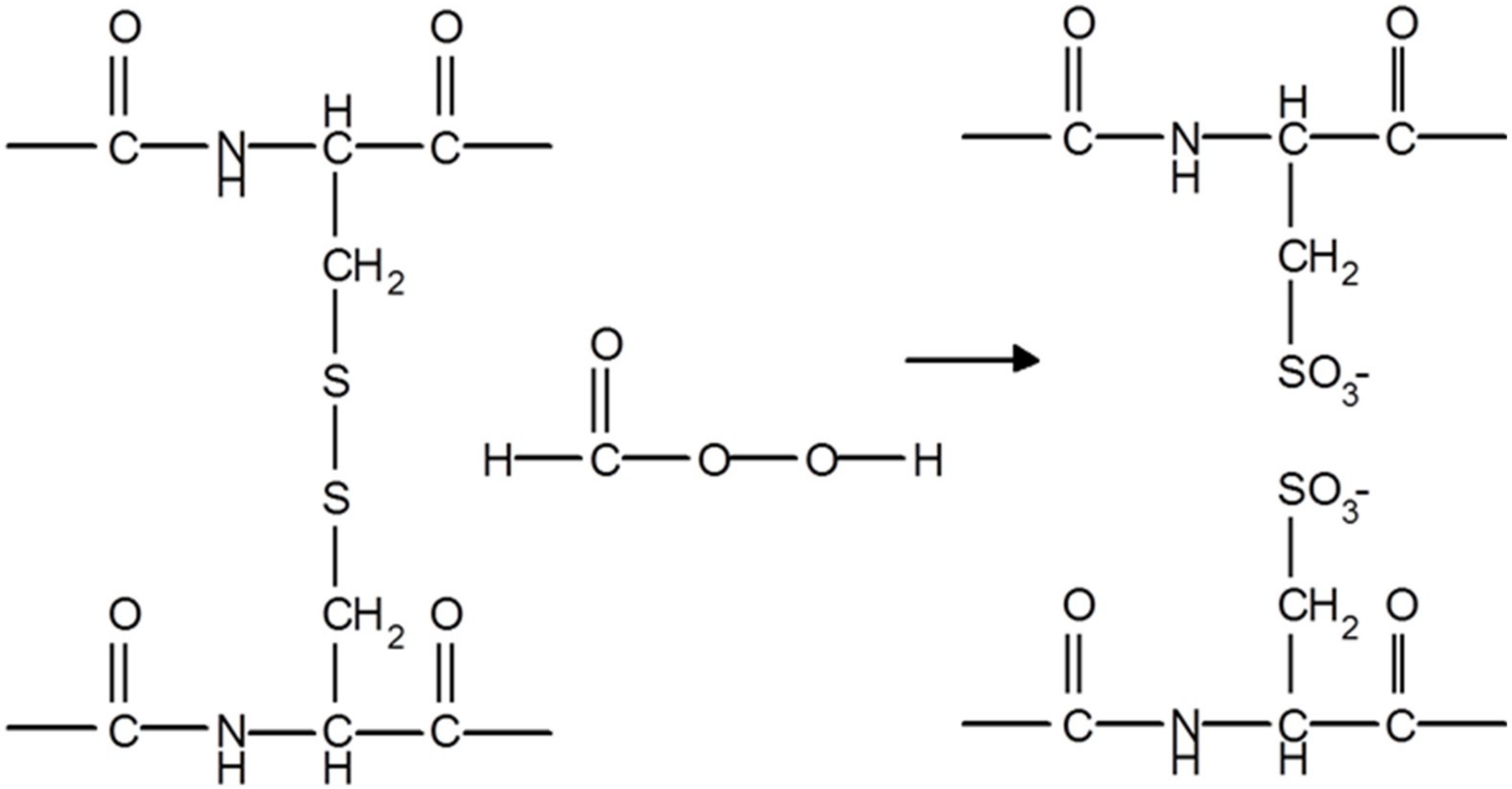








Perfomic Acid



- Performic acid oxidizes cysteine to negatively charged cysteic acid, so there is no need for alkylation

Protein Sequencing: Preliminary Steps

- After breaking disulfide bonds, the chains are separated by disrupting noncovalent interchain interactions with pH extremes, 8 M urea, 6 M guanidium hydrochloride, or high salt.
- Then the individual protein chains are separated by electrophoresis or chromatography on the basis of size or charge

Determining Amino Acid Sequence

- Once each protein is purified the amino acid sequence is determined by:
- 1) determining the amino acid composition (how many of each amino acid are in the protein)
- 2) identifying the amino and carboxyl terminal amino acids
- 3) cleaving the protein into two or more sets of peptides using specific enzymatic or chemical reagents such as trypsin or cyanogen bromide

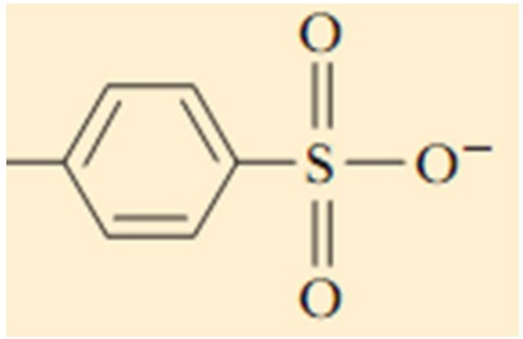
Determining Protein Sequence

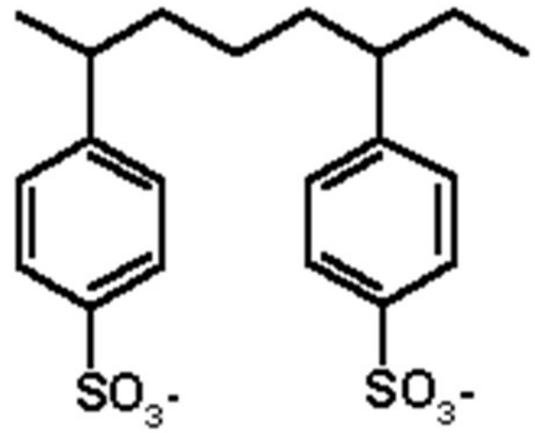
- 4) determining the amino acid sequence of each of the peptide fragments
- 5) determining the entire protein sequence from the sequences of overlapping peptide fragments
- 6) locating the position of disulfide bridges between cysteines

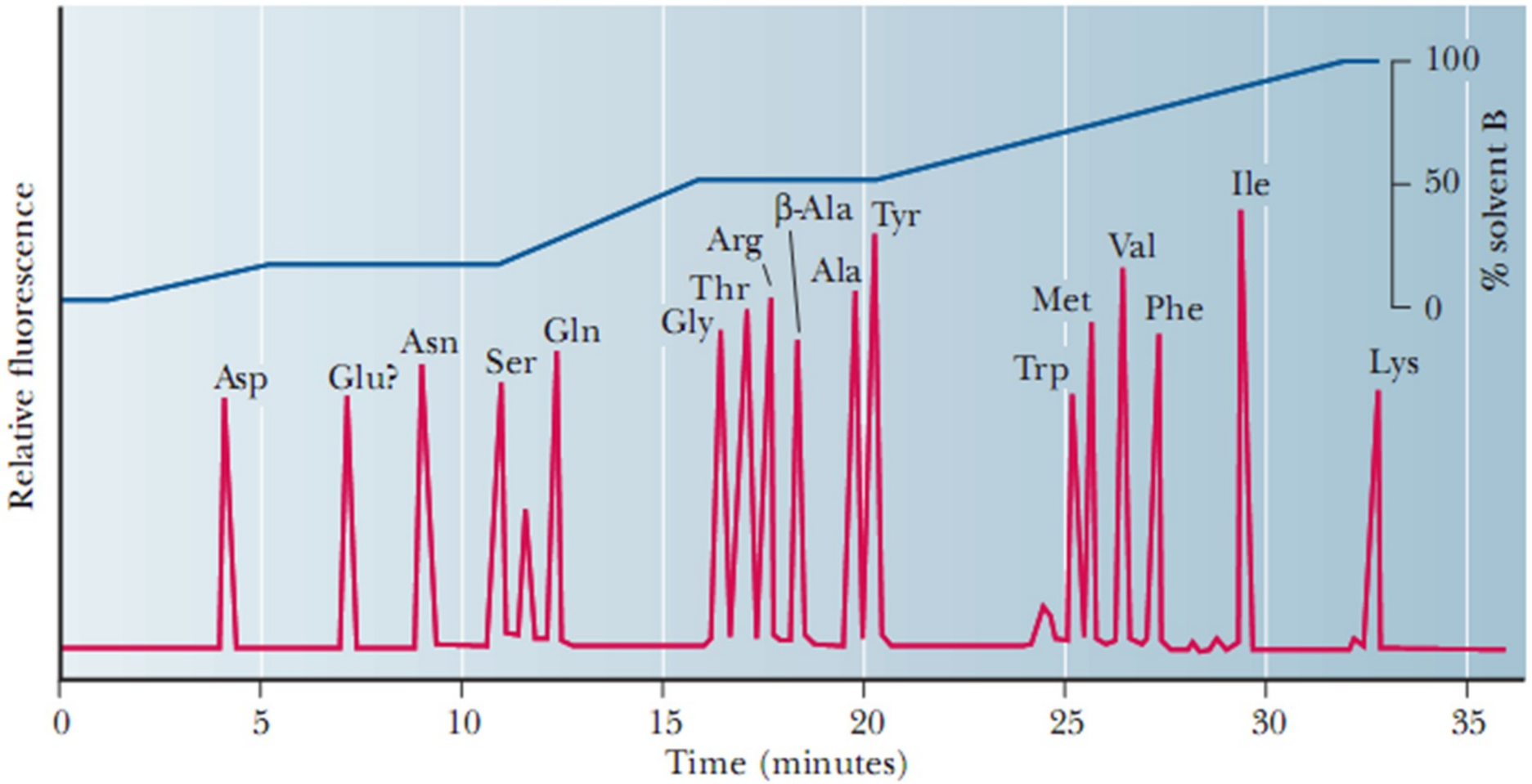
Determining Amino Acid Composition

- The amino acid composition is determined by:
- Hydrolysis with 6N HCl for 20-24 hours at 110ºC, *in vacuo*. (This treatment destroys Trp residues and converts "Asn & Gln" to "Asp & Glu + NH_3"
- Separating and quantifying individual amino acids by ion exchange HPLC (Dowex-50 resin) using an amino acid analyzer.

Sulfonated polystyrene
(Dowex-50)





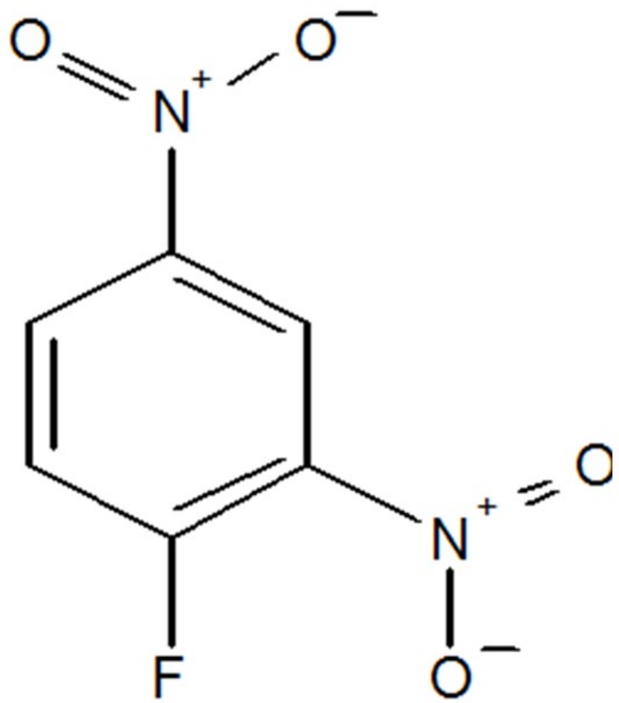


HPLC chromatogram of amino acid separation

Determining the N-Terminal Amino Acid

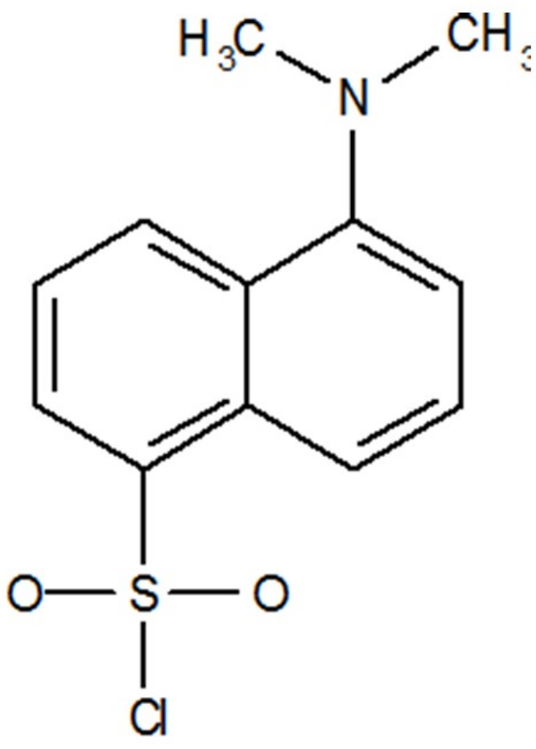
- The N-terminal amino acid is determined using either chemical reagents or enzymes
- Chemical reagents include:
- Sanger's reagent
- dansyl chloride
- Edman Degradation

Determining the N-Terminal Amino Acid



- Sanger's reagent
- Treat with dinitrofluorobenzene to form a dinitrophenyl (DNP) derivative of the amino-terminal amino acid
- Acid hydrolysis
- Extract the DNP-derivative from the acid hydrolysate with organic solvent
- Identify the DNP-derivative by chromatography and comparison with standards

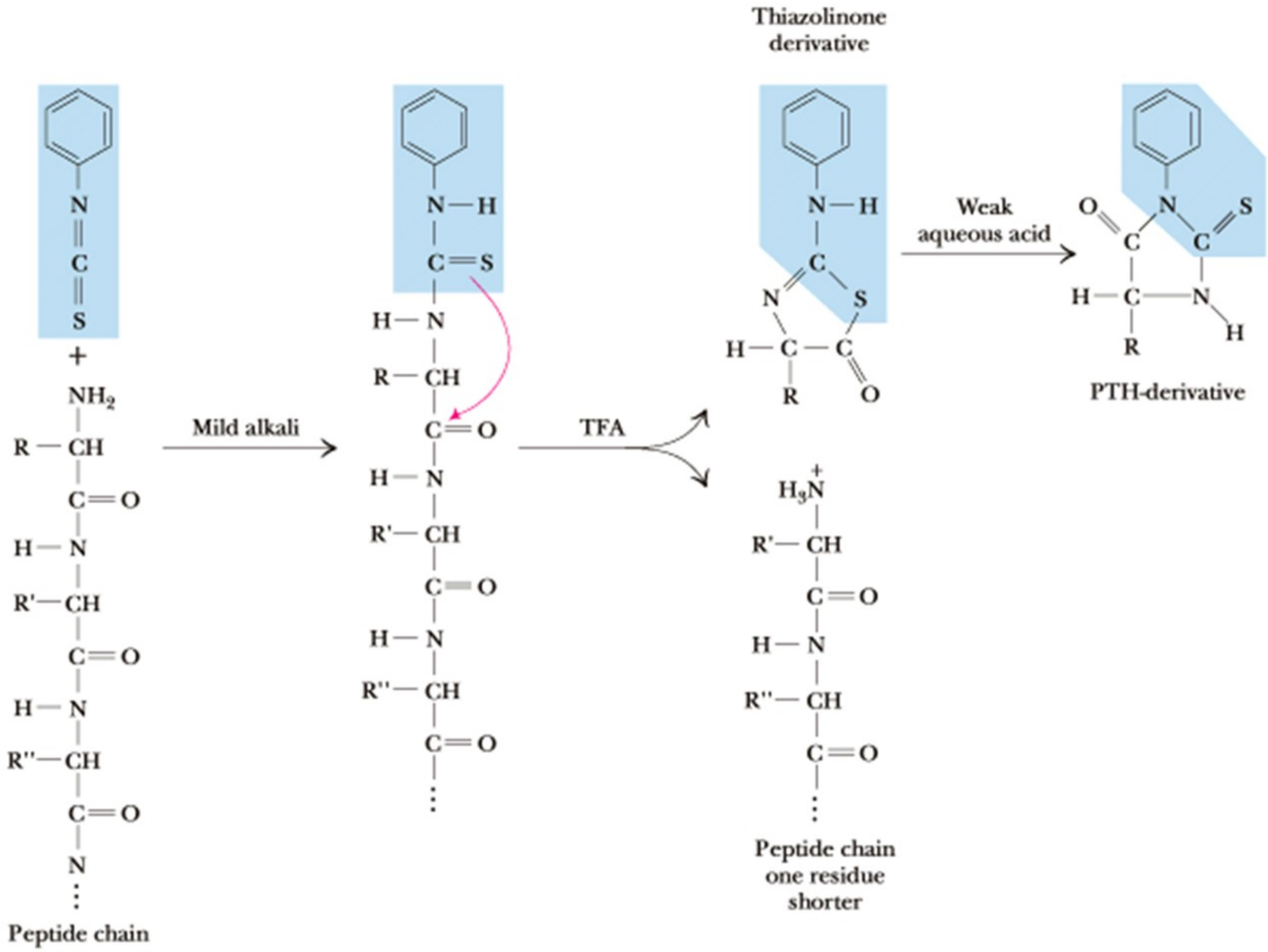
Determining the N-Terminal Amino Acid

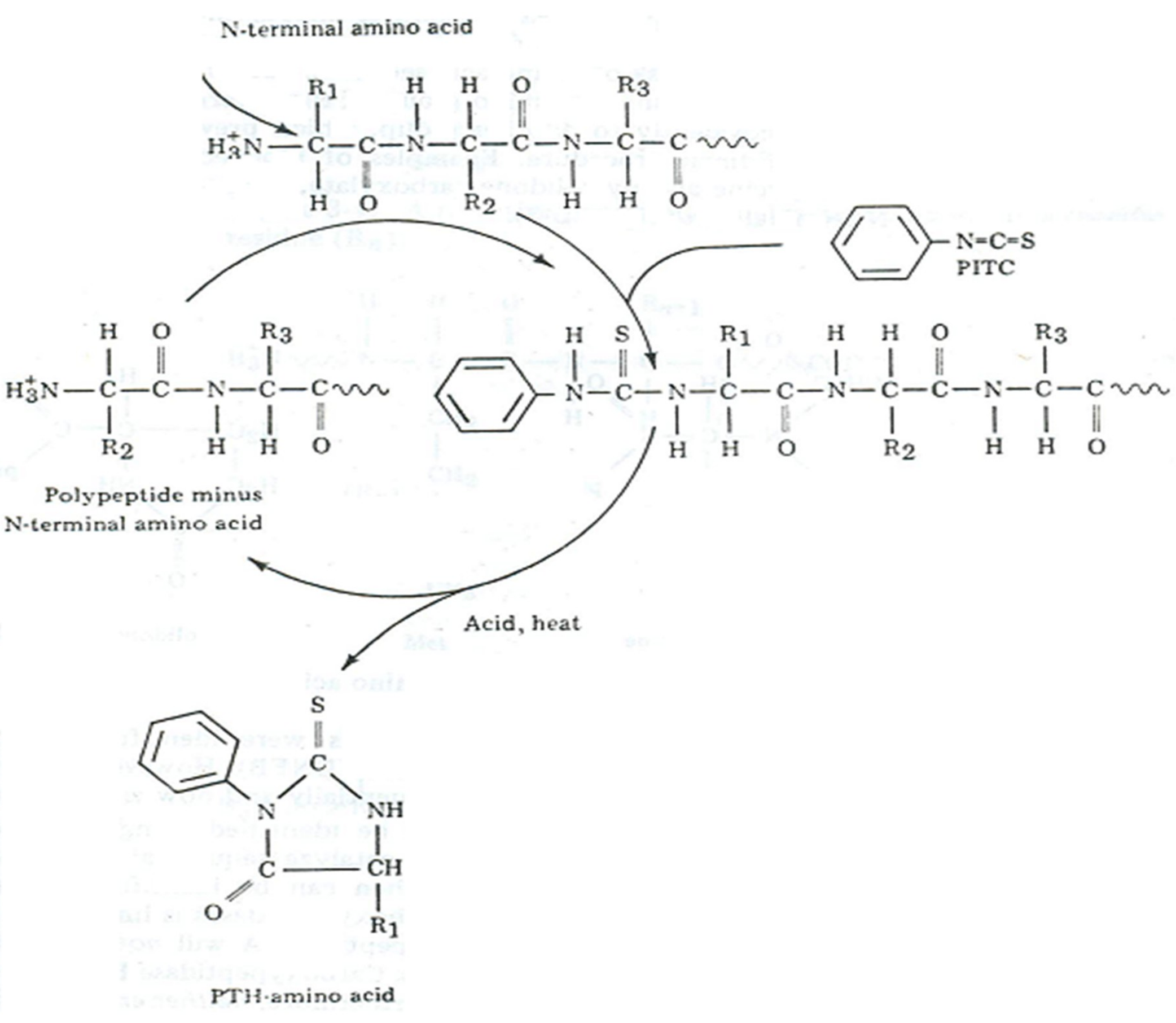


- Dansyl chloride (dimethylaminonaphthylenesulfonyl chloride)
- Forms a highly fluorescent derivative of the amino-terminal amino acid
- Identified by chromatography and fluorescence detection after acid hydrolysis
- Highly sensitive
- Best choice when the amount of protein is limited

Determining the N-Terminal Amino Acid

- Edman degradation
  - phenylisothiocyanate (phenyl-N=C=S) adds to N-terminus then acid treatment cleaves the N-terminal amino acid as a PTH-derivative
  - the remaining protein chain is intact and the cycle can be repeated (under ideal conditions the sequence of 30-60 amino acids can be determined




N-terminal amino acid
N=C=S
PITC
Polypeptide minus
N-terminal amino acid
Acid, heat
PTH-amino acid

Determining theC-Terminal Amino Acid

- Hydrazinolysis
  - hydrazine NH_2-NH_2 at 100°C cleaves all peptide bonds forming hydrazides except for the carboxyl terminal
- Carboxypeptidases
  - enzymatic removal of C-terminus amino acid R_n (provided $R_n \neq$ Pro and $R_{n-1} \neq$ Pro)

Peptide Fragments

- After determining the amino acid composition and the N & C-terminal amino acids, at least two different sets of protein fragments are needed for sequencing
- Why is the protein broken into fragments? Why isn't the protein sequenced directly?
- The sequencing methods currently available are only accurate for peptides up to about 20-30 amino acids, 60 under ideal conditions

Why 2 Sets of Fragments?

- Why can't the entire protein amino acid sequence be determined from a single set of peptide fragments obtained by cleavage with a single reagent?
- There's no way to determine how the fragments are connected with just one set
- A second or third set of fragments are used to deduce how the fragments are connected by identification and comparison of overlapping sequnces

Protein Cleavage Reagents

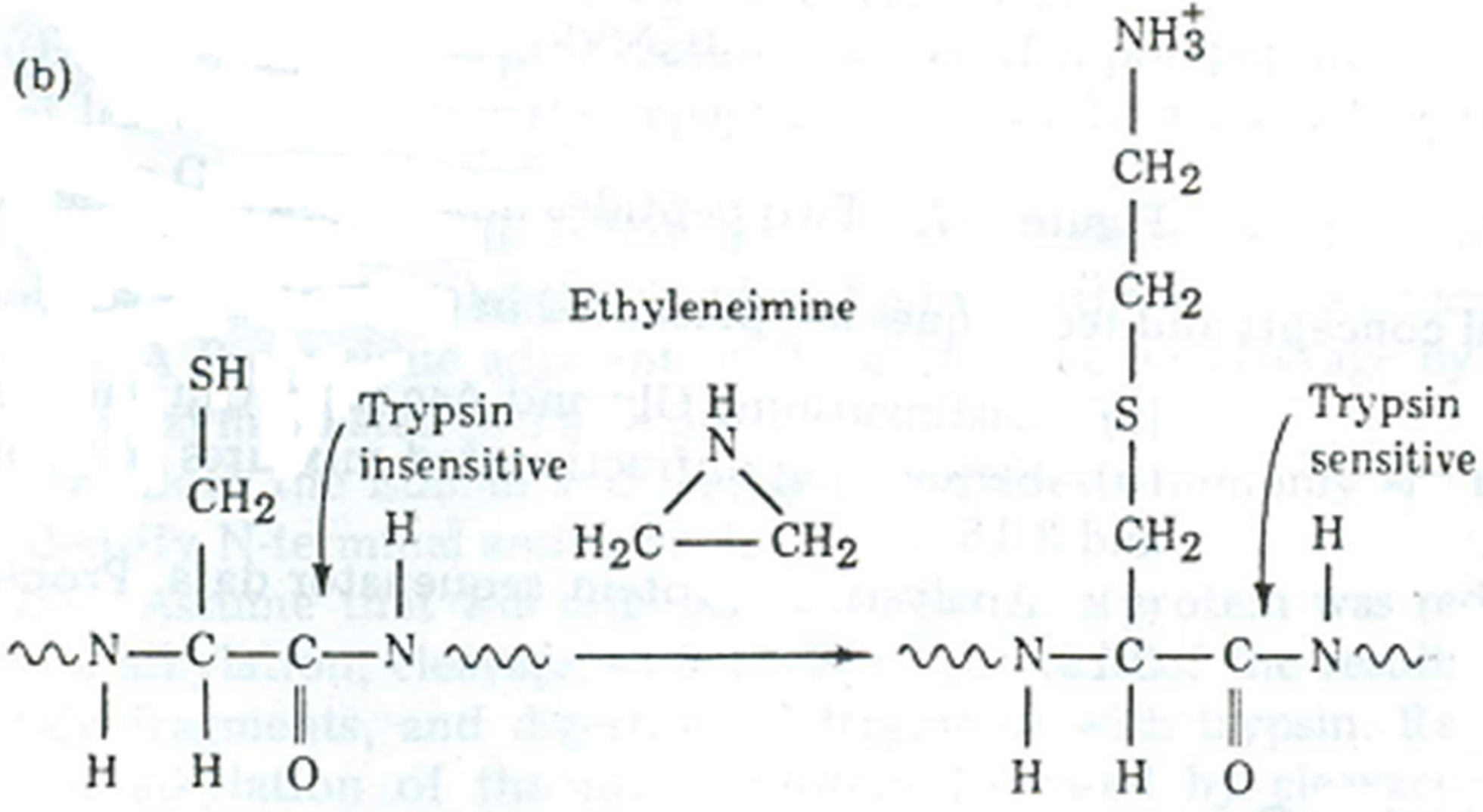
- Chemical or enzymatic reagents can be used to prepare protein fragments
- The most commonly used reagents are:
- cyanogen bromide (CNBr)
- various enzymes including
  - trypsin
  - chymotrypsin

Trypsin & Chymotrypsin

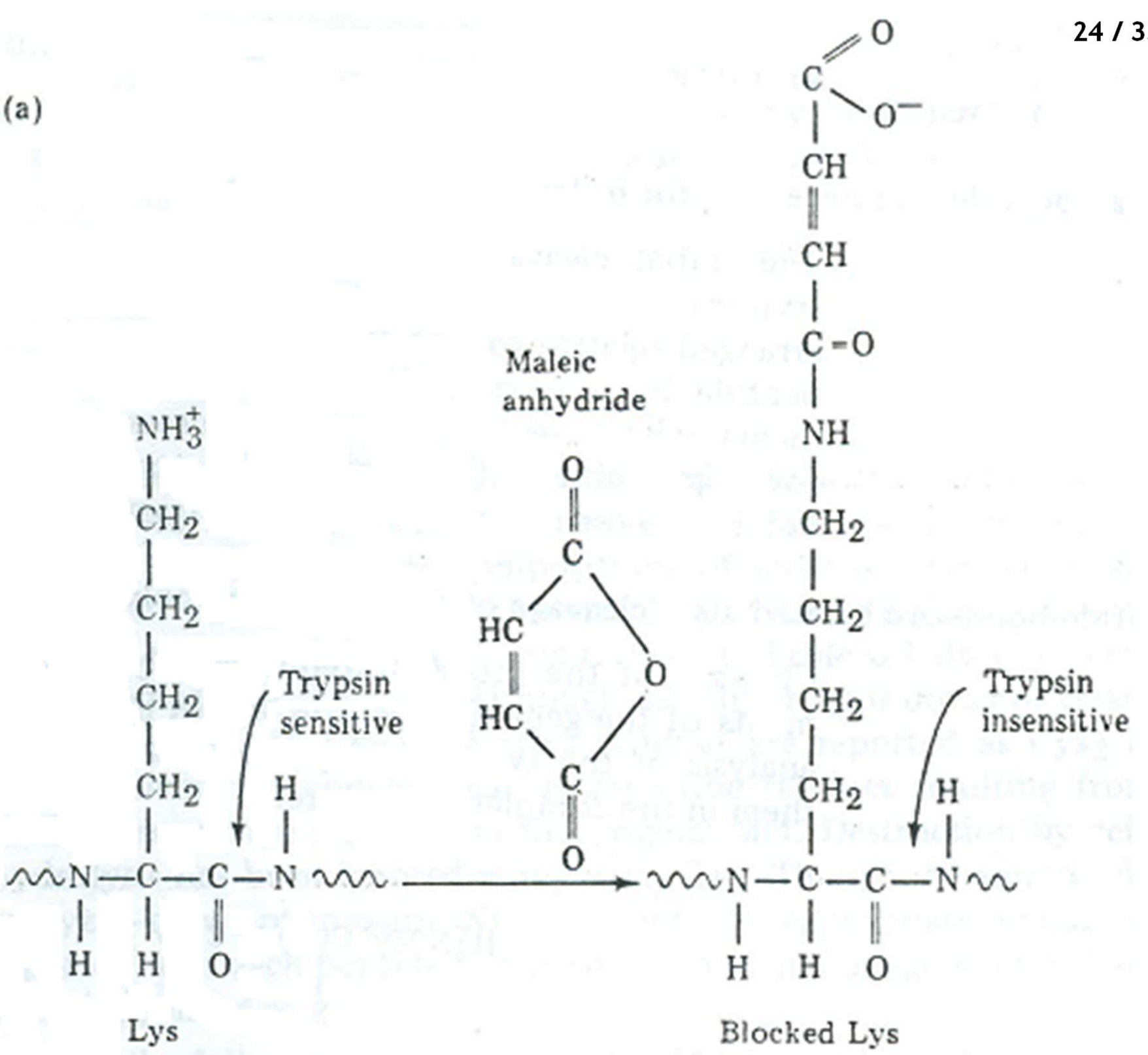
- Where in the protein sequence do the enzymes, trypsin and chymotrypsin cleave protein chains?
- trypsin cleaves at the carboxyl side of amino acids (R_n) with positively charged side chains such as lysine and arginine (provideed that $R_{n+1} \neq$ Pro)
- chymotrypsin cleaves at the carboxyl side of amino acids (R_n) with aromatic side chains such as phenylalanine, tyrosine and tryptophan in addition to Leu (provideed that $R_{n+1} \neq$ Pro)

Modifications on Trypsin Cleavage

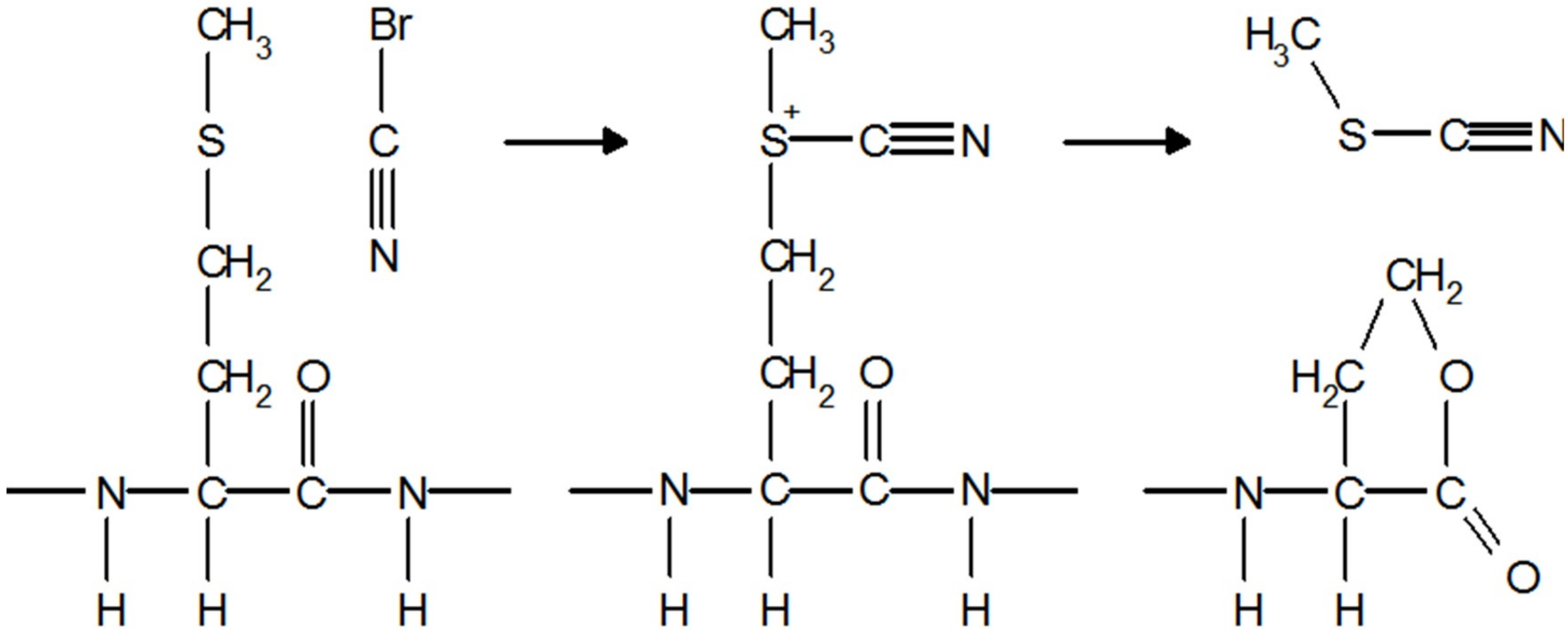
Modifications of the polypeptide substrate that alter the specificity of trypsin cleavage. (a) Reaction with maleic anhydride (maleylation) blocks trypsin attack at Lys residues. (b) Reaction with ethyleneimine (aminoethylation) permits trypsin attack at Cys residues.



(a)


Maleic anhydride
Trypsin sensitive
Trypsin insensitive
Lys
Blocked Lys

Cyanogen Bromide (CNBr)



- At which amino acid in the protein sequence does the reagent, cyanogen bromide, cleave protein chains?
- At internal methionines by reaction with the methionine sulfur as illustrated above (Met is converted to Homoserine Lactone(HSL)

Example Problem 1

Given an unknown peptide, UkP, determine the sequence from the following data.

1. Amino acid analysis (6N HCl, 24 hrs, 100° C) gave the results shown.

Glu	1	Ser	1	Ala	2
Ile	1	Tyr	1	Arg	2
Met	1	Val	1	Gly	2
Pro	1	Asp	2		

Example Problem 1, cont.

2. Treatment of UkP with α-chymotrypsin gave two peptides which were purified & analyzed.

C-1	Ala, Arg, Gly, Asp, Ser, Tyr	(DNP-Ala)
C-2	Ala, Arg, Glu, Gly, Ile, Asp, Met, Pro, Val	(DNP-Gly)

Ala-___- ___- ___-___-Tyr-Gly-___-___- ___-___-___- ___-___-___

3. Treatment of C-1 and C-2, separately with trypsin each gave two peptides which were purified & analyzed.

C-1:	T-1	Asp, Ser, Tyr	(DNP-Asp)
	T-2	Ala, Arg, Gly	(DNP-Ala)

Ala-Gly-Arg - ___-___-Tyr-Gly-___-___- ___-___-___- ___-___-___

Ala-Gly-Arg-Asp-Ser-Tyr-Gly-___-___- ___-___-___- ___-___-___

Example Problem 1, cont.

Ala-Gly-Arg-Leu-Ser-Tyr-Gly-___-___- ___-___-___- ___-___-___

3. Continued:

C-2:	T-3	Arg, Gly, Asp, Pro	(DNP-Gly)
	T-4	Ala, Glu, Ile, Met, Val	(DNP-Val)

Ala-Gly-Arg-Asp-Ser-Tyr-Gly-___-___-Arg-Val-___- ___-___-___

4. Treatment of T-4 with cyanogen bromide gave two peptides which were purified & analyzed.

T-4:	CN-1	Glu, Val, HSer	
	CN-2	Ile, Ala	(DNP-ILE)

Ala-Gly-Arg-Asp-Ser-Tyr-Gly-___-___-Arg-Val-Glu-Met-___-___

Ala-Gly-Arg-Asp-Ser-Tyr-Gly-___-___-Arg-Val-Glu-Met-Ile-Ala

Example Problem 1, cont.

5. Treatment of CN-2 with hydrazine and analysis gave Ala as the free amino acid (confirms previous data).

6. Mild acid hydrolysis of the original peptide gave six small peptides which were purified and analyzed.

A-1	Arg, Asp, Ser, Tyr	A-5	Ala, Ile, Met
A-2	Arg, Glu, Asp, Val	A-6	Arg ,Asp, Gly
A-3	Ala, Gly	A-7	Gly, Asp, Pro, Tyr
A-4	Gly, Pro	A-8	Glu, Met, Val

Ala-Gly-Arg-Asp-Ser-Tyr-Gly-___-___-Arg-Val-Glu-Met-Ile-Ala

-Arg-Asp-Ser-Tyr- -Asp-Arg-Val-Glu-
Ala-Gly- -Gly- Pro- -Met-Ile-Ala

-Gly-Arg-Asp- -Tyr-Gly-Pro-Asp- -Val-Glu-Met-

Ala-Gly-Arg-Asp-Ser-Tyr-Gly-Pro-Asp-Arg-Val-Glu-Met-Ile-Ala

Example Problem 2

Samples of an unknown peptide are subjected to the two treatments below. After each treatment the resultant peptides are purified and sequenced by Edman's method to provide the data given. Determine the sequence of the unknown peptide using the overlapping peptide method.

1. Cyanogen bromide:

Asp-Ile-Lys-Gln-Hser
Lys
Lys-Phe-Ala-Hser
Tyr-Arg-Gly-Hser

2. Trypsin:

Gln-Met-Lys
Gly-Met-Asp-Ile-Lys
Phe-Ala-Met-Lys
Tyr-Arg

___-___-___-___-___-___-___-___-___-___-___-___-___-___

Example Problem 2, cont.

1. Cyanogen bromide:

x Asp-Ile-Lys-Gln-Hser
x Lys
x Lys-Phe-Ala-Hser
x Tyr-Arg-Gly-Hser

2. Trypsin:

x Gln-Met-Lys
x Gly-Met-Asp-Ile-Lys
x Phe-Ala-Met-Lys
x Tyr-Arg

Asp-Ile-Lys-Gln-Hser

Lys

Gly-Met-Asp-Ile-Lys

Tyr-Arg-Gly-Hser

Gln-Met-Lys

Tyr-Arg

Lys-Phe-Ala-Hser

Phe-Ala-Met-Lys

Tyr-Arg-Gly-Met-Asp-Ile-Lys-Gln-Met-Lys-Phe-Ala-Met-Lys

Problem 3

Deduce the sequence of amino acids in a peptide from the following information:

(a) Complete acid hydrolysis yielded **ala+arg+2ser+lys+phe+met+trp+pro**.

(b) Treatment with **dinitrofluorobenzene (DNFB, the Sanger reagent)** followed by complete acid hydrolysis yielded dinitrophenylalanine (DNP-**ala**) and ε–dinitrophenyllysine (ε-DNP-**lys**) as the only DNP derivatives.

(c) No C-terminal a. acid was released upon treatment with **carboxypeptidases**.

(d) Treatment with **cyanogen bromide (CNBr)** yielded two peptides.

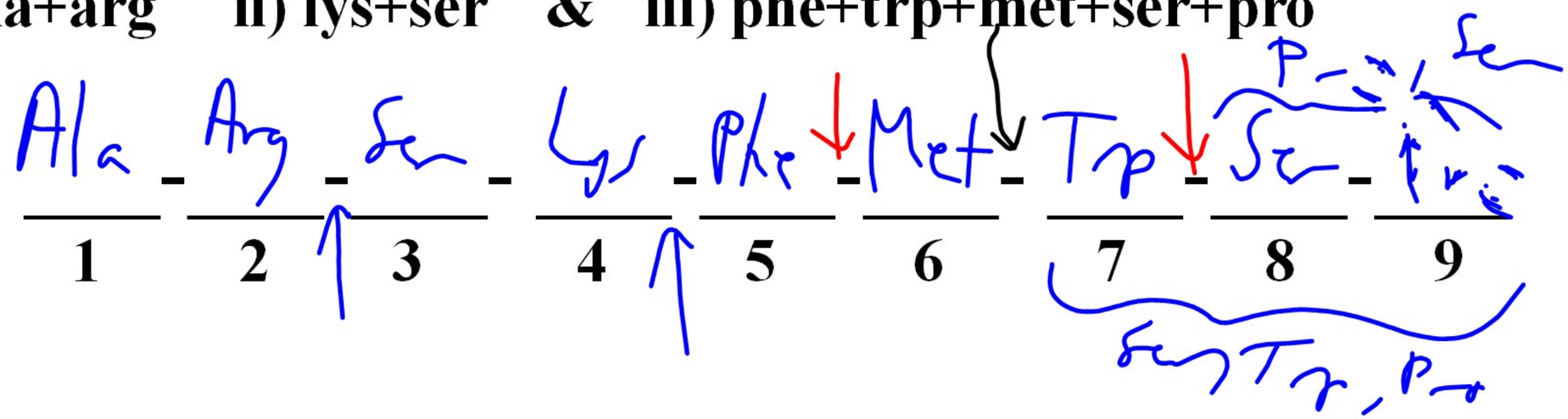
i) ser+trp+pro & **ii)** the remaining amino acids (including the second ser).

(e) Treatment with **chymotrypsin** yielded three peptides.

i) ser+pro **ii) met+trp** & **iii) phe+lys+ser+arg+ala**

(f) Treatment with **trypsin** yielded three peptides.

i) ala+arg **ii) lys+ser** & **iii) phe+trp+met+ser+pro**



Problem 4

Deduce the sequence of amino acids in a peptide from the following information:

(a) Composition = **phe+pro+glu+2lys**;

(b) Treatment with **Edman reagent** yielded PTH-**glu**tamate; and

(c) Trypsin and **carboxypeptidases** did not release any smaller peptides or amino acids.