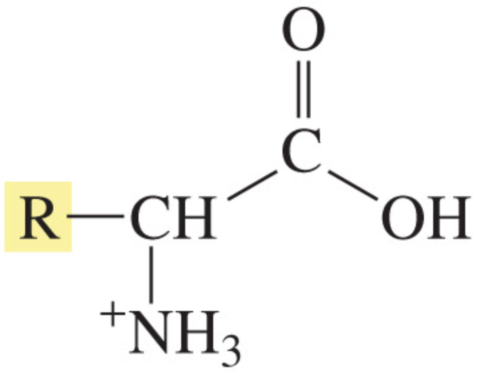
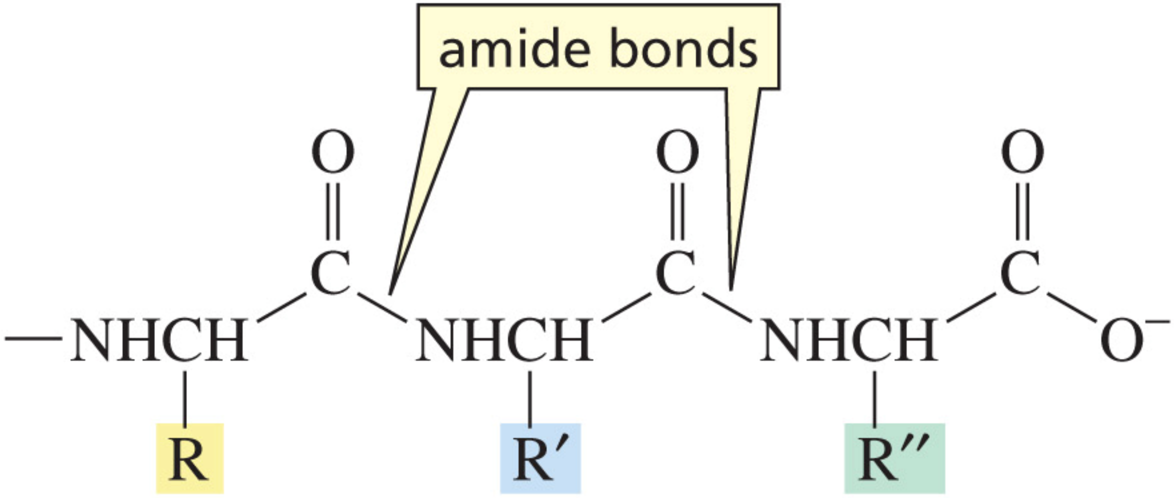
01. Amino Acids

Biomolecules

- **Protein**
- **Carbohydrate**
- **Nucleic acid**
- **Lipid**



a protonated
α-aminocarboxylic acid
an amino acid



amino acids are linked together by amide bonds

peptide → polypeptide → protein

di-, tri-, oligo-

Table 22.1 Examples of the Many Functions of Proteins in Biological Systems	
Structural proteins	These proteins impart strength to biological structures or protect organisms from their environment. For example, collagen is the major component of bones, muscles, and tendons; keratin is the major component of hair, hooves, feathers, fur, and the outer layer of skin.
Protective proteins	Snake venoms and plant toxins protect their owners from predators. Blood-clotting proteins protect the vascular system when it is injured. Antibodies and peptide antibiotics protect us from disease.
Enzymes	Enzymes are proteins that catalyze the reactions that occur in living systems.
Hormones	Some of the hormones, such as insulin, that regulate the reactions that occur in living systems are proteins.
Proteins with physiological functions	These proteins are responsible for physiological functions such as the transport and storage of oxygen in the body, the storage of oxygen in the muscles, and the contraction of muscles.

fibrous proteins

proteins

globular proteins

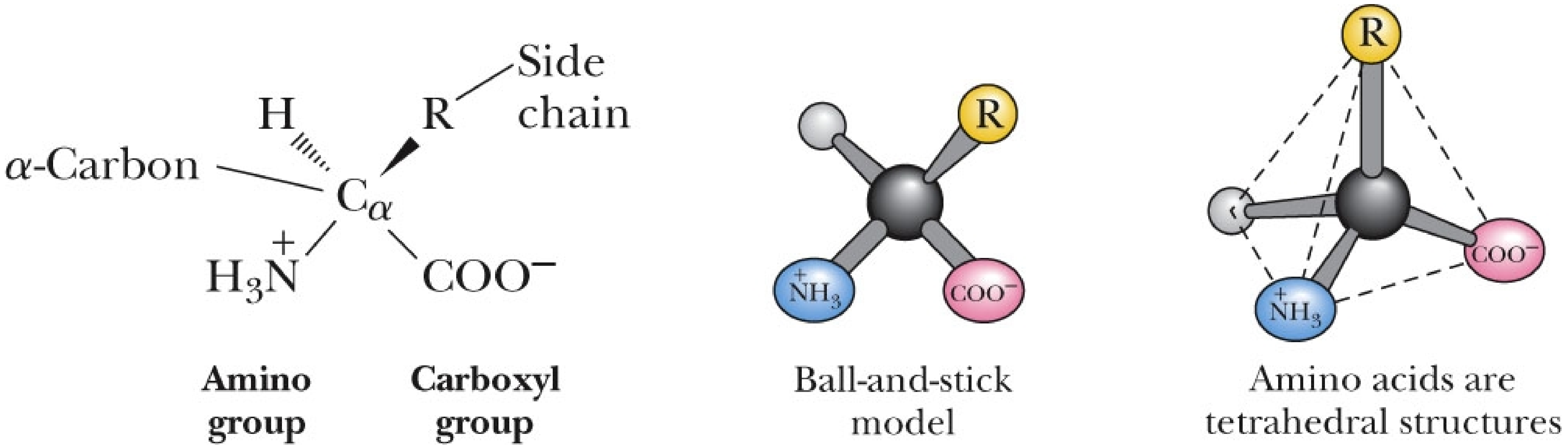
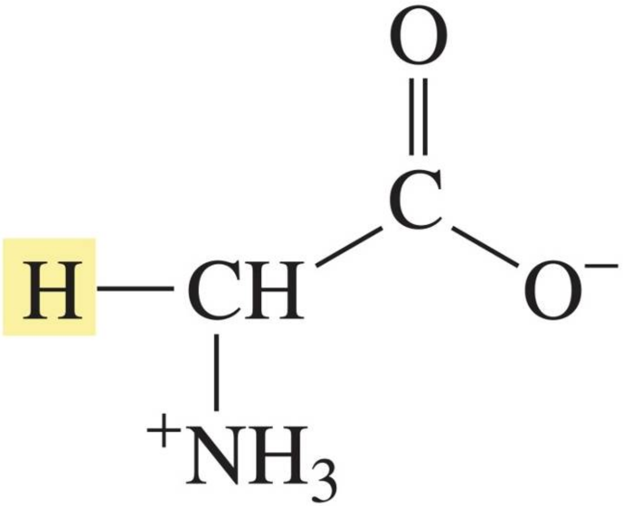
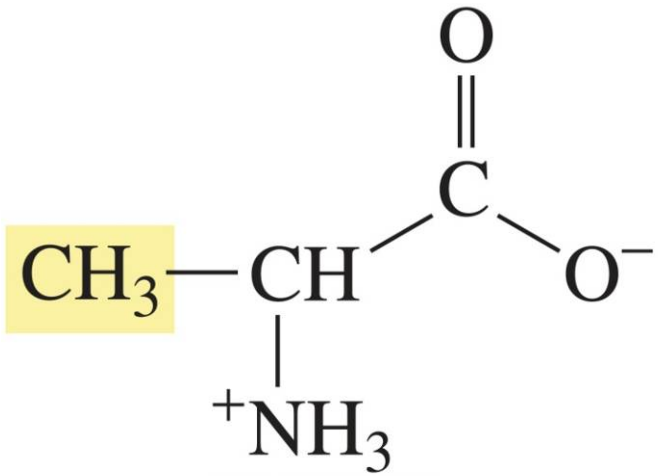
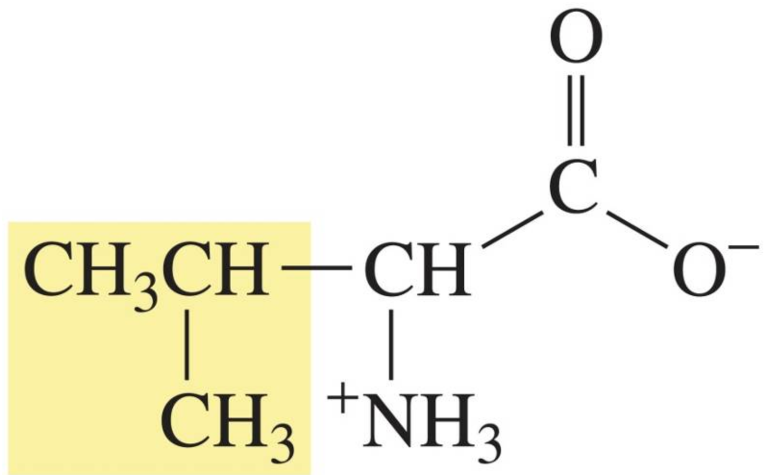


Figure 4.1 Anatomy of an amino acid. Except for proline and its derivatives, all of the amino acids commonly found in proteins possess this type of structure.



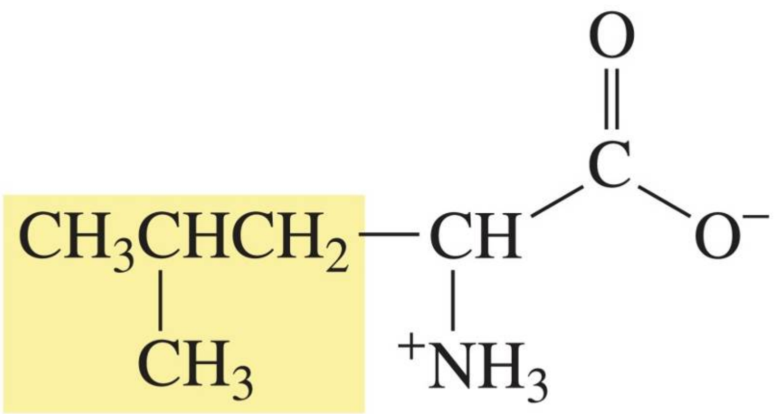


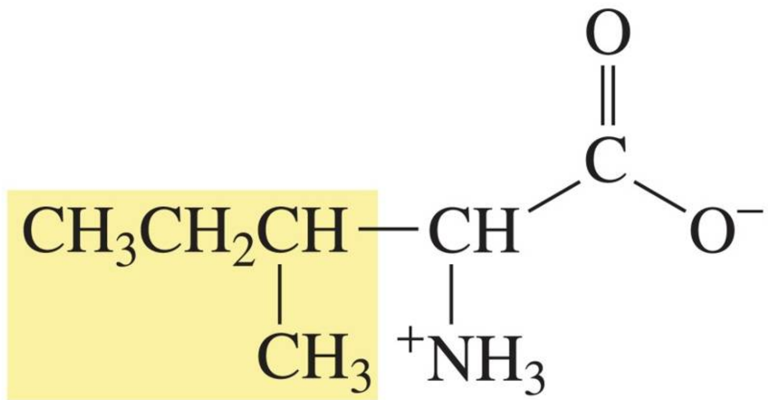


Glycine (Gly, G)

Alanine (Ala, A)

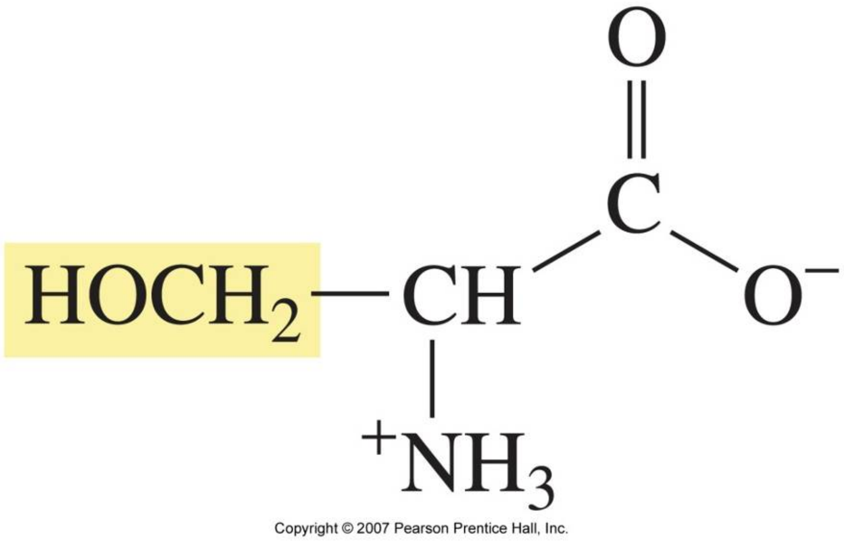
Valine (Val, V)*

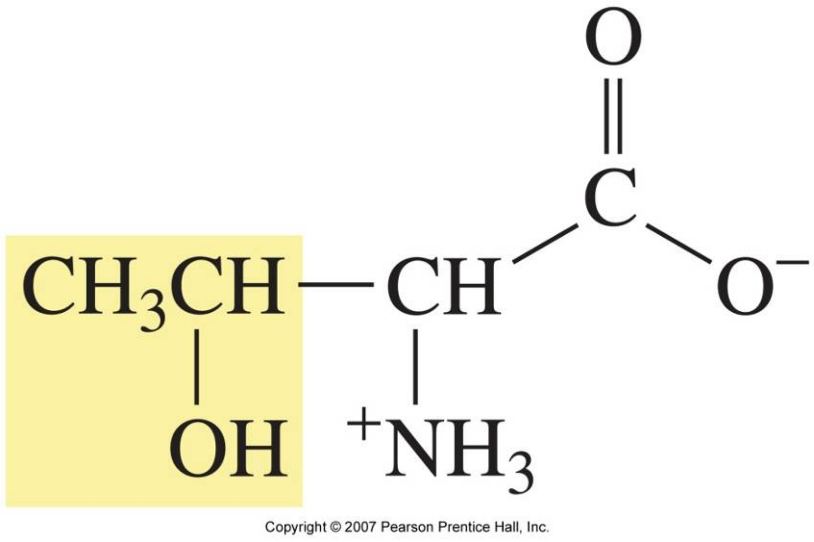




Leucine (Leu, L)*

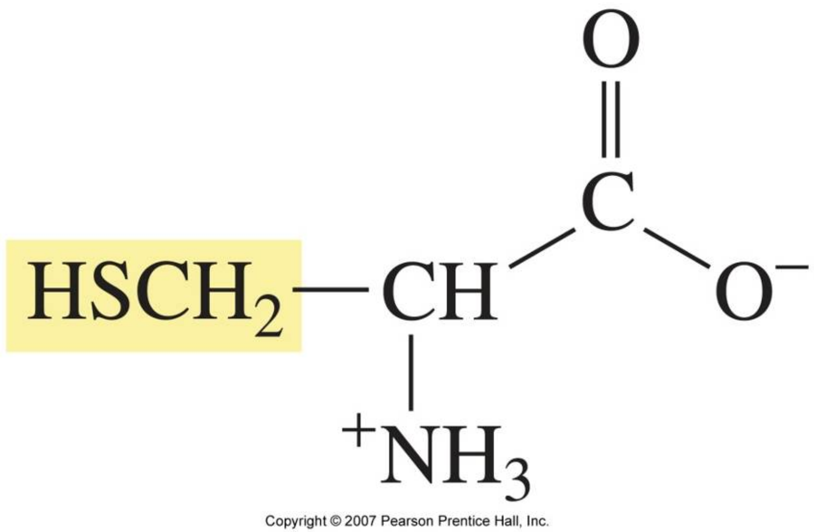
Isoleucine (Ile. I)*

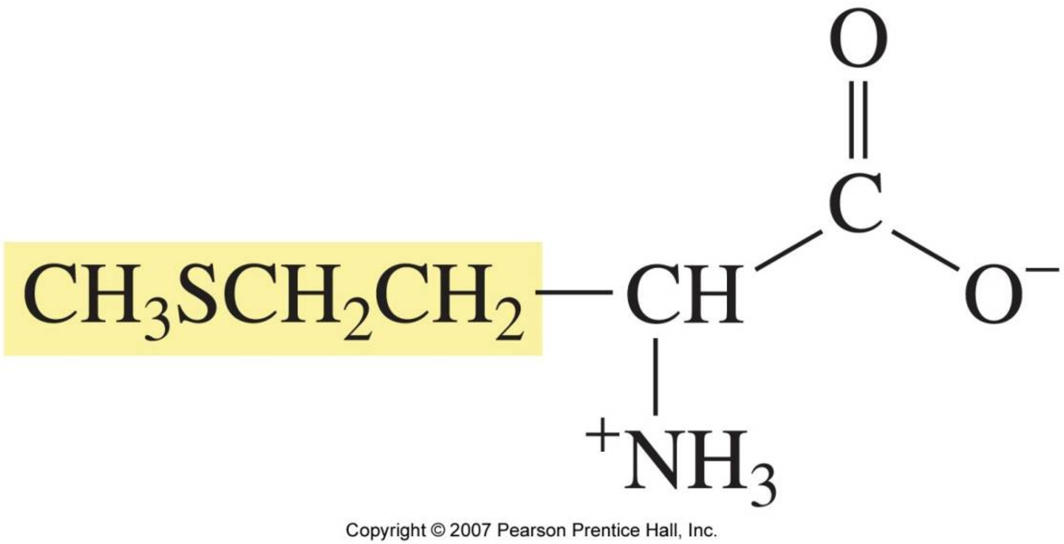




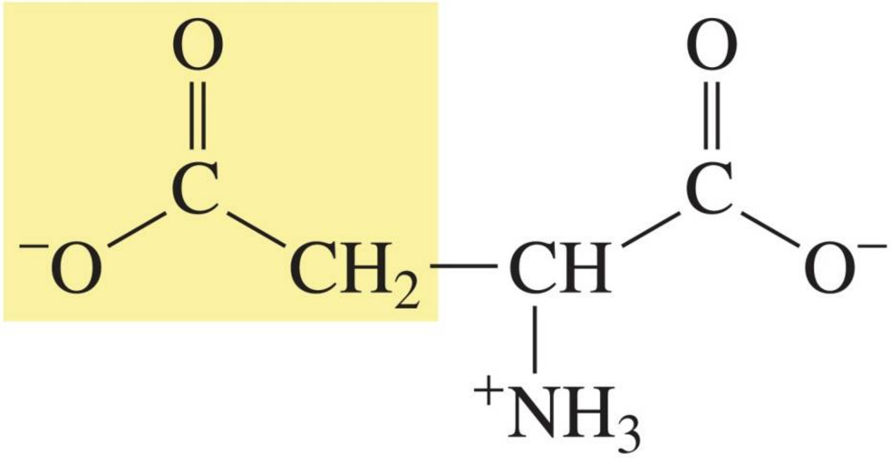
Serine (Ser, S)

Threonine (Thr, T)*

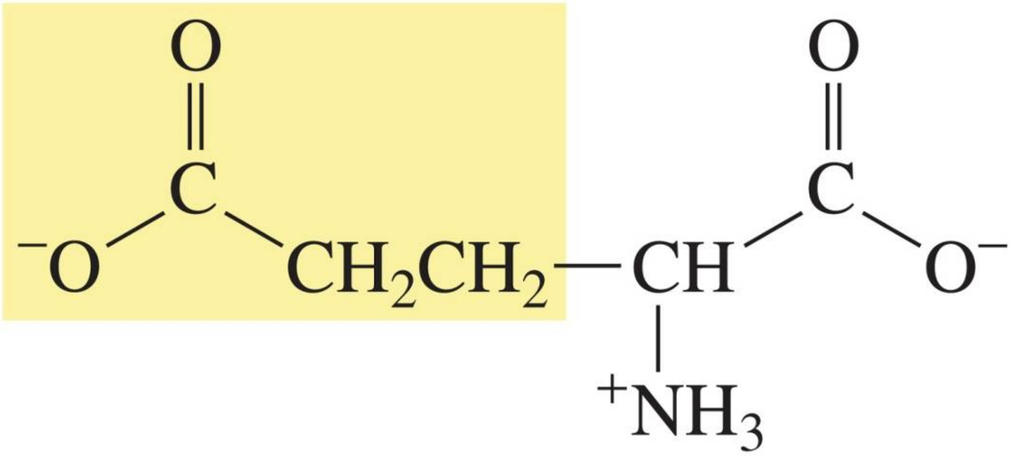




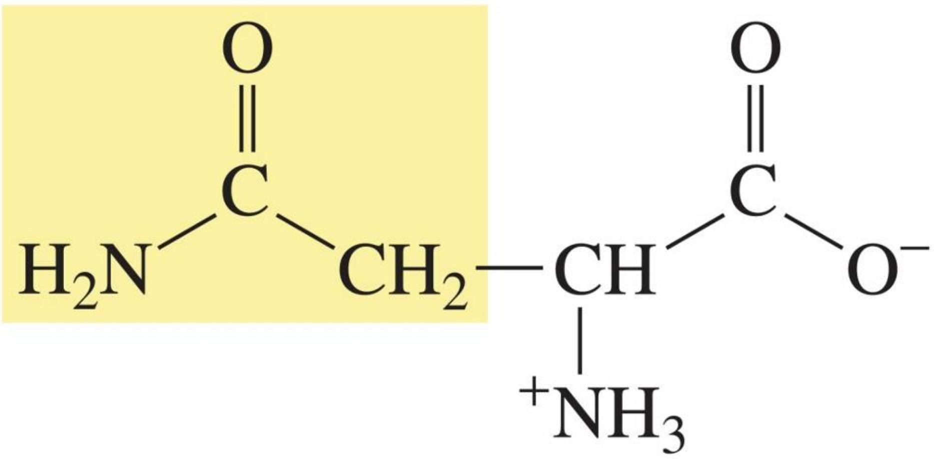
Cysteine (Cys, C) **cystine** **Methionine (Met, M)***



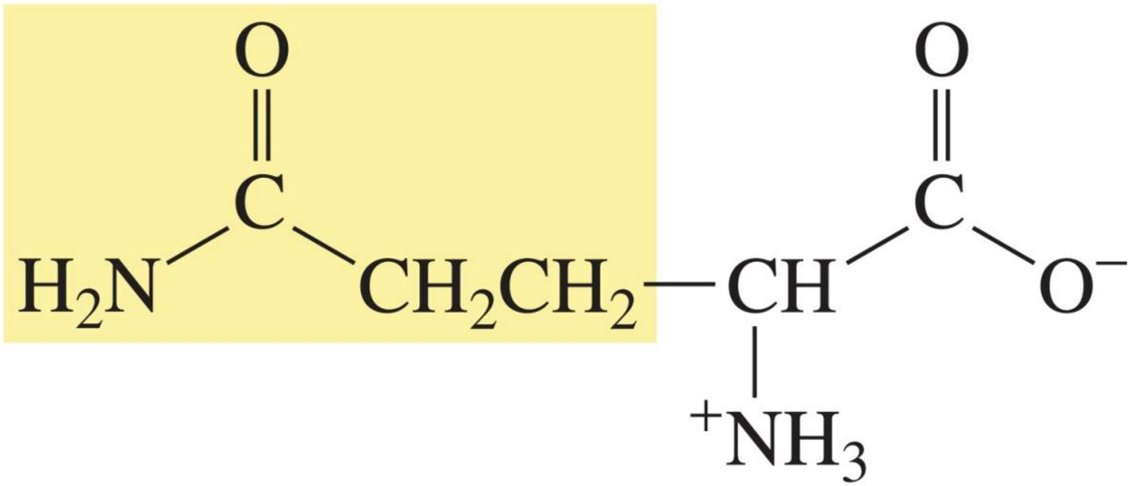
Aspartate (Asp, D)



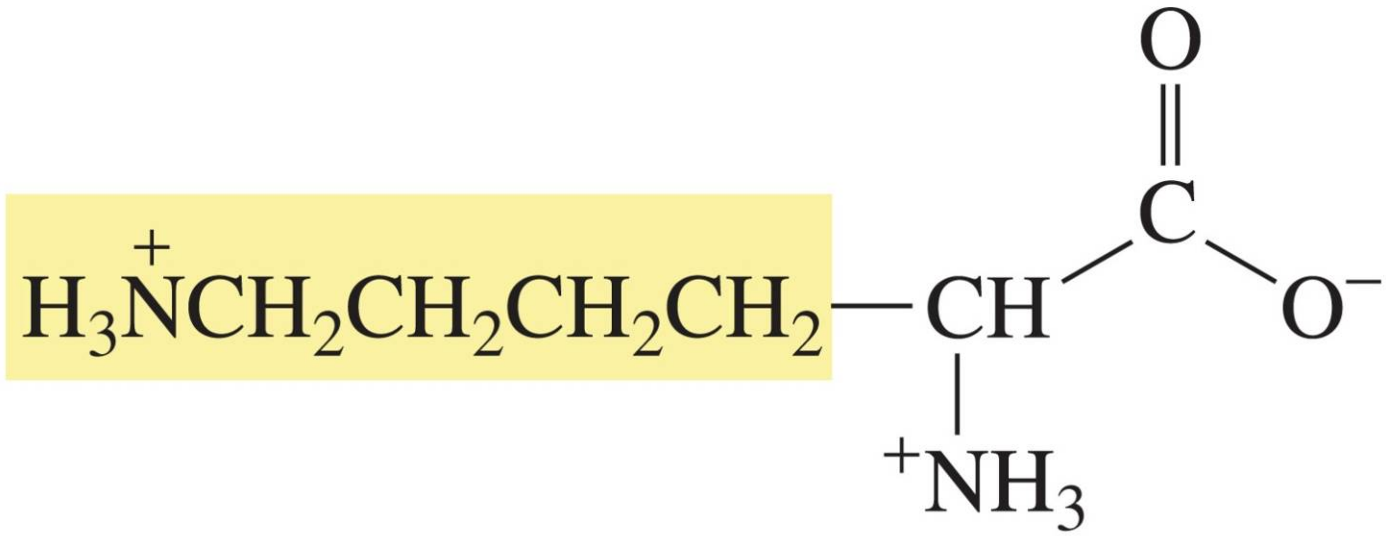
Glutamate (Glu, E)



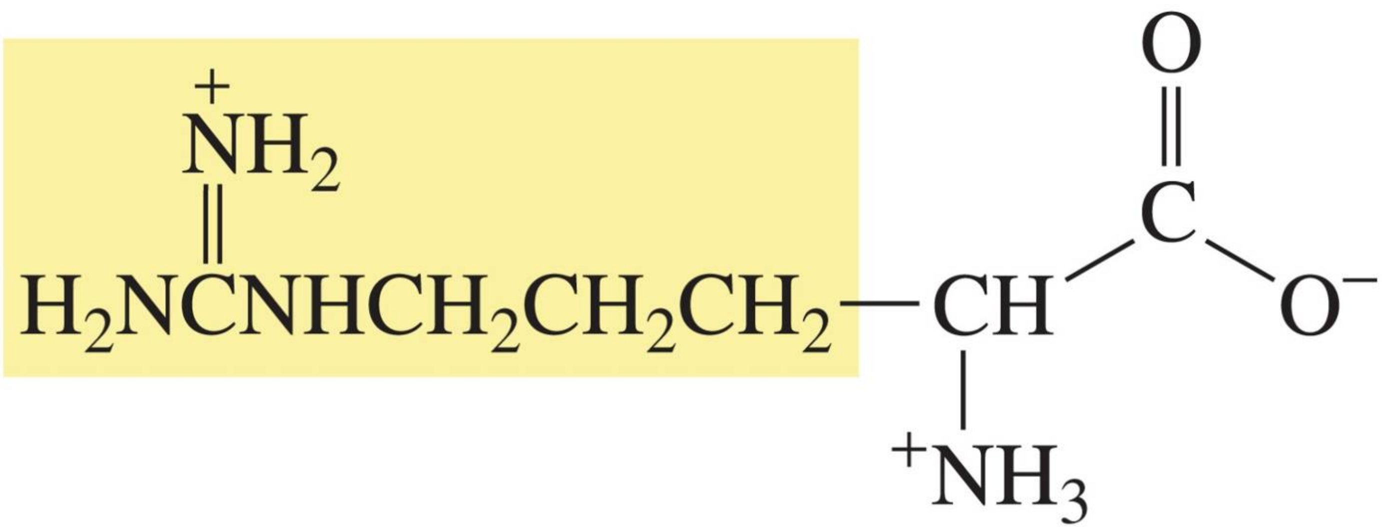
Asparagine (Asn, N)



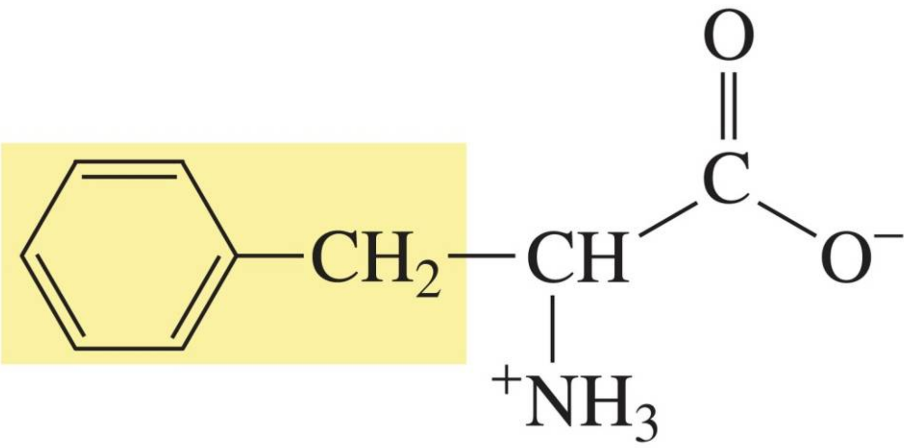
Glutamine (Gln, Q)



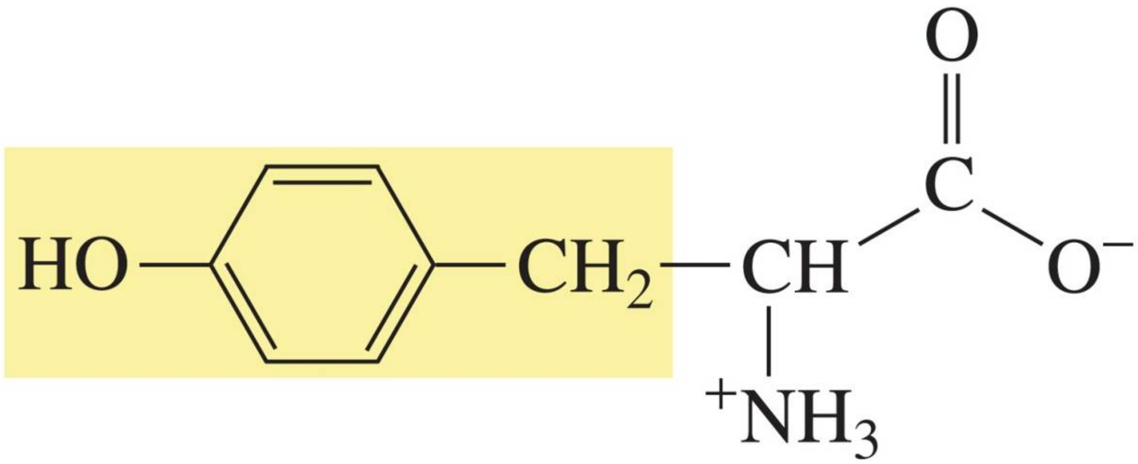
Lysine (Lys, K)*



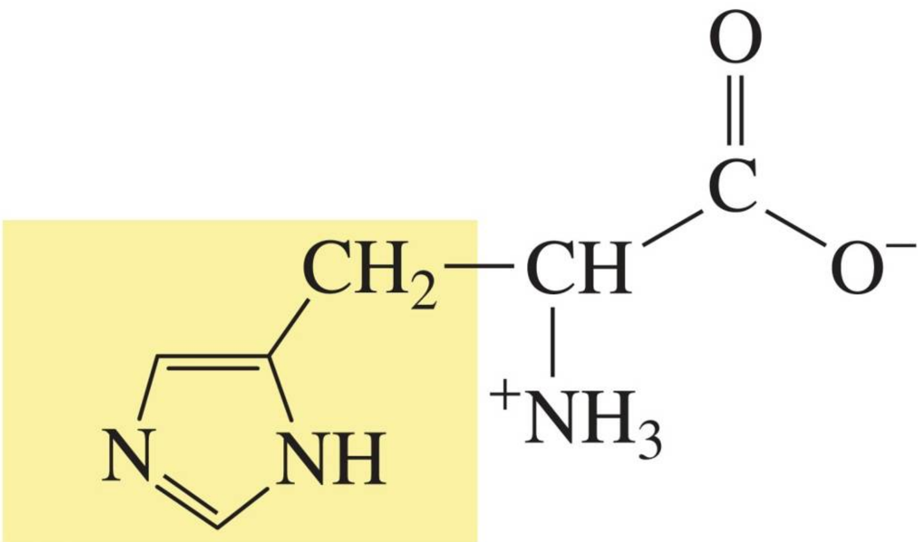
Arginine (Arg, R)*



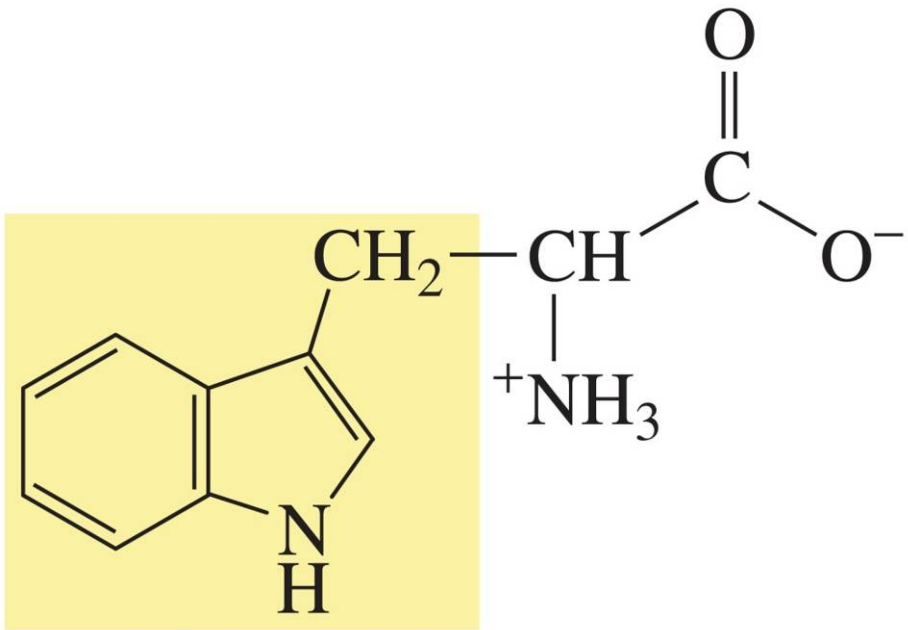
Phenylalanine (Phe, F)*



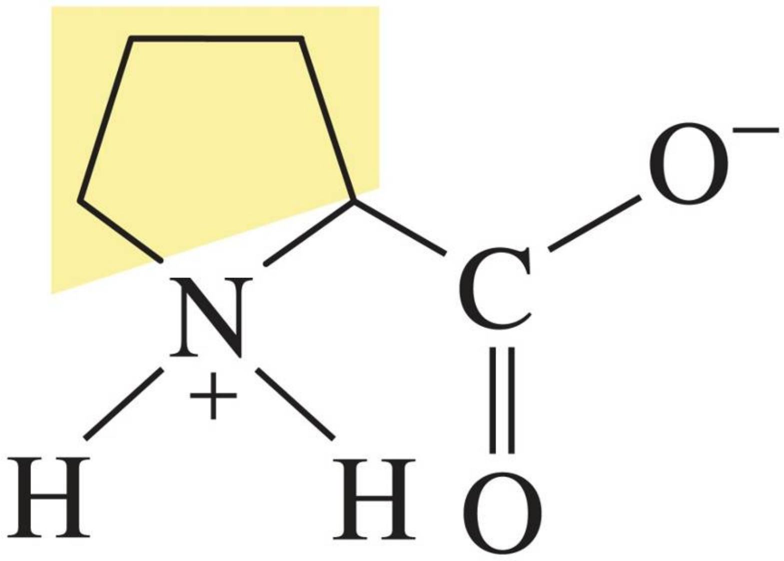
Tyrosine (Tyr, Y)



Histidine (His, H)*



Tryptophan (Trp, W)*



Proline (Pro, P)

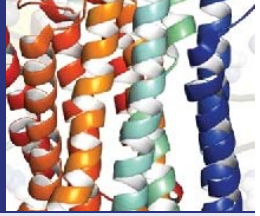
Hydrophobic (A, G, I, L, F, V, P)

Hydrophilic (D, E, R, S, T, C, N, Q, H)

Amphipathic (K, M, W, Y)

Essential amino acids: V, L, I, T, M, K, R, F, H, W

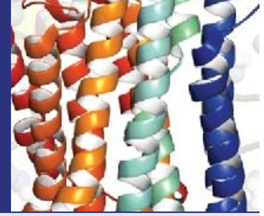
Several Amino Acids Occur Rarely in Proteins

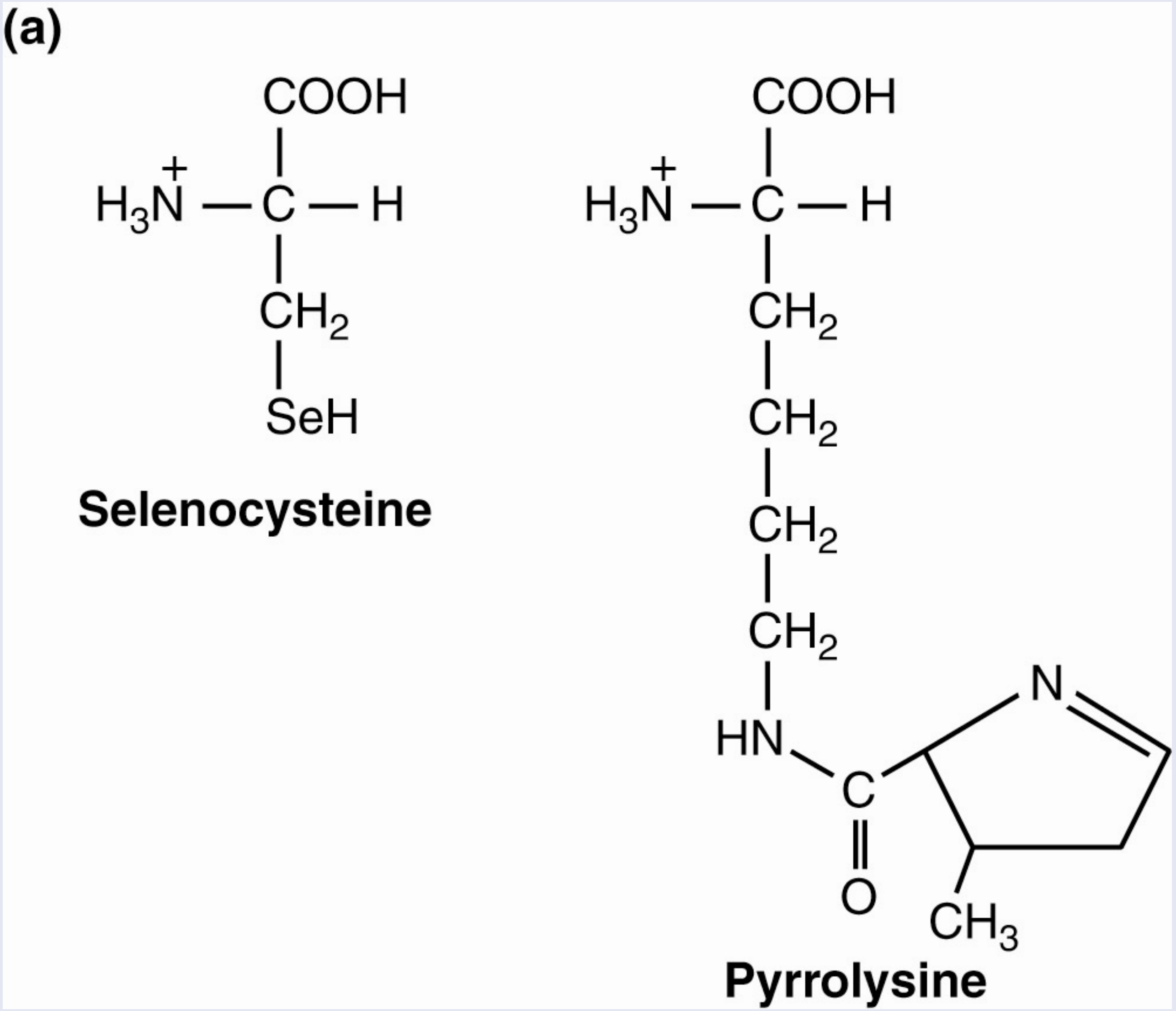


We'll see some of these in later chapters

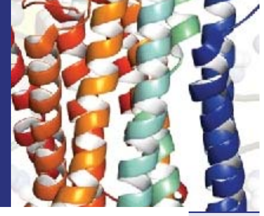
- **Selenocysteine** in many organisms
- **Pyrrolysine** in several archaeal species
- **Hydroxylysine**, hydroxyproline - collagen
- **Carboxyglutamate** - blood-clotting proteins
- **Pyroglutamate** – in bacteriorhodopsin
- **GABA**, epinephrine, histamine, serotonin act as neurotransmitters and hormones
- **Phosphorylated** amino acids – a signaling device

Several Amino Acids Occur Rarely in Proteins





Several Amino Acids Occur Rarely in Proteins



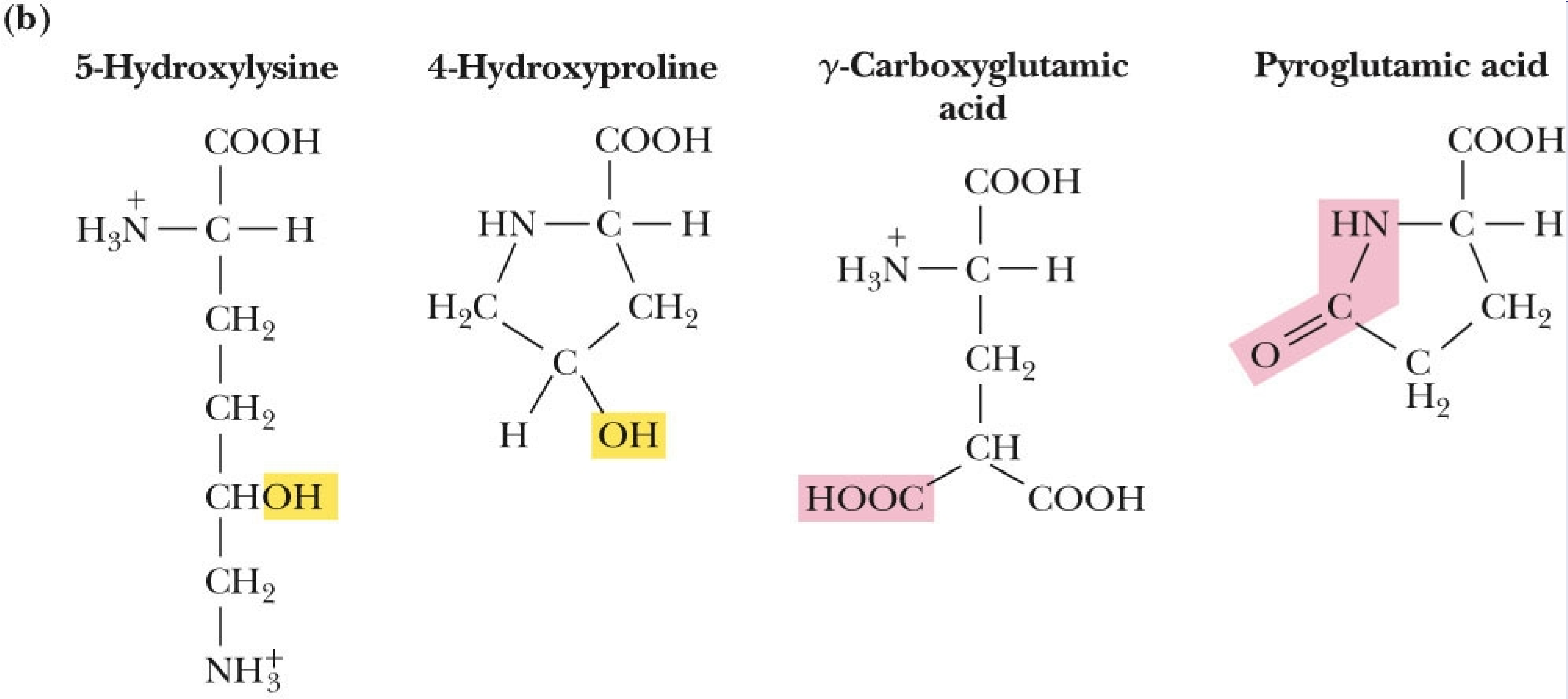
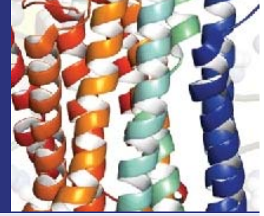


Figure 4.4 (b) Some amino acids are less common, but nevertheless found in certain proteins. Hydroxylysine and hydroxyproline are found in connective-tissue proteins; carboxyglutamate is found in blood-clotting proteins; pyroglutamate is found in bacteriorhodopsin (see Chapter 9).

Several Amino Acids Occur Rarely in Proteins



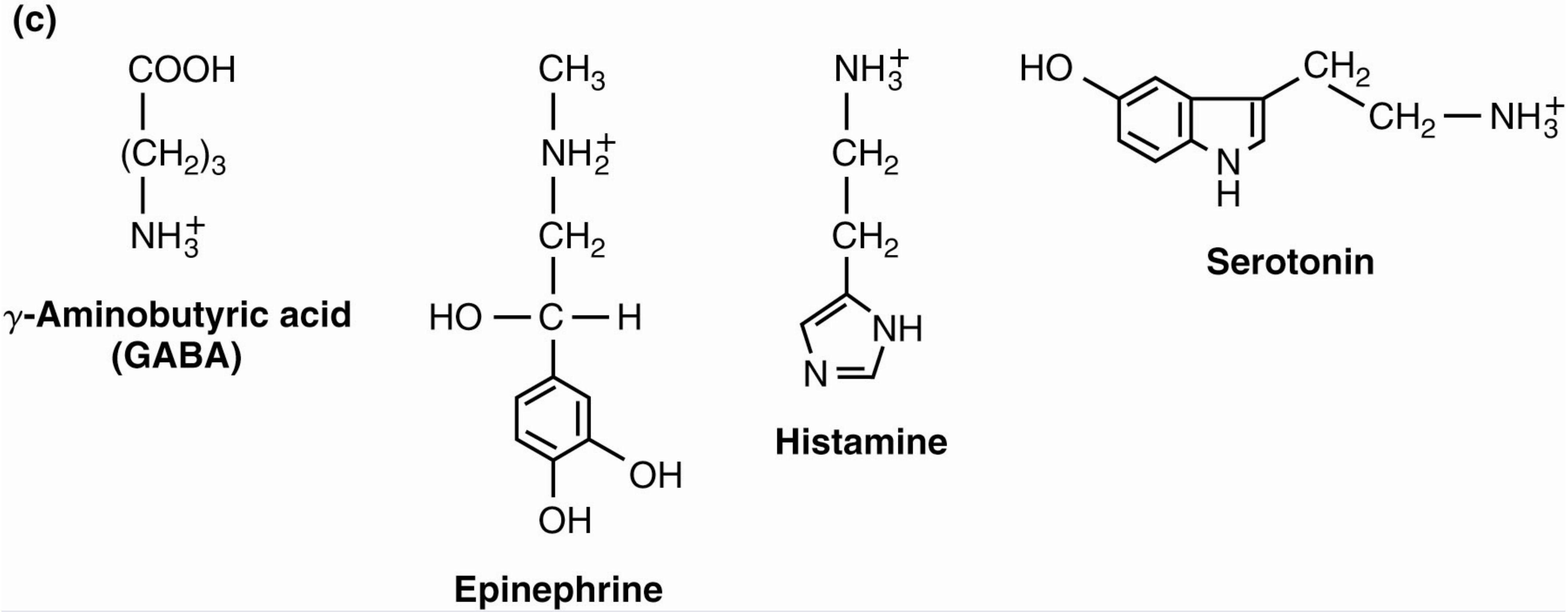
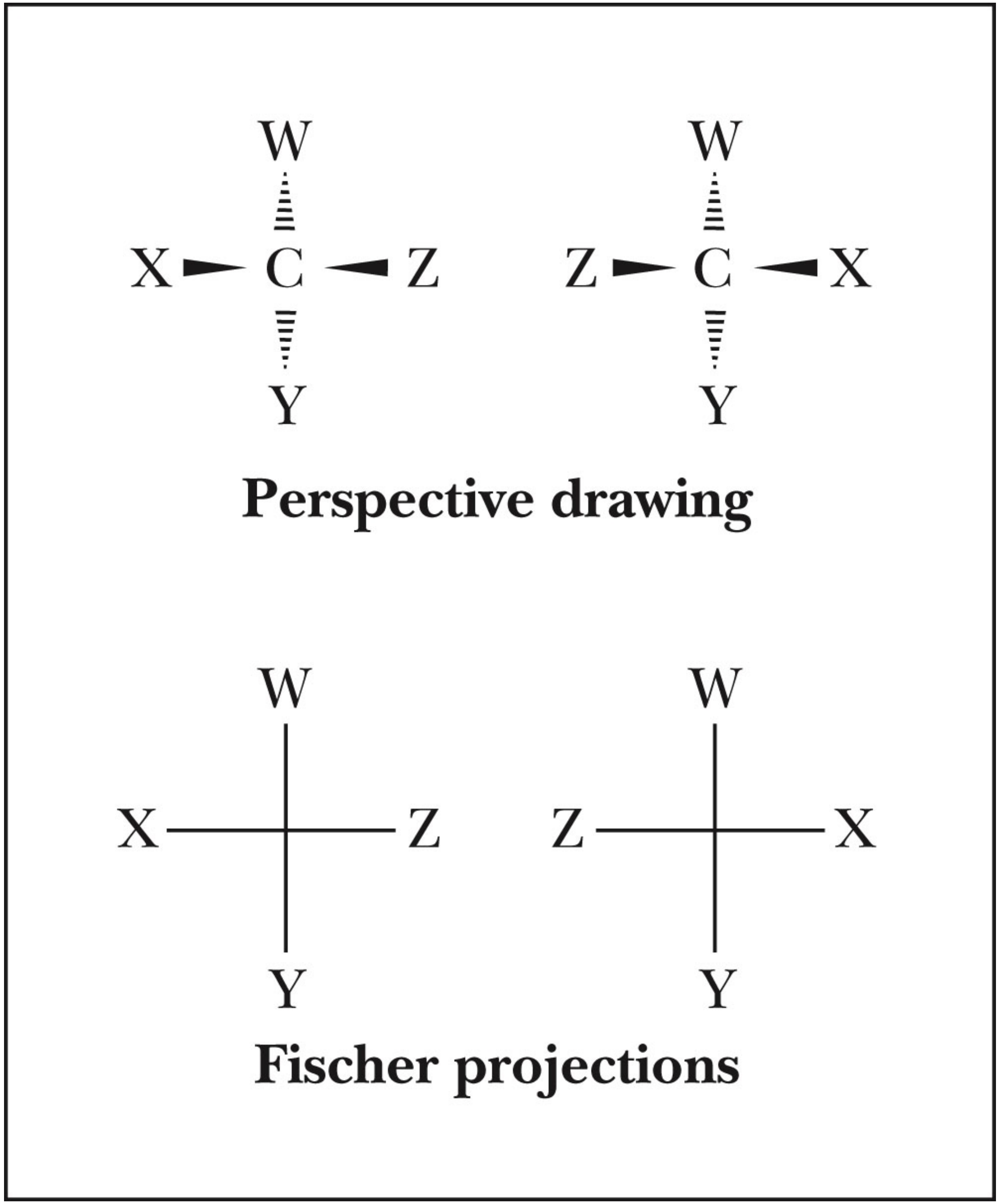
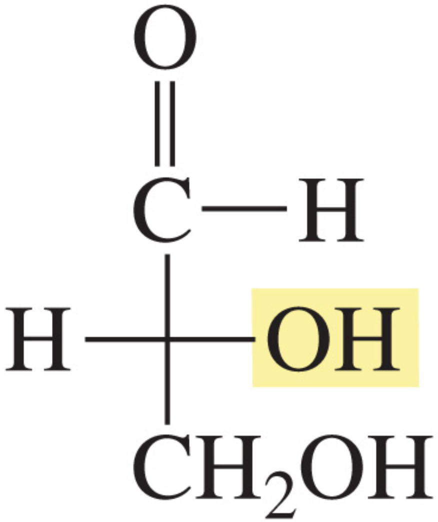


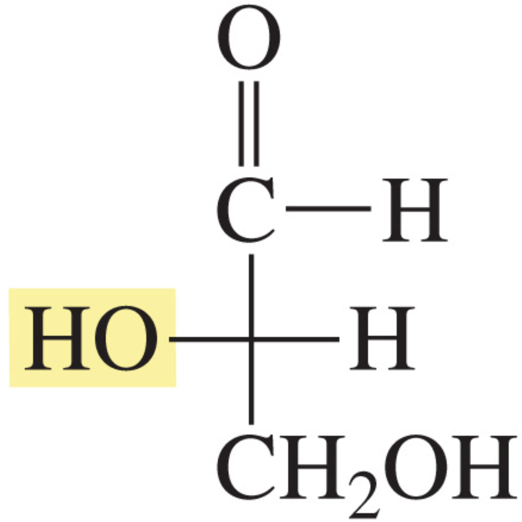
Figure 4.4 (c) Several amino acids that act as neurotransmitters and hormones.

Stereochemistry of Amino Acids

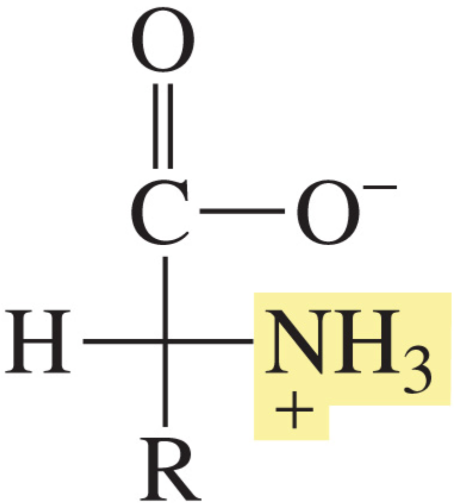




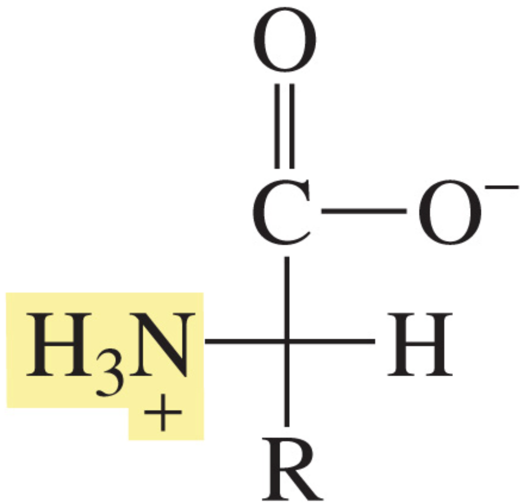
D-glyceraldehyde



L-glyceraldehyde

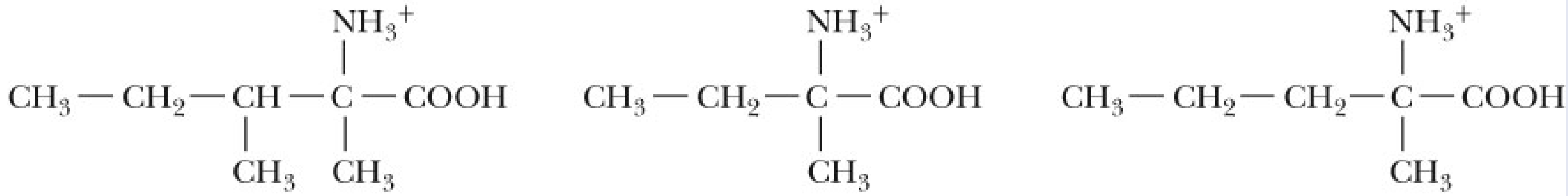


D-amino acid



L-amino acid

The Murchison Meteorite – Discovery of Extraterrestrial Handedness



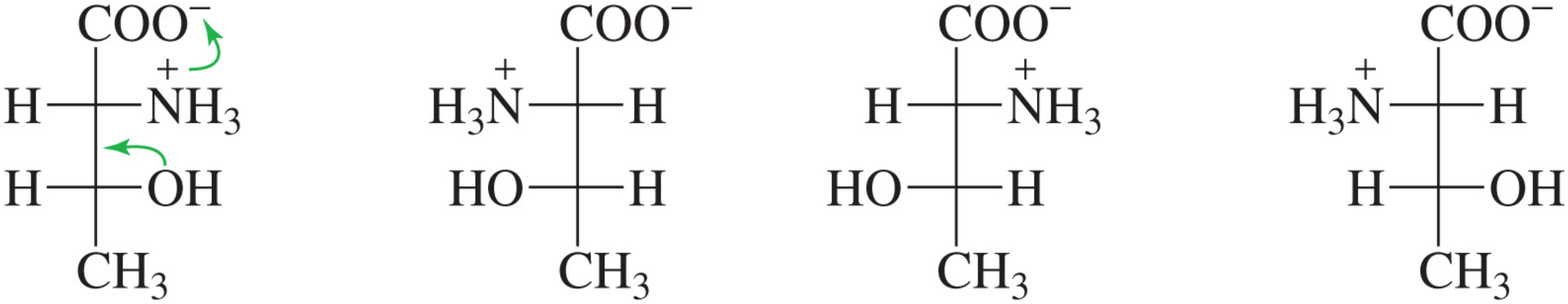
2-Amino-2,3-dimethylpentanoic acid* | Isovaline | α-Methylnorvaline

▲ Amino acids found in the Murchison meteorite.

Why do L-amino acids predominate in biological systems? What process might have selected L-amino acids over their D- counterparts?

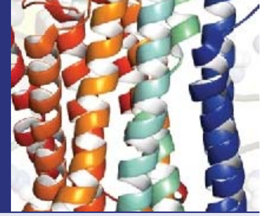
The meteorite found near Murchison, Australia may provide answers. Certain amino acids found in the meteorite have been found to have L-enantiomeric excesses of 2% to 9%.

Which is (2*S*,3*R*)-threonine?



1 2 3 4

4.2 What Are Acid-Base Properties of Amino Acids?

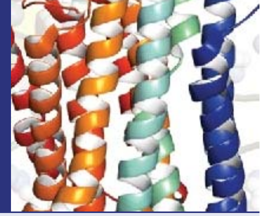


- *Amino Acids are Weak Polyprotic Acids*
- *The degree of dissociation depends on the pH of the medium*

- $H_2A^+ + H_2O \rightarrow HA^0 + H_3O^+$

$$K_{a1} = \frac{[HA^0][H_3 0^+]}{[H_2A^+]}$$

4.2 What Are Acid-Base Properties of Amino Acids?



The second dissociation (the amino group in the case of glycine):

- $HA^0 + H_2O \rightarrow A^- + H_3O^+$

$$K_{a2} = \frac{[A^-][H_3O^+]}{[HA^0]}$$

4.2 What Are Acid-Base Properties of Amino Acids?

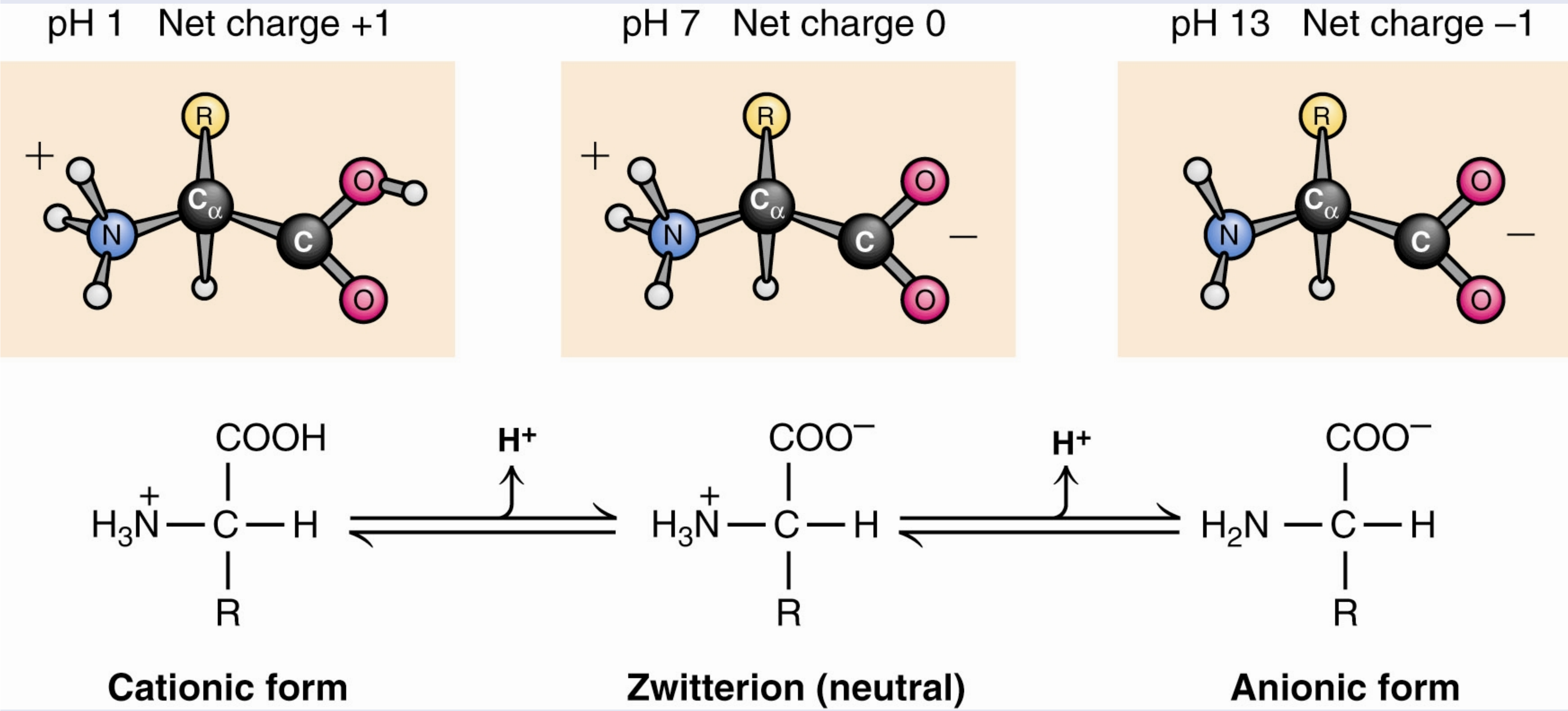
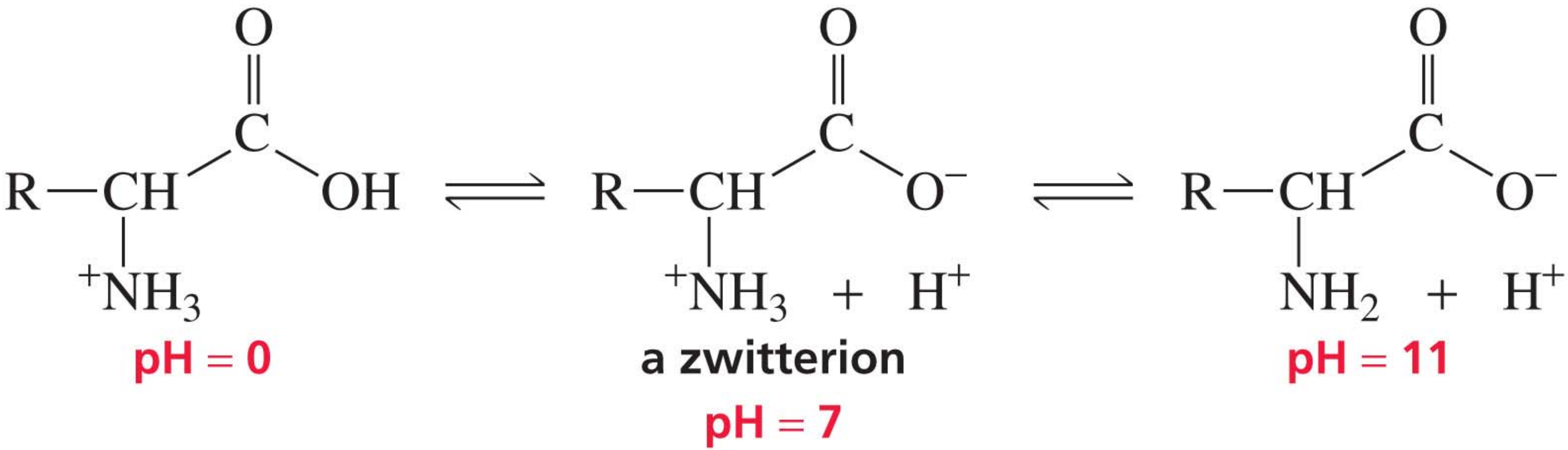
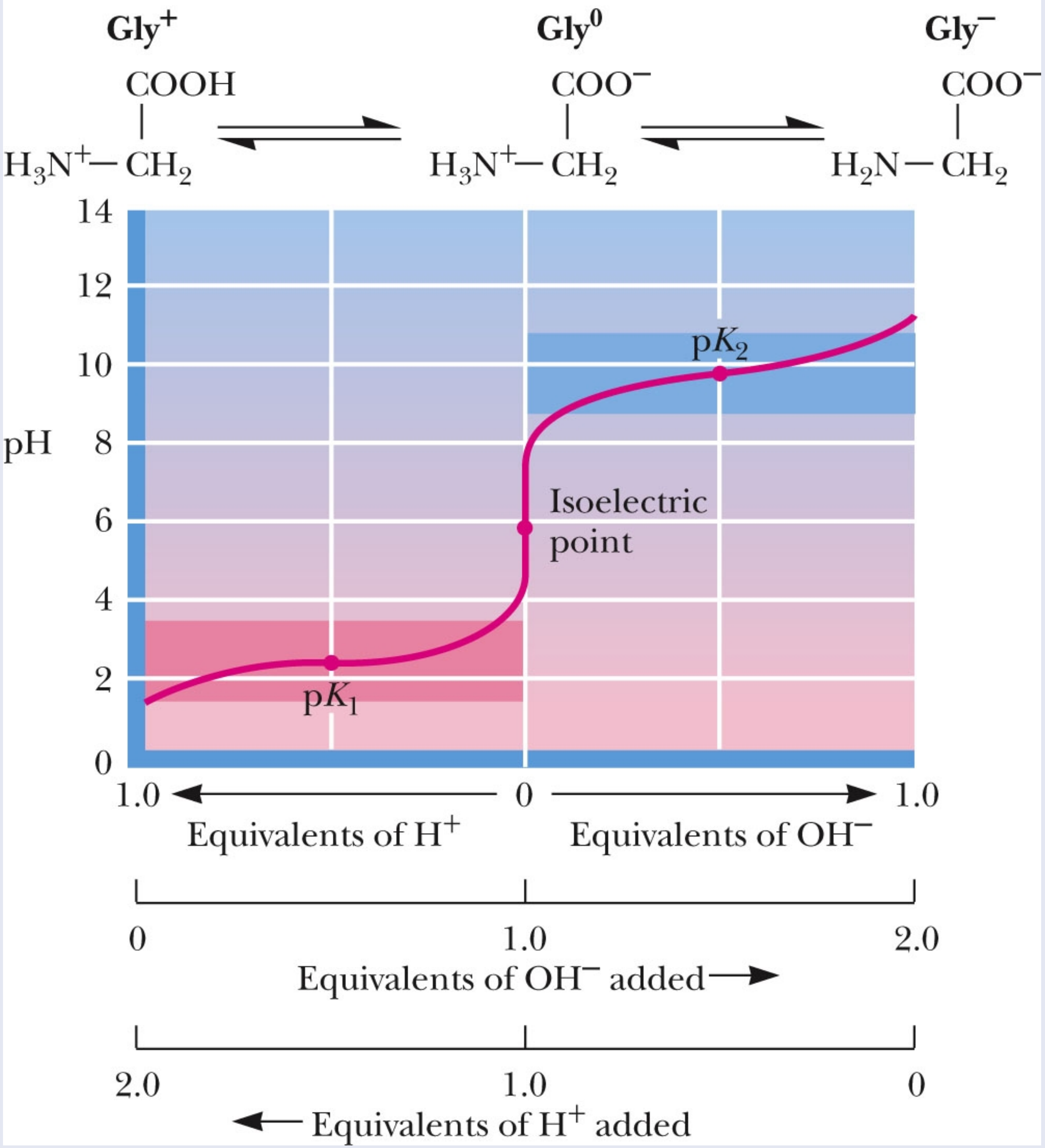


Figure 4.5 The ionic forms of the amino acids, shown without consideration of any ionizations on the side chain.

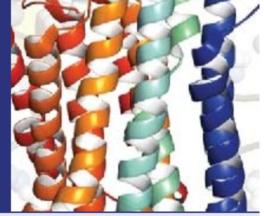


The adjacent α-amino group makes the α-COOH group more acidic.

4.2 What Are Acid-Base Properties of Amino Acids?



pK_a Values of the Amino Acids

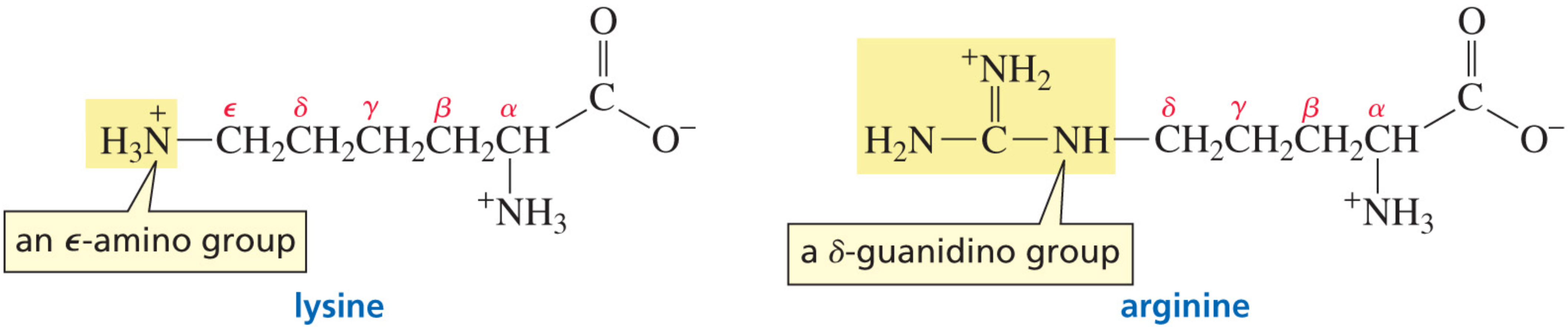


You should know these numbers and know what they mean

- Alpha carboxyl group - $pK_a = 2$
- Alpha amino group - $pK_a = 9$
- These numbers are approximate, but entirely suitable for our purposes.

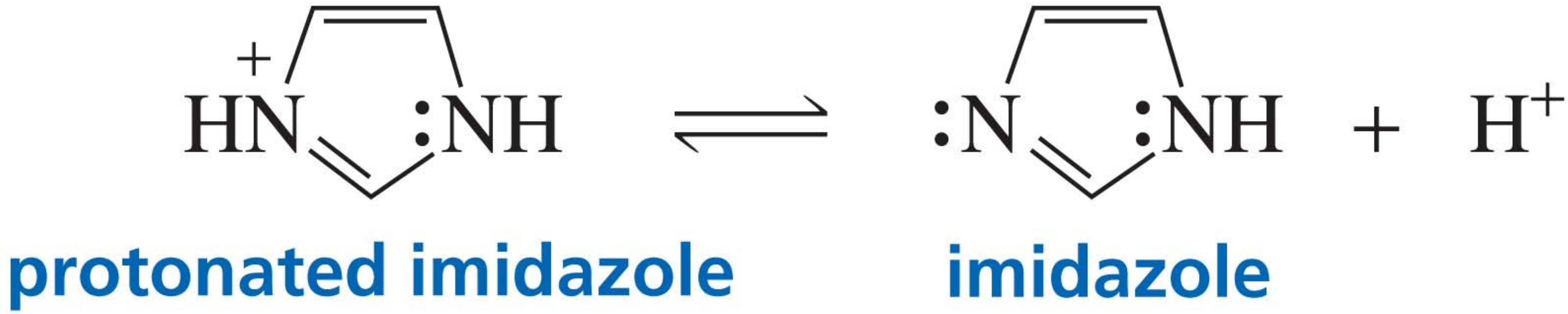
Table 22.3 The pK_a Values of Amino Acids

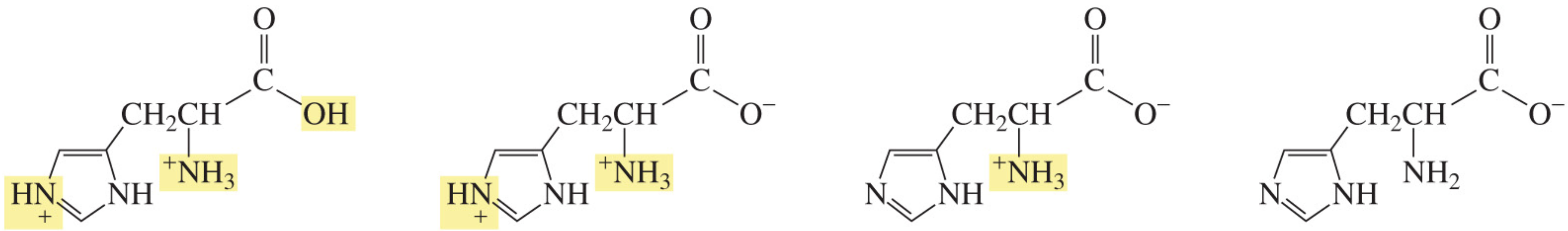
Amino acid	pK_a α-COOH	pK_a α-NH_3^+	pK_a side chain
Alanine	2.34	9.69	—
Arginine	2.17	9.04	12.48
Asparagine	2.02	8.84	—
Aspartic acid	2.09	9.82	3.86
Cysteine	1.92	10.46	8.35
Glutamic acid	2.19	9.67	4.25
Glutamine	2.17	9.13	—
Glycine	2.34	9.60	—
Histidine	1.82	9.17	6.04
Isoleucine	2.36	9.68	—
Leucine	2.36	9.60	—
Lysine	2.18	8.95	10.79
Methionine	2.28	9.21	—
Phenylalanine	2.16	9.18	—
Proline	1.99	10.60	—
Serine	2.21	9.15	—
Threonine	2.63	9.10	—
Tryptophan	2.38	9.39	—
Tyrosine	2.20	9.11	10.07
Valine	2.32	9.62	—



Protonated at physiological pH

p*K*a = 6.0

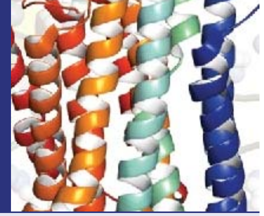




pH = 0 pH = 4 pH = 8 pH = 12

histidine

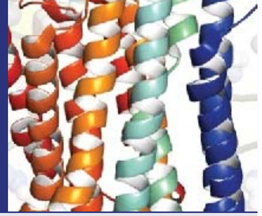
pK_a Values of the Amino Acids



You should know these numbers and know what they mean

- Arginine, Arg, R: pK_a(guanidino group) = 12.5
- Aspartic Acid, Asp, D: pK_a = 3.9
- Cysteine, Cys, C: pK_a = 8.3
- Glutamic Acid, Glu, E: pK_a = 4.3
- Histidine, His, H: pK_a = 6.0

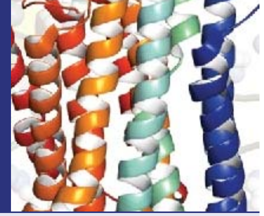
pK_a Values of the Amino Acids



You should know these numbers and know what they mean

- Lysine, Lys, K: $pK_a = 10.5$
- Serine, Ser, S: $pK_a = 13$
- Threonine, Thr, T: $pK_a = 13$
- Tyrosine, Tyr, Y: $pK_a = 10.1$

Another Sample Calculation



What is the pH of a lysine solution if the side chain amino group is 3/4 dissociated?

$$pH = 10.5 + \log_{10}\frac{[3]}{[1]}$$

- pH = 10.5 + (0.477)
- pH = 10.977 = 11.0

- Note that, when the group is ¾ dissociated, ¾ is dissociated and ¼ is not; thus the ratio in the log term is ¾ over ¼ or 3/1.

OC p60 Henderson-Hasselbalch eq

$$K_a = \frac{[H_3O^+][A^-]}{[HA]} = K_{eq}[H_2O]$$

$$pK_a = -\log K_a$$

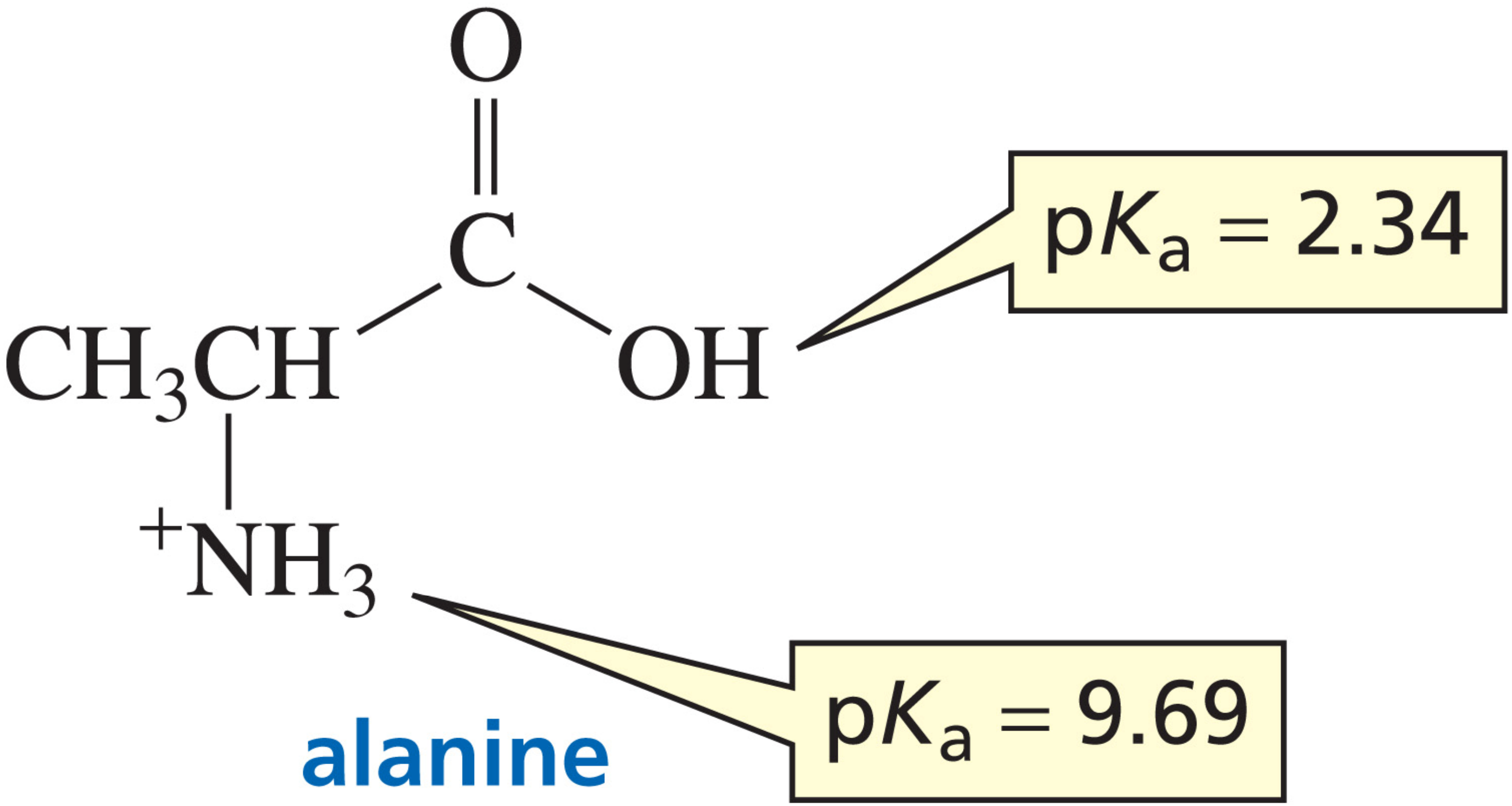
the Henderson–Hasselbalch equation

$$pK_a = pH + \log \frac{[HA]}{[A^-]}$$

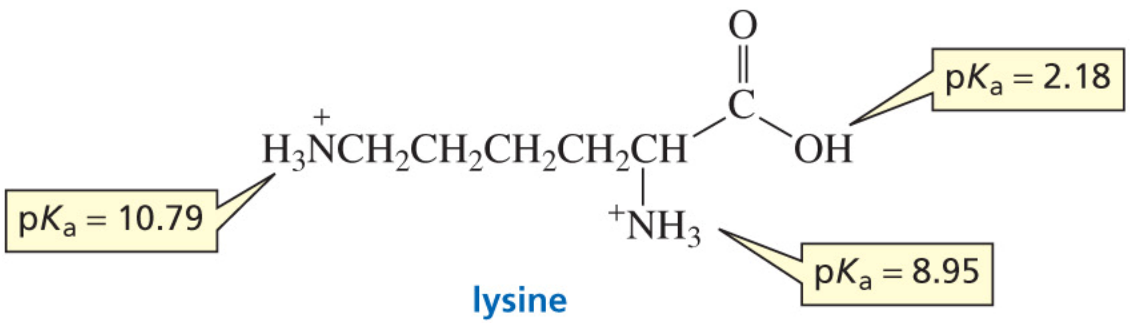
If pH < pK_a, acidic form
If pH > pKa, basic form

See p60

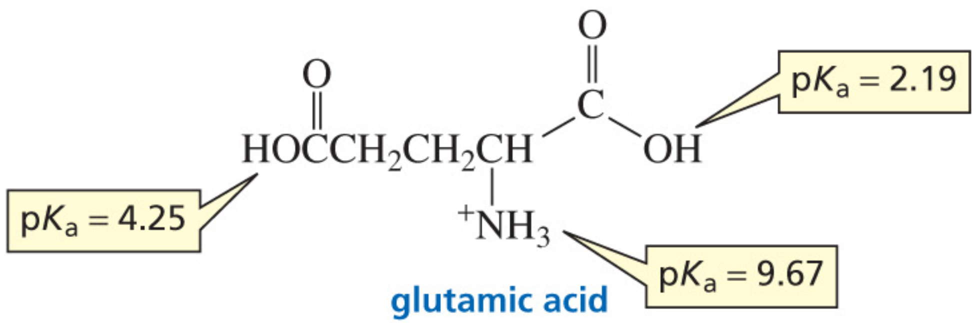
Isoelectric point (IP): no net charge



$$pI = \frac{2.34 + 9.69}{2} = \frac{12.03}{2} = 6.02$$



lysine

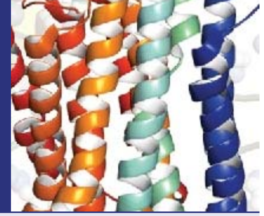


glutamic acid

$$pI = \frac{8.95 + 10.79}{2} = \frac{19.74}{2} = 9.87$$

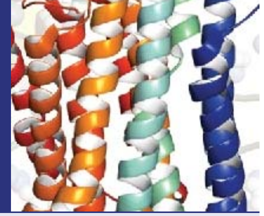
$$pI = \frac{2.19 + 4.25}{2} = \frac{6.44}{2} = 3.22$$

Reactions of Amino Acids

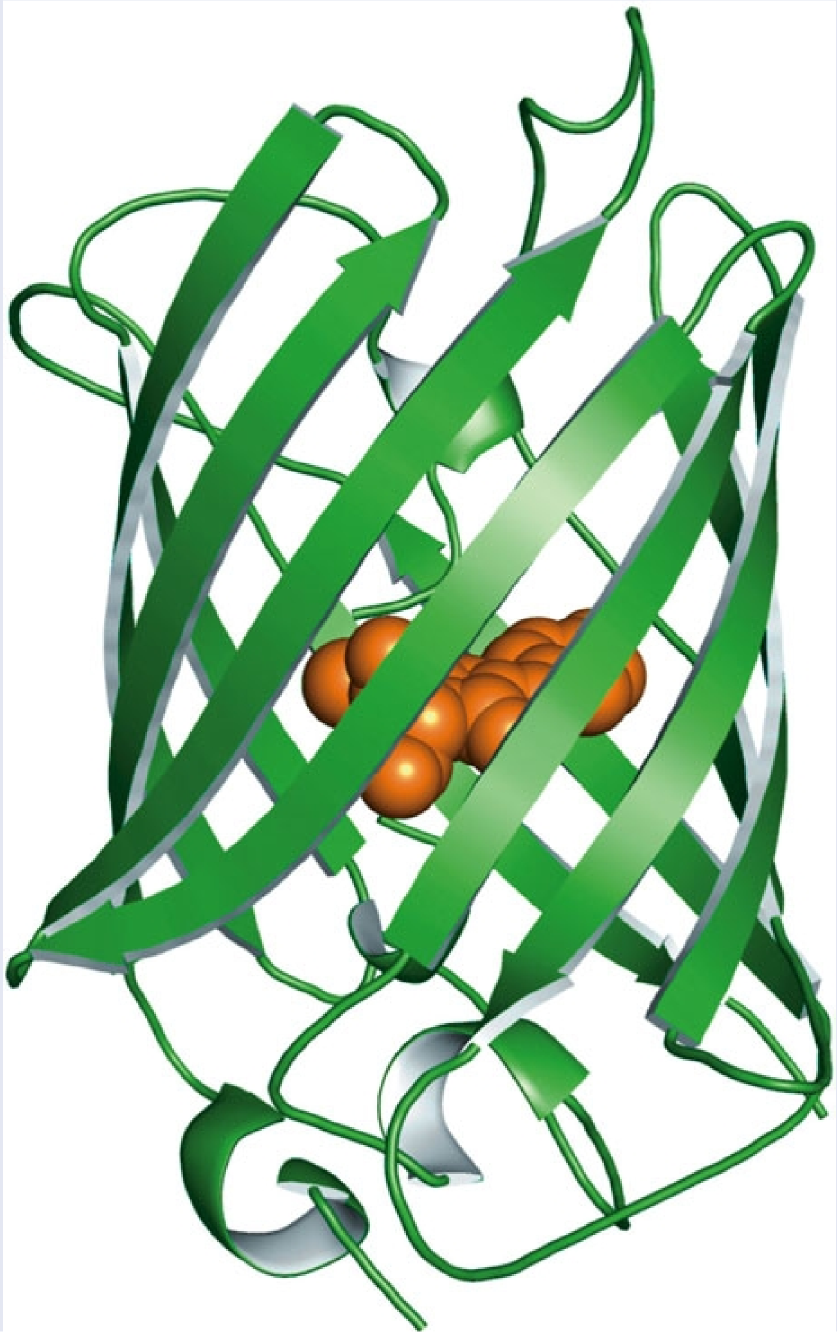


- Carboxyl groups form amides & esters
- Amino groups form Schiff bases and amides
- Edman reagent (phenylisothiocyanate) reacts with the α-amino group of an amino acid or peptide to produce a phenylthiohydantoin (PTH) derivative.
- Side chains show unique reactivities
  - Cys residues can form disulfides and can be easily alkylated
  - Few reactions are specific to a single kind of side chain

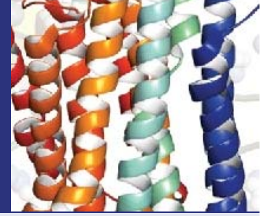
Green Fluorescent Protein

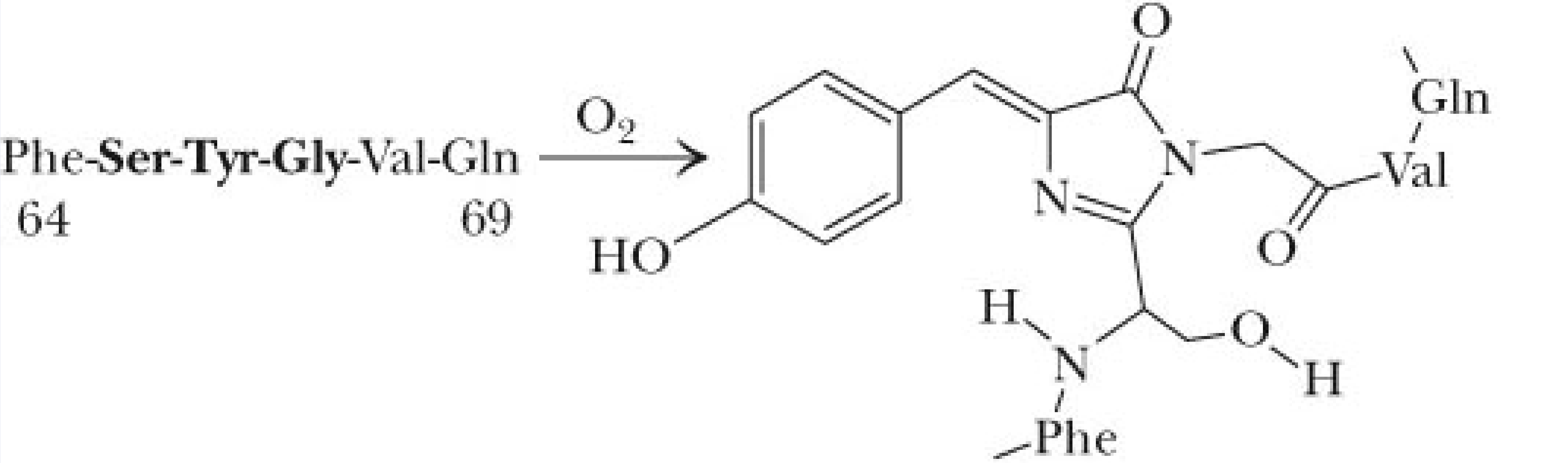


A jellyfish (Aequorea victoria) native to the northwest Pacific Ocean contains a **green fluorescent protein**. GFP is a naturally fluorescent protein. Genetic engineering techniques can be used to "tag" virtually any protein, structure, or organelle in a cell. The GFP chromophore lies in the center of a β-barrel protein structure.



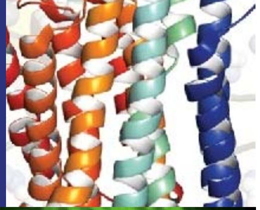
Green Fluorescent Protein



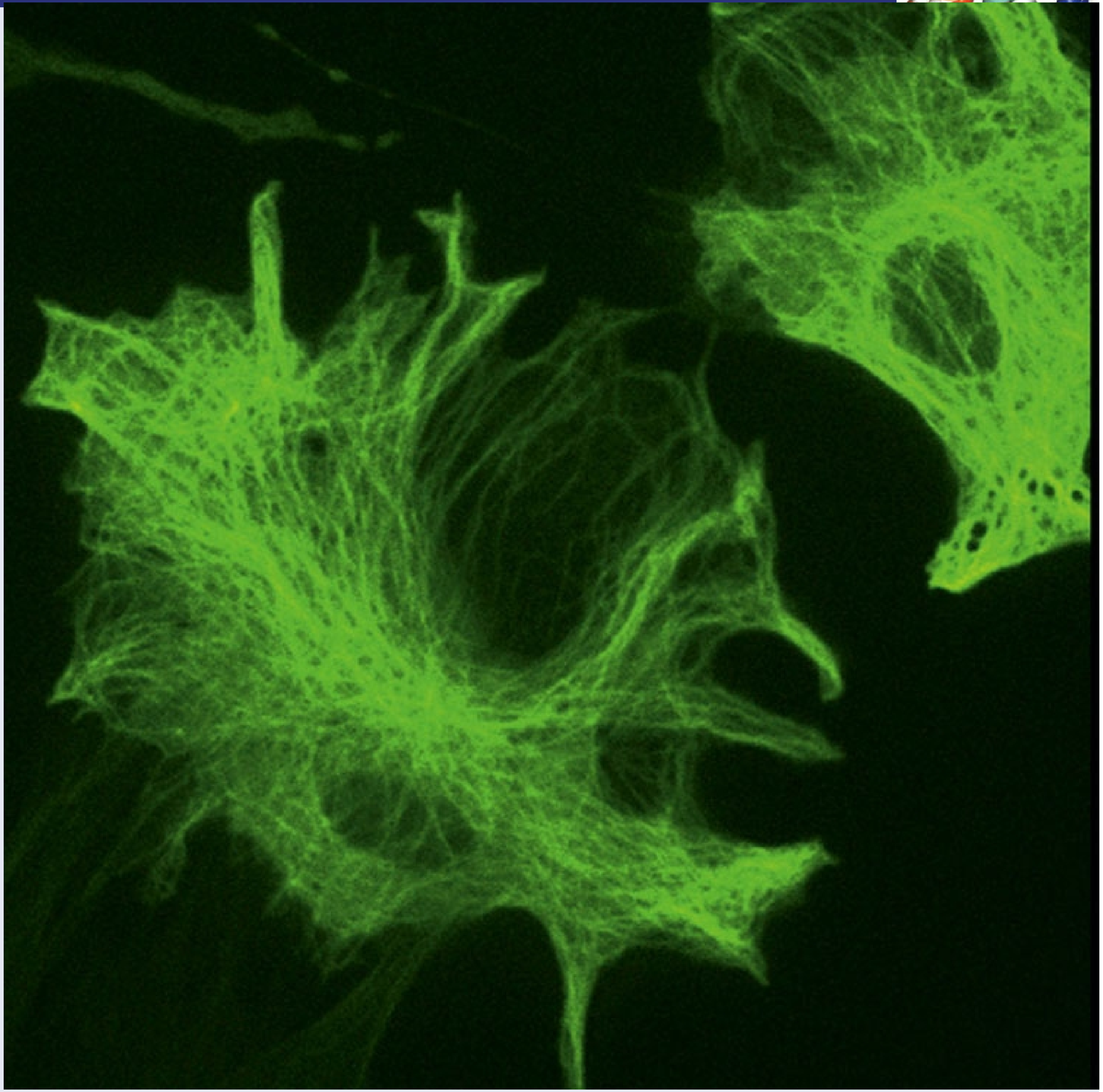


The prosthetic group of GFP is an oxidative product of the sequence –FSYGVQ-.

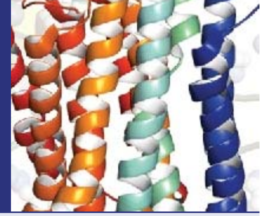
Yellow fluorescent protein



Amino acid substitutions in GFP can tune the color of emitted light. Shown here is an image of African green monkey kidney cells expressing yellow fluorescent protein (YFP) fused to α-tubulin, a cytoskeletal protein.



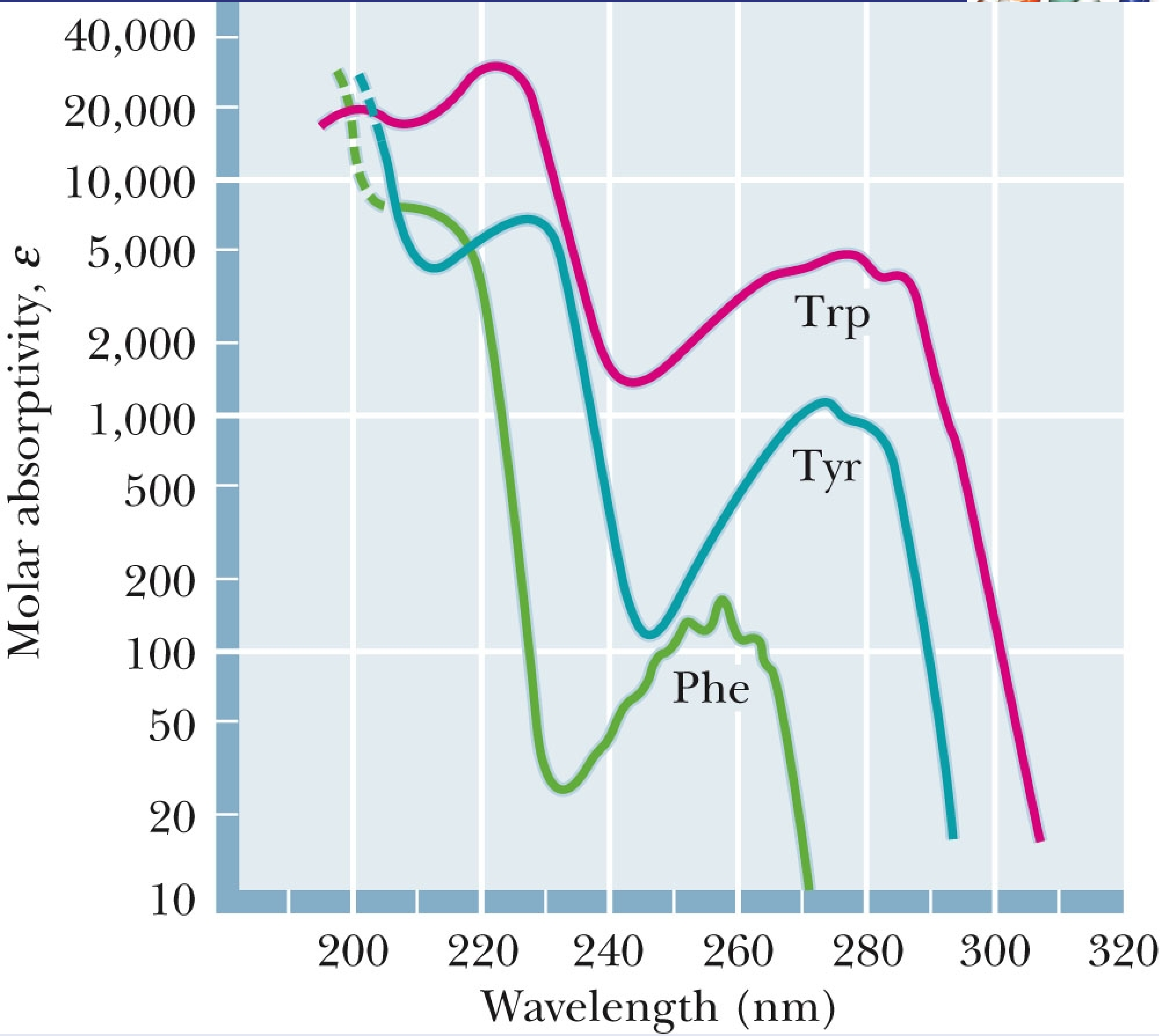
Spectroscopic Properties



- All amino acids absorb at infrared wavelengths
- Only Phe, Tyr, and Trp absorb UV
- Absorbance at 280 nm is a good diagnostic device for amino acids
- NMR spectra are characteristic of each residue in a protein, and high resolution NMR measurements can be used to elucidate three-dimensional structures of proteins

Spectroscopic Properties

Figure 4.10 The UV spectra of the aromatic amino acids at pH 6.



Spectroscopic Properties

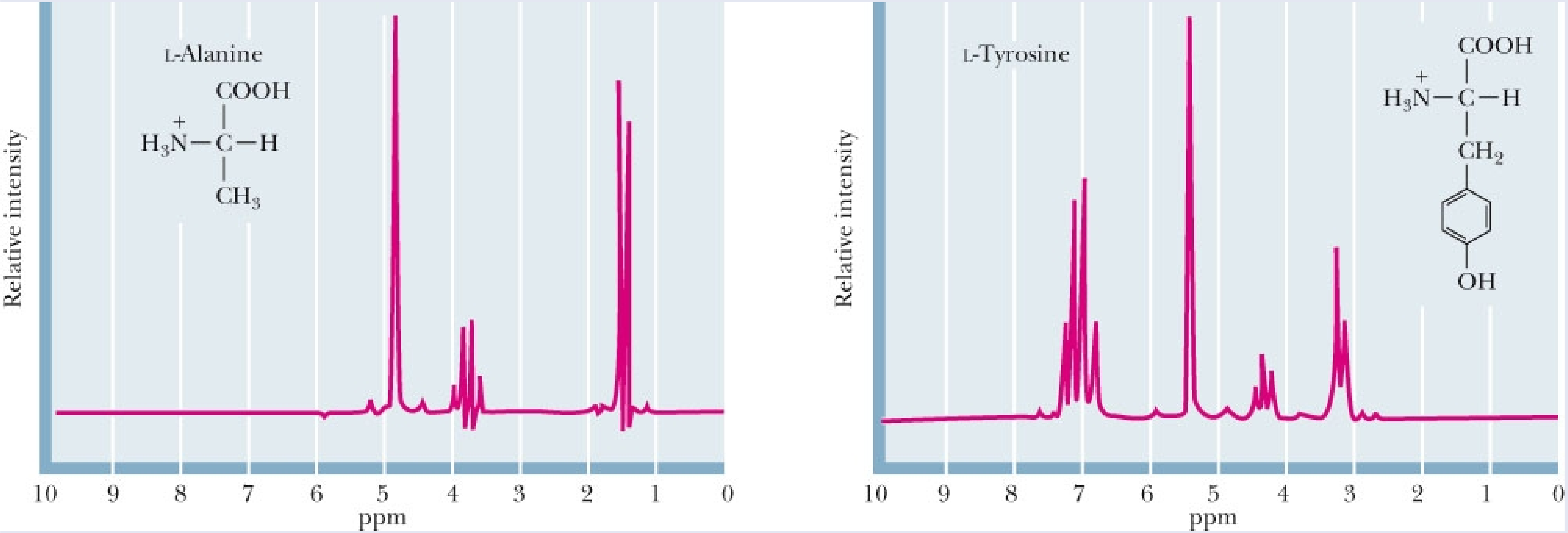
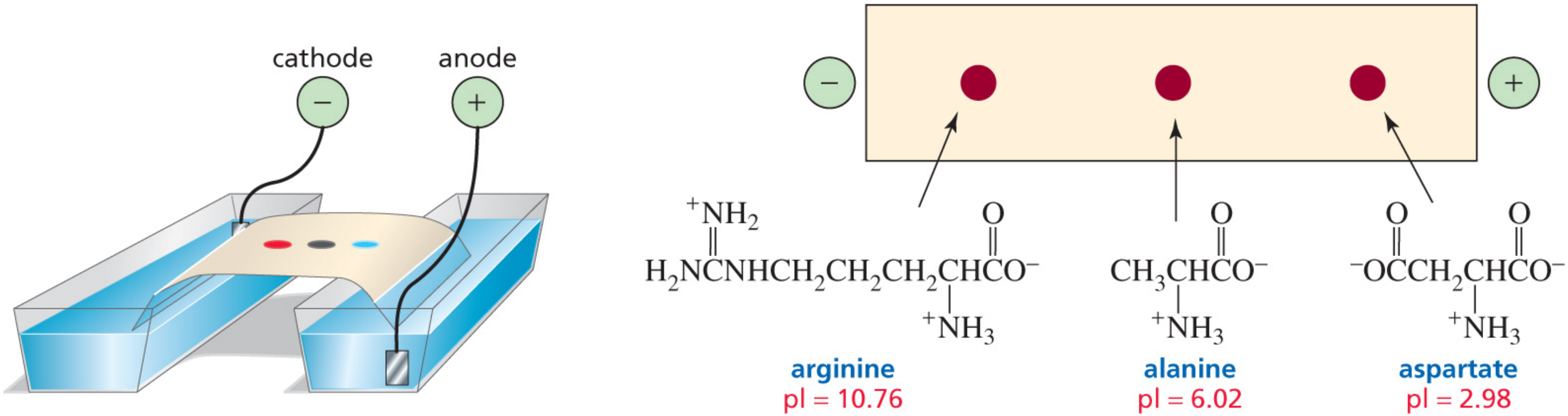


Figure 4.11 Proton NMR spectra of several amino acids.

Separation of AA

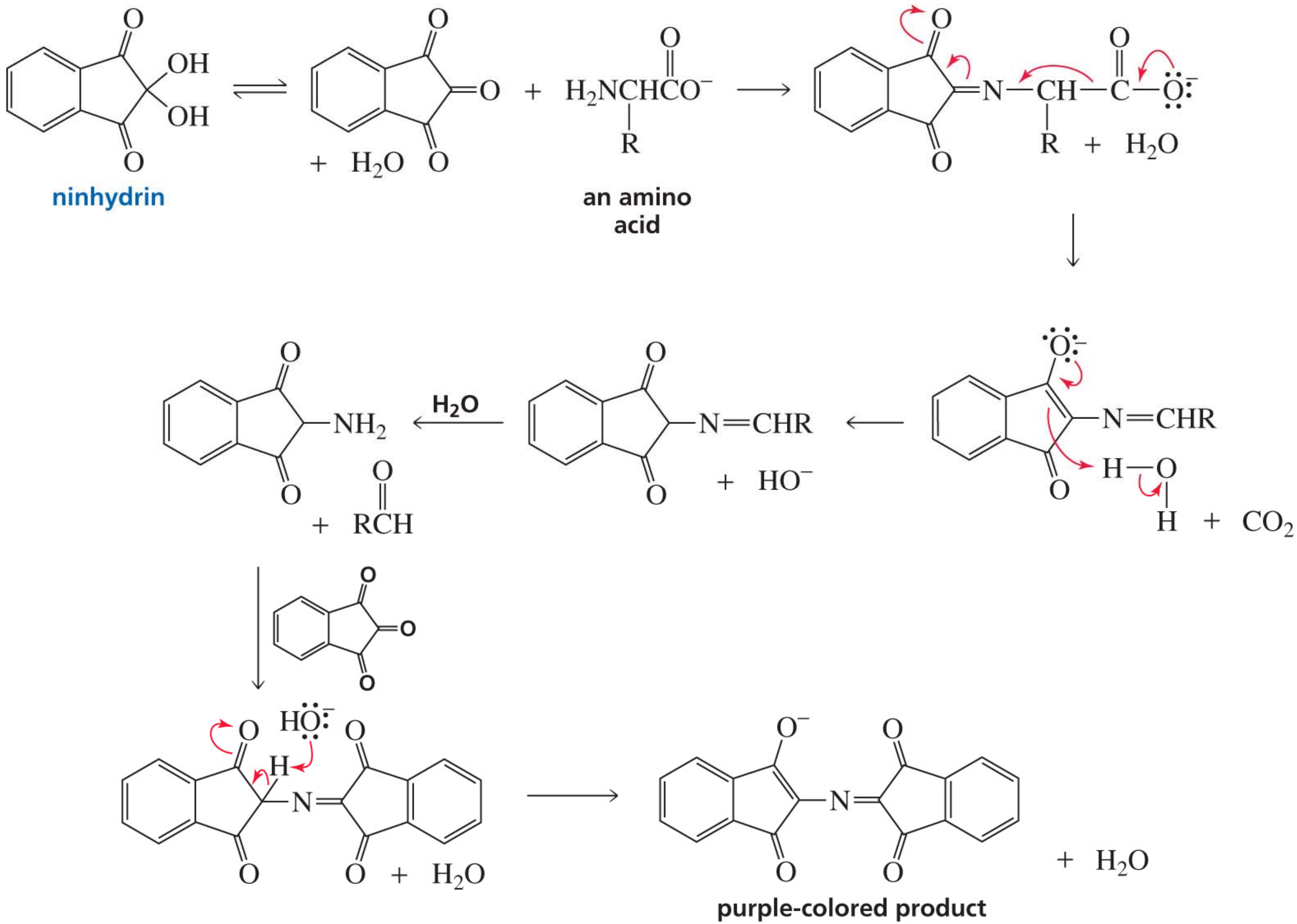
Analytical separation
Preparative separation

Electrophoresis (based on charge)



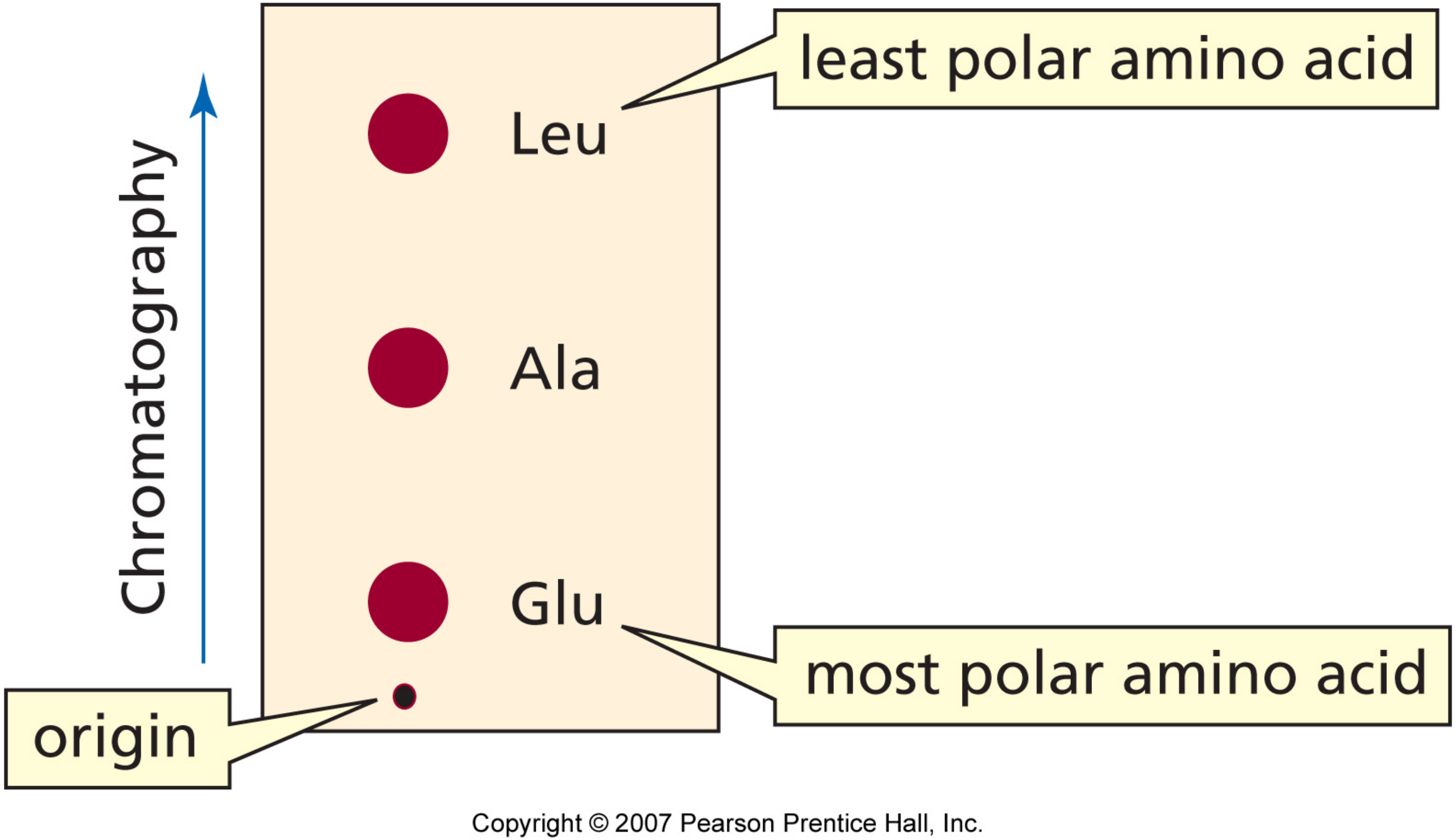
OC p1028

Ninhydrin test (AA is purple)



Paper Chromatograhpy and Thin-Layer Chromatography

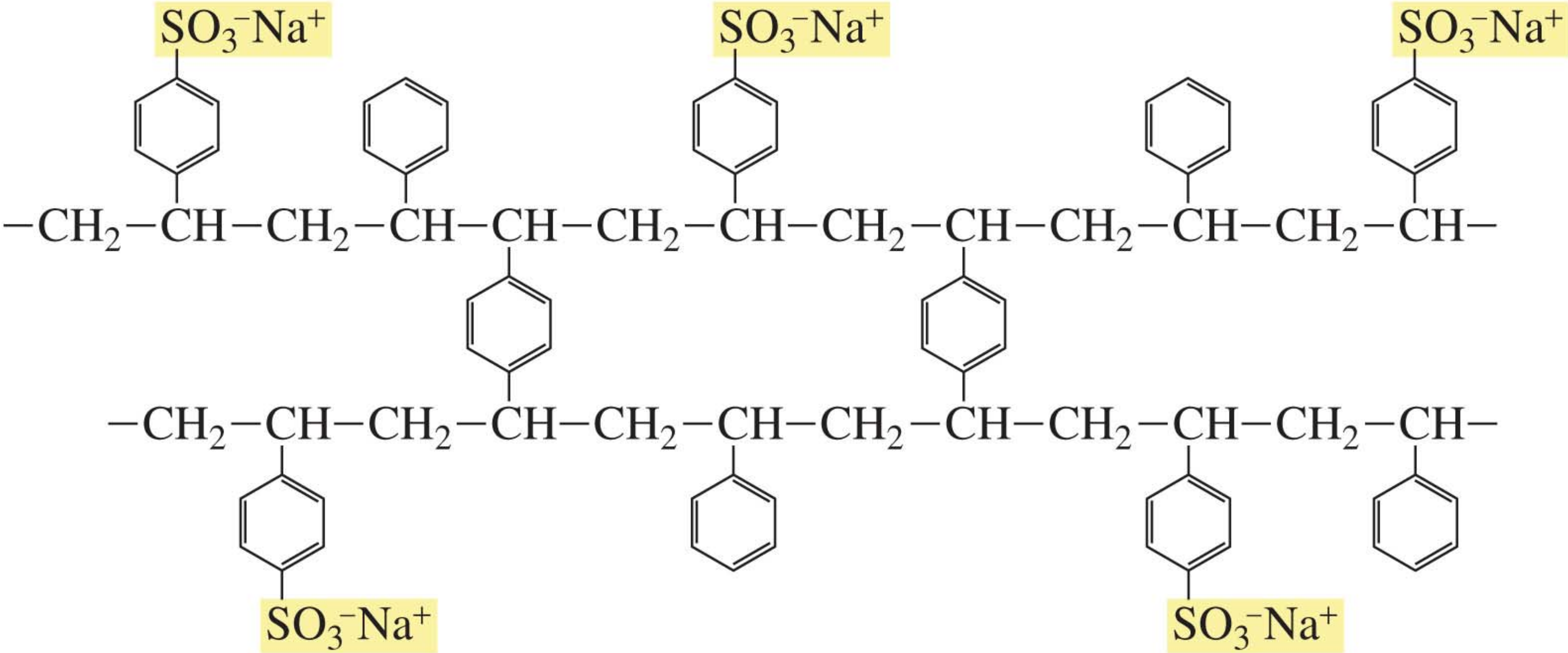
(based on polarity)



Solvent: H_2O/AcOH/BuOH

MIT 5.301 Chemistry Laboratory Techniques
http://www.youtube.com/watch?v=EUn2skAAjHk

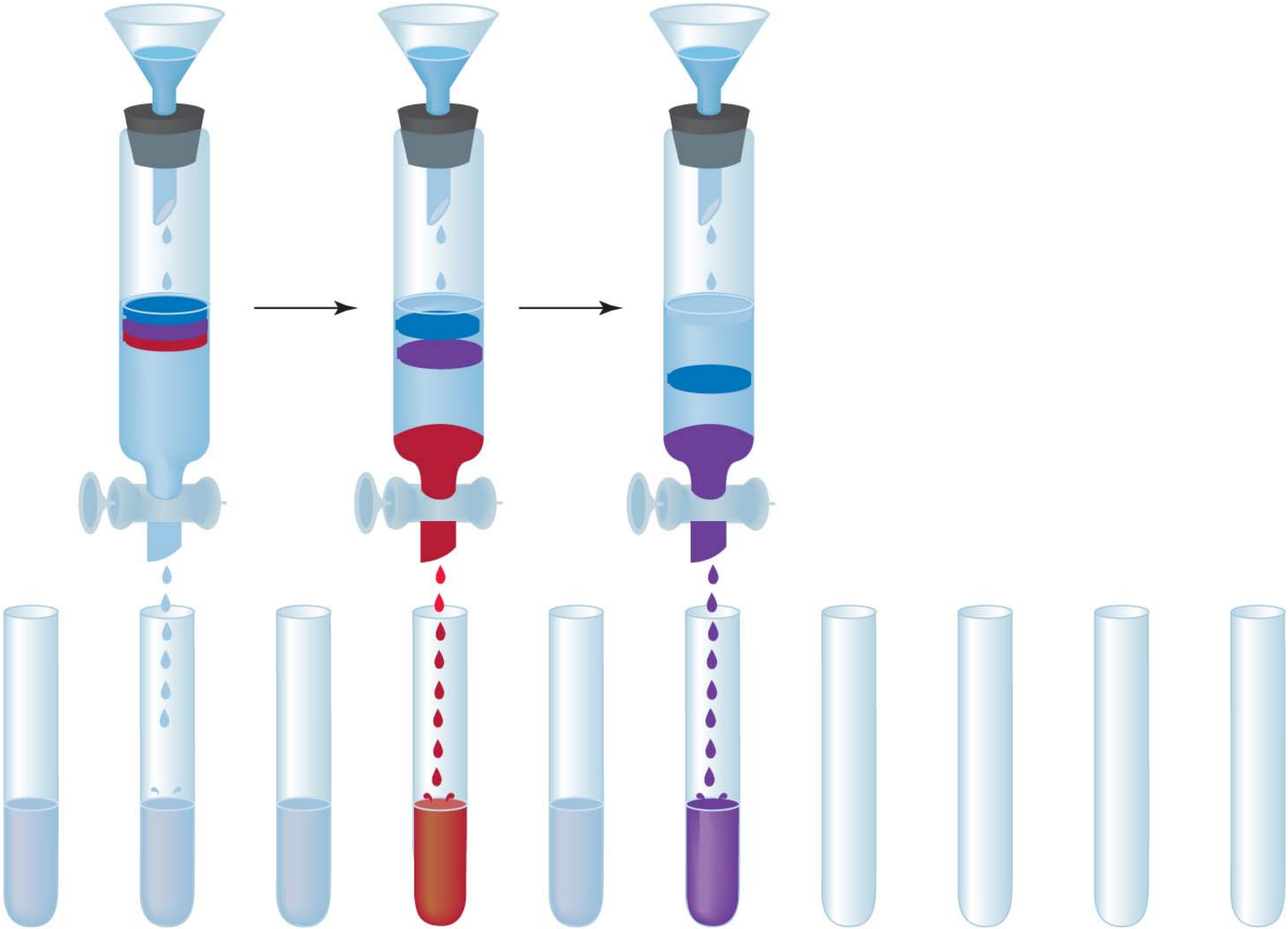
Ion-Exchange Chromatography



Cation-exchange resin

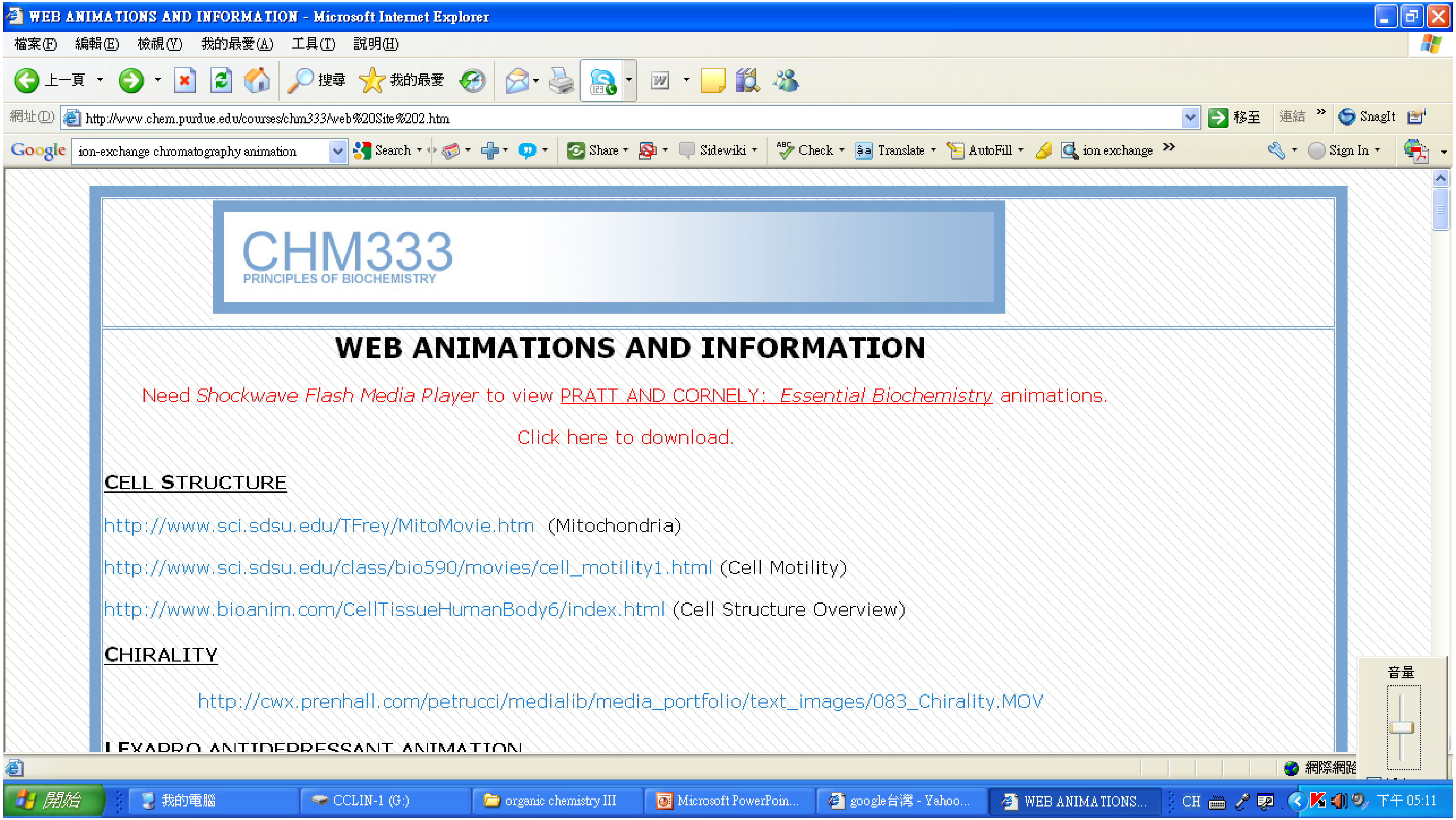
Nonpolar nature of the column caused it to retain nonpolar amino acids longer than polar amino acids.

Water softener OC p1032

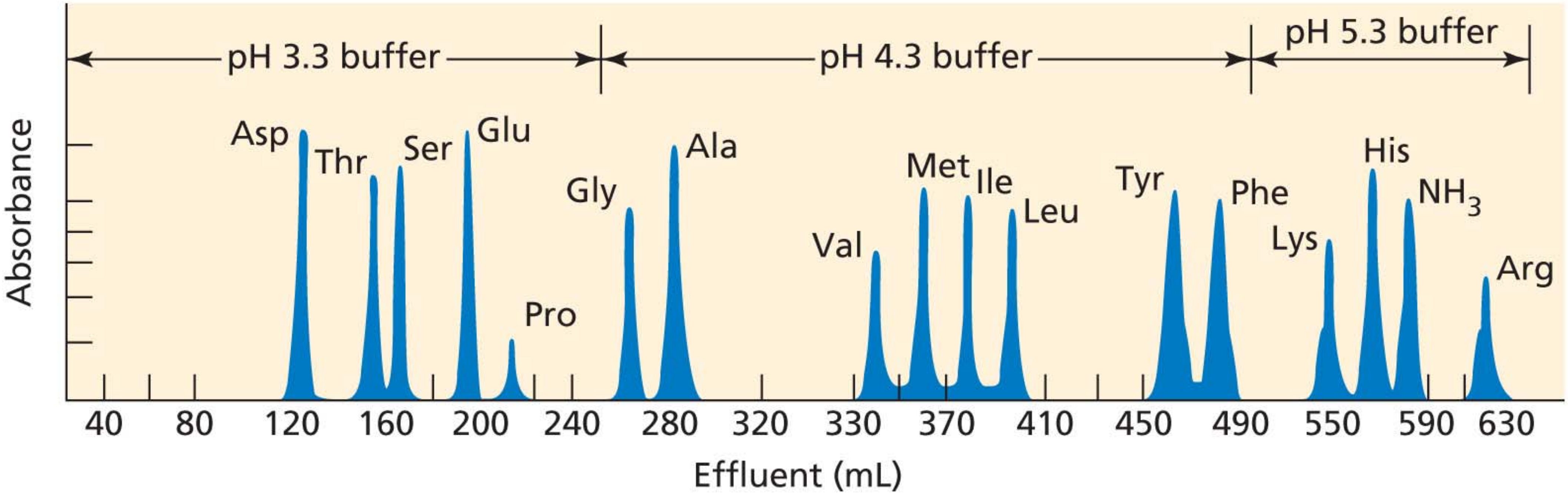


Fractions sequentially collected

http://www.chem.purdue.edu/courses/chm333/web%20Site%202.htm



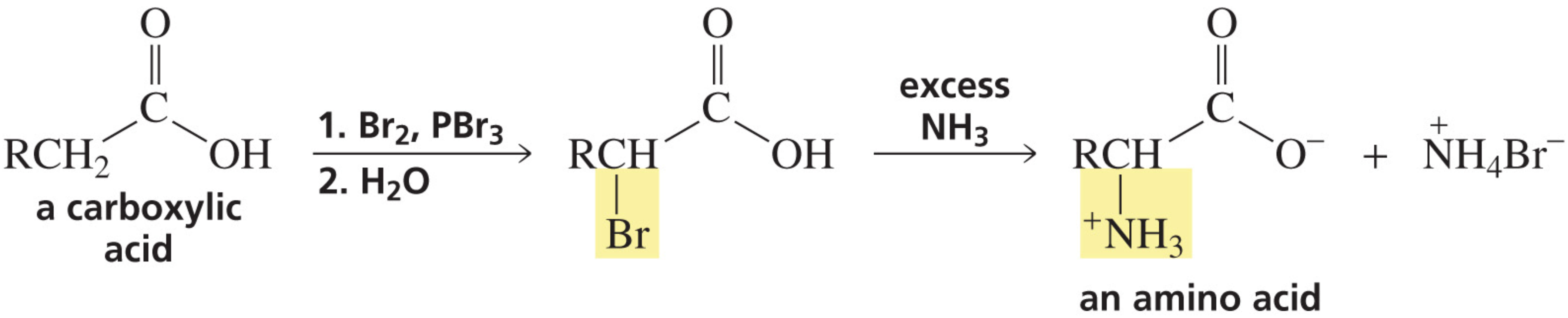
After addition of ninhydrin

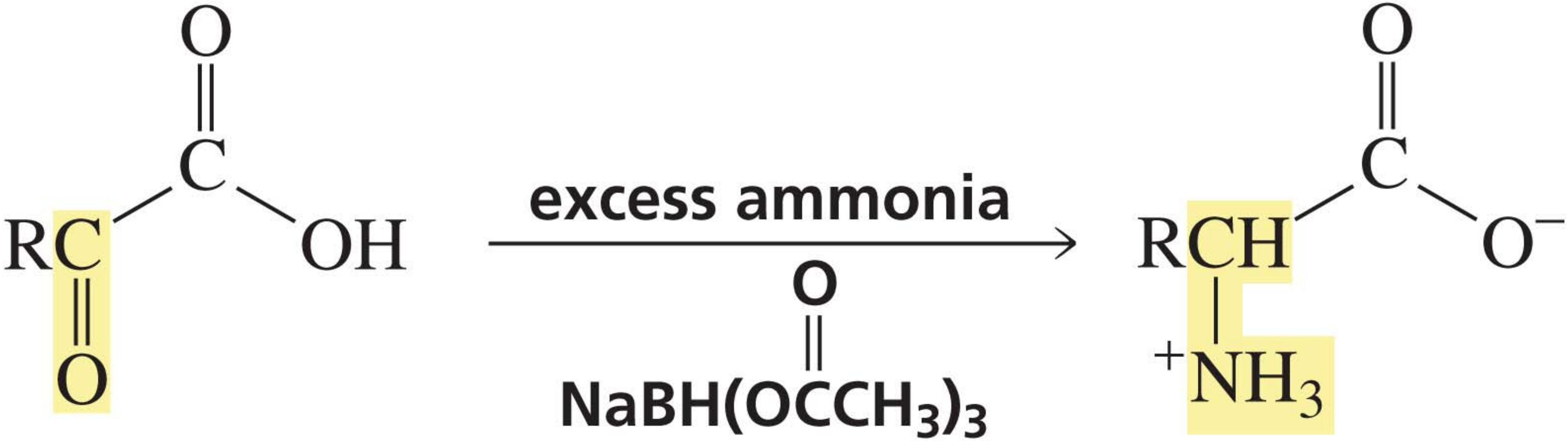


At 570 nm

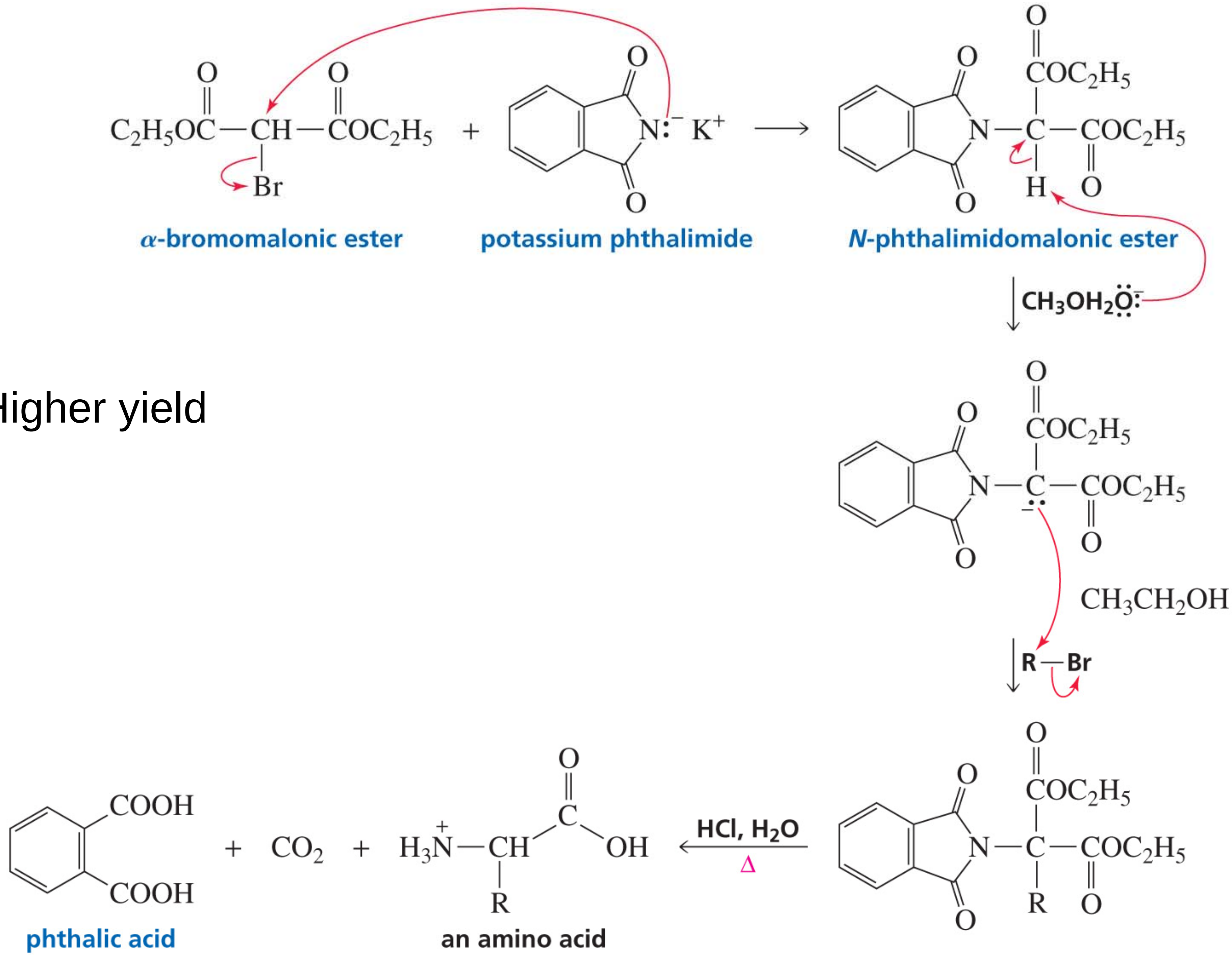
Synthesis of Amino acids

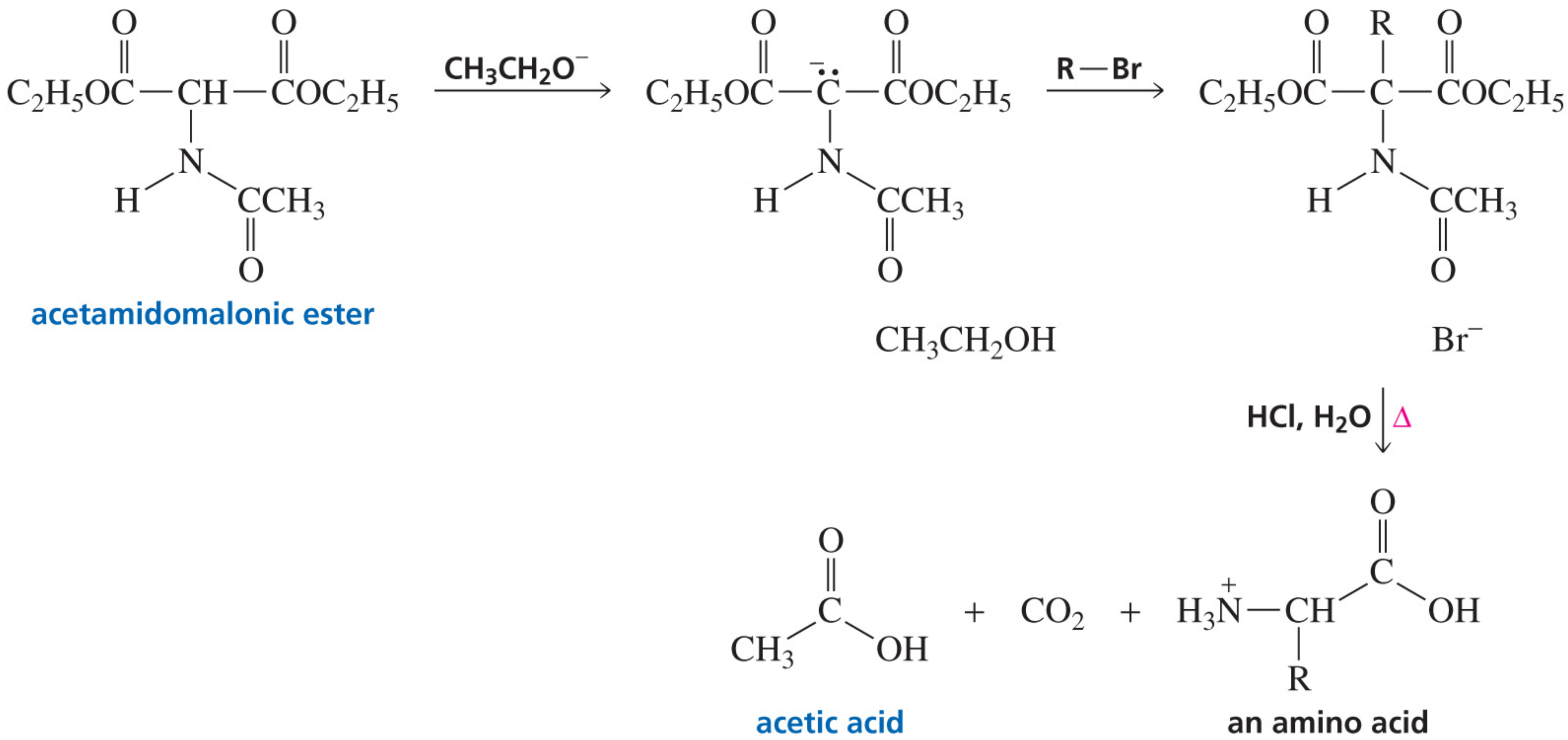
Hell-Volhard-Zelinski reaction




excess ammonia
NaBH(OCCH3)3
RC
OH
RCH
O⁻
+NH3

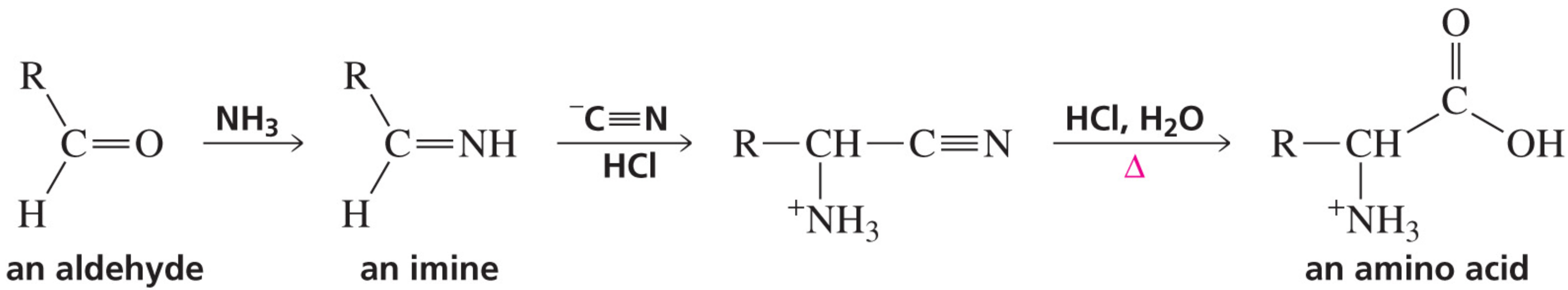
Higher yield





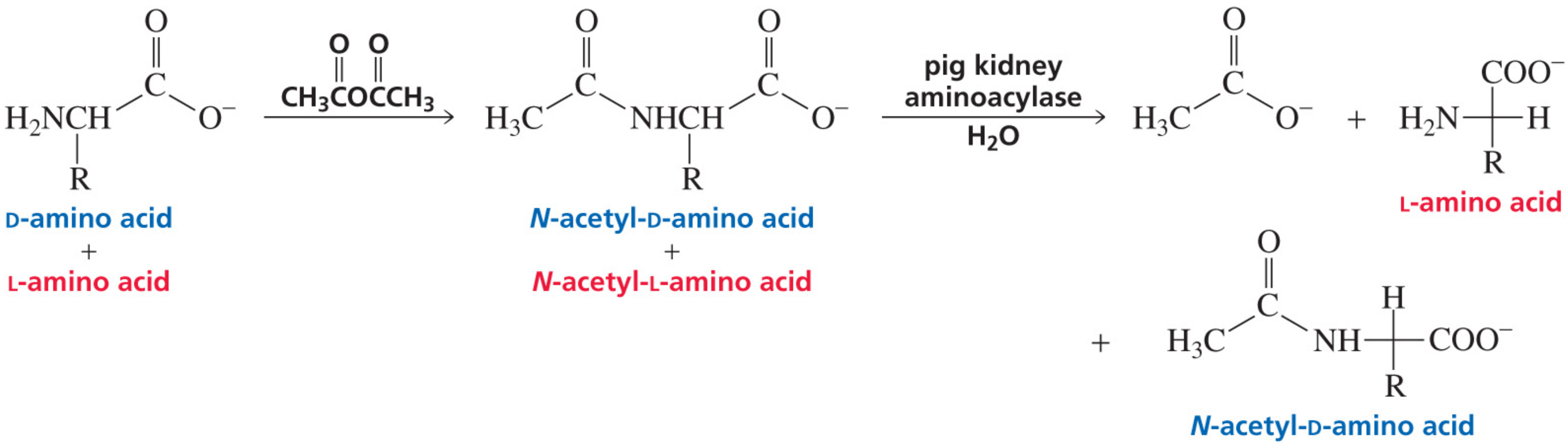
CH3CH2O−
R—Br
acetamidomalonic ester
CH3CH2OH
Br−
HCl, H2O
Δ
acetic acid
an amino acid

Strecker synthesis

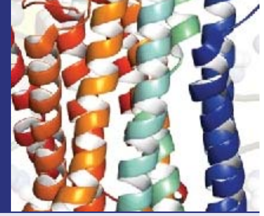


Kinetic resolution

In **kinetic resolution**, two enantiomers show different reaction rates in a chemical reaction, thereby creating an excess of the less reactive enantiomer. This excess goes through a maximum and disappears on full completion of the reaction. Kinetic resolution is a very old concept in organic chemistry and can be used in the organic synthesis of chiral molecules. It has been surpassed by other methods.

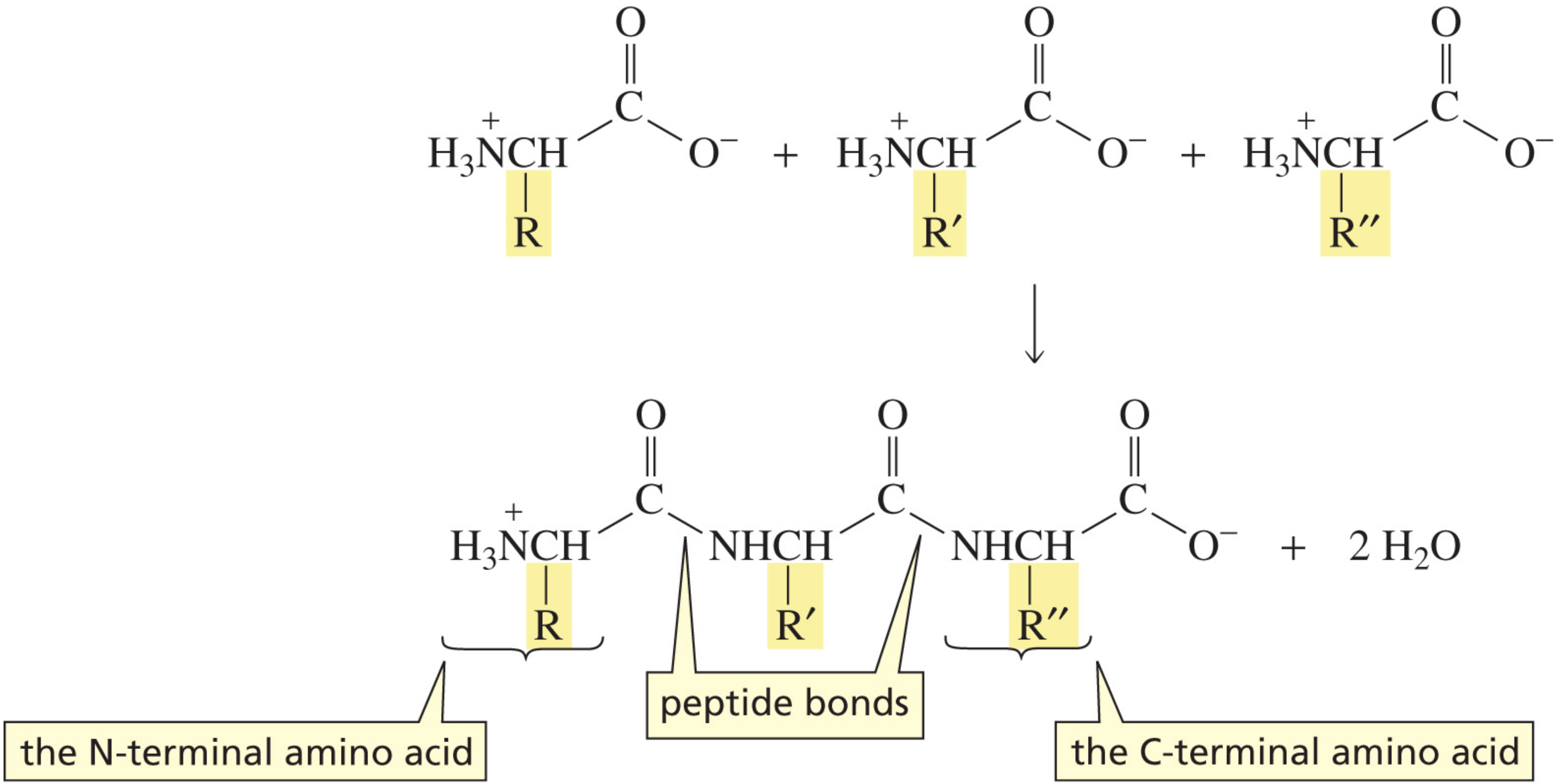


4.7 What is the Fundamental Structural Pattern in Proteins?



- Proteins are unbranched polymers of amino acids
- Amino acids join head-to-tail through formation of covalent peptide bonds
- Peptide bond formation results in release of water
- The peptide backbone of a protein consists of the repeated sequence $–N-C_{\alpha}-C_{o}-$
- “N” is the amide nitrogen of the amino acid
- “C_{α}” is the alpha-C of the amino acid
- “C_{o}” is the carbonyl carbon of the amino acid

peptide bond (amide bond)



a tripeptide

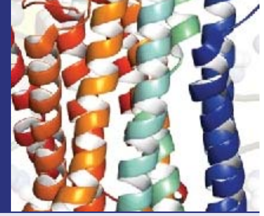
Glu, Cys, His, Val, Ala

the pentapeptide contains the indicated amino acids, but their sequence is not known

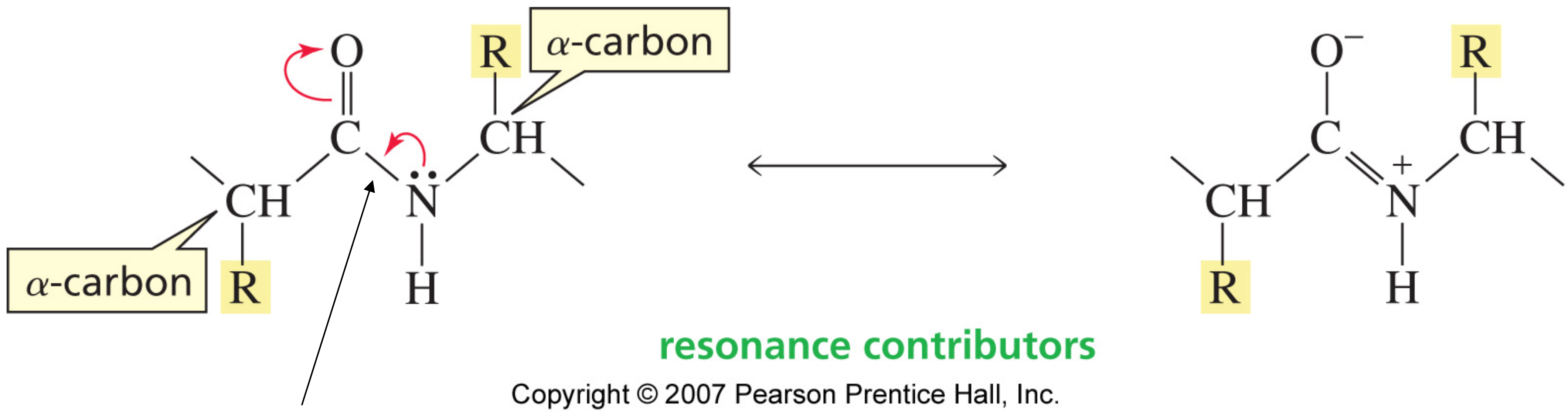
Val-Cys-Ala-Glu-His

the amino acids in the pentapeptide have the indicated sequence

The Peptide Bond

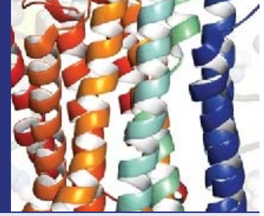


- Is usually found in the *trans* conformation
- Has partial (40%) double bond character
- Is about 0.133 nm long - shorter than a typical single bond but longer than a double bond
- Due to the double bond character, the six atoms of the peptide bond group are always planar
- N partially positive; O partially negative



40% double-bond character

4.7 What is the Fundamental Structural Pattern in Proteins?



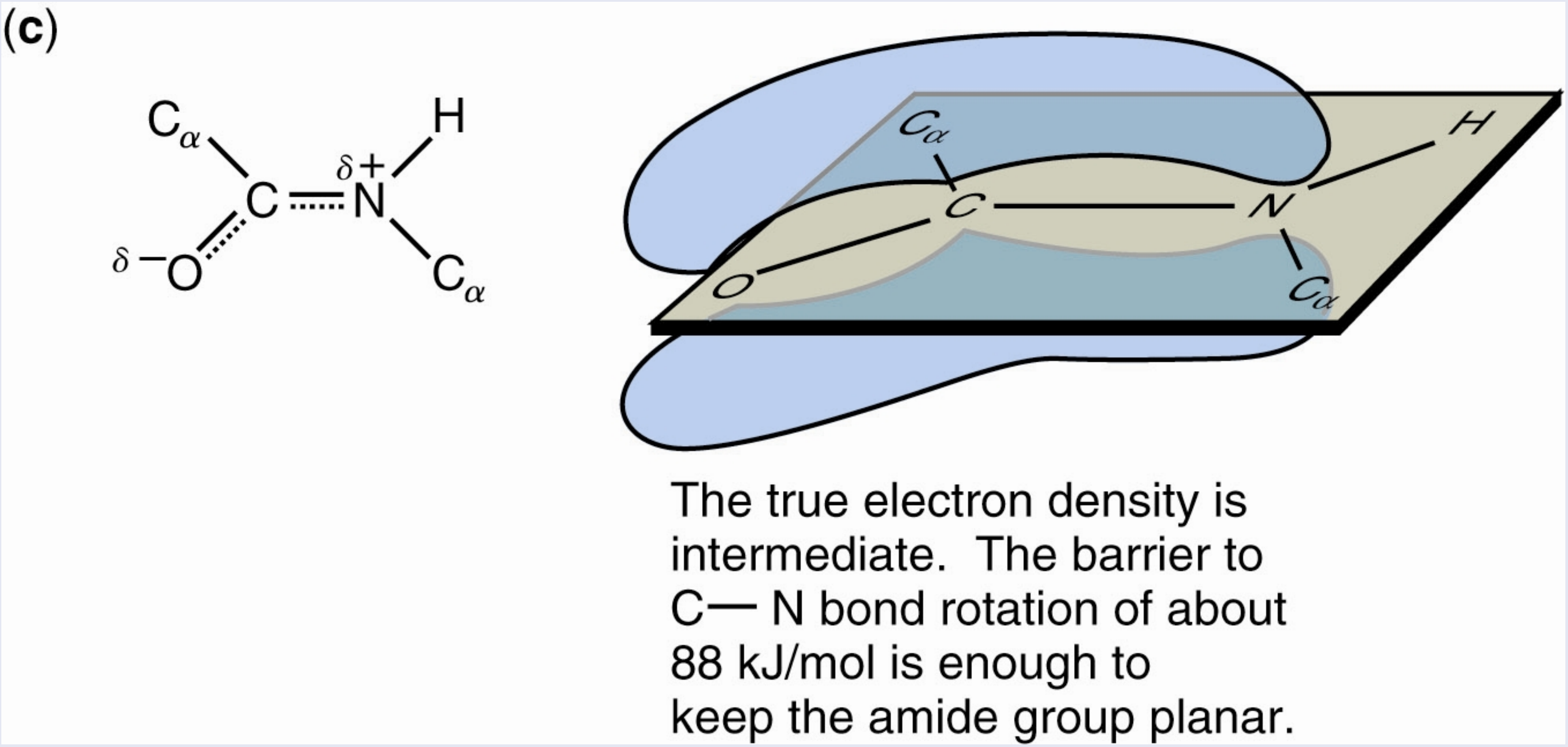


Figure 4.16 (c) The peptide bond is best described as a resonance hybrid of the forms shown on the two previous slides.

The Peptide Bond

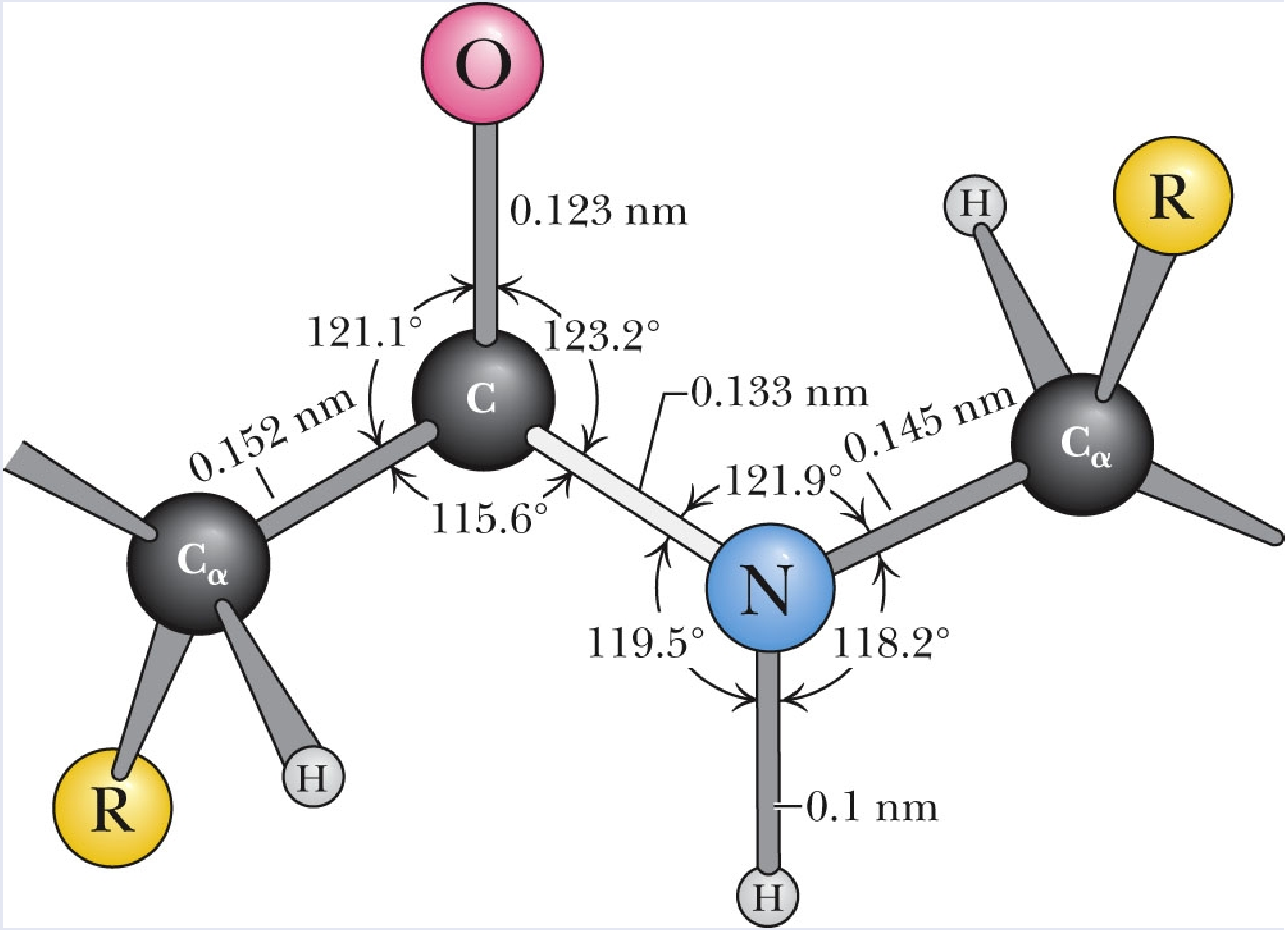
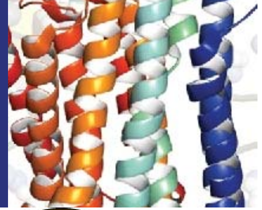
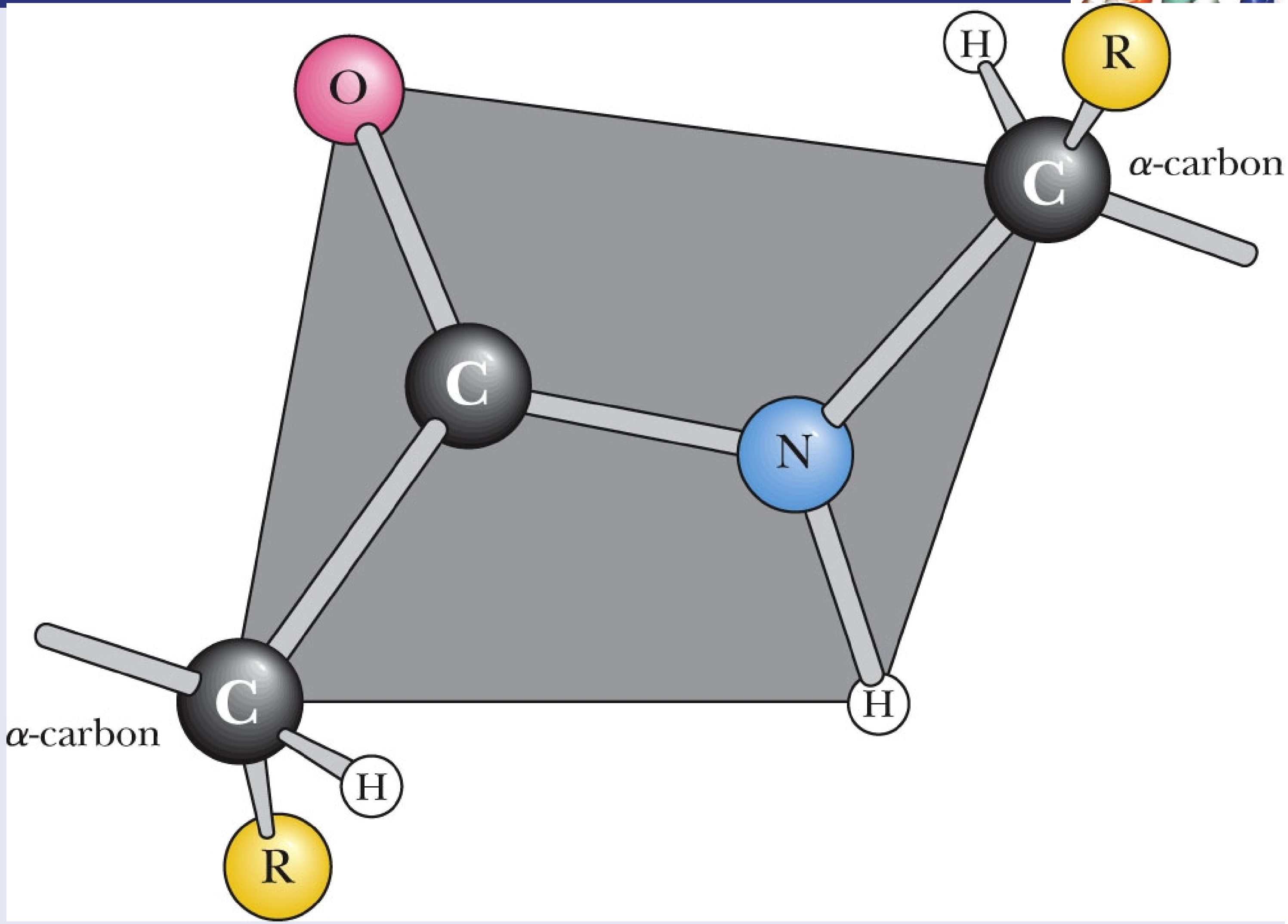


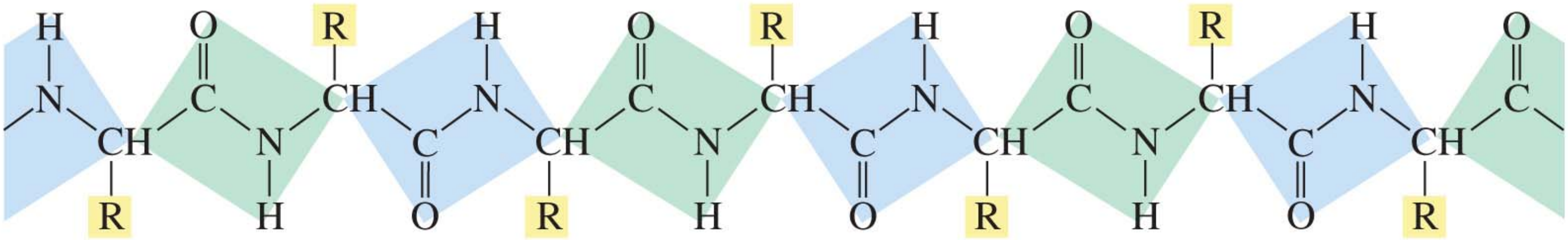
Figure 4.15 The *trans* conformation of the peptide bond.

4.7 What is the Fundamental Structural Pattern in Proteins?





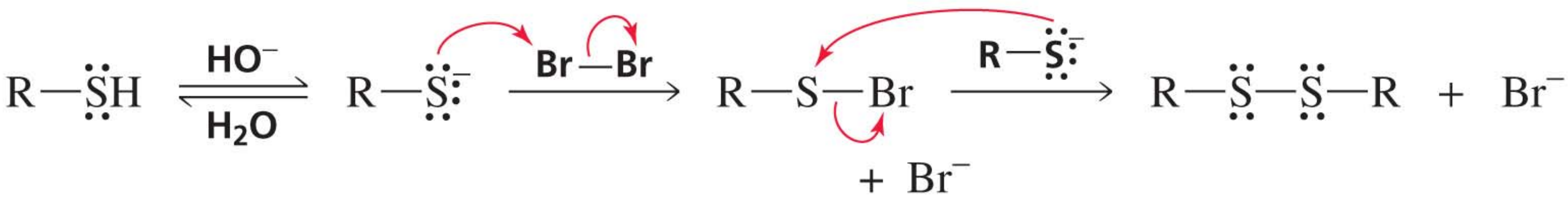
The coplanar relationship of the atoms in the amide group is highlighted here by an imaginary shaded plane lying between adjacent α-carbons.


H
O
R
H
O
R
H
O
R
H
O
N
C
CH
N
C
CH
N
C
CH
N
C
CH
N
C
CH
N
C
CH
N
C
CH
N
C
CH
R
H
O
R
H
O
R
H
O
R

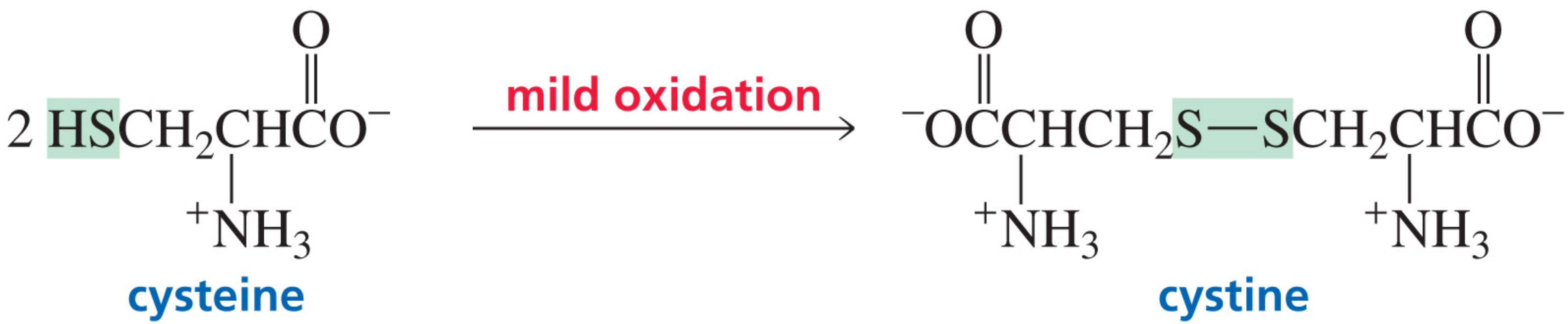
$$2\,R{-}SH \xrightarrow[\text{Br}_2\ (\text{or } I_2)]{\text{mild oxidation}} RS{-}SR$$

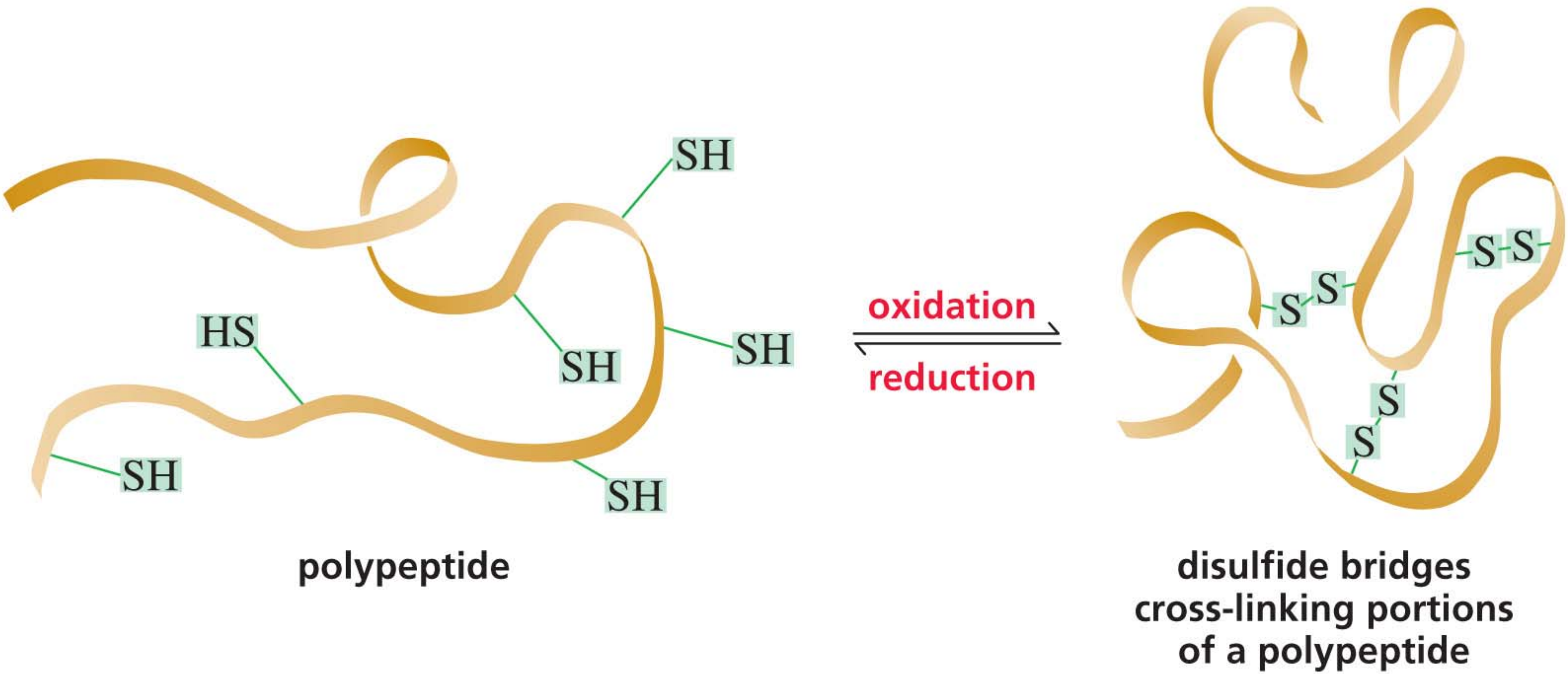
a thiol → a disulfide

mechanism for oxidation of a thiol to a disulfide

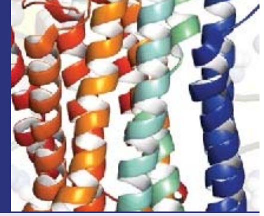


$$\underset{\text{a disulfide}}{RS-SR} \xrightarrow{\text{reduction}} \underset{\text{a thiol}}{2\,R-SH}$$


mild oxidation
cysteine
cystine

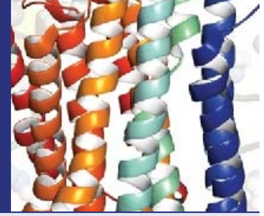

SH
HS
SH
SH
SH
SH
oxidation
reduction
S S
S S
S
S
polypeptide
disulfide bridges cross-linking portions of a polypeptide

“Peptides”



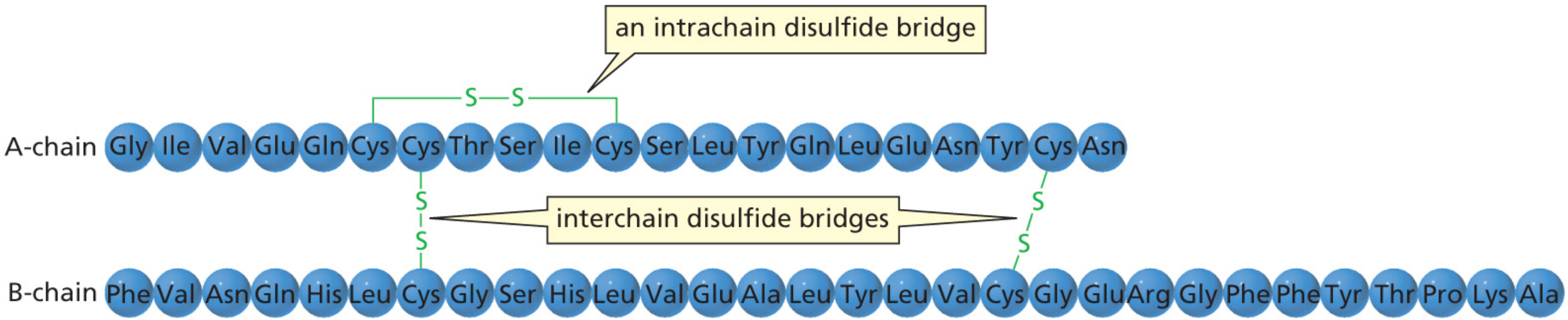
- Short polymers of amino acids
- Each unit is called a residue
- 2 residues - dipeptide
- 3 residues - tripeptide
- 12-20 residues - oligopeptide
- many - polypeptide

“Protein”



One or more polypeptide chains

- One polypeptide chain - a monomeric protein
- More than one - multimeric protein
- Homomultimer - one kind of chain
- Heteromultimer - two or more different chains
- Hemoglobin, for example, is a heterotetramer
- It has two alpha chains and two beta chains


an intrachain disulfide bridge
S—S
A-chain Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn
S
S
interchain disulfide bridges
S
S
B-chain Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Ala

insulin

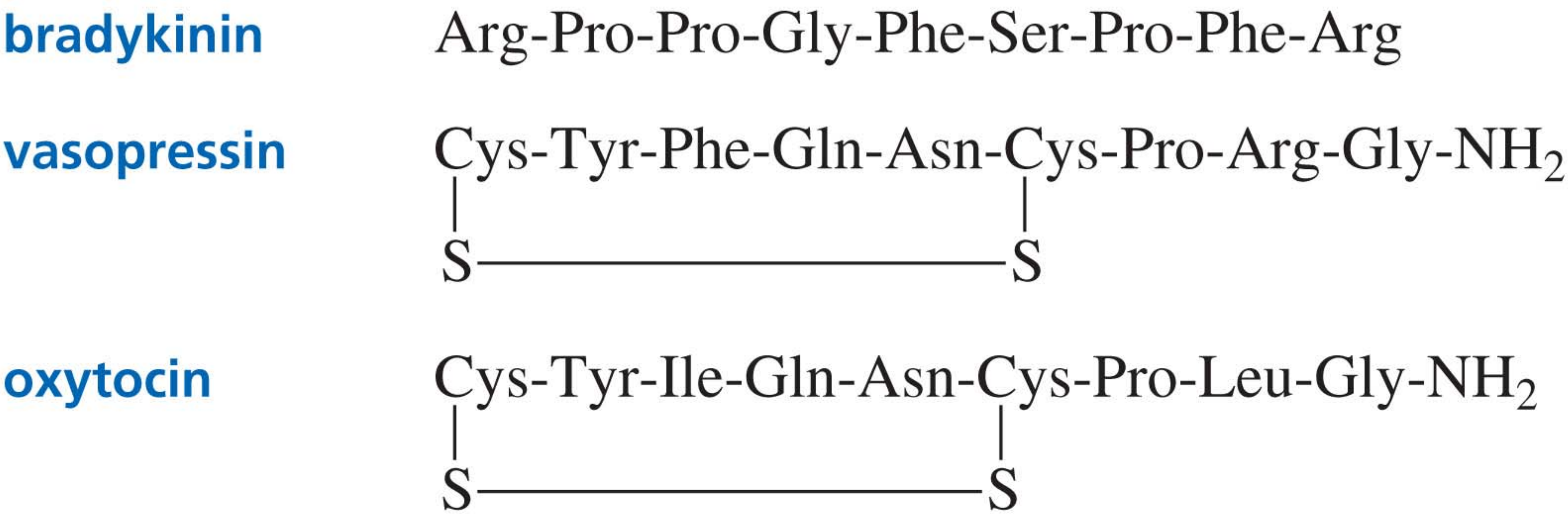
Enkephalins synthesized by the body to control pain

Tyr-Gly-Gly-Phe-Leu
leucine enkephalin

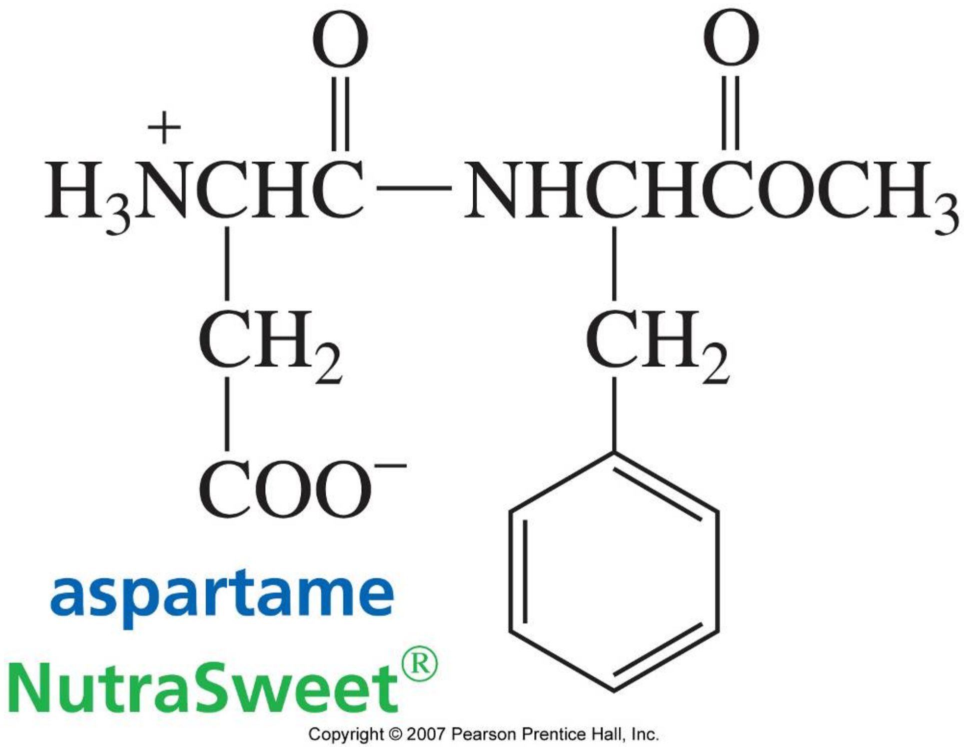
Tyr-Gly-Gly-Phe-Met
methionine enkephalin

morphine

Peptide hormones

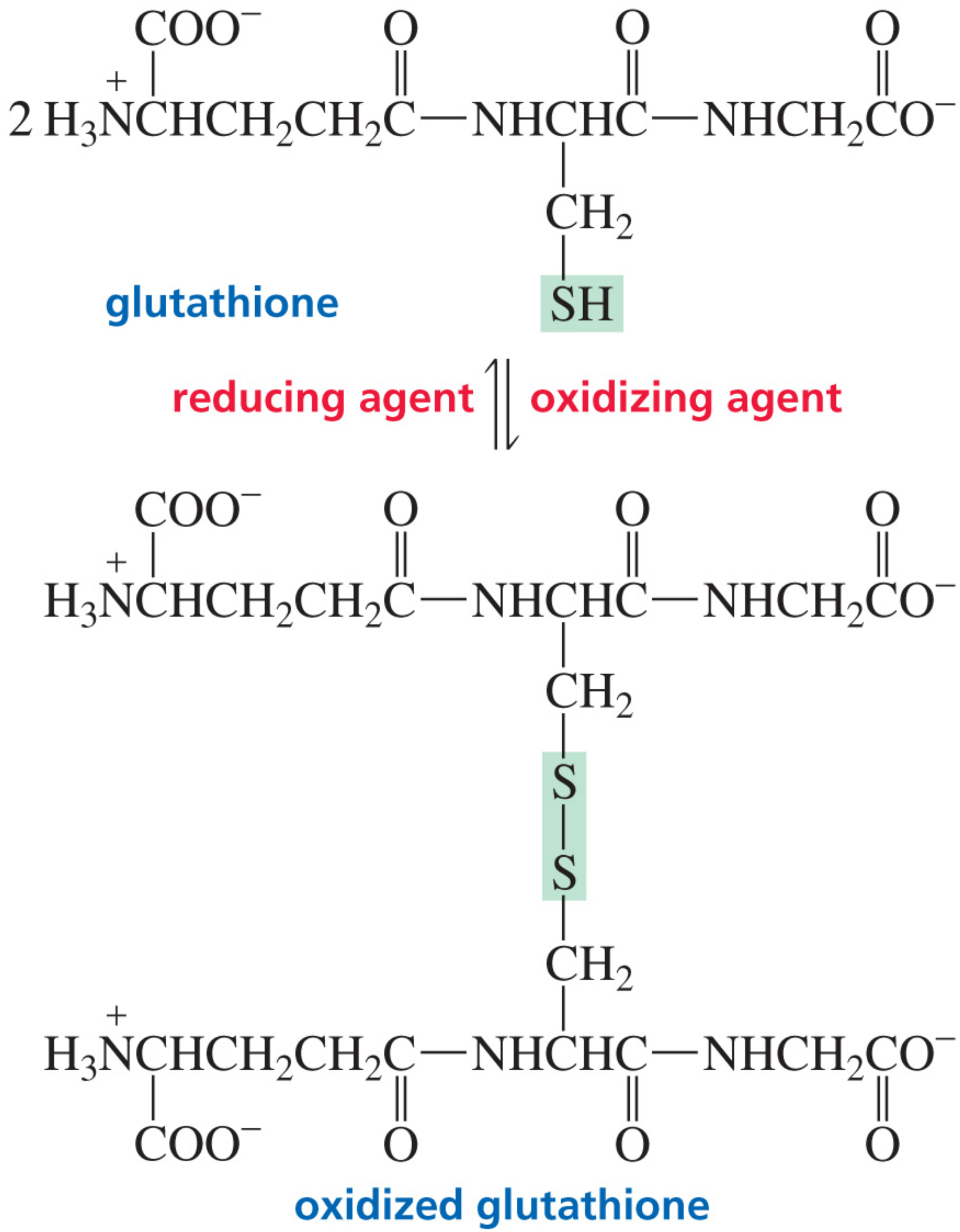


Bradykinin inhibits the inflammation of tissues
Vasopressin controls blood pressure
Oxyocin induces labor in pregnant women by stimulating the uterine muscle



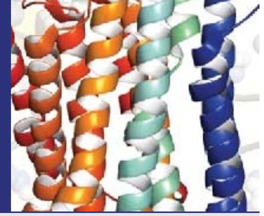
200 times sweeter than sucrose

D-amino acid is bitter.



Destroy harmful oxidizing agents in the body

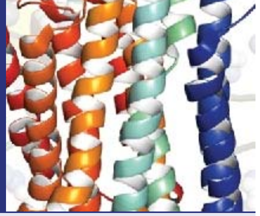
The Sequence of Amino Acids in a Protein



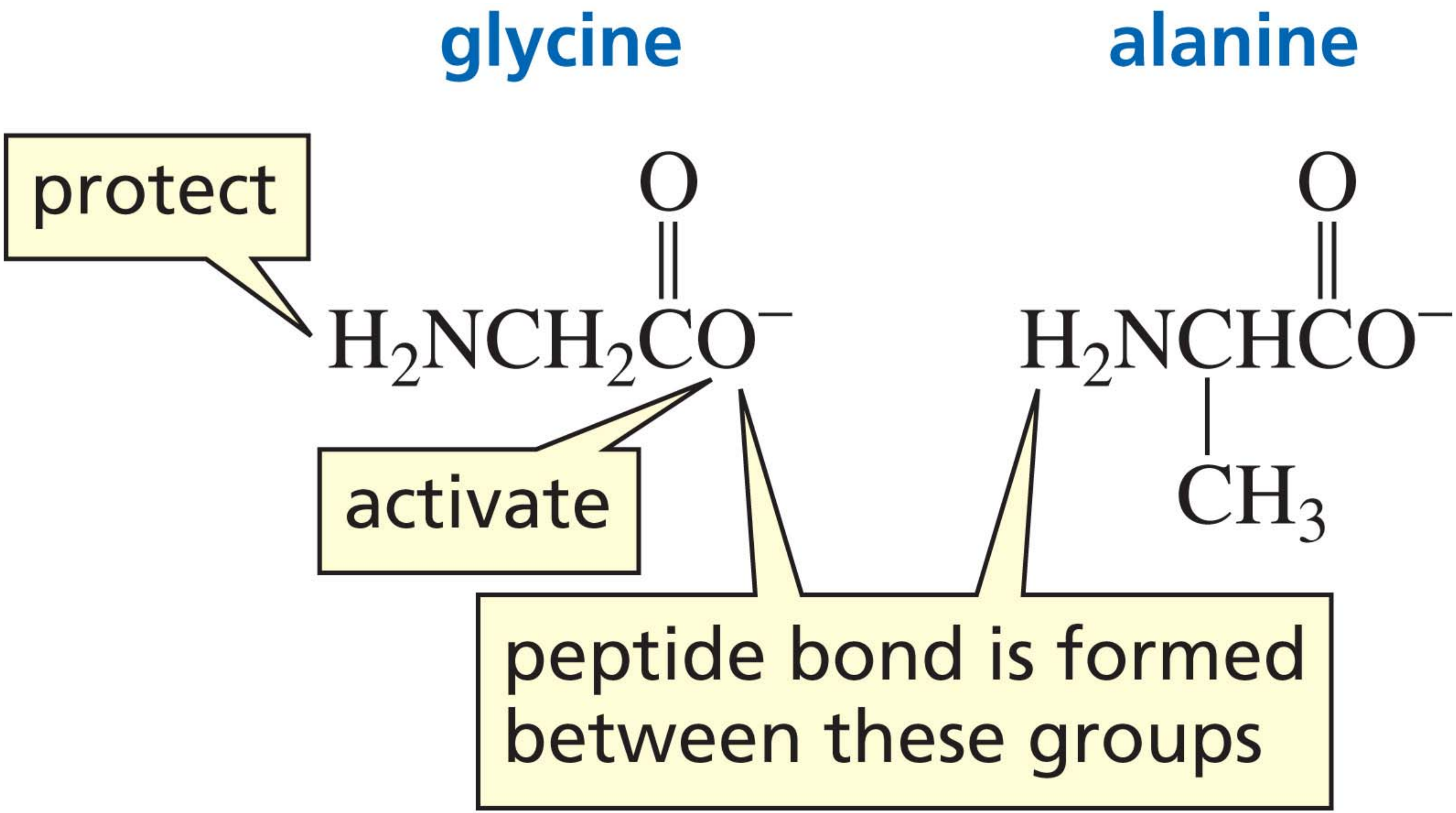
- Is a unique characteristic of every protein
- Is encoded by the nucleotide sequence of DNA
- Is thus a form of genetic information
- Is read from the amino terminus to the carboxyl terminus

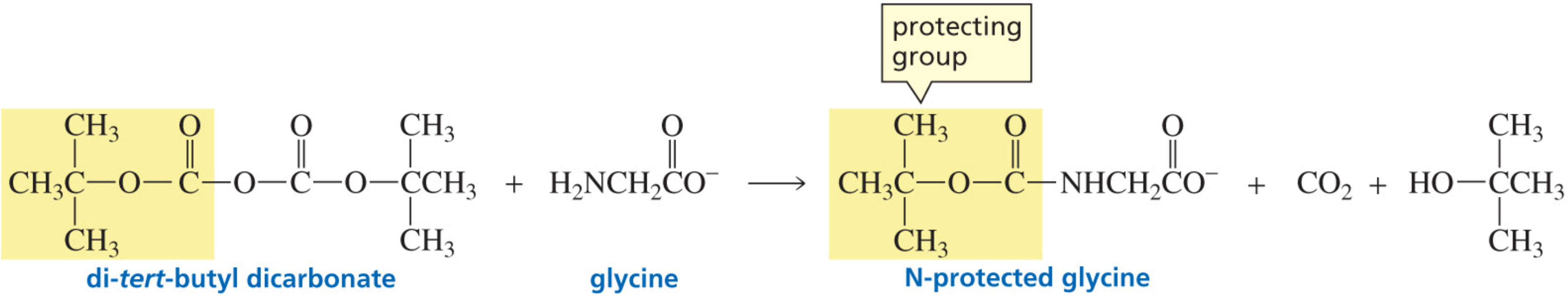
Chemical Synthesis of peptide

Solid Phase Synthesis of Peptides

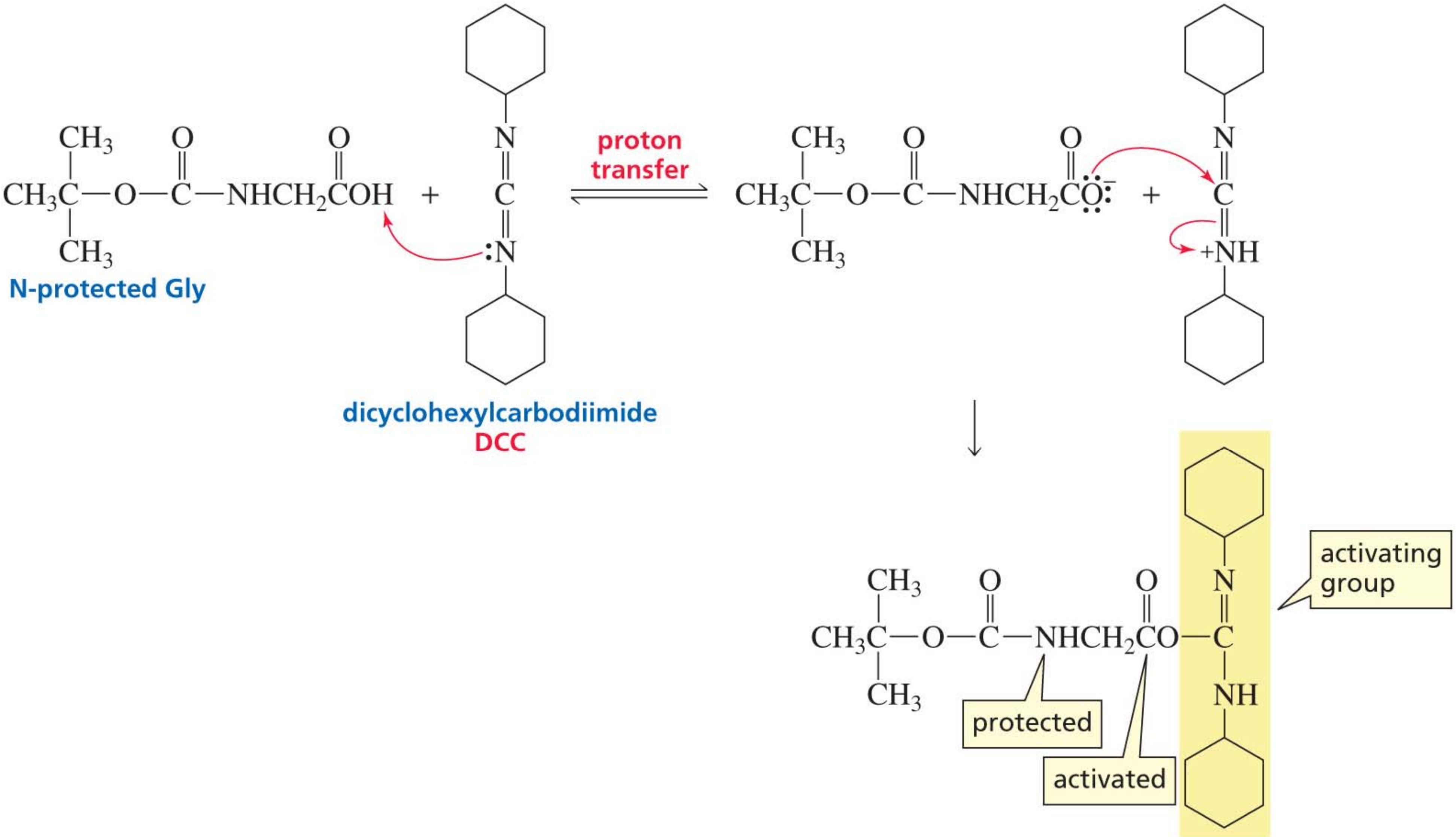


- R. Bruce Merrifield and his collaborators pioneered the solid-phase synthesis of polypeptides in the laboratory
- Carboxy terminus of a nascent peptide is covalently anchored to an insoluble resin
- After each addition of a residue, the resin particles are collected by filtration
- Automation and computer control now permit synthesis of peptides of 30 residues or more

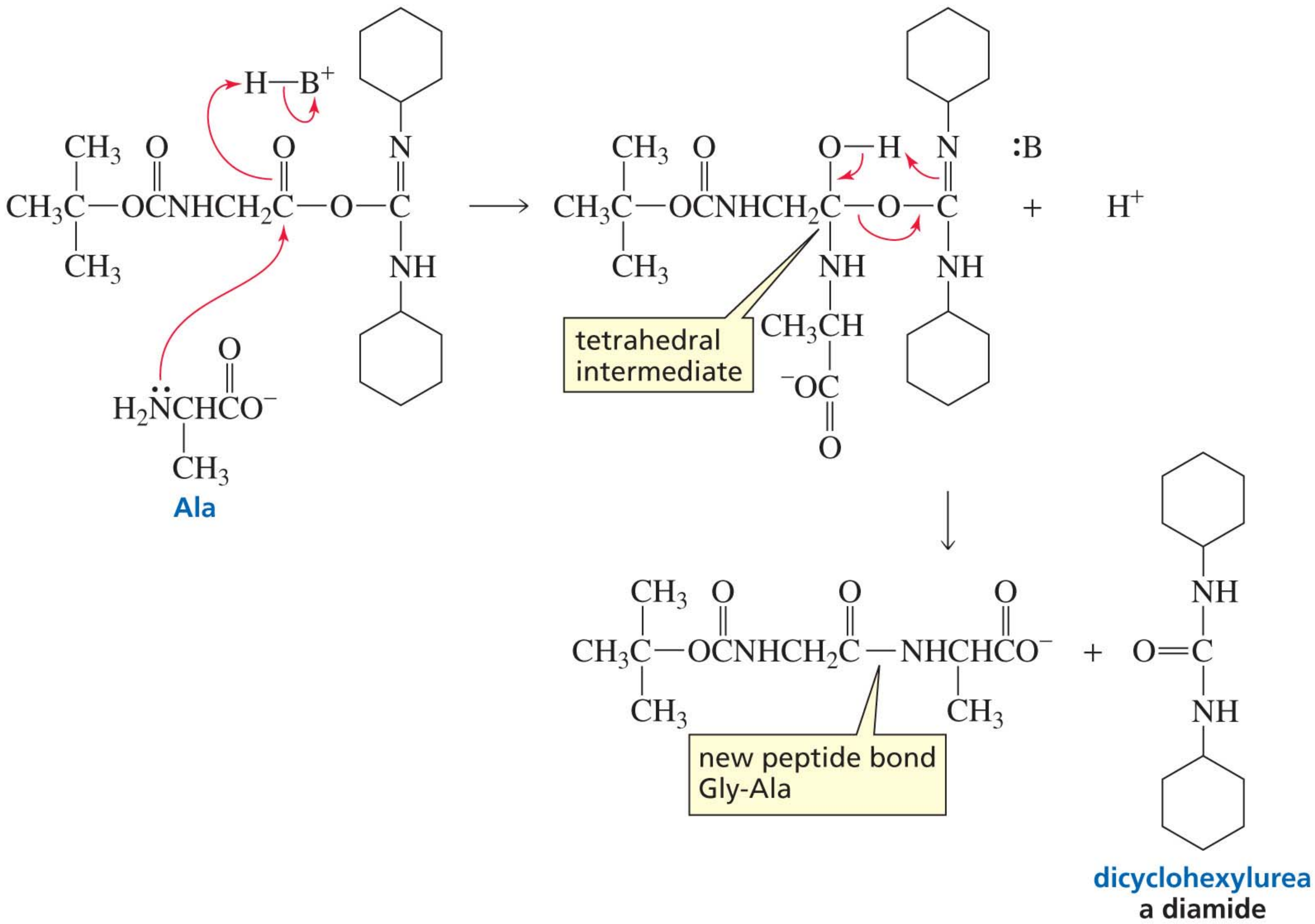

glycine
alanine
protect
activate
peptide bond is formed between these groups

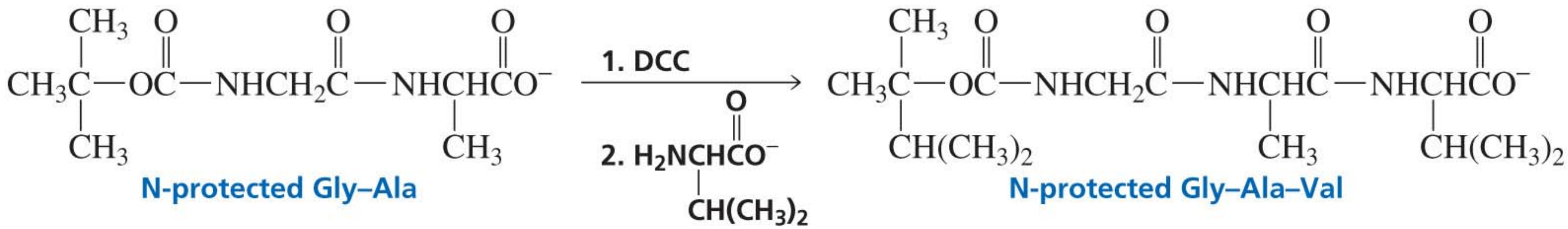

protecting group
di-*tert*-butyl dicarbonate
glycine
N-protected glycine

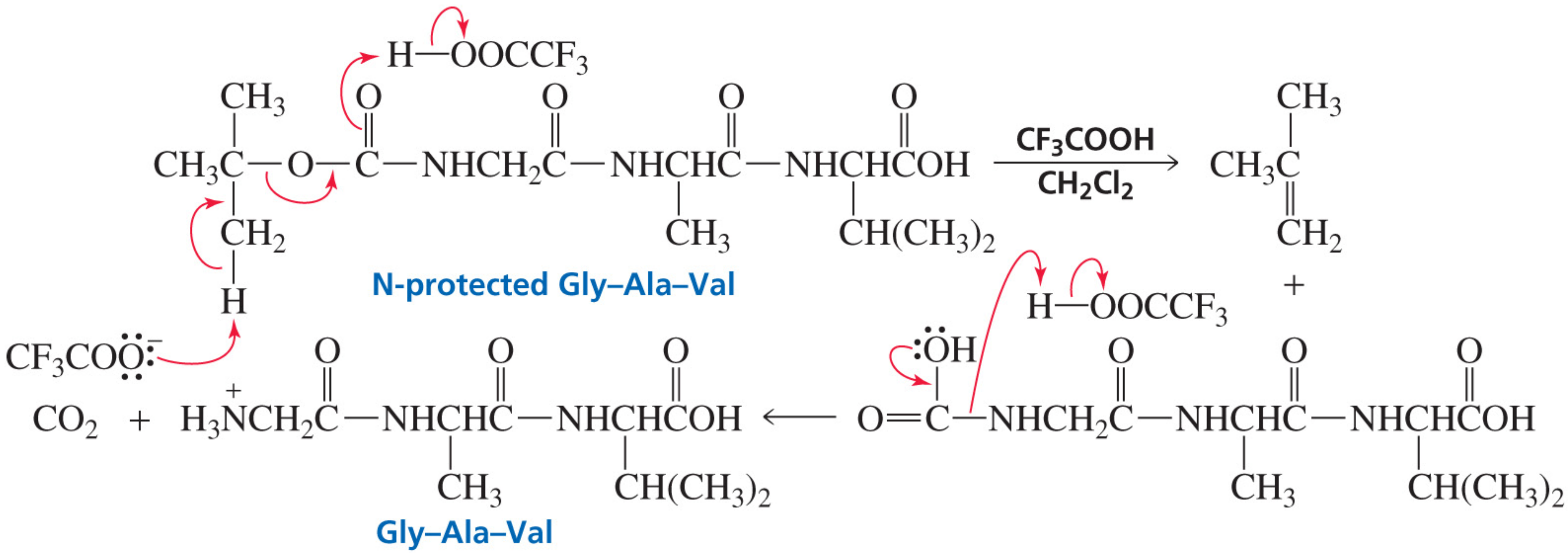
Boc



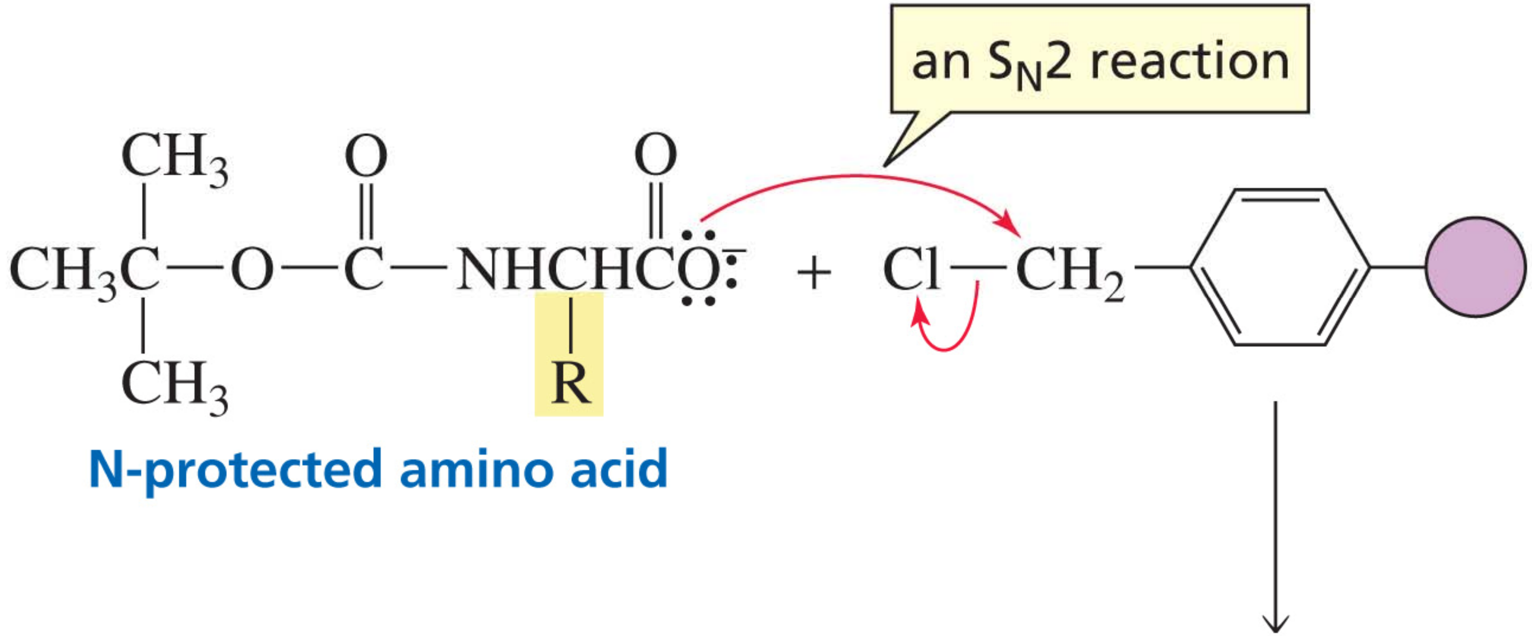
N-protected Gly
dicyclohexylcarbodiimide
DCC
proton
transfer
activating
group
protected
activated

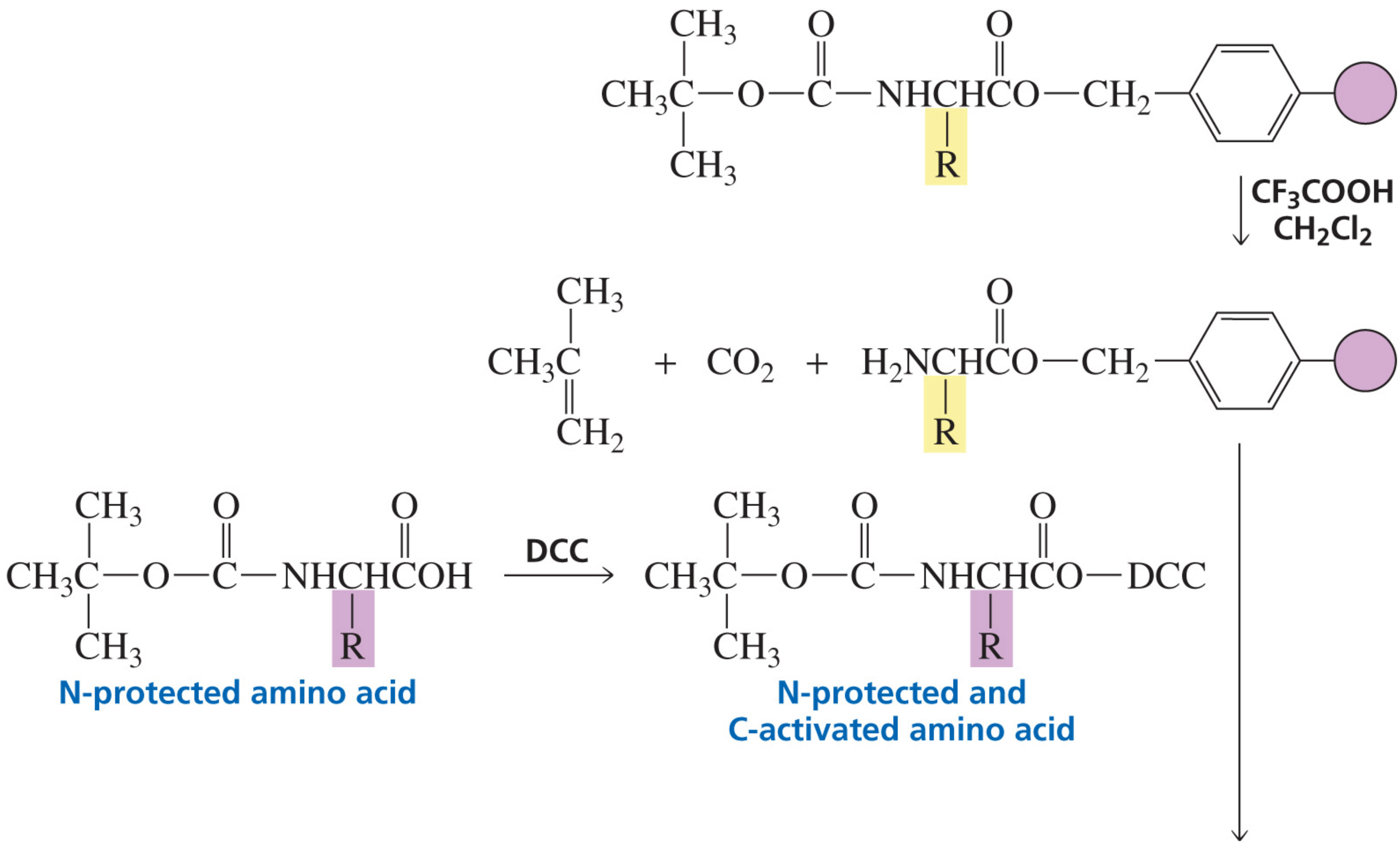

Ala
tetrahedral intermediate
new peptide bond Gly-Ala
dicyclohexylurea
a diamide

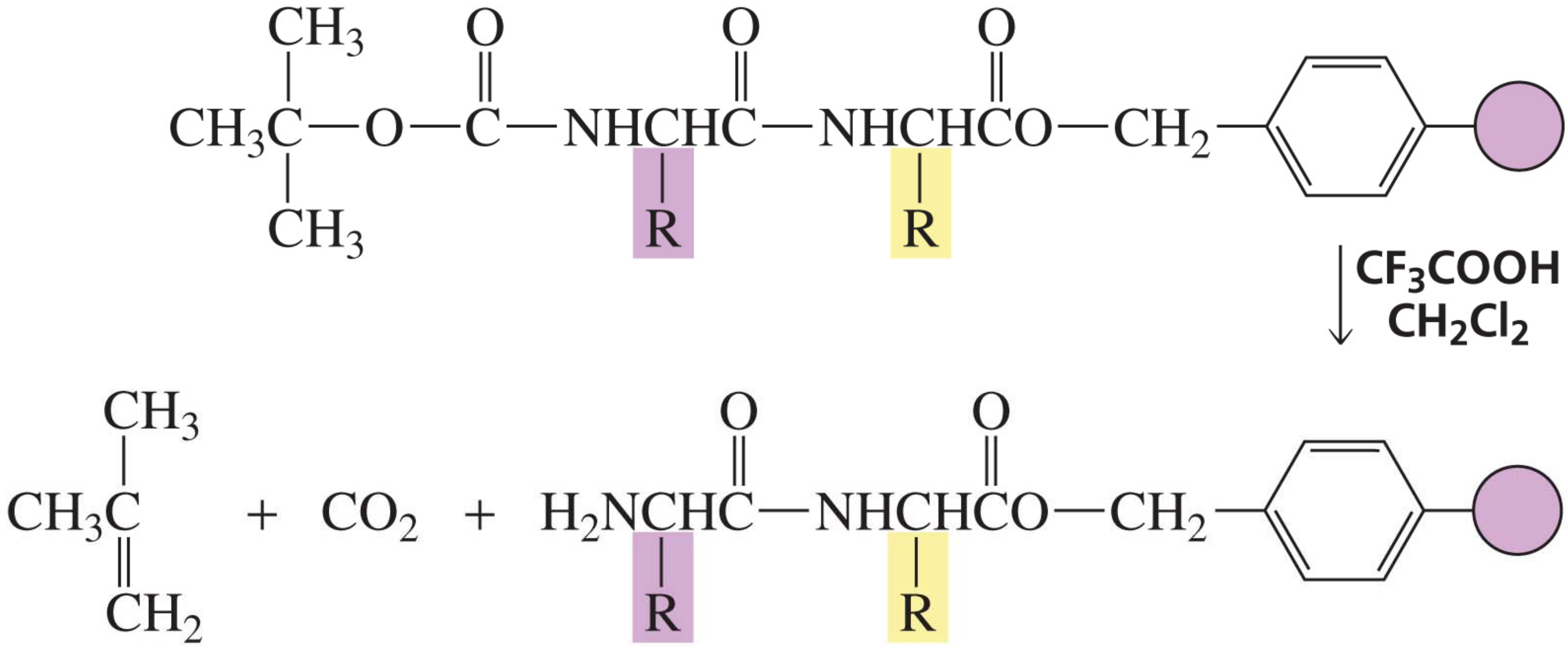

1. DCC
2. H2NCHCO−
CH(CH3)2
N-protected Gly–Ala
N-protected Gly–Ala–Val

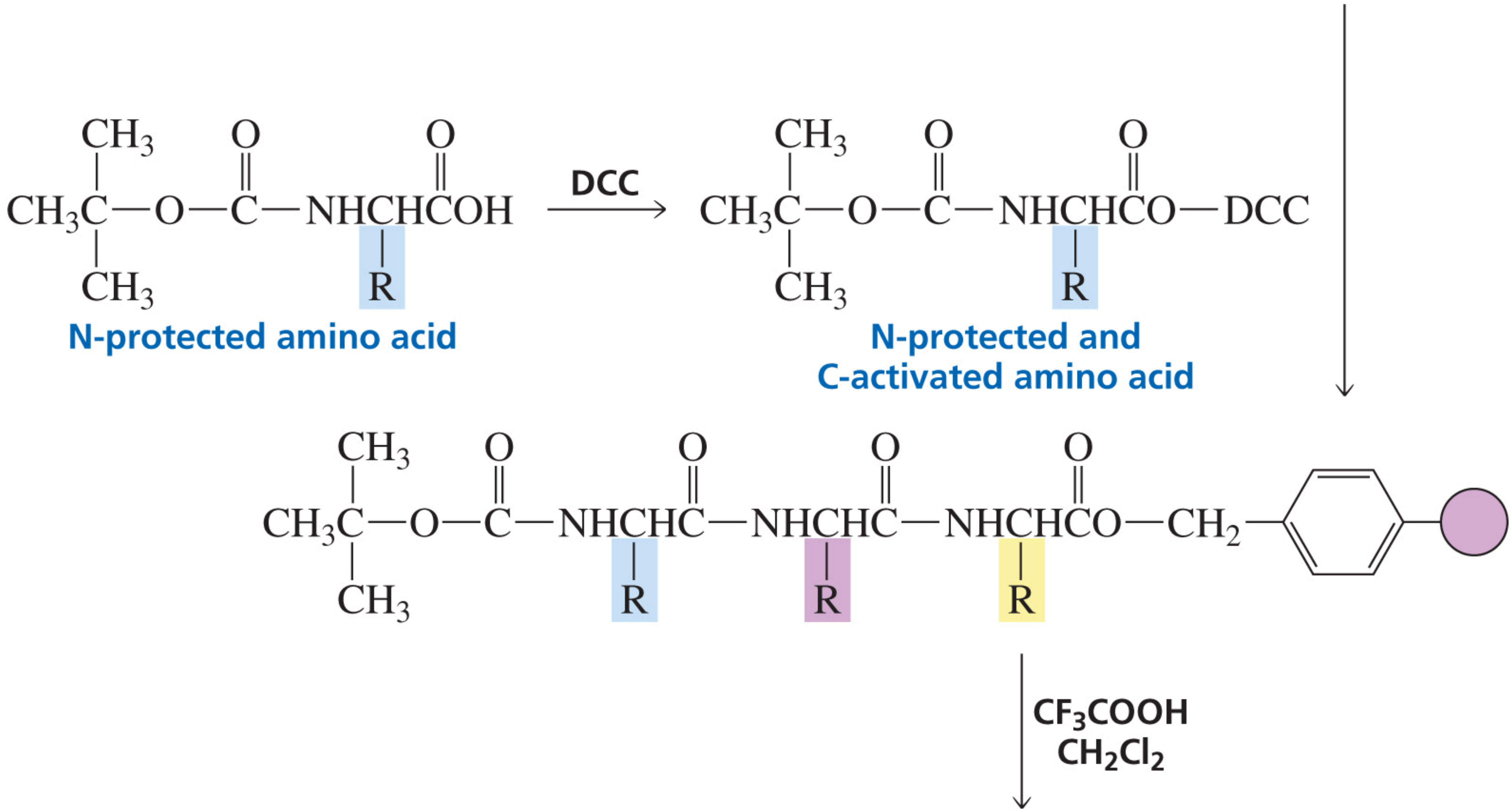

CF3COOH
CH2Cl2
N-protected Gly–Ala–Val
Gly–Ala–Val

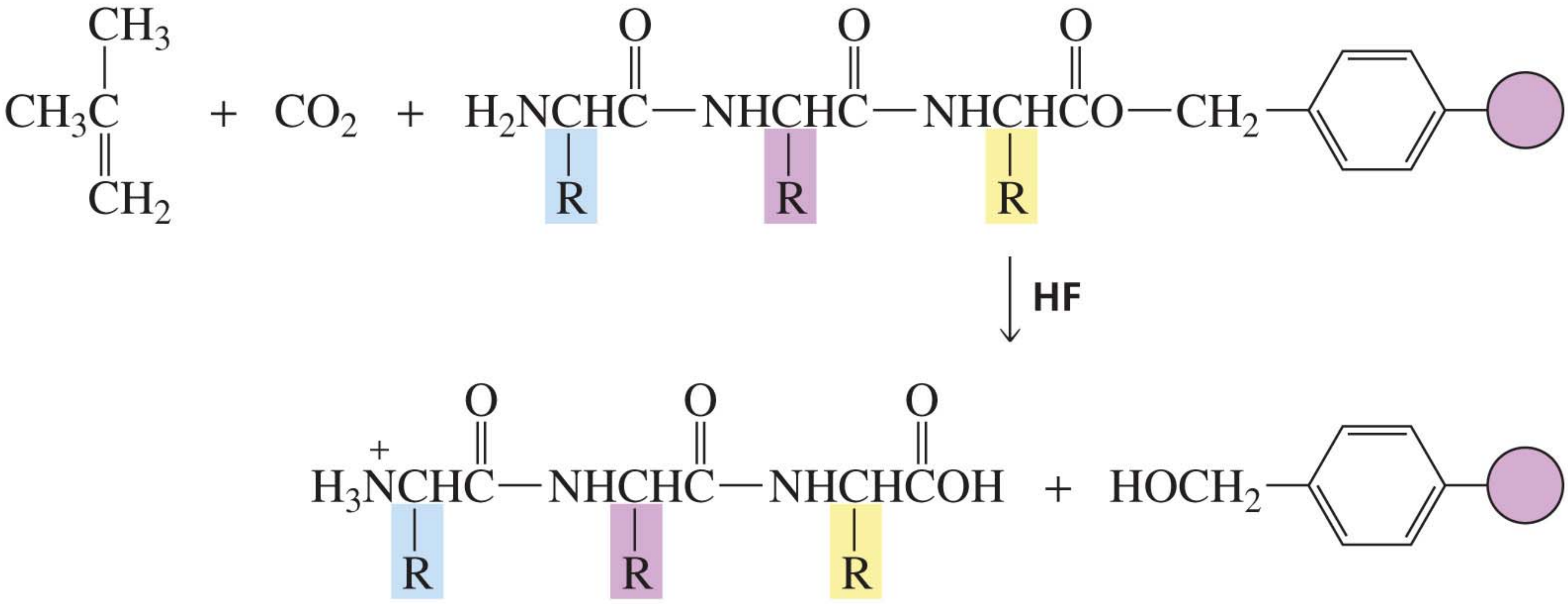
Merrifield automated solid-phase synthesis of a tripeptide




CF3COOH
CH2Cl2
DCC
N-protected amino acid
N-protected and
C-activated amino acid


CF3COOH
CH2Cl2


DCC
N-protected amino acid
N-protected and
C-activated amino acid
CF3COOH
CH2Cl2



Automated solid-phase peptide synthesis


PS3
Protein