LEHNINGER PRINCIPLES OF BIOCHEMISTRY Sixth Edition

CHAPTER 3

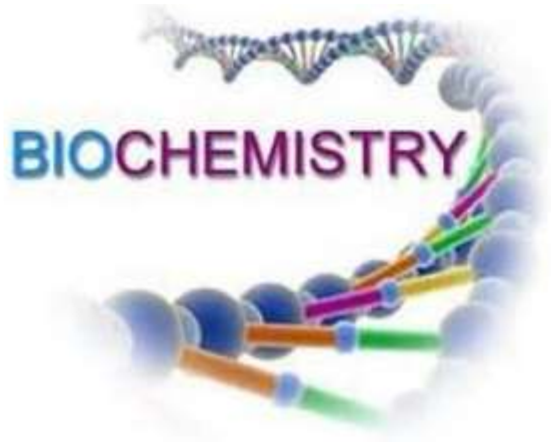
Amino Acids, Peptides, and Proteins

3.1 Amino Acids

3.2 Peptides and Proteins

3.3 Working with Proteins

3.4 The Structure of Proteins: Primary Structure



3.1 Amino Acids

Amino acids: building blocks of proteins

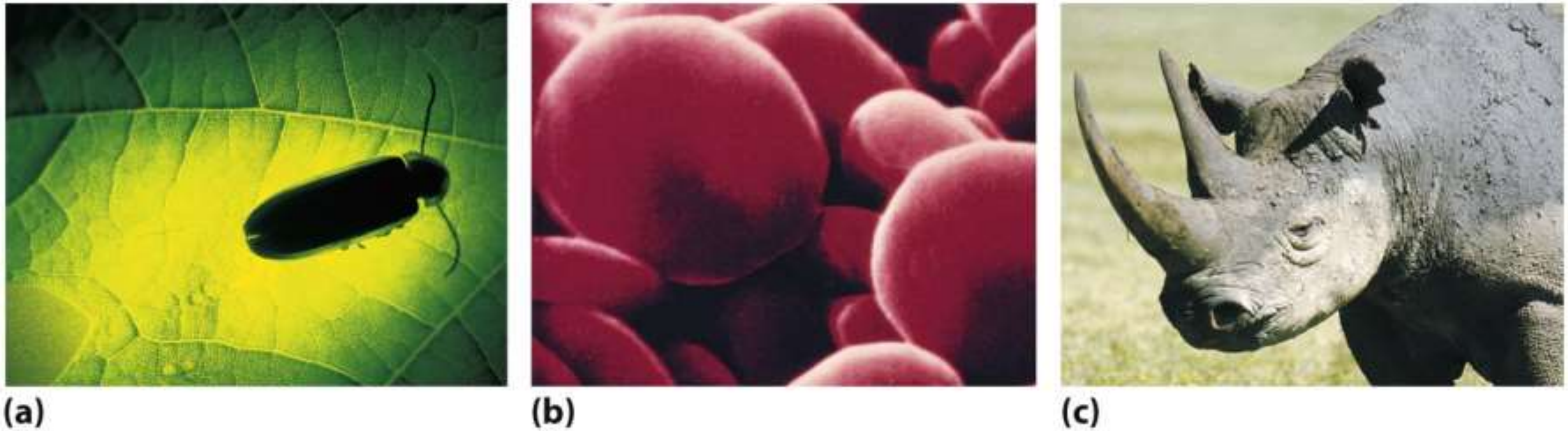
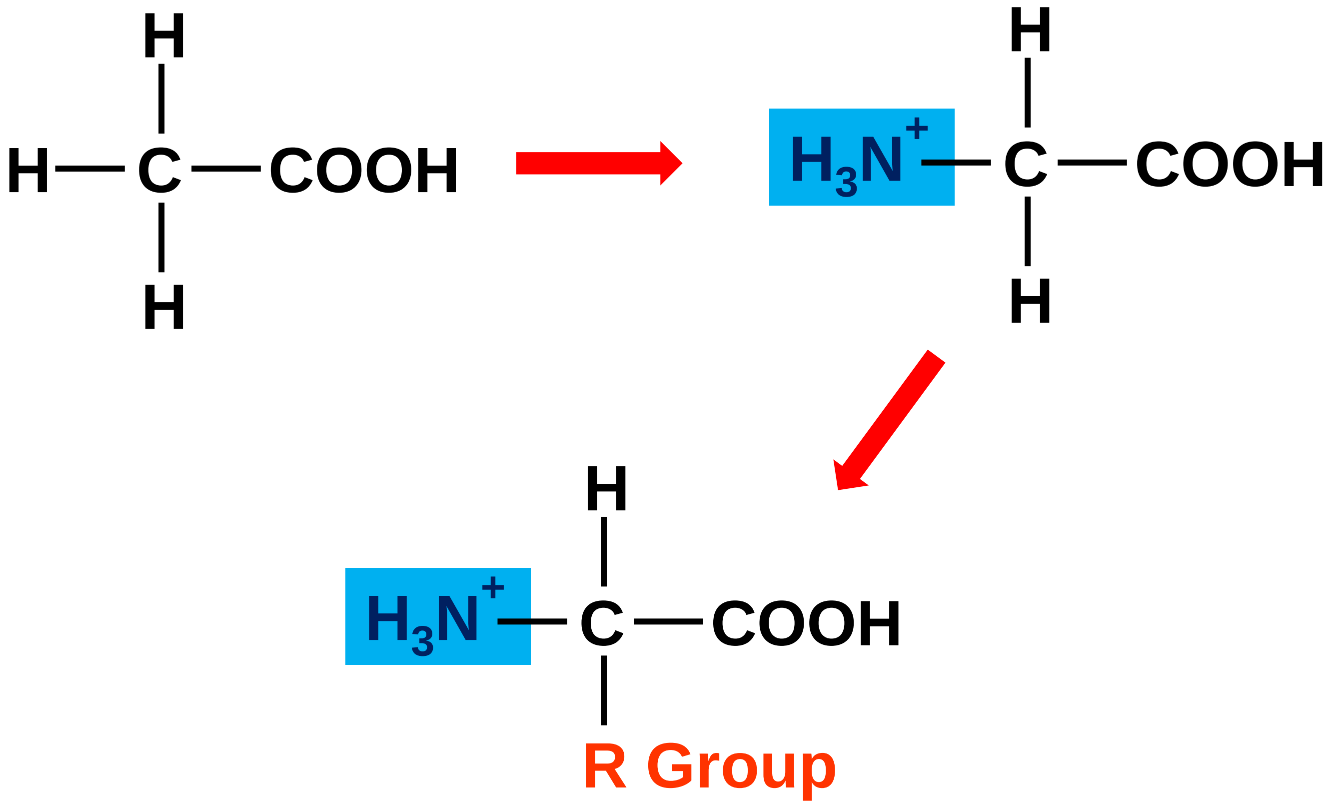

(a) (b) (c)

Figure 3-1
Lehninger Principles of Biochemistry, Fifth Edition

20 amino acids & millions of proteins

3.1 Amino Acids

- **General structure of amino acids**



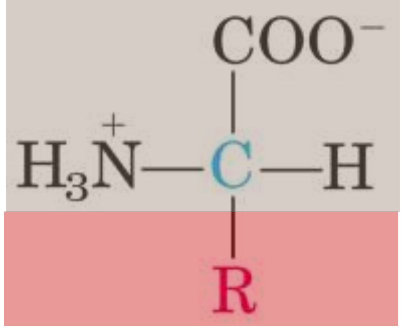
$$\mathrm{H}-\overset{\mathrm{H}}{\underset{\mathrm{H}}{\mathrm{C}}}-\mathrm{COOH} \longrightarrow \mathrm{H_3\overset{+}{N}}-\overset{\mathrm{H}}{\underset{\mathrm{H}}{\mathrm{C}}}-\mathrm{COOH}$$

$$\mathrm{H_3\overset{+}{N}}-\overset{\mathrm{H}}{\underset{\text{R Group}}{\mathrm{C}}}-\mathrm{COOH}$$

3.1 Amino Acids

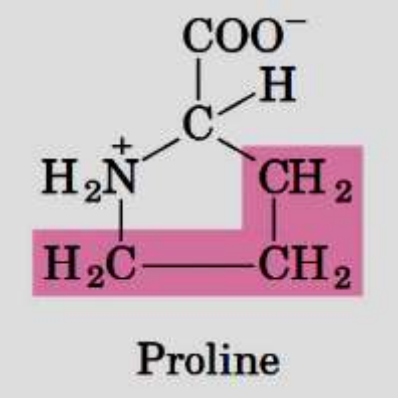
- **Amino acids share common structural features**

Any amino acid



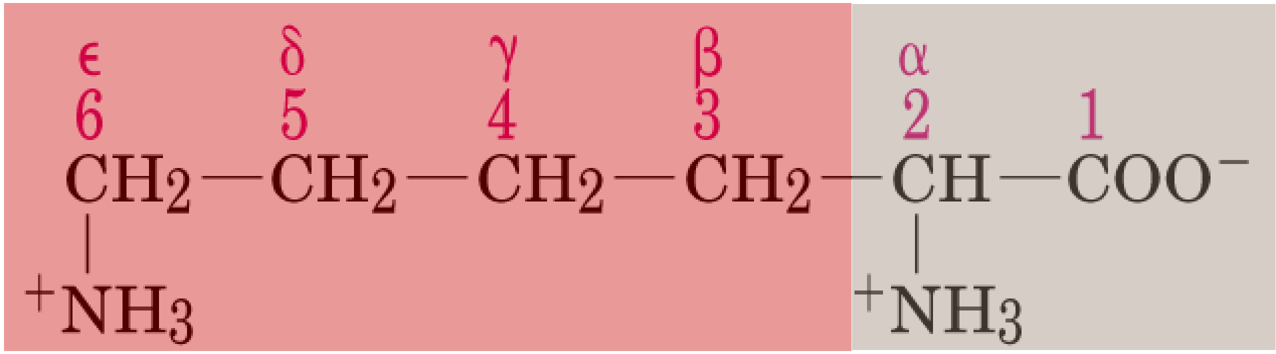
$$H_3\overset{+}{N}-\overset{COO^-}{\underset{R}{C}}-H$$

except



Proline

Lysine

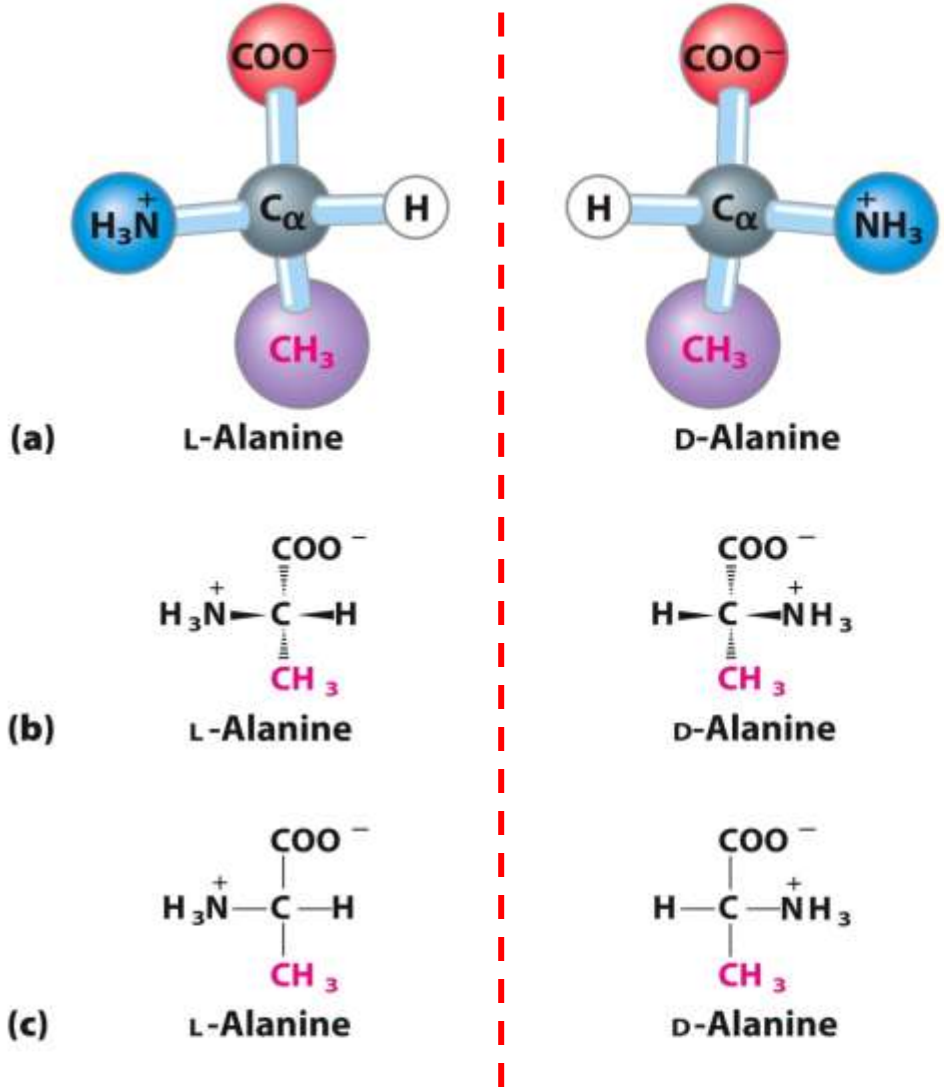


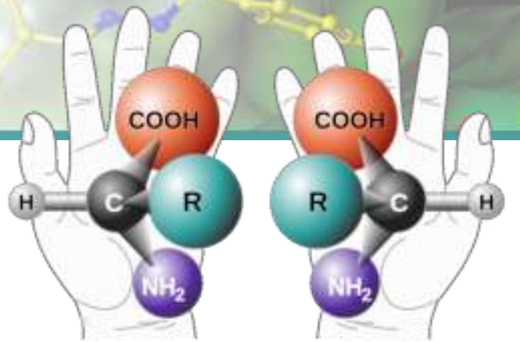
$$\underset{^+NH_3}{CH_2}-CH_2-CH_2-CH_2-\underset{^+NH_3}{CH}-COO^-$$

R Group

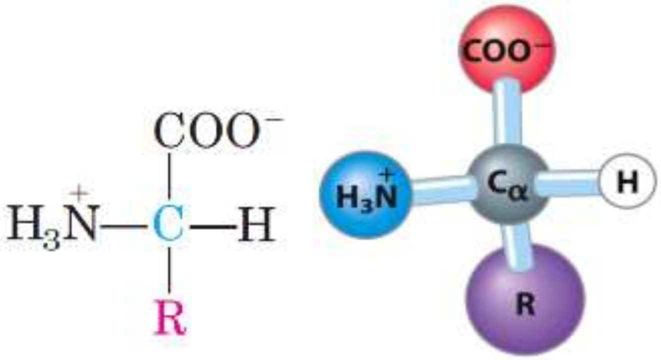
3.1 Amino Acids

- **Chirality of amino acids**





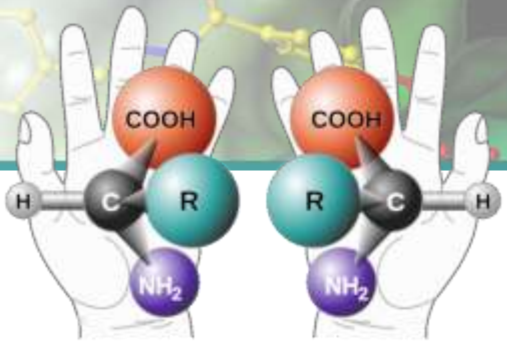
Chirality of amino acid is termed by the side of its **α-amino** group.

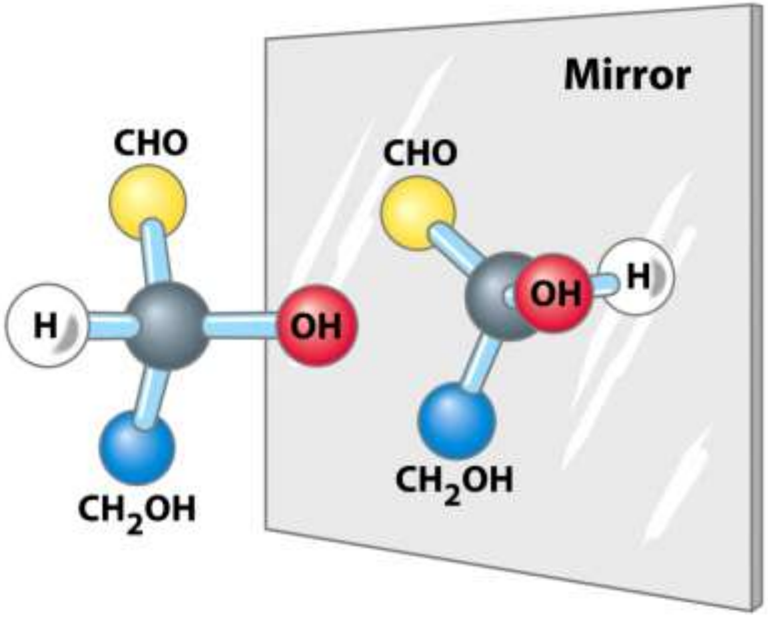


All amino acid residues in protein molecules are **L** stereoisomers.

3.1 Amino Acids

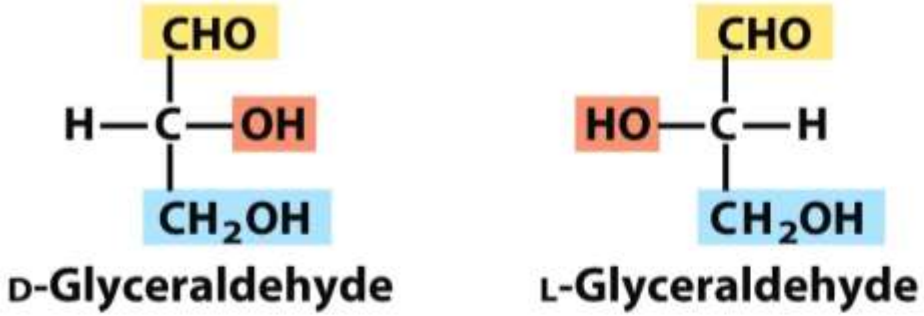
- **Chirality（手性）**



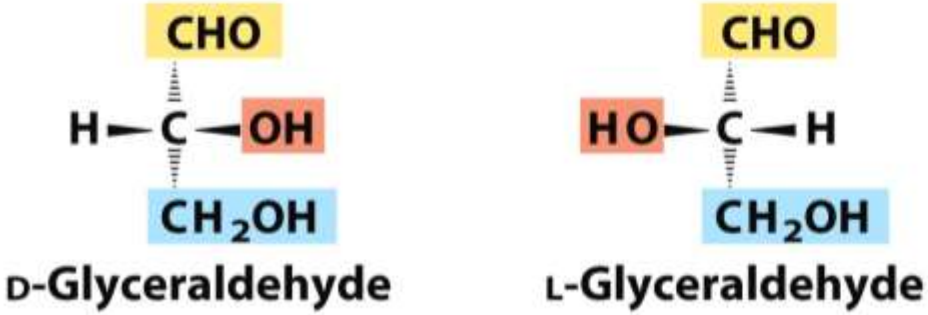


Ball-and-stick models

Figure 7-2 part 1
Lehninger Principles of Biochemistry, Sixth Edition



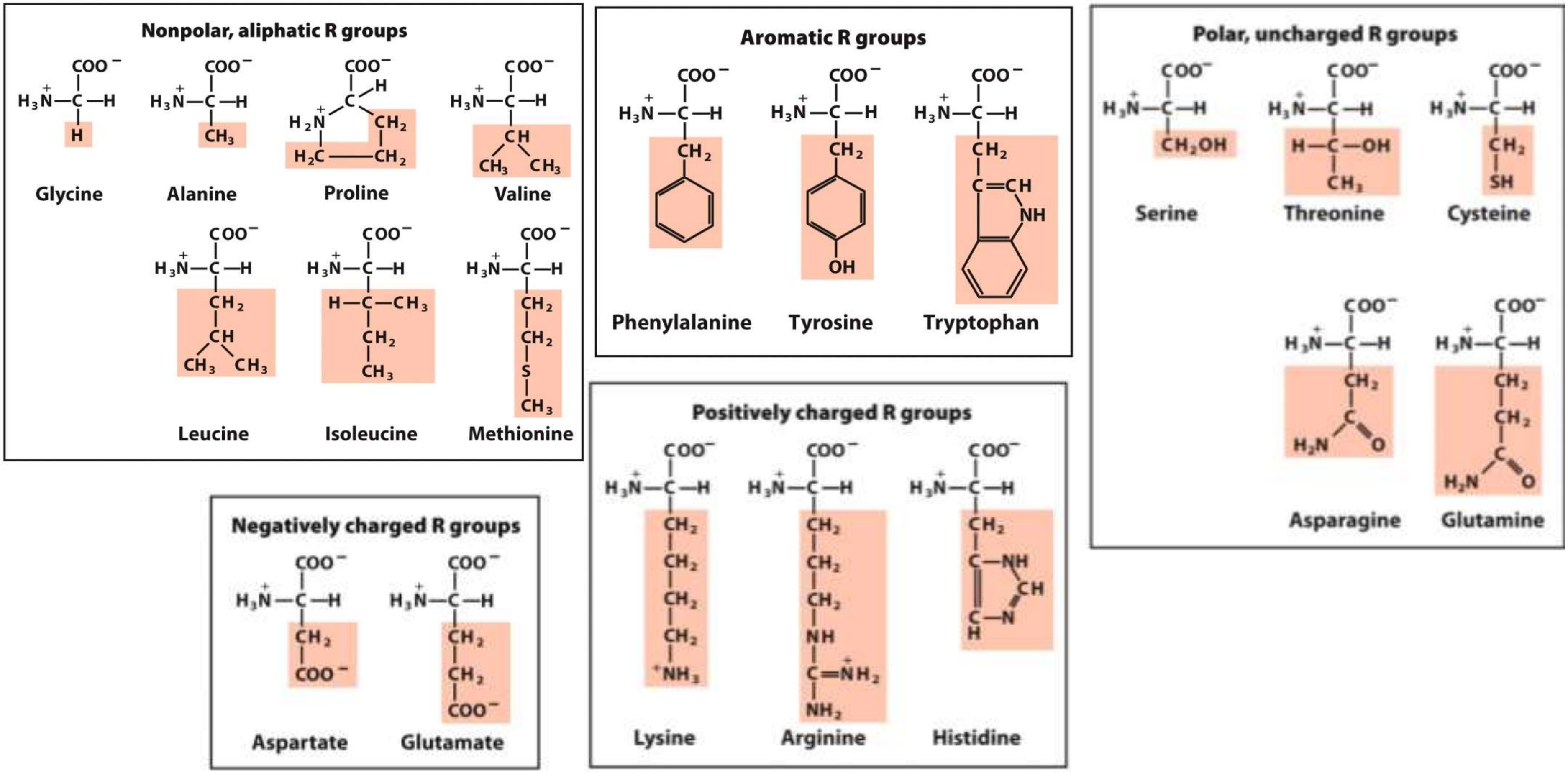
Fischer projection formulas



Perspective formulas

3.1 Amino Acids

- **The 20 common amino acids of proteins**



3.1 Amino Acids

I. Nonpolar aliphatic R groups amino acids:

Glycine (Gly, G, 甘氨酸)

Alanine (Ala，A, 丙氨酸)

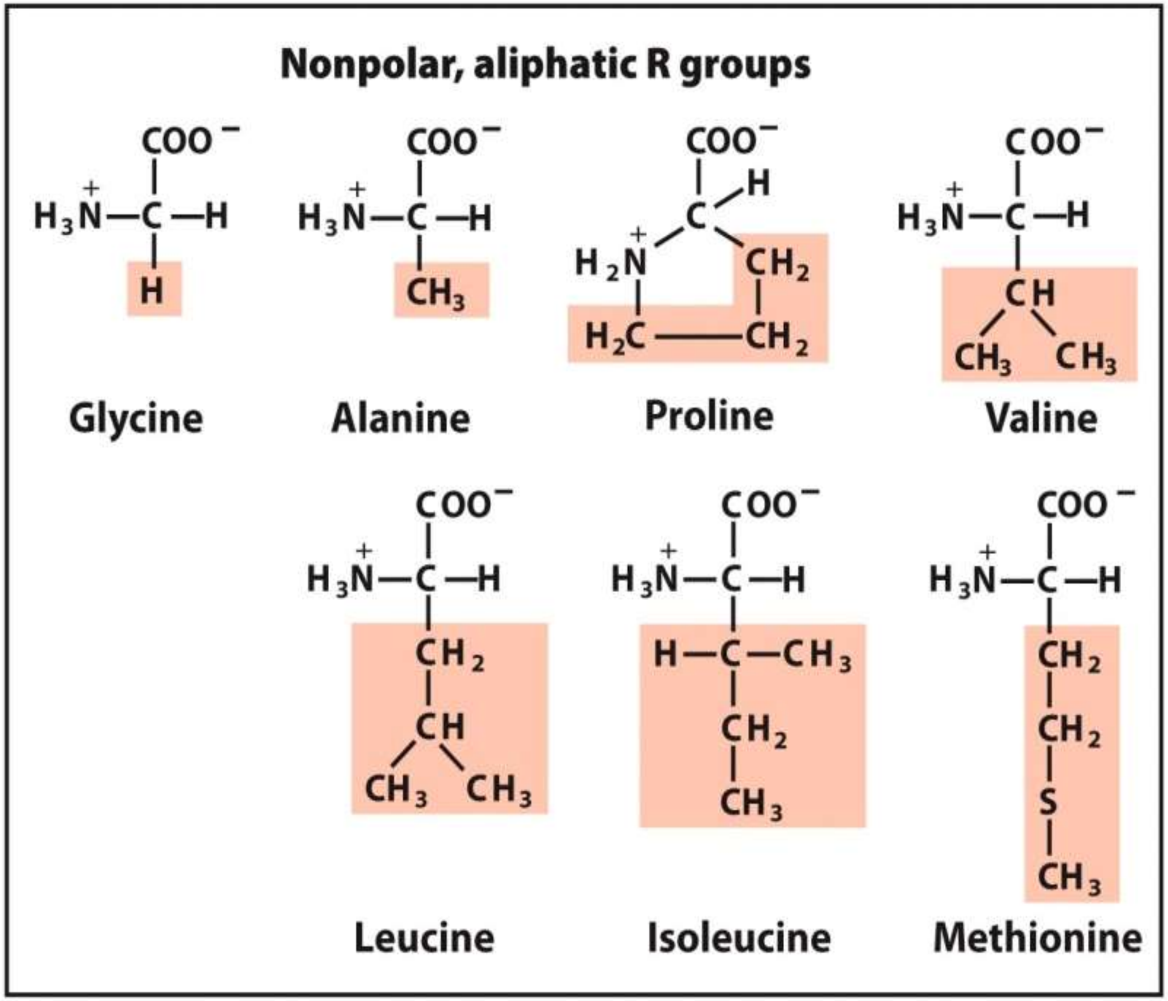
Proline (Pro, P, 脯氨酸)

Valine (Val， V, 缬氨酸)

Leucine (Leu, L, 亮氨酸)

Isoleucine (Ile, I, 异亮氨酸)

Methionine (Met, M,蛋氨酸,
甲硫氨酸)



3.1 Amino Acids

II. Polar, uncharged amino acids

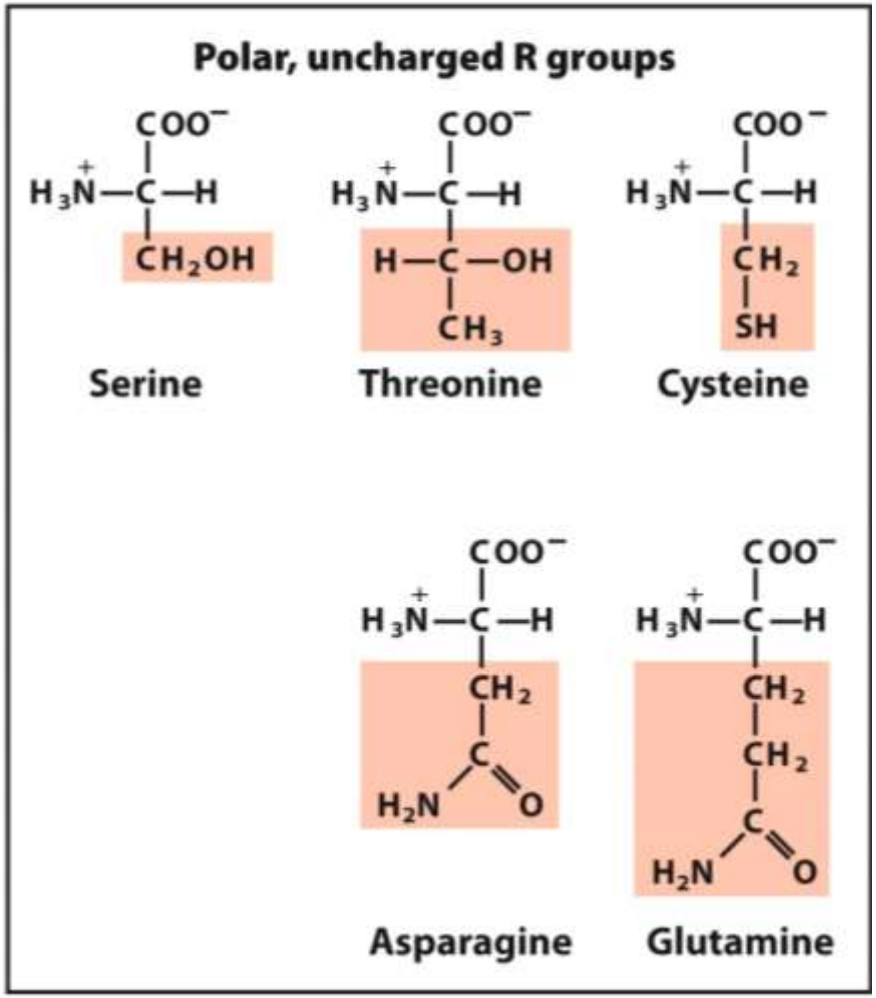
Serine (Ser, S, 丝氨酸)

Threonine (Thr, T, 苏氨酸)

Cysteine (Cys，C, 半胱氨酸)

Asparagine (Asn, N, 天冬酰胺)

Glutamine (Gln, Q, 谷氨酰胺)



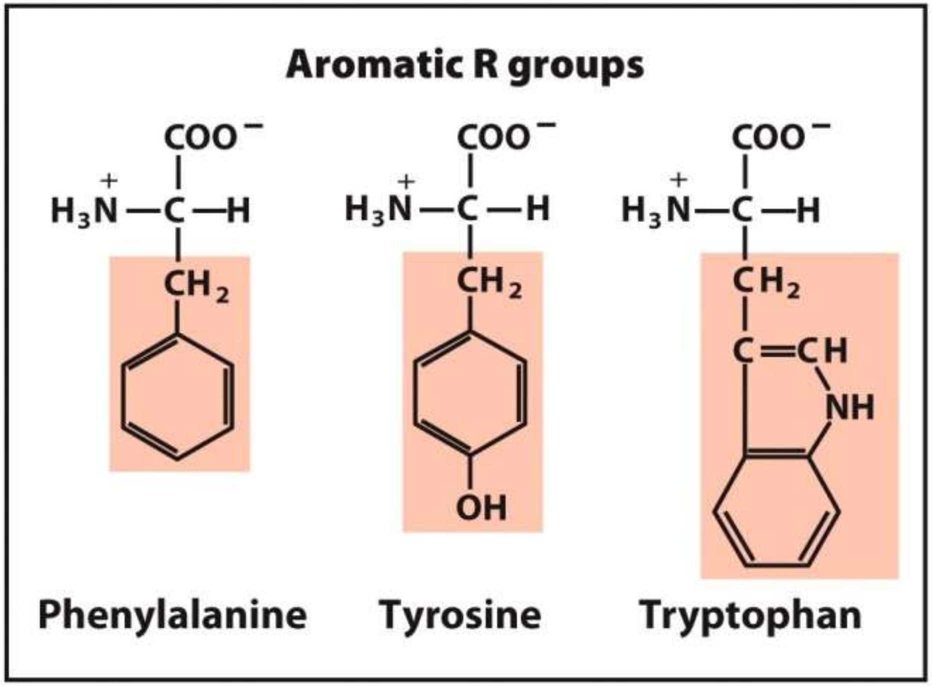
3.1 Amino Acids

III. Aromatic R groups

Phenylalanine (Phe, F, 苯丙氨酸)

Tyrosine (Tyr, Y, 酪氨酸)

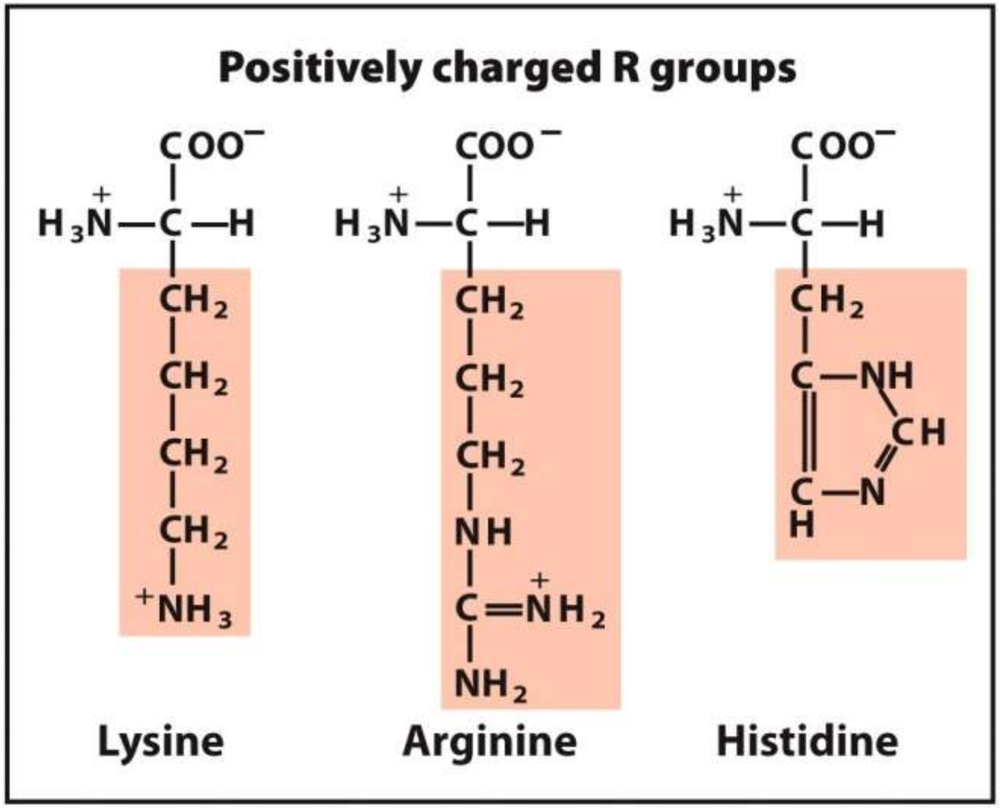
Tryptophan (Trp, W, 色氨酸)



3.1 Amino Acids

IV. Positively charged amino acids

--Basic amino acids



Histidine (His, H, 组氨酸)

Lysine (Lys, K, 赖氨酸)

Arginine (Arg, R, 精氨酸)

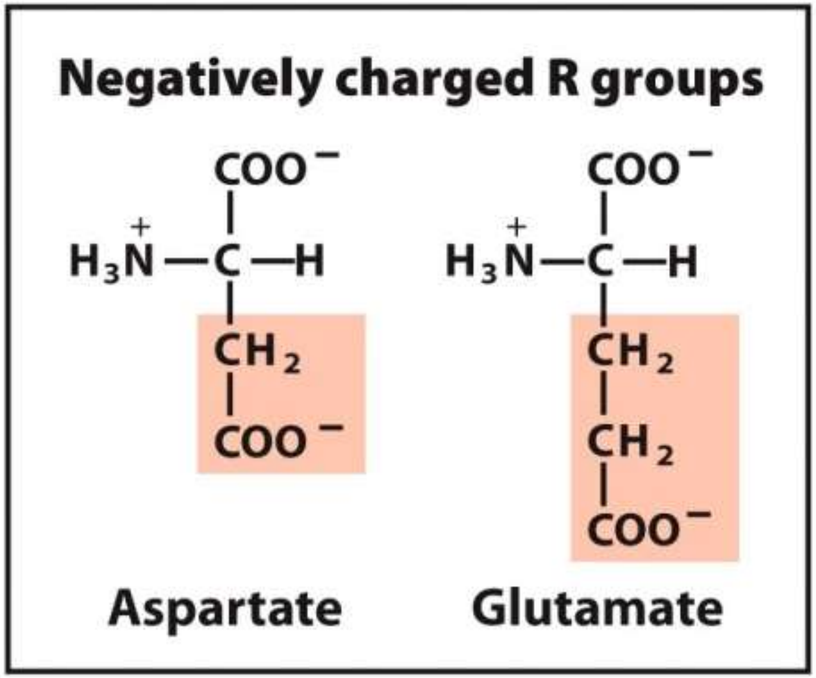
3.1 Amino Acids

V. Negatively charged amino acids

--Acidic amino acids

Aspartic acid (Asp, D, 天冬氨酸)

Glutamic acid (Glu, E, 谷氨酸)



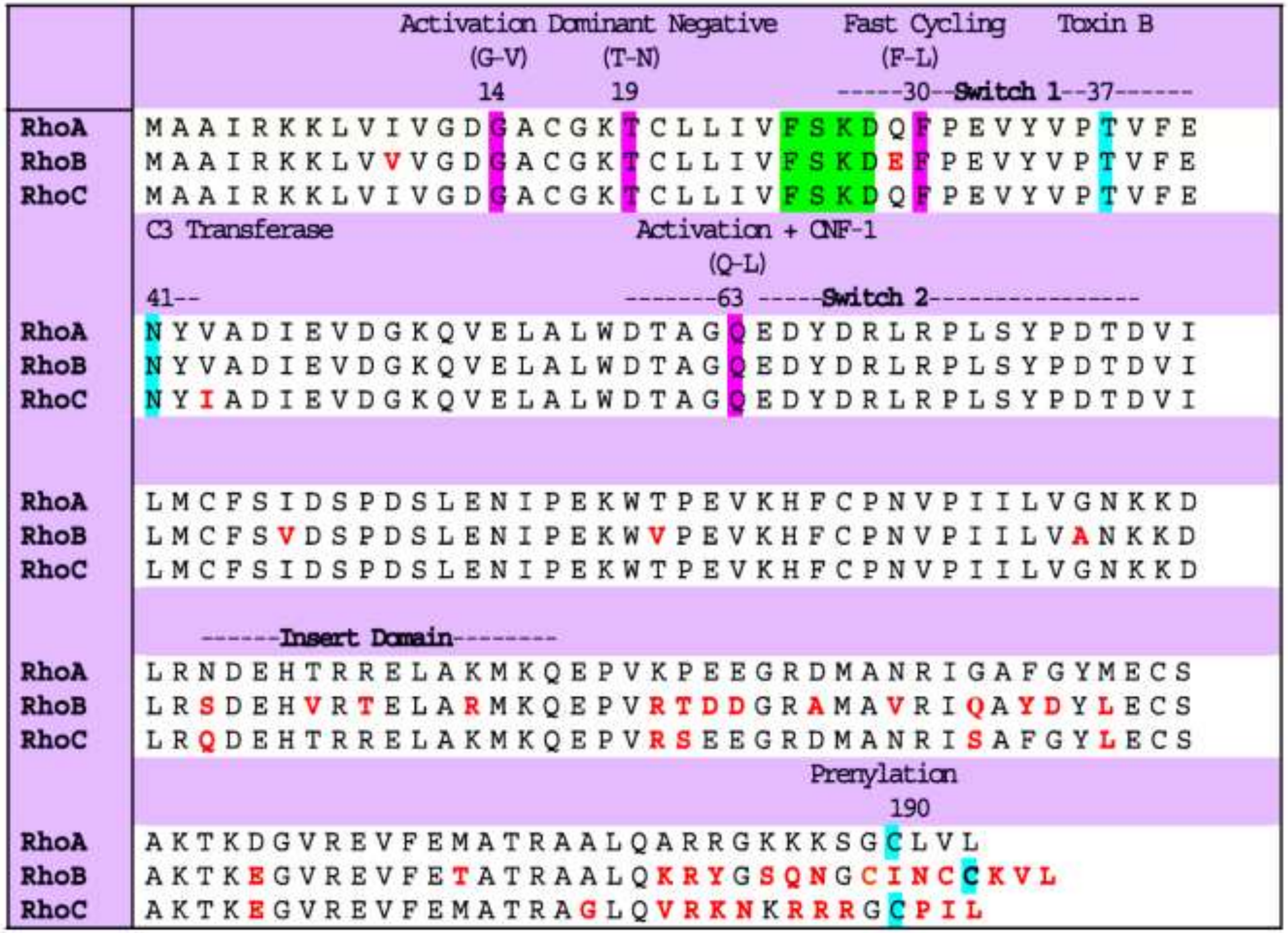
3.1 Amino Acids

- **Different ways to classify the 20 amino acids**

1. **Alkyl chain R groups: Gly、Ala、Val、Leu、Ile**
2. **Sulfur-containing amino acids: Met、Cys**
3. **Hydroxyl-containing amino acids: Ser、Thr 、Tyr**
4. **Aromatic amino acids: Phe、Tyr、Trp**
5. **Unusual cyclic amino acid: Pro、His**
6. **Amide groups : Asn、Gln**
7. **Acidic amino acids: Asp、Glu**
8. **Basic amino acids: Lys、Arg**

3.1 Amino Acids

- **Abbreviation symbols for amino acids**



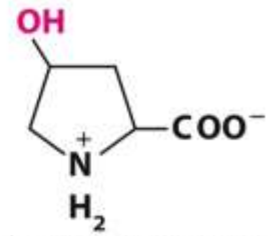
```
                          Activation Dominant Negative      Fast Cycling    Toxin B
                                (G-V)     (T-N)                (F-L)
                                  14        19              -----30--Switch 1--37------
RhoA    M A A I R K K L V I V G D G A C G K T C L L I V F S K D Q F P E V Y V P T V F E
RhoB    M A A I R K K L V V V G D G A C G K T C L L I V F S K D E F P E V Y V P T V F E
RhoC    M A A I R K K L V I V G D G A C G K T C L L I V F S K D Q F P E V Y V P T V F E
        C3 Transferase                      Activation + CNF-1
                                                  (Q-L)
        41--                               --------63 -----Switch 2----------------
RhoA    N Y V A D I E V D G K Q V E L A L W D T A G Q E D Y D R L R P L S Y P D T D V I
RhoB    N Y V A D I E V D G K Q V E L A L W D T A G Q E D Y D R L R P L S Y P D T D V I
RhoC    N Y I A D I E V D G K Q V E L A L W D T A G Q E D Y D R L R P L S Y P D T D V I

RhoA    L M C F S I D S P D S L E N I P E K W T P E V K H F C P N V P I I L V G N K K D
RhoB    L M C F S V D S P D S L E N I P E K W V P E V K H F C P N V P I I L V A N K K D
RhoC    L M C F S I D S P D S L E N I P E K W T P E V K H F C P N V P I I L V G N K K D

            ------Insert Domain--------
RhoA    L R N D E H T R R E L A K M K Q E P V K P E E G R D M A N R I G A F G Y M E C S
RhoB    L R S D E H V R T E L A R M K Q E P V R T D D G R A M A V R I Q A Y D Y L E C S
RhoC    L R Q D E H T R R E L A K M K Q E P V R S E E G R D M A N R I S A F G Y L E C S
                                                          Prenylation
                                                                190
RhoA    A K T K D G V R E V F E M A T R A A L Q A R R G K K K S G C L V L
RhoB    A K T K E G V R E V F E T A T R A A L Q K R Y G S Q N G C I N C C K V L
RhoC    A K T K E G V R E V F E M A T R A G L Q V R K N K R R R G C P I L
```

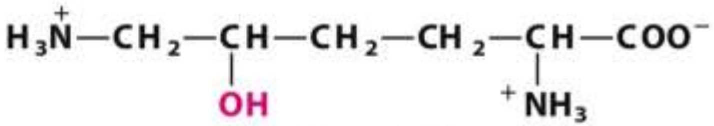
Margaret Oakley Dayhoff
1925–1983

The one-letter code reduces the data file sizes to describe amino acid sequences, which was important for the progression of bioinformatics field.

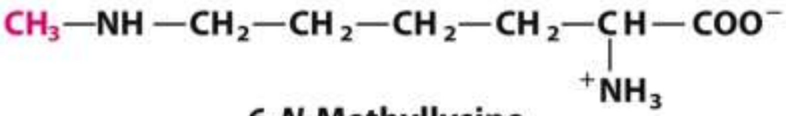
3.1 Amino Acids

- **Uncommon amino acids**

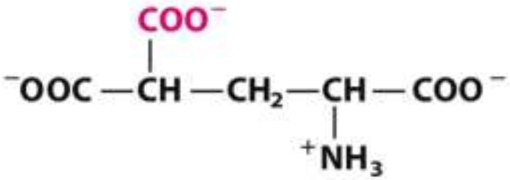

4-Hydroxyproline


$H_3\overset{+}{N}-CH_2-\underset{OH}{CH}-CH_2-CH_2-\underset{^{+}NH_3}{CH}-COO^-$

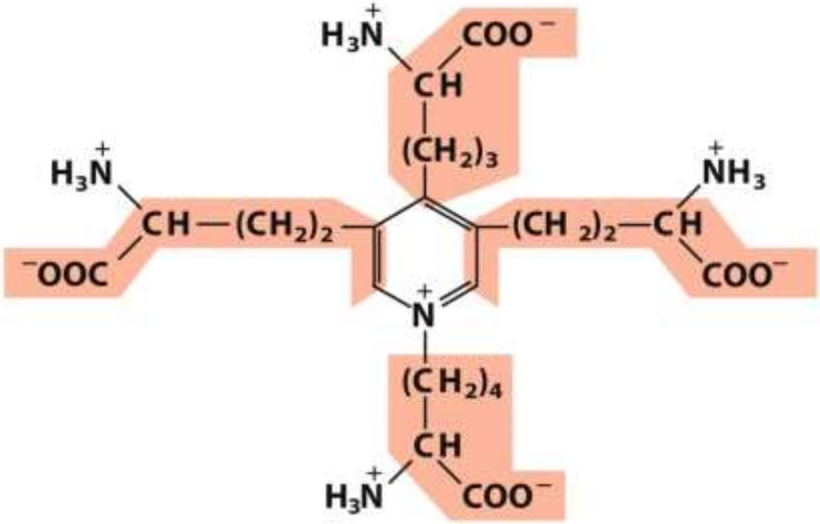
5-Hydroxylysine

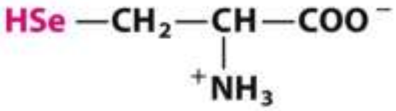

$CH_3-NH-CH_2-CH_2-CH_2-CH_2-\underset{^{+}NH_3}{CH}-COO^-$

6-*N*-Methyllysine


$^-OOC-\overset{COO^-}{CH}-CH_2-\underset{^{+}NH_3}{CH}-COO^-$

γ-Carboxyglutamate

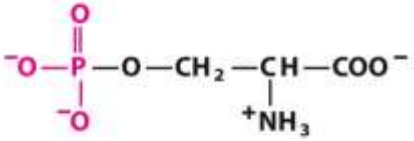

Desmosine


$HSe-CH_2-\underset{^{+}NH_3}{CH}-COO^-$

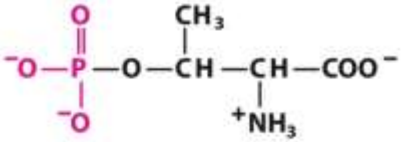
Selenocysteine

3.1 Amino Acids

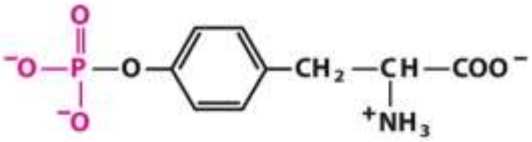
- **Amino acids with reversible modifications**



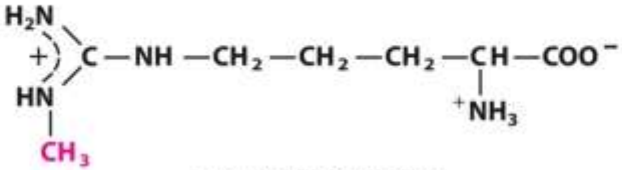
Phosphoserine



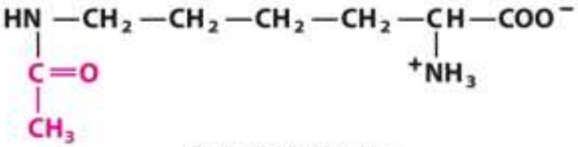
Phosphothreonine



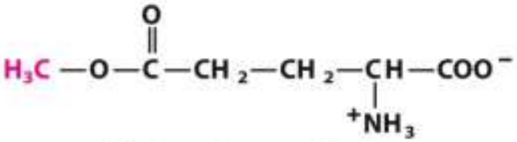
Phosphotyrosine



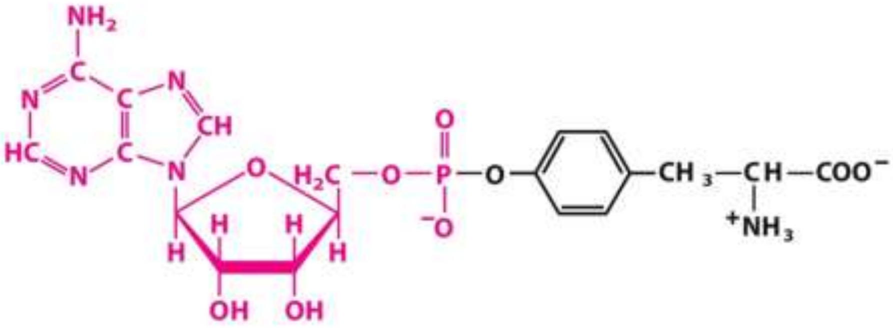
σ-*N*-Methylarginine



6-*N*-Acetyllysine



Glutamate γ-methyl ester



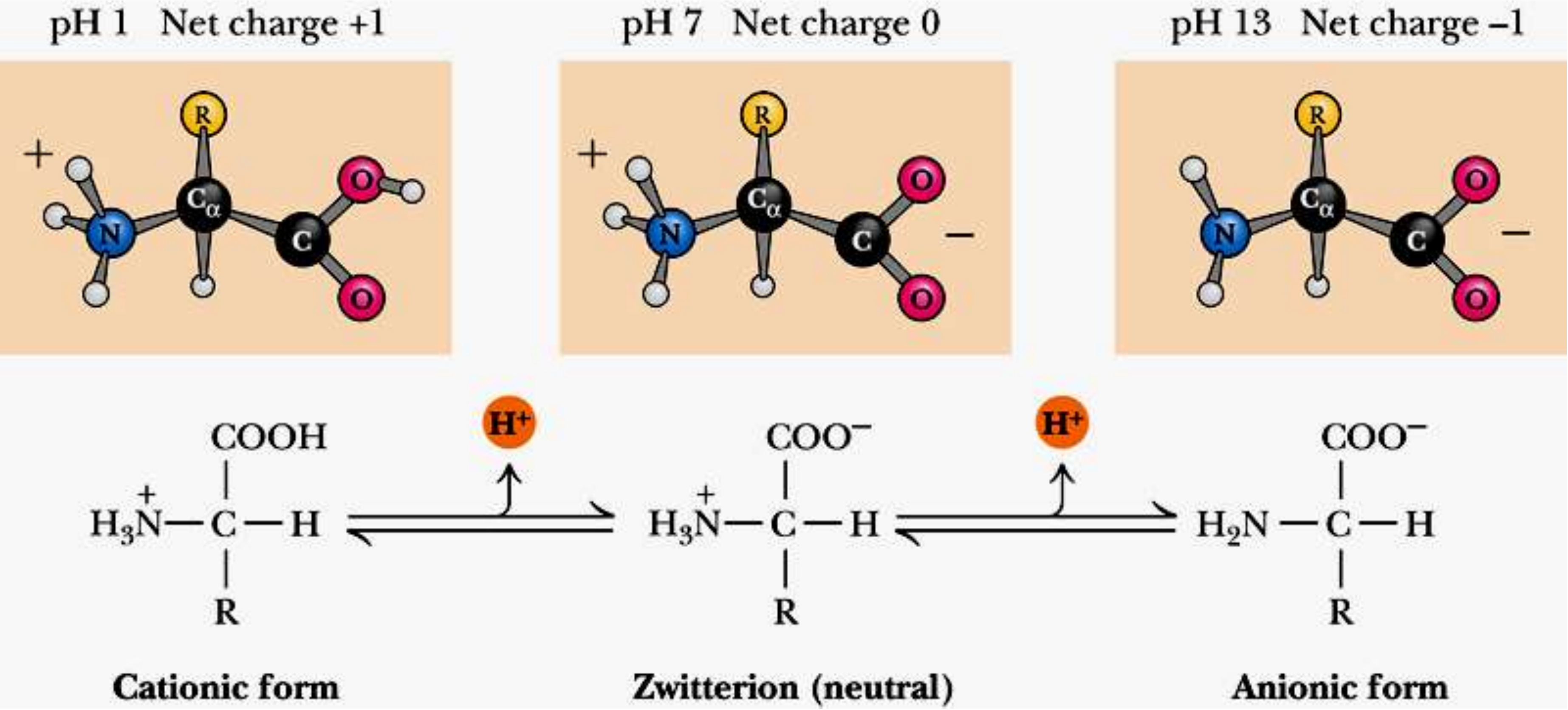
Adenylyltyrosine

3.1 Amino Acids

■ Properties of the 20 common amino acids

Acid-base properties of amino acids

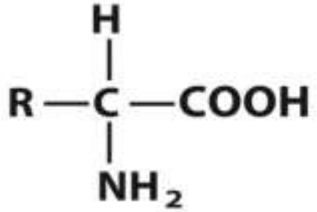
– Zwitterion of amino acids



3.1 Amino Acids

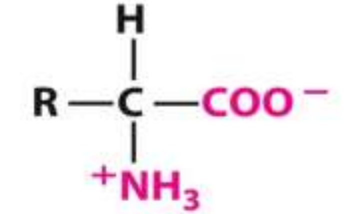
- **Nonionic and Zwitterionic forms of amino acids**

The nonionic form does not occur in significant amounts in aqueous solutions. The zwitterion predominates at neutral pH. A zwitterion can act as either an acid (proton donor) or a base (proton acceptor).



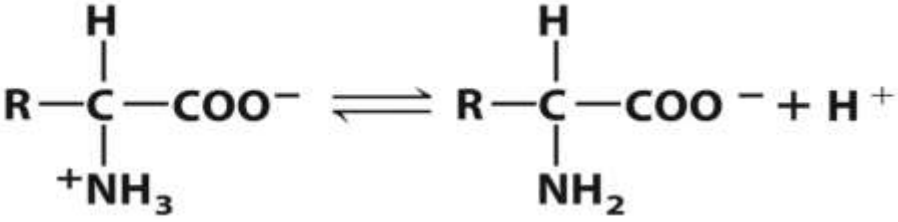
$R-\overset{H}{\underset{NH_2}{C}}-COOH$

Nonionic form



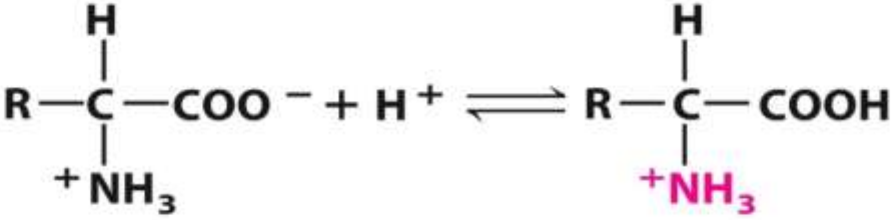
$R-\overset{H}{\underset{^{+}NH_3}{C}}-COO^{-}$

Zwitterionic form



$R-\overset{H}{\underset{^{+}NH_3}{C}}-COO^{-} \rightleftharpoons R-\overset{H}{\underset{NH_2}{C}}-COO^{-} + H^{+}$

Zwitterion as acid



$R-\overset{H}{\underset{^{+}NH_3}{C}}-COO^{-} + H^{+} \rightleftharpoons R-\overset{H}{\underset{^{+}NH_3}{C}}-COOH$

Zwitterion as base

Zwitter (德语) = hybrid

3.1 Amino Acids

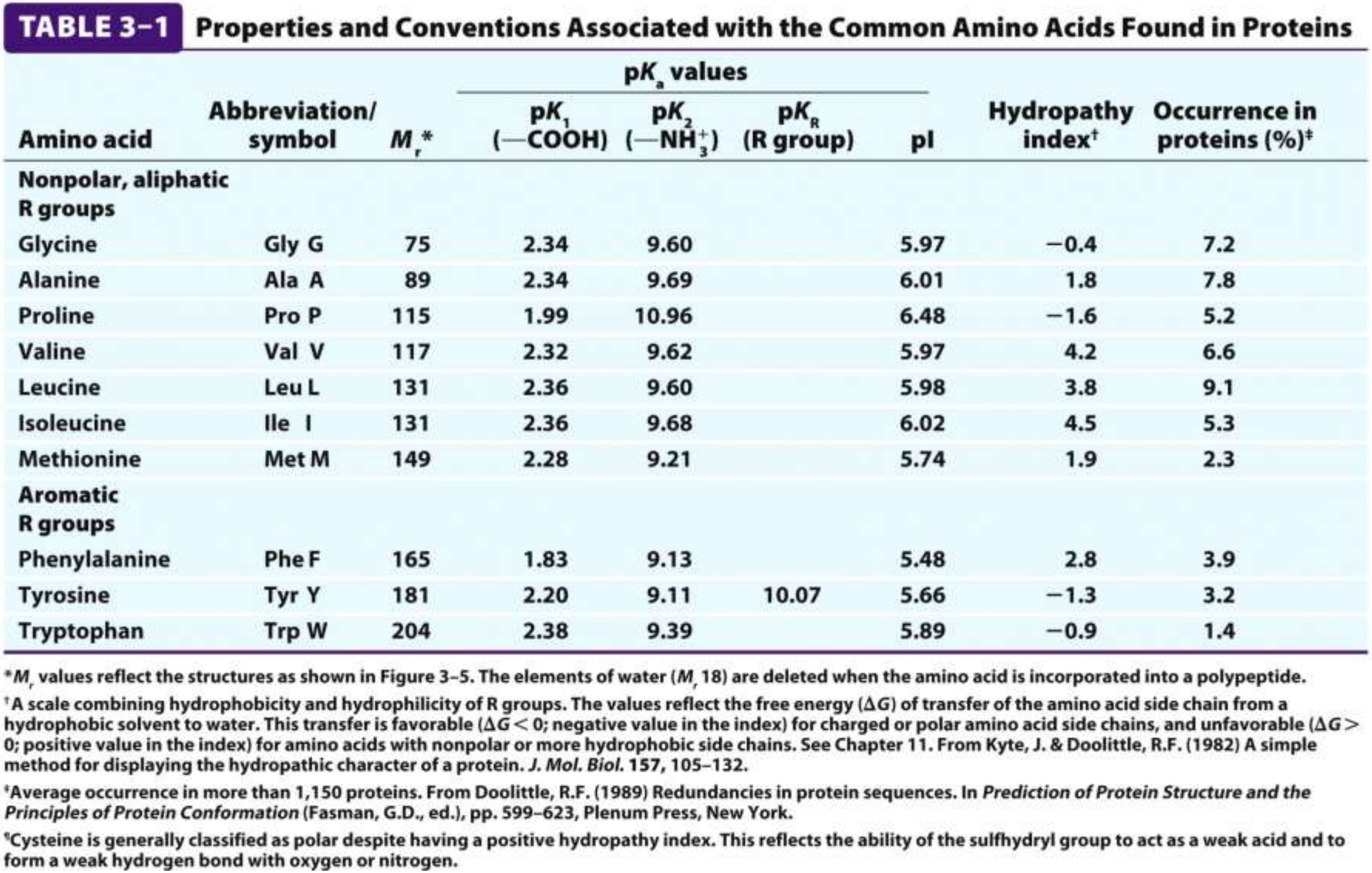


TABLE 3–1 **Properties and Conventions Associated with the Common Amino Acids Found in Proteins**

			pK_a values					
Amino acid	**Abbreviation/ symbol**	M_r*	pK_1 (—COOH)	pK_2 ($—NH_3^+$)	pK_R (R group)	**pI**	**Hydropathy index**†	**Occurrence in proteins (%)**‡
Nonpolar, aliphatic R groups								
Glycine	Gly G	75	2.34	9.60		5.97	−0.4	7.2
Alanine	Ala A	89	2.34	9.69		6.01	1.8	7.8
Proline	Pro P	115	1.99	10.96		6.48	−1.6	5.2
Valine	Val V	117	2.32	9.62		5.97	4.2	6.6
Leucine	Leu L	131	2.36	9.60		5.98	3.8	9.1
Isoleucine	Ile I	131	2.36	9.68		6.02	4.5	5.3
Methionine	Met M	149	2.28	9.21		5.74	1.9	2.3
Aromatic R groups								
Phenylalanine	Phe F	165	1.83	9.13		5.48	2.8	3.9
Tyrosine	Tyr Y	181	2.20	9.11	10.07	5.66	−1.3	3.2
Tryptophan	Trp W	204	2.38	9.39		5.89	−0.9	1.4

*M_r values reflect the structures as shown in Figure 3–5. The elements of water (M_r 18) are deleted when the amino acid is incorporated into a polypeptide.

†A scale combining hydrophobicity and hydrophilicity of R groups. The values reflect the free energy (ΔG) of transfer of the amino acid side chain from a hydrophobic solvent to water. This transfer is favorable ($\Delta G < 0$; negative value in the index) for charged or polar amino acid side chains, and unfavorable ($\Delta G >$ 0; positive value in the index) for amino acids with nonpolar or more hydrophobic side chains. See Chapter 11. From Kyte, J. & Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105–132.

‡Average occurrence in more than 1,150 proteins. From Doolittle, R.F. (1989) Redundancies in protein sequences. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G.D., ed.), pp. 599–623, Plenum Press, New York.

§Cysteine is generally classified as polar despite having a positive hydropathy index. This reflects the ability of the sulfhydryl group to act as a weak acid and to form a weak hydrogen bond with oxygen or nitrogen.

3.1 Amino Acids

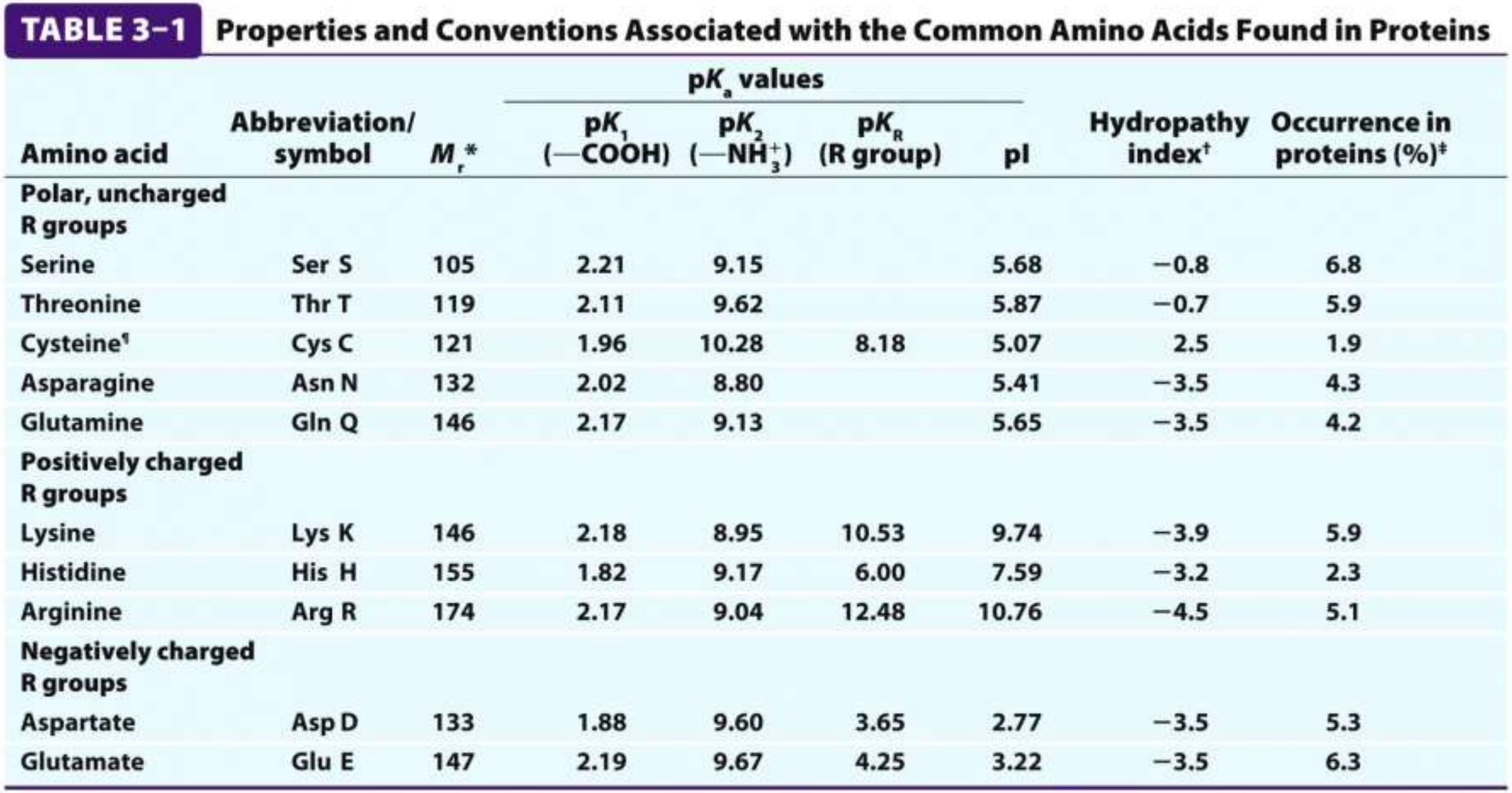


TABLE 3–1 Properties and Conventions Associated with the Common Amino Acids Found in Proteins

Amino acid	Abbreviation/ symbol	M_r*	pK_a values			pI	Hydropathy index†	Occurrence in proteins (%)‡
			pK_1 (—COOH)	pK_2 (—NH_3^+)	pK_R (R group)			
Polar, uncharged R groups								
Serine	Ser S	105	2.21	9.15		5.68	−0.8	6.8
Threonine	Thr T	119	2.11	9.62		5.87	−0.7	5.9
Cysteine¶	Cys C	121	1.96	10.28	8.18	5.07	2.5	1.9
Asparagine	Asn N	132	2.02	8.80		5.41	−3.5	4.3
Glutamine	Gln Q	146	2.17	9.13		5.65	−3.5	4.2
Positively charged R groups								
Lysine	Lys K	146	2.18	8.95	10.53	9.74	−3.9	5.9
Histidine	His H	155	1.82	9.17	6.00	7.59	−3.2	2.3
Arginine	Arg R	174	2.17	9.04	12.48	10.76	−4.5	5.1
Negatively charged R groups								
Aspartate	Asp D	133	1.88	9.60	3.65	2.77	−3.5	5.3
Glutamate	Glu E	147	2.19	9.67	4.25	3.22	−3.5	6.3

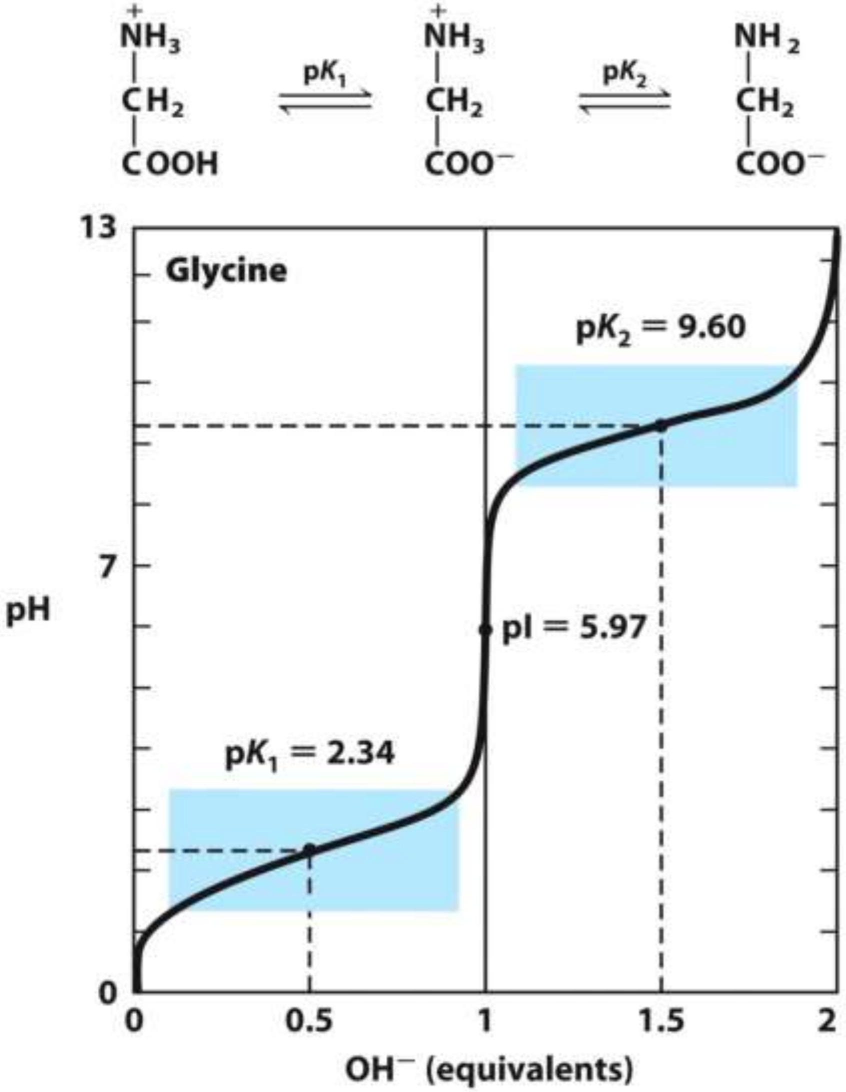
3.1 Amino Acids

- **Ionization and isoelectric point of amino acids**

Titration of Glycine

$$K_1 = \frac{[Gly^o][H^+]}{[Gly^+]}$$

$$K_2 = \frac{[Gly^-][H^+]}{[Gly^o]}$$



3.1 Amino Acids

$$K_1 = \frac{[Gly^o][H^+]}{[Gly^+]} \qquad K_2 = \frac{[Gly^-][H^+]}{[Gly^o]}$$

$$K_1 \bullet K_2 = [H^+]^2 \frac{[Gly^-]}{[Gly^+]}$$

At isoelectric point $[Gly^-]=[Gly^+]$

$$K_1 \cdot K_2 = [H^+]^2$$

$$[H^+] = (K_1 \cdot K_2)^{1/2}$$

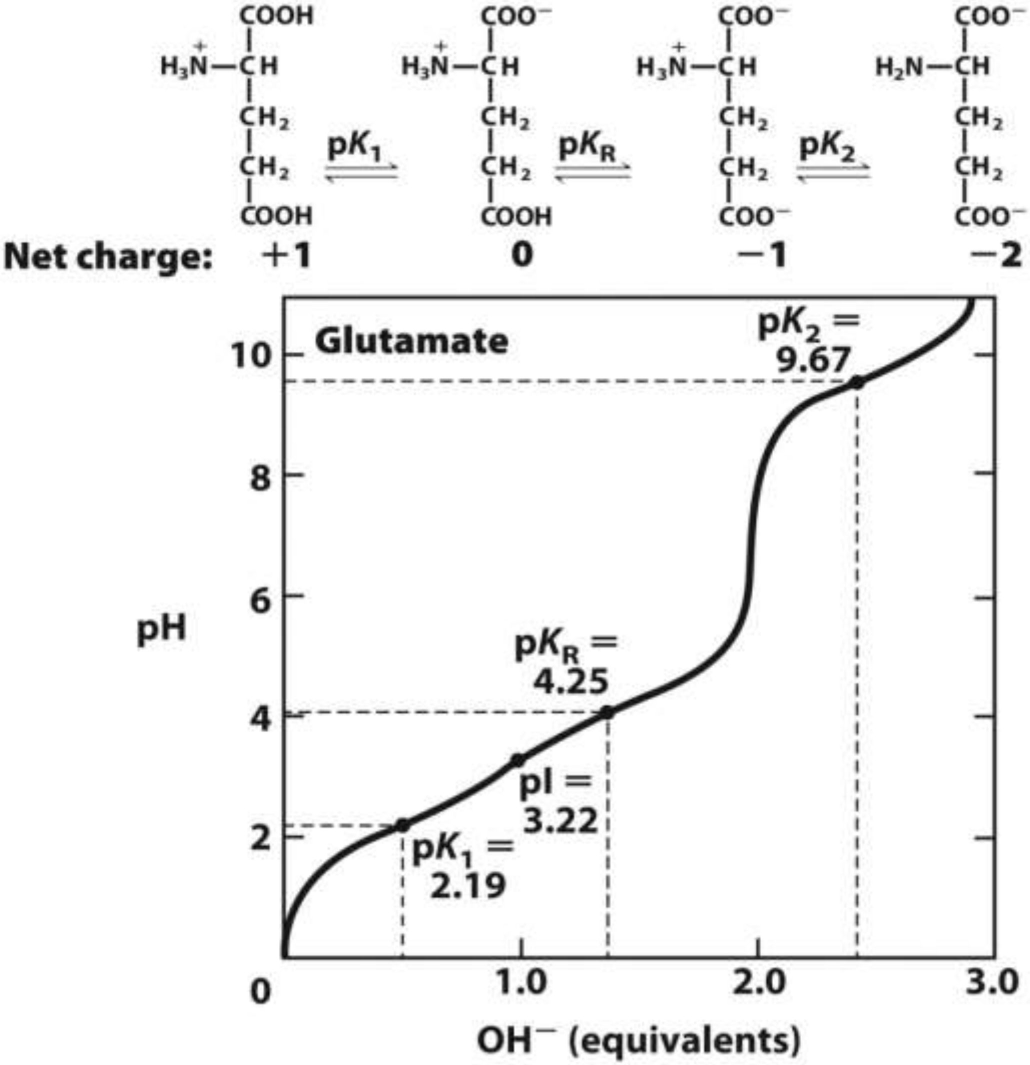
$$I = (K_1 \cdot K_2)^{1/2}$$

$$pI = (p\,K_1 + p\,K_2)/2$$

$$pI = (2.34 + 9.60)/2 = 5.97$$

3.1 Amino Acids

- **Ionization and isoelectric point of amino acids with dissociable groups on side chains**



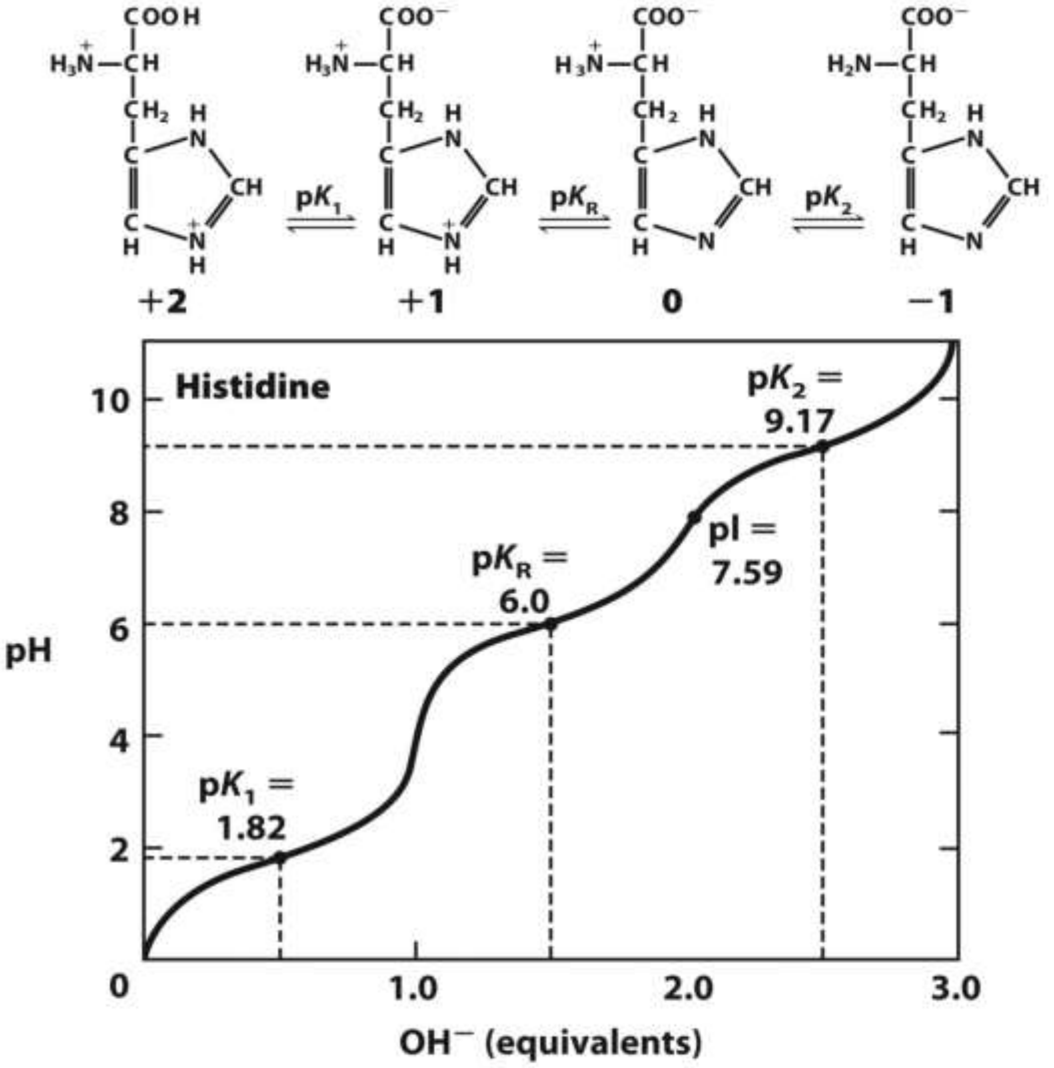
$$\mathrm{p}I = (\mathrm{p}K_1 + \mathrm{p}K_R)/2$$

$$= (2.19+4.25)/2$$

$$= 3.22$$

3.1 Amino Acids

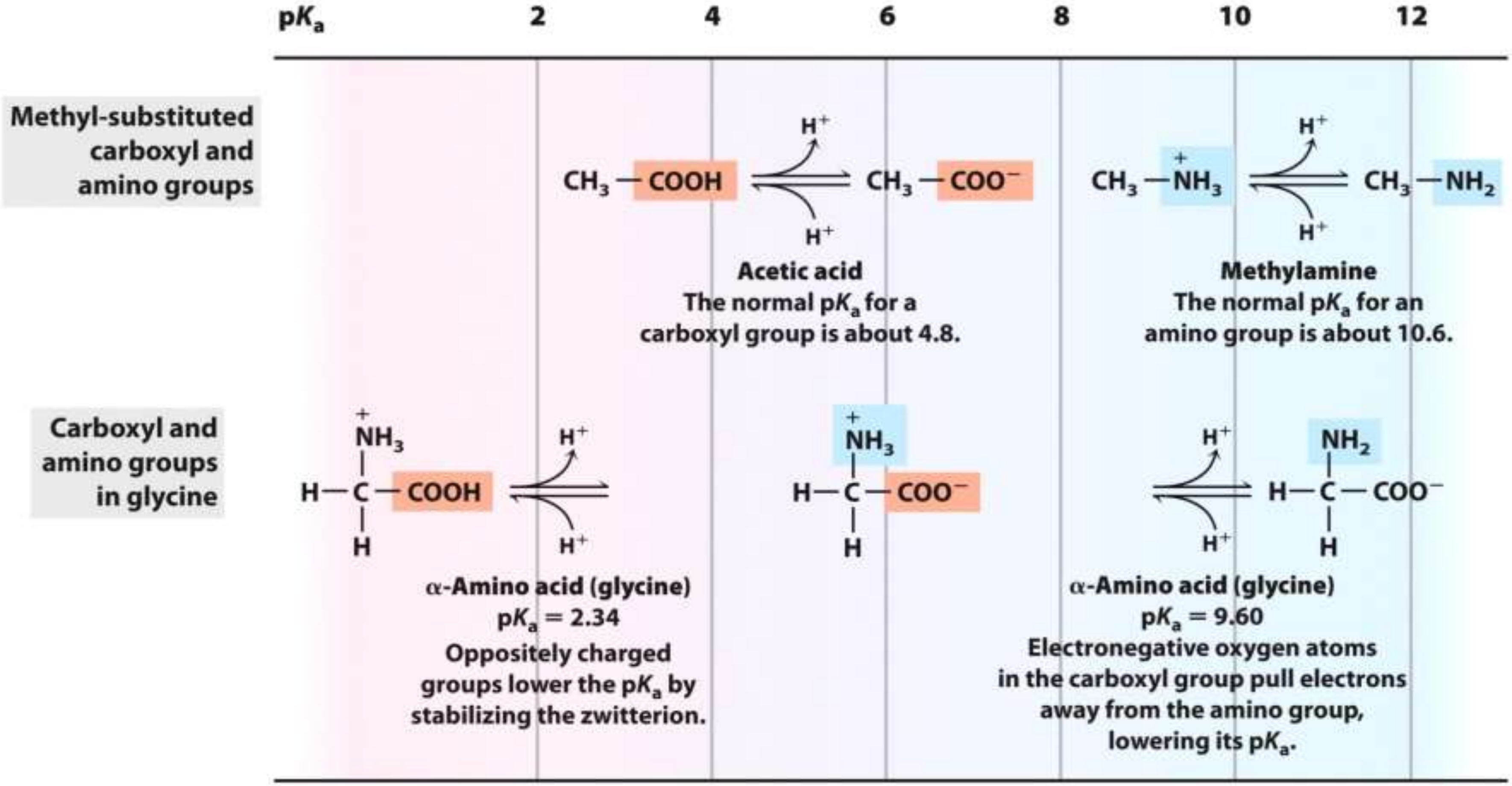
- **Ionization and isoelectric point of amino acids with dissociable groups on side chains**



$$pI = (pK_2 + pK_R)/2$$
$$= (9.17+6.0)/2$$
$$= 7.59$$

3.1 Amino Acids

- **Effect of the chemical environment on pK_a**



Acetic acid
The normal pK_a for a carboxyl group is about 4.8.

Methylamine
The normal pK_a for an amino group is about 10.6.

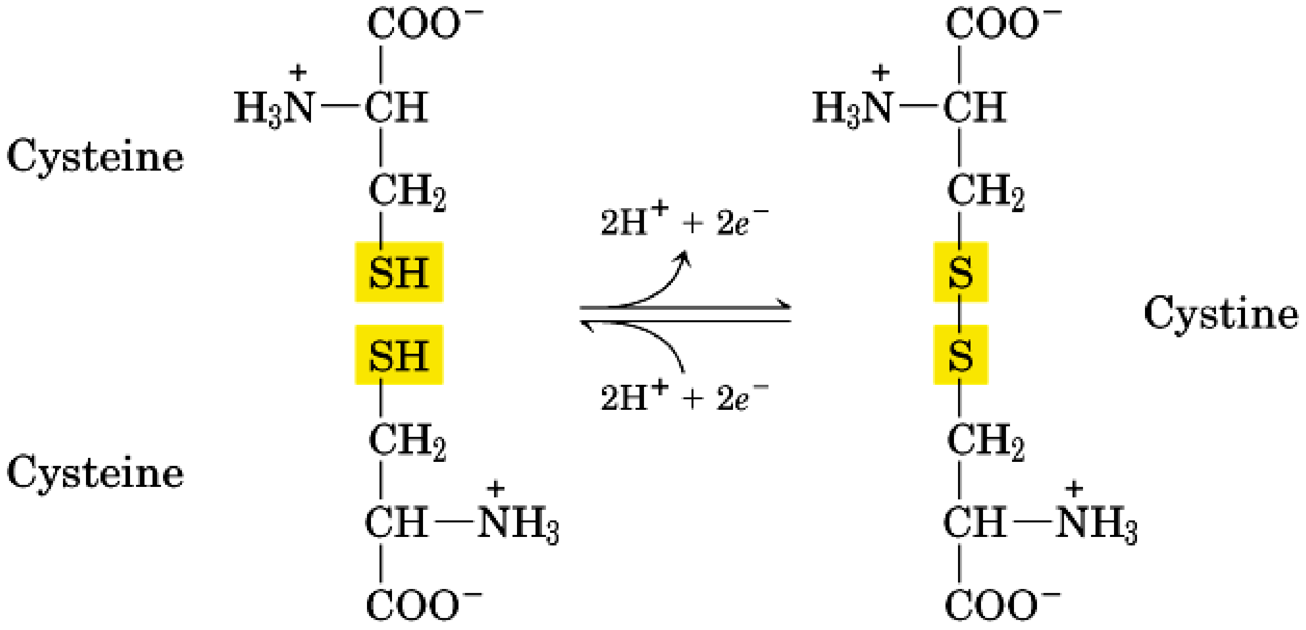
α-Amino acid (glycine)
pK_a = 2.34
Oppositely charged groups lower the pK_a by stabilizing the zwitterion.

α-Amino acid (glycine)
pK_a = 9.60
Electronegative oxygen atoms in the carboxyl group pull electrons away from the amino group, lowering its pK_a.

3.1 Amino Acids

- **Specific reactions of amino acid side chains**

Cysteine



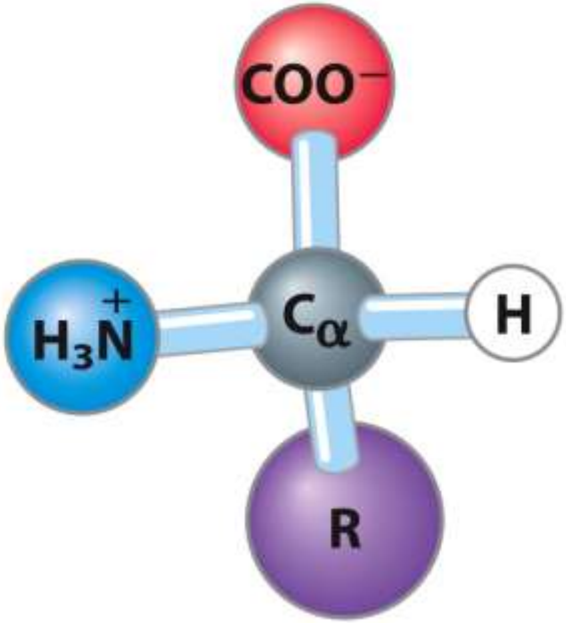
巯基乙醇 (mercaptoethanol, ME)

二硫苏糖醇 (dithiothrcitol, DTT)

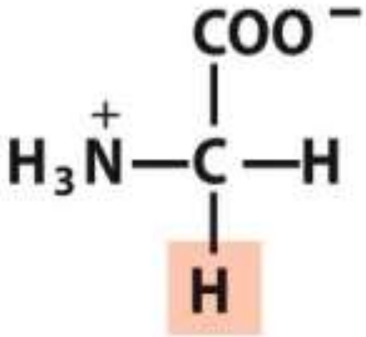
3.1 Amino Acids

- **Optical properties of amino acids**

Optical activity: α-carbon of amino acid is asymmetric or chiral



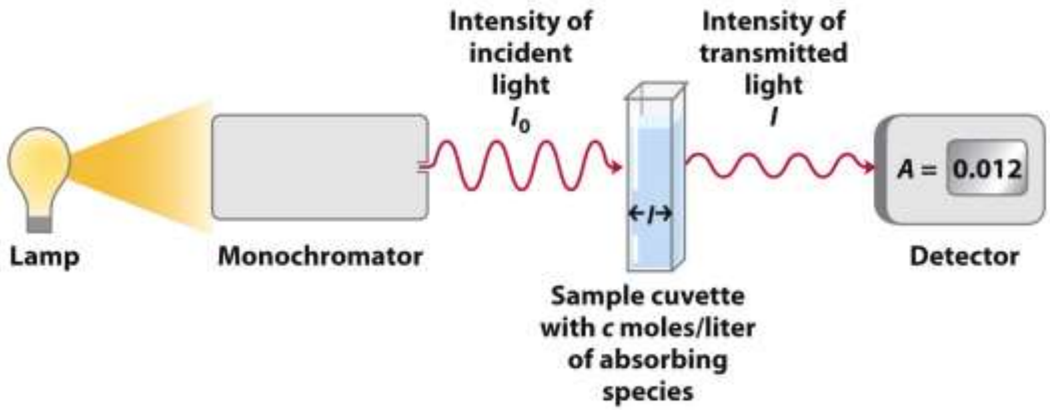
Except glycine

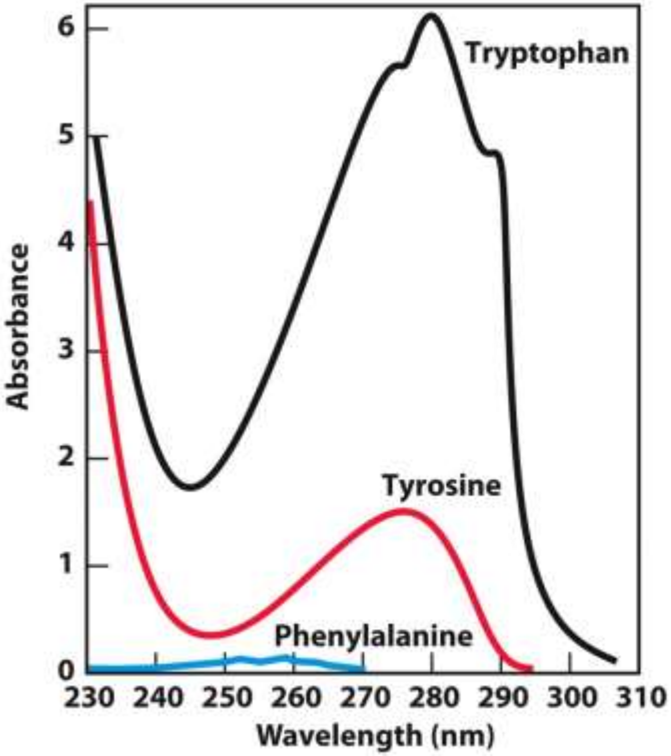


$$H_3\overset{+}{N}-\underset{H}{\overset{COO^-}{C}}-H$$

3.1 Amino Acids

- **Ultraviolet spectra of amino acids**

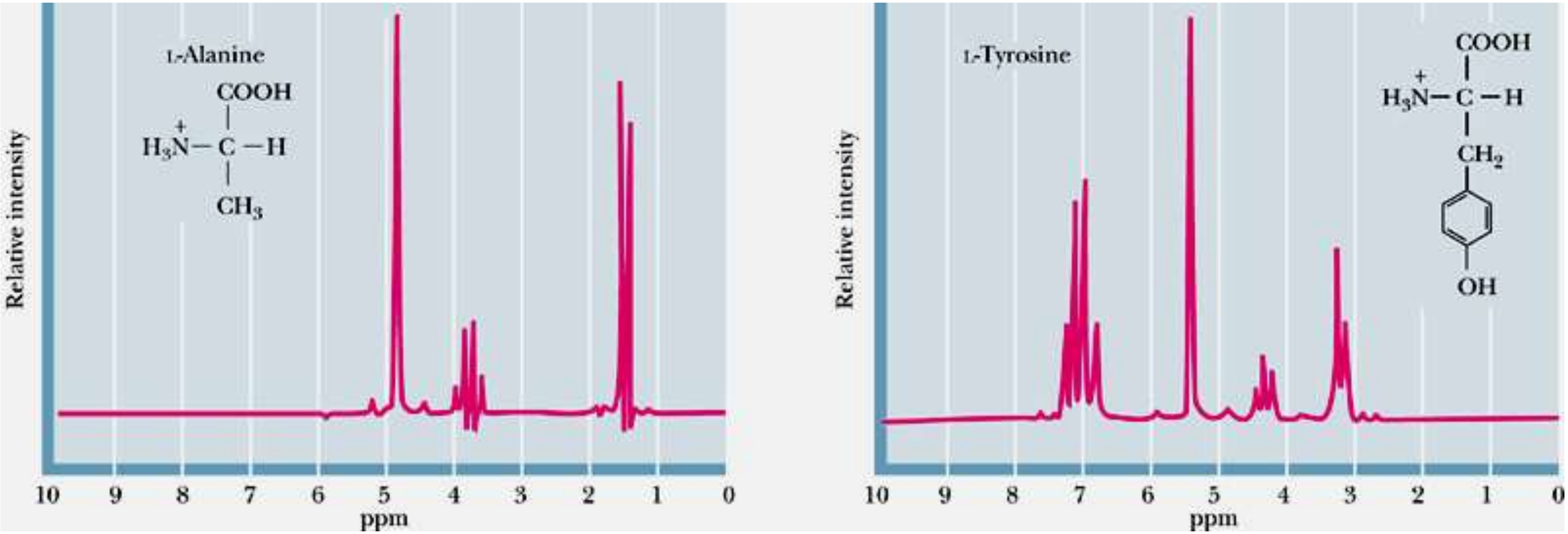




Measurement of light absorption by a spectrophotometer is used to detect and identify molecules and to measure their concentration in solution.

3.1 Amino Acids

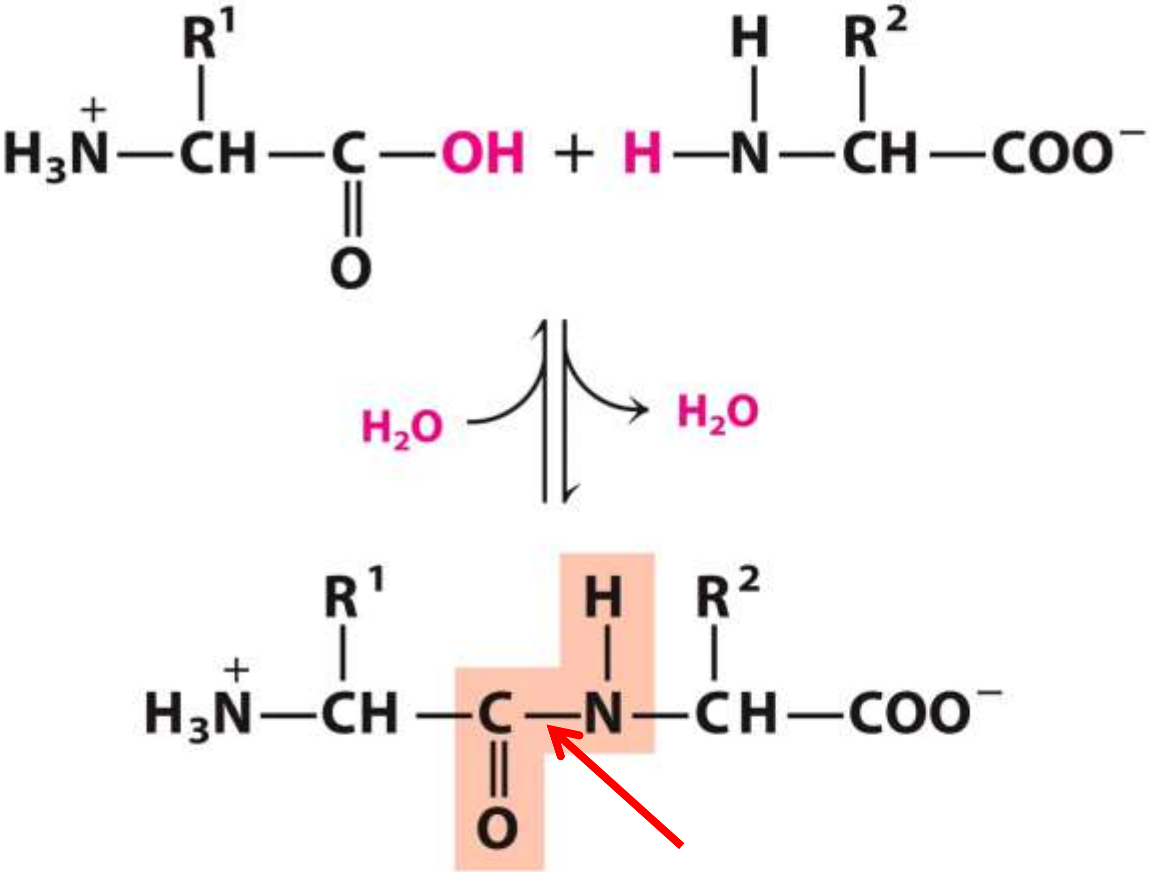
- **Nuclear magnetic resonance spectra of amino acids**



Nuclear Magnetic Resonance (NMR, 核磁共振谱)

3.2 Peptides and Proteins

- **Formation of peptide bond**



Peptide bond

3.2 Peptides and Proteins

■ Peptide classification

Dipeptide

Tripeptide

Tetrapeptide

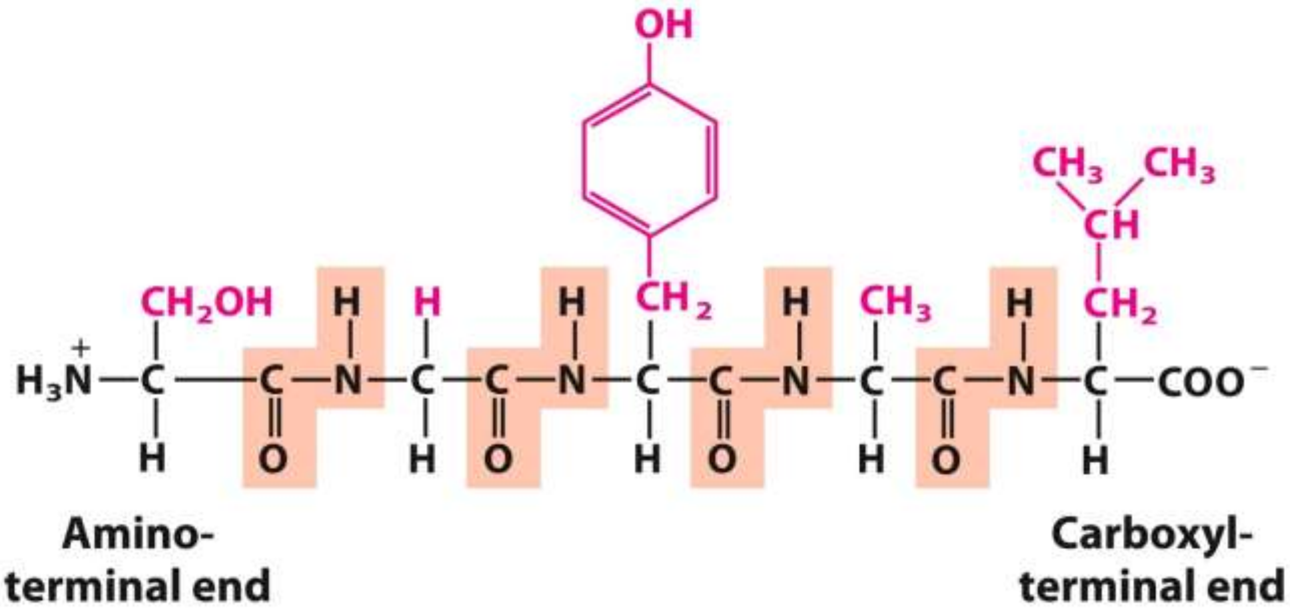
⋮

Oligopeptide (寡肽)

Polypeptide (多肽),

Protein

3.2 Peptides and Proteins



Seryl glycyl tyrosyl alanyl leucine

(丝氨酰甘氨酰酪氨酰丙氨酰亮氨酸)

Ser-G1y-Tyr-A1a-Leu

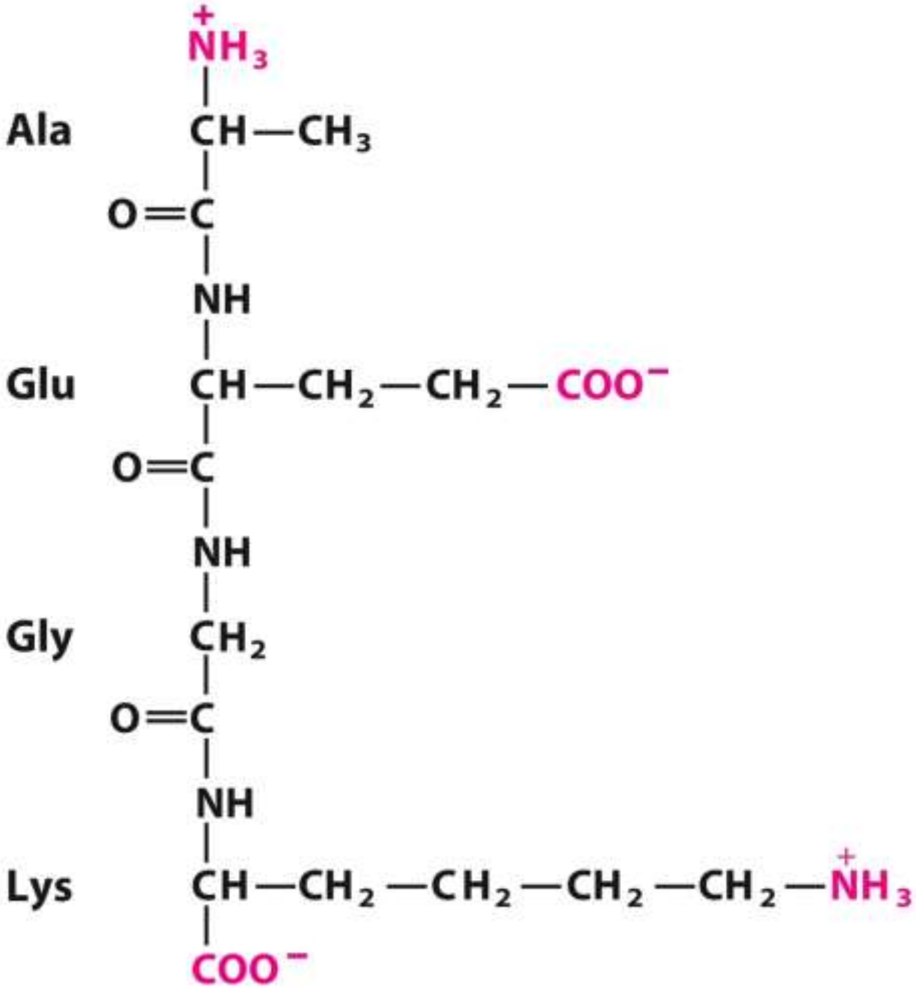
SGYAL

Amino acid residue (氨基酸残基)

3.2 Peptides and Proteins

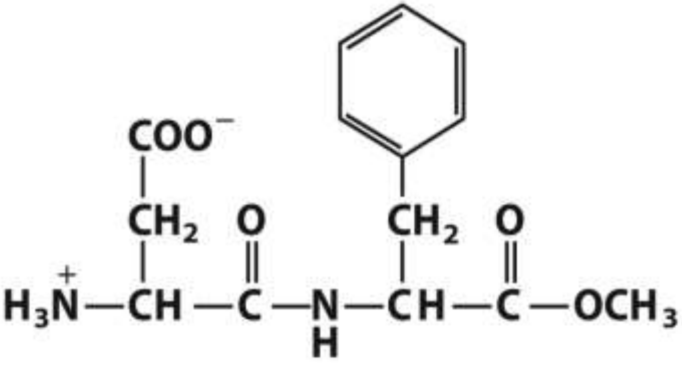
- **Acid-basic chemistry of oligopeptides**

The amino acids with ionizable R groups in a peptide contribute to the overall acid-base properties of the peptide.

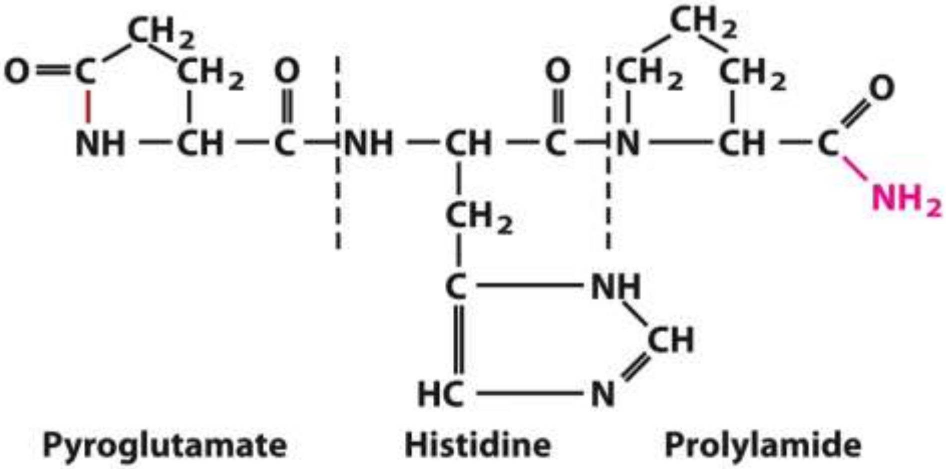


3.2 Peptides and Proteins

- **Biologically active peptides and polypeptides occur in a vast range of sizes and compositions**



L-Aspartyl-L-phenylalanine methyl ester
(aspartame)



pyroGlu–His–Pro–NH_2

NutraSweet

Thyrotropin-releasing hormone (TRH)
(促甲状腺素释放激素)

3.2 Peptides and Proteins

- **Proteins occur in a vast range of sizes**

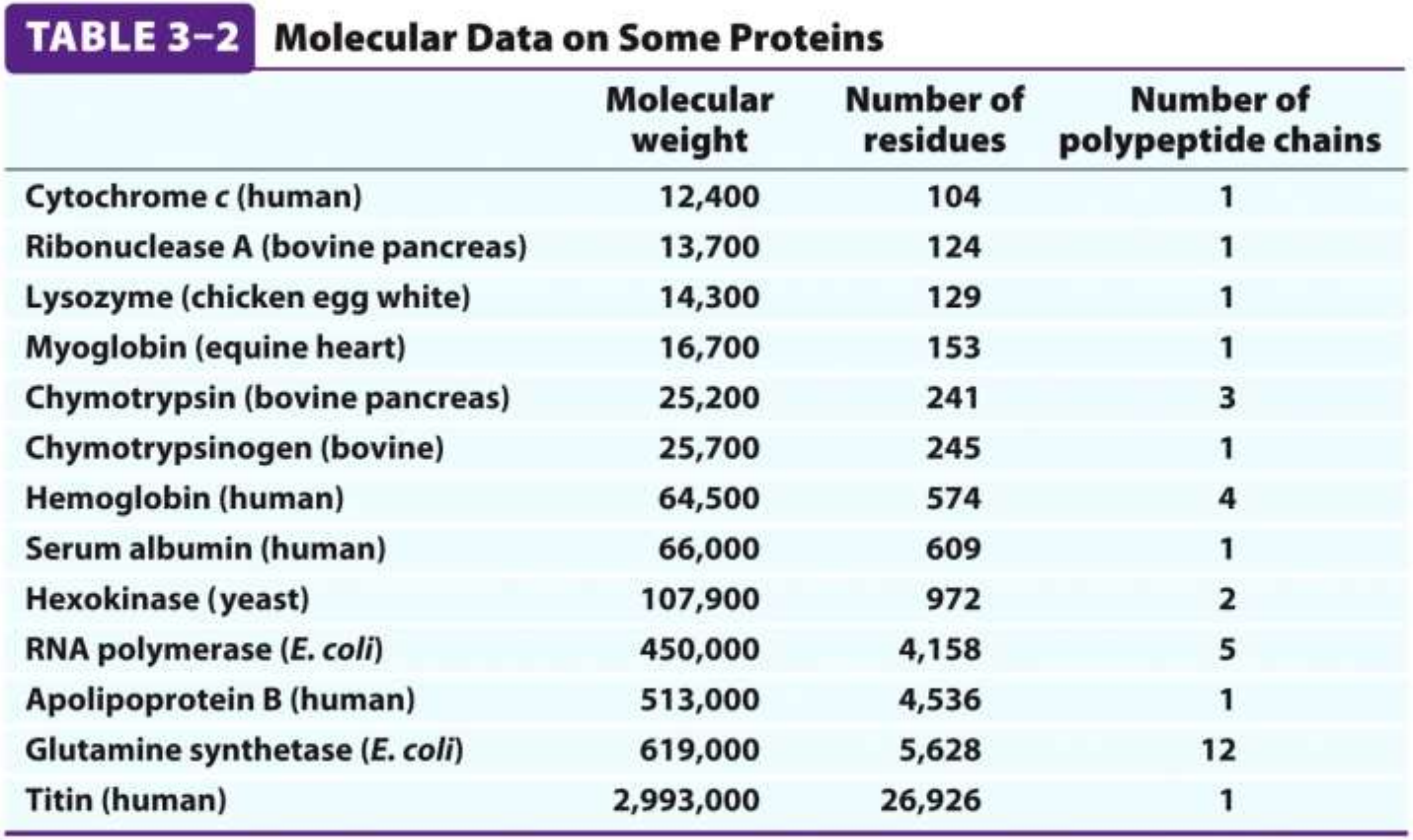


TABLE 3–2 Molecular Data on Some Proteins

	Molecular weight	Number of residues	Number of polypeptide chains
Cytochrome *c* (human)	12,400	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (chicken egg white)	14,300	129	1
Myoglobin (equine heart)	16,700	153	1
Chymotrypsin (bovine pancreas)	25,200	241	3
Chymotrypsinogen (bovine)	25,700	245	1
Hemoglobin (human)	64,500	574	4
Serum albumin (human)	66,000	609	1
Hexokinase (yeast)	107,900	972	2
RNA polymerase (*E. coli*)	450,000	4,158	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamine synthetase (*E. coli*)	619,000	5,628	12
Titin (human)	2,993,000	26,926	1

Average molecular weight of protein amino acid residues is nearer to **110**. Therefore, molecular weight of a protein ≈ number of amino acid residues X 110.

3.2 Peptides and Proteins

- **Amino acid composition of proteins can be different**

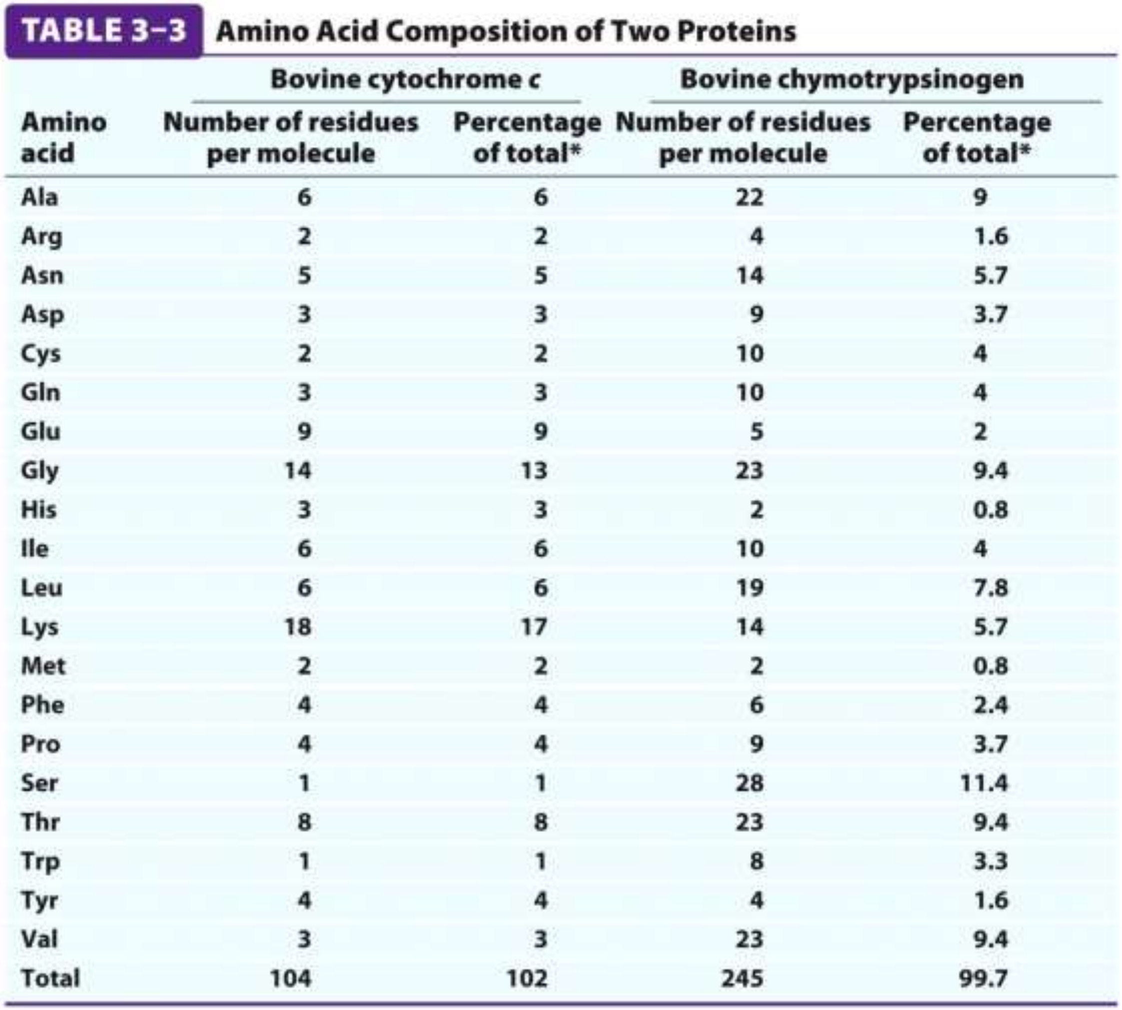


TABLE 3–3 Amino Acid Composition of Two Proteins

Amino acid	Bovine cytochrome *c* Number of residues per molecule	Bovine cytochrome *c* Percentage of total*	Bovine chymotrypsinogen Number of residues per molecule	Bovine chymotrypsinogen Percentage of total*
Ala	6	6	22	9
Arg	2	2	4	1.6
Asn	5	5	14	5.7
Asp	3	3	9	3.7
Cys	2	2	10	4
Gln	3	3	10	4
Glu	9	9	5	2
Gly	14	13	23	9.4
His	3	3	2	0.8
Ile	6	6	10	4
Leu	6	6	19	7.8
Lys	18	17	14	5.7
Met	2	2	2	0.8
Phe	4	4	6	2.4
Pro	4	4	9	3.7
Ser	1	1	28	11.4
Thr	8	8	23	9.4
Trp	1	1	8	3.3
Tyr	4	4	4	1.6
Val	3	3	23	9.4
Total	104	102	245	99.7

3.2 Peptides and Proteins

- **Proteins may contain chemical groups other than amino acids**

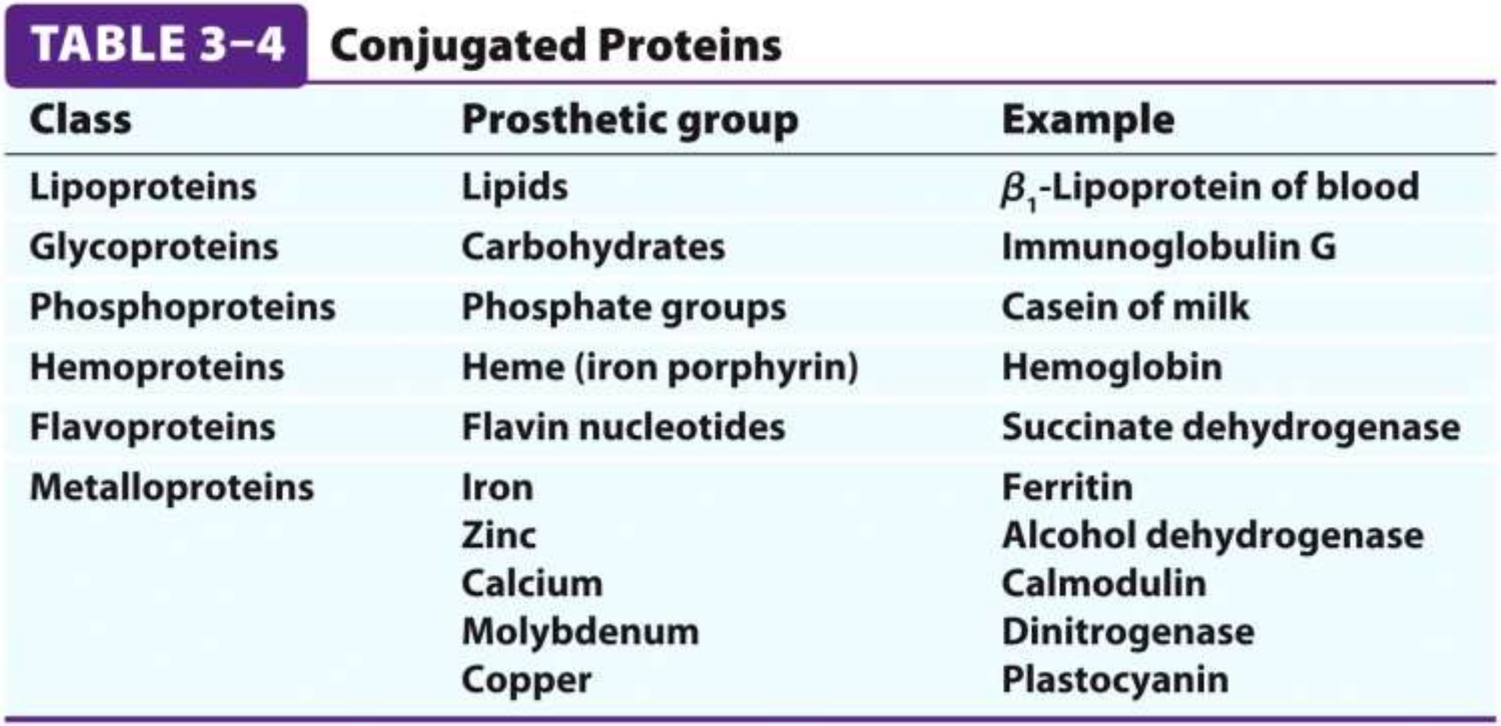


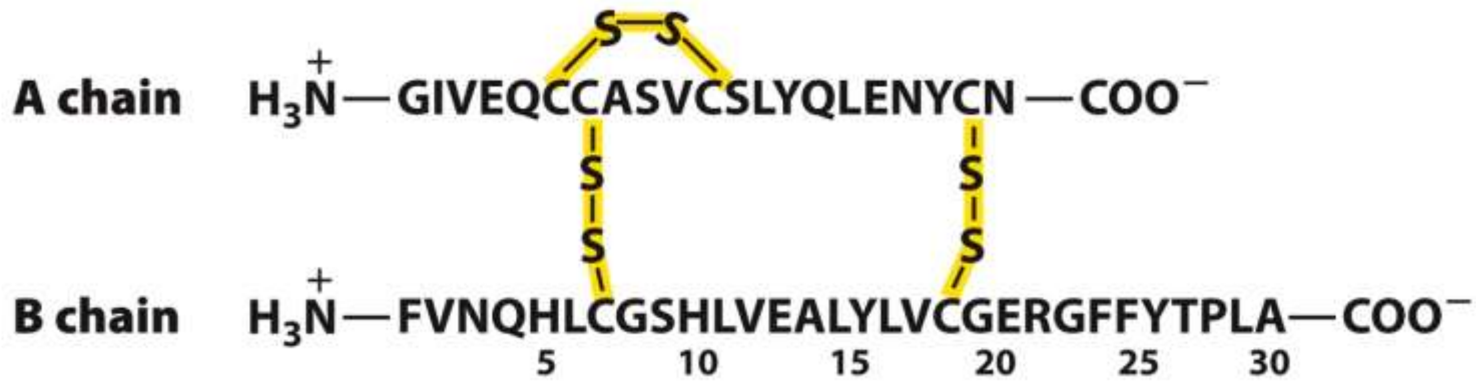
TABLE 3–4 Conjugated Proteins

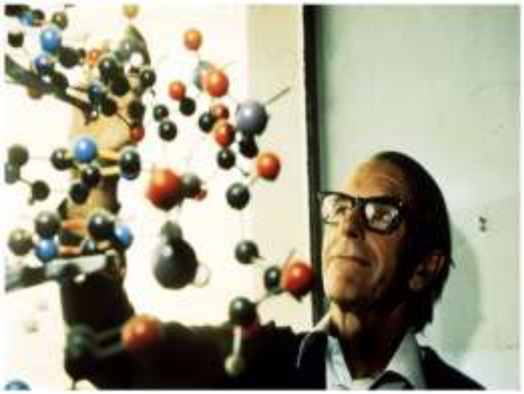
Class	Prosthetic group	Example
Lipoproteins	Lipids	β_1-Lipoprotein of blood
Glycoproteins	Carbohydrates	Immunoglobulin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron Zinc Calcium Molybdenum Copper	Ferritin Alcohol dehydrogenase Calmodulin Dinitrogenase Plastocyanin

3.2 Peptides and Proteins

- **Determining the amino acid sequence**

Bovine insulin, first peptide sequence determined





Frederick Sanger
1958 Nobel Prize in Chemistry
1980 Nobel Prize in Chemistry

3.2 Peptides and Proteins

- **The procedure developed by Fred Sanger to sequence insulin**

1. Purify the target protein
2. Determine amino terminus
3. Amino acids analysis
4. Breaking disulfide bonds
5. Fragmentation of the polypeptide chain
6. Determine the sequence
7. Location of the disulfide bonds

Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

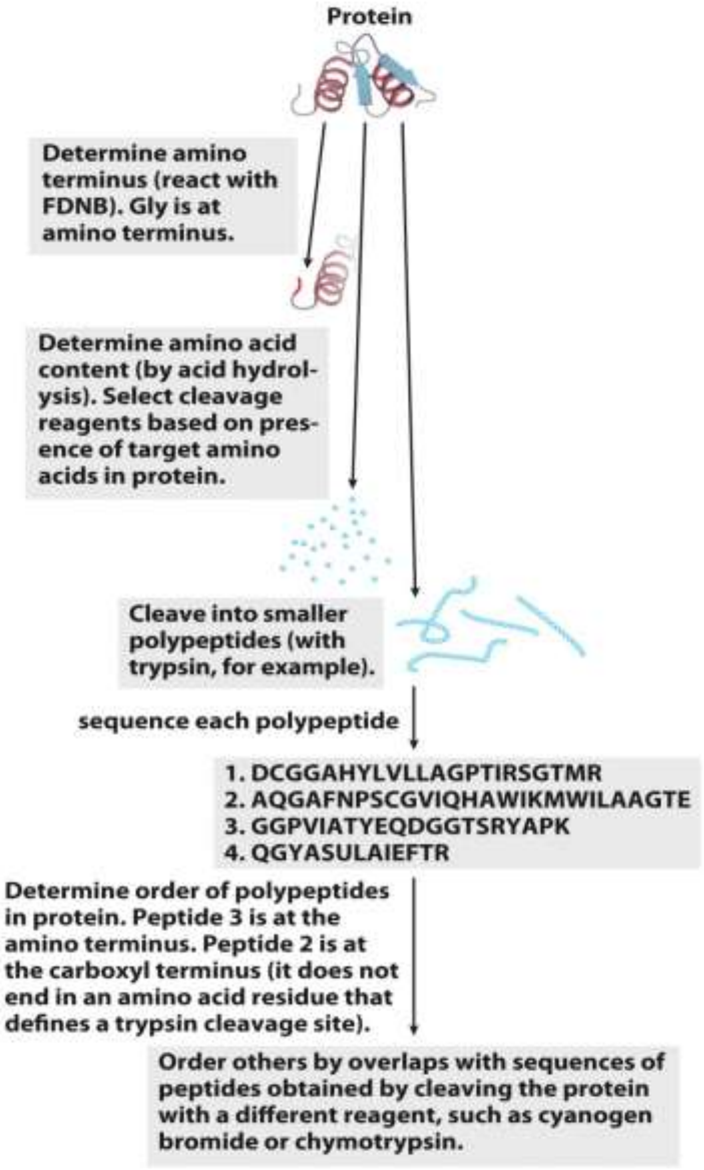
Edited by James F. Crow and William F. Dove

The First Sequence: Fred Sanger and Insulin

Antony O. W. Stretton

Department of Zoology, University of Wisconsin, Madison, Wisconsin 53706

***Genetics* 162:** 527–532 (October 2002)



Determine amino terminus (react with FDNB). Gly is at amino terminus.

Determine amino acid content (by acid hydrolysis). Select cleavage reagents based on presence of target amino acids in protein.

Cleave into smaller polypeptides (with trypsin, for example).

sequence each polypeptide

1. DCGGAHYLVLLAGPTIRSGTMR
2. AQGAFNPSCGVIQHAWIKMWILAAGTE
3. GGPVIATYEQDGGTSRYAPK
4. QGYASULAIEFTR

Determine order of polypeptides in protein. Peptide 3 is at the amino terminus. Peptide 2 is at the carboxyl terminus (it does not end in an amino acid residue that defines a trypsin cleavage site).

Order others by overlaps with sequences of peptides obtained by cleaving the protein with a different reagent, such as cyanogen bromide or chymotrypsin.

3.2 Peptides and Proteins

- **N-Terminal analysis:**

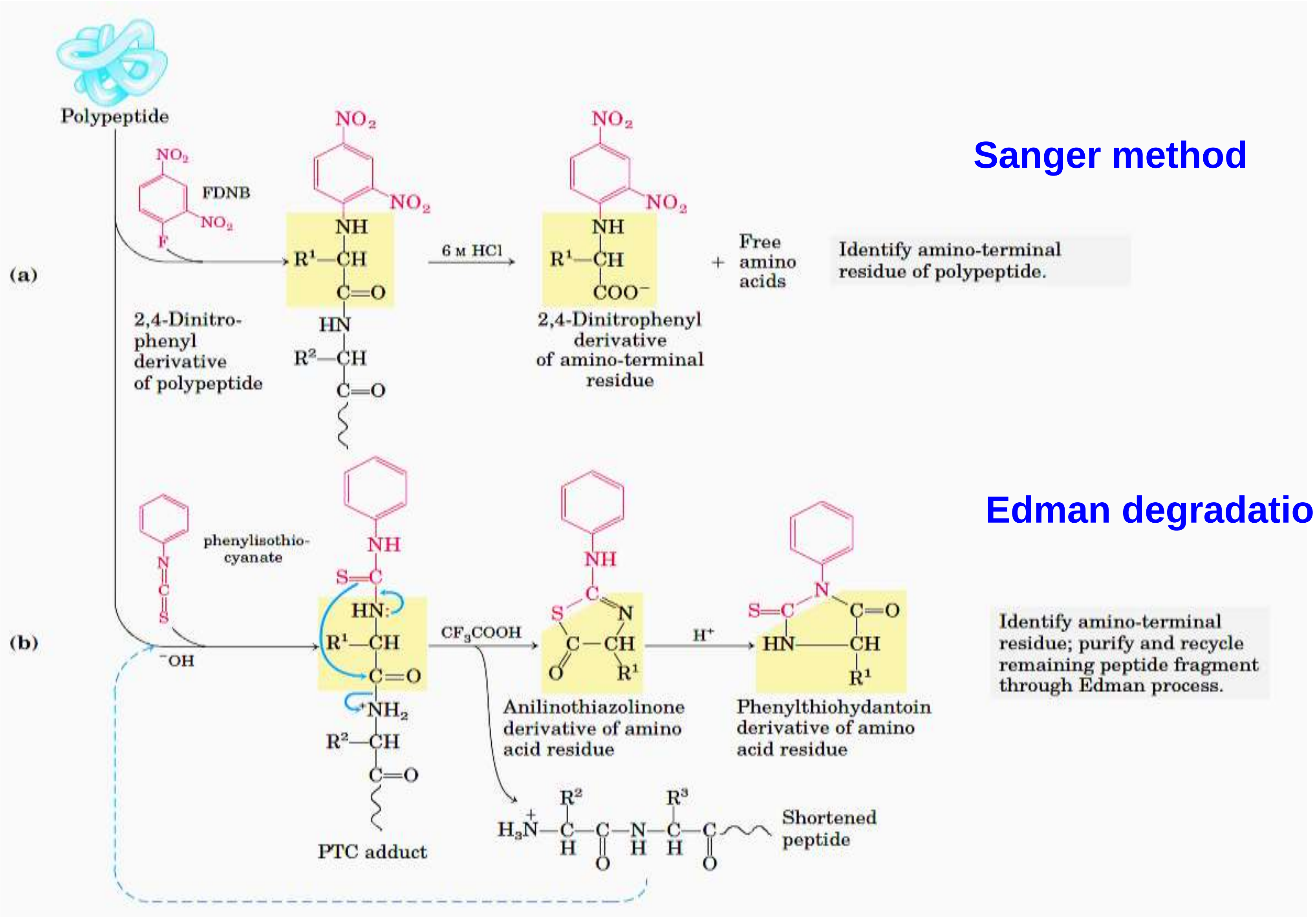
Sanger method

FDNB (1-fluoro-2,4-dinitrobenzene)

Edman degradation

phenyl-N=C=S （异硫氰酸苯酯）

3.2 Peptides and Proteins



Sanger method

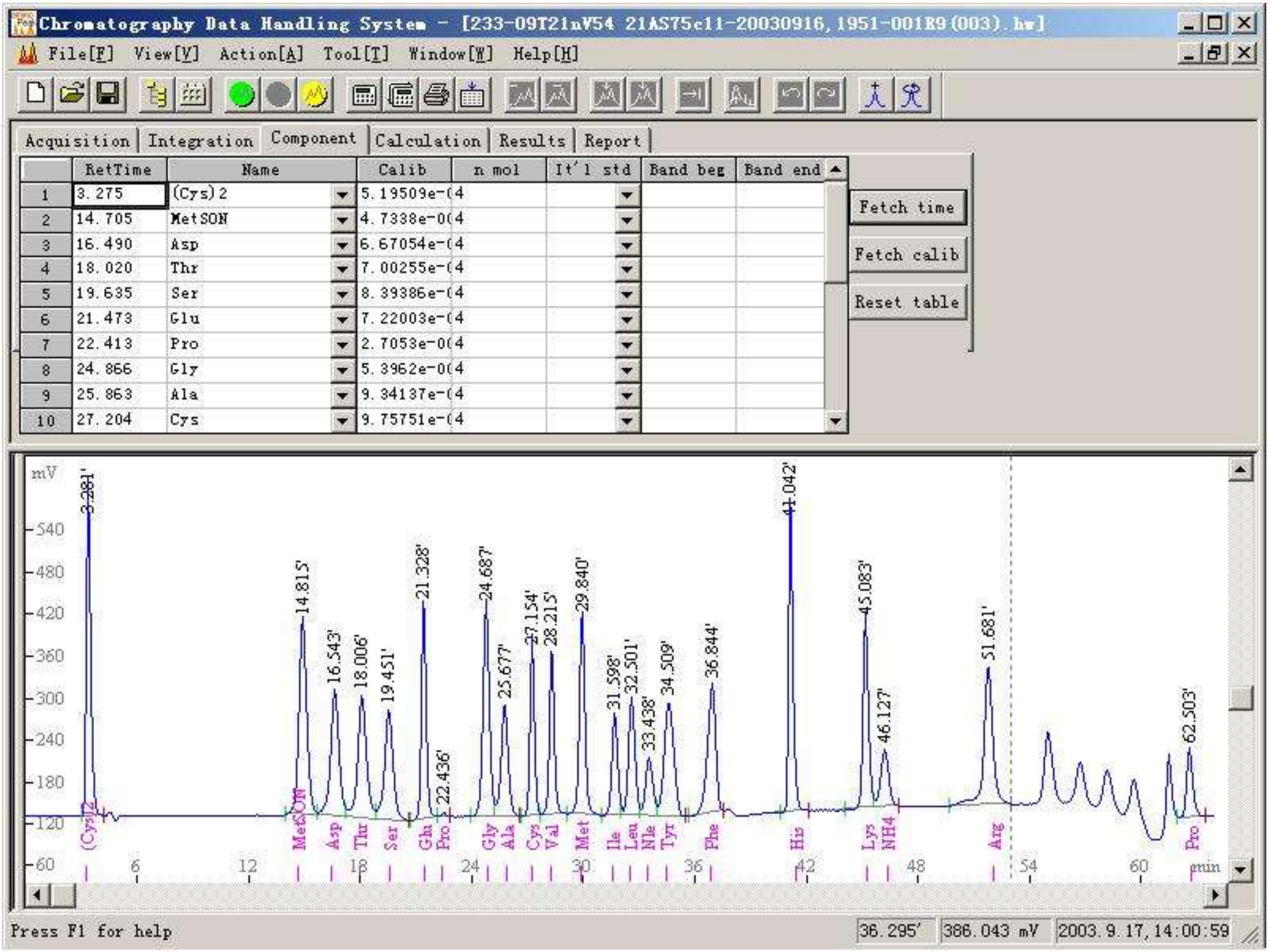
Edman degradation

3.2 Peptides and Proteins

- **Amino acids analysis**

Hydrolysis of peptide

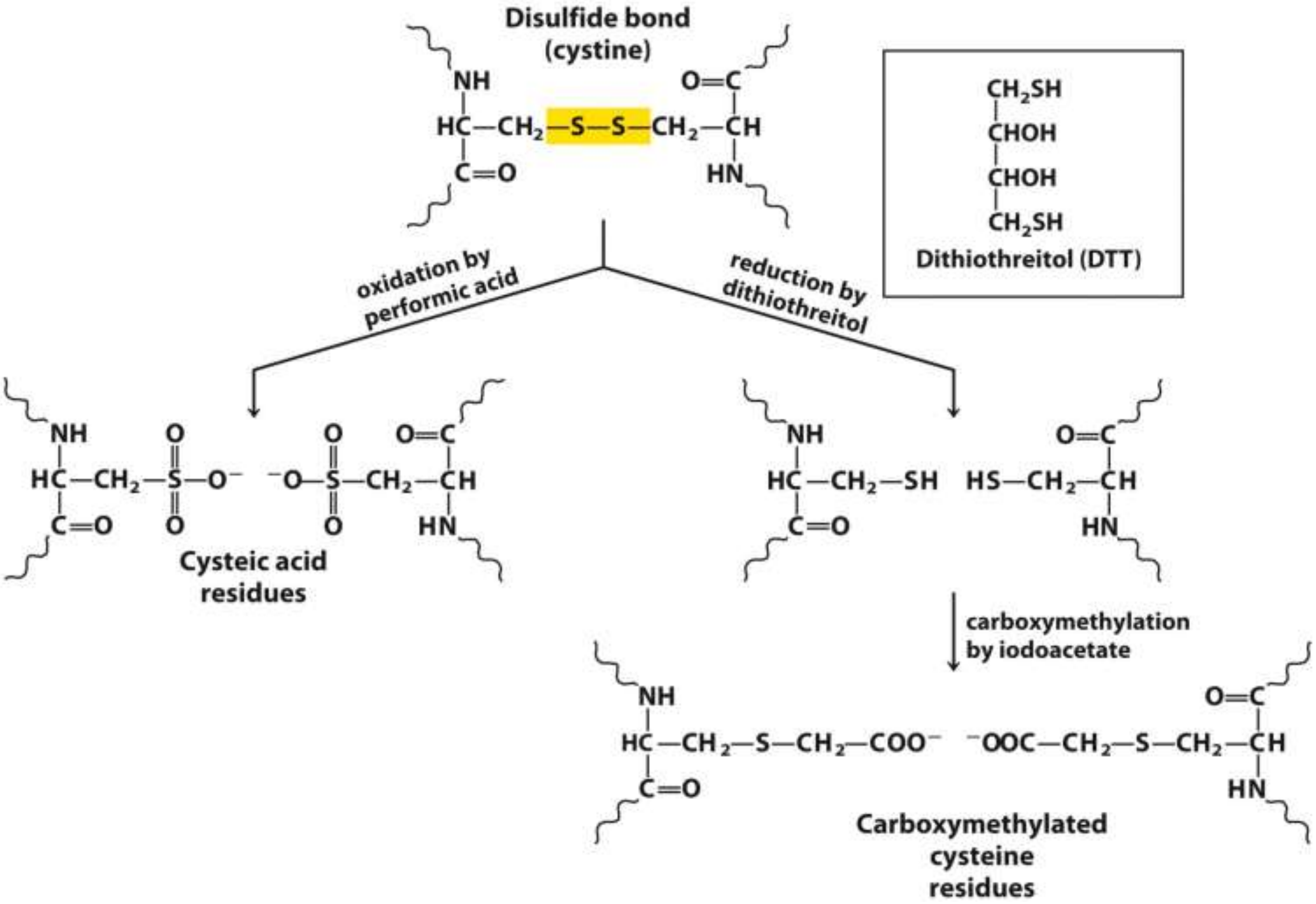
6N HCl, 110°C, overnight



	RetTime	Name	Calib	n mol	It'l std	Band beg	Band end
1	3.275	(Cys)2	5.19509e-(	4			
2	14.705	MetSON	4.7338e-0(	4			
3	16.490	Asp	6.67054e-(	4			
4	18.020	Thr	7.00255e-(	4			
5	19.635	Ser	8.39386e-(	4			
6	21.473	Glu	7.22003e-(	4			
7	22.413	Pro	2.7053e-0(	4			
8	24.866	Gly	5.3962e-0(	4			
9	25.863	Ala	9.34137e-(	4			
10	27.204	Cys	9.75751e-(	4			

3.2 Peptides and Proteins

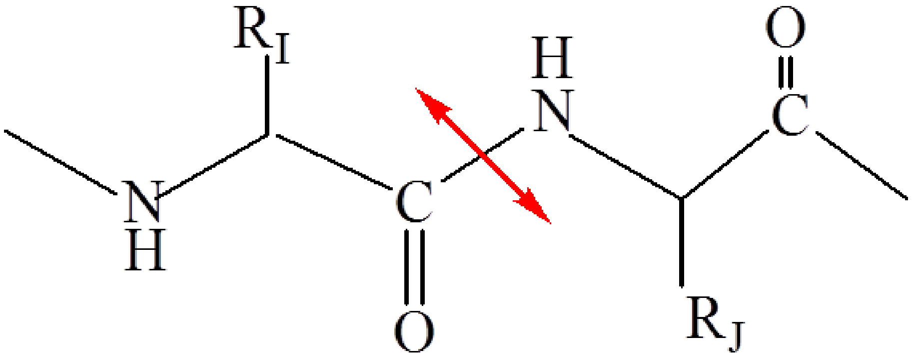
- **Cleavage of disulfide bonds**



3.2 Peptides and Proteins

- **Fragmentation of the polypeptide chain**

1) Fragmentation of the polypeptide by protease

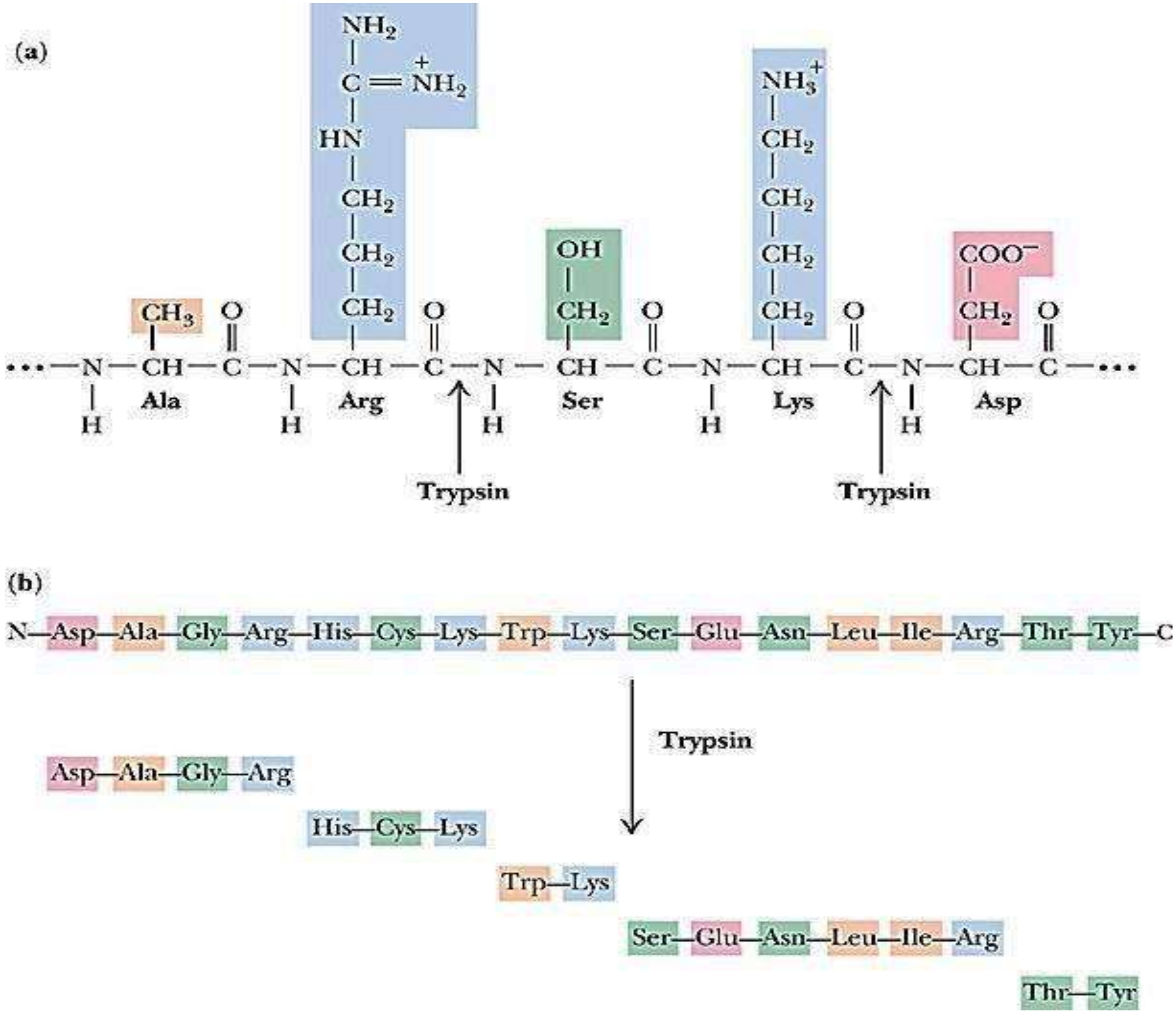


Trypsin (胰蛋白酶)：R_I = Lys、Arg，$R_J \neq$ Pro

Chymotrypsin (糜蛋白酶)：R_I = Phe、Trp、Tyr，$R_J \neq$ Pro

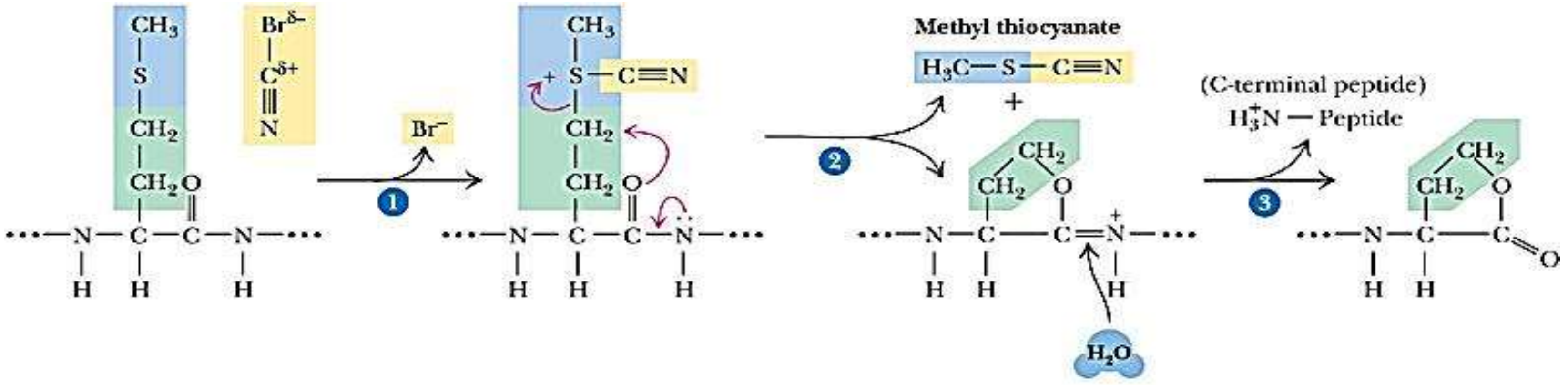
Pepsin (胃蛋白酶)：R_I、R_J = 芳香族或较大的疏水基团；$R_J \neq$ Pro

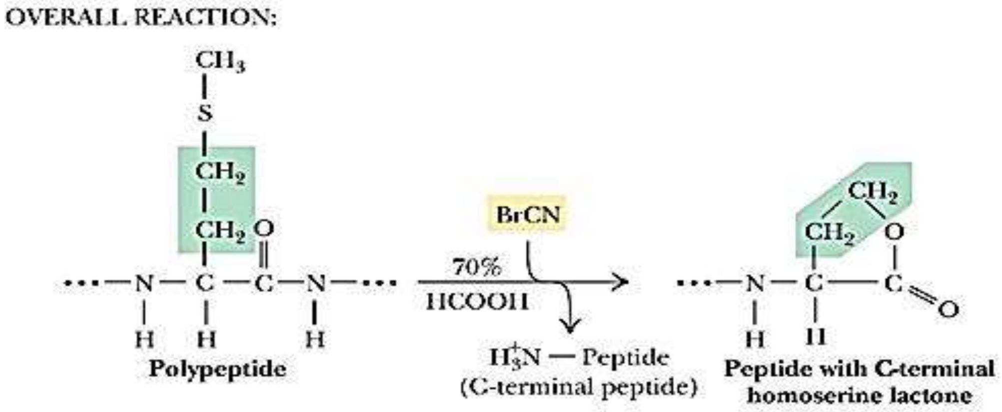
3.2 Peptides and Proteins

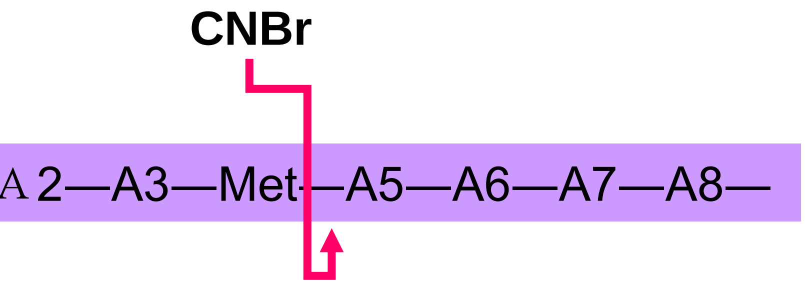


3.2 Peptides and Proteins

2) Fragmentation of the polypeptide chain by CNBr







CNBr

A1— A 2—A3—Met—A5—A6—A7—A8—

3.2 Peptides and Proteins

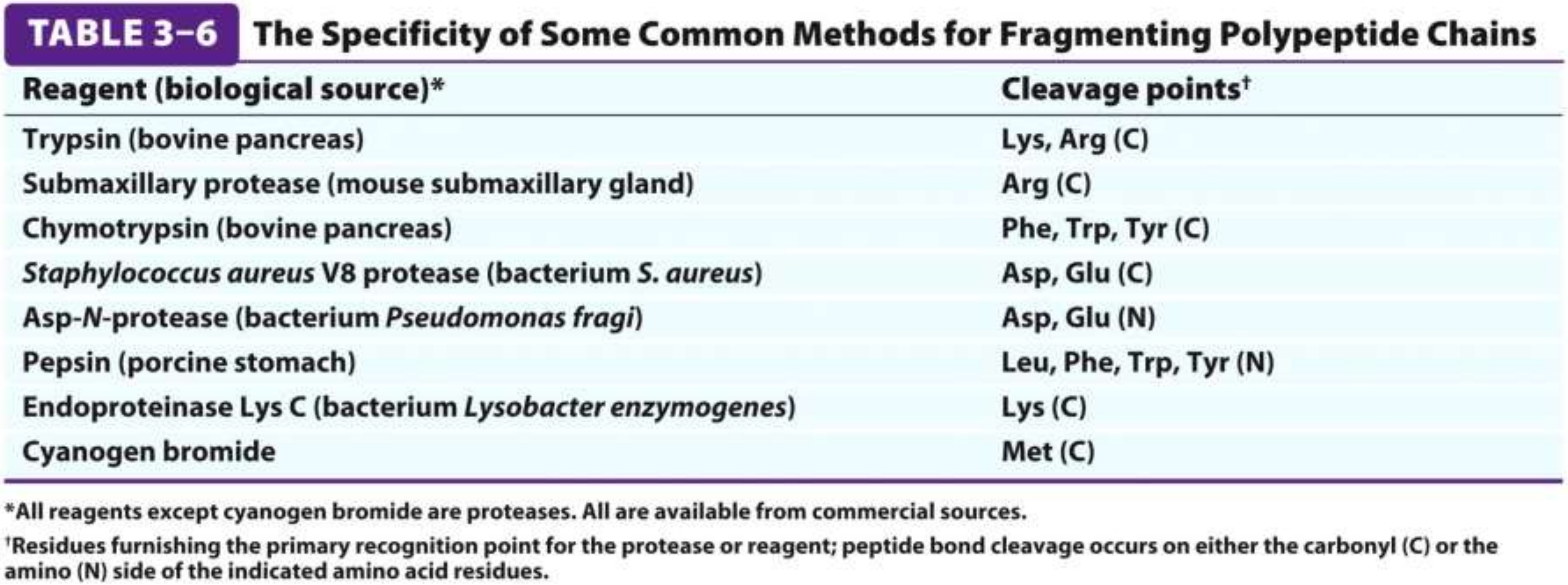


TABLE 3–6 **The Specificity of Some Common Methods for Fragmenting Polypeptide Chains**

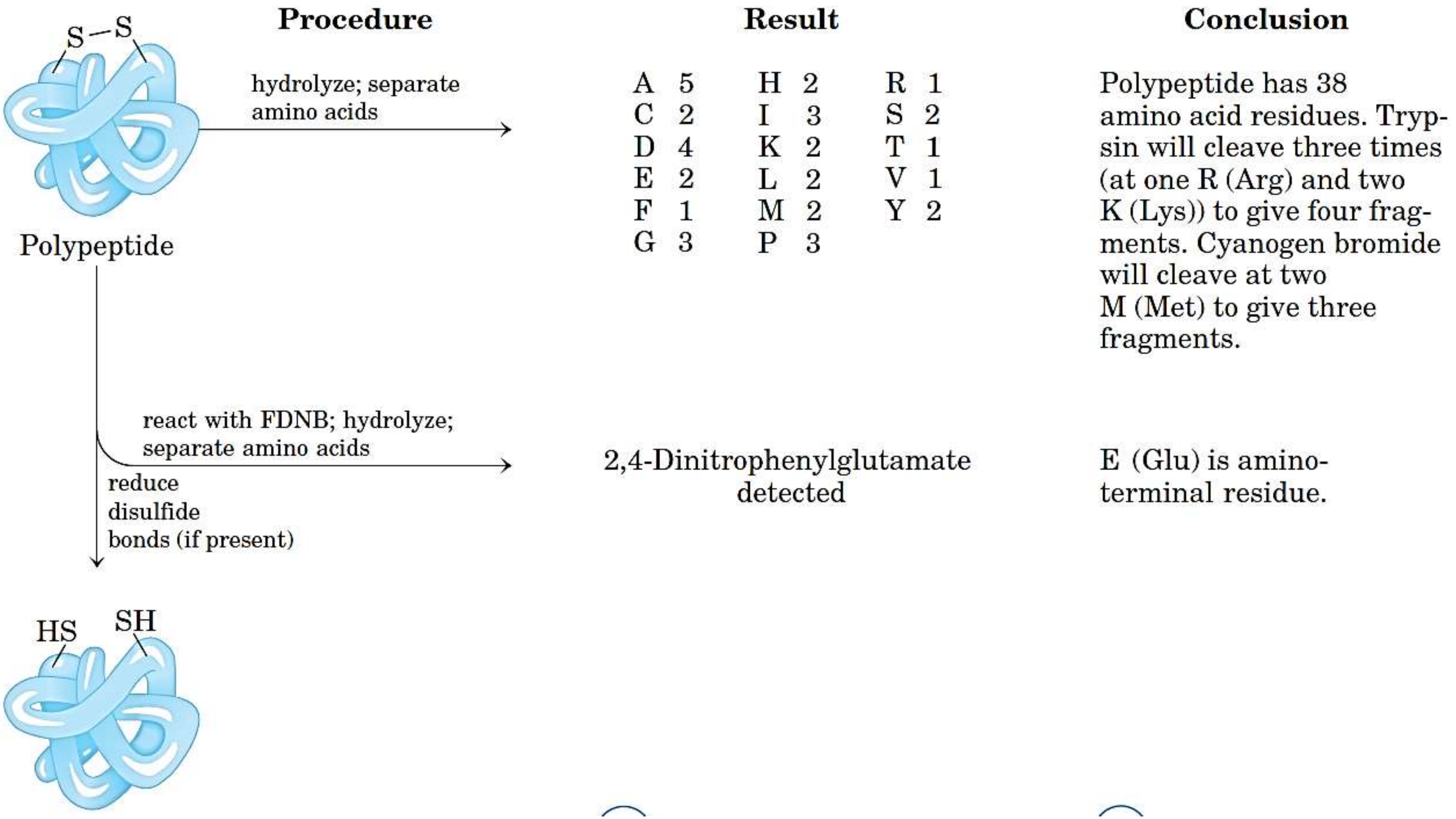
Reagent (biological source)*	Cleavage points†
Trypsin (bovine pancreas)	Lys, Arg (C)
Submaxillary protease (mouse submaxillary gland)	Arg (C)
Chymotrypsin (bovine pancreas)	Phe, Trp, Tyr (C)
Staphylococcus aureus V8 protease (bacterium *S. aureus*)	Asp, Glu (C)
Asp-*N*-protease (bacterium *Pseudomonas fragi*)	Asp, Glu (N)
Pepsin (porcine stomach)	Leu, Phe, Trp, Tyr (N)
Endoproteinase Lys C (bacterium *Lysobacter enzymogenes*)	Lys (C)
Cyanogen bromide	Met (C)

*All reagents except cyanogen bromide are proteases. All are available from commercial sources.

†Residues furnishing the primary recognition point for the protease or reagent; peptide bond cleavage occurs on either the carbonyl (C) or the amino (N) side of the indicated amino acid residues.

3.2 Peptides and Proteins

Sample： **Deduce the amino acid sequence of peptide chain**



Procedure

hydrolyze; separate amino acids

react with FDNB; hydrolyze; separate amino acids

reduce disulfide bonds (if present)

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		

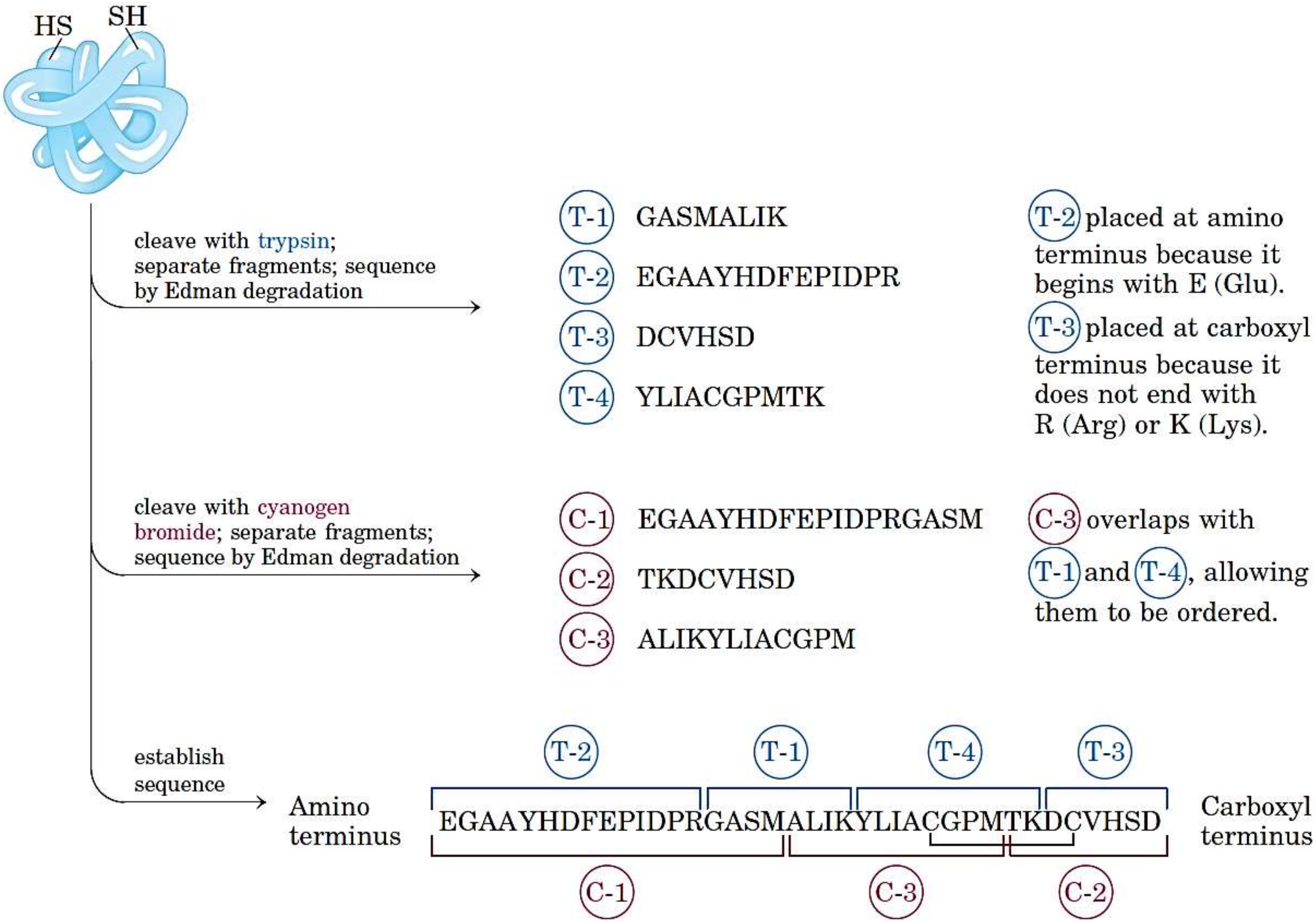
2,4-Dinitrophenylglutamate detected

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.

3.2 Peptides and Proteins



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

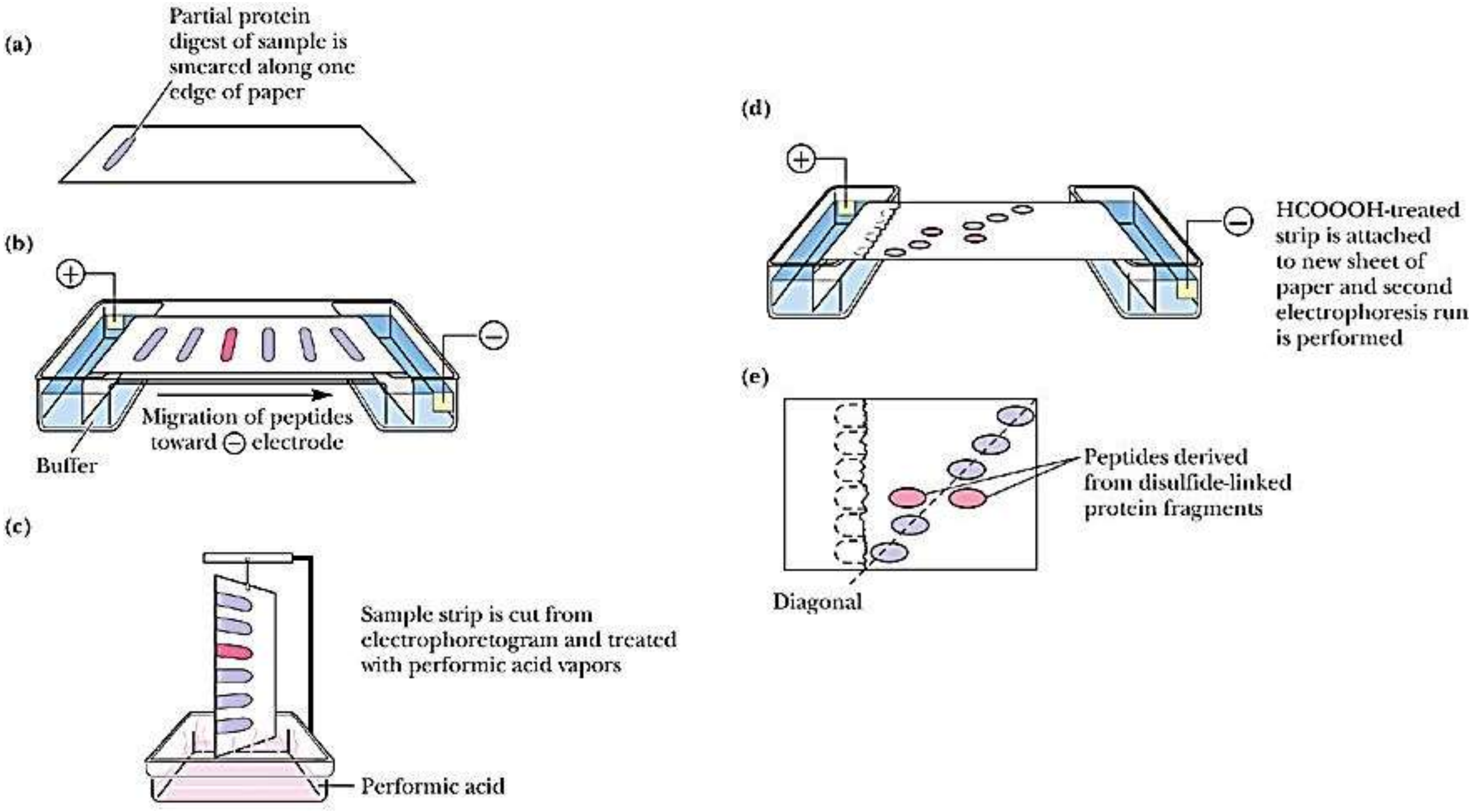
C-3 overlaps with T-1 and T-4, allowing them to be ordered.

establish sequence

Amino terminus EGAAYHDFEPIDPRGASMALIKYLIACGPMTKDCVHSD Carboxyl terminus

3.2 Peptides and Proteins

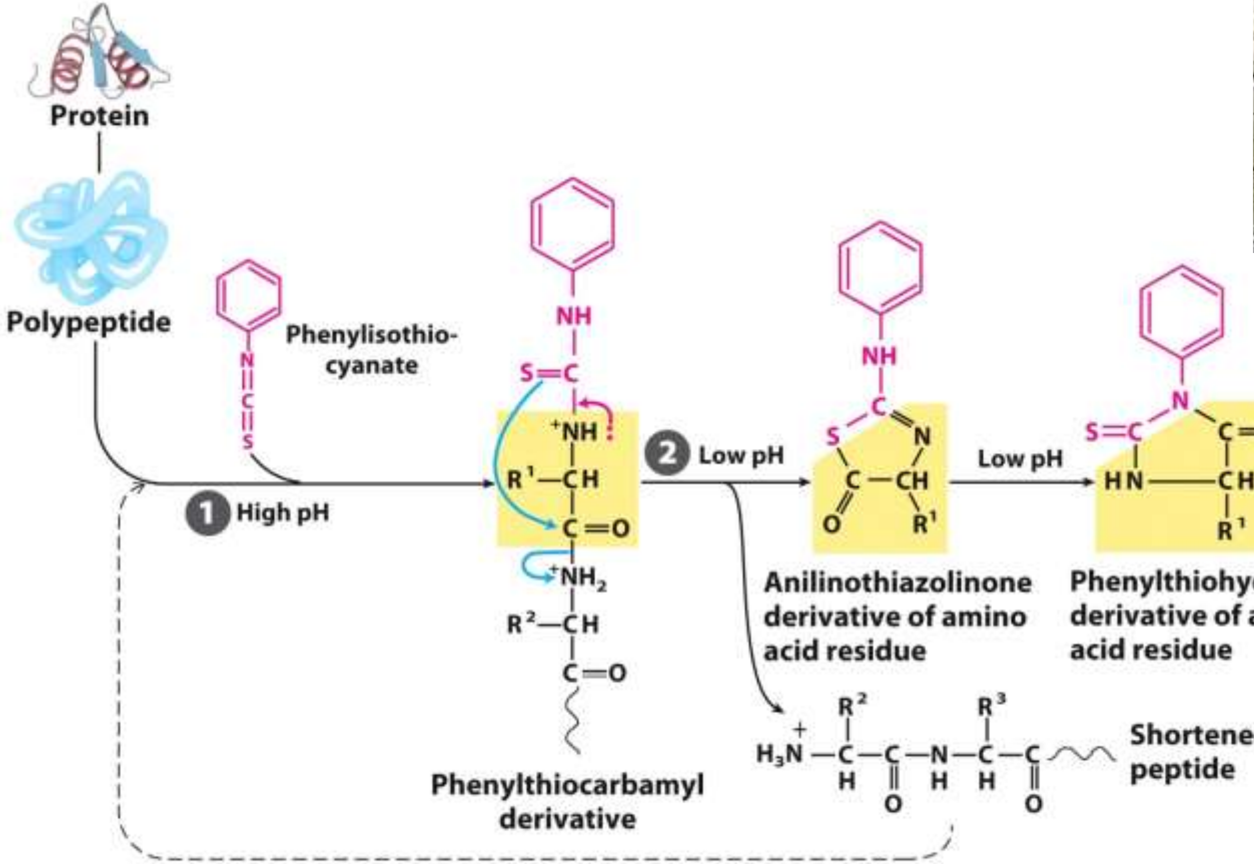
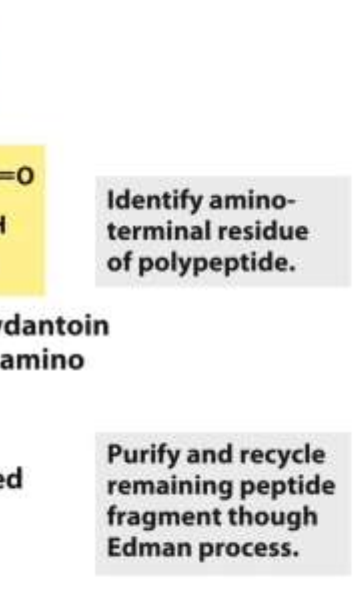
- **Location of disulfide bonds**



3.2 Peptides and Proteins

- **Sequencing of amino acids in polypeptide today**

Edman sequenator

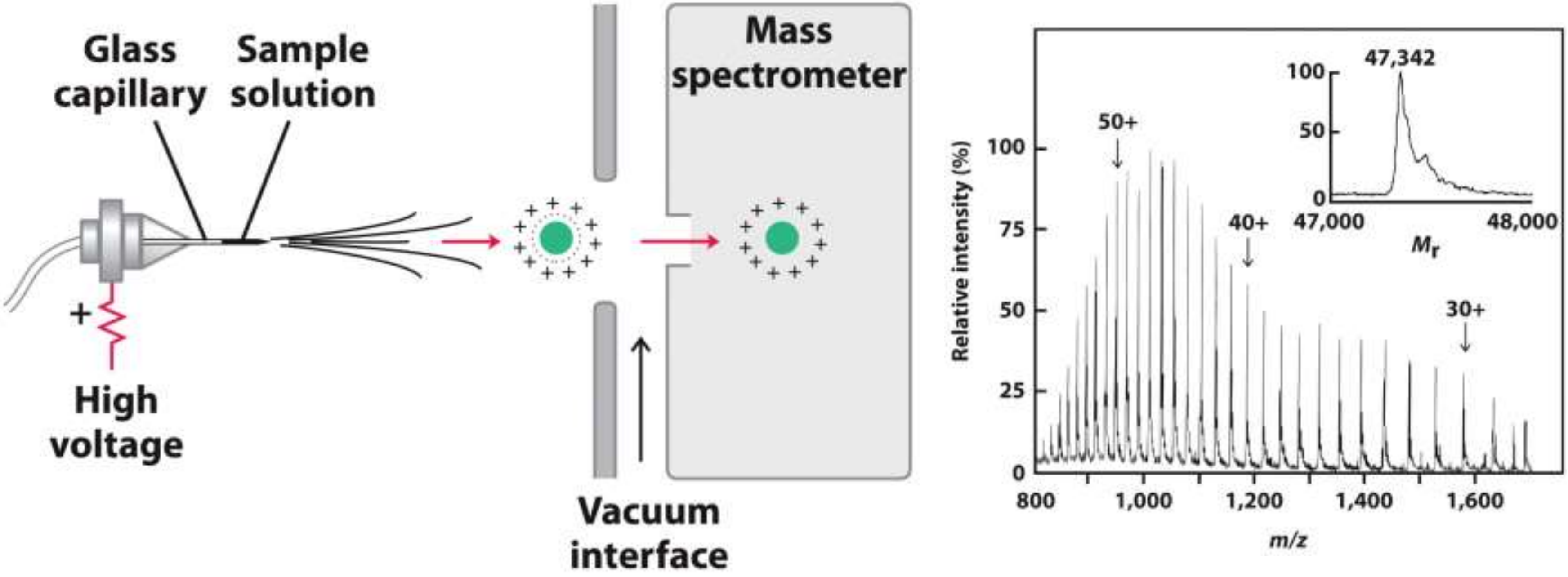



Identify amino-terminal residue of polypeptide.

Purify and recycle remaining peptide fragment though Edman process.

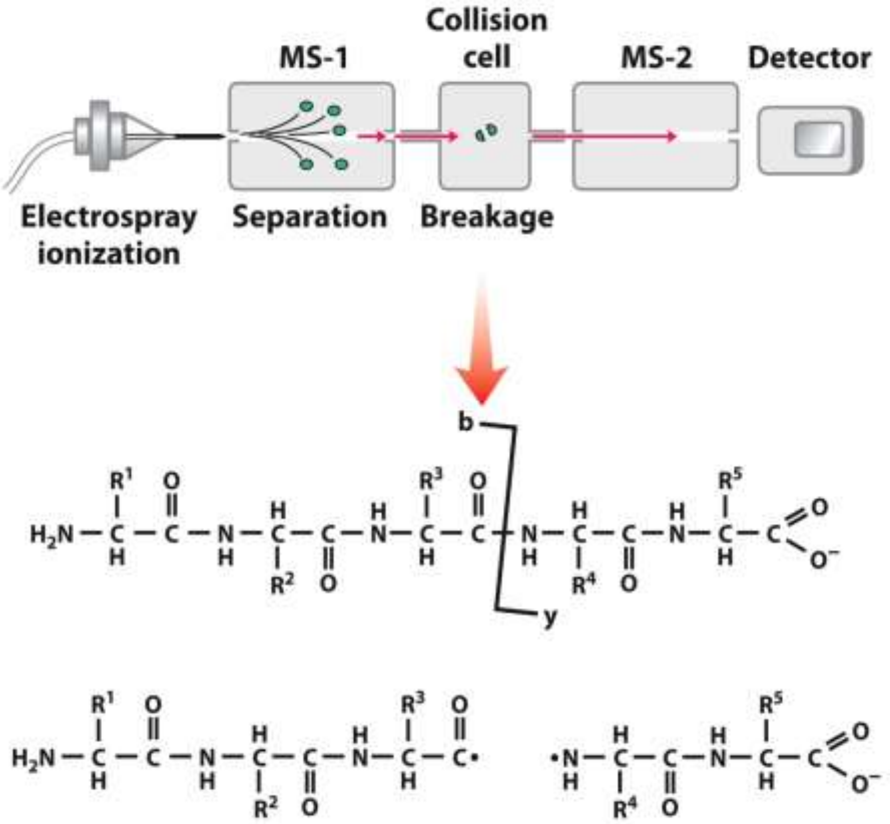
3.2 Peptides and Proteins

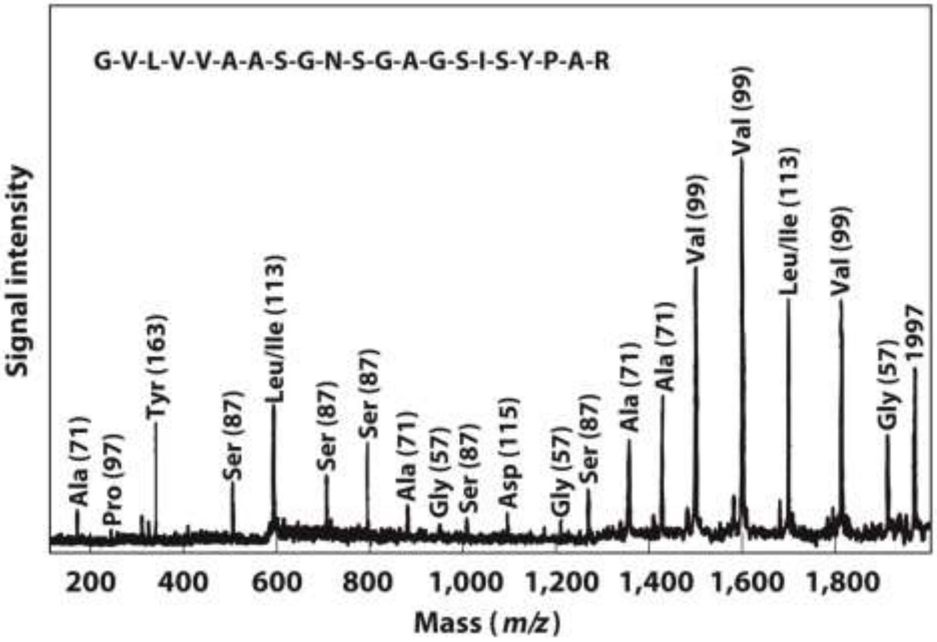
Mass spectrometry



3.2 Peptides and Proteins

Obtaining protein sequence information with tandem MS





3.2 Peptides and Proteins

■ Chemical synthesis of peptides

R. Bruce Merrifield
1921–2006

Unnumbered 3 p101
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

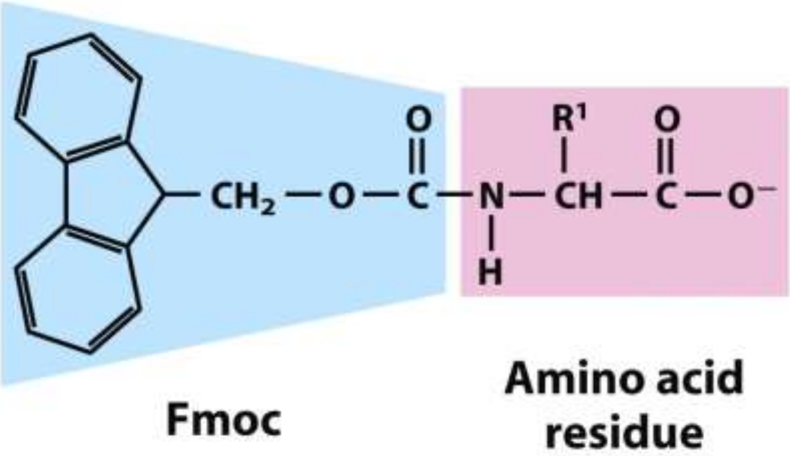
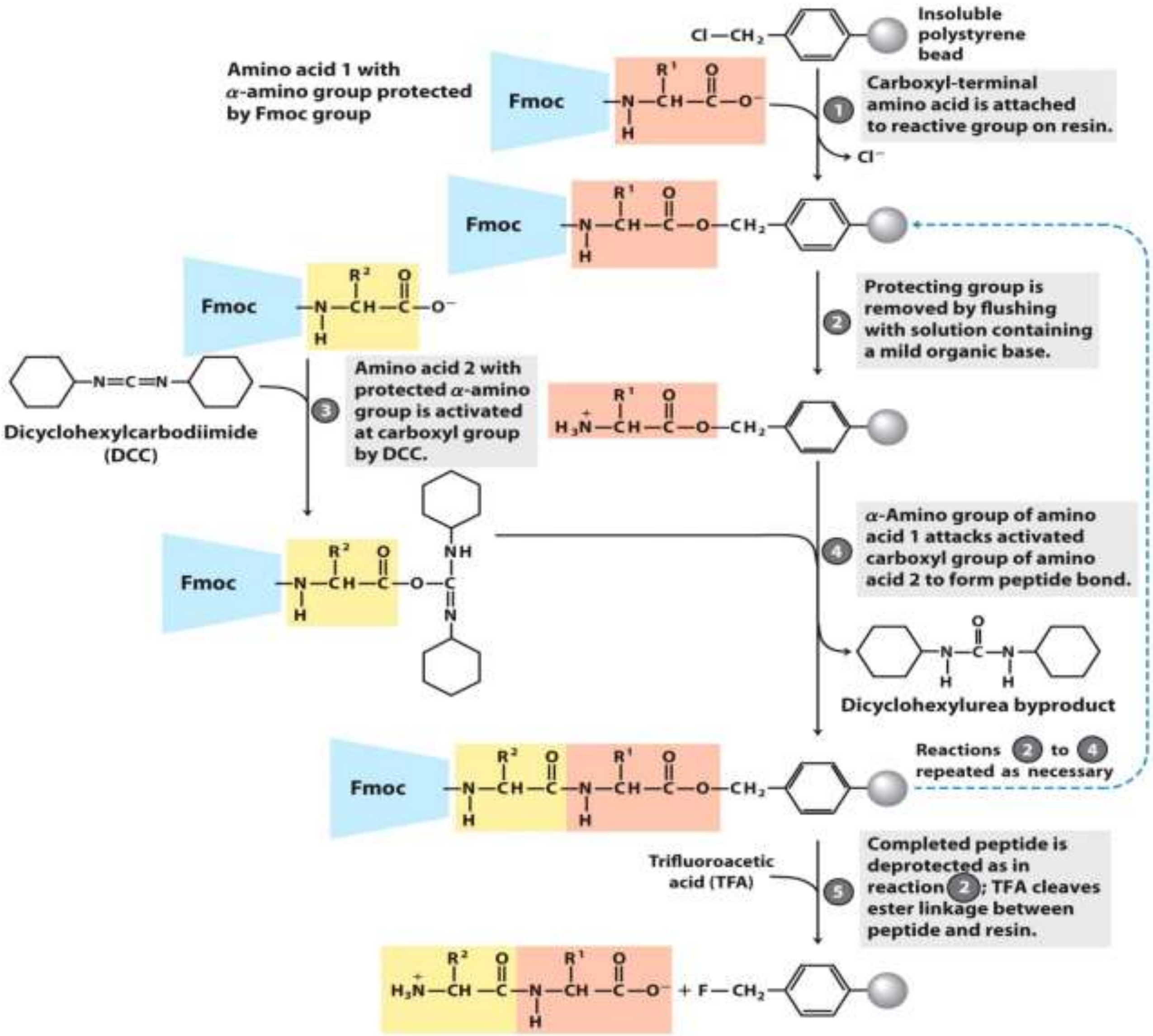


Figure 3-29a
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W.H. Freeman and Company

9-fluorenylmethoxycarbonyl (Fmoc)

3.2 Peptides and Proteins



① Carboxyl-terminal amino acid is attached to reactive group on resin.

② Protecting group is removed by flushing with solution containing a mild organic base.

③ Amino acid 2 with protected α-amino group is activated at carboxyl group by DCC.

④ α-Amino group of amino acid 1 attacks activated carboxyl group of amino acid 2 to form peptide bond.

⑤ Completed peptide is deprotected as in reaction ②; TFA cleaves ester linkage between peptide and resin.

3.2 Peptides and Proteins

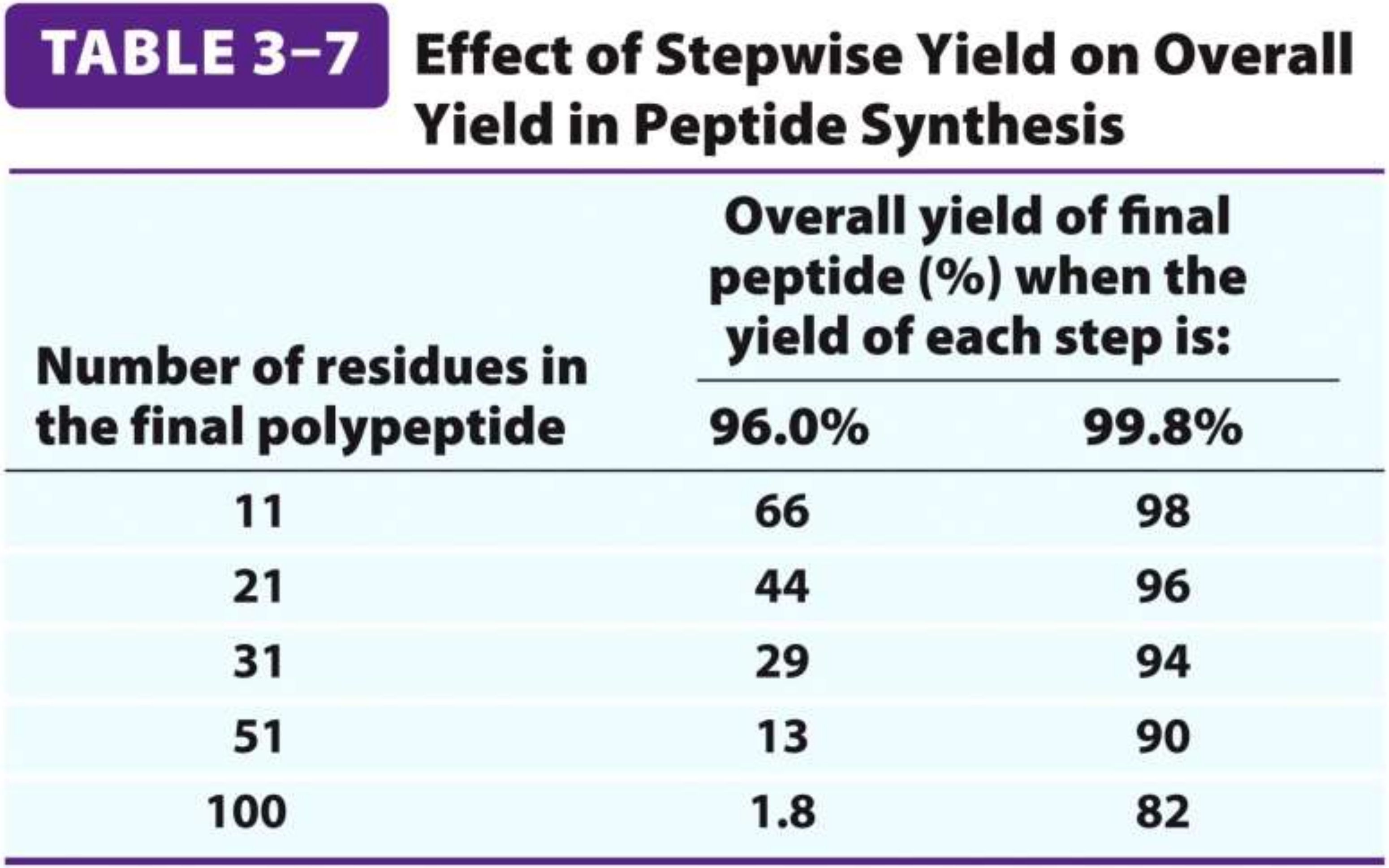
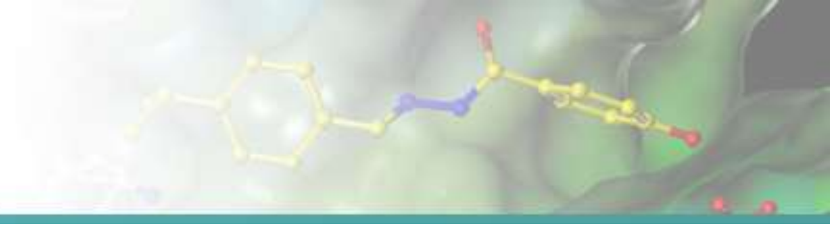


TABLE 3–7 Effect of Stepwise Yield on Overall Yield in Peptide Synthesis

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	1.8	82



3.2 Peptides and Proteins

- **Amino acid sequences provide important biochemical information**

Domain prediction

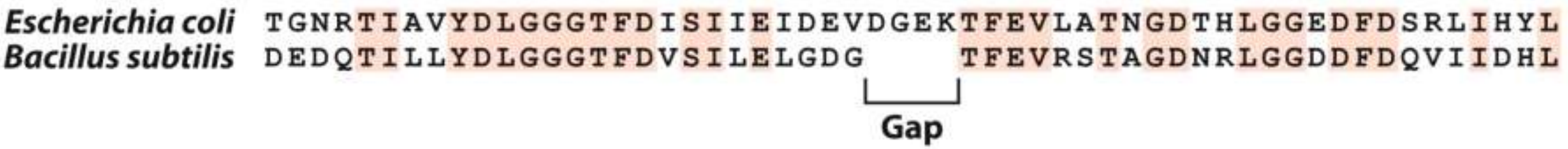
Signal peptide prediction

Function prediciton

....

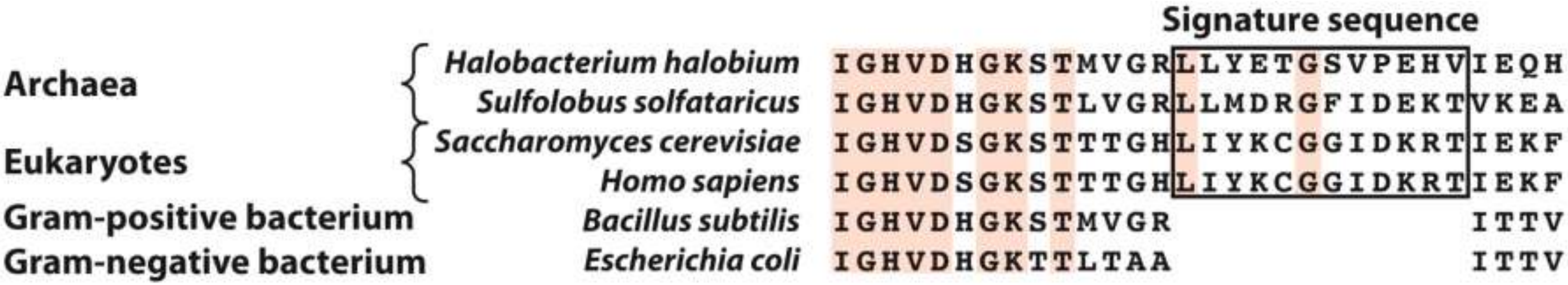
3.2 Peptides and Proteins

Alignment of protein sequences



```
Escherichia coli   TGNRTIAVYDLGGGTFDISIIEIDEVDGEKTFEVLATNGDTHLGGEDFDSRLIHYL
Bacillus subtilis  DEDQTILLYDLGGGTFDVSILELGDG    TFEVRSTAGDNRLGGDDFDQVIIDHL
```

Figure 3-33
Lehninger Principles of Biochemistry, Sixth Edition

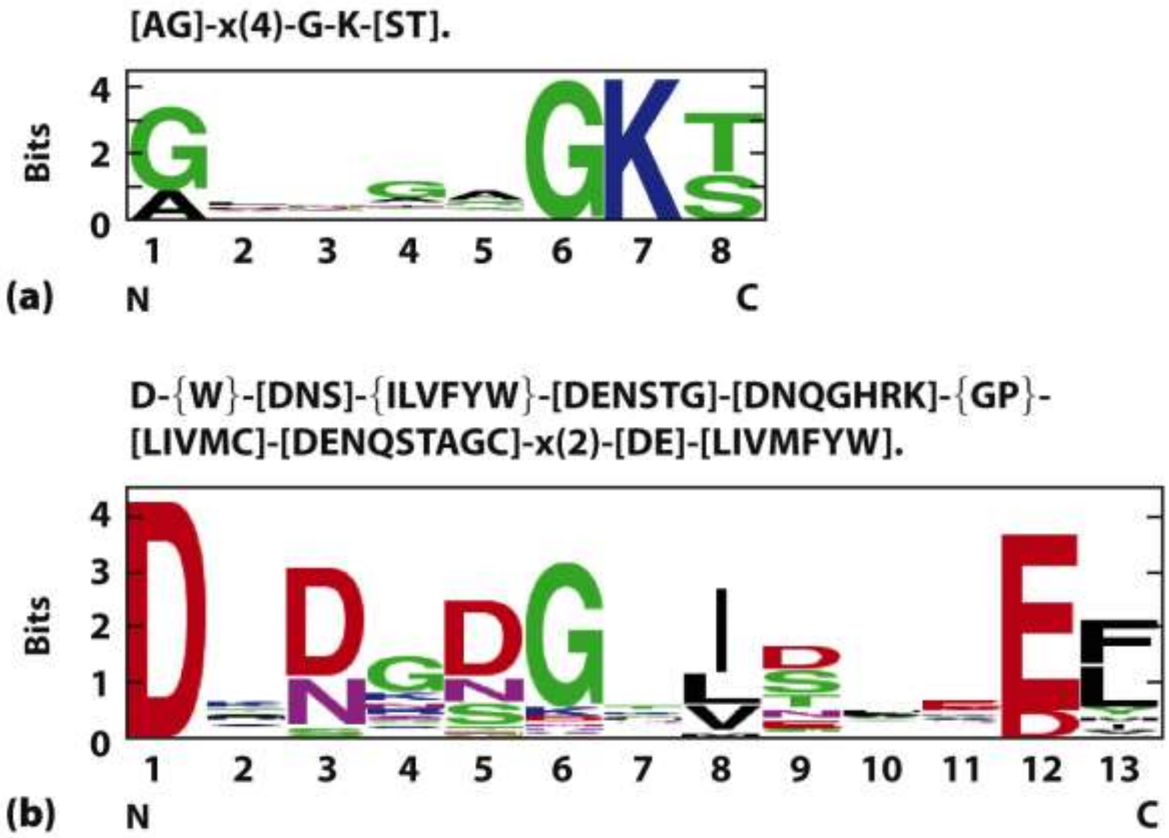


```
Archaea                      Halobacterium halobium    IGHVDHGKSTMVGRLLYETGSVPEHVIEQH
                             Sulfolobus solfataricus   IGHVDHGKSTLVGRLLMDRGFIDEKTVKEA
Eukaryotes                   Saccharomyces cerevisiae  IGHVDSGKSTTTGHLIYKCGGIDKRTIEKF
                             Homo sapiens              IGHVDSGKSTTTGHLIYKCGGIDKRTIEKF
Gram-positive bacterium      Bacillus subtilis         IGHVDHGKSTMVGR            ITTV
Gram-negative bacterium      Escherichia coli          IGHVDHGKTTLTAA            ITTV
```

Figure 3-34
Lehninger Principles of Biochemistry, Sixth Edition

3.2 Peptides and Proteins

Consensus sequence



[AG]-x(4)-G-K-[ST].

D-{W}-[DNS]-{ILVFYW}-[DENSTG]-[DNQGHRK]-{GP}-[LIVMC]-[DENQSTAGC]-x(2)-[DE]-[LIVMFYW].

Take home messages …

✓ **Amino Acids**

The 20 common amino acids, Chirality, Zwitterion of amino acids

✓ **Peptides and Proteins**

Peptide & Protein, Determining the amino acid sequence, Chemical synthesis of peptides