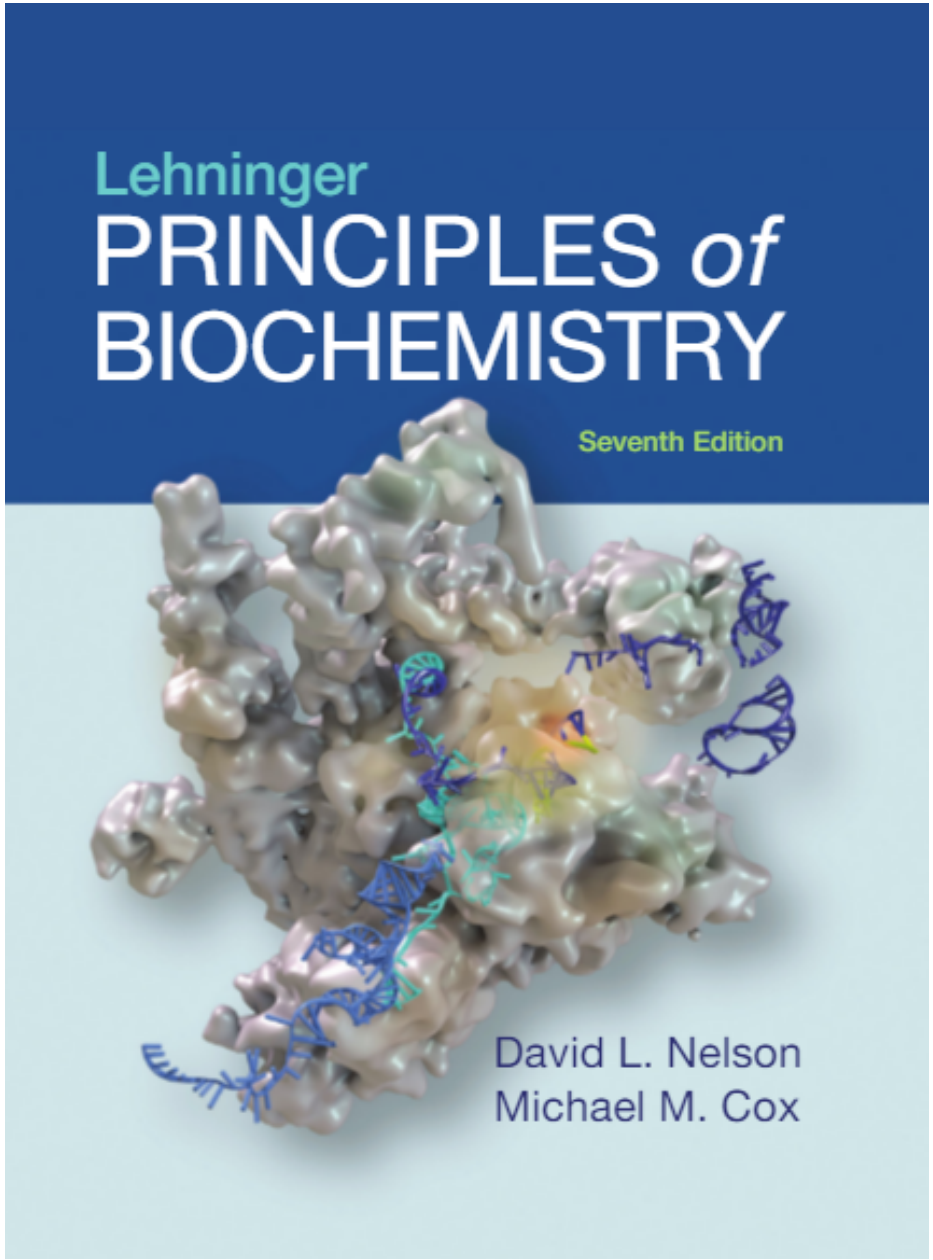
3 | Amino Acids, Peptides, and Proteins



CHAPTER 3
Amino Acids, Peptides, and Proteins

Learning goals:

- Structure and naming of amino acids
- Structure and properties of peptides
- Ionization behavior of amino acids and peptides
- Methods to characterize peptides and proteins

Proteins: Main Agents of Biological Function

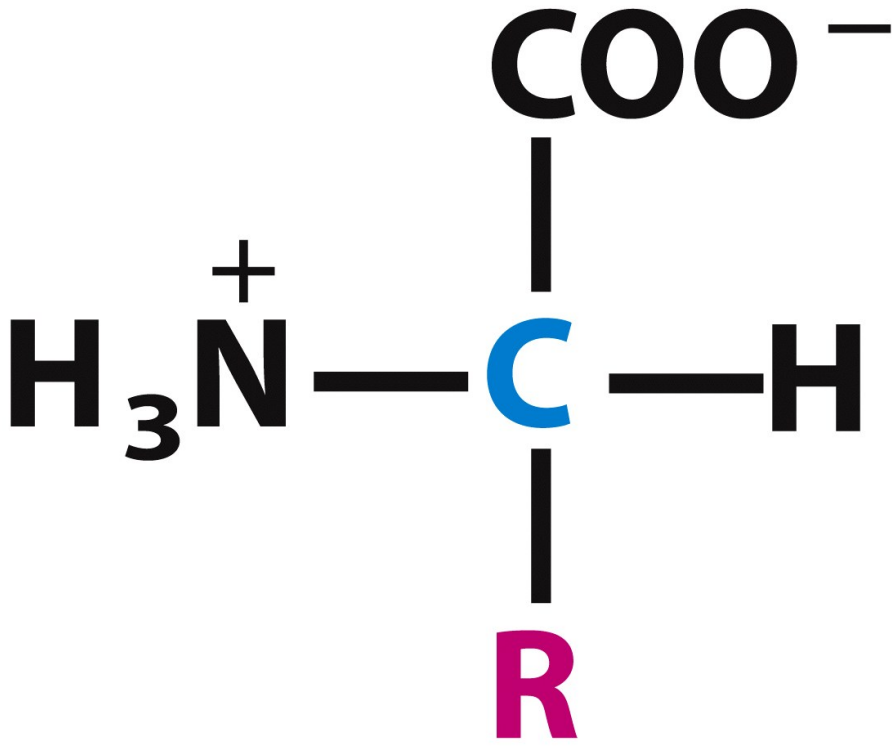
- Catalysis:
  - –enolase (in the glycolytic pathway)
  - –DNA polymerase (in DNA replication)
- Transport:
  - –hemoglobin (transports O_2 in the blood)
  - –lactose permease (transports lactose across the cell membrane)
- Structure:
  - –collagen (connective tissue)
  - –keratin (hair, nails, feathers, horns)
- Motion:
  - –myosin (muscle tissue)
  - –actin (muscle tissue, cell motility)

Figure 3-1c
Lehninger Principles of Biochemistry, Fifth Edition

Amino Acids: Building Blocks of Protein

- Proteins are heteropolymers of α-amino acids
- Amino acids have properties that are well suited to carry out a variety of biological functions:
  - Capacity to polymerize
  - Useful acid-base properties
  - Varied physical properties
  - Varied chemical functionality

Amino acids share many features, differing only at the R substituent



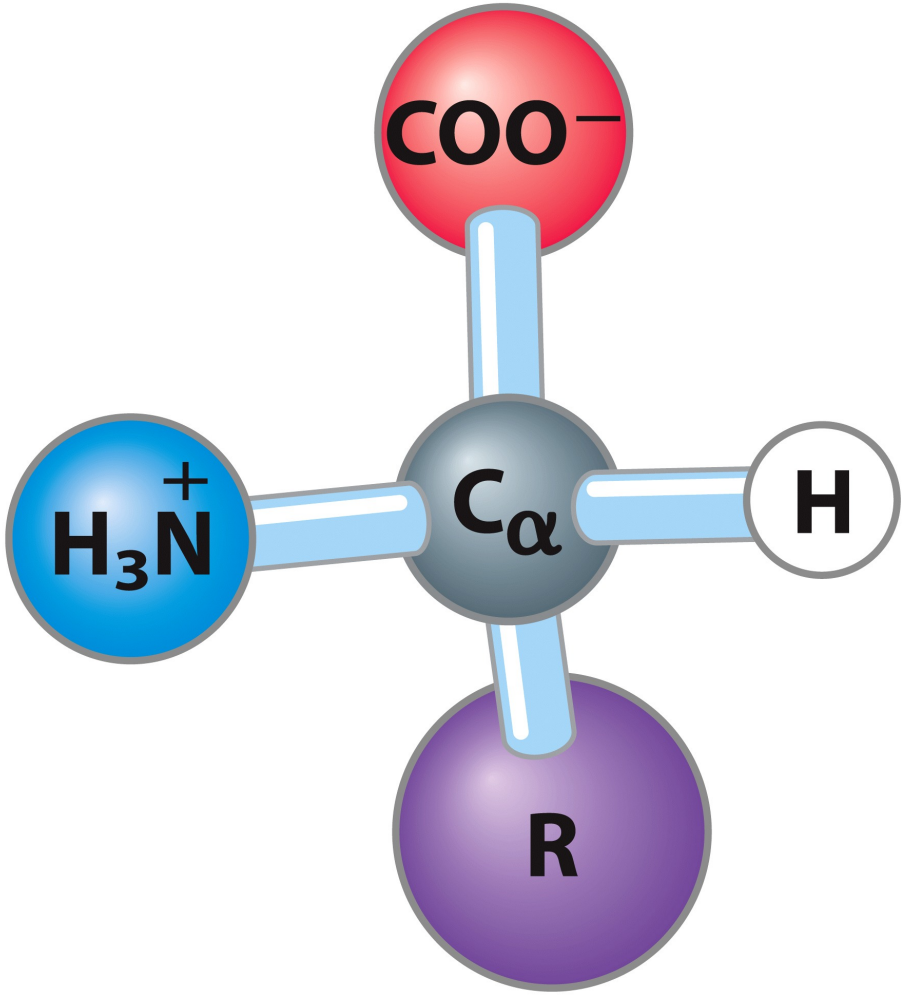
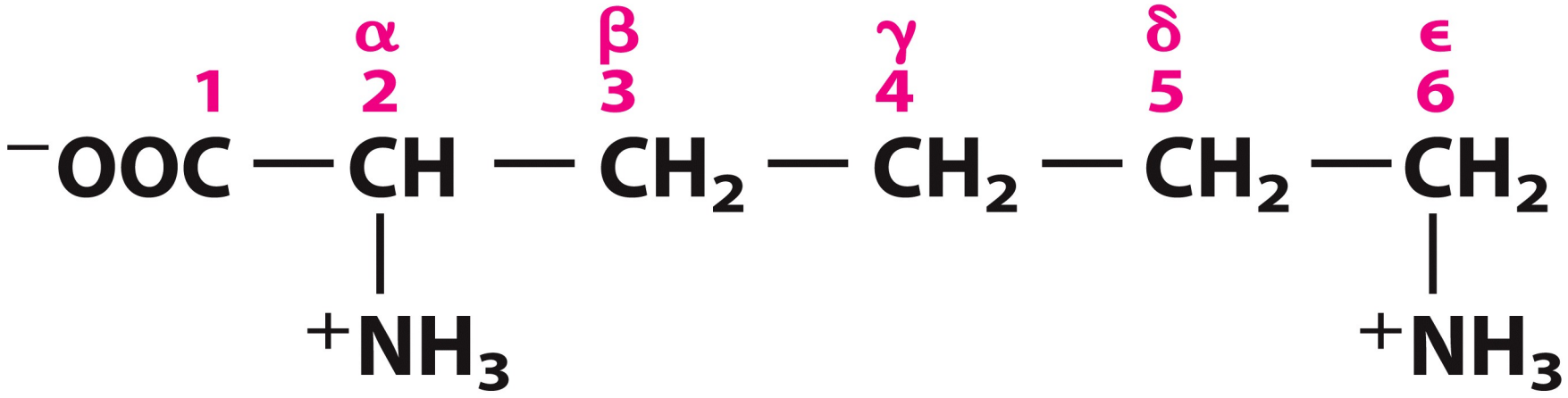


Figure 3-2
Lehninger Principles of Biochemistry, Sixth Edition

Amino Acids: Atom Naming

- Organic nomenclature: start from one end
- Biochemical designation:
  - start from α-carbon and go down the R-group



Lysine

Unnumbered 3 p78
Lehninger Principles of Biochemistry, Sixth Edition

Amino Acids Have Three Common Functional Groups Attached to the α Carbon

- The α carbon always has four substituents and is tetrahedral.
- All (except proline) have:
  - an acidic carboxyl group connected to the α carbon
  - a basic amino group connected to the α carbon
  - an α hydrogen connected to the α carbon
- The fourth substituent (R) is unique in glycine, the simplest amino acid. The fourth substituent is also hydrogen.

All Amino Acids Are Chiral (Except Glycine)

Proteins only contain L amino acids

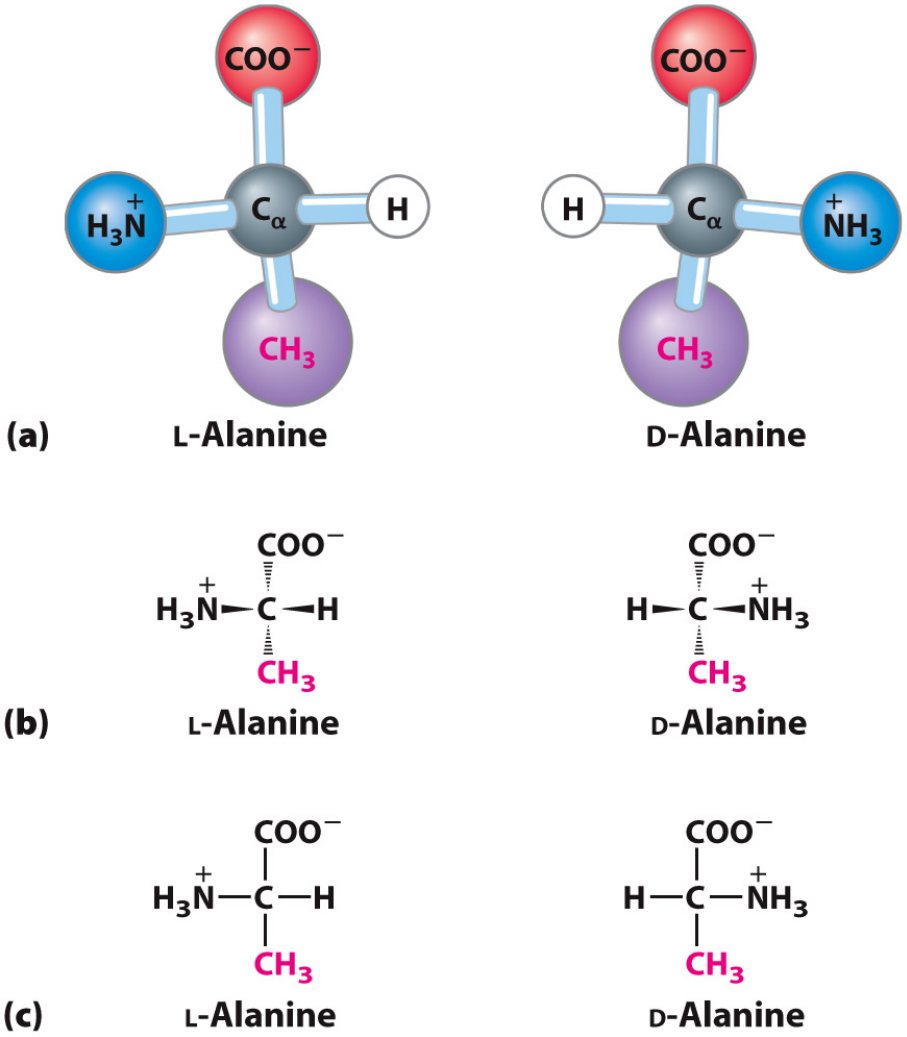


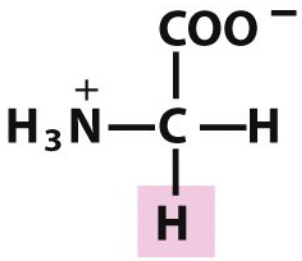
Figure 3-3
Lehninger Principles of Biochemistry, Seventh Edition

Amino Acids: Classification

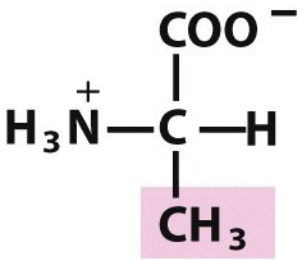
Common amino acids can be placed in five basic groups depending on their R substituents:

- Nonpolar, aliphatic (7)
- Aromatic (3)
- Positively charged (3)
- Negatively charged (2)
- Polar, uncharged (5)

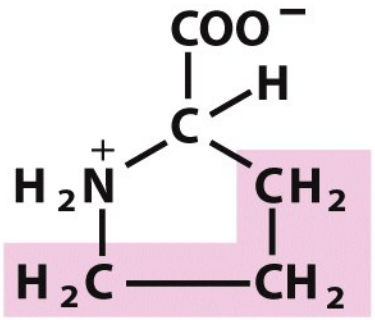
Nonpolar, aliphatic R groups



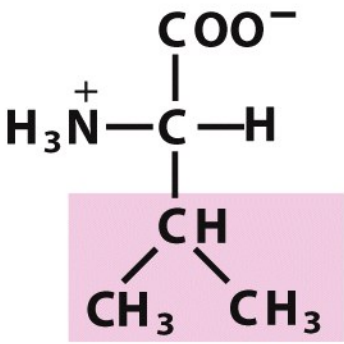
Glycine



Alanine

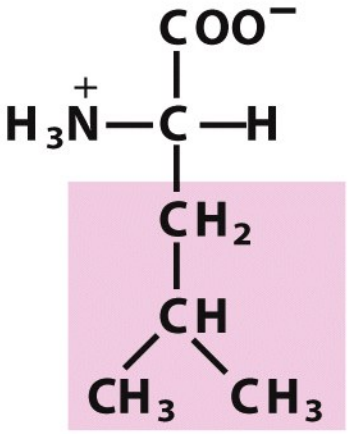


Proline

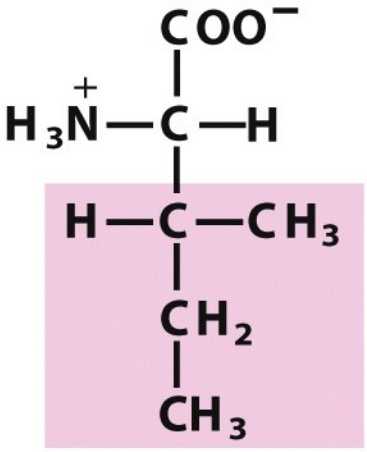


Valine

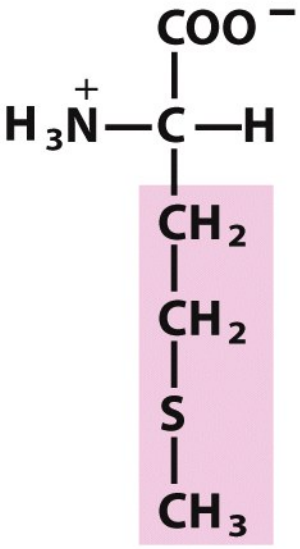
These amino acid side chains are ***hydrophobic***



Leucine



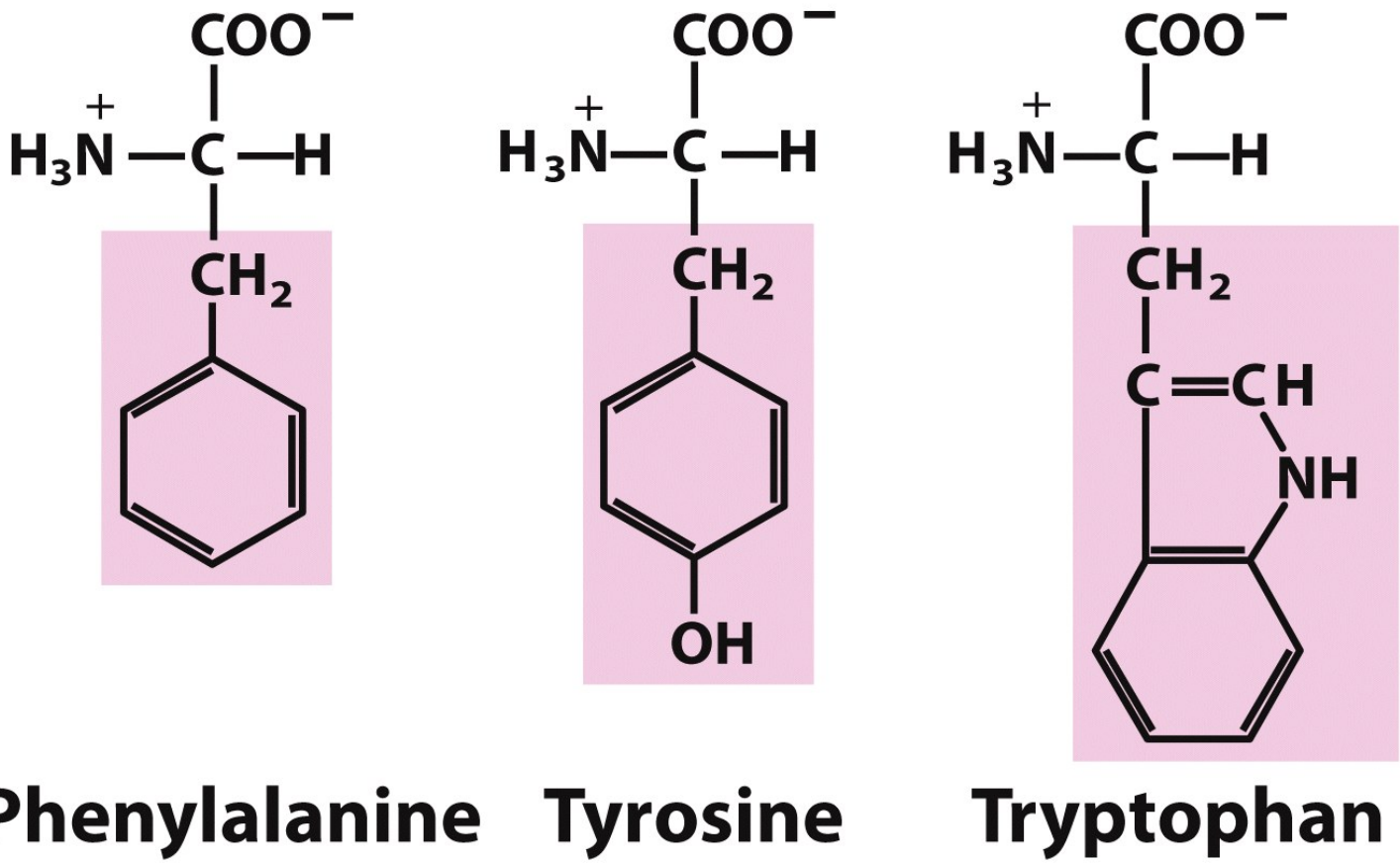
Isoleucine



Methionine

Figure 3-5 part 1
Lehninger Principles of Biochemistry, Fifth Edition

Aromatic R groups



Phenylalanine **Tyrosine** **Tryptophan**

Figure 3-5 part 2
Lehninger Principles of Biochemistry, Fifth Edition

These amino acid side chains ***absorb UV light at 270-280 nm***

Positively charged R groups

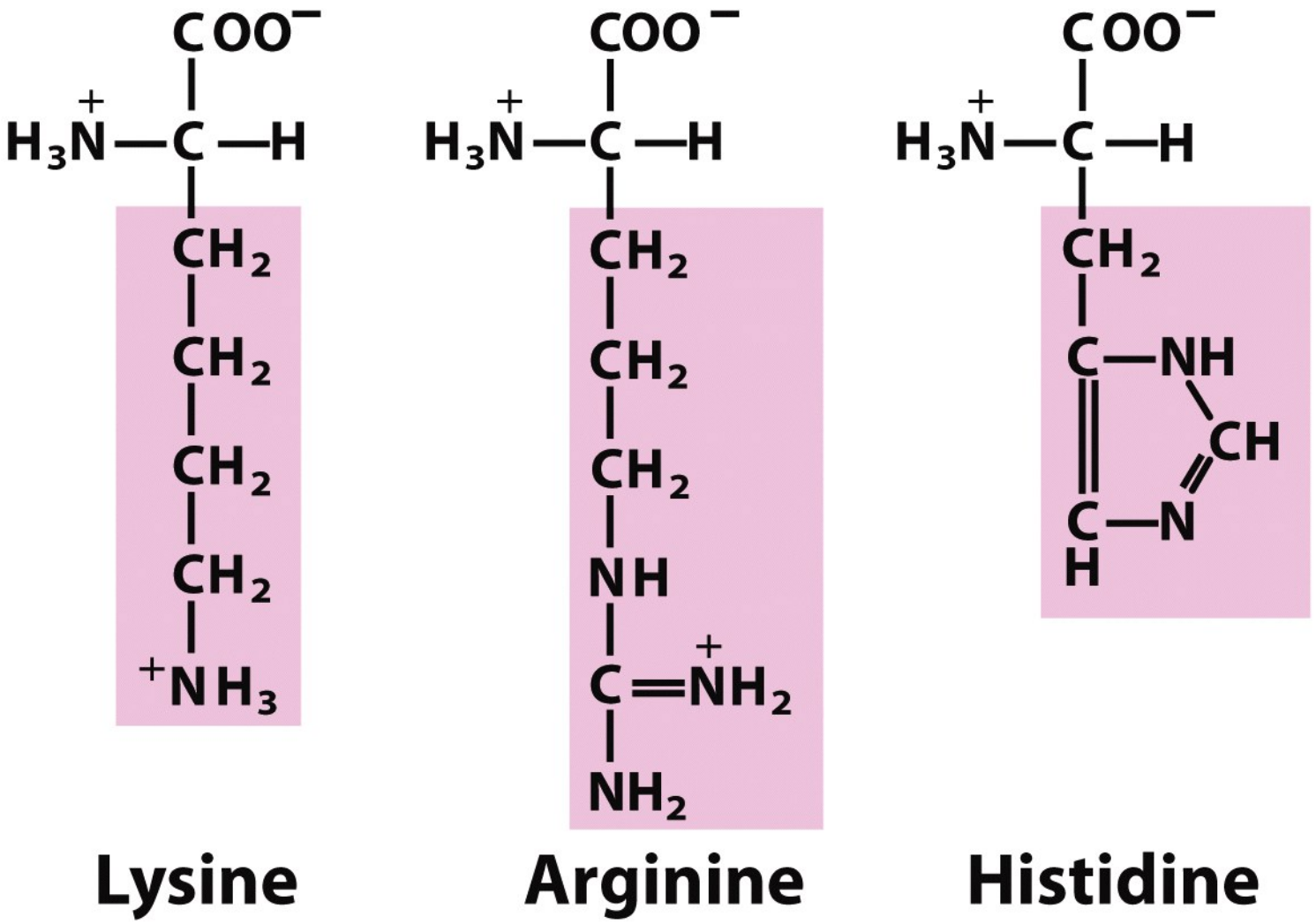
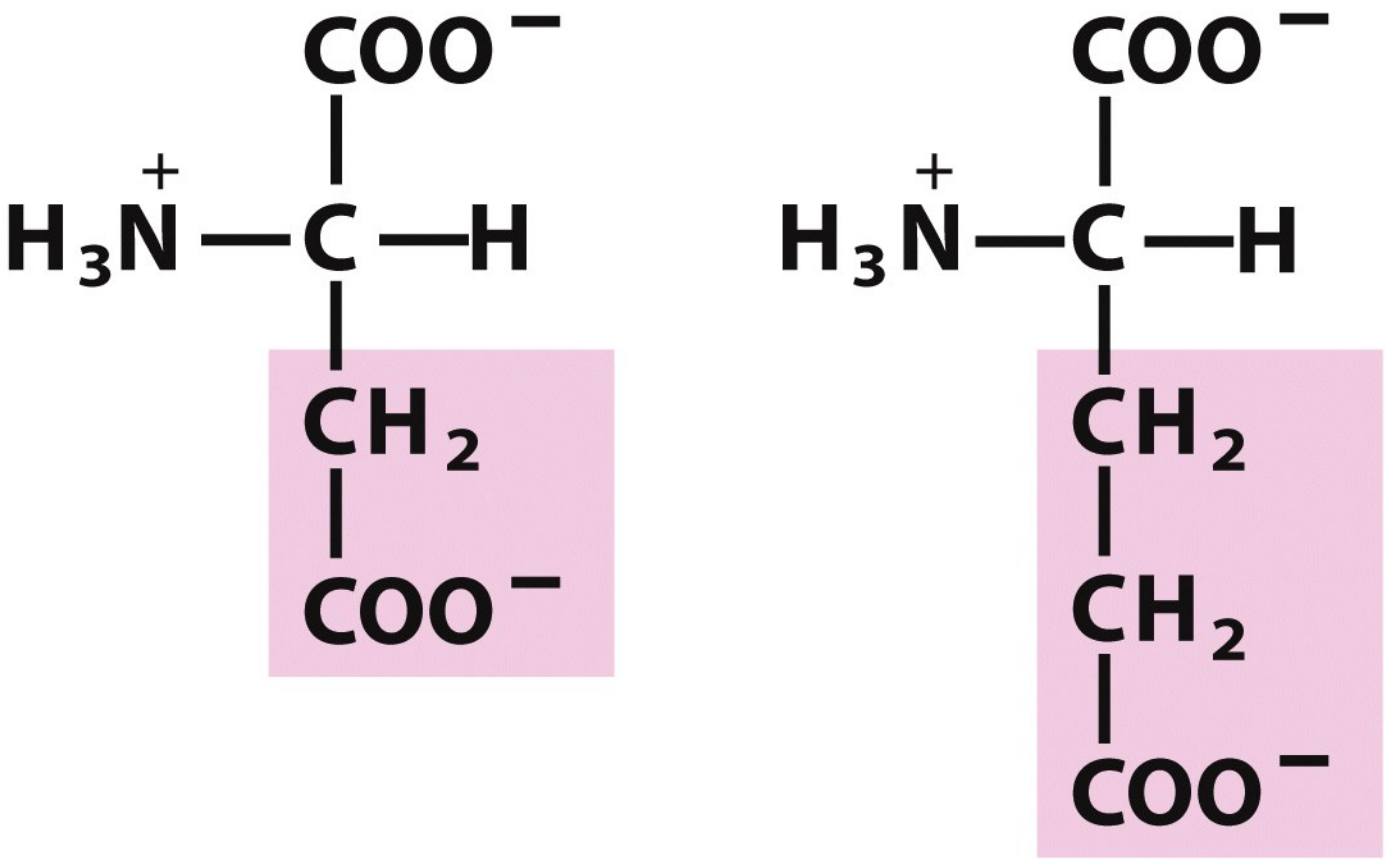


Figure 3-5 part 4
Lehninger Principles of Biochemistry, Fifth Edition

These are ***basic amino acids***

Negatively charged R groups

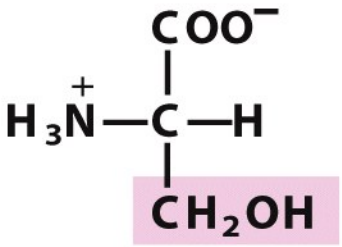


Aspartate **Glutamate**

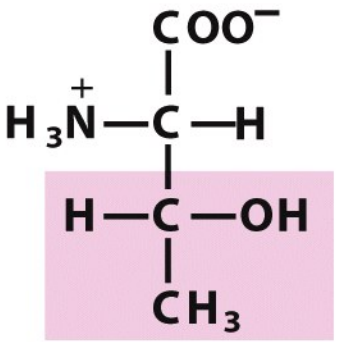
Figure 3-5 part 5
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W. H. Freeman and Company

These are ***acidic amino acids***

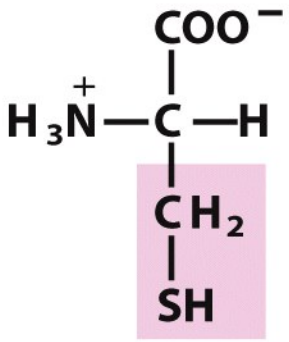
Polar, uncharged R groups



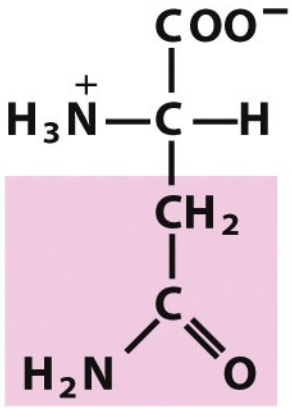
Serine



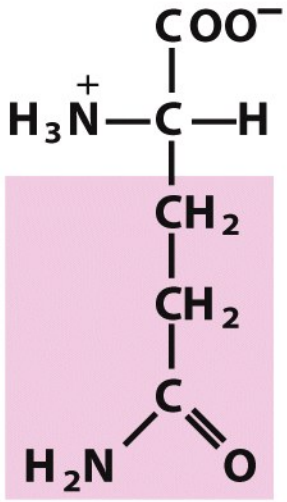
Threonine



Cysteine



Asparagine



Glutamine

These amino acids side chains can form ***hydrogen bonds***

Cysteine can form ***disulfide bonds***

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.

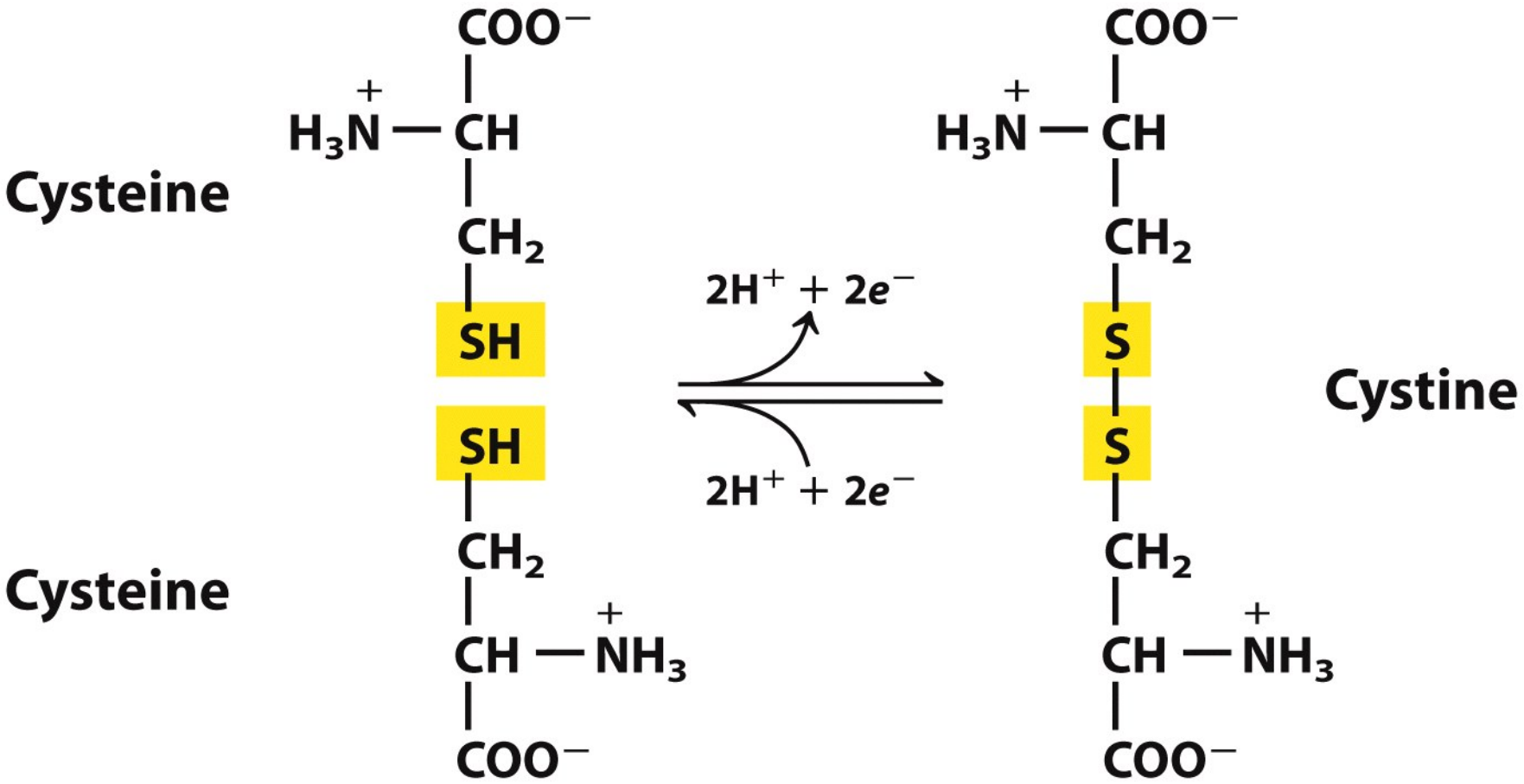
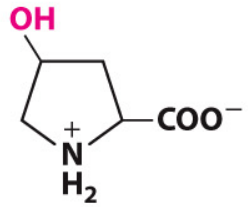


Figure 3-7
Lehninger Principles of Biochemistry, Fifth Edition

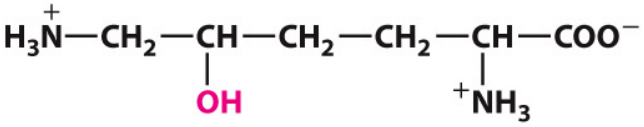
Uncommon Amino Acids in Proteins

- Not incorporated by ribosomes
  – except for selenocysteine
- Arise by posttranslational modifications of proteins
- Reversible modifications, especially phosphorylation, are important in regulation and signaling

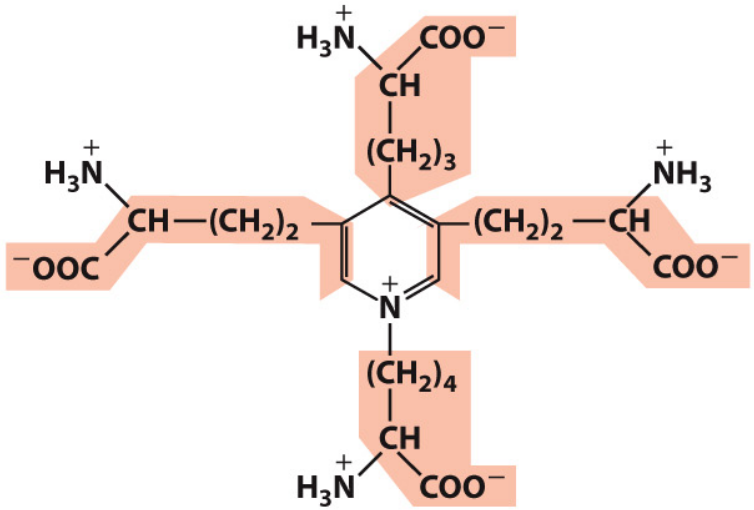
Modified Amino Acids Found in Proteins



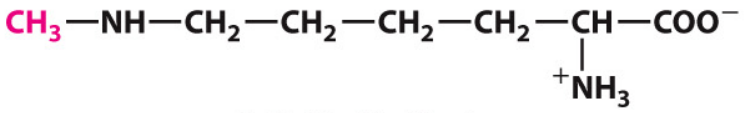
4-Hydroxyproline



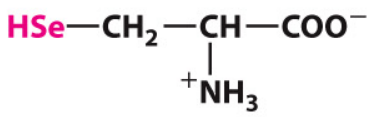
5-Hydroxylysine



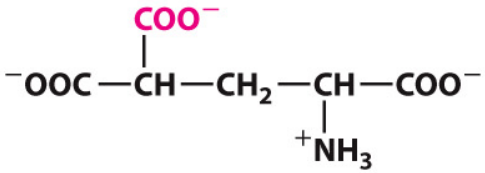
Desmosine



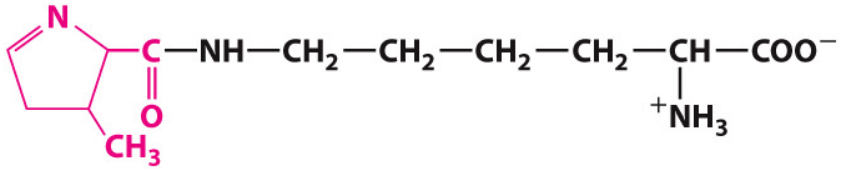
6-*N*-Methyllysine



Selenocysteine



γ-Carboxyglutamate

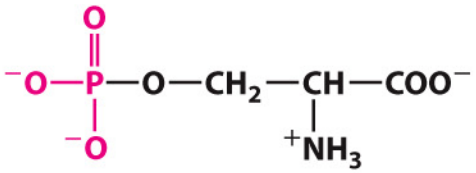


Pyrrolysine

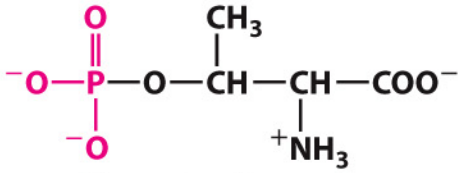
Figure 3-8a

Lehninger Principles of Biochemistry, Seventh Edition

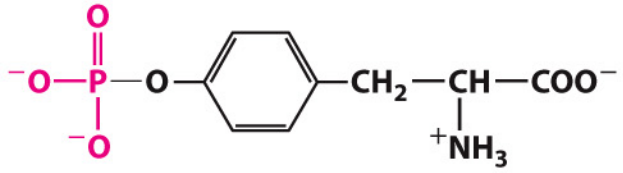
Reversible Modifications of Amino Acids



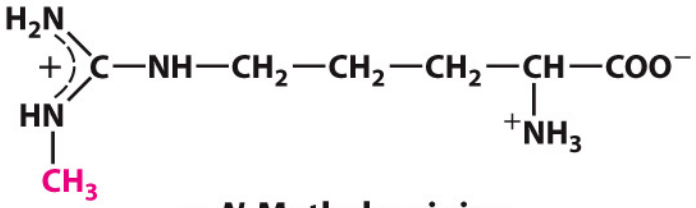
Phosphoserine



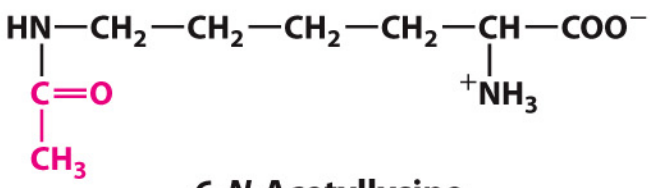
Phosphothreonine



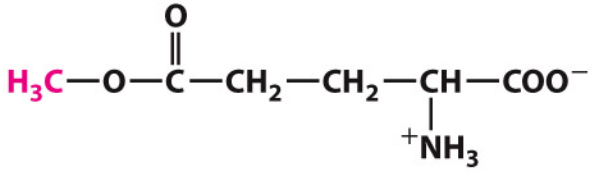
Phosphotyrosine



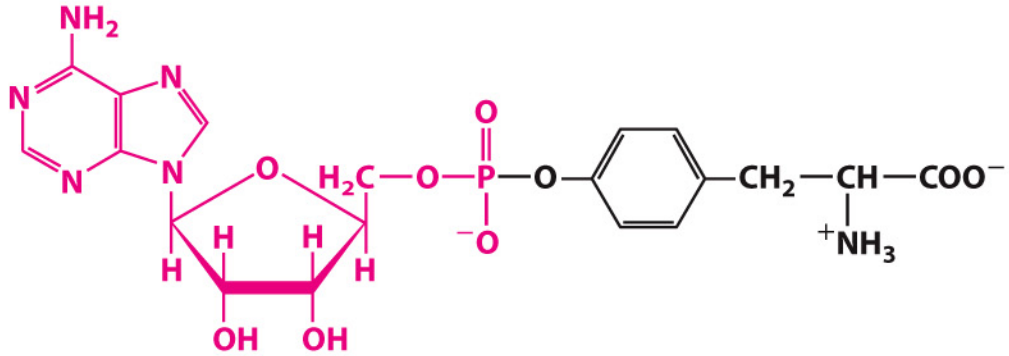
ω-*N*-Methylarginine



6-*N*-Acetyllysine



Glutamate γ-methyl ester

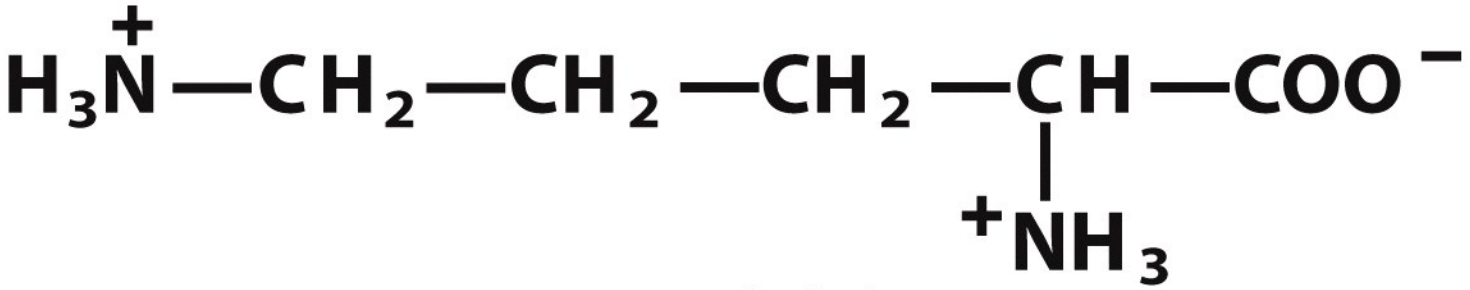


Adenylyltyrosine

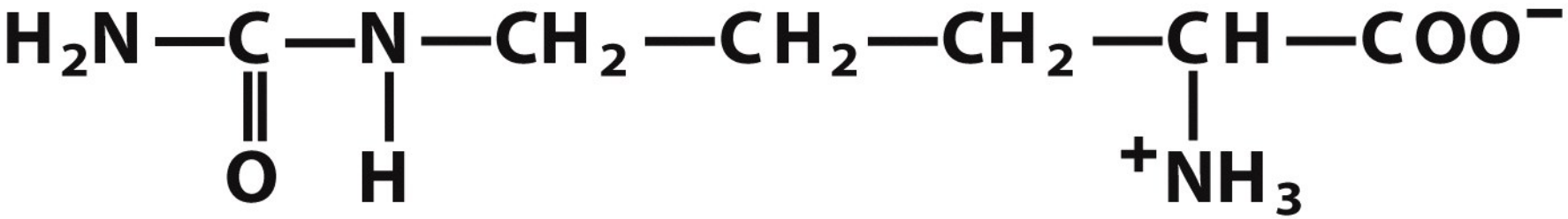
Figure 3-8b

Lehninger Principles of Biochemistry, Seventh Edition

Important Amino Acids in Urea Metabolism



Ornithine

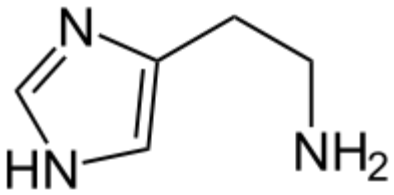


Citrulline

Figure 3-8c
Lehninger Principles of Biochemistry, Fifth Edition

A Derivative of amino acids

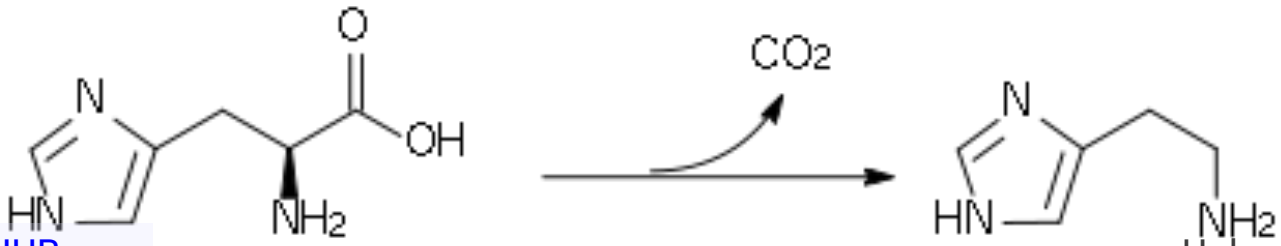
• **Histamine** is an organic nitrogen compound involved in local immune responses as well as regulating physiological function in the gut and acting as a neurotransmitter.



• Histamine triggers the inflammatory response

• Histamine increases the permeability of the capillaries to white blood cells and some proteins, to allow them to engage pathogens in the infected tissues

• Histamine is derived from the decarboxylation of the amino acid histidine, a reaction catalyzed by the enzyme *L-histidine decarboxylase*



Ionization of Amino Acids

- Amino acids contain at least two ionizable protons, each with its own pK_a.
- The carboxylic acid has an acidic pK_a and will be protonated at an acidic (low) pH:

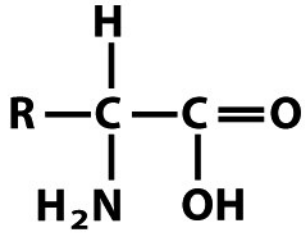
$$-COOH \leftrightarrow COO^- + H^+$$

- The amino group has a basic pK_a and will be protonated until basic pH (high) is achieved:

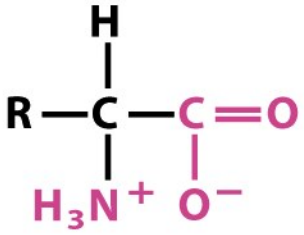
$$-NH_4^+ \leftrightarrow NH_3 + H^+$$

Ionization of Amino Acids

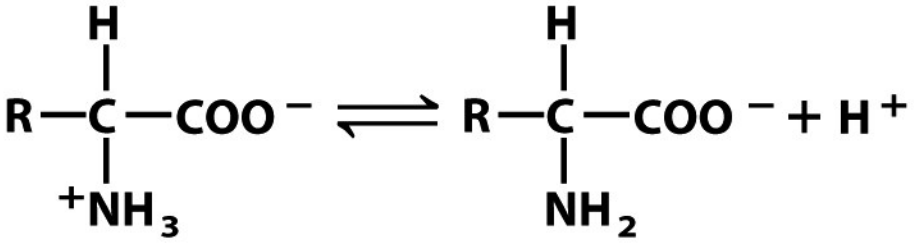
- At low (acidic) pH, the amino acid exists in a positively charged form (cation).
- At high (basic) pH, the amion acid exists in a negatively charged form (anion).
- Between the pK_a for each group, the amino acid exists in a zwitterion form, in which a single molecule has both a positive and negative charge.



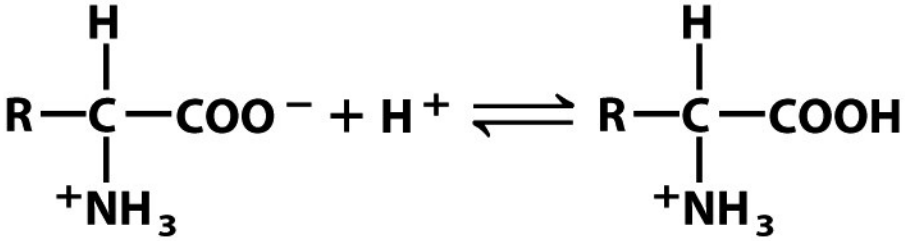
Nonionic form



Zwitterionic form



Zwitterion as acid



Zwitterion as base

Figure 3-9

Lehninger Principles of Biochemistry, Fifth Edition

Amino acids carry a net charge of zero at a specific pH (the pI)

- Zwitterions predominate at pH values between the pK_a values of the amino and carboxyl groups
- For amino acids without ionizable side chains, the Isoelectric Point (equivalence point, pI) is

$$pI = \frac{pK_1 + pK_2}{2}$$

- At this point, the net charge is zero
  - AA is least soluble in water
  - AA does not migrate in electric field

Cation → Zwitterion → Anion

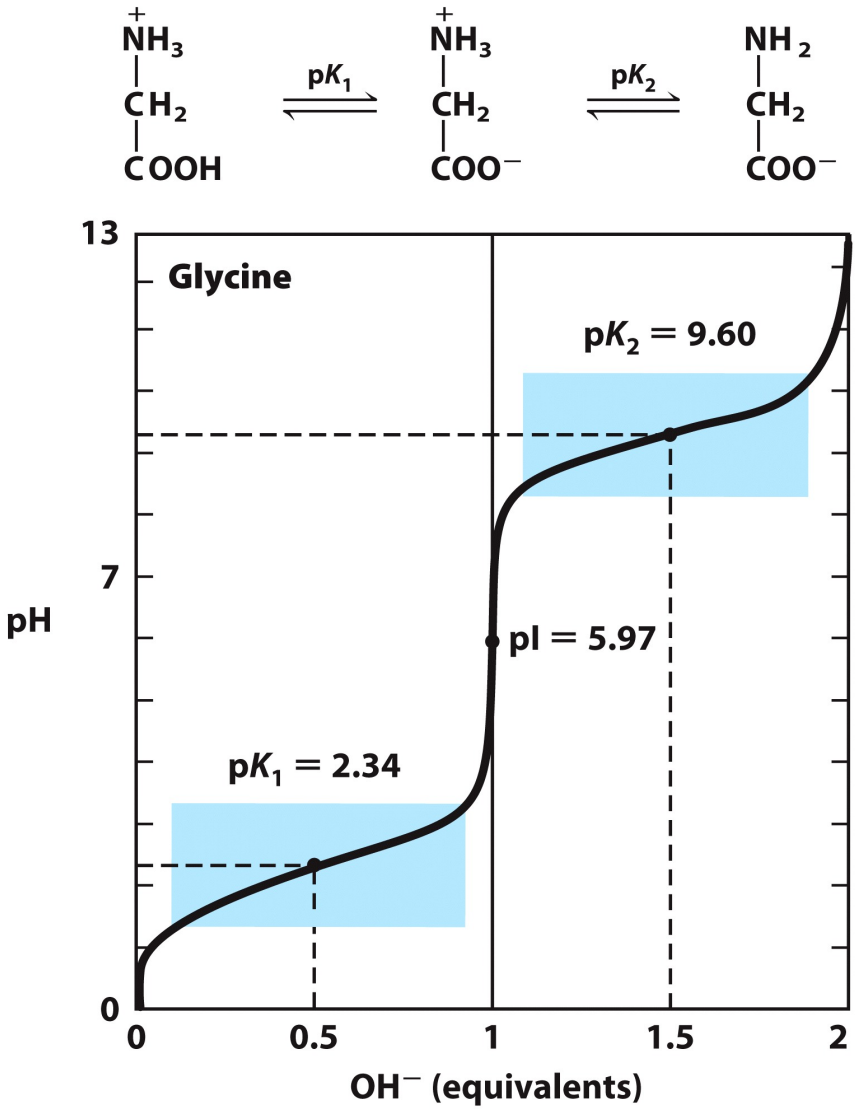


Figure 3-10
Lehninger Principles of Biochemistry, Sixth Edition

Chemical Environment Affects pK_a Values

α-carboxy group is much more acidic than in carboxylic acids

α-amino group is slightly less basic than in amines

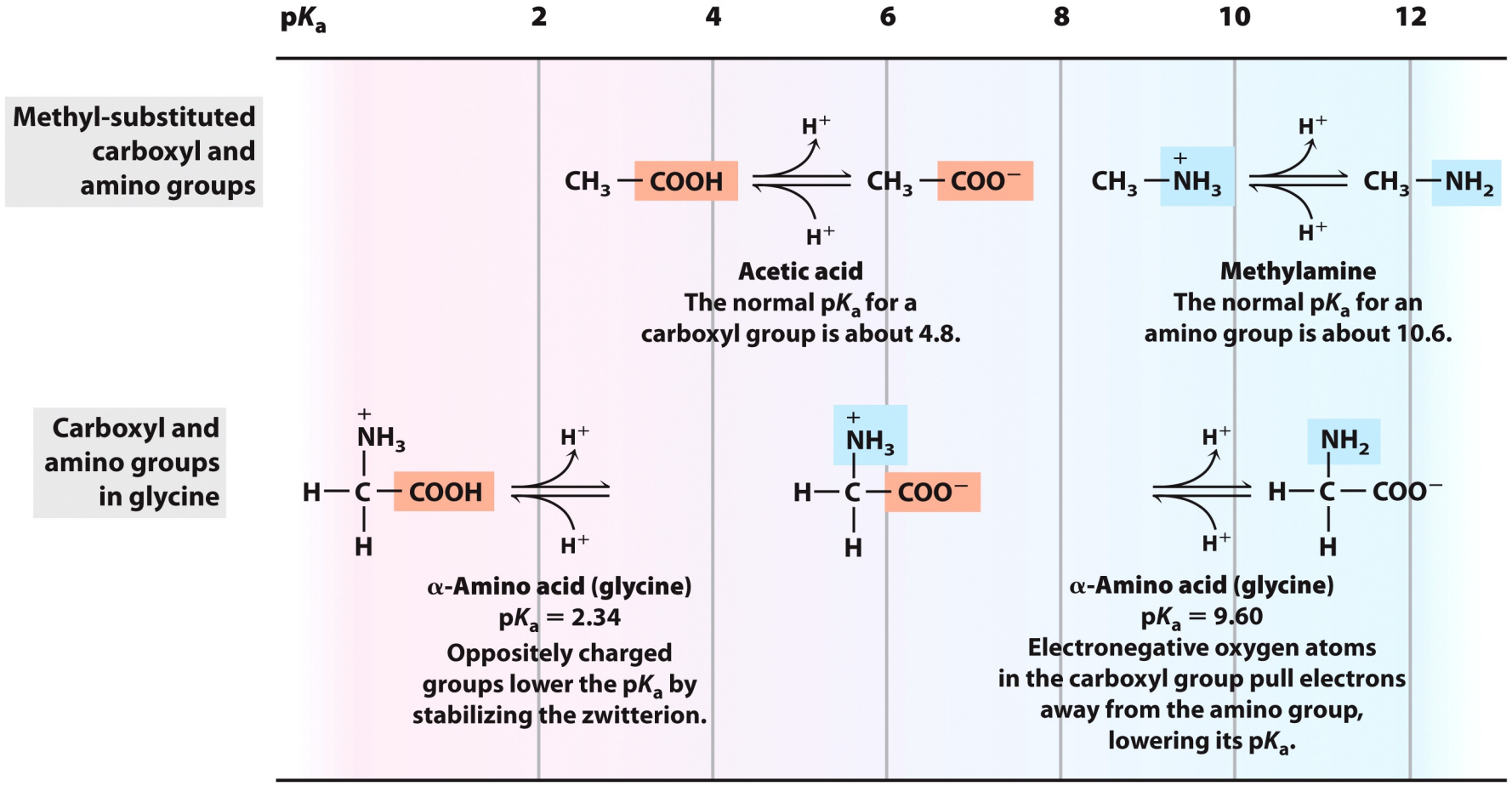


Figure 3-11

Lehninger Principles of Biochemistry, Sixth Edition

Amino Acids Can Act as Buffers

Amino acids with uncharged side chains, such as glycine, have two pK_a values:

- The pK_a of the α-carboxyl group is 2.34.
- The pK_a of the α-amino group is 9.6.

As buffers prevent change in pH close to the pK_a, glycine can act as a buffer in two pH ranges.

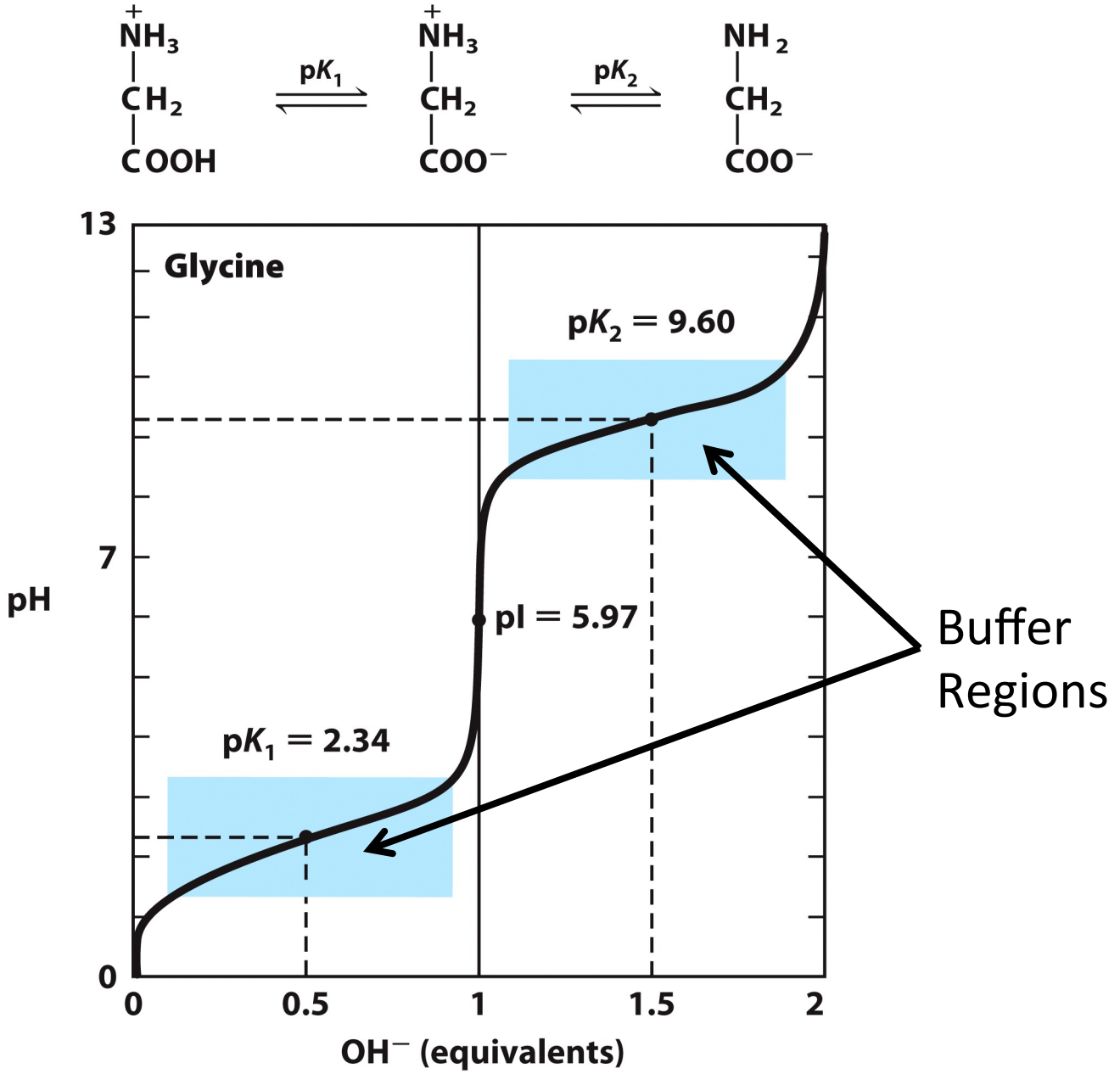


Figure 3-10
Lehninger Principles of Biochemistry, Sixth Edition

Ionizable Side Chains Also Have pK_a and Act as Buffers

- Ionizable side chains influence the pI of the amino acid.
- Ionizable side chains can be also titrated.
- Titration curves are now more complex, as each pK_a has a buffering zone of 2 pH units.

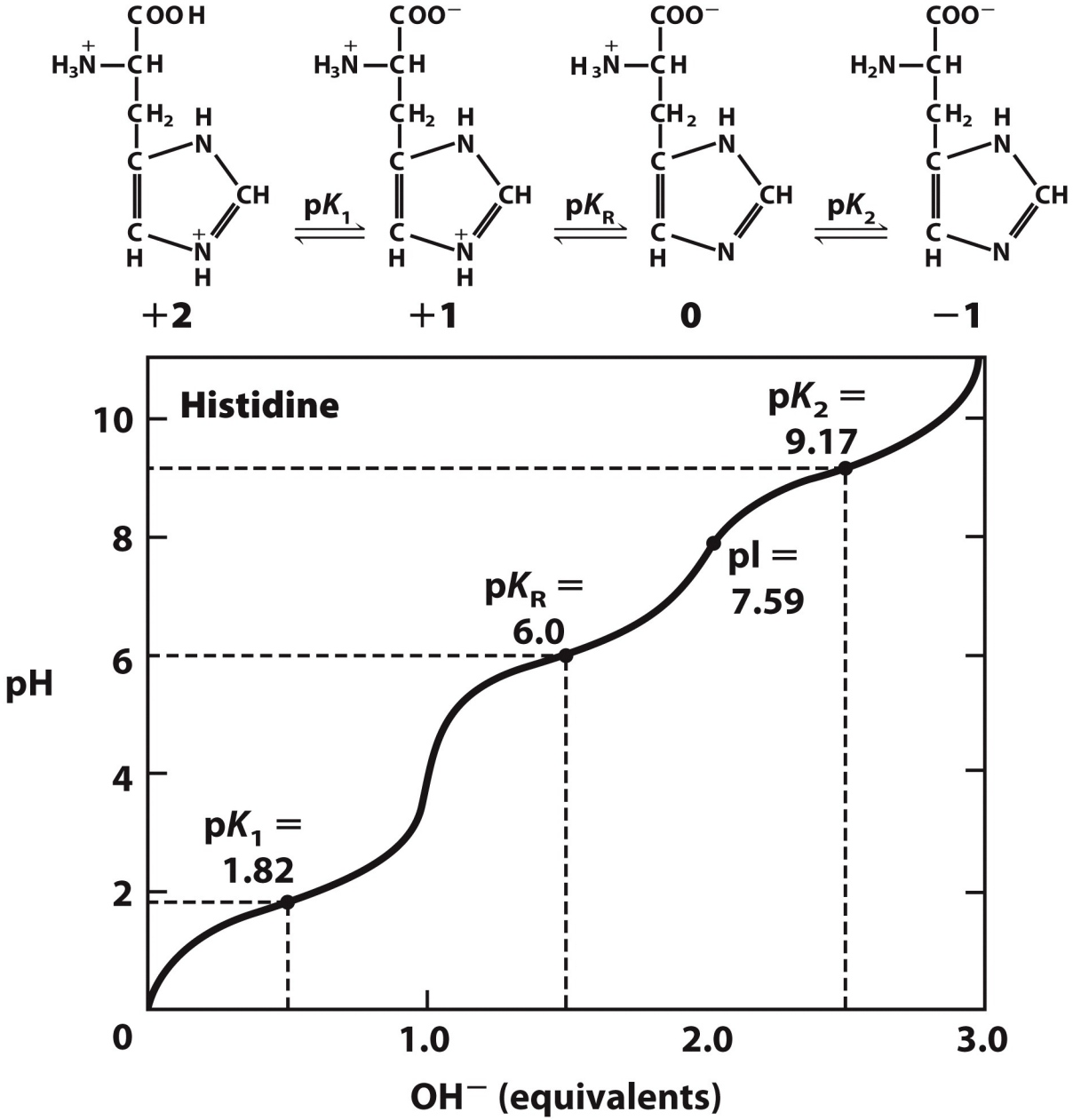


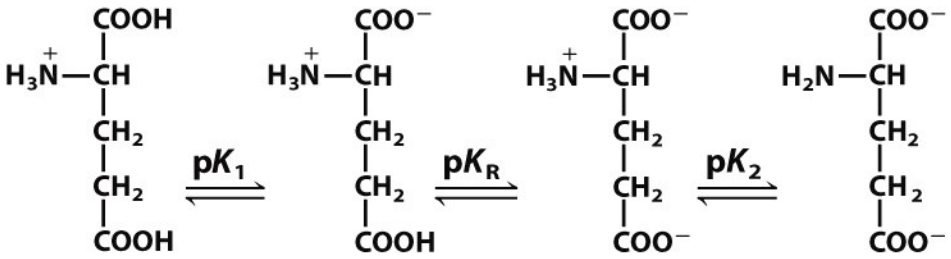
Figure 3-12b
Lehninger Principles of Biochemistry, Sixth Edition

How to Calculate the pI When the Side Chain is Ionizable

- At the pI, *the net charge of the molecule is zero*
- Identify species that carries a net zero charge
- Identify the species on either side of the neutral form (0 charge)
- Take average the two pK_a values

What is the pI of histidine?

- $(pK_R + pK_2)/2 = pI$
- $(6 + 9.17)/2 = 7.59$



What is the pI of glutamate?

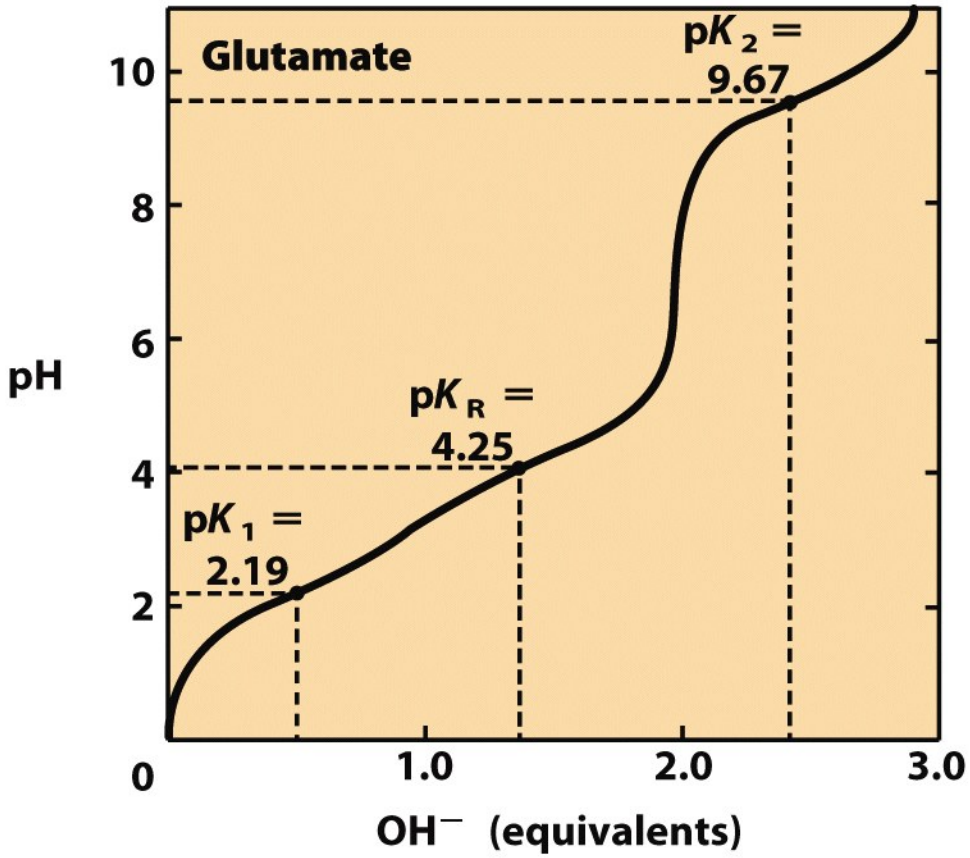


Figure 3-12a
Lehninger Principles of Biochemistry, Fifth Edition

Amino Acids Polymerize to Form Peptides

- Peptides are small condensation products of amino acids
- They are "small" compared to proteins (M_w < 10 kDa)

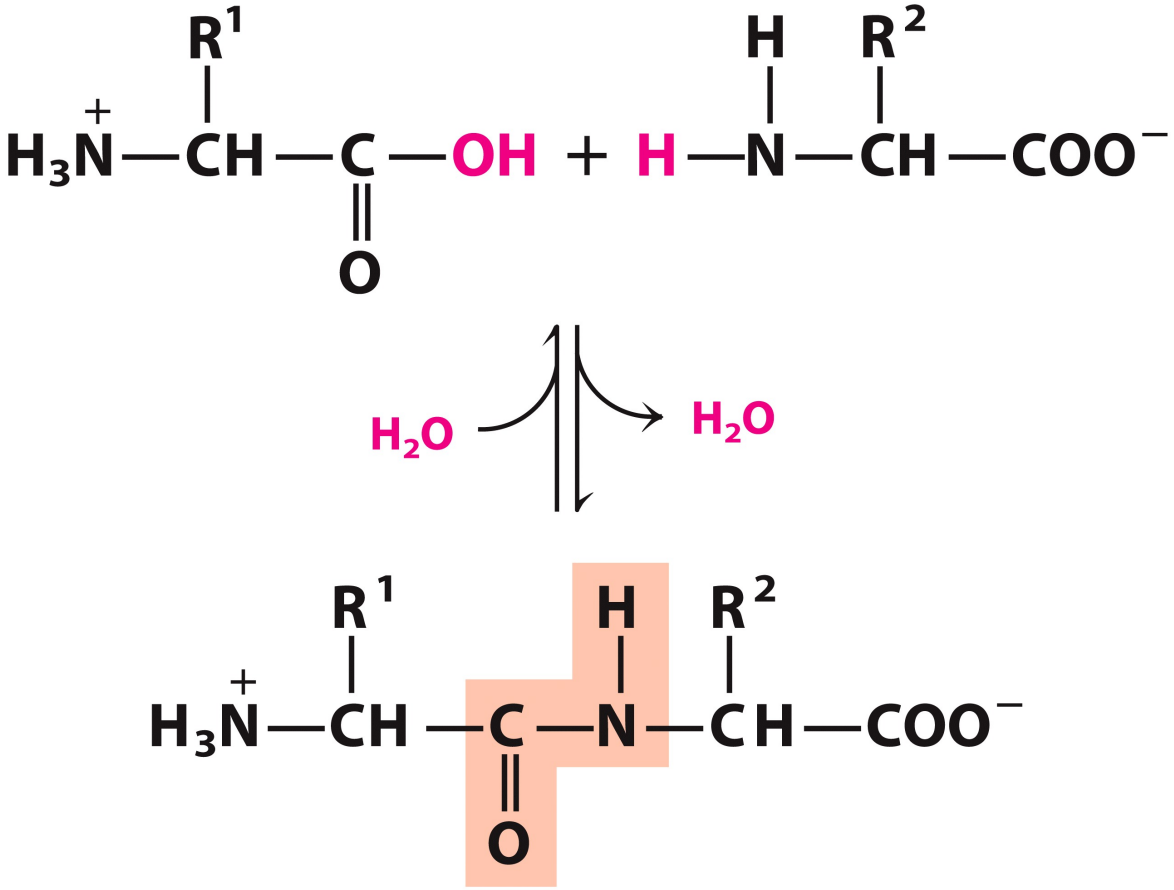


Figure 3-13
Lehninger Principles of Biochemistry, Sixth Edition
© 2013 W. H. Freeman and Company

Peptide ends are not the same

Numbering (and naming) starts from the amino terminus

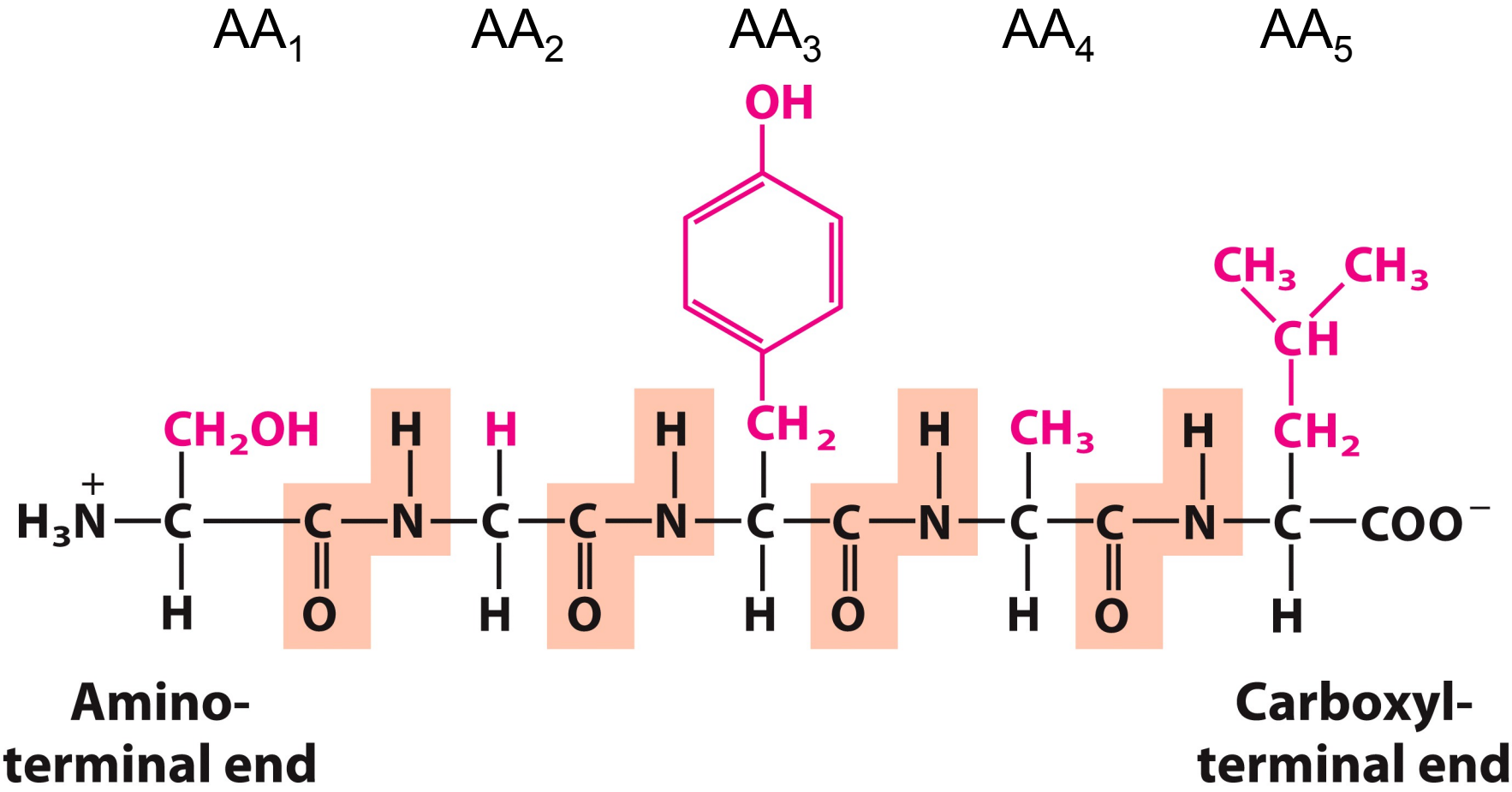
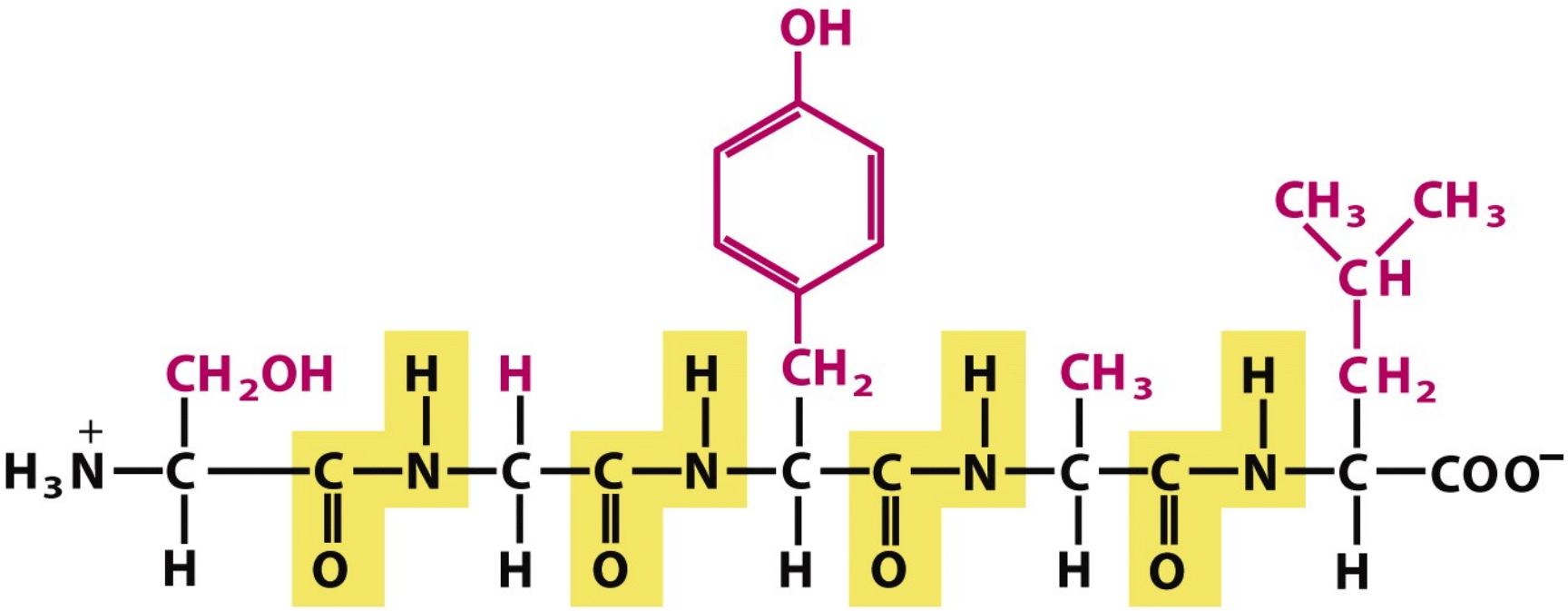


Figure 3-14
Lehninger Principles of Biochemistry, Sixth Edition

Naming peptides: start at the N-terminus

The pentapeptide **serylglycyltyrosylalanylleucine**,
Ser–Gly–Tyr–Ala–Leu, or **SGYAL**



Alanylglutamylglycyllysine

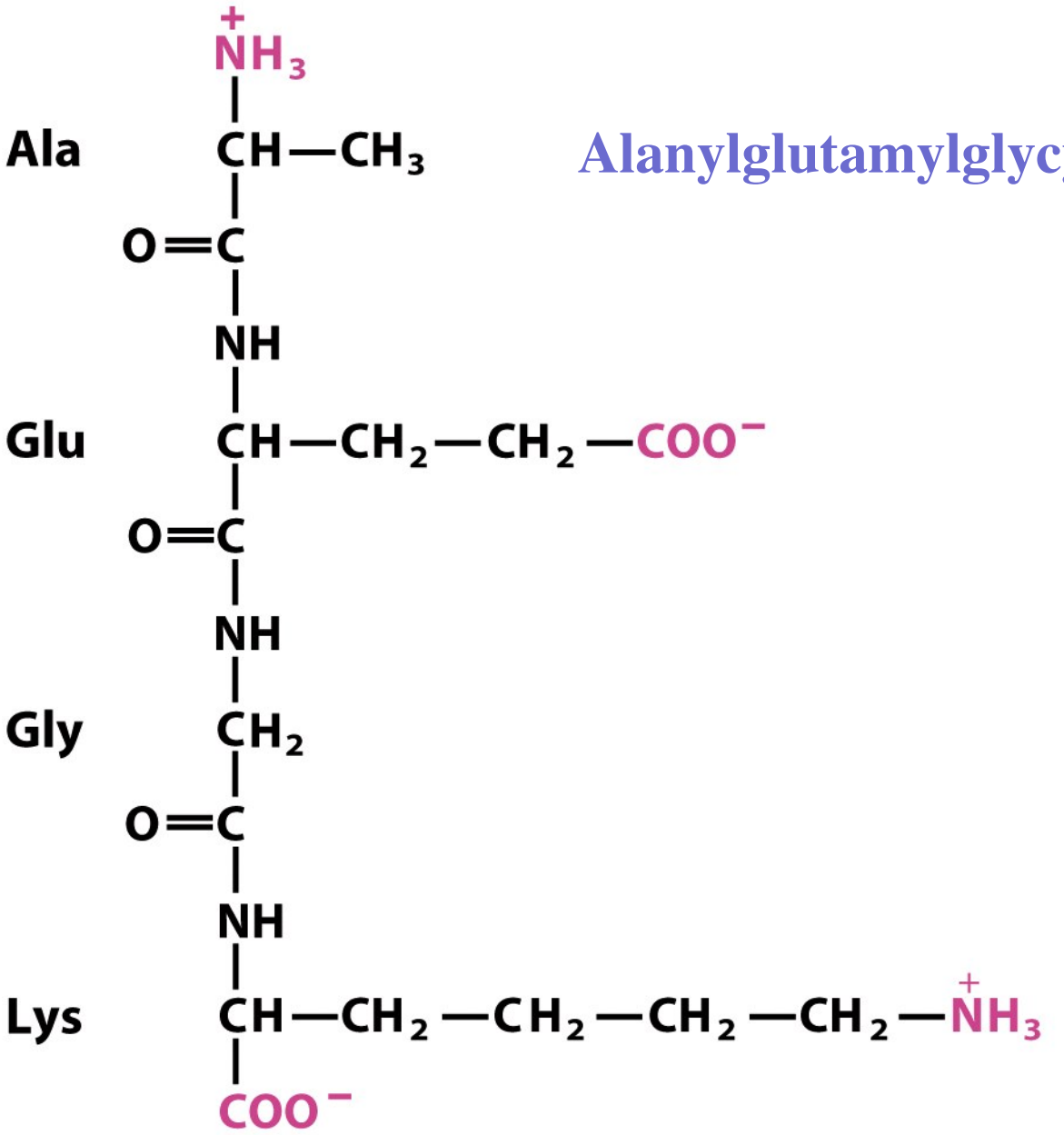
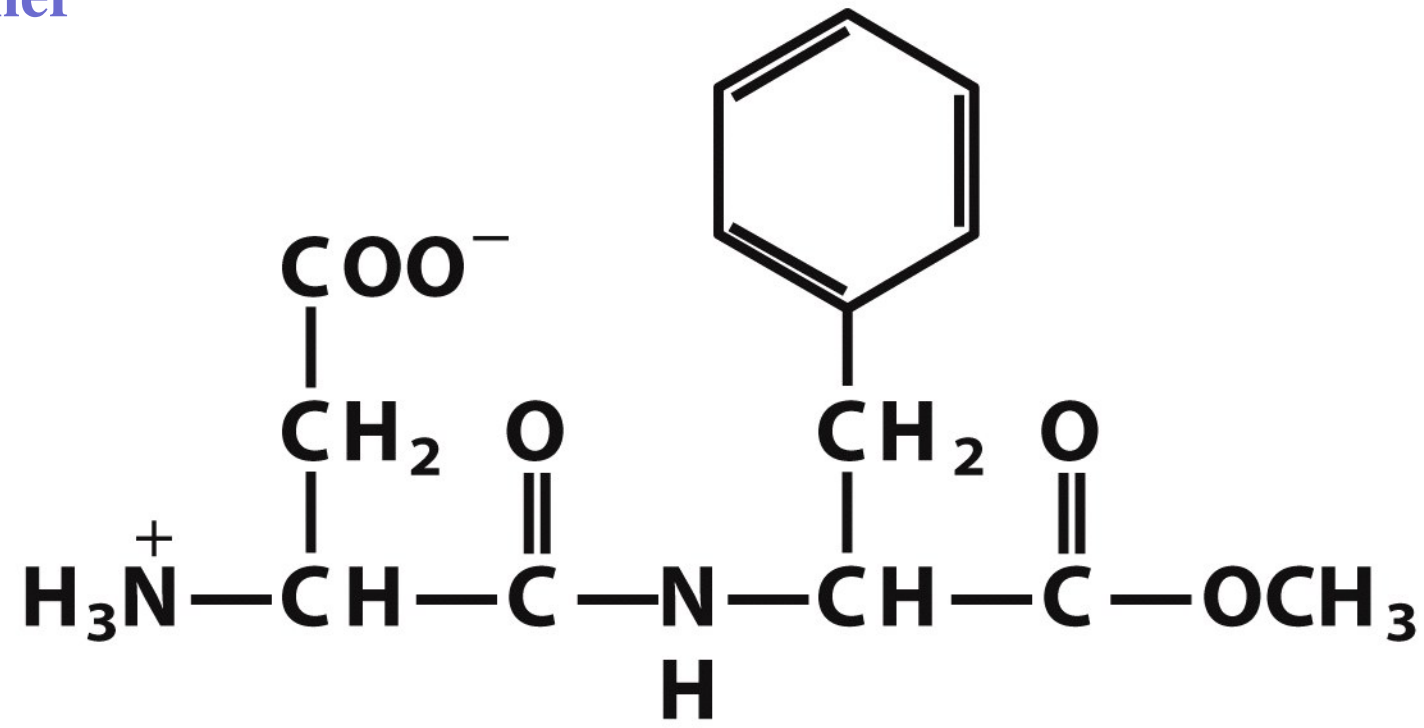


Figure 3-15
Lehninger Principles of Biochemistry, Fifth Edition

Peptides: A Variety of Functions

- Hormones and pheromones:
  - insulin (sugar uptake)
  - oxytocin (childbirth)
  - sex-peptide (fruit fly mating)
- Neuropeptides
  - substance P (pain mediator)
- Antibiotics:
  - polymyxin B (for Gram - bacteria)
  - bacitracin (for Gram + bacteria)
- Protection, e.g. toxins
  - amanitin (mushrooms)
  - conotoxin (cone snails)
  - chlorotoxin (scorpions)

Artificial sweetener



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

Unnumbered 3 p83
Lehninger Principles of Biochemistry, Fifth Edition

Proteins are comprised of:

- Polypeptides (covalently linked α-amino acids) + possibly:
  - **cofactors**
    - functional non-amino acid component
    - metal ions or organic molecules
  - **coenzymes**
    - organic cofactors
    - NAD^+ in lactate dehydrogenase
  - **prosthetic groups**
    - covalently attached cofactors
    - heme in myoglobin
  - **other modifications** (posttranslational modifications)

Polypeptide Size and Number Varies Greatly in Proteins

TABLE 3-2 Molecular Data on Some Proteins

Protein	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
Myglobin (equine heart)	16,700	153	1
Chymotrypsin (bovine pancreas)	25,700	245	1
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

Estimation of protein MW: 110 x aa = MW in Da

Conjugated Proteins Are Covalently Bound to a Nonprotein Entity

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

Studying Proteins and Peptides Sometimes Requires Purification from a Mixture

- Polypeptides contain differing amino acid sequences.
- The sequence and arrangement of amino acids gives the polypeptide a chemical character (i.e., charged, polar, hydrophobic, etc.).
- Some polypeptides bind specific targets, which can be used to “fish them out” of a complex mixture.

Proteins Can Be Separated and Purified

- Protein source is normally a tissue or cells

1. Open these cells releasing their proteins into solution (crude extract)
2. Differential centrifugation
3. Once extract or organelle fraction is ready, many techniques can be used

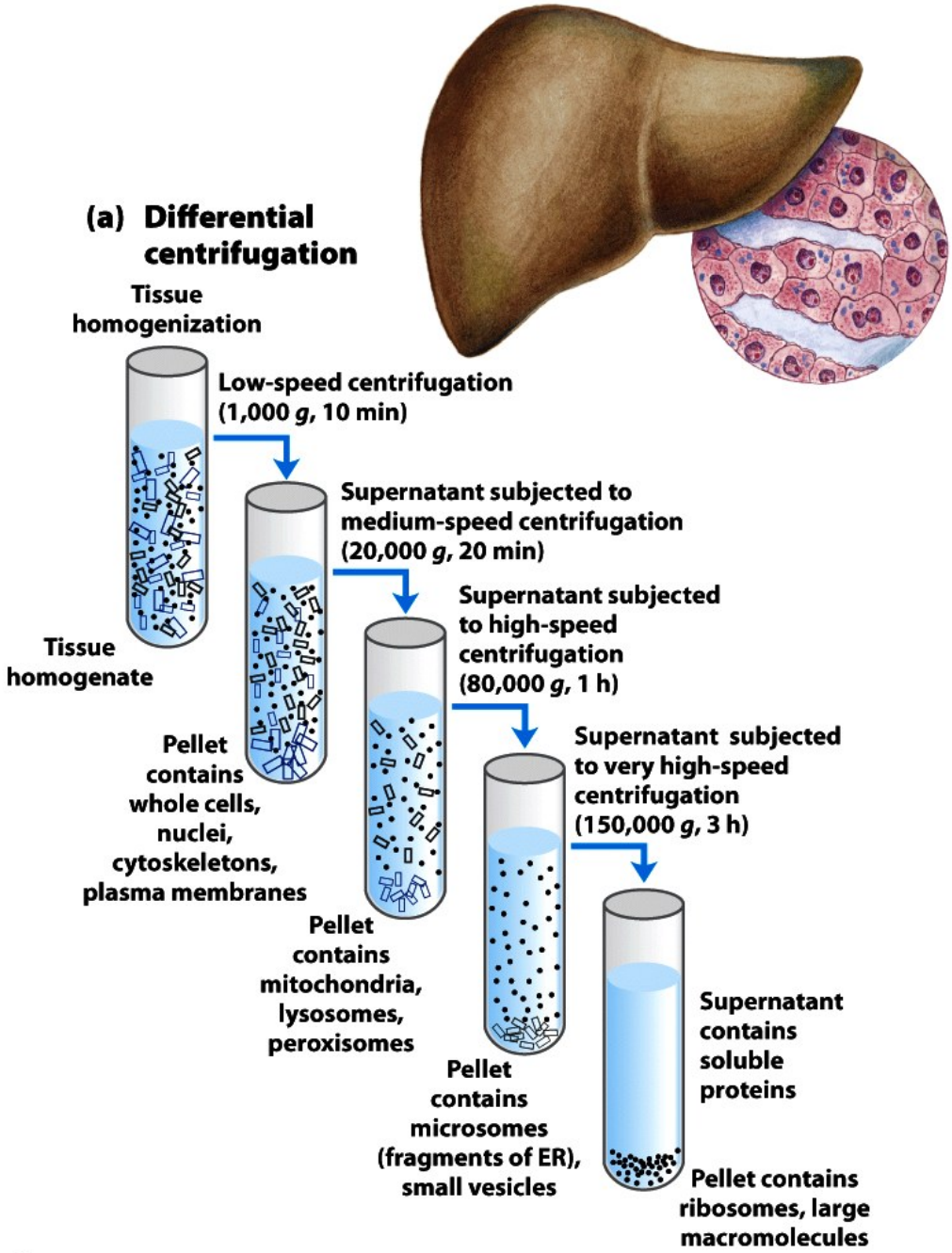


Figure 1-8

Lehninger Principles of Biochemistry, Fifth Edition

A mixture of proteins can be separated

- Separation relies on differences in physical and chemical properties
  - Charge
  - Size
  - Affinity for a ligand
  - Solubility
  - Hydrophobicity
  - Thermal stability
- Chromatography is commonly used for preparative separation in which the protein is often able to remain fully folded

Protein separation

- *“Salting out”:* some proteins come out of solution (precipitate) at high salt concentration (while others stay in solution). $(NH_4)_2SO_4$ is normally used
- *Dialysis:* separation of proteins from solvent because proteins are large. Proteins are put in a semi-permeable bag which is soaked in a larger volume of the correct buffer and salt concentration. Ions and buffer will equilibrate (going in) while proteins cannot go out. Can be used to remove $(NH_4)_2SO_4$

Column Chromatography

Chromatography is commonly used for preparative separation

- Stationary phase
- Mobile phase
- Effluent

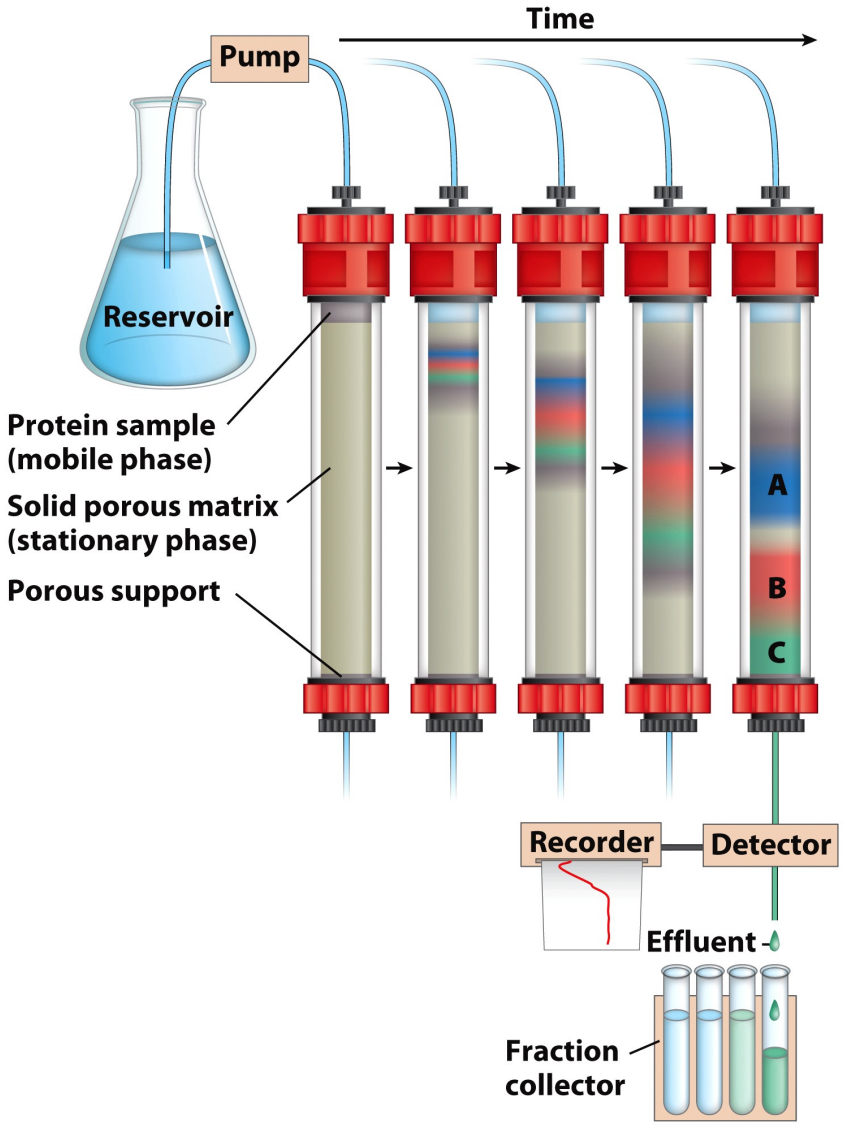


Figure 3-16
Lehninger Principles of Biochemistry, Sixth Edition

Separation by Charge: Ion Exchange

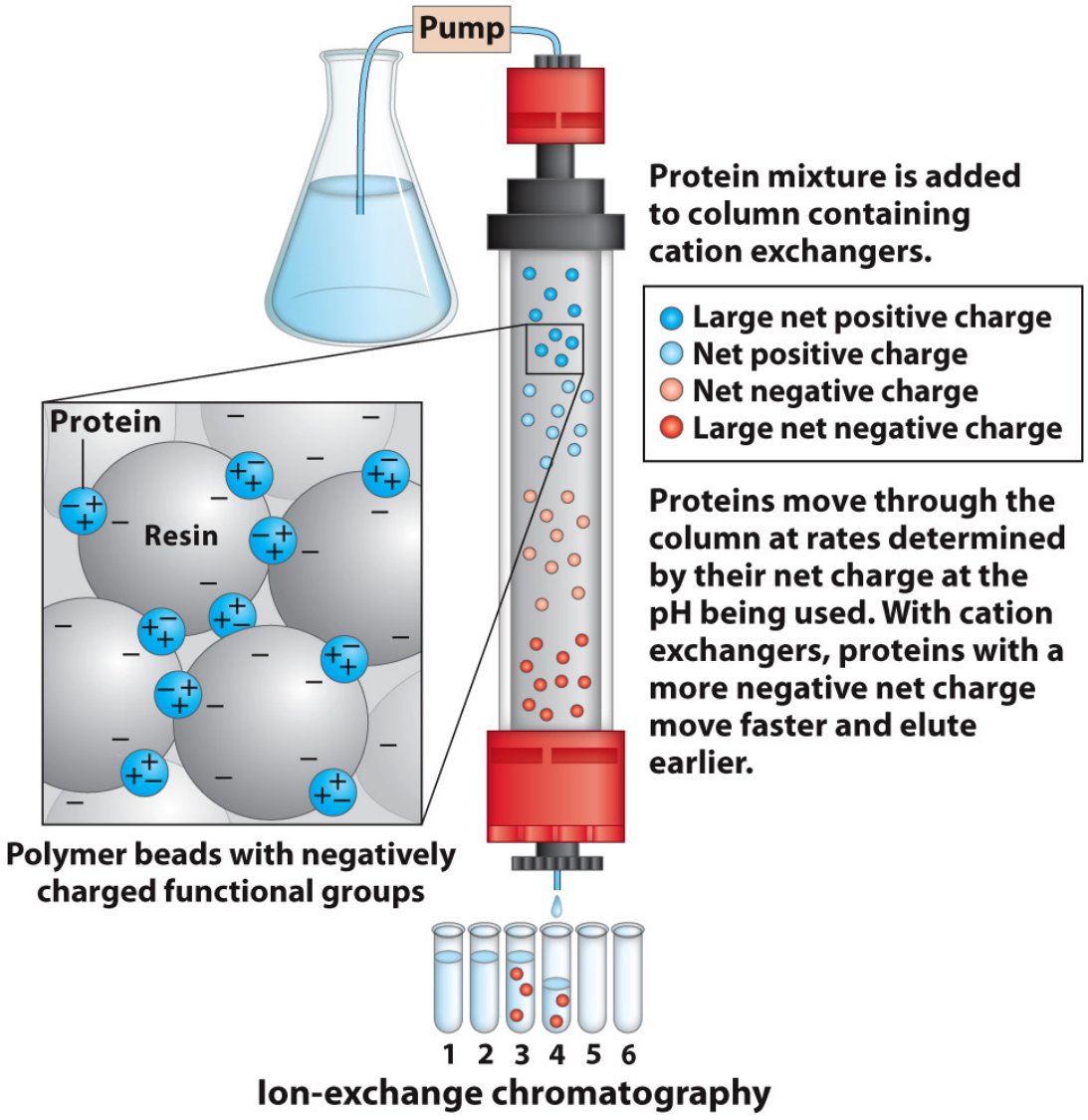


Figure 3-17a
Lehninger Principles of Biochemistry, Seventh Edition

Separation by Size: Size Exclusion

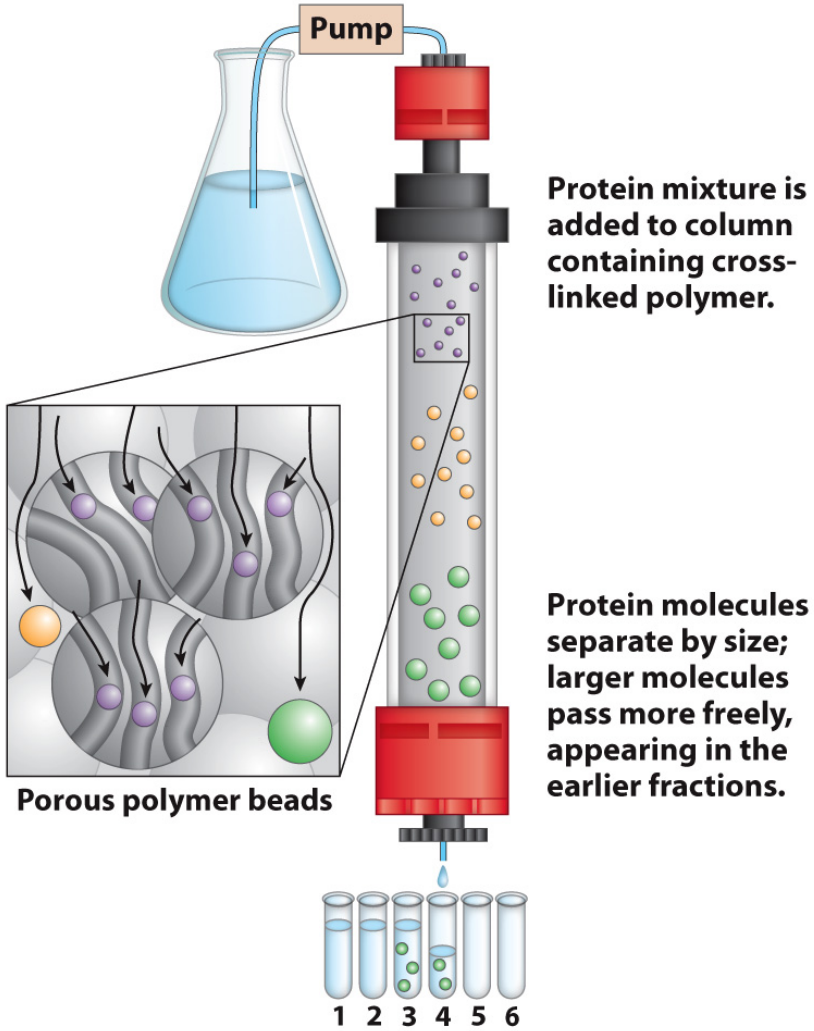


Figure 3-17b
Lehninger Principles of Biochemistry, Seventh Edition

Separation by Binding: Affinity

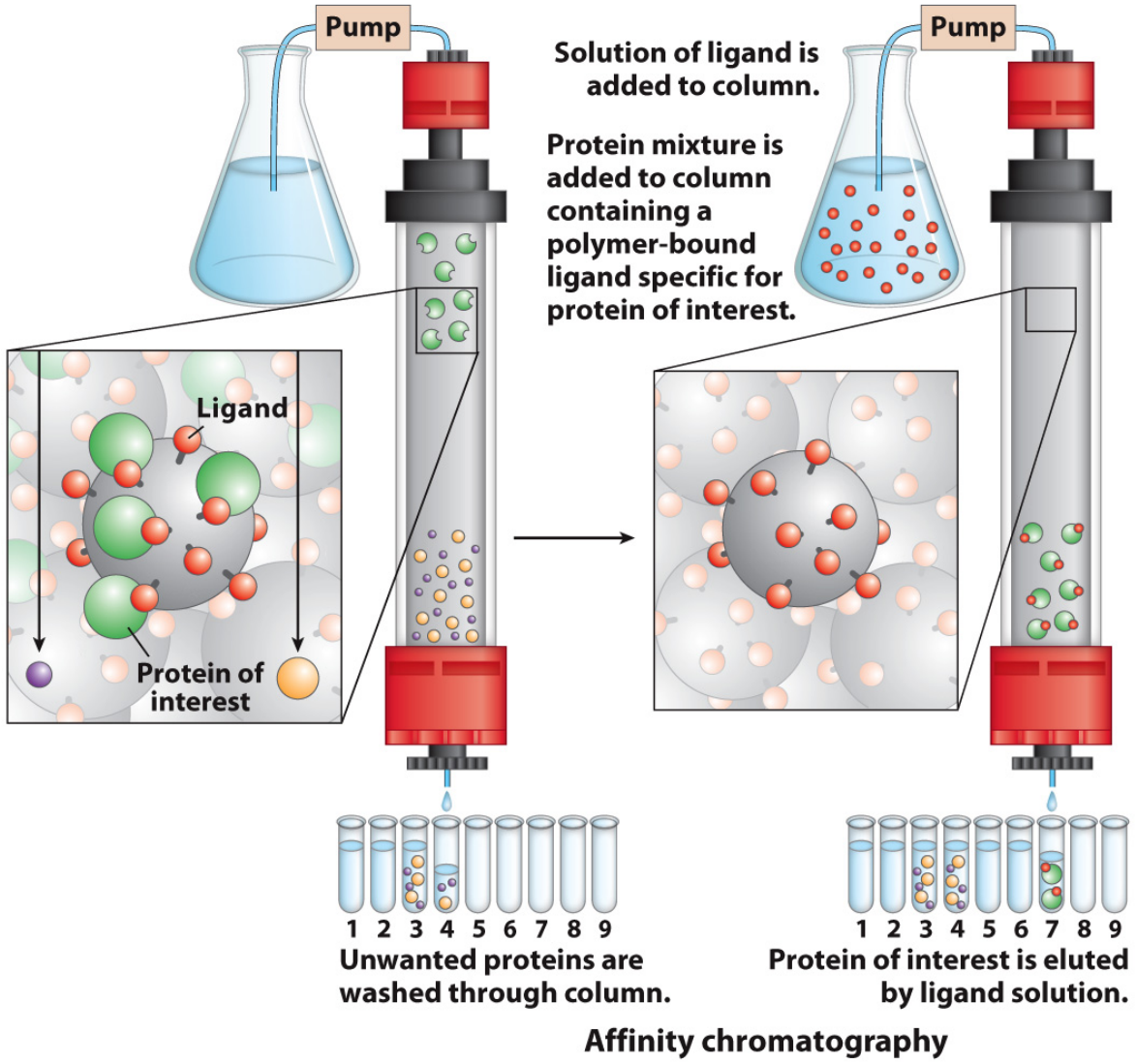


Figure 3-17c
Lehninger Principles of Biochemistry, Seventh Edition

Electrophoresis for Protein Analysis

- Separation in analytical scale is commonly done by electrophoresis
- Electric field pulls proteins according to their charge
- Gel matrix hinders mobility of proteins according to their size and shape
- The gel is commonly polyacrylamide, so separation of proteins via electrophoresis is often called polyacrylamide gel electrophoresis, or PAGE

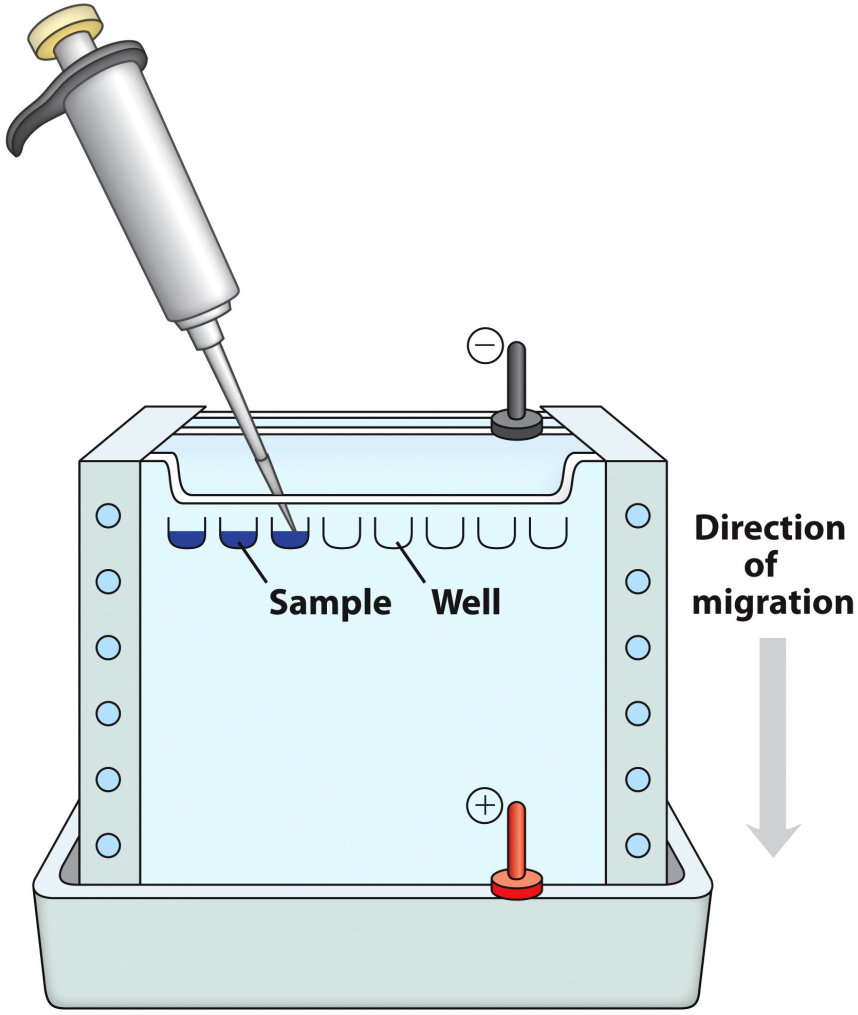
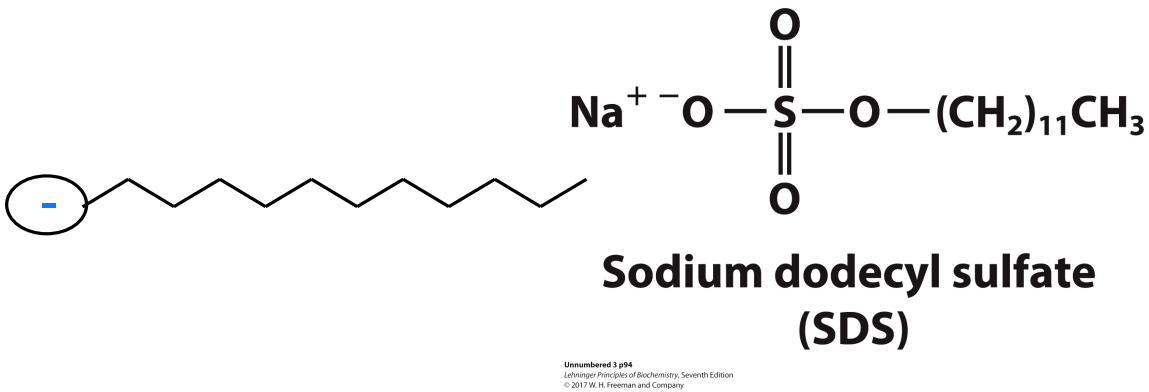


Figure 3-18a
Lehninger Principles of Biochemistry, Sixth Edition

SDS PAGE Separates Proteins by Molecular Weight

- SDS – sodium dodecyl sulfate – a detergent



- SDS micelles bind to proteins and facilitate unfolding.
  - SDS gives all proteins a uniformly negative charge.
  - The native shape of proteins does not matter.
  - The rate of movement will only depend on size: small proteins will move faster.

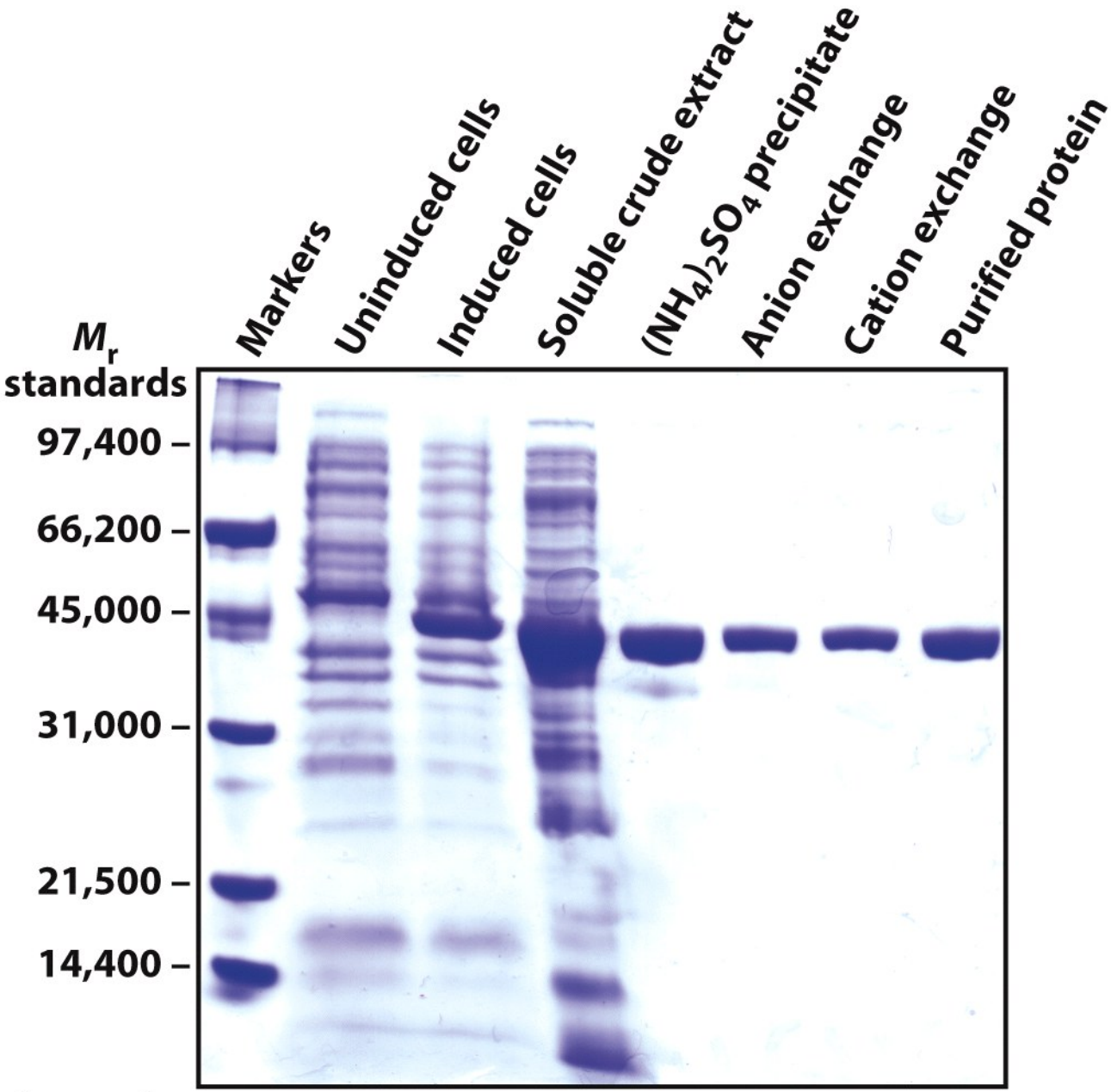


Figure 3-18b
Lehninger Principles of Biochemistry, Fifth Edition

SDS-PAGE can be used to calculate the molecular weight of a protein

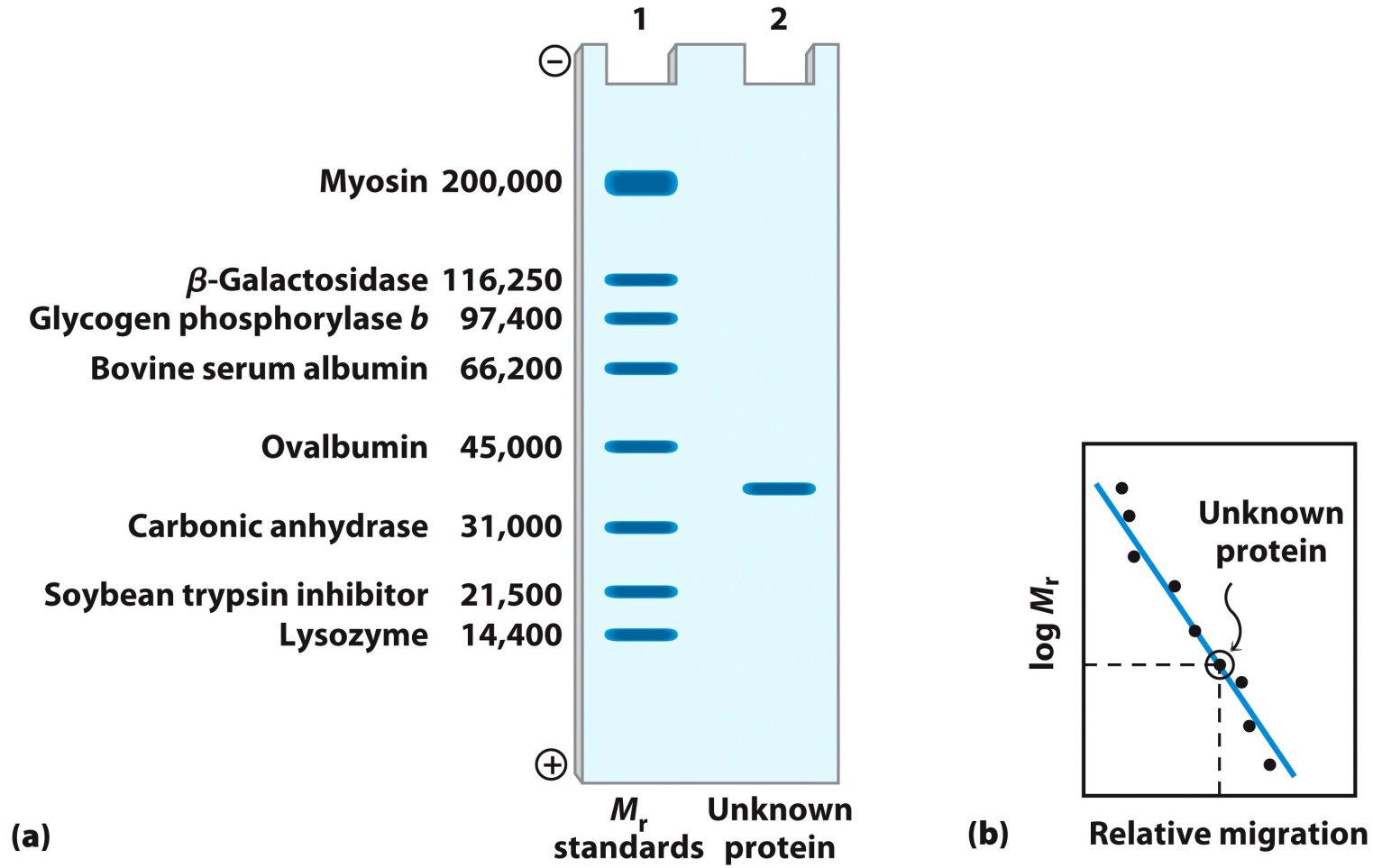


Figure 3-19
Lehninger Principles of Biochemistry, Sixth Edition

Specific Activity Describes the Purity of the Protein of Interest

- Proteins in a complex mixture often require more than one purification to completely isolate the protein of interest.
- During purification, determination of the location of the protein of interest can be performed by tracking its behavior.
- If a protein has a specific function (e.g., binding insulin), the fraction that binds insulin best after each purification step will contain the most of the protein of interest.
- The function of the protein is called the "**activity**."
- The ratio of activity to total protein concentration is called the "**specific activity**."

Specific Activity

- 1 unit of enzyme activity: amount of enzyme causing transformation of 1 μmol of substrate / min at 25 °C

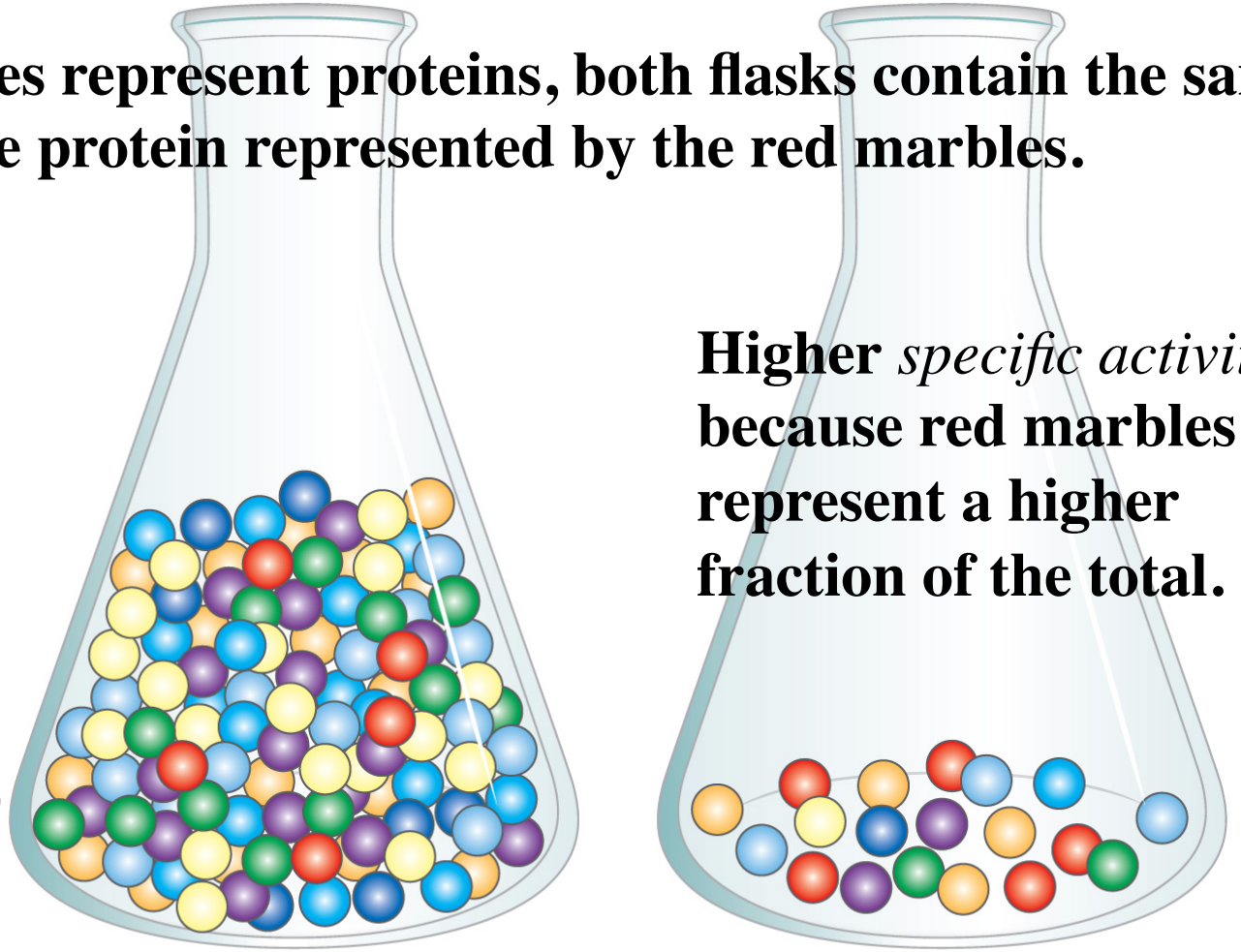
- ***Activity:*** Total units of enzyme in a solution
- ***Specific Activity:*** number of enzyme units / mg of total protein
- In a purification, many steps are used
- After each step, total protein ↓ (sometimes activity ↓) but specific activity ↑

Specific Activity Describes the Purity of the Protein of Interest

If the marbles represent proteins, both flasks contain the same *activity* of the protein represented by the red marbles.

Higher *specific activity* because red marbles represent a higher fraction of the total.

5 red in 100



5 red in 20

Figure 3-22
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

TABLE 3–5 A Purification Table for a Hypothetical Enzyme

Procedure or step	Fraction volume (mL)	Total protein (mg)	Activity (units)	Specific activity (units/mg)
1. Crude cellular extract	1,400	10,000	100,000	10
2. Precipitation with ammonium sulfate	280	3,000	96,000	32
3. Ion-exchange chromatography	90	400	80,000	200
4. Size-exclusion chromatography	80	100	60,000	600
5. Affinity chromatography	6	3	45,000	15,000

Note: All data represent the status of the sample *after* the designated procedure has been carried out. Activity and specific activity are defined on page 91.

Table 3-5

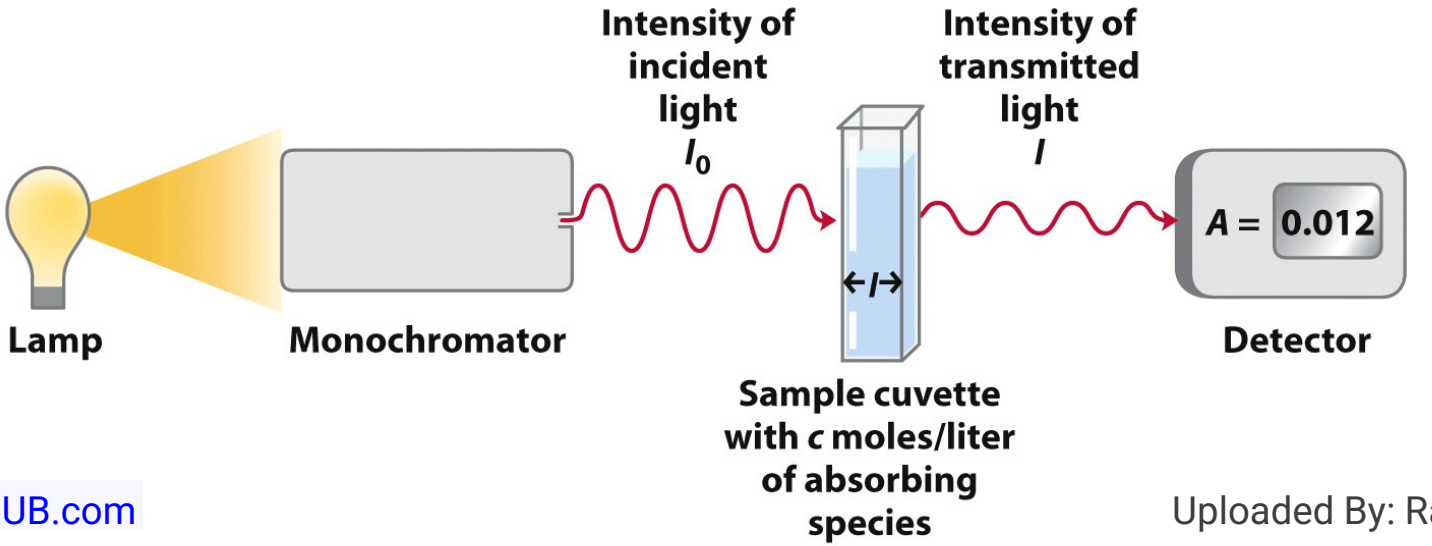
Lehninger Principles of Biochemistry, Fifth Edition

Protein Sequencing

- It is essential to further biochemical analysis that we know the sequence of the protein we are studying.
- The actual sequence is generally determined from the DNA sequence.
- **Edman degradation** (classical method):
  - successive rounds of *N*-terminal modification, cleavage, and identification
  - can be used to identify protein with known sequence
- **Mass spectrometry** (modern method):
  - MALDI MS and ESI MS can precisely identify the mass of a peptide, and thus the amino acid sequence
  - can be used to determine posttranslational modifications

Spectroscopic Detection of Aromatic Amino Acids

- The aromatic amino acids absorb light in the UV region
- Proteins typically have UV absorbance maxima around 275-280 nm
- Tryptophan and tyrosine are the strongest chromophores
- Concentration can be determined by UV-visible spectrophotometry using Beer-Lambert's law: $A = \varepsilon \cdot c \cdot l$



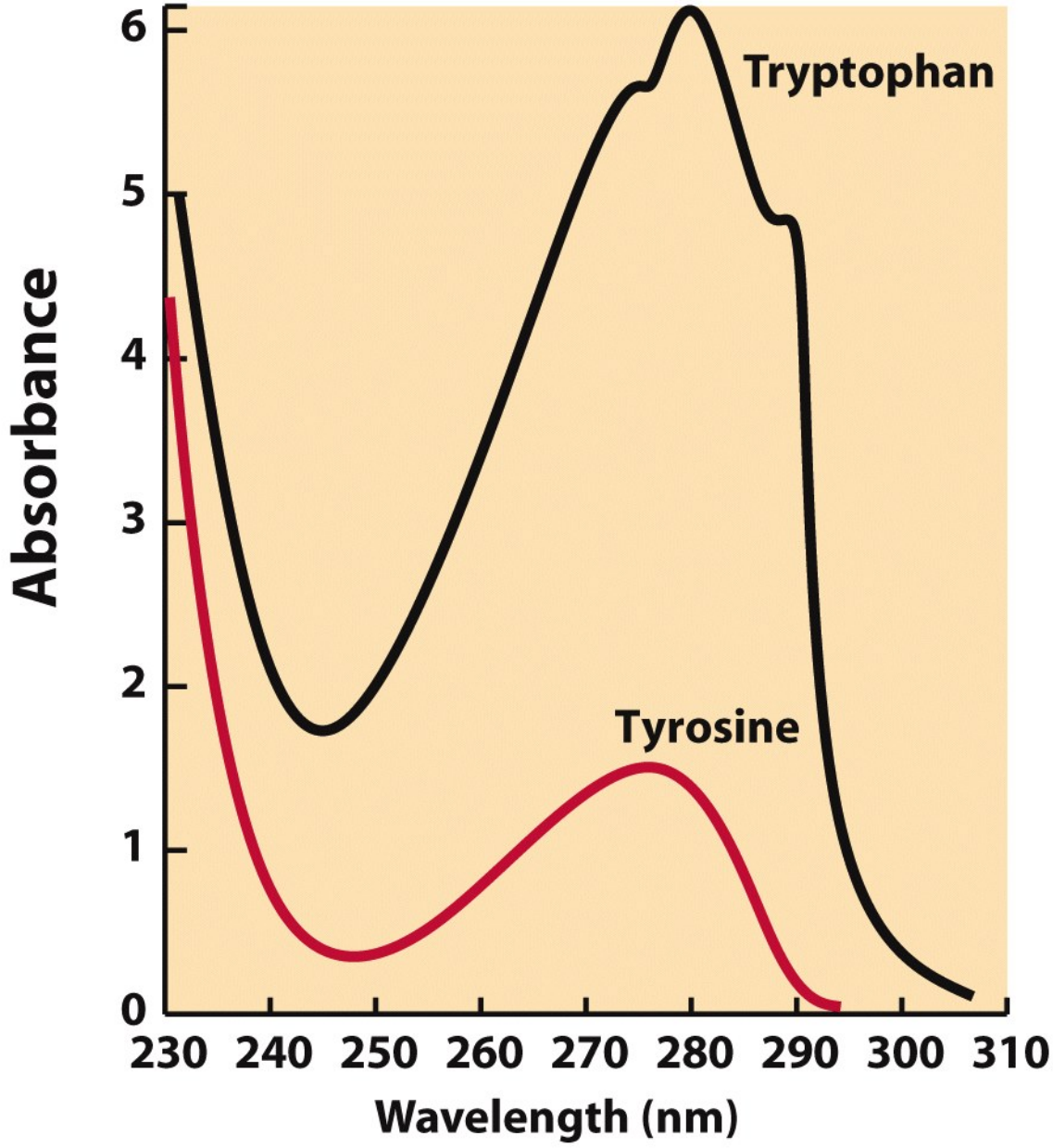


Figure 3-6
Lehninger Principles of Biochemistry, Fifth Edition

Proteases Can Be Used to Cleave Proteins

- Enzymes that catalyze the hydrolysis of peptide bonds are **proteases**
- Different kinds of proteases
- Trypsin, cleavage points: K,R (C)
  peptide: WTRCTTSRLPLKSSWSSRWSET

  will be cleaved by trypsin into:
  WTR + CTTSR + LPLK + SSWSSR + WSET

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)	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.

Table 3-7
Lehninger Principles of Biochemistry, Fifth Edition

Protein Sequences as Clues to Evolutionary Relationships

- Sequences of homologous proteins from a wide range of species can be aligned and analyzed for differences
- Differences indicate evolutionary divergences
- Analysis of multiple protein families can indicate evolutionary relationships between organisms, ultimately the history of life on Earth