Chapter 3

3.4 The structure of proteins

Amino acid sequence of bovine insulin

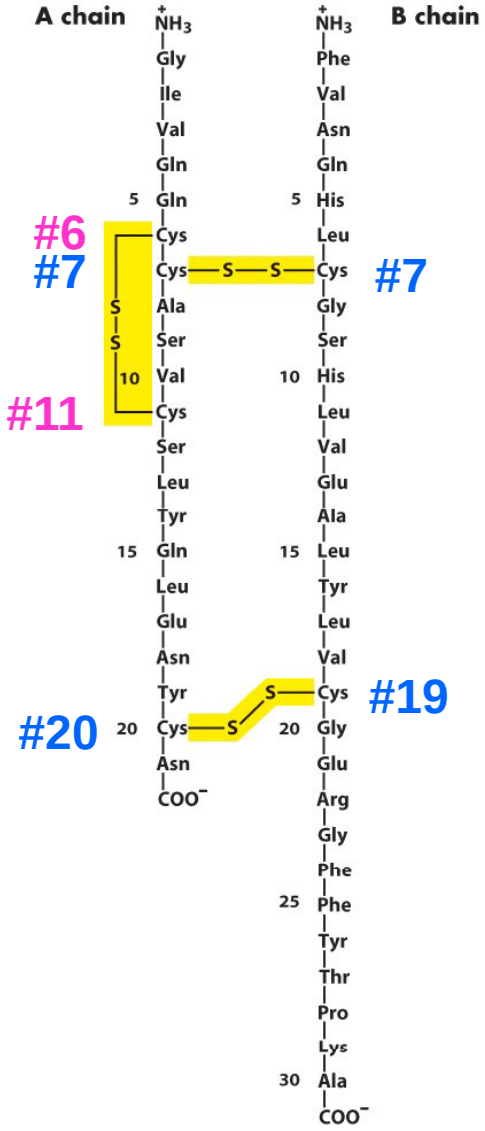
Sanger and colleagues took more than 10 years to know the primary structure of insulin.

1945

1955

1 intra-chain disulfide bond

2 inter-chain disulfide bonds



Glycyl chain

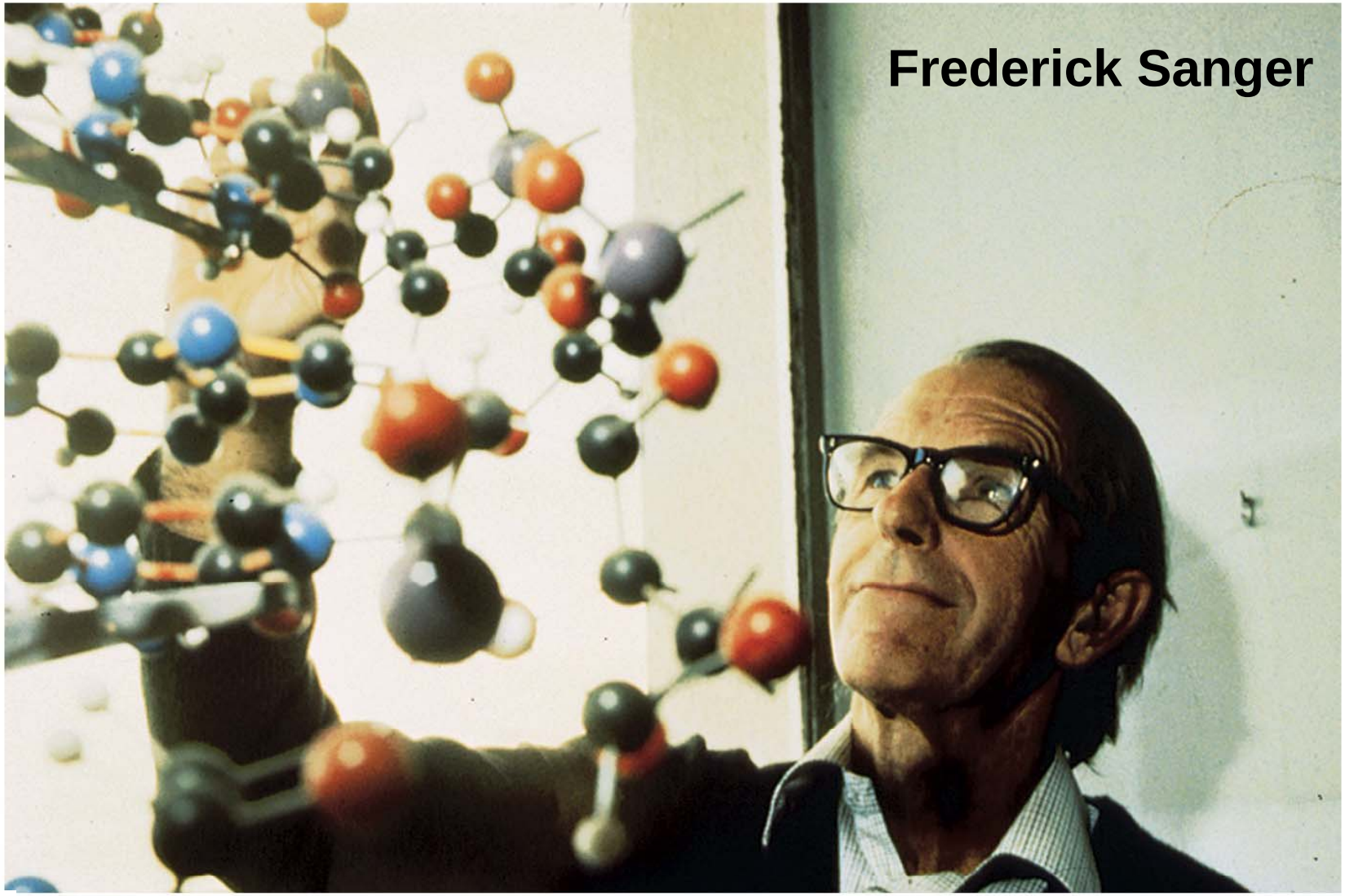
A chain

Gly-Ile-Val-Glu-Gln-Cys-Cys-Ala-Ser-Val-Cys-Ser-Leu-Tyr-Gln-Leu-Glu-Asn-Tyr-Cys-Asn

Phenyl chain

B chain

Phe-Val-Asn-Gln-His-Leu-Cys-Gly-Ser-His-Leu-Val-Glu-Ala-Leu-Tyr-Leu-Val-Cys-Gly-Glu-Arg-Gly-Phe-Phe-Tyr-Thr-Pro-Lys-Ala



1958 The Nobel Prize in Chemistry 1958 was awarded to Frederick Sanger *"for his work on the structure of proteins, especially that of insulin*

1980 the other half jointly to Walter Gilbert and Frederick Sanger *"for their contributions concerning the determination of base sequences in nucleic acids"*

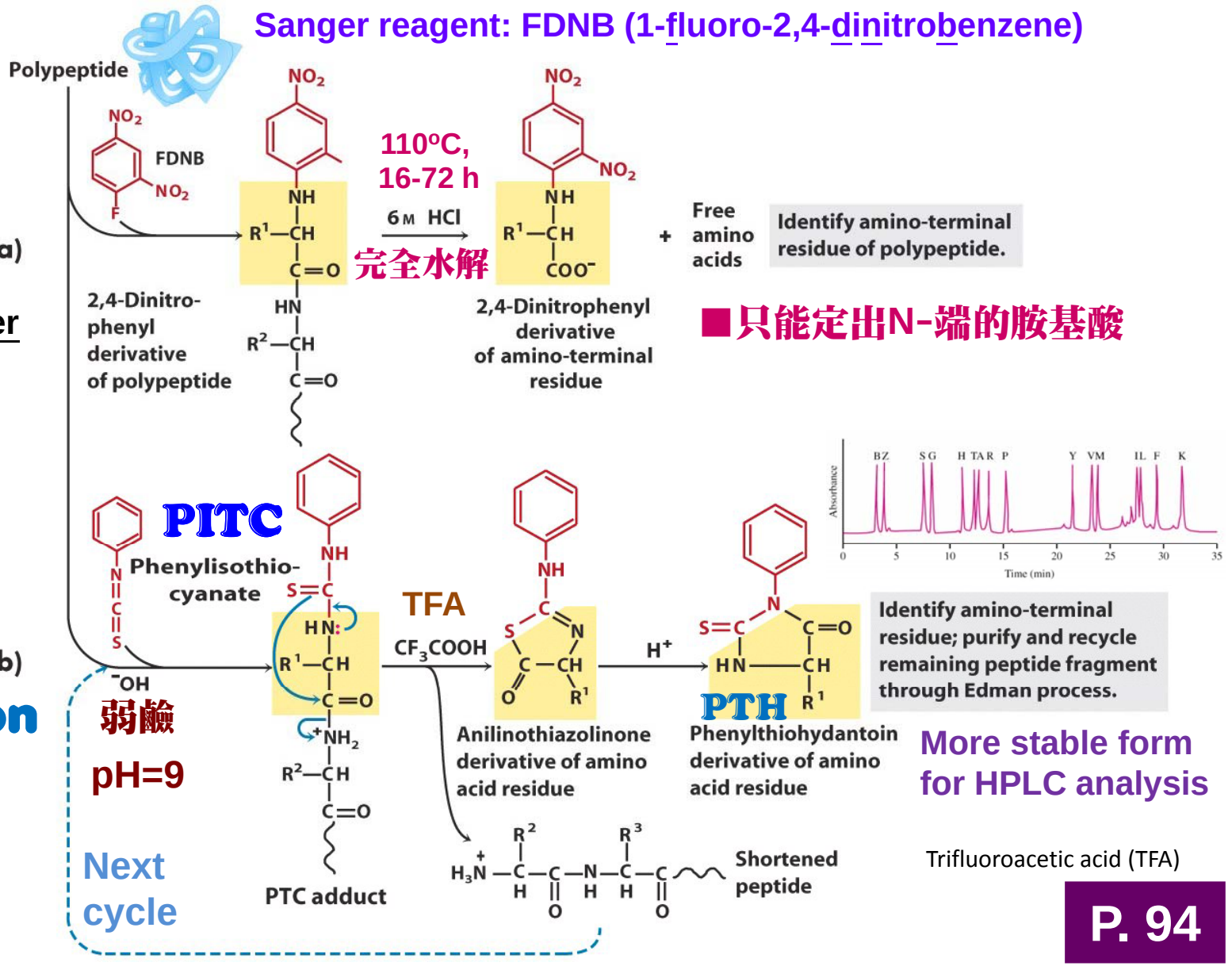
Steps in sequencing a polypeptide

Sanger's method (a)

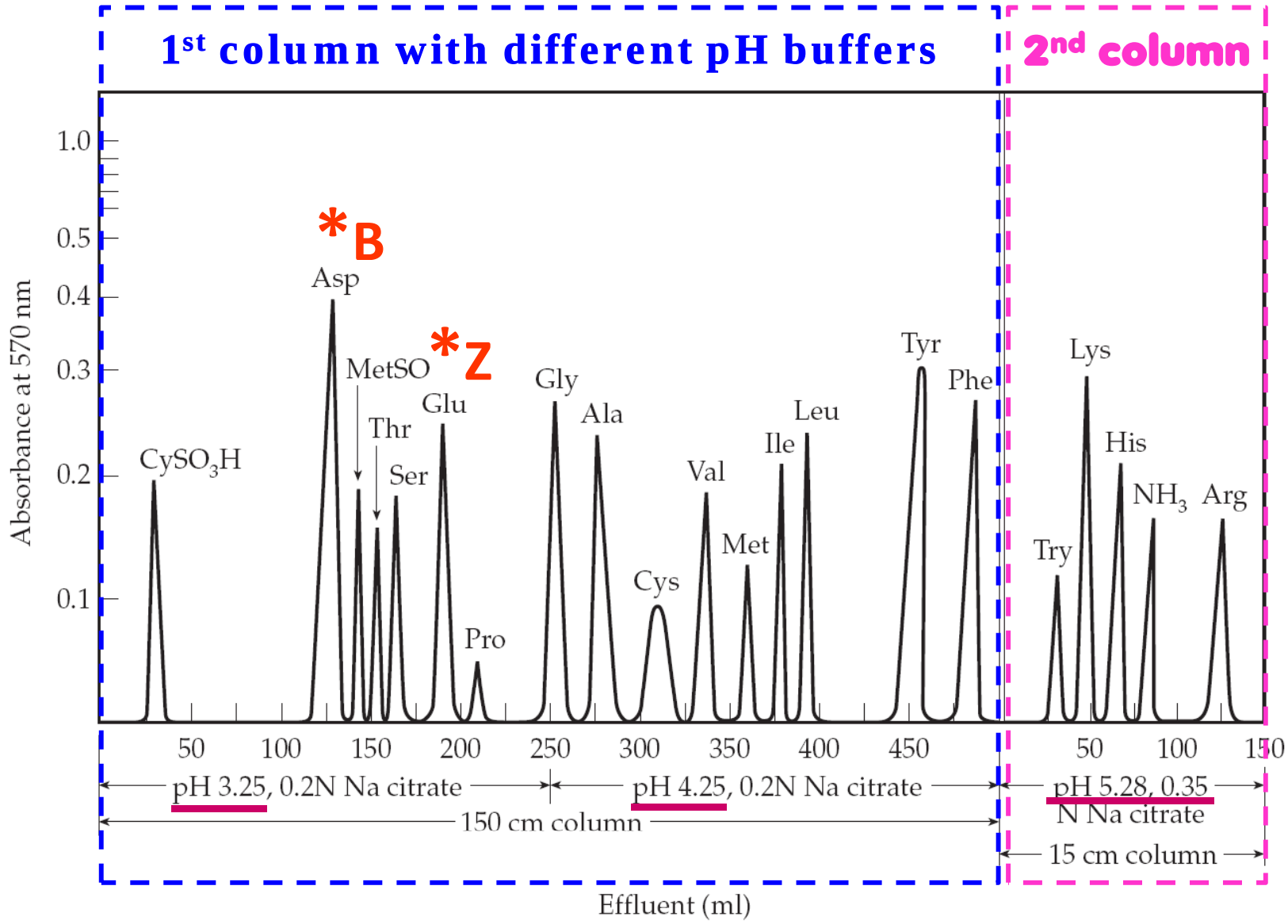
Frederick Sanger

Edman degradation (b)

Pehr Edman

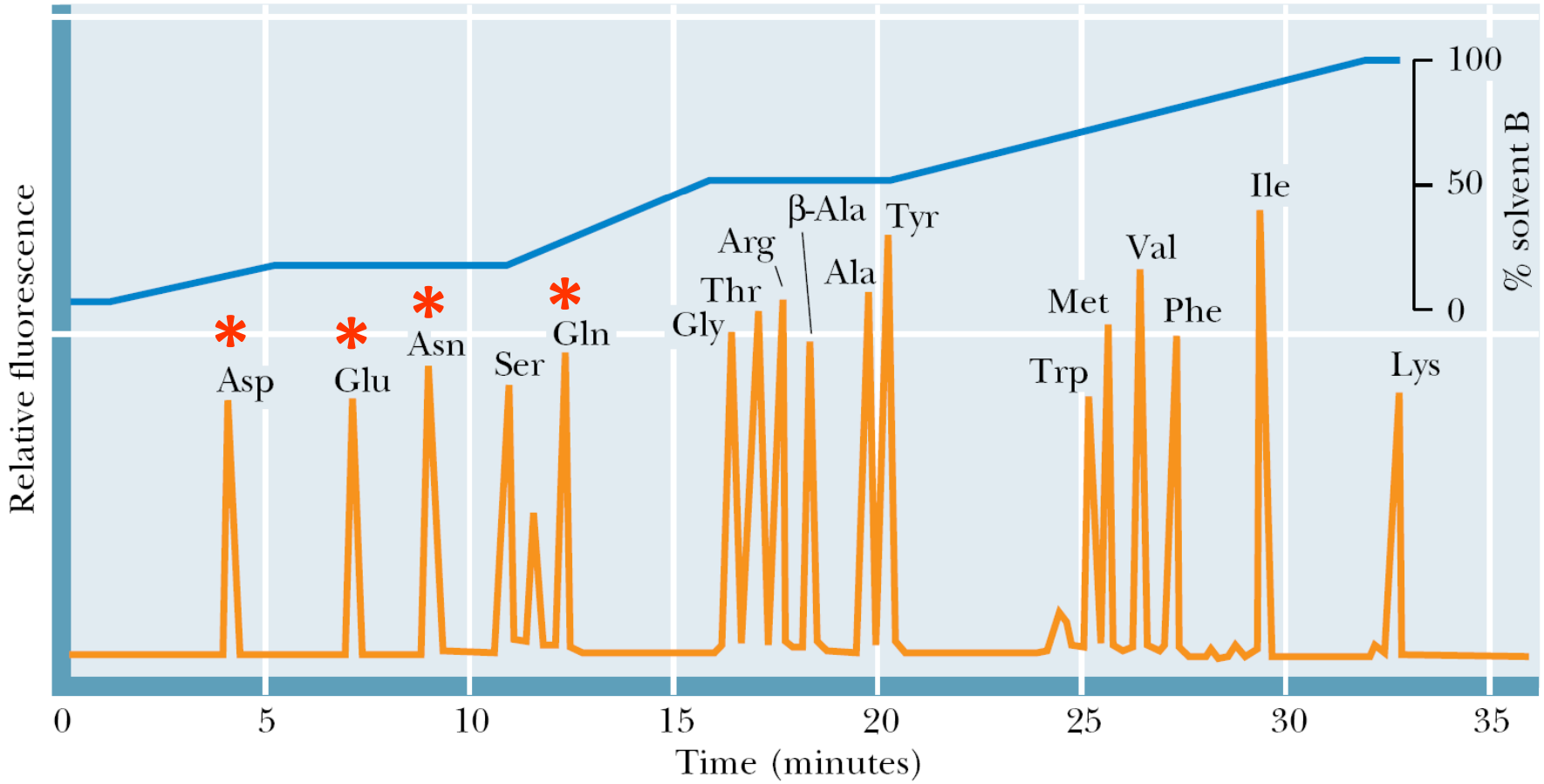


Separation of amino acids by cation-exchange chromatography on a sulfonated polystyrene



Biochemistry: The Chemical Reactions of Living Cells (2nd edition, 2001), David E. Metzler

Amino acid analysis by HPLC after precolumn derivatization with OPA



Retention time

elapsed time between the time of injection of a solute and the time of elution of the peak maximum of that solute.

Biochemistry 2nd edition (1999), Garrett and Grisham

- **Edman degradation procedure** - Determining one residue at a time from the N-terminus

(1) Treat peptide at weak alkaline with PITC which reacts with the N-terminus to form a PTC-peptide

(2) Treat with trifluoroacetic acid (TFA) to selectively cleave the N-terminal peptide bond

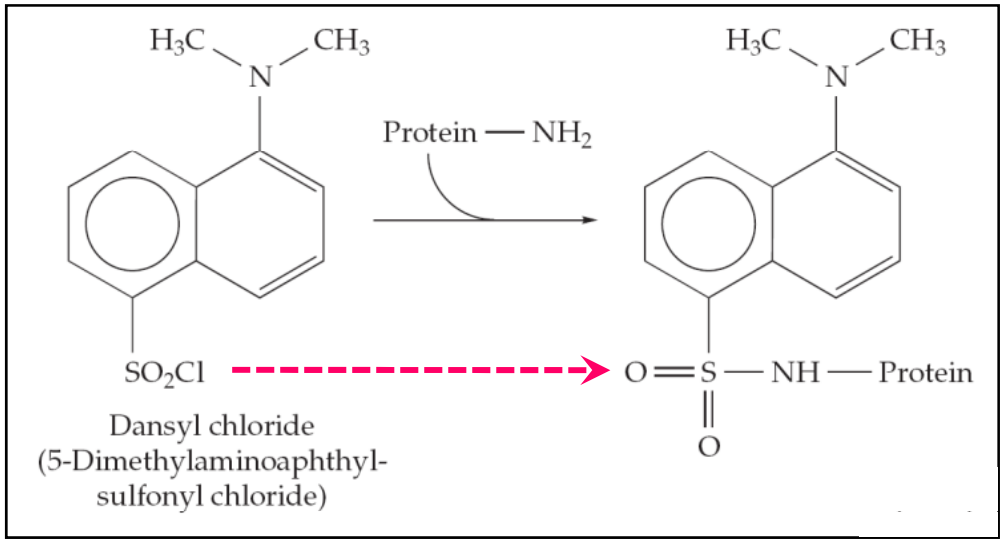
(3) Separate N-terminal derivative from peptide

(4) Convert derivative to PTH-amino acid

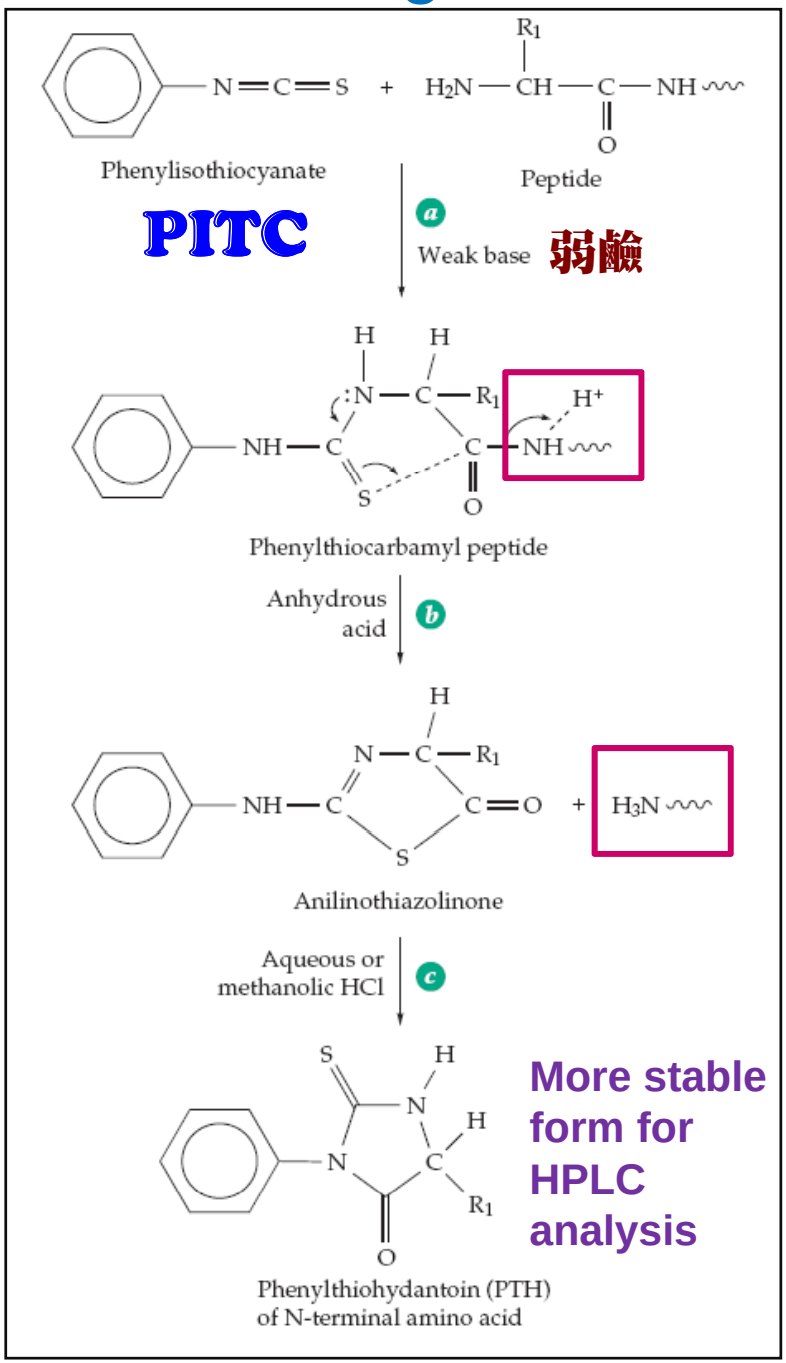
(5) Identify the amino acid by HPLC

Reactions of amino groups

Dansyl chloride

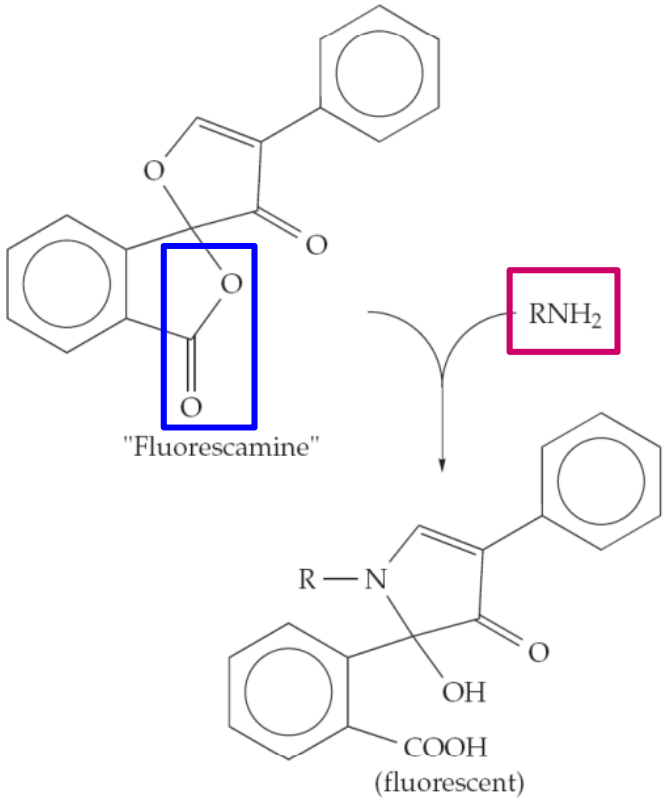


Edman degradation

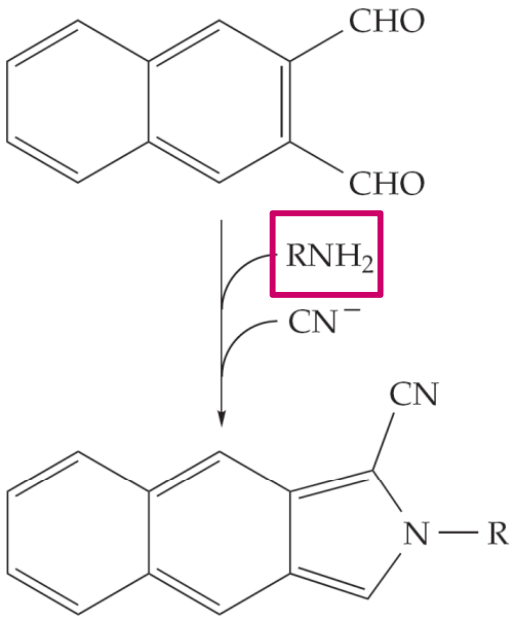


Reactions of amino groups

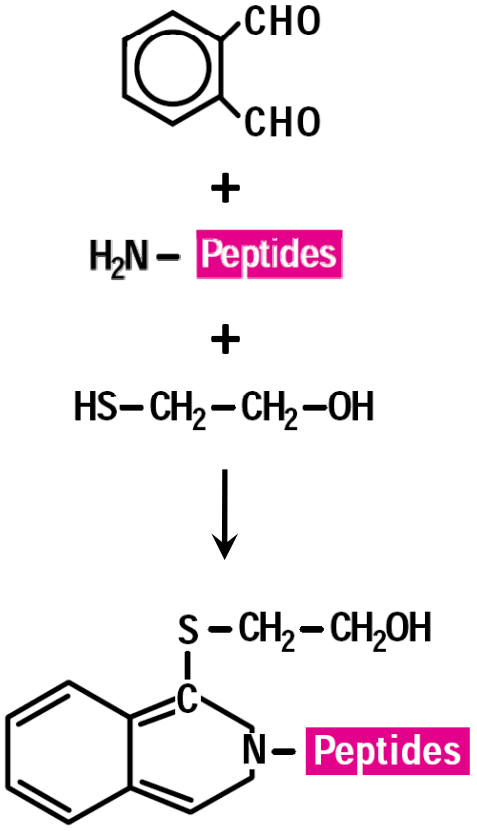
Fluorescamine



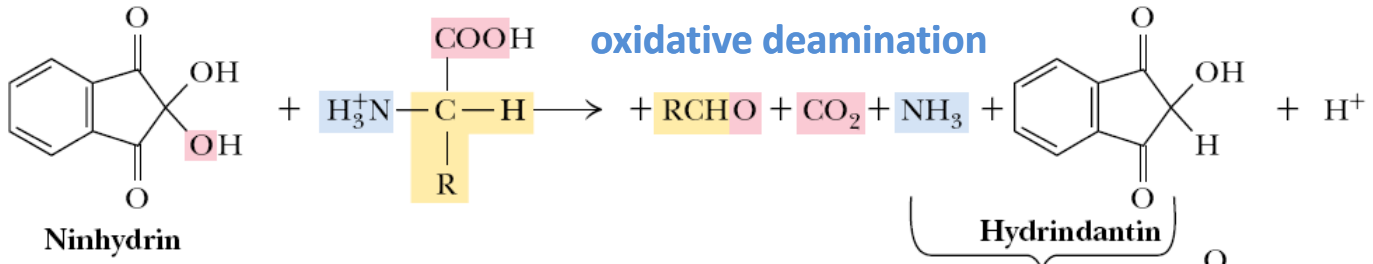
Naphthalene 2,3 dicarboxaldehyde



O-Phthaldialdehyde OPA

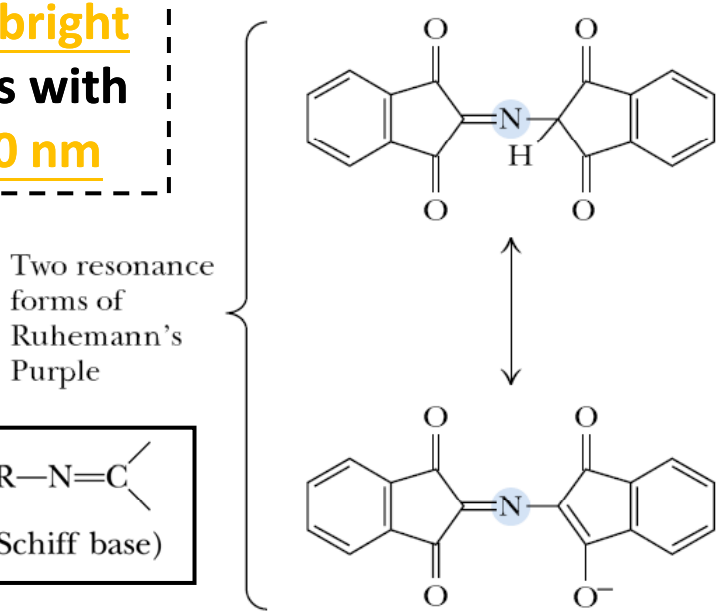


The pathway of the ninhydrin Reaction



consumes two molecules of ninhydrin

α-Imino acids, such as proline and hydroxyproline, give bright yellow ninhydrin products with absorption maxima at 440 nm



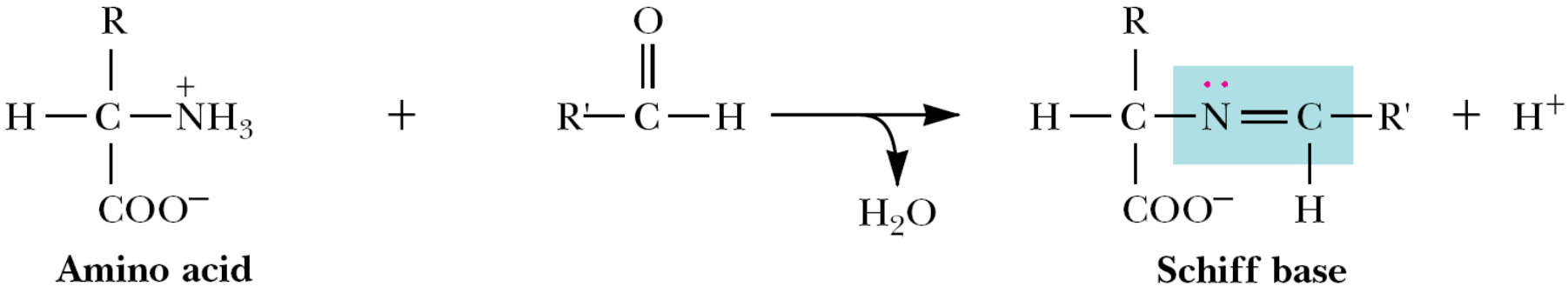
Ruhemann's Purple that absorbs light at 570 nm.

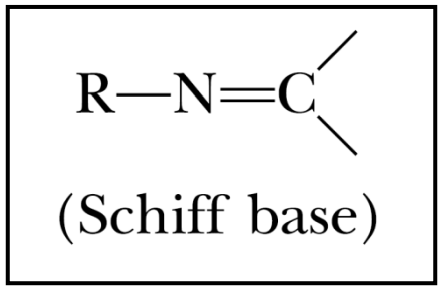
R—N=C (Schiff base)

Discovered by Siegfried Ruhemann in 1910

Biochemistry 2nd edition (1999), Garrett and Grisham

SCHIFF BASE

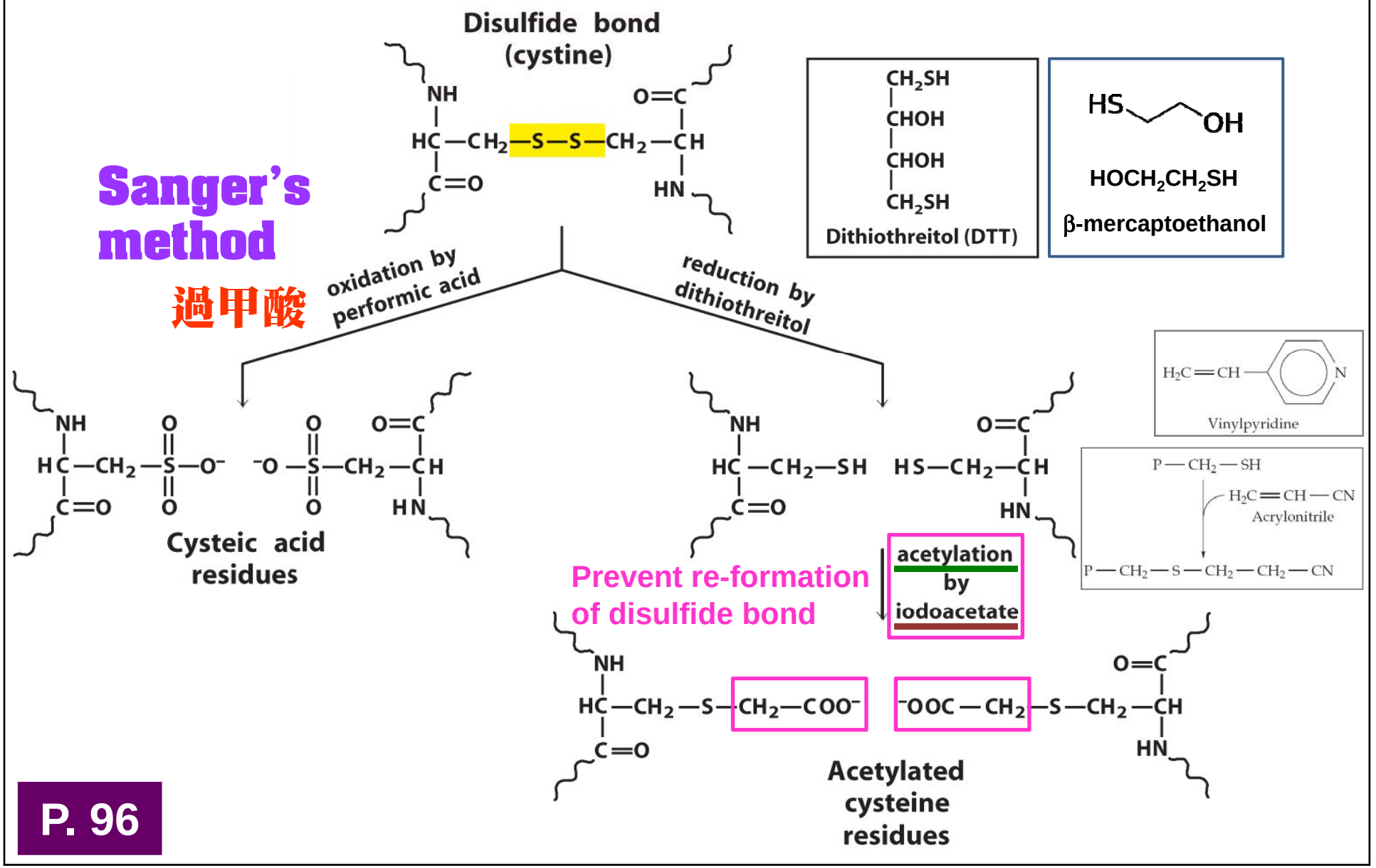

R
H — C — NH_3^+
COO^-
Amino acid
+
O
R' — C — H
H_2O
R
H — C — N = C — R'
COO^-
H
Schiff base
+ H^+


R—N=C
(Schiff base)

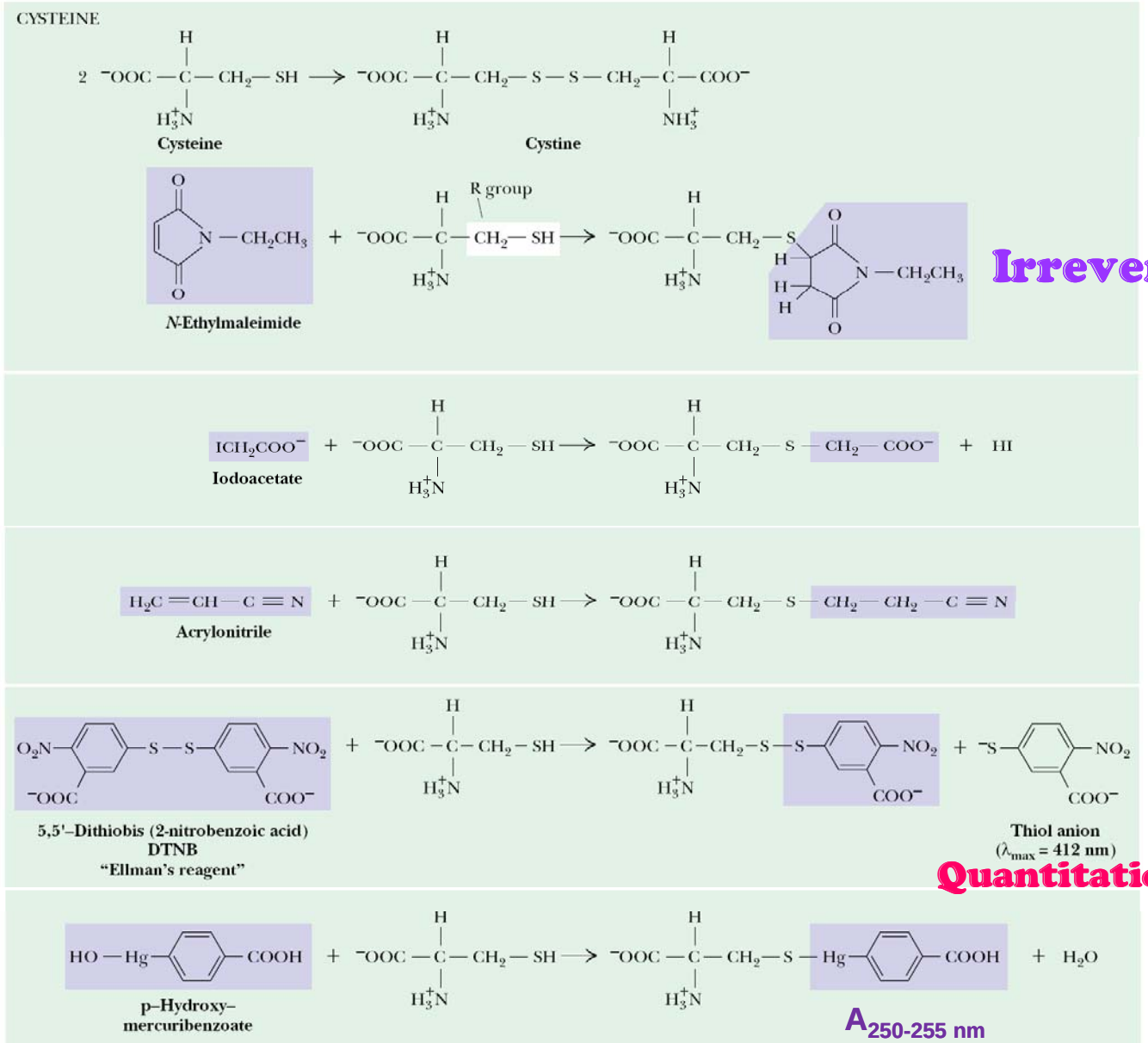
DETERMINING AMINO ACID COMPOSITION AND SEQUENCE

- **Breaking disulfide bonds**
  To permit isolation of the PTH-cysteine during the Edman procedure, and to separate peptide chains
- **Cleaving the polypeptide chain**
  Proteins may be too large to be sequenced completely by the Edman method
- **Sequencing the peptide**
- **Mapping the peptide fragments**
- **Locating disulfide bonds**

Breaking disulfide bonds in proteins



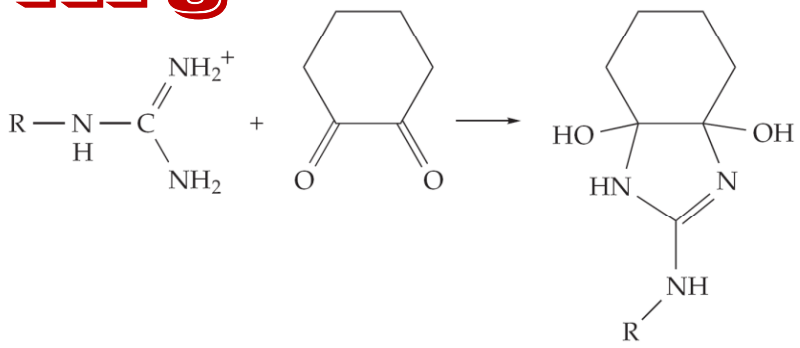
Reactions of SH groups



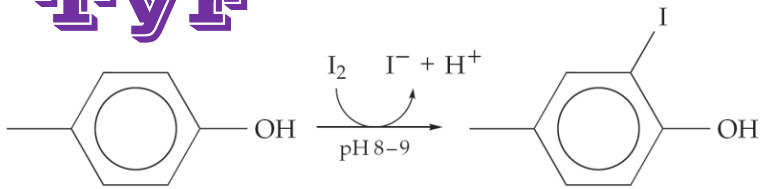
Biochemistry 2nd edition (1999), Garrett and Grisham

Reactions of other side chains

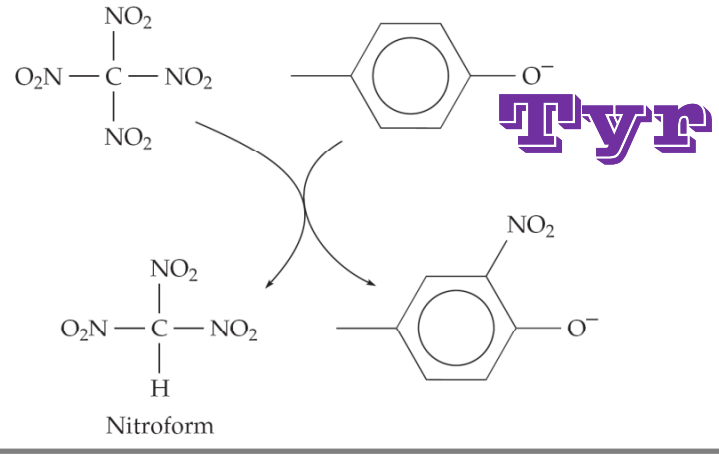
Arg



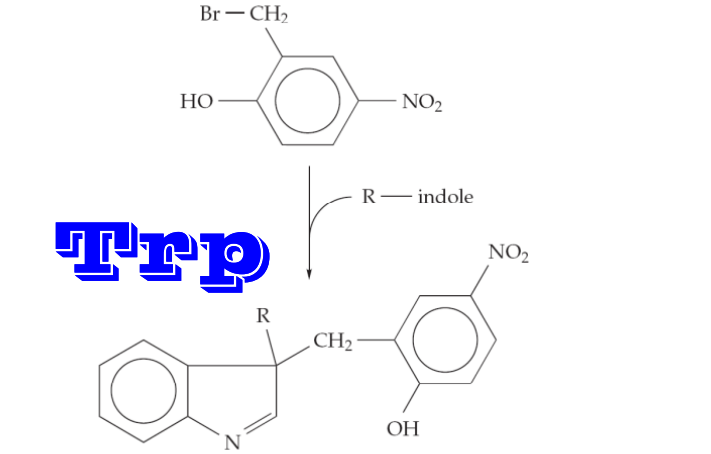
Tyr



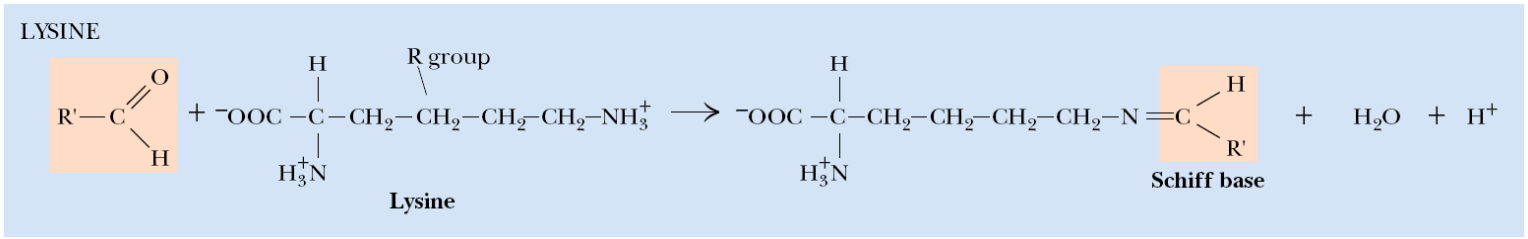
Tyr



Trp



Lys



Biochemistry 2nd edition (1999), Garrett and Grisham

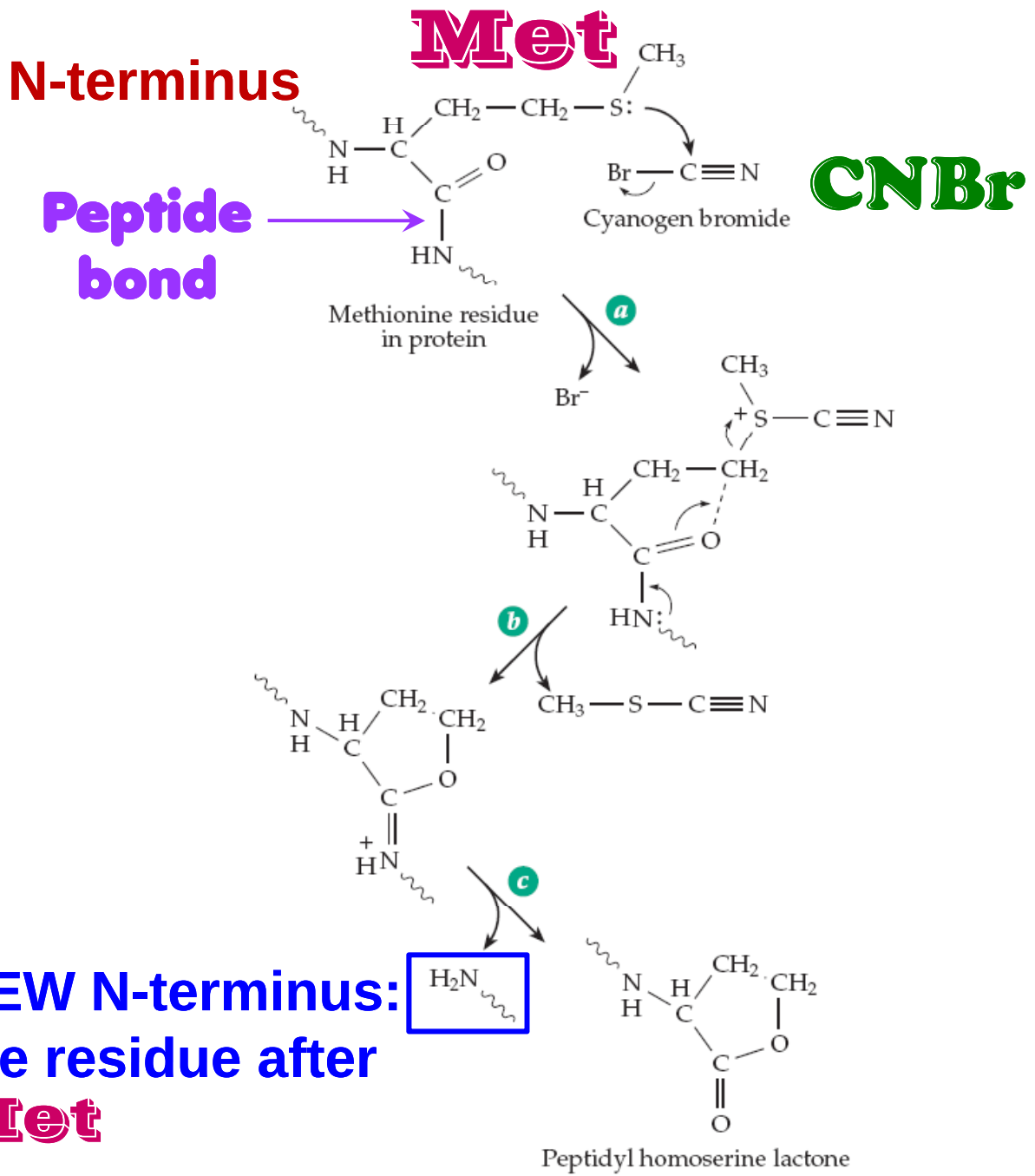
The specificity of some common methods for fragmenting polypeptide chains

P. 96

*Reagent (biological source)**	*Cleavage points*†	
Trypsin (bovine pancreas)	Lys, Arg (C)	
Submaxillarus protease (mouse submaxillary gland)	Arg (C)	
Chymotrypsin (bovine pancreas)	Phe, Trp, Tyr (C)	
Staphylococcus aureus V8 protease (bacterium *S. aureus*)	Asp, Glu (C)	**in 50 mM ammonium bicarbonate V8 cleaves only at Glu**
Asp-*N*-protease (bacterium *Pseudomonas fragi*)	Asp, Glu (N)	
Pepsin (porcine stomach)	Phe, Trp, Tyr (N)	
Endoproteinase Lys C (bacterium *Lysobacter enzymogenes*)	Lys (C)	
Cyanogen bromide N≡C—Br	Met (C)	**CNBr**

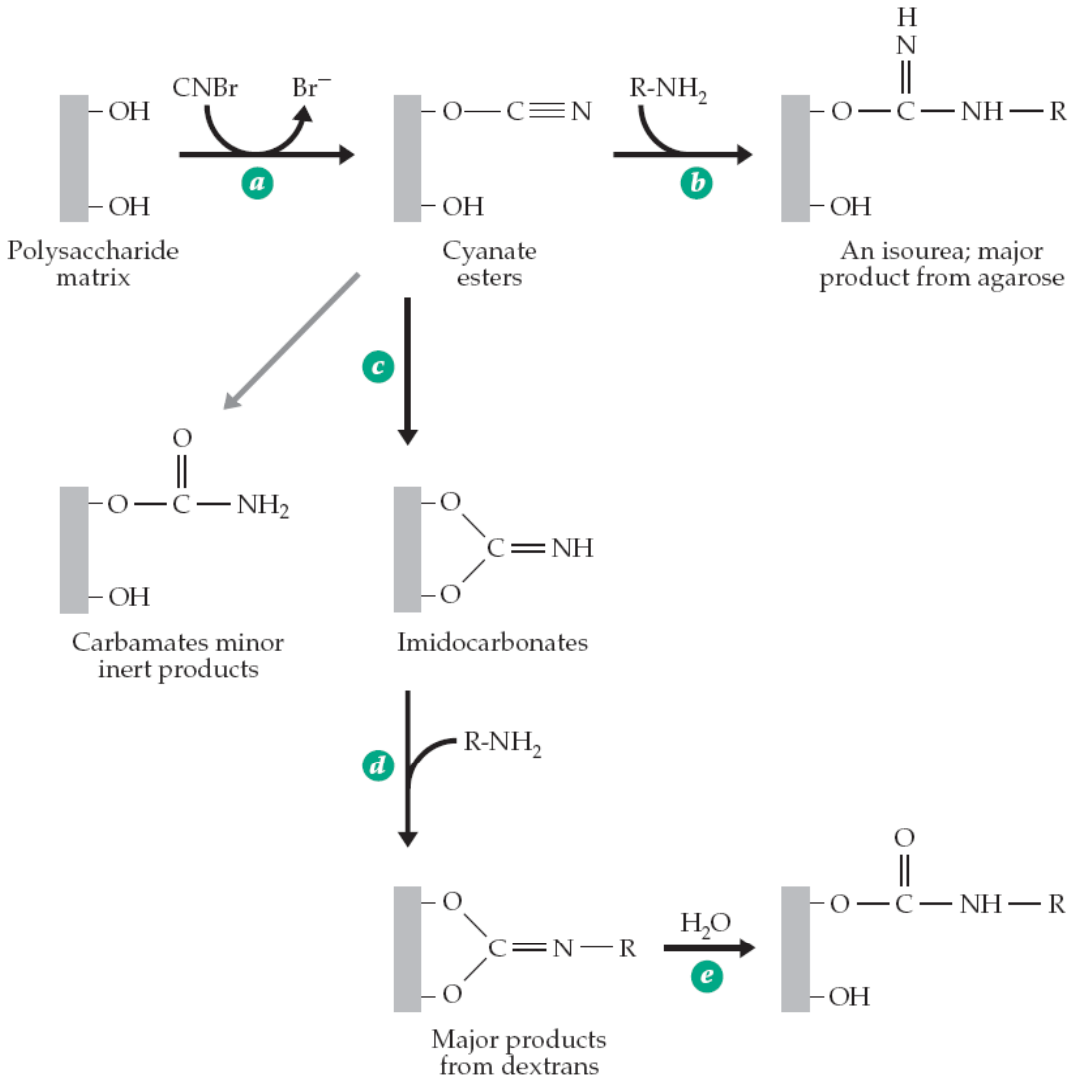
Enteropeptidase/enterokinase : D-D-D-D-K-X

Protein cleavage by CNBr



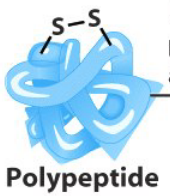
Conjugation of proteins with carbohydrate by CNBr

Affinity chromatography



Biochemistry: The Chemical Reactions of Living Cells (2nd edition, 2001), David E. Metzler

Cleaving protein and sequencing and ordering the peptide fragments



Polypeptide

Procedure

hydrolyze; separate amino acids →

Result

A	5	H	2	R	1
C	2	I	3	S	2
D	4	K	2	T	1
E	2	L	2	V	1
F	1	M	2	Y	2
G	3	P	3		

a.a. composition

Conclusion

Polypeptide has 38 amino acid residues. Trypsin will cleave three times (at one R (Arg) and two K (Lys)) to give four fragments. Cyanogen bromide will cleave at two M (Met) to give three fragments.

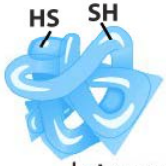
E (Glu) is amino-terminal residue.

react with FDNB; hydrolyze; separate amino acids →

2,4-Dinitrophenylglutamate detected

Sanger's method for N-terminus

reduce disulfide bonds (if present)



cleave with trypsin; separate fragments; sequence by Edman degradation →

- T-1 GASMALIK
- T-2 EGAAYHDFEPIDPR
- T-3 DCVHSD
- T-4 YLIACGPMTK

T-2 placed at amino terminus because it begins with E (Glu).

T-3 placed at carboxyl terminus because it does not end with R (Arg) or K (Lys).

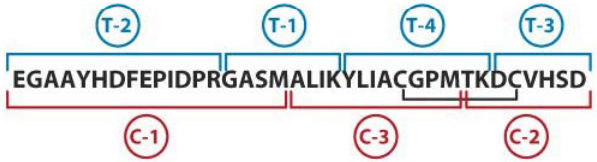
cleave with cyanogen bromide; separate fragments; sequence by Edman degradation →

- C-1 EGAAYHDFEPIDPRGASM
- C-2 TKDCVHSD
- C-3 ALIKYLIACGPM

C-3 overlaps with T-1 and T-4, allowing them to be ordered.

establish sequence →

Amino terminus



Carboxyl terminus

DETERMINING AMINO ACID COMPOSITION AND SEQUENCE

- Breaking disulfide bonds
- Cleaving the polypeptide chain
- Sequencing the peptide
- Mapping the peptide fragments
- Locating disulfide bonds

The Amino-acid Sequence in the Glycyl Chain of Insulin

1. THE IDENTIFICATION OF LOWER PEPTIDES FROM PARTIAL HYDROLYSATES

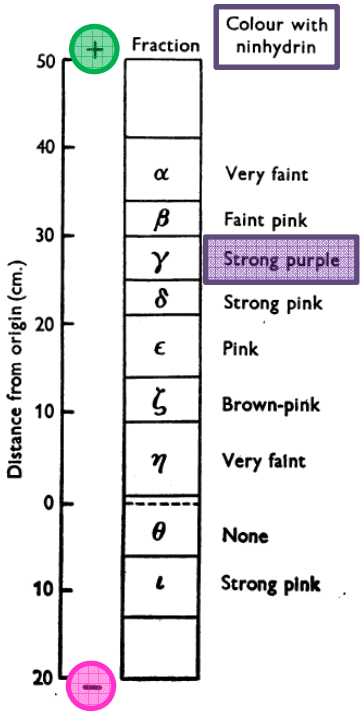
By F. SANGER* AND E. O. P. THOMPSON†
Biochemical Laboratory, University of Cambridge

(*Received* 16 *July* 1952)

Insulin has been shown to be composed of two types of polypeptide chains held together by —S—S— bridges. These chains may be separated after breaking the —S—S— bridges by oxidation to —SO_3H groups (Sanger, 1949*a*). The more acidic glycyl chain is found in the fraction (*A*) which is soluble at pH 6·5 and the more basic phenylalanyl chain in the insoluble fraction (*B*). By making use of the methods of Consden, Gordon & Martin (1947) for fraction-peptide, two spots were present, one with tyrosine and one with TyrX. In the original method a large excess of performic acid was used to oxidize the cystine residues. It has now been found that if only a slight excess is used the formation of TyrX can be avoided. The yield of the fraction *A* is somewhat lower, but for most purposes this drawback is offset by the advantage of having a simpler peptide mixture.

Insulin (1 g.) was dissolved in 8 ml. formic acid and treated with 0·48 ml. 30% (w/w) H_2O_2 (1·6 equiv. on the basis of the S content). After standing for 1 hr. 8 ml. water

To determine the *N*-terminal residues of the peptides, the DNP method has been used in most cases. Sanger & Tuppy (1951*a*) carried out the condensation between the peptide and 1:2:4-fluorodinitrobenzene (FDNB) in the presence of $NaHCO_3$. The salt had to be removed before paper chromatography could be carried out and this was achieved by extraction of the DNP peptide into ethyl acetate. In order to avoid this somewhat tedious step the reaction is now carried out in the presence of trimethylamine, which may be removed *in vacuo*. The identity of the *N*-terminal residue was deduced by Sanger & Tuppy from its absence in the hydrolysate of the DNP peptide. This clearly is less decisive than the identification of the actual DNP amino-acid



The Amino-acid Sequence in the Glycyl Chain of Insulin

2. THE INVESTIGATION OF PEPTIDES FROM ENZYMIC HYDROLYSATES

By F. SANGER* and E. O. P. THOMPSON†
Biochemical Laboratory, University of Cambridge

(*Received* 16 *July* 1952)

Experiment A_p (pepsin)

In a typical experiment 50 mg. fraction *A* were dissolved in 5 ml. 0·1 N-acetic acid and 2 mg. pepsin added. The mixture was incubated at 37° for 48 hr. After boiling to inactivate the enzyme, the solution was centrifuged to remove denatured pepsin and evaporated to dryness *in vacuo*. In some experiments the digestion was carried out in 0·01 N-HCl, but no difference in the peptides obtained could be detected.

No preliminary fractionation was necessary and samples of the hydrolysate were investigated directly on paper chromatograms. Where Whatman no. 4 filter paper was used an amount equivalent to 5 mg. of the original fraction *A* was applied, and for Whatman no. 3 paper 10–12·5 mg.

The peptides were investigated essentially by the methods described in the previous paper.

Experiment A_c (chymotrypsin)

Fraction *A* (50 mg.) was dissolved in 5 ml. water, 2·5 mg. chymotrypsin were added and the pH adjusted to 7·5 with dilute NH_3. The mixture was incubated at 37° for 24 hr. It was boiled to inactivate the enzyme and evaporated to dryness *in vacuo*. Before application to a chromatogram, the residue was taken up in a small volume of water, and centrifuged to remove insoluble material.

Experiments with carboxypeptidase

In preliminary experiments it was evident that the carboxypeptidase preparation contained traces of chymo-

Experiment A_p (pepsin)

Two chromatograms of the peptic hydrolysate are shown in Figs. 1 and 2. When each solvent was run only one length of the paper (Fig. 1) spots 3–9 were not sufficiently resolved to give clear results. It was therefore necessary to clamp a second folded sheet of paper on to the bottom edge during each run so that the chromatograms were developed two lengths in each direction (Fig. 2). The combined results are shown in Table 1.

Partial hydrolysis of peptides A_{p5}, A_{p14}, A_{p15}. Eluates of spot A_{p5} from two chromatograms on Whatman no. 3 paper, were hydrolysed with 12 N-HCl at 37° for 3 days and the hydrolysate was chromatographed on Whatman no. 4 paper. The results are shown in Fig. 3 and Table 2.

Eluates of spots A_{p14} and A_{p15} were treated with 1:2:4-fluorodinitrobenzene and the dinitrophenyl (DNP) derivatives hydrolysed 3 days with 12 N-HCl at 37°. The hydrolysates were extracted with ether, and in both cases bis-DNP-tyrosine was identified in the extract by using reversed phase paper chromatography. The aqueous solutions were chromatographed on Whatman no. 4 paper. Peptide A_{p14} gave on partial hydrolysis Glu.Leu ($A_{1\varepsilon 8}$) Leu.Glu ($A_{1\varepsilon 7}$) and Glu.Leu.Glu ($A_{1\delta 6}$) in addition to the free amino-acids and consequently had the structure Tyr.Glu.Leu.Glu. Peptide

Correspondence of DNA and amino acid sequences

P. 98

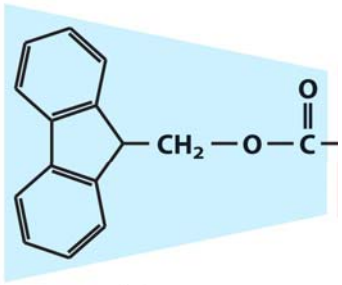
Amino acid sequence (protein) Gln – Tyr – Pro – Thr – Ile – Trp

DNA sequence (gene) CAGTATCCTACGATTTCG

	U		C		A		G	
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly

1069

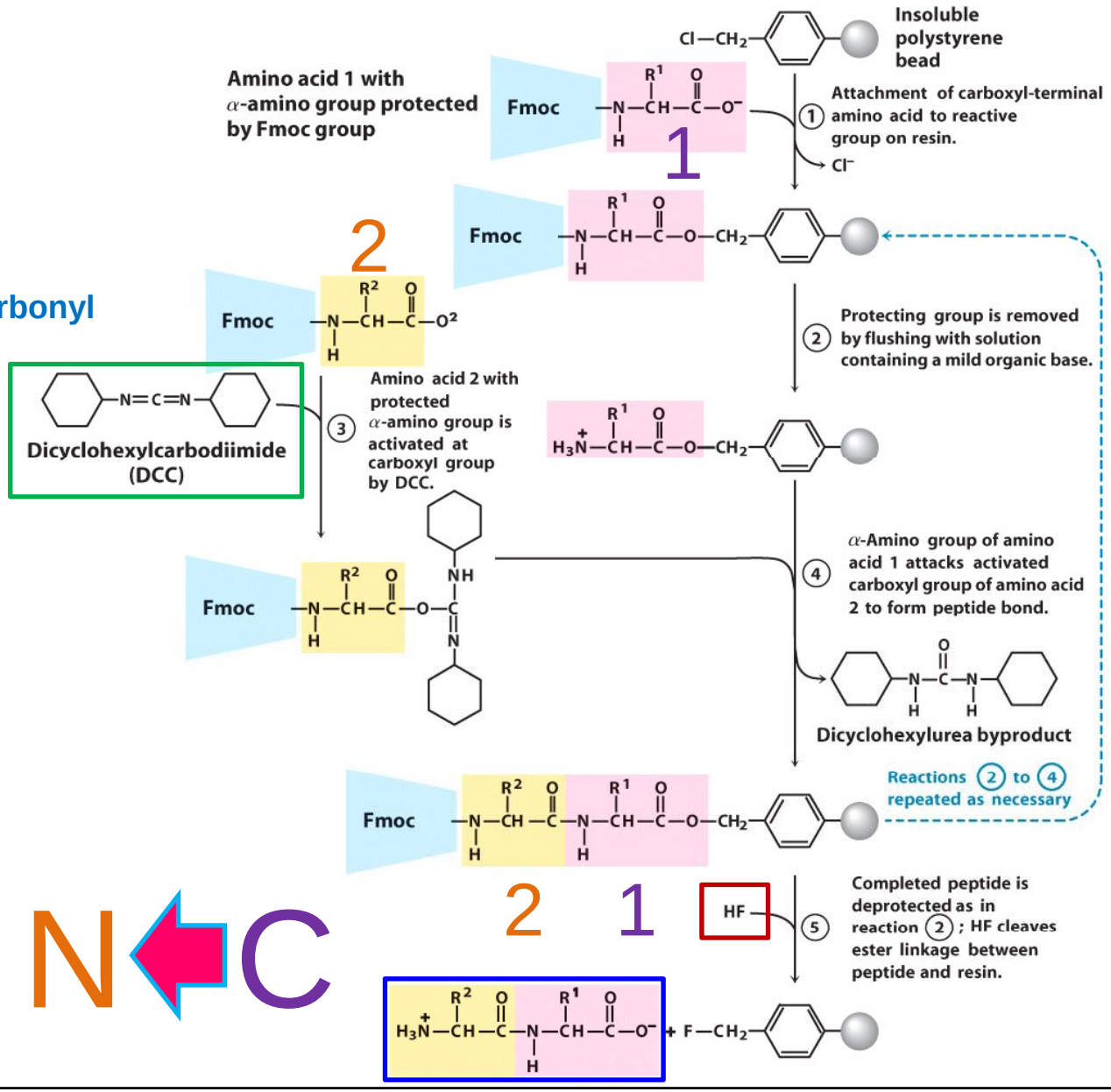
Chemical synthesis of a peptide on an insoluble polymer support



Fmoc

9-Fluorenylmethoxycarbonyl

R. Bruce Merrifield



Solid-Phase Peptide Synthesis

TABLE 3–8 Effect of Stepwise Yield on Overall Yield in Peptide Synthesis

Number of residues in the final polypeptide	*Overall yield of final peptide (%) when the yield of each step is:*	
	96.0%	*99.8%*
11	66	98
21	44	96
31	29	94
51	13	90
100	1.7	82