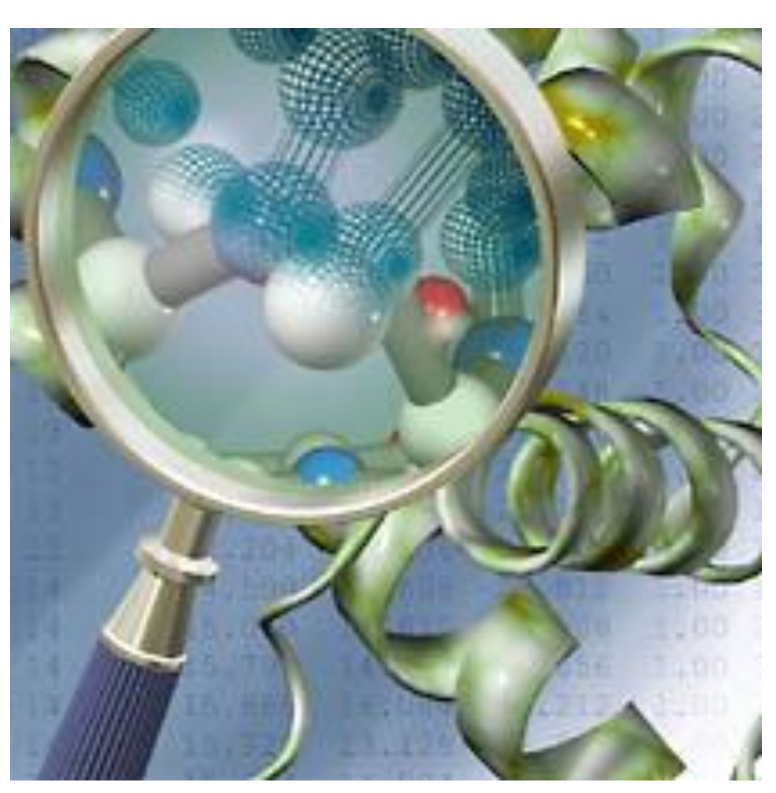
Biochemistry



2.1. Amino acids and peptides

Protein structure & function

Proteins serve crucial functions in essentially all biological processes

1 Proteins are built from a repertoire of 21 amino acids
2 **<u>Primary structure</u>**: amino acids linked by peptide bonds form polypeptide chains
3 Secondary structure: polypeptide chains fold into regular structures such as alpha helix, beta sheet, & turns & loops
4 Tertiary structure: water-soluble proteins fold into compact structures with nonpolar cores
5 Quaternary structure: polypeptide chains can assemble into multisubunit structures
6 The amino acid sequence of a protein determines its three-dimensional structure

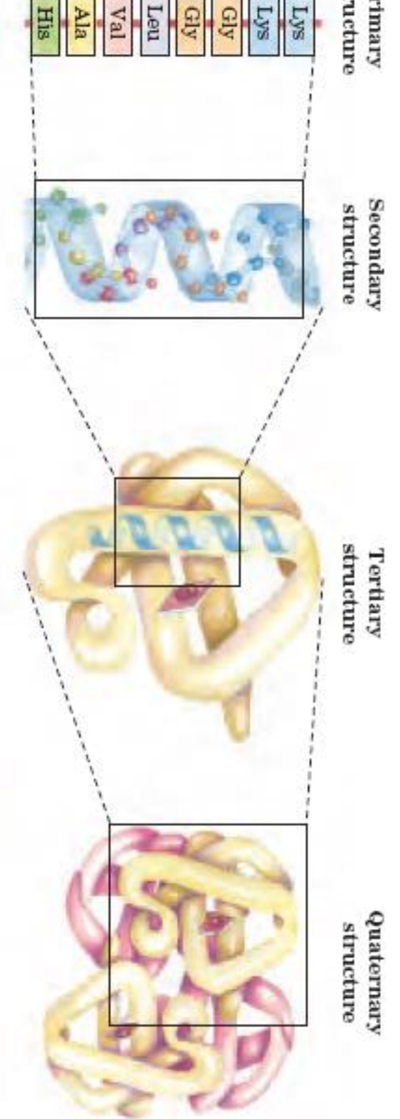


FIGURE 3–16 Levels of structure in proteins. The *primary structure* consists of a sequence of amino acids linked together by peptide bonds and includes any disulfide bonds. The resulting polypeptide can be coiled into units of *secondary structure*, such as an α helix. The helix is a part of the *tertiary structure* of the folded polypeptide, which is itself one of the subunits that make up the *quaternary structure* of the multisubunit protein, in this case hemoglobin.

Proteins - Key properties - a wide range of functions

1. Proteins are linear polymers built of monomer units called amino acids - spontaneously fold into 3-dimensional structures

2. Proteins contain a wide range of functional groups - alcohols, thiols, thioethers, carboxylic acids, carboxamides, & a variety of basic groups - eg. chemical reactivity essential to function of *enzymes*

3. Proteins can interact with one another, & with other biological macromolecules to form complex assemblies - macromolecular machines

4. Some proteins are quite rigid, whereas others display limited flexibility - structural elements in the cytoskeleton *v* parts that act as hinges, springs, & levers *etc*

Amino acids

Proteins: Essential for all organisms

- AA
- Peptide
- Polypeptide
- Proteins more 50 AA

- **Proteins** are polymers of amino acids, with each amino acid residue joined to its neighbor by a specific type of covalent bond.
- L-α-amino acids and their derivatives participate in cellular functions as diverse as nerve transmission and the biosynthesis of porphyrins, purines, pyrimidines, and urea.

Amino acids and Proteins

2.1) Amino acids

- **Amino acid classes**
- **Modified AA in proteins**
- **AA stereo isomers**
- **Titration of AA**
- **AA reactions**

2.2) Peptides

2.3) Protein structure

- **Protein structure**
- **Fibrous proteins**
- **Globular proteins**

AA functions (300 AA)

- Primary function (components of proteins)
- Chemical messenger:
  - **Neurotransmiters** (substance released from one nerve cell
  - that influence function second nerve cell)
  - GABA (g-aminobutyric acids), glycin, serotonin (tryptophan)
  - **Hormones** (chemical messanger ….produced by one type cell and regulate function of other type of cell)
  - Thyroxine (tyrosin)
  - Indole acetic acid (plant)

- Precursores for nitrogen containing molecules
  - Nucleotides, heme, chlorophyl

- Metabolic intermediates: arginine, citruline, ornithine – urea cycle

GABA

Serotonin

Thyroxine

Indole acetic acid

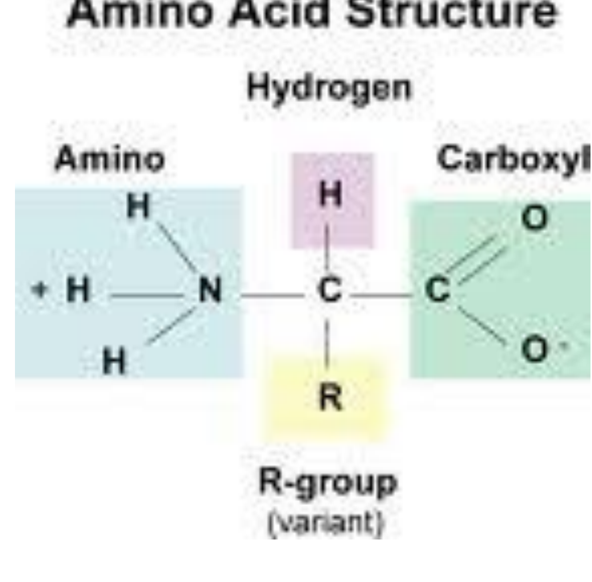
Amino acids

300 AA

20 AA (proteins, gene code…)

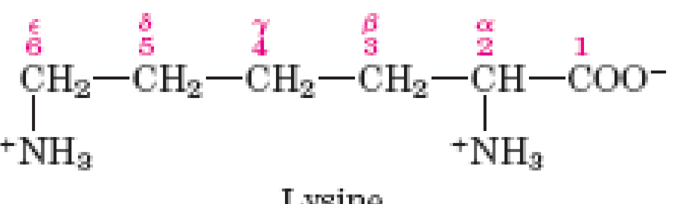
- **Standard AA** (in proteins 21)
- **Nonstandard AA**

(modified after incorporation to polypeptide)



General structure

L- α– amino acids



NH OH O

proline

Alpha-imidoacids

NON-POLAR

Glycine (Gly / G)

Alanine (Ala / A)

Valine (Val / V)

Cysteine (Cys / C)

Proline (Pro / P)

Leucine (Leu / L)

Isoleucine (Ile / I)

Methionine (Met / M)

Tryptophan (Trp / W)

Phenylalanine (Phe / F)

+ CHARGE

Lysine (Lys / K)

Arginine (Arg / R)

Histidine (His / H)

POLAR

Serine (Ser / S)

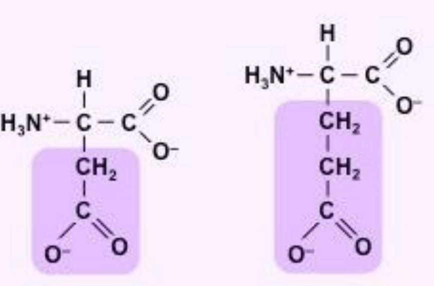
Threonine (Thr / T)

Tyrosine (Tyr / Y)

Asparagine (Asn / N)

Glutamine (Gln / Q)

− CHARGE

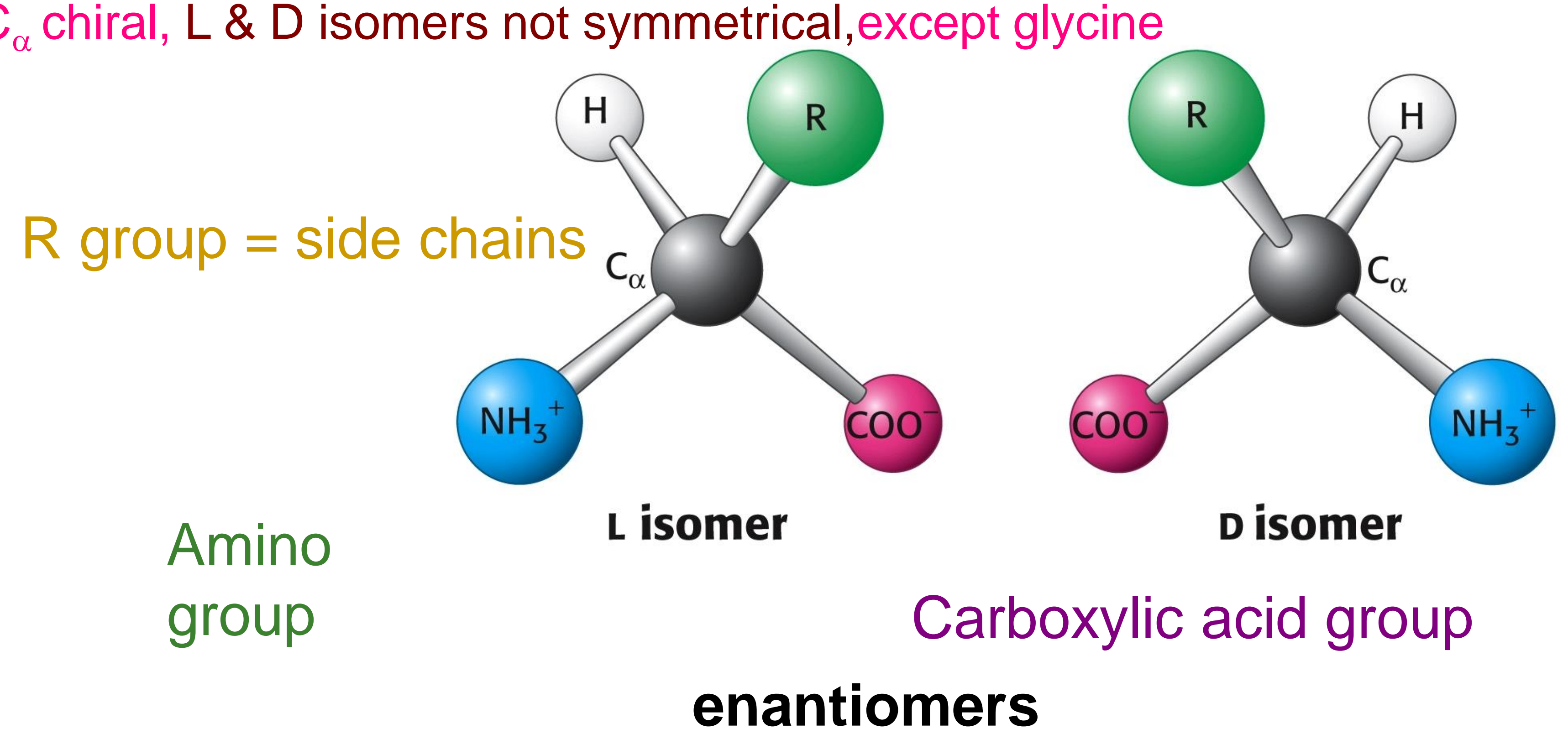


Amino acid stereoisomers

Protein subunits: αamino acids: L & D isomers

Only L amino acids found in proteins.

C_α chiral, L & D isomers not symmetrical,except glycine



R group = side chains

Amino group

Carboxylic acid group

enantiomers

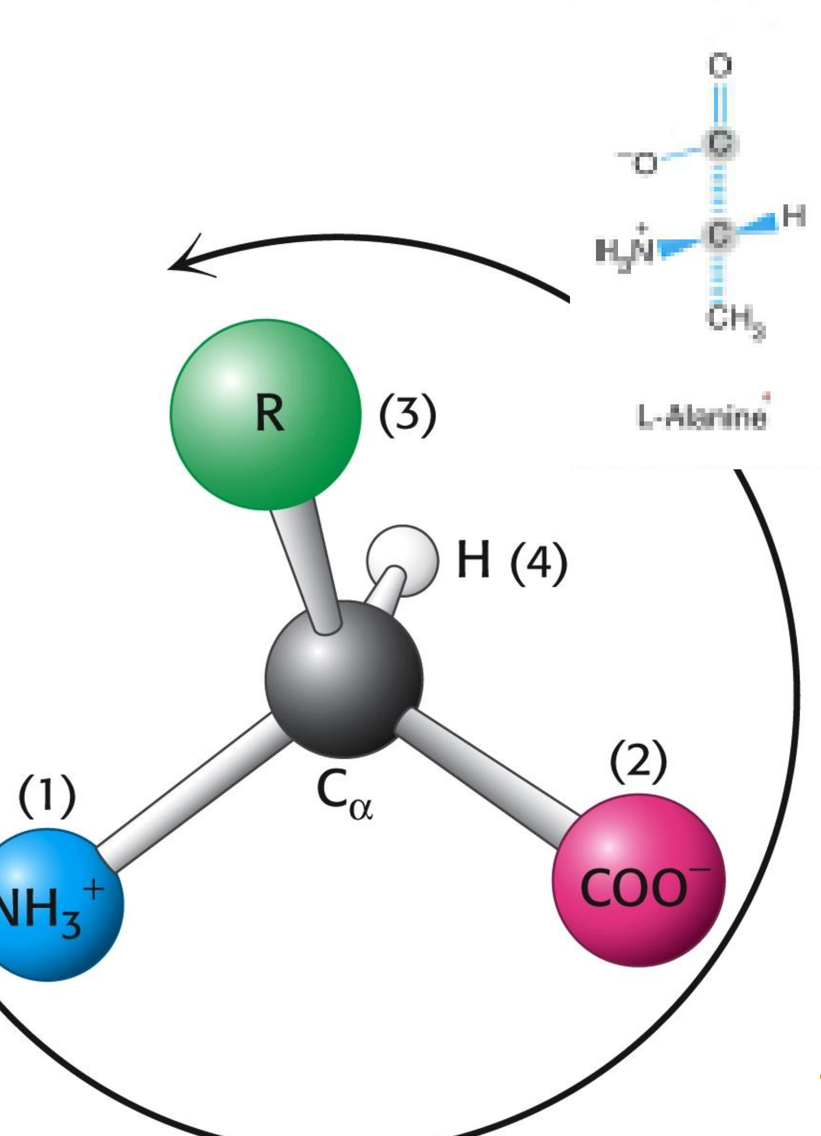
Mirror images of each other

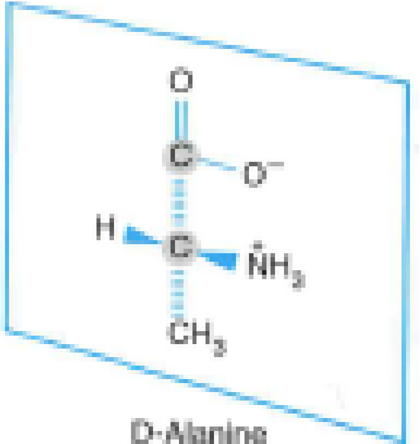
Only L amino acids in proteins

Tetrahedral
α-carbon atom,

C_α **chiral center** is
S configuration.
(sinister for left)
Counterclockwise
arrow indicates
chiral center
is S configuration

Amino group:
highest priority
substituent
(according to
atomic #)





Amino Acids May Have Positive, Negative, or Zero Net Charge - Titration of Amino acids

- **-COOH and -NH3 weak acid groups exist in solution in protonic equilibrium:**

$$R-COOH = R-COO^- + H^+$$

$$R-NH_3^+ = R-NH_2 + H^+$$

as either an acid (proton donor):

$$R-C(H)(^+NH_3)-COO^- \rightleftharpoons R-C(H)(NH_2)-COO^- + H^+$$

Zwitterion

or a base (proton acceptor):

$$R-C(H)(^+NH_3)-COO^- + H^+ \rightleftharpoons R-C(H)(^+NH_3)-COOH$$

Zwitterion

amphoteric

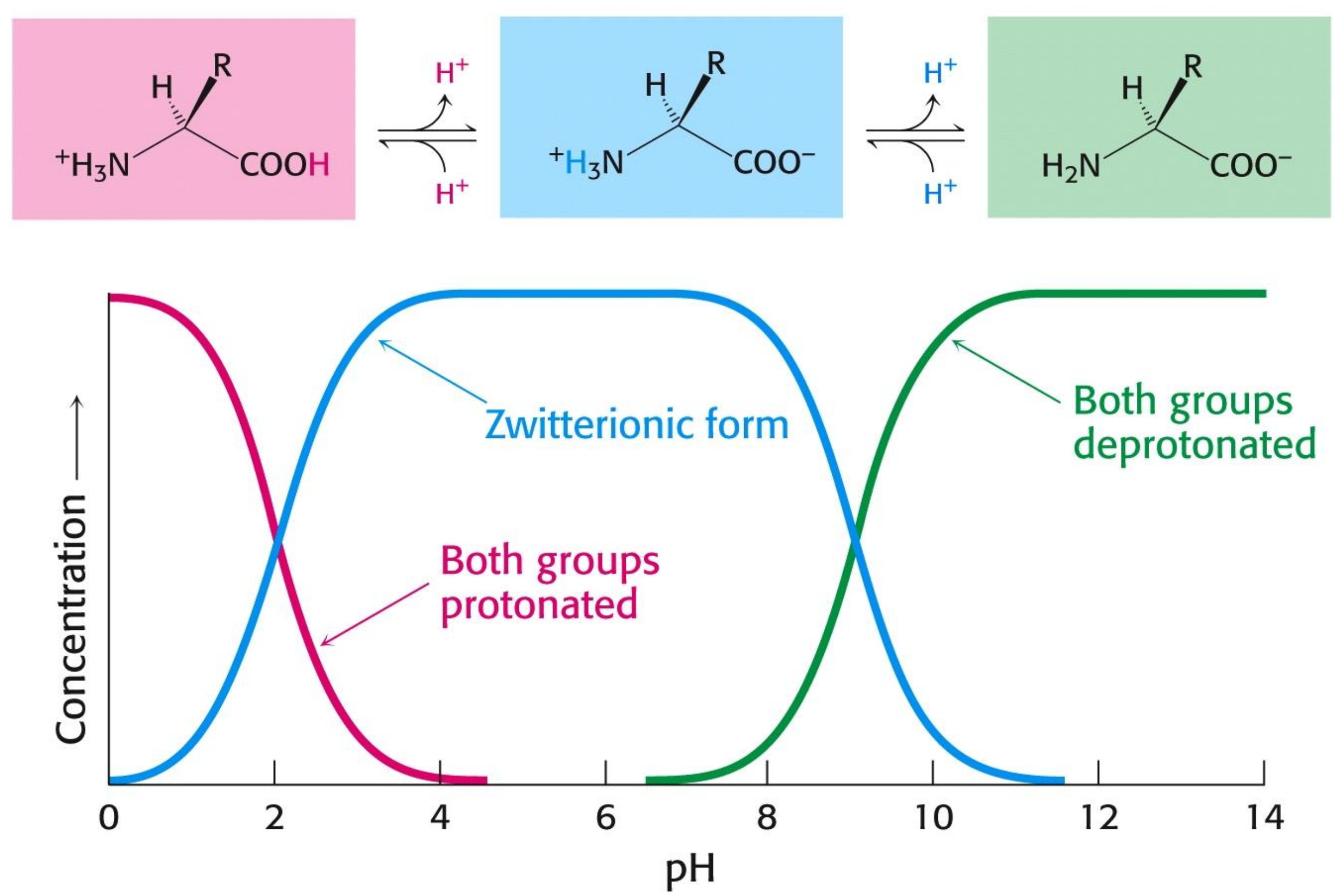
zwitterions (neutral, physiological pH)

1: $H_2N - {}^{\alpha}C(H)(R) - C(=O)OH$

2: $H_3N^+ - {}^{\alpha}C(H)(R) - C(=O)O^{\ominus}$

- Full disociation COOH and NH2 groups at physiological pH

Ionization state as a function of pH



Physiological pH (measure of $[H^+]$)

pK_a express the strengths of weak acids

TABLE 3.4 **pK_a values of some amino acids**

Amino acid	pK_a values (25°C)		
	α-COOH group	α-NH_3^+ group	Side chain
Alanine	2.3	9.9	
Glycine	2.4	9.8	
Phenylalanine	1.8	9.1	
Serine	2.1	9.2	
Valine	2.3	9.6	
Aspartic acid	2.0	10.0	3.9
Glutamic acid	2.2	9.7	4.3
Histidine	1.8	9.2	6.0
Cysteine	1.8	10.8	8.3
Tyrosine	2.2	9.1	10.9
Lysine	2.2	9.2	10.8
Arginine	1.8	9.0	12.5

After J. T. Edsall and J. Wyman, *Biophysical Chemistry* (Academic Press, 1958), Chapter 8.

pK_a of ionizable side chains

TABLE 3.1 Typical pK_a values of ionizable groups in proteins

Group	Acid ⇌ Base	Typical pK_a*
Terminal α-carboxyl group	$-COOH \rightleftharpoons -COO^-$	3.1
Aspartic acid Glutamic acid	$-COOH \rightleftharpoons -COO^-$	4.1
Histidine	imidazolium (N–H, N⁺–H) ⇌ imidazole (N, N–H)	6.0
Terminal α-amino group	$-NH_3^+ \rightleftharpoons -NH_2$	8.0
Cysteine	$-SH \rightleftharpoons -S^-$	8.3
Tyrosine	$-C_6H_4-OH \rightleftharpoons -C_6H_4-O^-$	10.9
Lysine	$-NH_3^+ \rightleftharpoons -NH_2$	10.8
Arginine	$-NH-C(NH_2)_2^+ \rightleftharpoons -NH-C(=NH)NH_2$	12.5

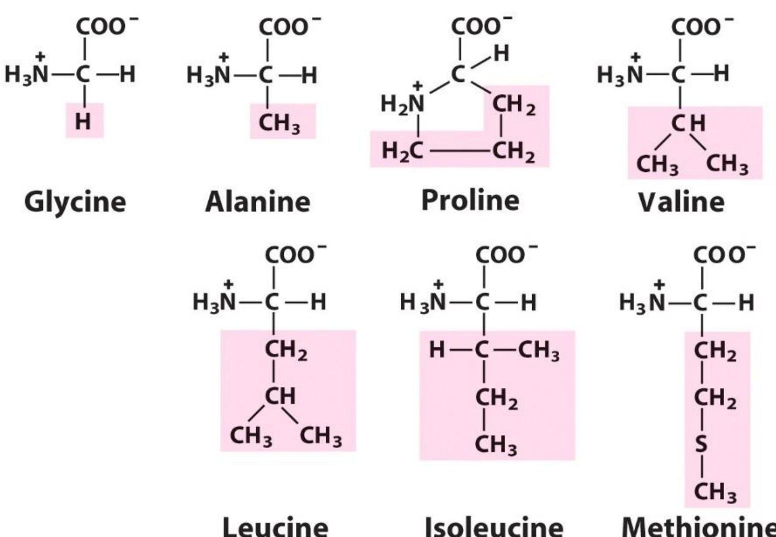
*pK_a values depend on temperature, ionic strength, and the microenvironment of the ionizable group.

pK_a = pH for 50% dissociation,

Note range

The 20 Amino Acids Found in Proteins

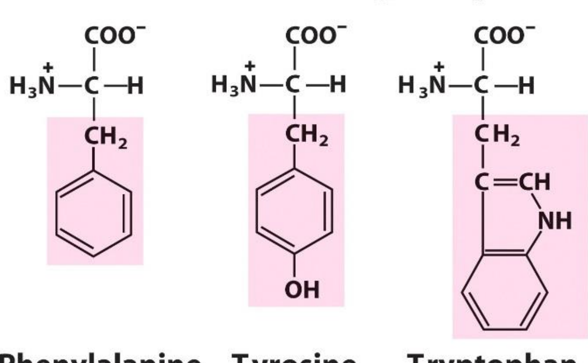
Nonpolar, aliphatic R groups



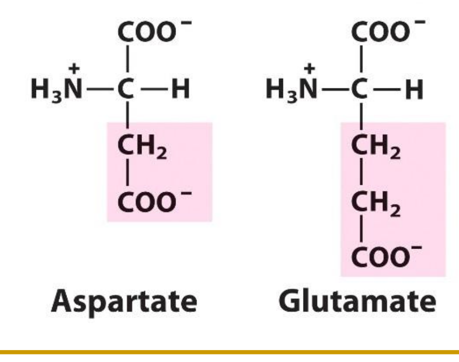
Polar, uncharged R groups

Serine
Threonine
Cysteine
Asparagine
Glutamine

Aromatic R groups



Negatively charged R groups



Positively charged R groups

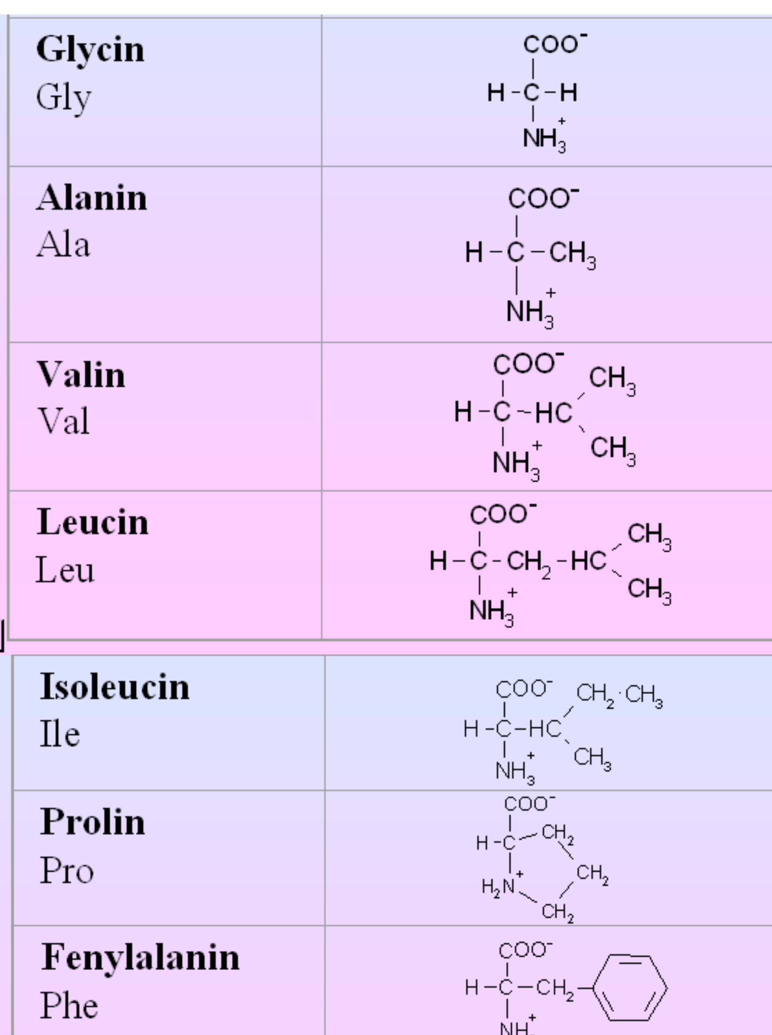
Lysine
Arginine
Histidine

Classification of AA

- Neutral nonpolar
- Neutral polar
- Acidic
- Basic

Neutral nonpolar

- Interact poorly with water
- Hydrophobic
- 3D structure of proteins
- Aliphatic and aromatic



Name	Structure
Glycin Gly	COO^-, H-C-H, NH_3^+
Alanin Ala	COO^-, H-C-CH_3, NH_3^+
Valin Val	COO^-, H-C-HC, CH_3, CH_3, NH_3^+
Leucin Leu	COO^-, H-C-CH_2-HC, CH_3, CH_3, NH_3^+
Isoleucin Ile	COO^-, H-C-HC, CH_2-CH_3, CH_3, NH_3^+
Prolin Pro	COO^-, H-C, CH_2, CH_2, CH_2, H_2N^+
Fenylalanin Phe	COO^-, H-C-CH_2-(benzene ring), NH_3^+

Name	Structure
Methionin Met	COO^-, H-C-CH_2-CH_2-S-CH_3, NH_3^+
Tryptofan Trp	COO^-, H-C-CH_2-(indole ring, N, H), NH_3^+

Neutral polar

Name	Structure	
Serin Ser	$H-C(COO^-)(NH_3^+)-CH_2-OH$	P
Threonin Thr	$H-C(COO^-)(NH_3^+)-CH(CH_3)(OH)$	P
Cystein Cys	$H-C(COO^-)(NH_3^+)-CH_2-SH$	P
Tyrosin Tyr	$H-C(COO^-)(NH_3^+)-CH_2-C_6H_4-OH$	P

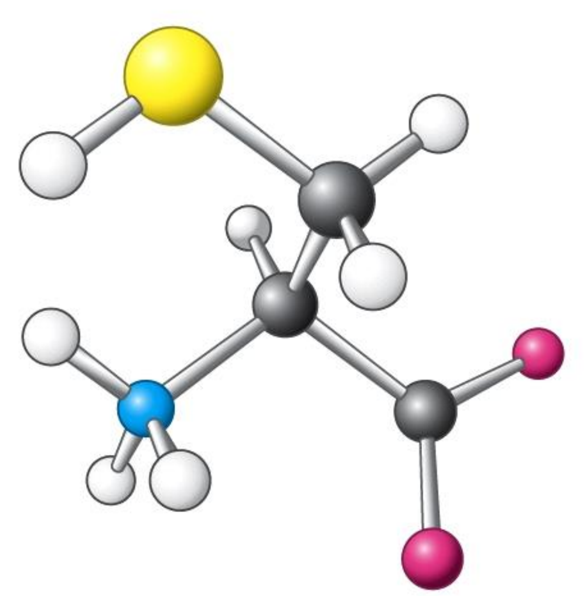
Name	Structure	
Asparagin Asn	$H-C(COO^-)(NH_3^+)-CH_2-CONH_2$	P
Glutamin Gln	$H-C(COO^-)(NH_3^+)-CH_2-CH_2-CONH_2$	P

Amide group is highly polar, affect protein stability….

Cysteine

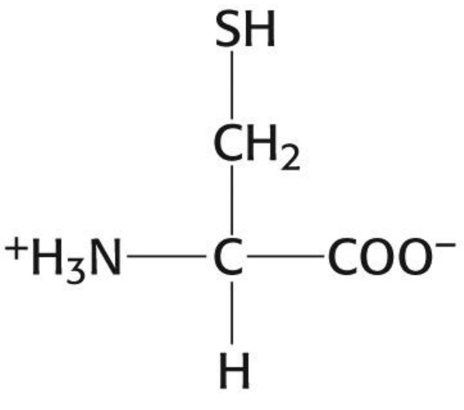
Similar to Serine with sulfhydryl, or thiol (-SH) group replacing hydroxyl (-OH) group

-SH more reactive than -OH. -SH pairs form disulfide bonds (aka bridges), key role stabilizing proteins



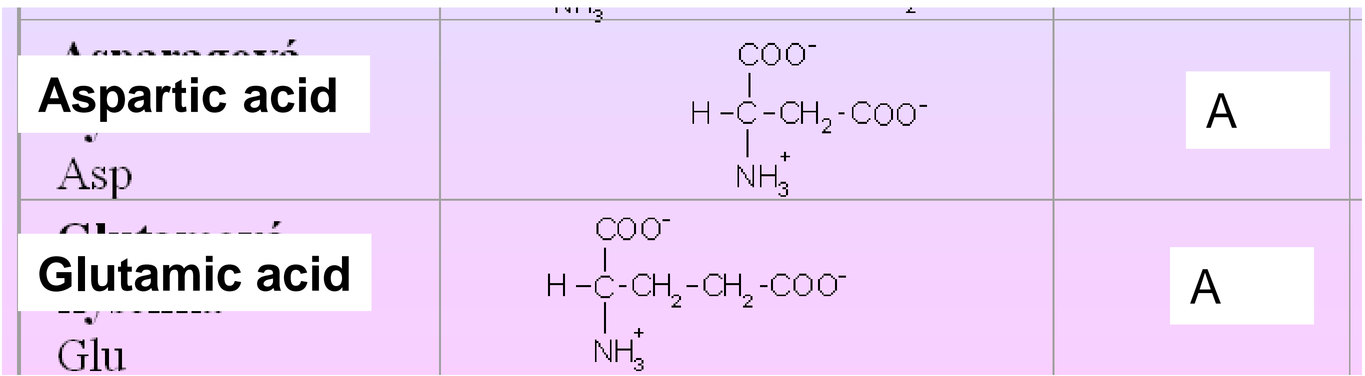
H
S
CH_2
H
$^{+}H_3N$
COO^-

Cysteine (Cys, C)



Acidic AA

- Negative charge at physiological pH



Basic AA

- Positive charge at physiological pH
- Ionic bonds with amino acids
- Arginine - strong base, no function in acide/base reaction
- Lysin – ammonium ion, oxidation of lysines side chain – collagen, linkage
- Histidine – weak base, ony partial disociation at ph 7, react as buffer,
- Catalytic activity of enzymes

Name	Structure	
Histidin His	$H-C(COO^-)(NH_3^+)-CH_2-C$ (imidazole ring: $C=CH-NH^+=CH-NH$)	B
Lysin Lys	$H-C(COO^-)(NH_3^+)-CH_2-CH_2-CH_2-CH_2-NH_3^+$	B
Arginin Arg	$H-C(COO^-)(NH_3^+)-CH_2-CH_2-CH_2-HN-C(=NH_2^+)-NH_2$	B

Aromatic side chains

Aromatic rings have delocalized π electrons,
Absorb UV light

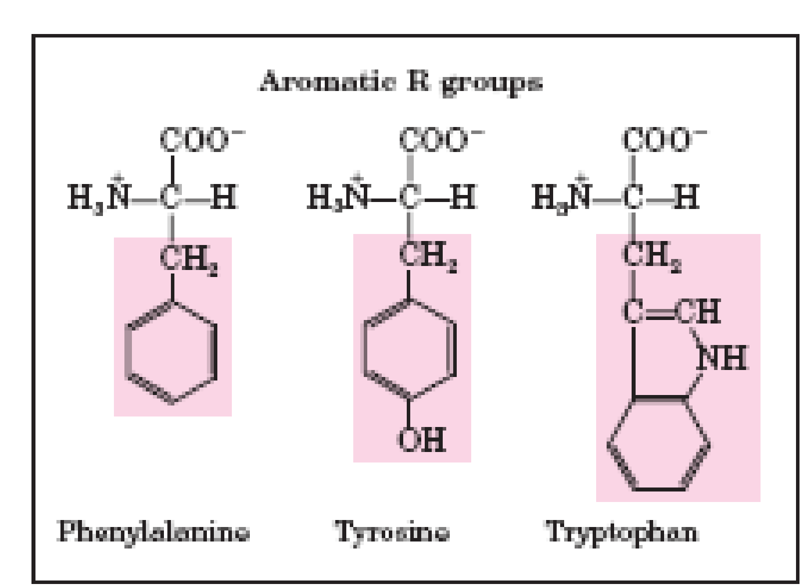
Phenylalanine (Phe, F)

Tyrosine (Tyr, Y)

Tryptophan (Trp, W)

Hydrophobic & Hydrophilic properties

Optical properties of AA and proteins



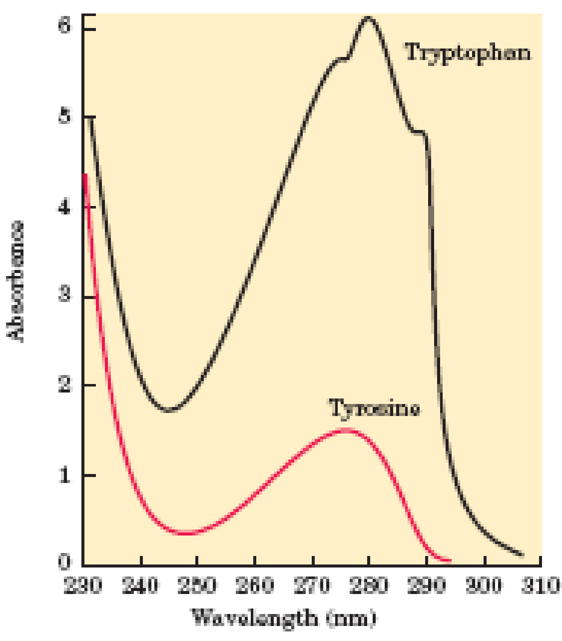
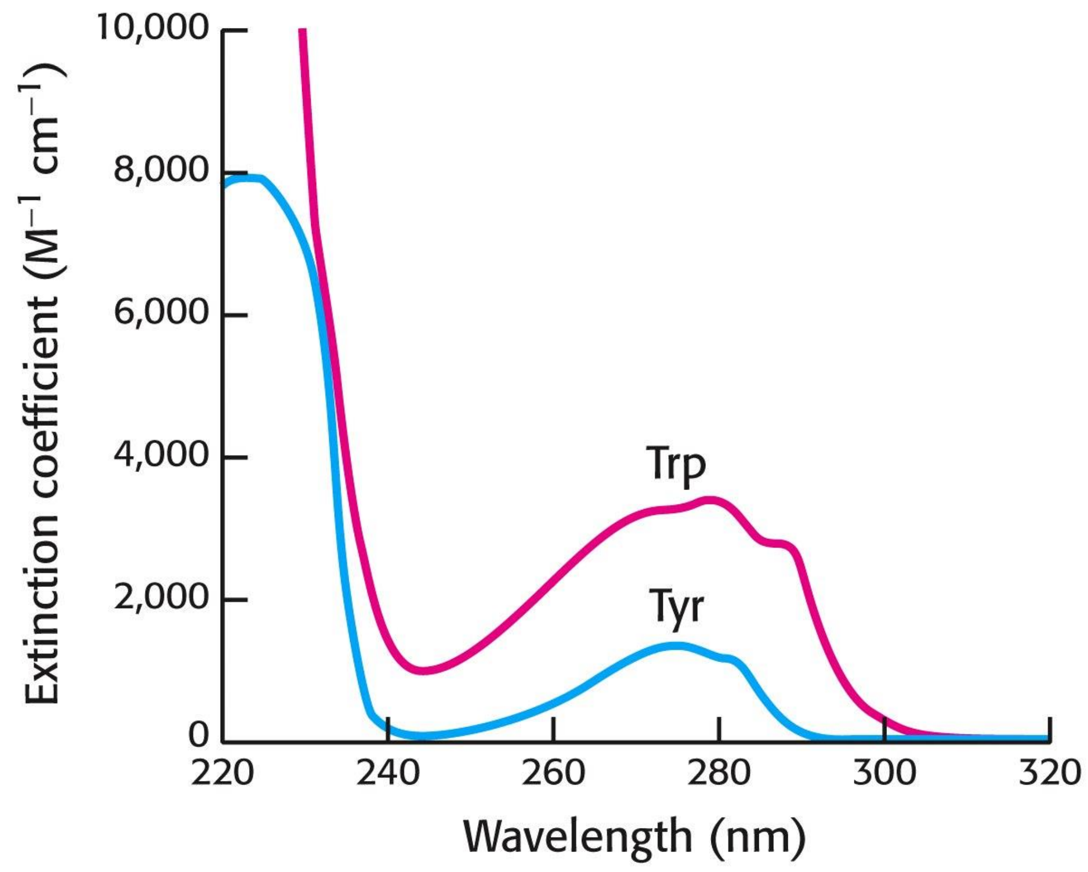


FIGURE 3–6 Absorption of ultraviolet light by aromatic amino acids. Comparison of the light absorption spectra of the aromatic amino acids tryptophan and tyrosine at pH 6.0. The amino acids are present in equimolar amounts (10^{-3} M) under identical conditions. The measured absorbance of tryptophan is as much as four times that of tyrosine. Note that the maximum light absorption for both tryptophan and tyrosine occurs near a wavelength of 280 nm. Light absorption by the third aromatic amino acid, phenylalanine (not shown), generally contributes little to the spectroscopic properties of proteins.

TEST

Absorption spectra of Trp & Tyr

Beer's law: A = εcl. Used to estimate protein concentration



TEST

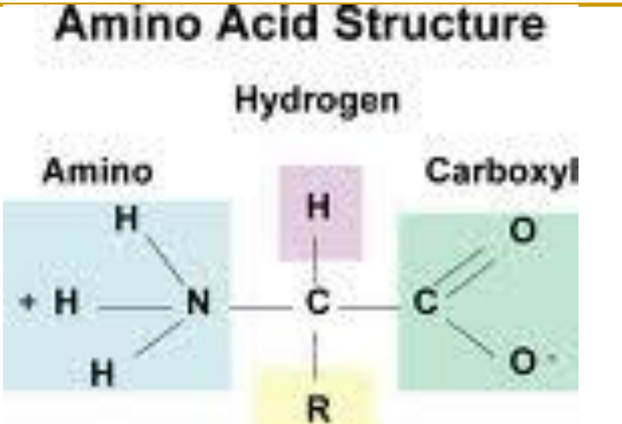
TABLE 18–1 Nonessential and Essential Amino Acids for Humans and the Albino Rat

Nonessential	*Conditionally essential**	*Essential*
Alanine	Arginine	Histidine
Asparagine	Cysteine	Isoleucine
Aspartate	Glutamine	Leucine
Glutamate	Glycine	Lysine
Serine	Proline	Methionine
	Tyrosine	Phenylalanine
		Threonine
		Tryptophan
		Valine

*Required to some degree in young, growing animals, and/or sometimes during illness.

Amino acid reactions

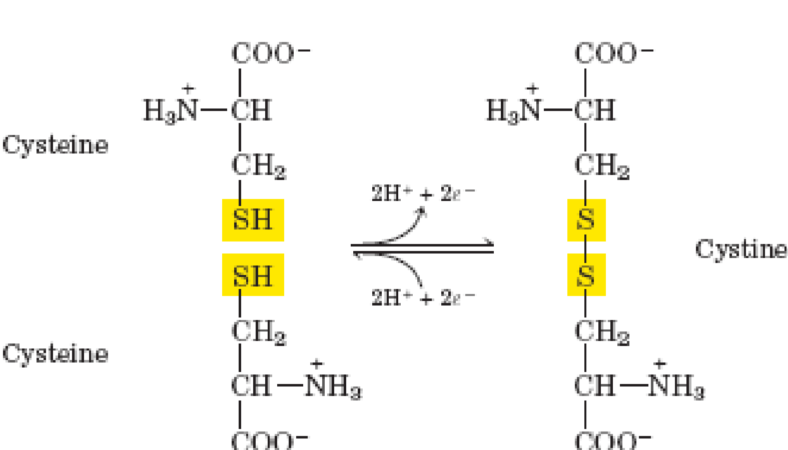
- peptide bond
- disulfide bridge



$$H_3\overset{+}{N}-CH(R^1)-C(=O)-OH + H-N(H)-CH(R^2)-COO^-$$

$$\rightleftharpoons \; (H_2O)$$

$$H_3\overset{+}{N}-CH(R^1)-C(=O)-N(H)-CH(R^2)-COO^-$$



Primary structure: Peptide bond, between AAs

Between α-carboxyl group of one AA & α-amino group of another

$$^{+}H_3N-CH(R_1)-COO^- + {}^{+}H_3N-CH(R_2)-COO^- \rightleftharpoons {}^{+}H_3N-CH(R_1)-CO-NH-CH(R_2)-COO^- + H_2O$$

Peptide bond

2 amino acids

Dipeptide

Loss of H_2O

Formation of a peptide bond by condensation. The amino group of one amino acid (with R2 group) acts as a **nucleophile** to displace the hydroxyl group of another amino acid (with R1 group),forming a peptide bond (shaded in yellow). Amino groups are good nucleophiles, but the hydroxyl group is a poor leaving group and is not readily displaced. At physiological pH, the reaction shown does not occur to any appreciable extent.

Equilibrium favors hydrolysis, hence,
biosynthesis of peptide bonds require free energy input

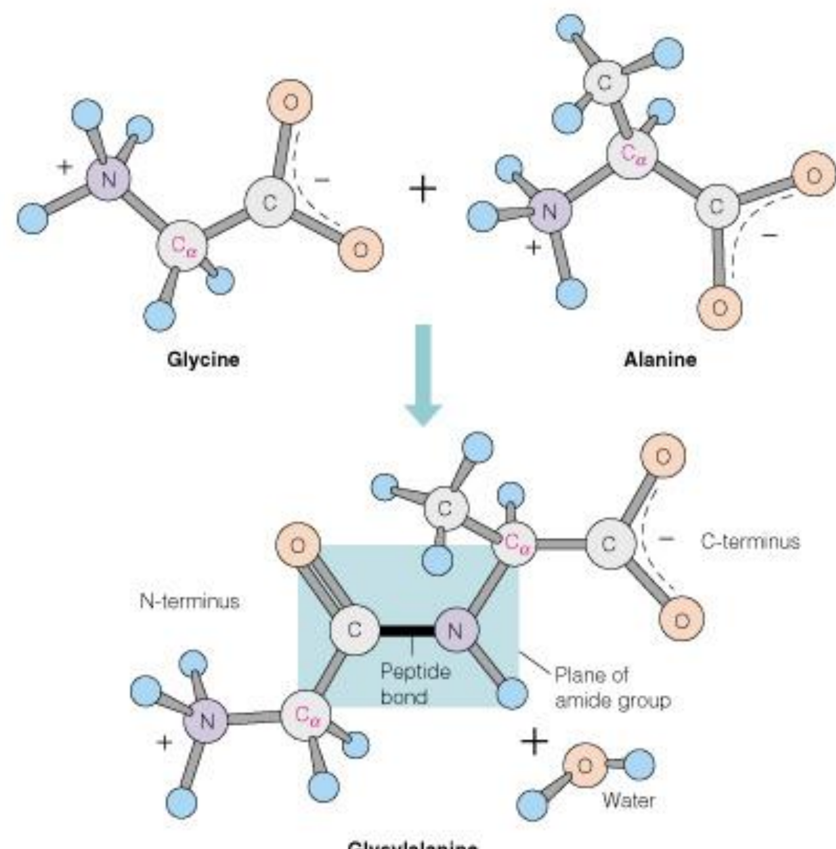
Peptide bonds are stable kinetically

Formation of a Peptide

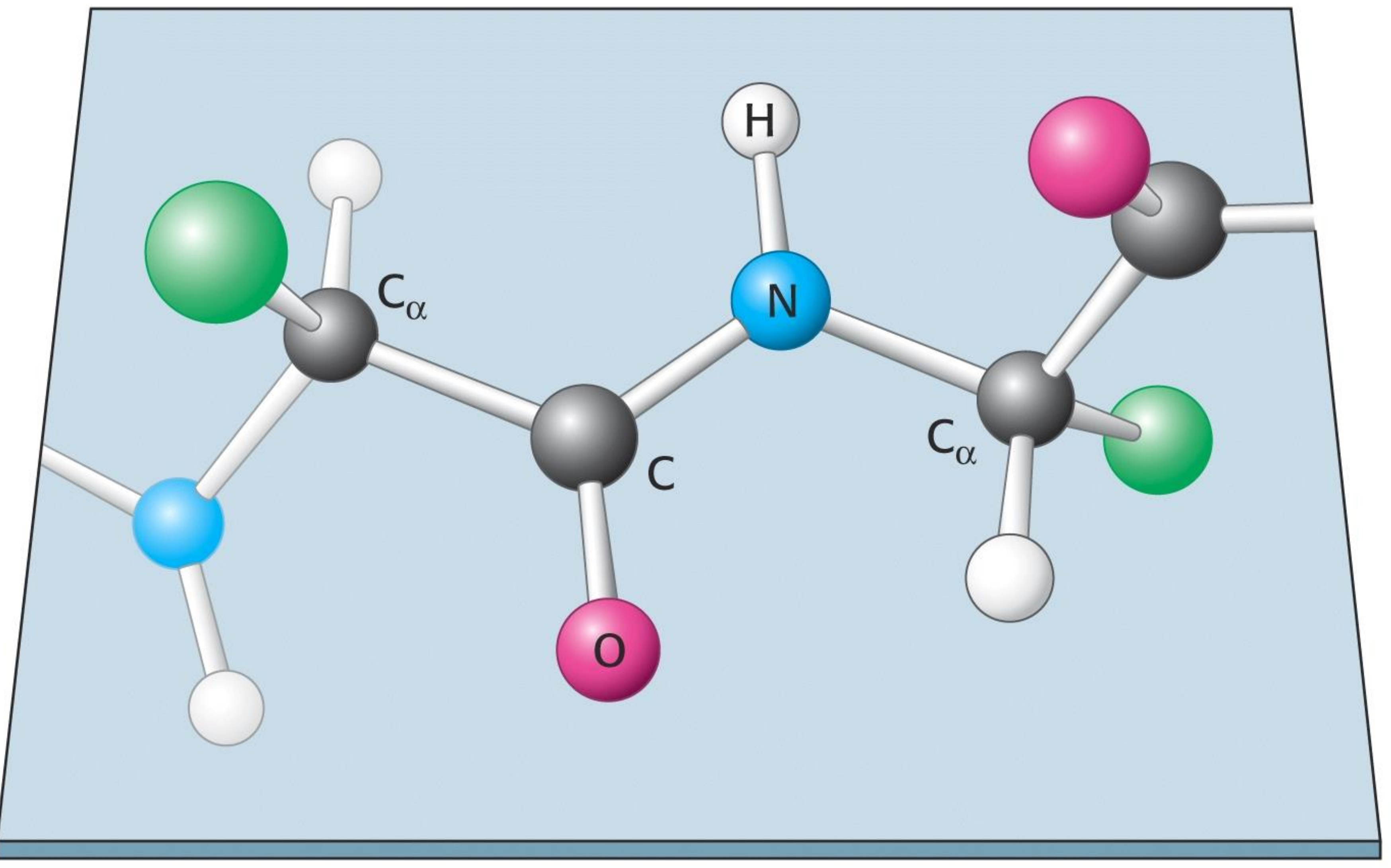
$$H_3\overset{+}{N}-CH(R^1)-C(=O)-OH + H-NH-CH(R^2)-COO^-$$

$$\rightleftharpoons \quad (H_2O \text{ in}, \ H_2O \text{ out})$$

$$H_3\overset{+}{N}-CH(R^1)-C(=O)-NH-CH(R^2)-COO^-$$

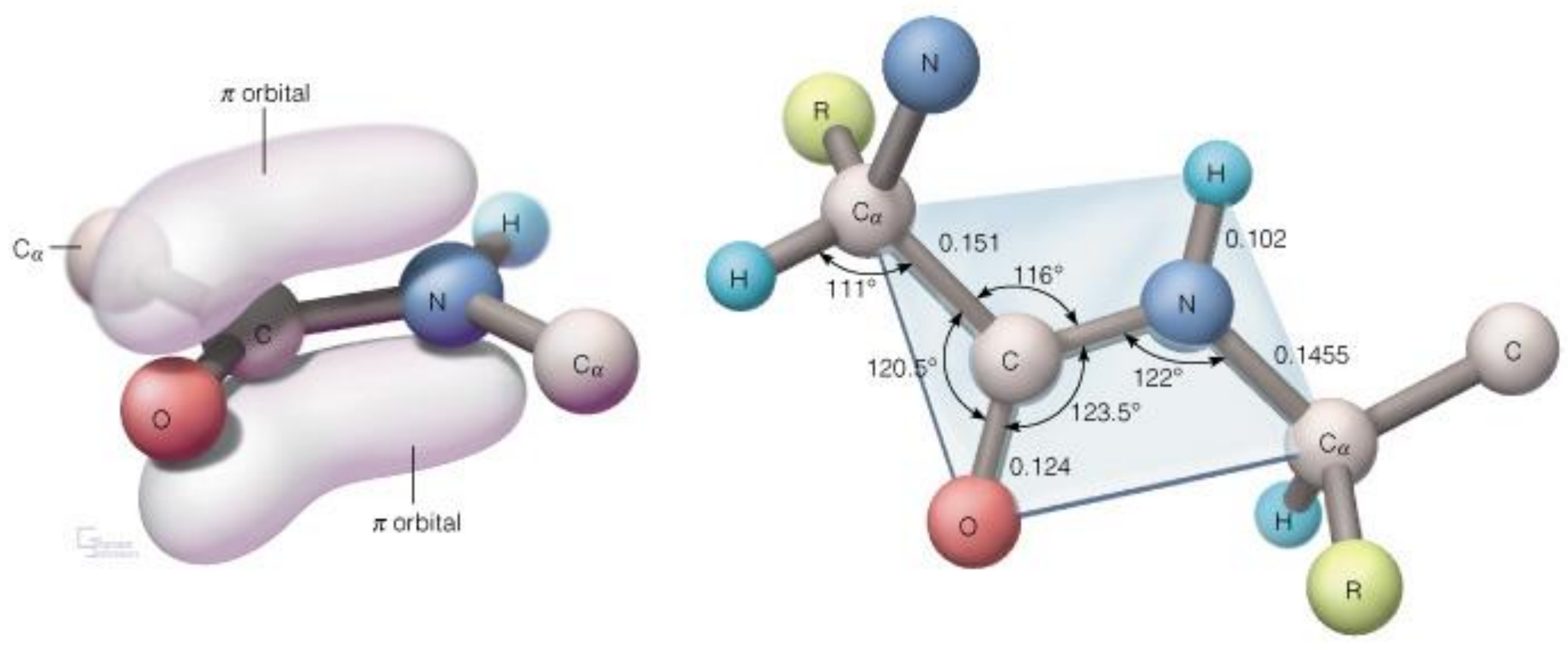


Polypeptide bonds are planar



Six atoms (C_a, C, O, N, H, C_a) lie in a plane, in a pair of aa

Planarity of Peptide (Amide) Bond



(a) Partial double-bond character of peptide bond

(b) Bond angles and lengths

Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

The carbonyl oxygen has a partial negative charge and the amide nitrogen a partial positive charge, setting up a small electric dipole. Virtually all peptide bonds in proteins occur in this trans configuration .

Main chain or backbone

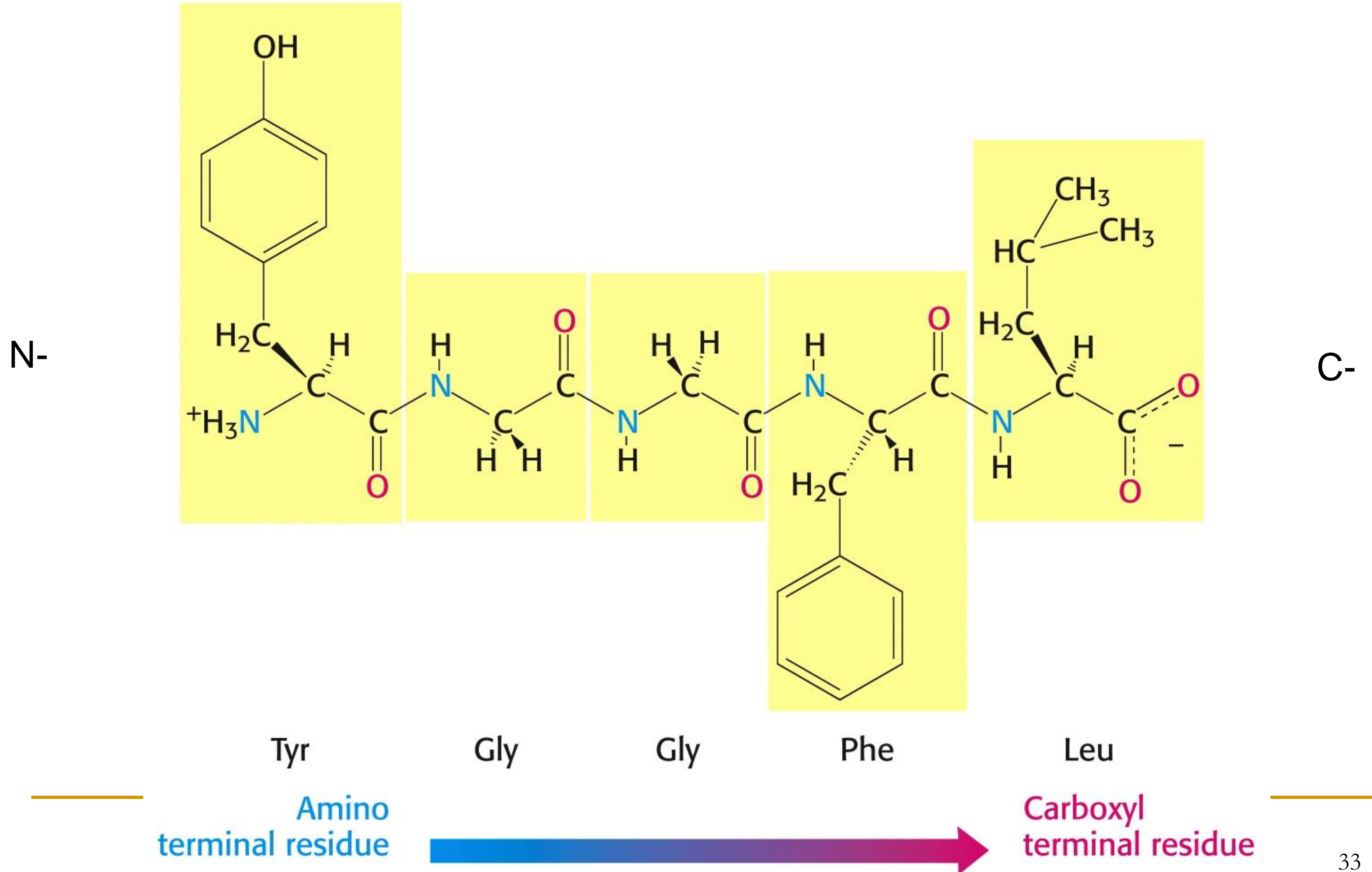
Constant backbone: regularly repeating part

Distinctive side chains (R-groups): variable part

AA unit in a polypeptide is called a residue, which contains,
a carbonyl group; good hydrogen-bond acceptor,
an NH group (except Pro); good hydrogen-bond donor

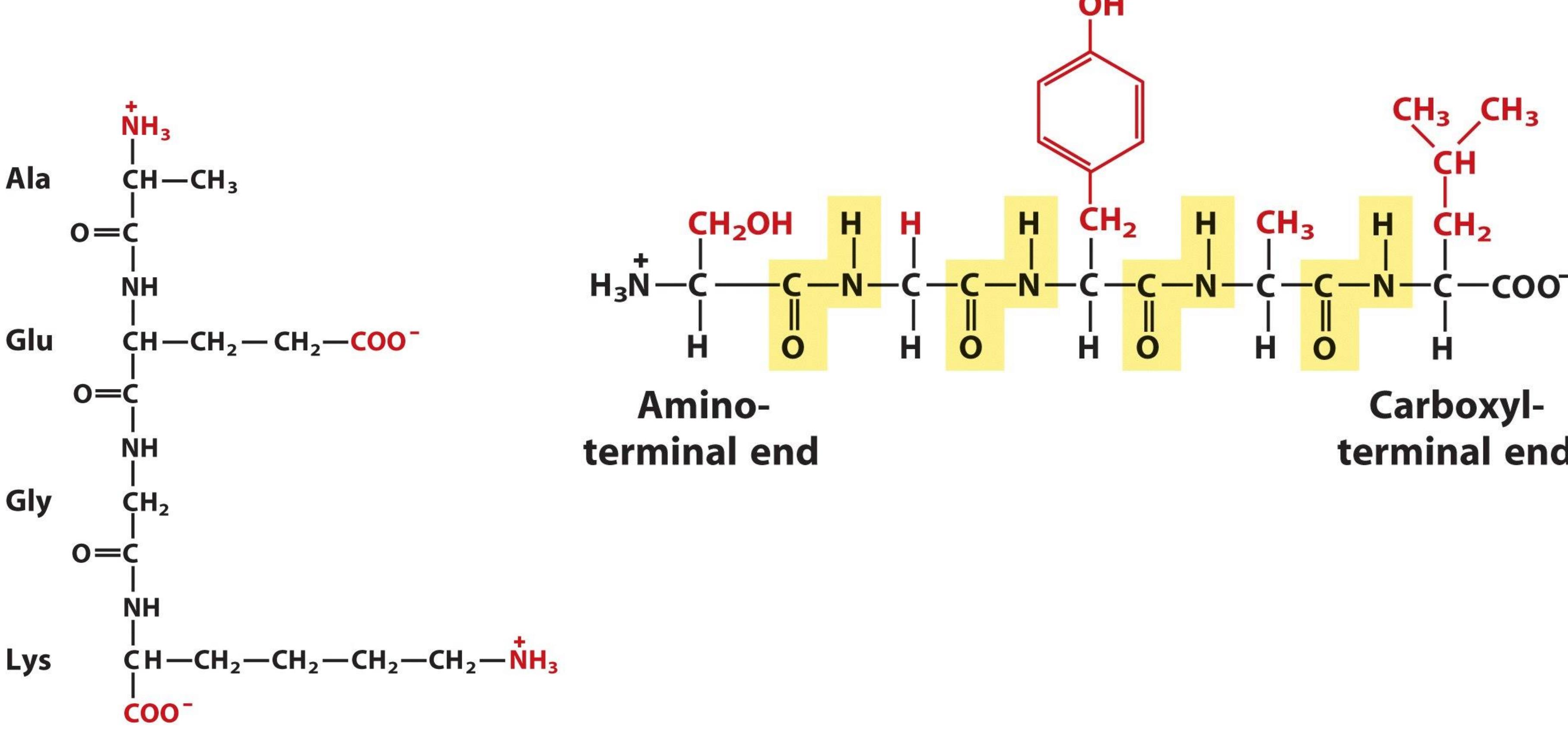
Polypeptide chain has direction

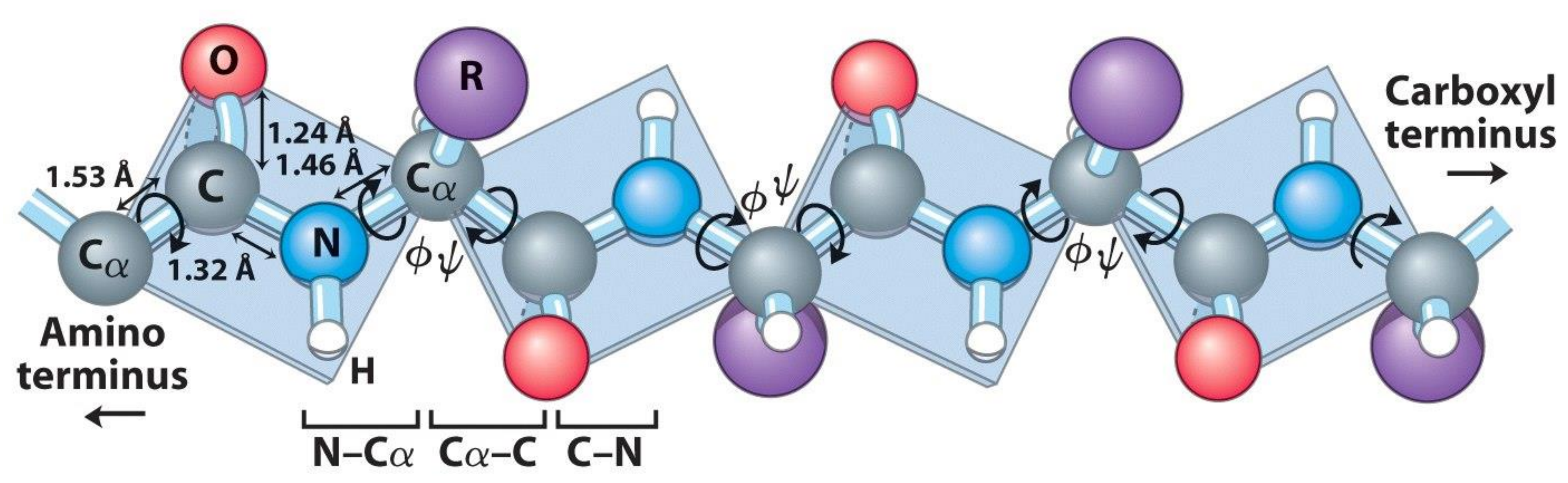
N-



C-

Examples of Oligopeptides




O
R
1.24 Å
1.46 Å
1.53 Å
C
Cα
N
1.32 Å
ϕ ψ
ϕ ψ
ϕ ψ
H
Carboxyl terminus
Amino terminus
N–Cα
Cα–C
C–N

N- and C-Termini May Be Modified in Proteins

N-Formyl group

N-Acetyl group

C-terminal amide

Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Cysteine oxidation

-Highly reactive SH- (sulphydryl group of Cysteine)

- reversible oxidation – to form disulfide

-CYSTINE (disulfide bond) –

-**DISULFIDE BRIDGE**

-- singe chain, 2 separate chains,

-In proteins: stability

Cysteine

Cysteine

$2H^+ + 2e^-$

$2H^+ + 2e^-$

Cystine

FIGURE 3–7 Reversible formation of a disulfide bond by the oxidation of two molecules of cysteine. Disulfide bonds between Cys residues stabilize the structures of many proteins.

Cross links (disulfide bridges)

Prevalent mainly in extracellular proteins

Cysteine + Cysteine $\underset{\text{Reduction}}{\overset{\text{Oxidation}}{\rightleftharpoons}}$ Cystine $+ \; 2\,H^+ + 2\,e^-$

TEST

Glutathione –

- (γ-glutamyl-L-cysteinnylglycine)

(γ-amide bond, γ-carboxy group contributes on peptide bond)

Function:

- Protein and DNA synthesis,
drug and environmental toxin metabolism,
amino acid transport

- Reducing agent
- Protects cells from destructive effects oxidation
- GSH/GSSG is high –normally present in cells
- Important intracellular reducing agent
- ***Glutathione peroxidase***

HO, O, NH₂, O, N H, SH, H N, O, OH

$$2\ GSH + H_2O_2 \rightarrow GSSG + 2\ H_2O$$

(methemoglobin)

Peptide hormones

•**Vertebrate hormones: m**any small peptides exert their effects at very low concentrations - small peptides.

•**<u>Hypothalamus:</u>**

Oxytocin (nine amino acid residues), which is secreted by the posterior pituitary and stimulates uterine contractions;

Vasopressin

bradykinin (nine residues), which inhibits inflammation of tissues; and

thyrotropin-releasing factor (three residues), which is formed in the hypothalamus and stimulates the release of another hormone, thyrotropin, from the anterior pituitary gland.

Oxytocin (b) Cys—Tyr—Ile—Gln—Asn—Cys—Pro—Leu—Gly—NH_2 (S—S)

Vasopressin (c) Cys—Tyr—Phe—Gln—Asn—Cys—Pro—Arg—Gly—NH_2 (S—S)

Met-enkephalin Tyr—Gly—Gly—Phe—Met

Leu-enkephalin Tyr—Gly—Gly—Phe—Leu

Atrial natriuretic factor Ser[1]—Leu—Arg—Arg—Ser—Ser—Cys—Phe—Gly—Gly[10]—Arg—Met—Asp—Arg—Ile—Gly—Ala—Gln—Ser—Gly—Leu—Gly—Cys—Asn—Ser—Phe—Arg—Tyr[28]

Substance P Arg—Pro—Lys—Pro—Gln—Phe—Phe—Gly—Leu—Met—NH_2

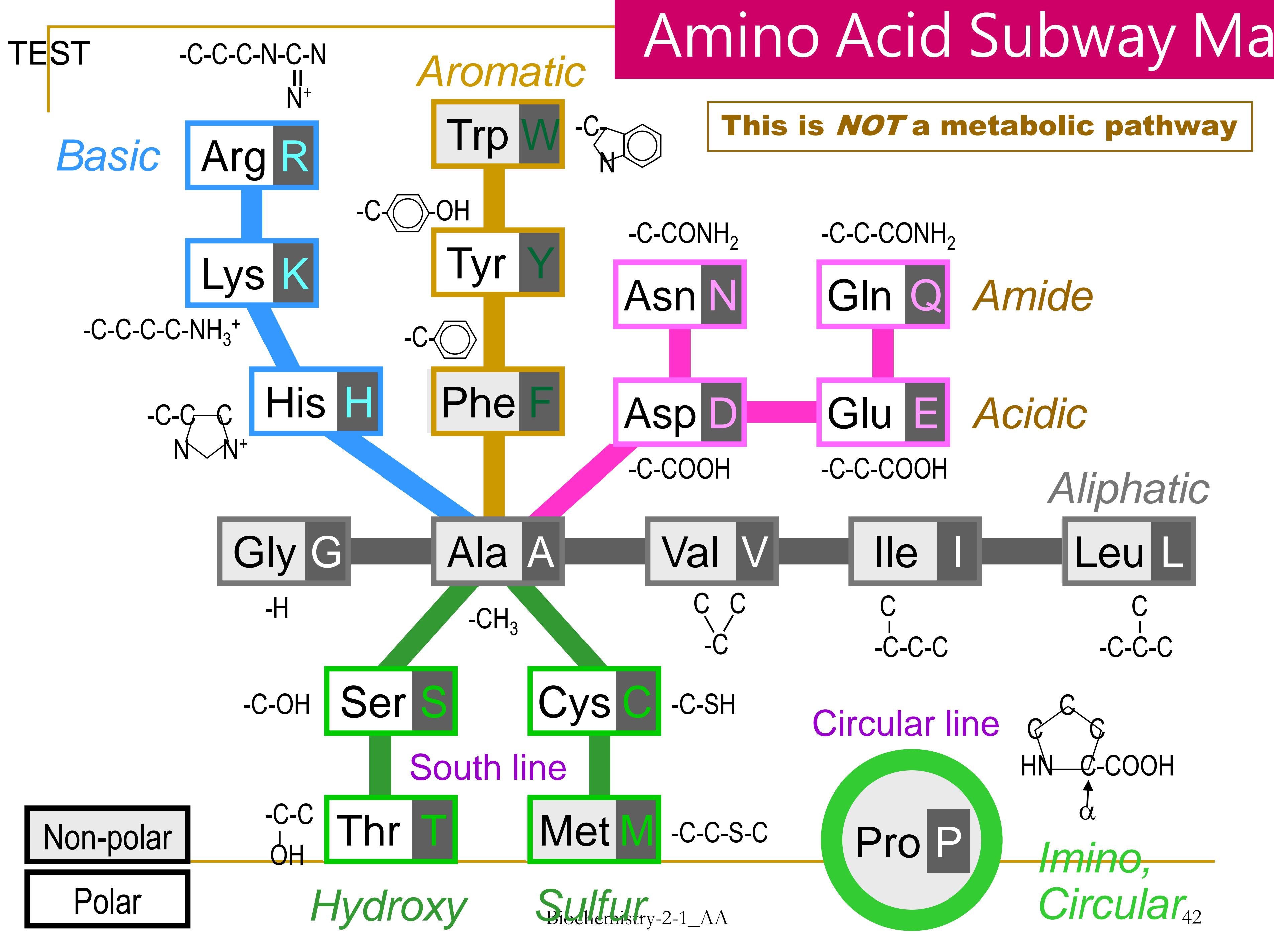
Bradykinin Arg—Pro—Pro—Gly—Phe—Ser—Pro—Phe—Arg

Peptides

- Aspartame

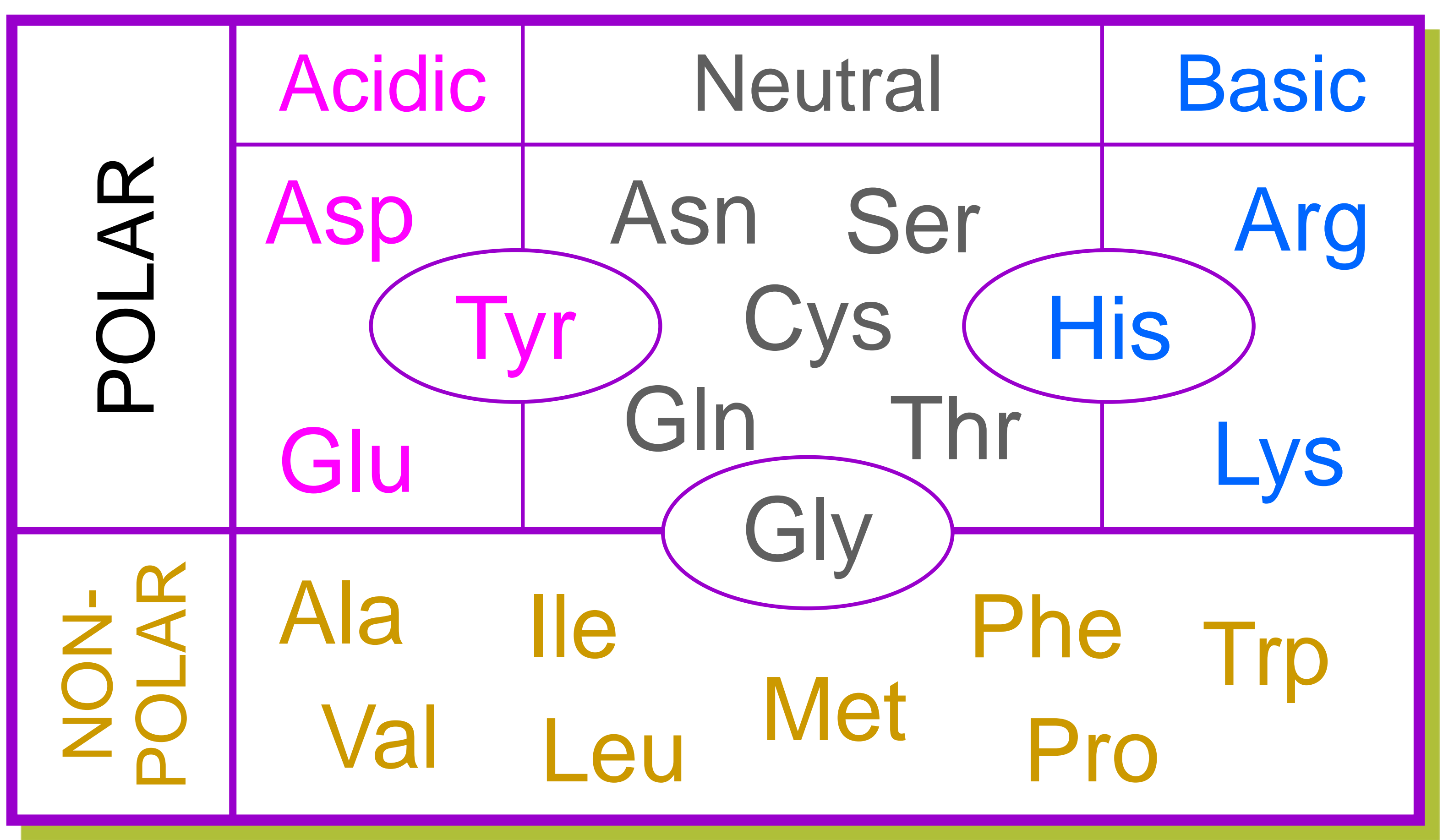
L-aspartyl-L-phenylalanine methyl ester, the artificial sweetener better known as aspartame or NutraSweet.

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


TEST
Amino Acid Subway Ma
This is NOT a metabolic pathway
Basic
Aromatic
Arg R
-C-C-C-N-C-N
N+
Lys K
-C-C-C-C-NH3+
His H
-C-C C
N N+
Trp W
-C
N
-C- -OH
Tyr Y
-C-
Phe F
-C-CONH2
Asn N
-C-C-CONH2
Gln Q
Amide
Asp D
Glu E
Acidic
-C-COOH
-C-C-COOH
Aliphatic
Gly G
-H
Ala A
-CH3
Val V
C C
-C
Ile I
C
-C-C-C
Leu L
C
-C-C-C
Ser S
-C-OH
Cys C
-C-SH
South line
Thr T
-C-C
OH
Met M
-C-C-S-C
Hydroxy
Sulfur
Circular line
C
C C
HN C-COOH
α
Pro P
Imino,
Circular
Non-polar
Polar

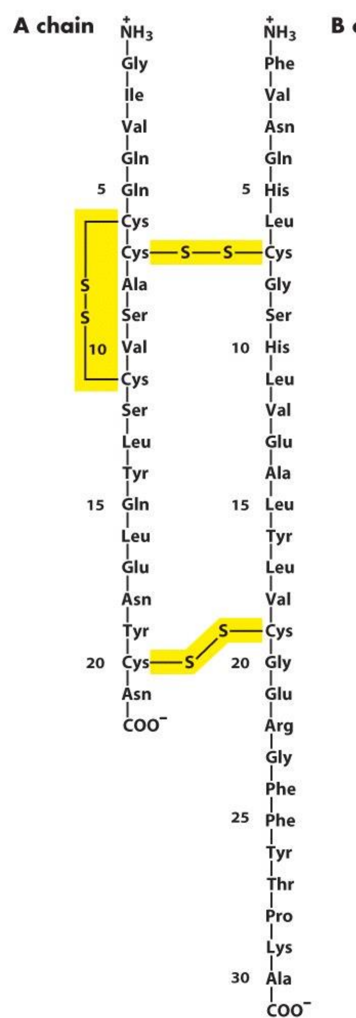
Classification of Amino Acids by Polarity

TEST

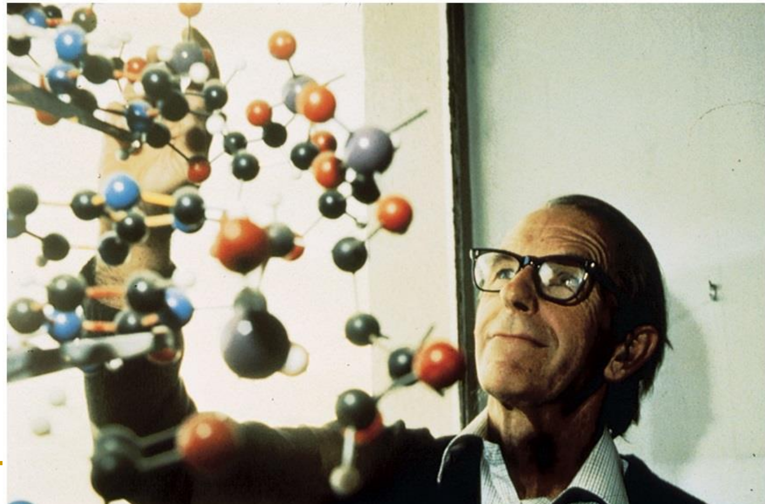


Polar or non-polar, it is the bases of the amino acid properties.

Primary Structure of Bovine Insulin



First protein to be fully sequenced (by Fred Sanger in 1953). For this, he won his first Nobel Prize (his second was for the Sanger dideoxy method of DNA sequencing).

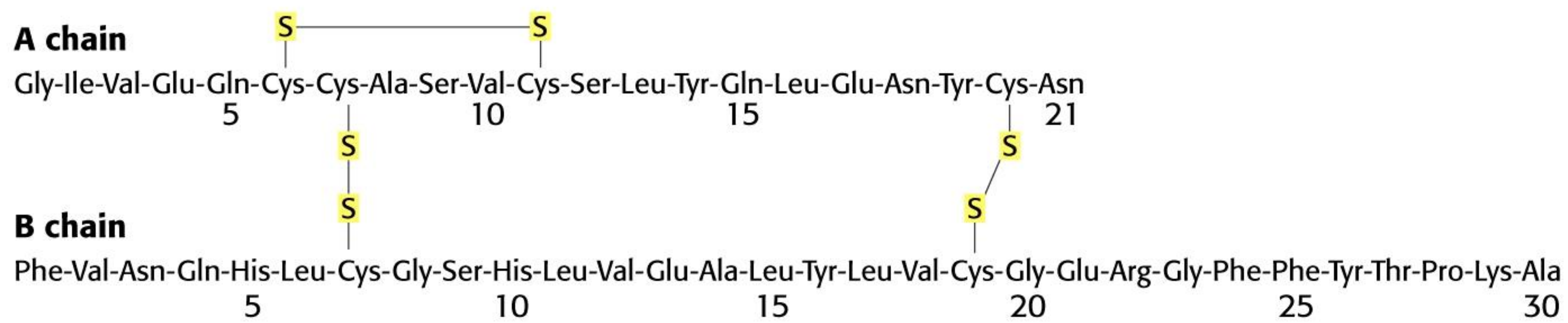


Bovine insulin: AA sequence

1953, Fred Sanger determined aa sequence of insulin, landmark!

Showed for 1st time, protein has precisely defined aa sequence
Also showed that only L-amino acids were present, linked by peptide bonds

Now, aa sequence of > 100,000 proteins are known



1950s-1960s studies showed aa sequence genetically determined
Each of 20 aa encoded by one or more specific sequences of 3 nucleotides.

Peptid

Some extremely **toxic mushroom poisons**, such as **amanitin**,
are also small peptides, as are many antibiotics (**neomycin, kanamycin**).

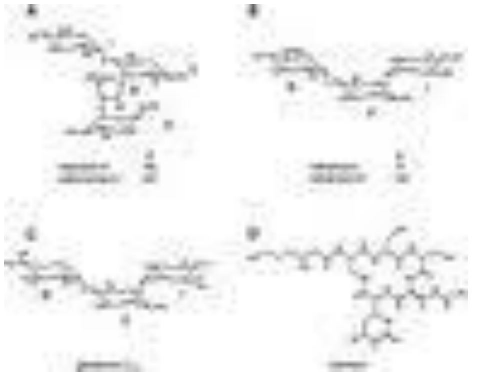


TABLE 3–2 Molecular Data on Some Proteins

	Molecular weight	*Number of residues*	*Number of polypeptide chains*
Cytochrome *c* (human)	13,000	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (chicken egg white)	13,930	129	1
Myoglobin (equine heart)	16,890	153	1
Chymotrypsin (bovine pancreas)	21,600	241	3
Chymotrypsinogen (bovine)	22,000	245	1
Hemoglobin (human)	64,500	574	4
Serum albumin (human)	68,500	609	1
Hexokinase (yeast)	102,000	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

TABLE 3–3 Amino Acid Composition of Two Proteins

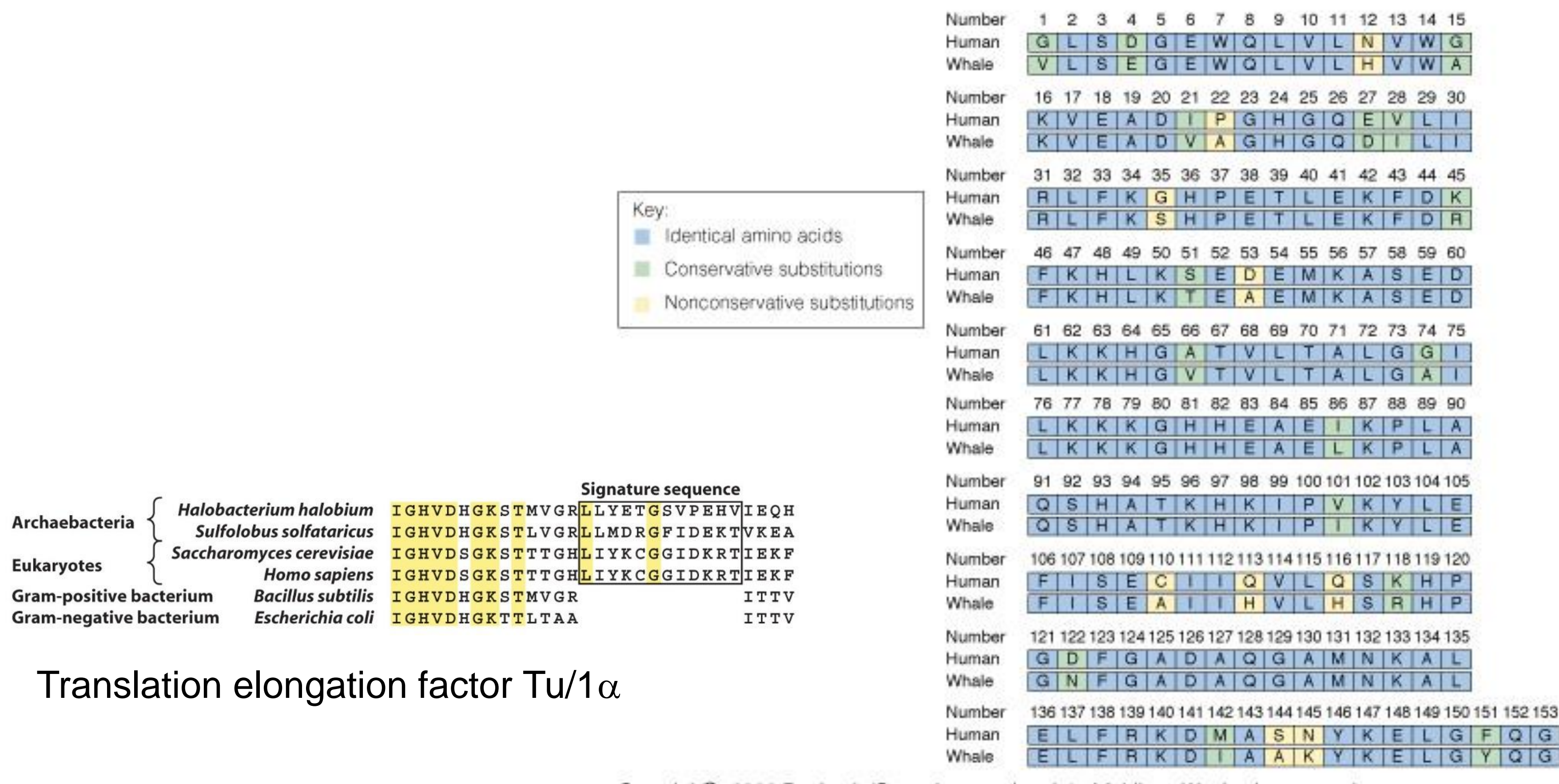
Amino acid	*Number of residues per molecule of protein**	
	Bovine cytochrome c	*Bovine chymotrypsinogen*
Ala	6	22
Arg	2	4
Asn	5	15
Asp	3	8
Cys	2	10
Gln	3	10
Glu	9	5
Gly	14	23
His	3	2
Ile	6	10
Leu	6	19
Lys	18	14
Met	2	2
Phe	4	6
Pro	4	9
Ser	1	28
Thr	8	23
Trp	1	8
Tyr	4	4
Val	3	23
Total	104	245

*In some common analyses, such as acid hydrolysis, Asp and Asn are not readily distinguished from each other and are together designated Asx (or B). Similarly, when Glu and Gln cannot be distinguished, they are together designated Glx (or Z). In addition, Trp is destroyed. Additional procedures must be employed to obtain an accurate assessment of complete amino acid content.

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	Ferritin
	Zinc	Alcohol dehydrogenase
	Calcium	Calmodulin
	Molybdenum	Dinitrogenase
	Copper	Plastocyanin

Evolution and Conservation of Protein Sequences

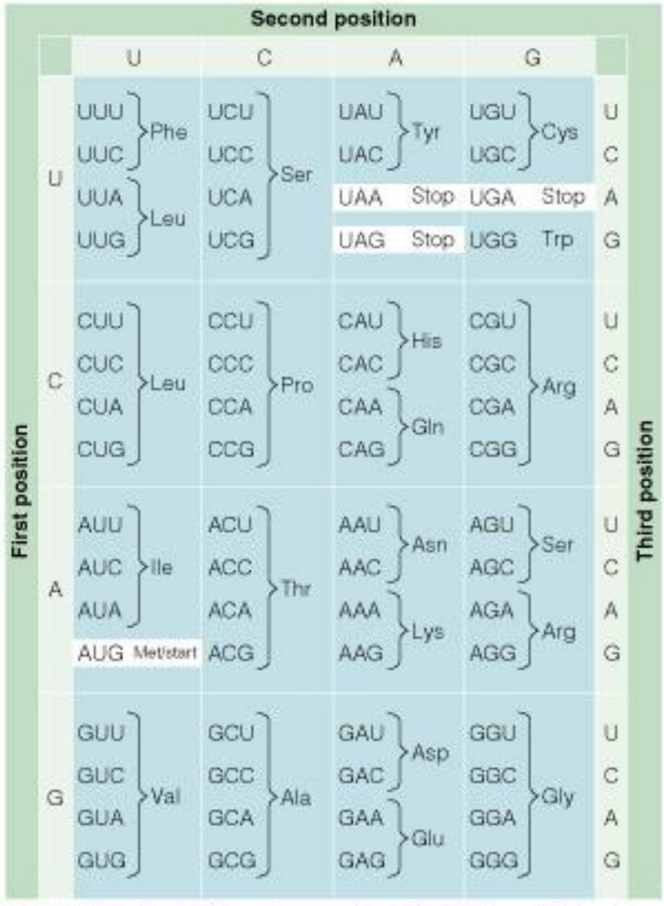


Copyright© 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Translation elongation factor Tu/1α

Myoglobin

The Genetic Code



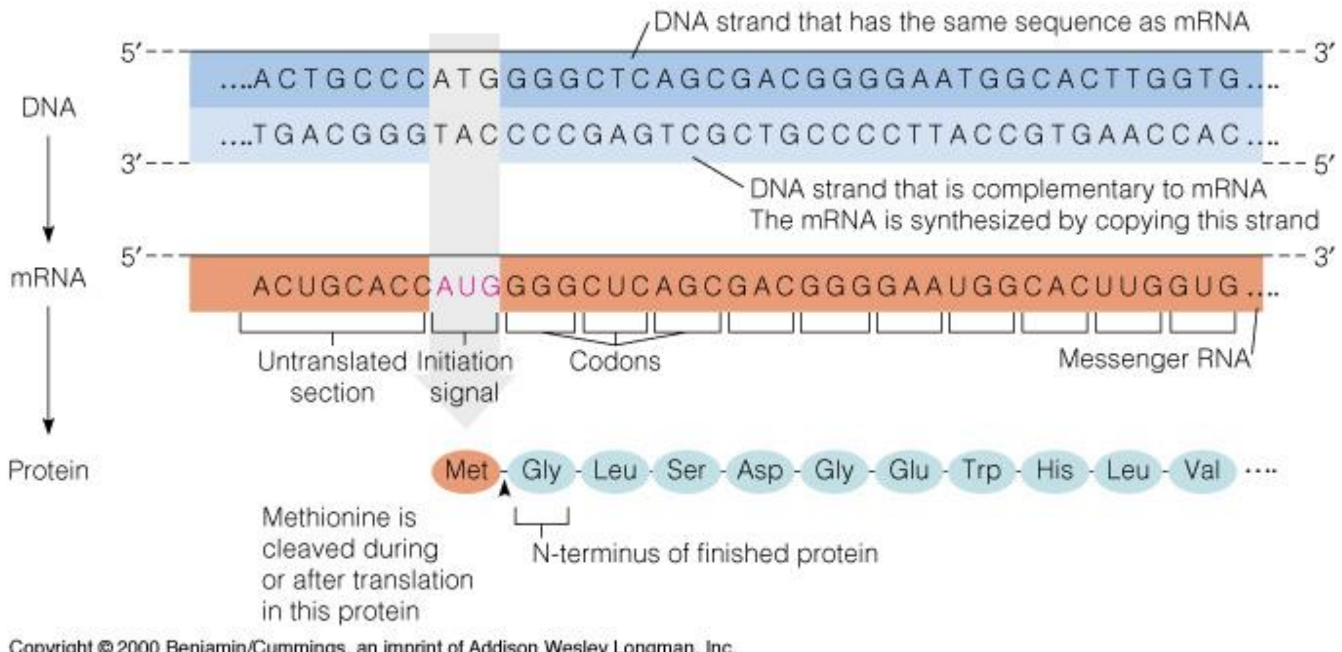
Second position

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

Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

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

DNA sequence (gene) **CAGTATCCTACGATTTCG**



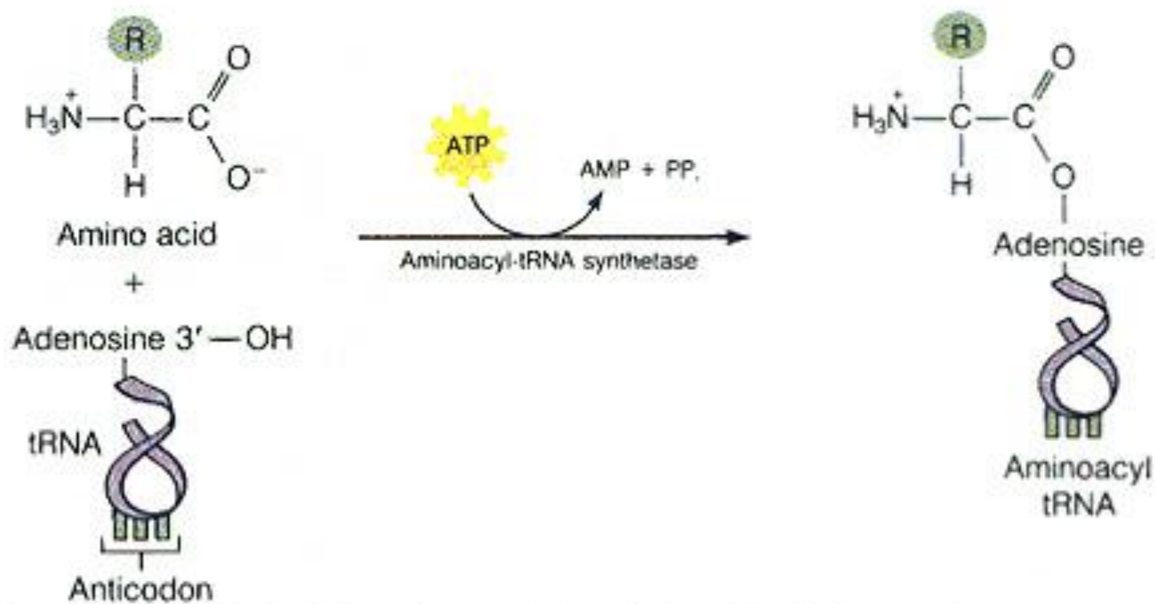
Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Initiating Amino Acid in Translation

N-Formylmethionine in prokaryotes

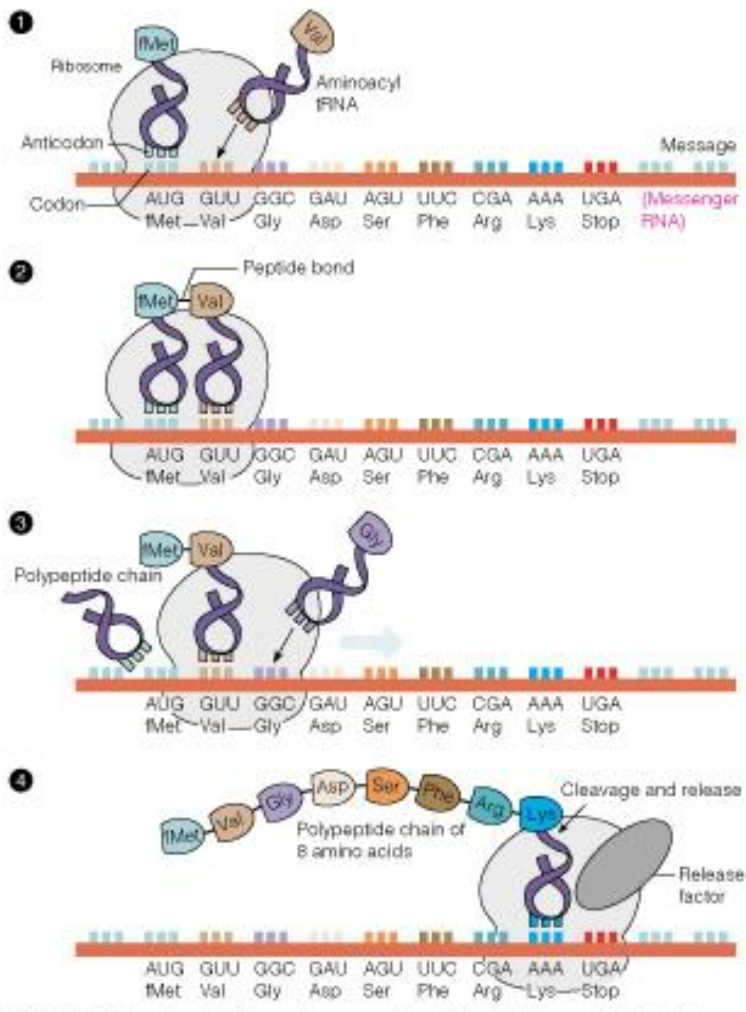
Just methionine in eukaryotes

Charging of tRNAs with Specific Amino Acids



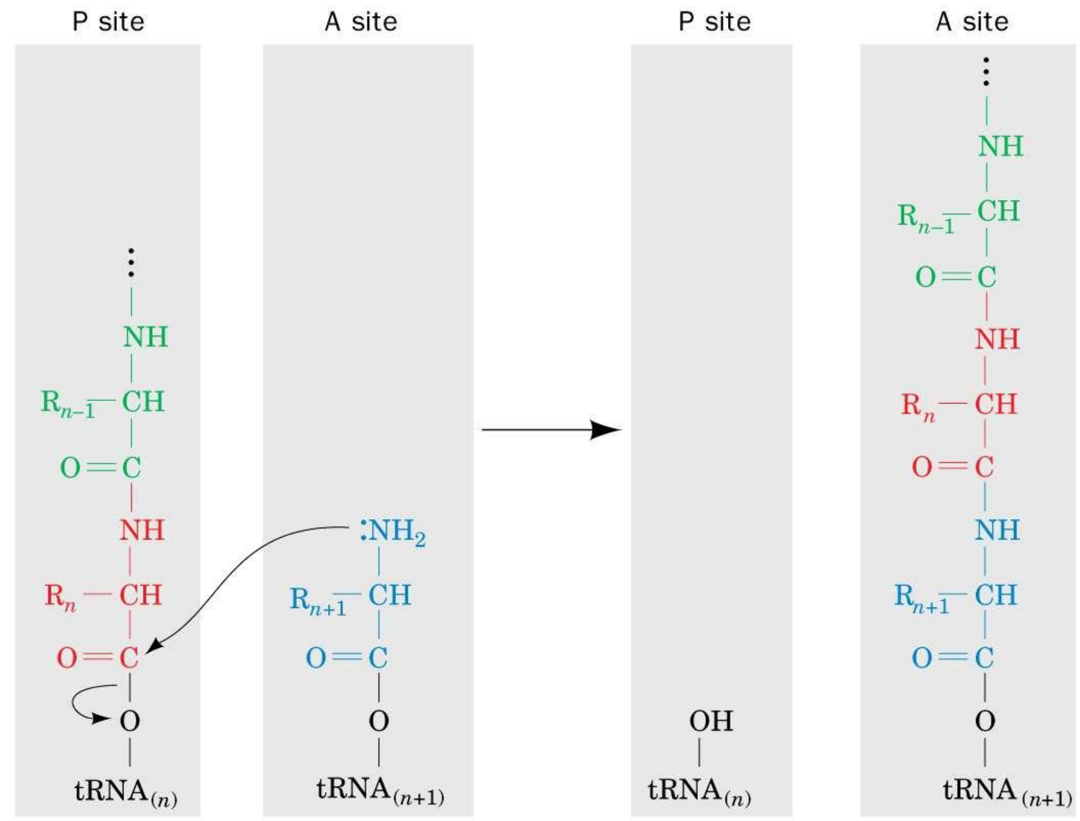
Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Translation of mRNA into Protein



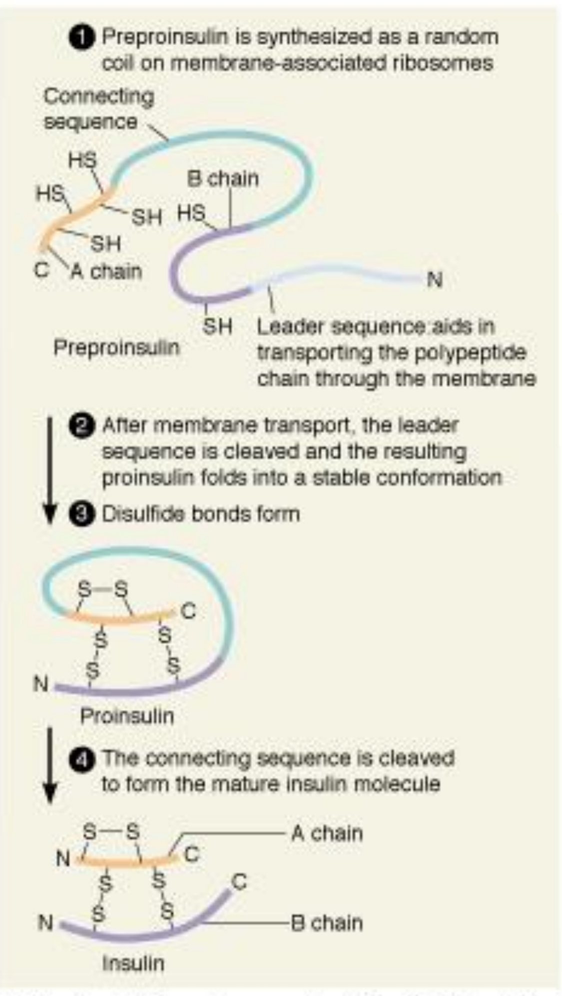
Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Ribosomal Peptidyl Transferase Activity



Note: the catalytic component of the ribosome's peptidyl transferase activity is RNA; it's an example of a catalytic RNA or ribozyme.

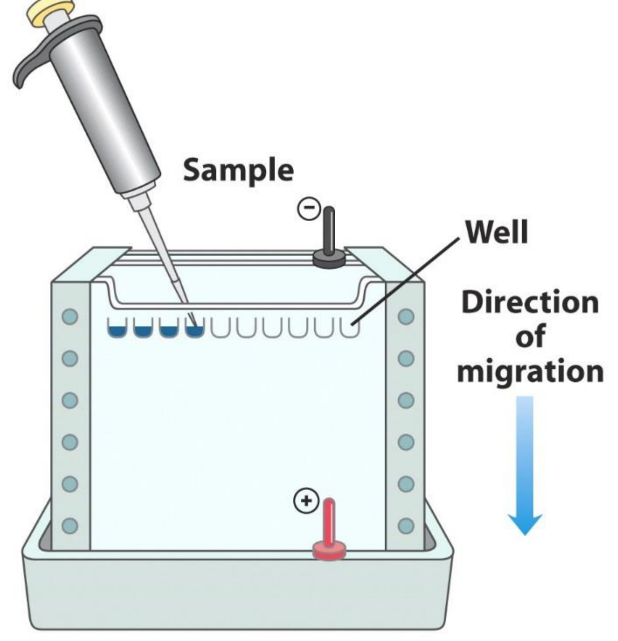
Disulfide Bond Formation in Insulin

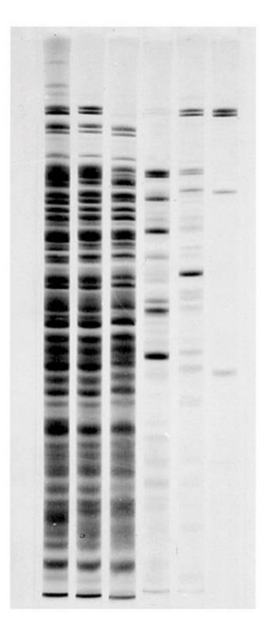


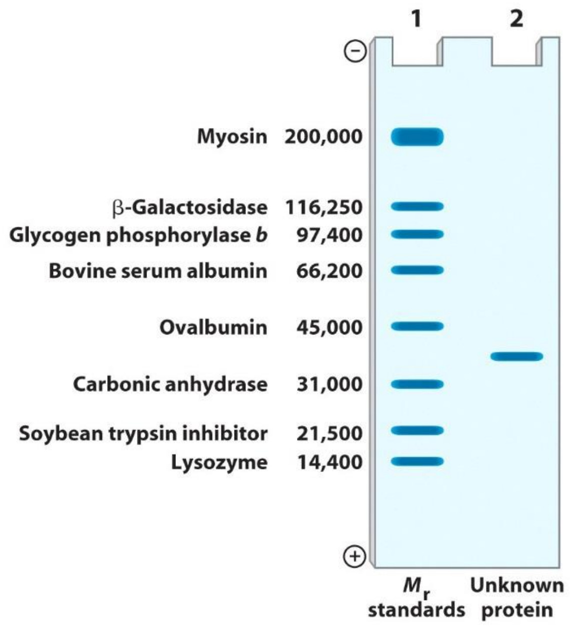
Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

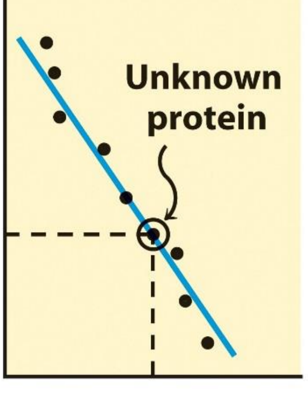
Methods in Protein Biochemistry

Gel Electrophoresis

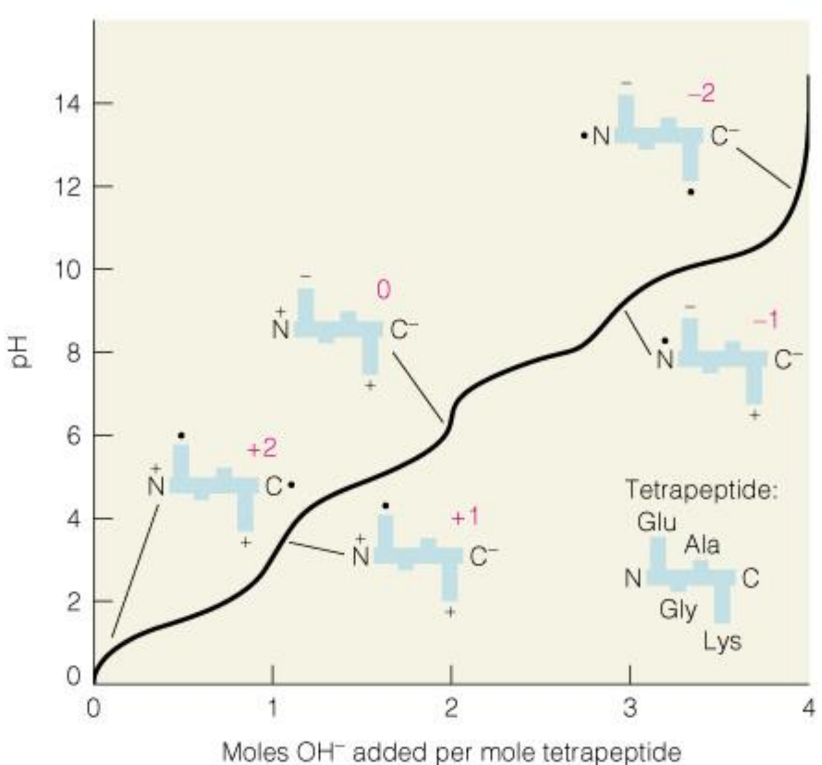








Polyampholyte Character of a Tetrapeptide and Isoelectric Points



Group	pKa
α-NH_3^+	9.7
Glu γ-COOH	4.2
Lys ε-NH_3^+	10.0
α-COOH	2.2

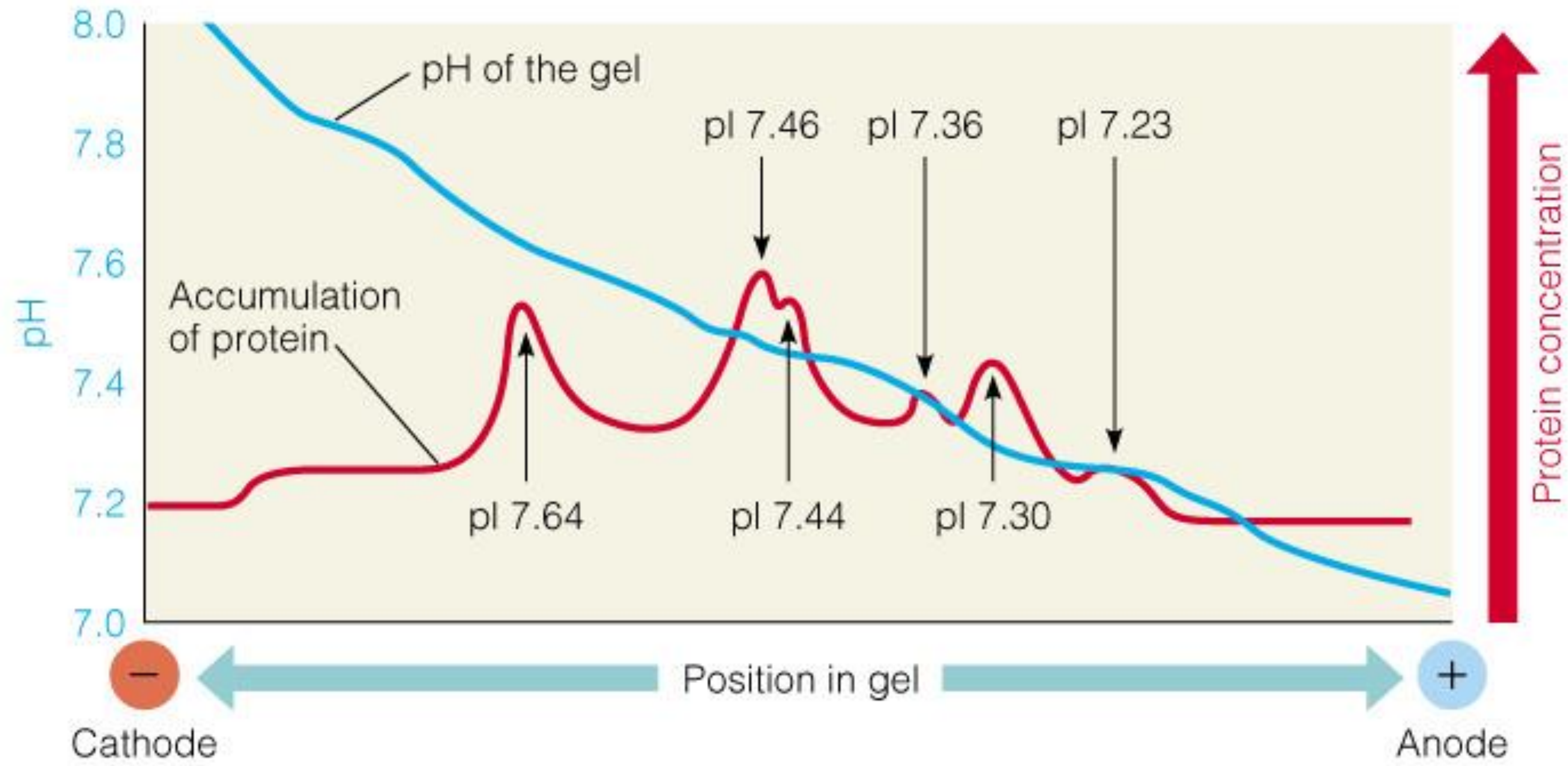
Isoelectric Point (pI), pH at which molecule has net zero charge, determined using computer program for known sequence or empirically (by isoelectric focusing).

TABLE 3–6 The Isoelectric Points of Some Proteins

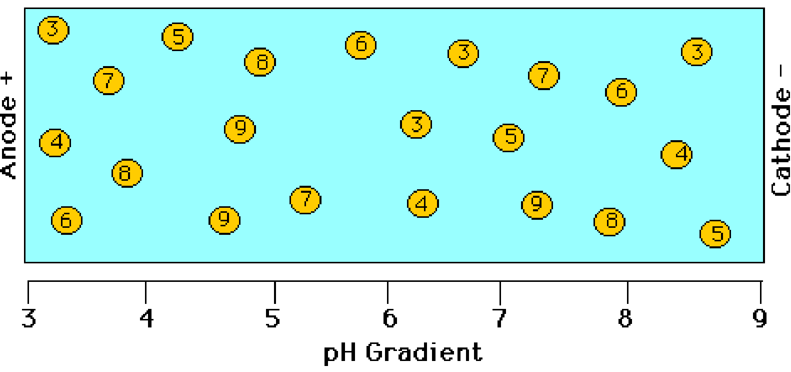
Protein	*pI*
Pepsin	<1.0
Egg albumin	4.6
Serum albumin	4.9
Urease	5.0
β-Lactoglobulin	5.2
Hemoglobin	6.8
Myoglobin	7.0
Chymotrypsinogen	9.5
Cytochrome *c*	10.7
Lysozyme	11.0

Isoelectric Focusing

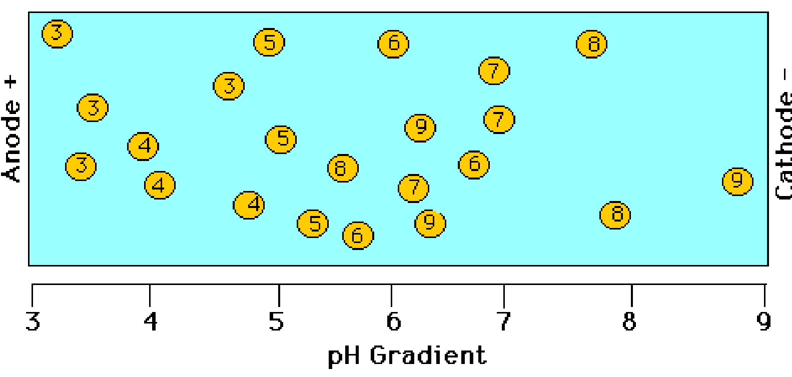
Electrophoresis through polyacrylamide gel in which there is a pH gradient.



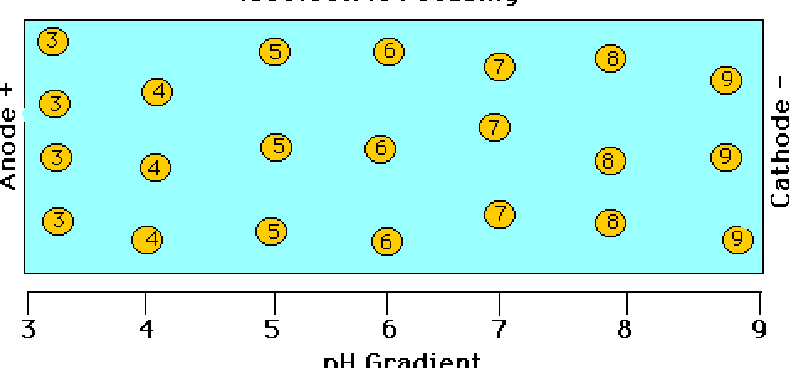
Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.



Isoelectric Focusing

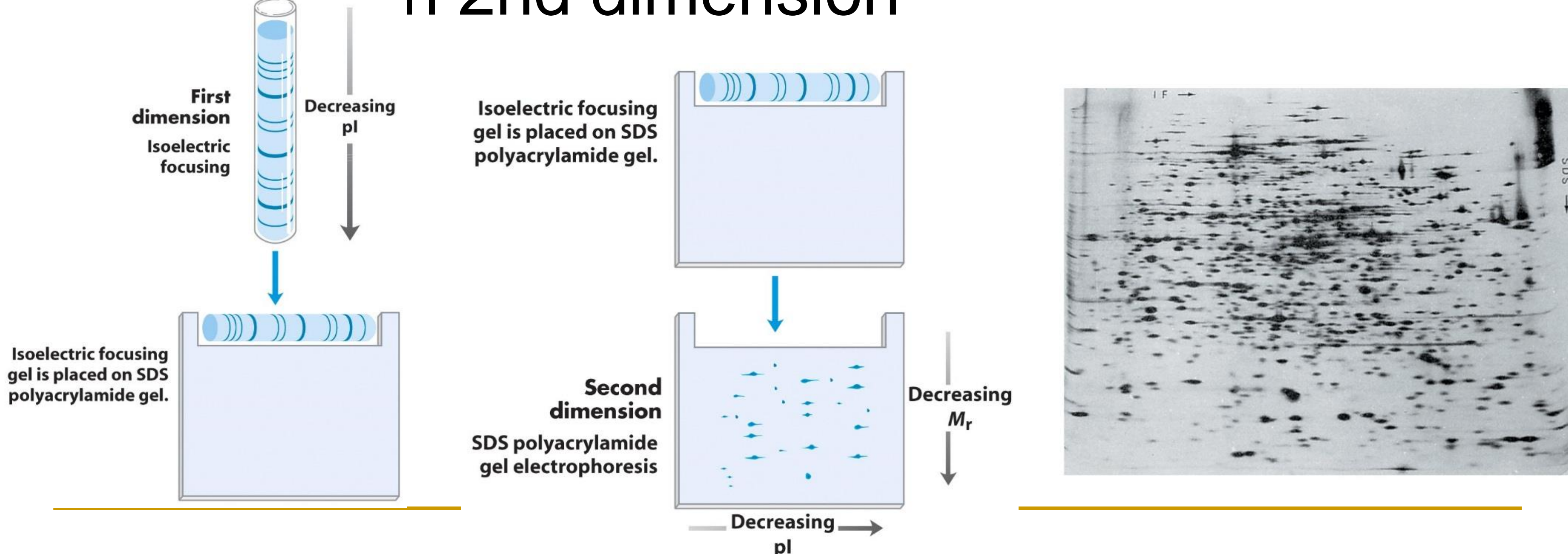


Isoelectric Focusing

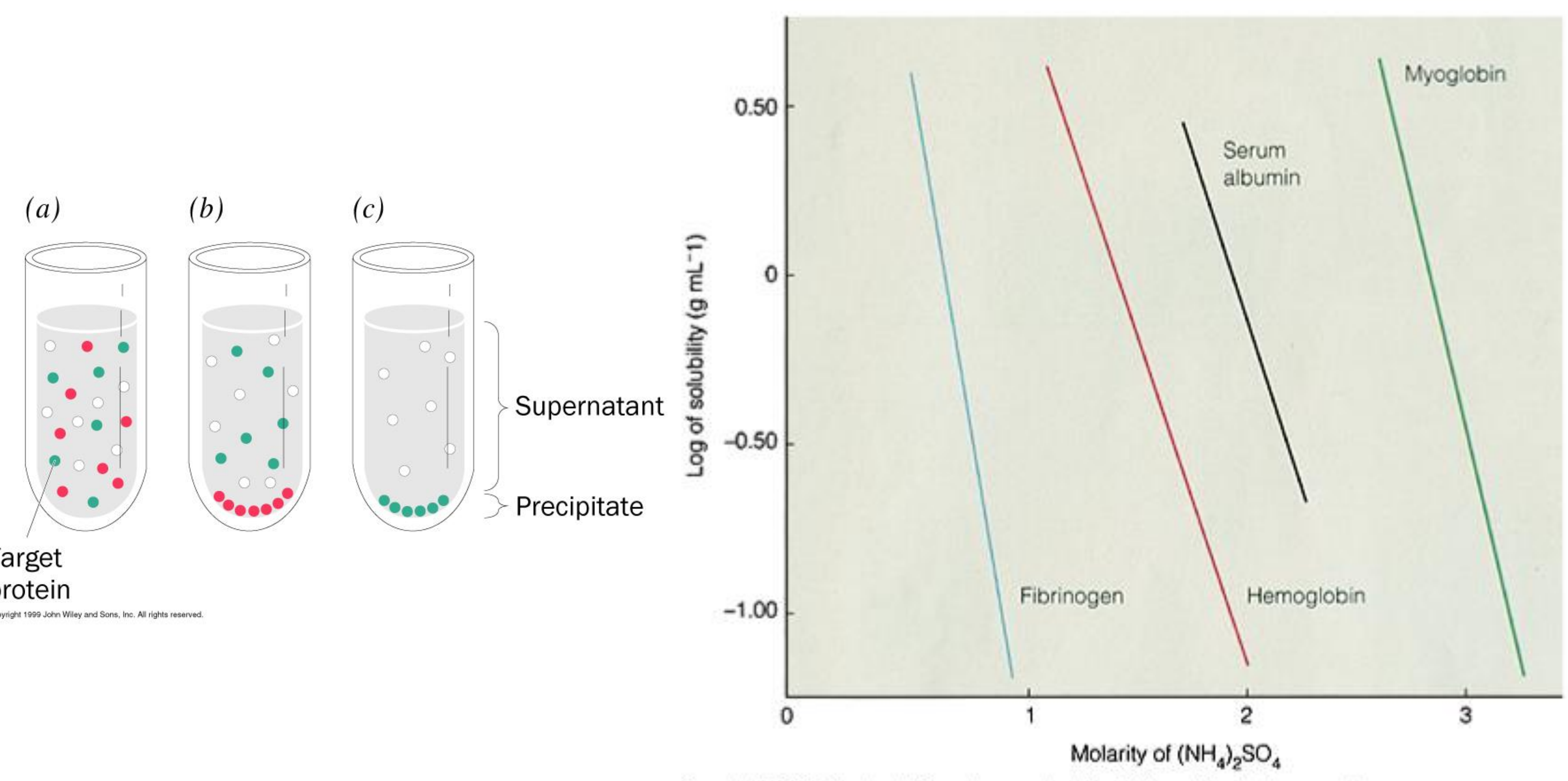


Two-Dimensional Gel Electrophoresis

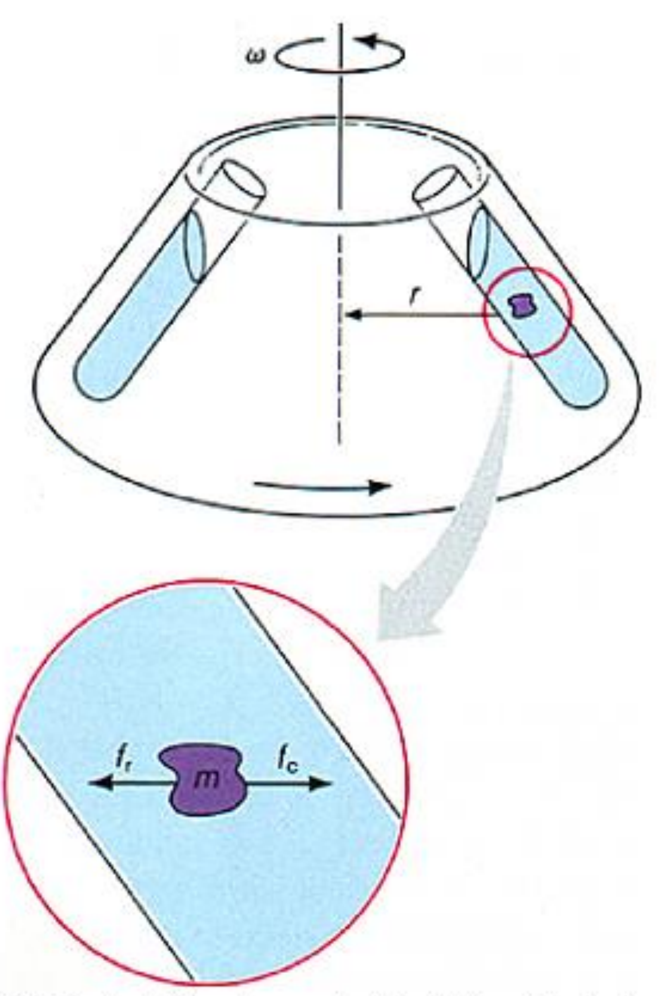
- Separate proteins based on isolectric point in 1st dimension
- Separate proteins based on molecular weight in 2nd dimension



"Salting Out": Ammonium Sulfate Precipitation in Protein Fractionation



Centrifugation



Low-speed, high-speed, or ultracentrifugation: different spin speeds and *g* forces

Centrifugation Methods

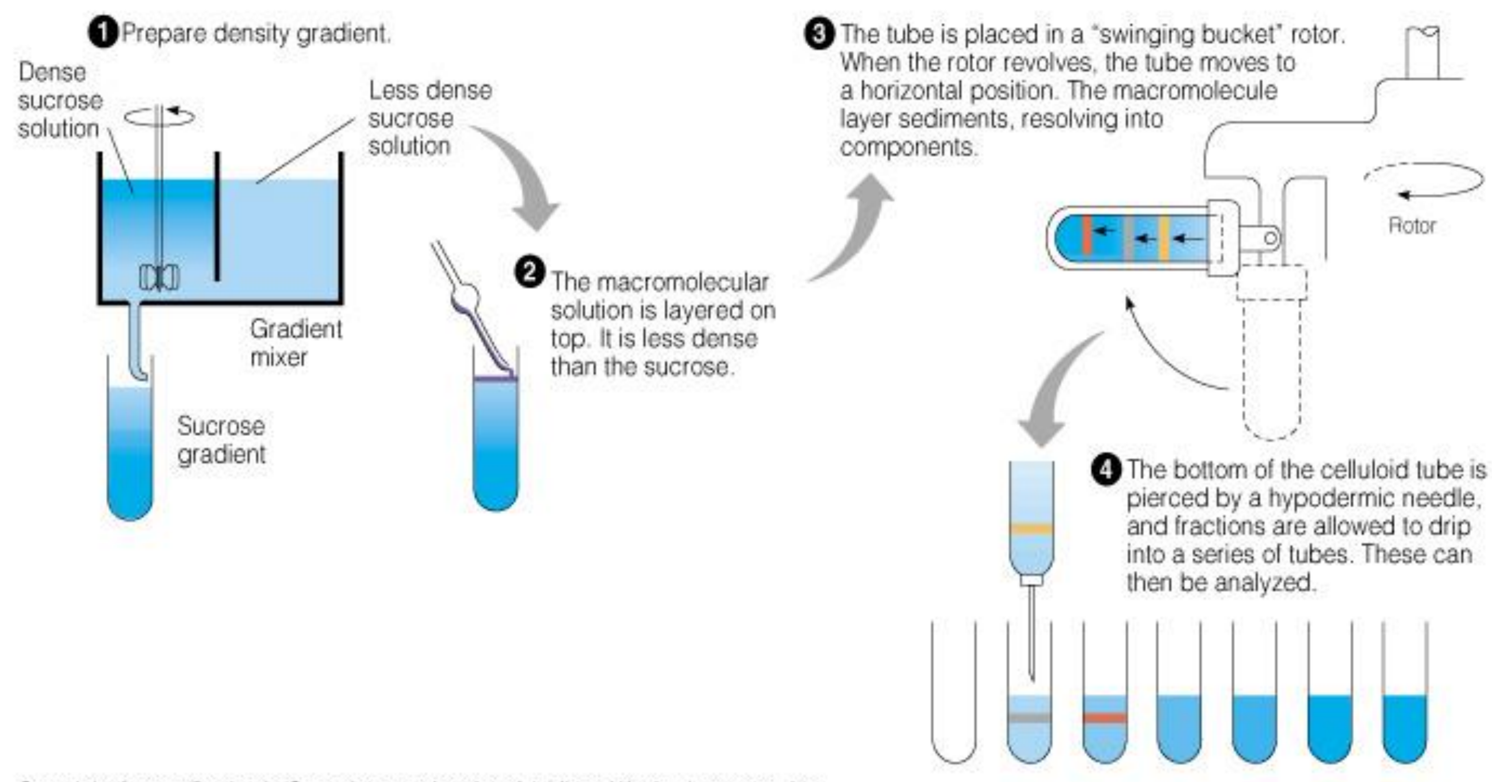
•Differential (Pelletting) – simple method for pelleting large particles using fixed-angle rotor (pellet at bottom of tube vs. supernatant solution above)

•Zonal ultracentrifugation (*e.g.*, sucrose-gradient) – swinging-bucket rotor

•Equilibrium-density gradient ultracentrifugation (*e.g.*, CsCl) – swinging-bucket or fixed-angle rotor

Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

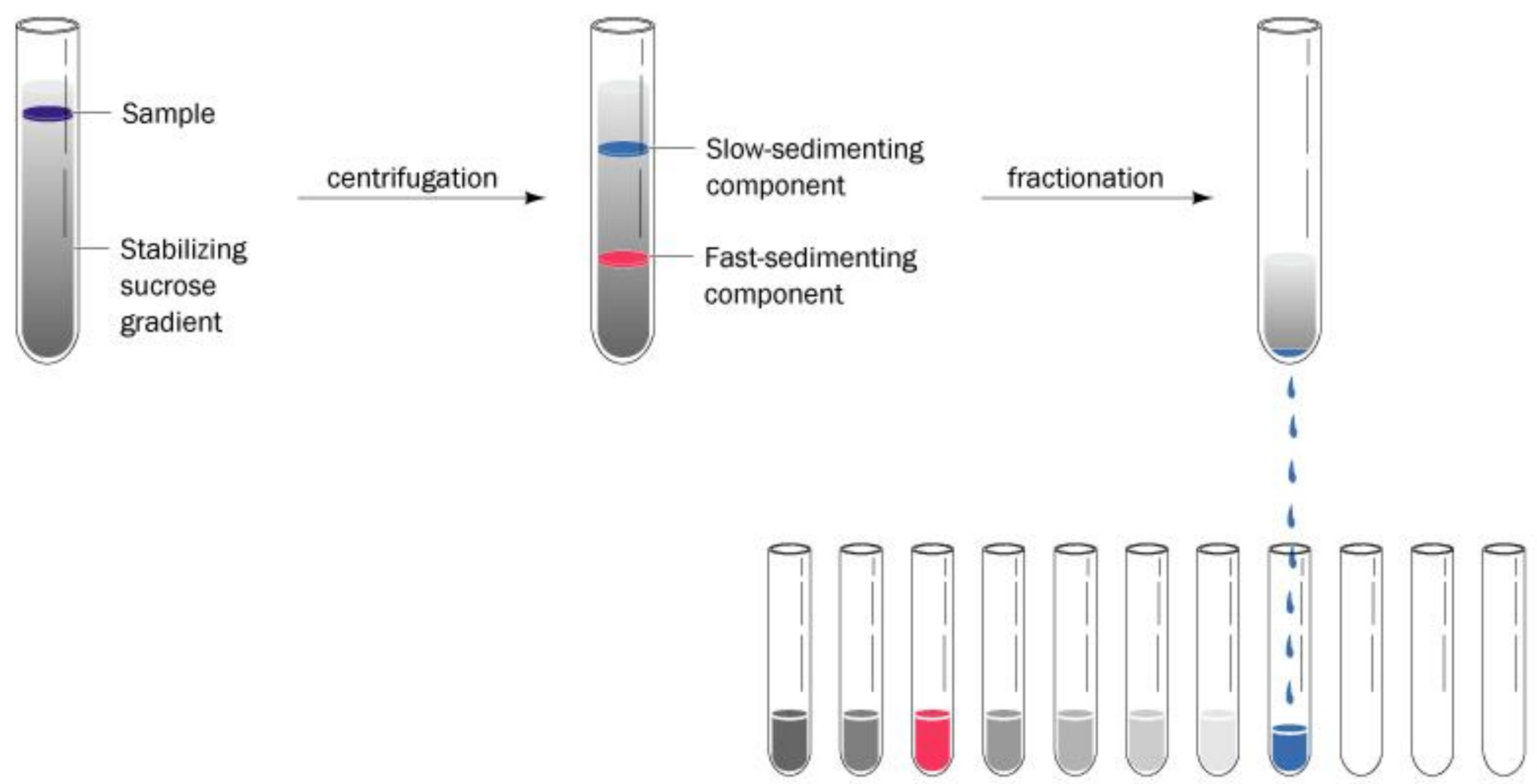
Zonal Centrifugation: Sucrose-Gradient Preparative Ultracentrifugation



Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Separates by sedimentation coefficient (determined by size and shape of solutes)

Sucrose-Gradient Preparative Ultracentrifugation

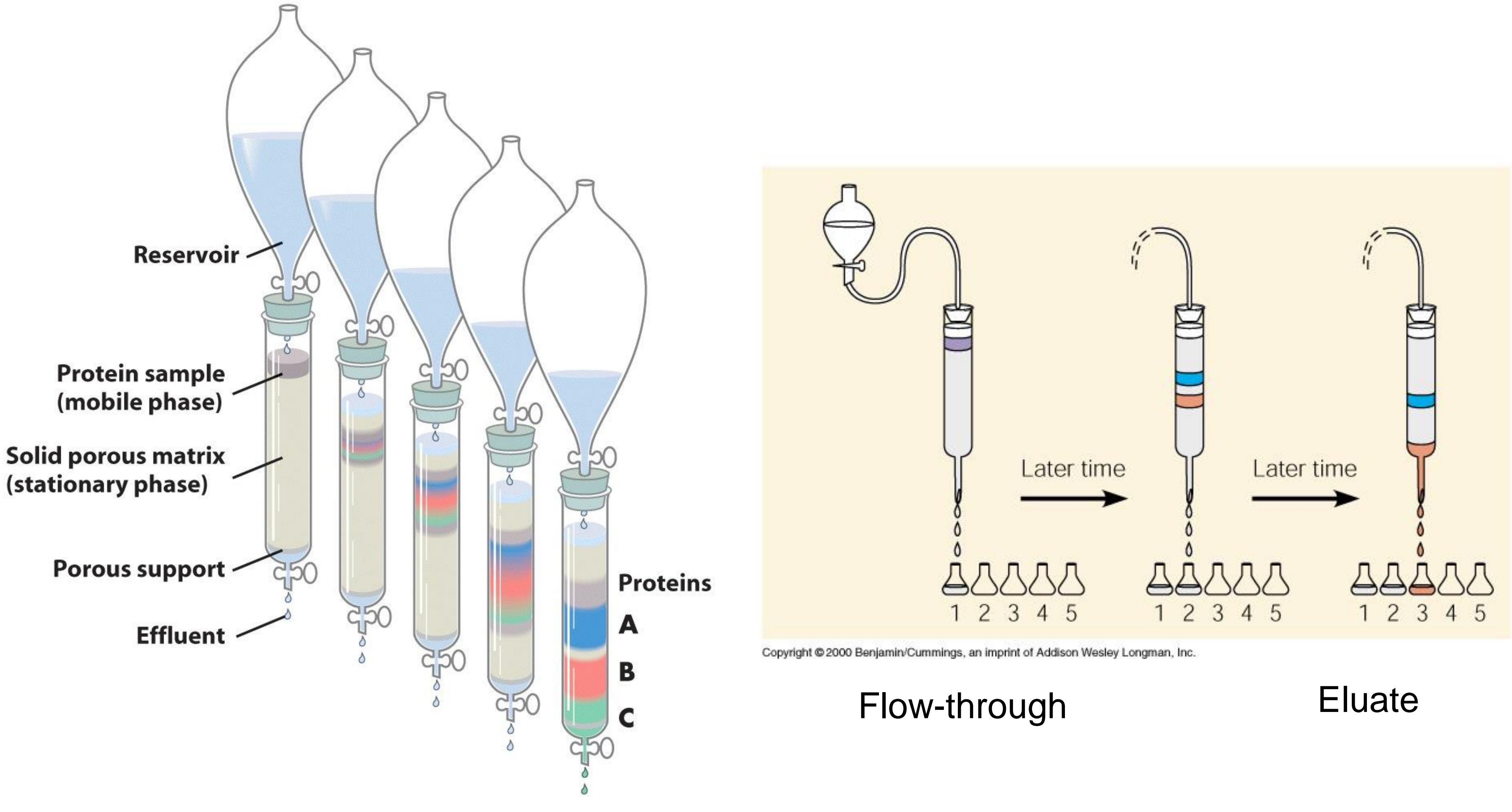


Copyright 1999 John Wiley and Sons, Inc. All rights reserved.

Equilibrium Density Gradient Ultracentrifugation

- Used in Meselsen-Stahl experiment.
- Separates based on densities of solutes.
- Does not require premade gradient.
- Pour dense solution of rapidly diffusing substance in tube (usually CsCl).
- Density gradient forms during centrifugation ("self-generating gradient").
- Solutes migrate according to their buoyant density (where density of solute = density of CsCl solution).

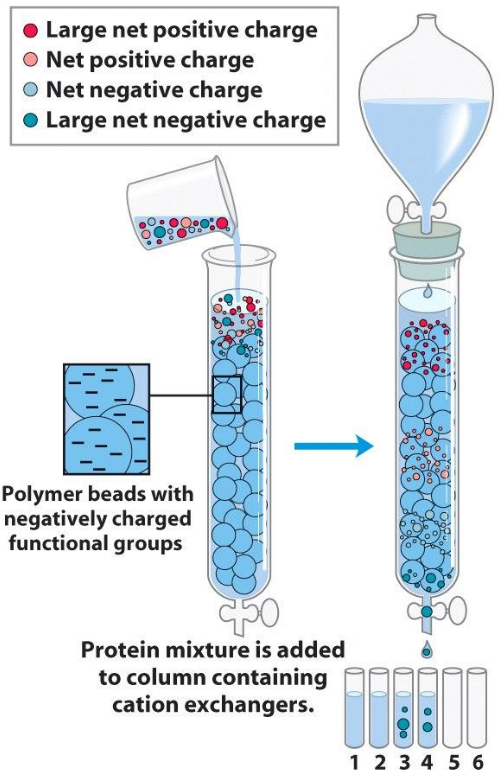
Column Chromatography

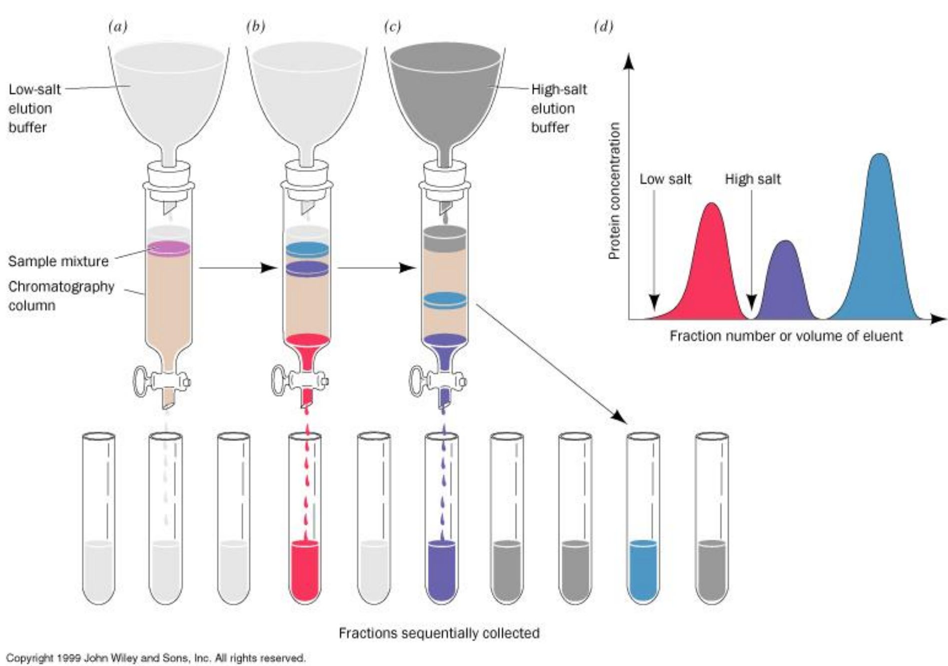


Different Types of Chromatography

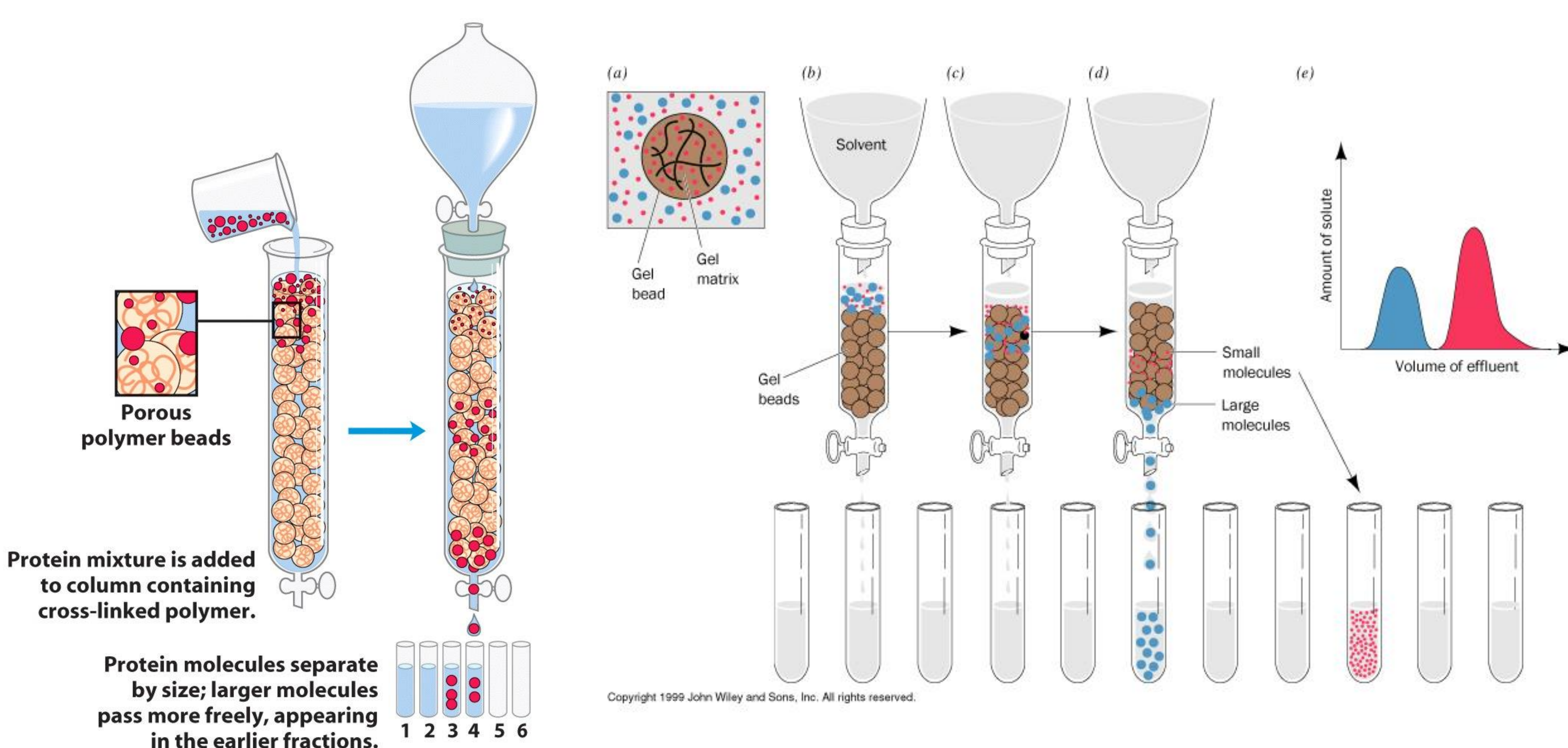
- Gel filtration/size exclusion/molecular sieve - separates by size (molecular weight) of proteins
- Ion exchange (cation exchange and anion exchange) - separates by surface charge on proteins
  - Cation exchange: separates based on positive charges of solutes/proteins, matrix is negatively charged
  - Anion exchange: separates based on negative charges of solutes/proteins, matrix is positively charged
- Hydrophobic interaction - separates by hydrophobicity of proteins
- Affinity - separates by some unique binding characteristic of protein of interest for affinity matrix

Ion-Exchange Chromatography



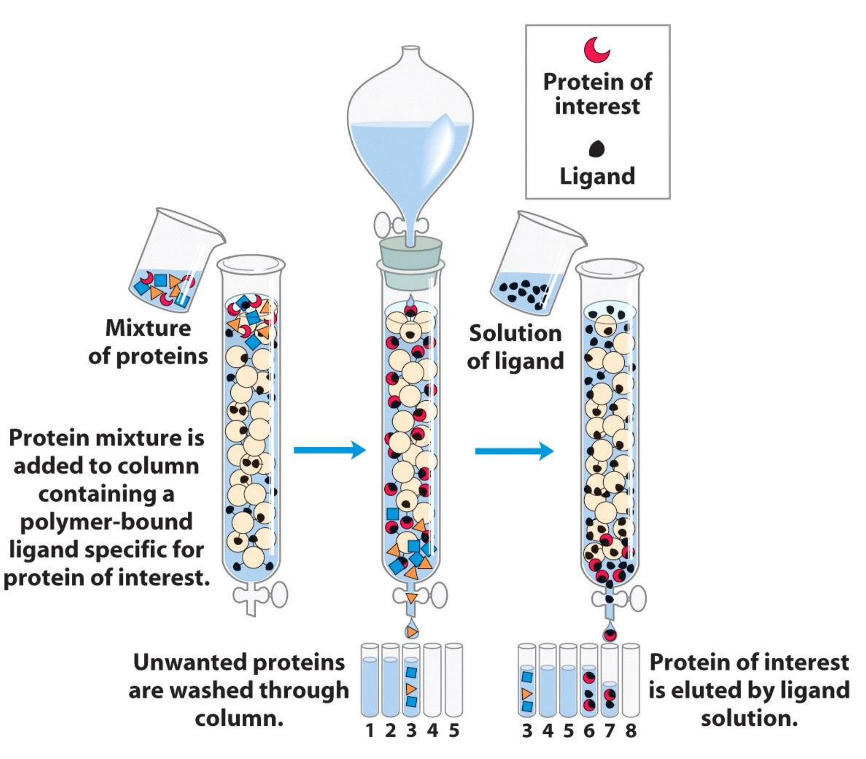


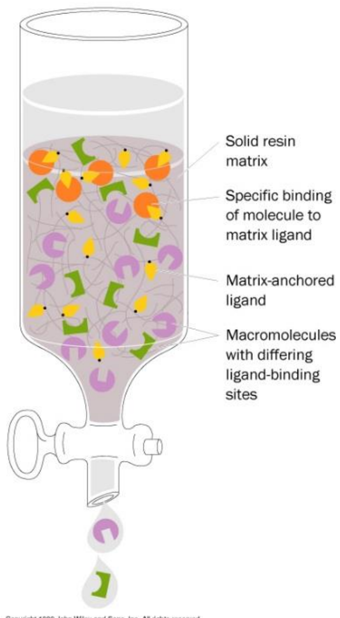
Gel Filtration Chromatography

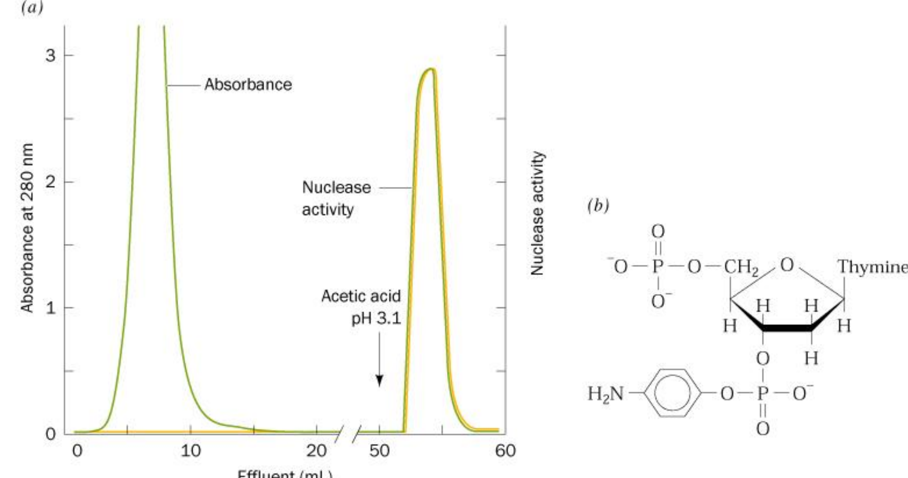


Copyright 1999 John Wiley and Sons, Inc. All rights reserved.

Affinity Chromatography







Cleavage of Polypeptides for Analysis

- Strong acid (*e.g.*, 6 M HCl) - not sequence specific
- Sequence-specific proteolytic enzymes (proteases)
- Sequence-specific chemical cleavage (*e.g.*, cyanogen bromide cleavage at methionine residues)

Protease Specificities

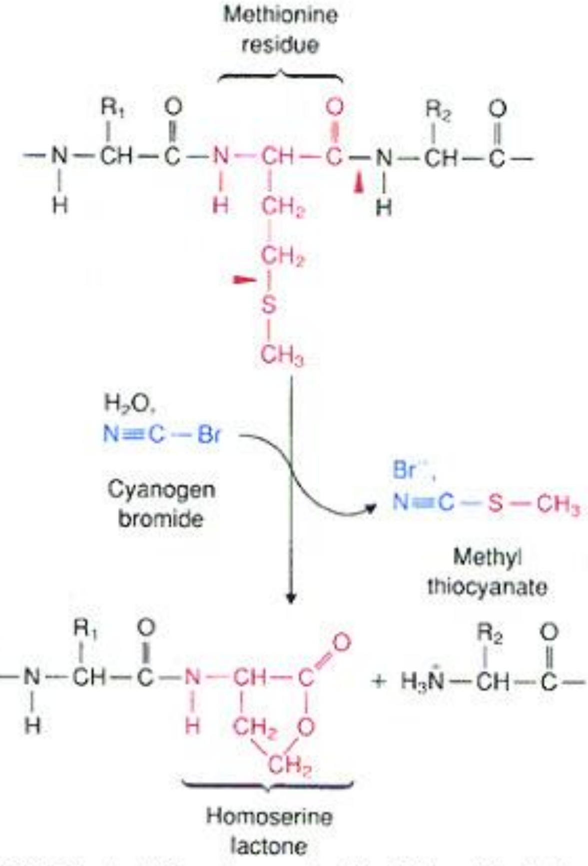
TABLE 3–7 The Specificity of Some Common Methods for Fragmenting Polypeptide Chains

*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)
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	Met (C)

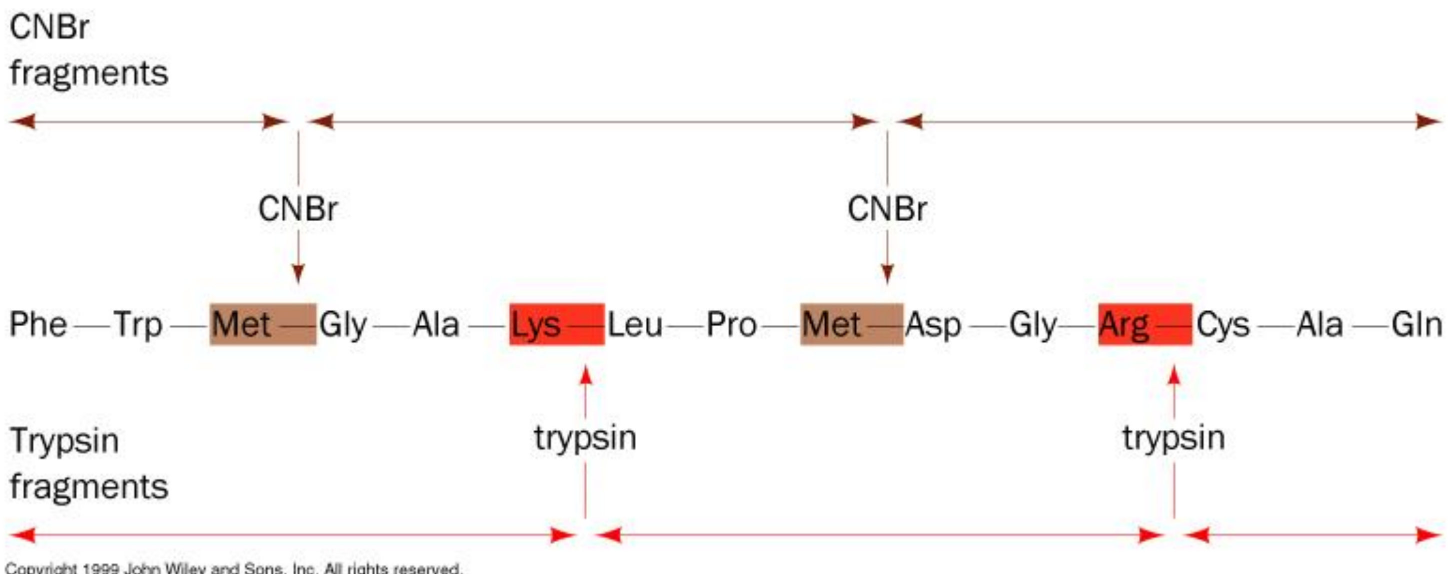
*All reagents except cyanogen bromide are proteases. All are available from commercial sources.

†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.

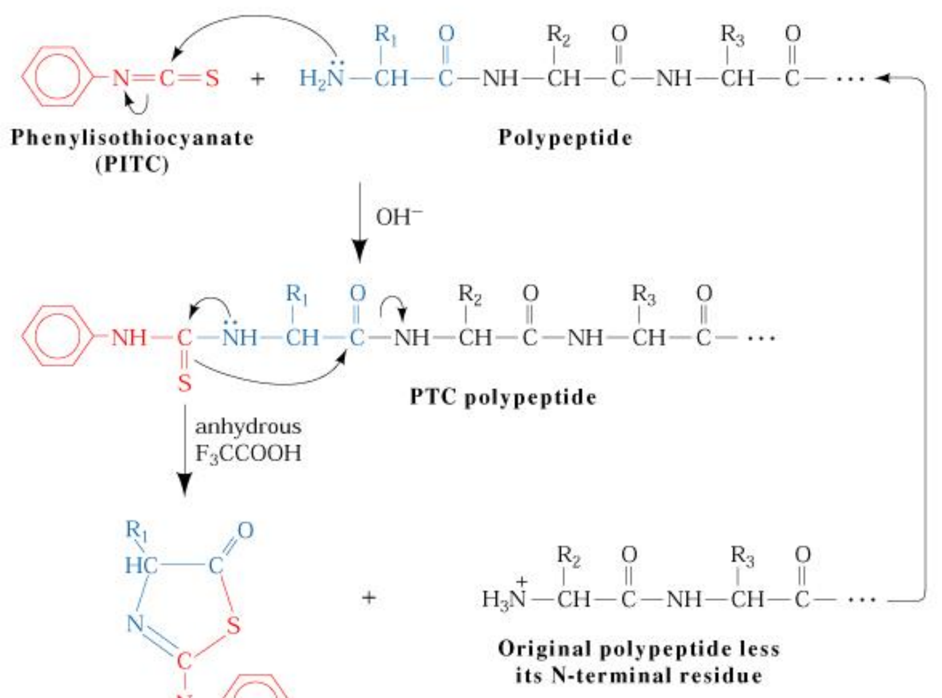
Cyanogen Bromide Cleavage at Methionine Residues



Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.


CNBr
fragments
CNBr
CNBr
Phe—Trp—Met—Gly—Ala—Lys—Leu—Pro—Met—Asp—Gly—Arg—Cys—Ala—Gln
trypsin
trypsin
Trypsin
fragments
Copyright 1999 John Wiley and Sons, Inc. All rights reserved.

Protein Sequencing: Edman Degradation



Thiazolinone derivative

H^+

PTH-amino acid

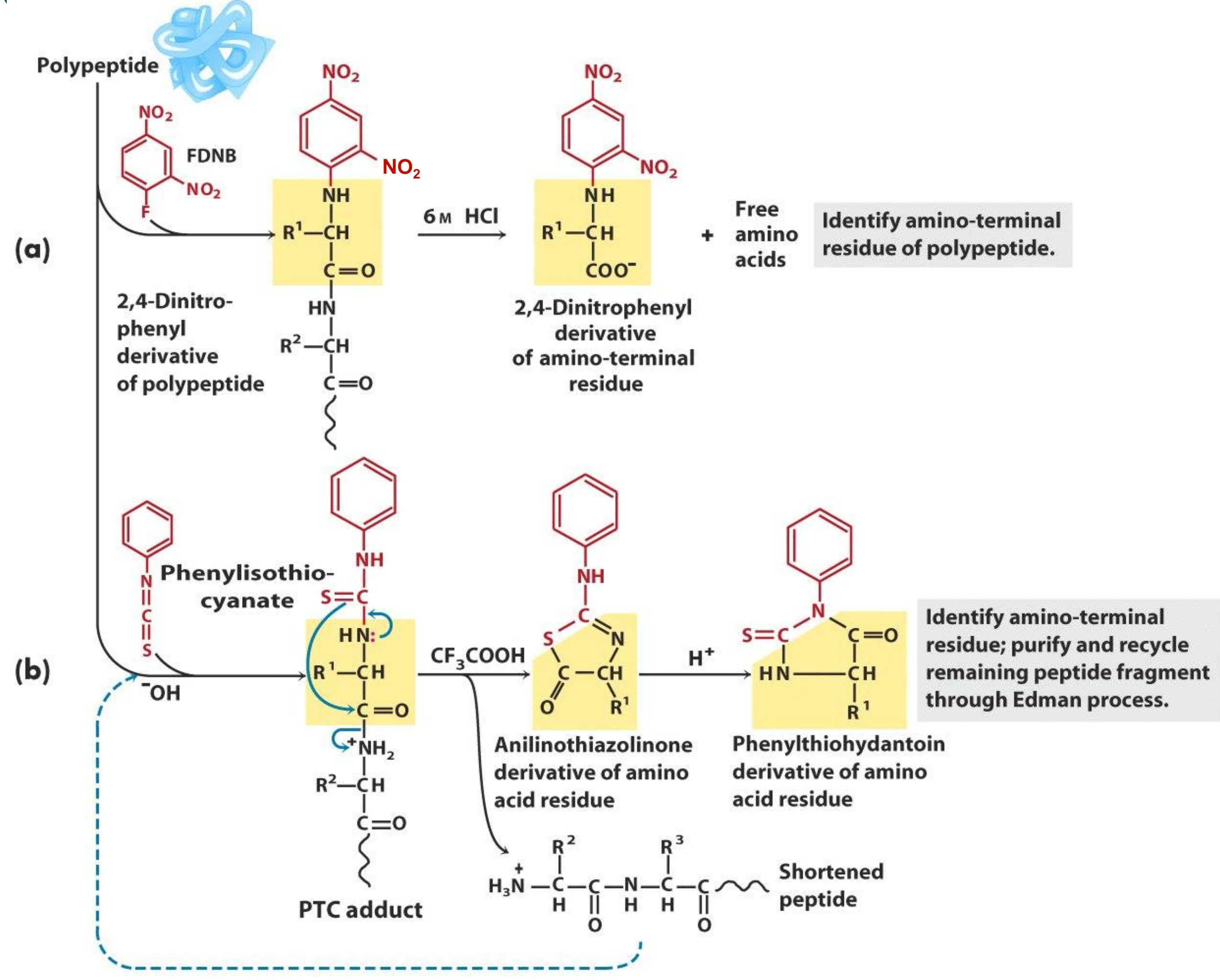
Copyright 1999 John Wiley and Sons, Inc. All rights reserved.

PTC = phenylthiocarbamyl

F_3CCOOH = trifluoroacetic acid

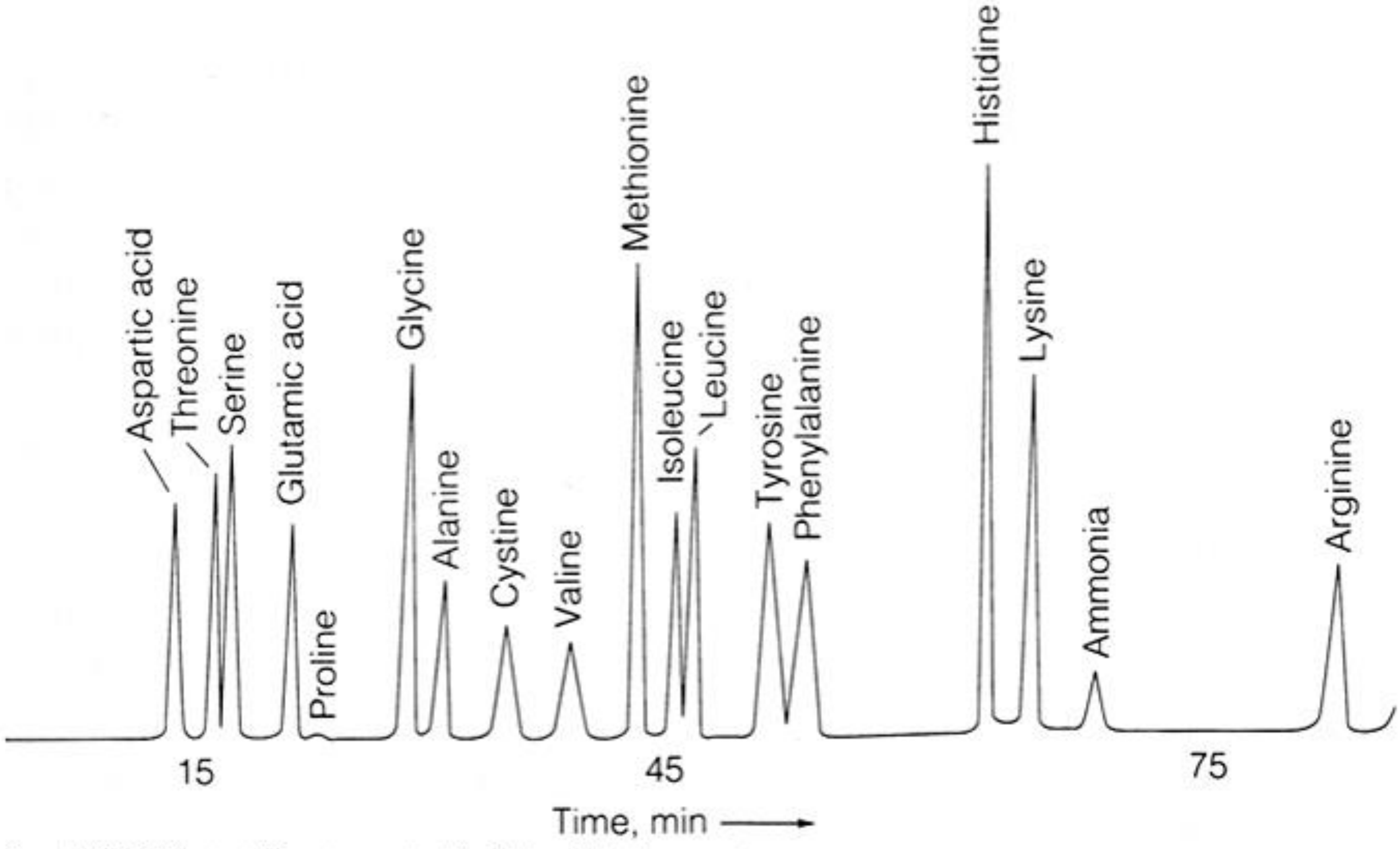
PTH = phenylthiohydantion

Identification of N-Terminal Residue



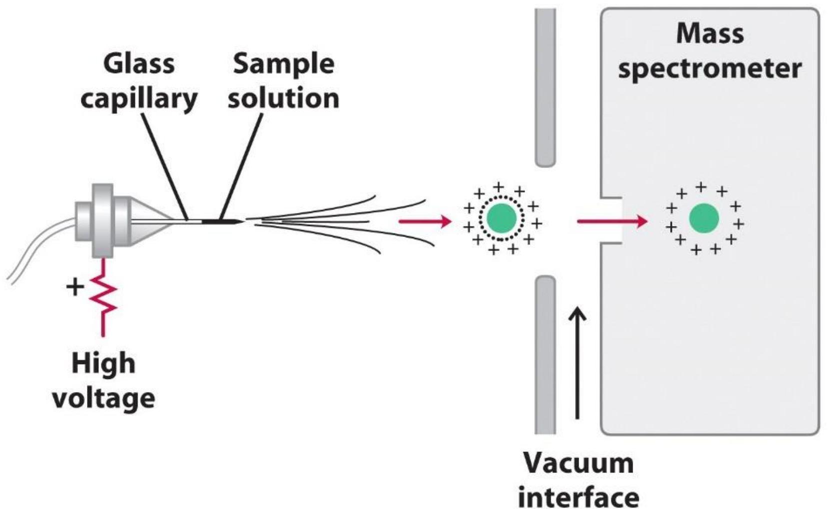
Note: Identification of C-terminal residue done by hydrazinolysis (reaction with anhydrous hydrazine in presence of mildly acidic ion exchange resin) or with a C-terminus-specific exopeptidase (carboxypeptidase).

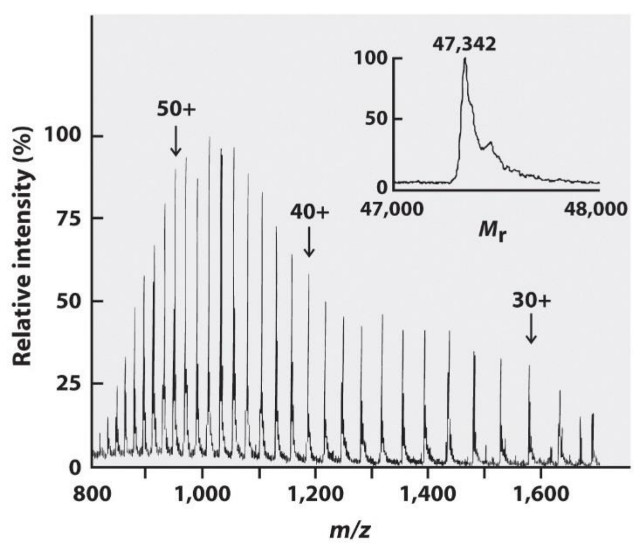
Separation of Amino Acids by HPLC

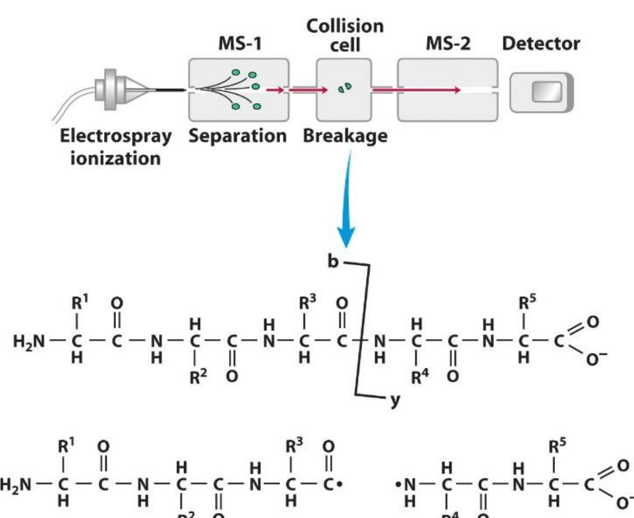


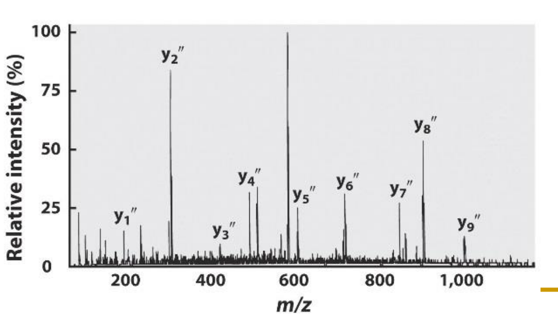
Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Protein Identification by Mass Spectrometry









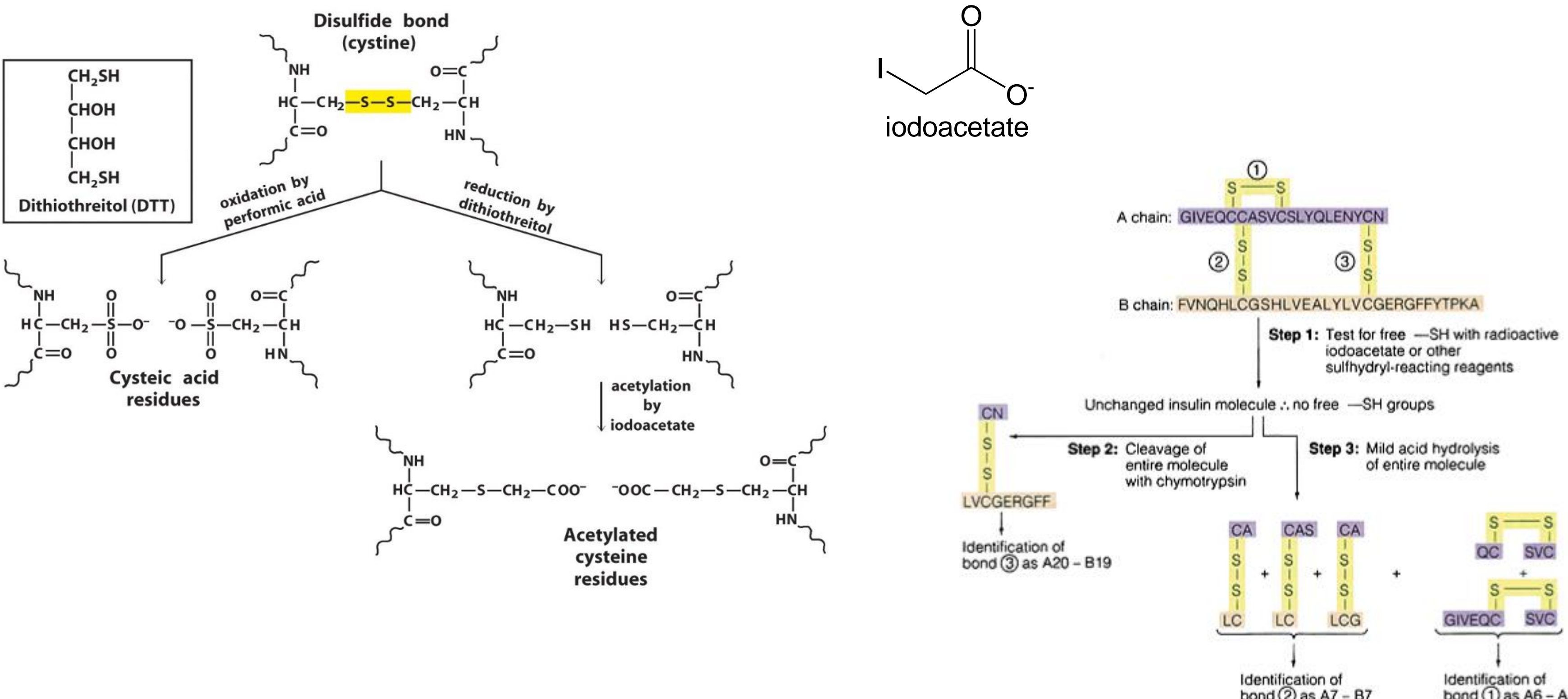
Protein Identification by Mass Spectrometry

Two main approaches:

1. Peptide mass fingerprinting: Proteolytic digestion of protein, then determination *m*/*z* of peptides by MS (*e.g.*, MALDI-TOF or ESI-TOF), search "fingerprint" against database. Success of ID depends on quality/ completeness of database for specific proteome.

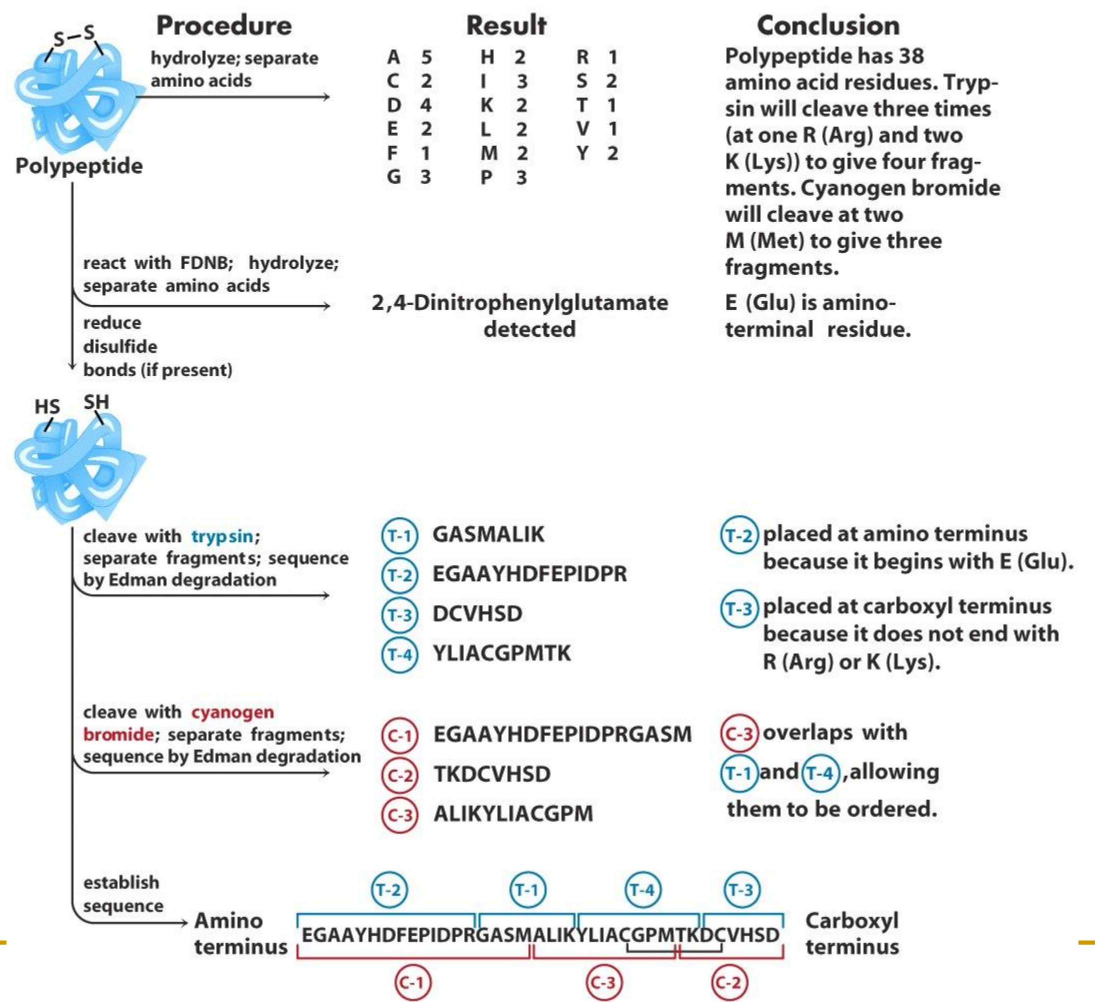
2. Tandem MS (MS/MS – *e.g.*, nanoLC-ESI-MS/MS): Proteolytic digestion of protein, separation and determination of *m*/*z* of each (MS-1), then determination of collision-induced dissociation fragment spectrum for each peptide (MS-2). Gives context/sequence-dependent information, so more of a *do novo* sequencing method.

Locating Disulfide Bonds



Copyright © 2000 Benjamin/Cummings, an imprint of Addison Wesley Longman, Inc.

Determining Primary Structure of an Entire Protein



Reactions in Solid-Phase Peptide Synthesis

Insoluble polystyrene bead

Amino acid 1 with α-amino group protected by Fmoc group

① Attachment of carboxyl-terminal amino acid to reactive group on resin.

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

Dicyclohexylcarbodiimide (DCC)

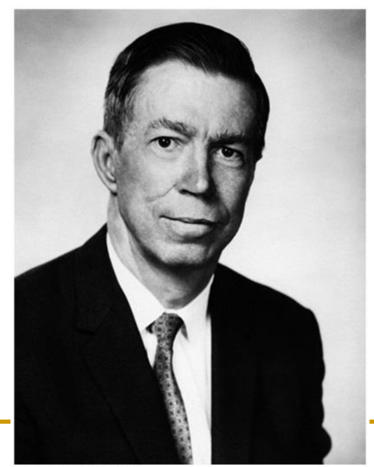
③ Amino acid 2 with protected α-amino group is activated at carboxyl group by DCC.

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

Dicyclohexylurea byproduct

Reactions ② to ④ repeated as necessary

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



R. Bruce Merrifield

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

Summary

1) Amino acids can be joined covalently through peptide bonds to form peptides and proteins. Cells generally contain thousands of different proteins, each with a different biological activity.

2. Proteins can be very long polypeptide chains of 100 to several thousand amino acid residues. However, some naturally occurring peptides have only a few amino acid residues. Some proteins are composed of several no covalently associated polypeptide chains, called subunits. Simple proteins yield only amino acids on hydrolysis; conjugated proteins contain in addition some other component, such as a metal or organic prosthetic group.

3. The sequence of amino acids in a protein is characteristic of that protein and is called its primary structure. This is one of four generally recognized levels of protein structure.