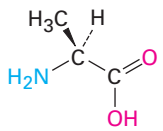
26

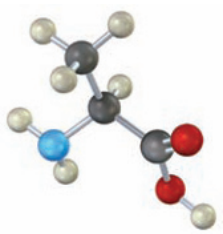
The building blocks of life that we call proteins are aptly named after Proteus, the early Greek sea-god whose name means "first" or "primordial."

Biomolecules: Amino Acids, Peptides, and Proteins

Proteins occur in every living organism, are of many different types, and have many different biological functions. The keratin of skin and fingernails, the fibroin of silk and spider webs, and the estimated 50,000 or so enzymes that catalyze the biological reactions in our bodies are all proteins. Regardless of their function, all proteins have a fundamentally similar structure and are made up of many *amino acids* linked together in a long chain.

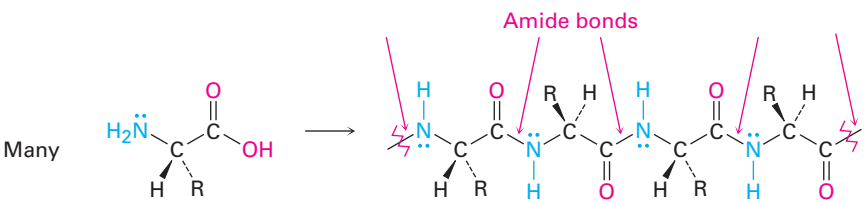
Amino acids, as their name implies, are difunctional. They contain both a basic amino group and an acidic carboxyl group.





Alanine, an amino acid

Their value as building blocks to make proteins stems from the fact that amino acids can join together into long chains by forming amide bonds between the $-NH_2$ of one amino acid and the $-CO_2H$ of another. For classification purposes, chains with fewer than 50 amino acids are often called **peptides**, while the term **protein** is generally used for larger chains.

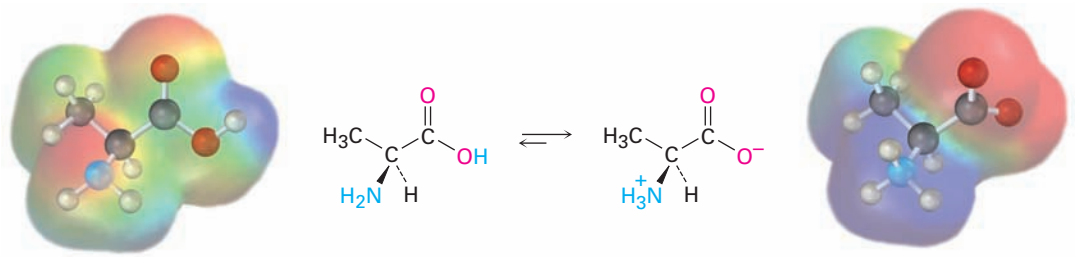


OWL Sign in to OWL for Organic Chemistry at **www.cengage.com/owl** to view tutorials and simulations, develop problem-solving skills, and complete online homework assigned by your professor.

Why This Chapter? Continuing our look at the main classes of biomolecules, we'll focus in this chapter on amino acids, the fundamental building blocks from which the 150,000 or so different proteins in our bodies are made. We'll then see how amino acids are incorporated into proteins and the structures of those proteins. Any understanding of biological chemistry would be impossible without this knowledge.

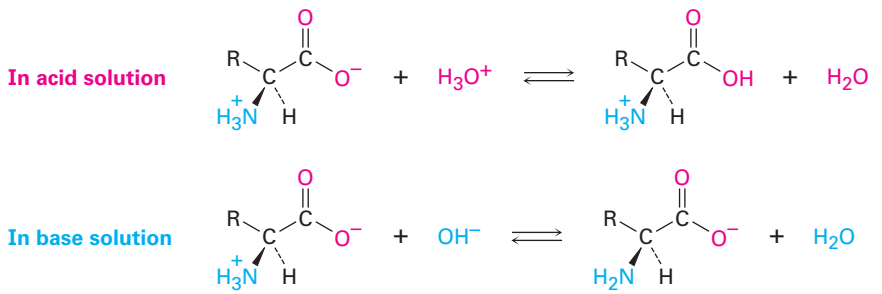
26.1 Structures of Amino Acids

We saw in **Sections 20.3 and 24.5** that a carboxyl group is deprotonated and exists as the carboxylate anion at a physiological pH of 7.3, while an amino group is protonated and exists as the ammonium cation. Thus, amino acids exist in aqueous solution primarily in the form of a dipolar ion, or **zwitterion** (from the German *zwitter,* meaning "hybrid").



Alanine

Amino acid zwitterions are internal salts and therefore have many of the physical properties associated with salts. They have large dipole moments, are relatively soluble in water but insoluble in hydrocarbons, and are crystalline with relatively high melting points. In addition, amino acids are *amphiprotic;* they can react either as acids or as bases, depending on the circumstances. In aqueous acid solution, an amino acid zwitterion is a base that accepts a proton onto its $-CO_2^-$ group to yield a cation. In aqueous base solution, the zwitterion is an acid that loses a proton from its $-NH_3^+$ group to form an anion.



The structures, abbreviations (both three- and one-letter), and pK_a values of the 20 amino acids commonly found in proteins are shown in Table 26.1. All are **α-amino acids**, meaning that the amino group in each is a substituent on the α carbon—the one next to the carbonyl group. Nineteen of the twenty

Table 26.1 The 20 Common Amino Acids in Proteins

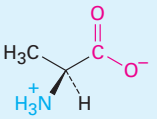
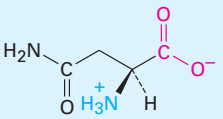
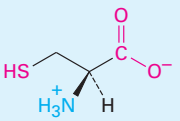
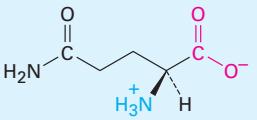
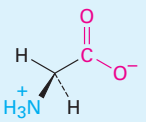
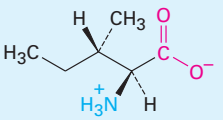
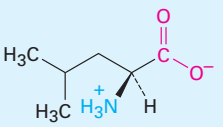
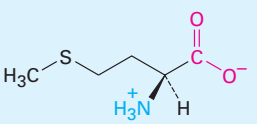
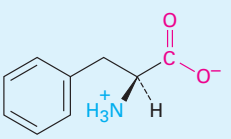
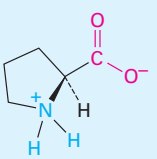
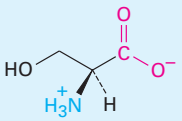
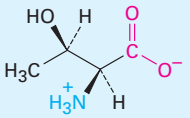
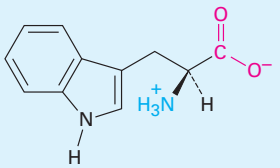
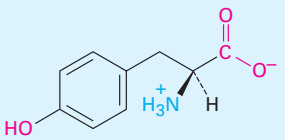
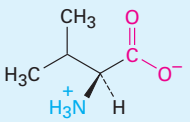
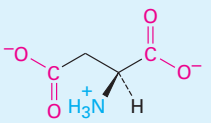
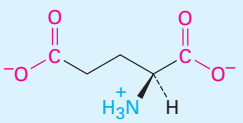
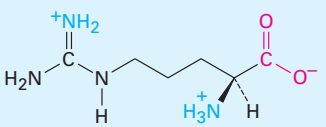
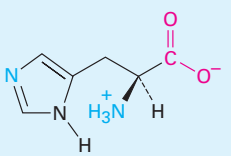
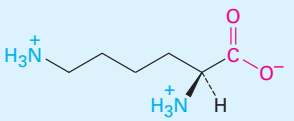
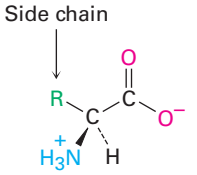
Name	Abbreviations		MW	Stucture	pK_a α-CO_2H	pK_a α-NH_3^+	pK_a side chain	p*I*
Neutral Amino Acids								
Alanine	Ala	A	89		2.34	9.69	—	6.01
Asparagine	Asn	N	132		2.02	8.80	—	5.41
Cysteine	Cys	C	121		1.96	10.28	8.18	5.07
Glutamine	Gln	Q	146		2.17	9.13	—	5.65
Glycine	Gly	G	75		2.34	9.60	—	5.97
Isoleucine	Ile	I	131		2.36	9.60	—	6.02
Leucine	Leu	L	131		2.36	9.60	—	5.98
Methionine	Met	M	149		2.28	9.21	—	5.74
Phenylalanine	Phe	F	165		1.83	9.13	—	5.48
Proline	Pro	P	115		1.99	10.60	—	6.30

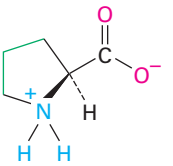
Table 26.1 The 20 Common Amino Acids in Proteins *(continued)*

Name	Abbreviations		MW	Stucture	pK_a α-CO_2H	pK_a α-NH_3^+	pK_a side chain	p*I*
Neutral Amino Acids *(continued)*								
Serine	Ser	S	105		2.21	9.15	—	5.68
Threonine	Thr	T	119		2.09	9.10	—	5.60
Tryptophan	Trp	W	204		2.83	9.39	—	5.89
Tyrosine	Tyr	Y	181		2.20	9.11	10.07	5.66
Valine	Val	V	117		2.32	9.62	—	5.96
Acidic Amino Acids								
Aspartic acid	Asp	D	133		1.88	9.60	3.65	2.77
Glutamic acid	Glu	E	147		2.19	9.67	4.25	3.22
Basic Amino Acids								
Arginine	Arg	R	174		2.17	9.04	12.48	10.76
Histidine	His	H	155		1.82	9.17	6.00	7.59
Lysine	Lys	K	146		2.18	8.95	10.53	9.74

amino acids are primary amines, RNH_2, and differ only in the nature of the **side chain**—the substituent attached to the α carbon. Proline is a secondary amine whose nitrogen and α carbon atoms are part of a five-membered pyrrolidine ring.

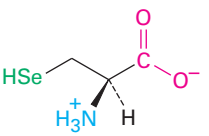


A primary α-amino acid

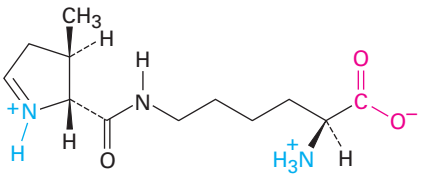


Proline, a secondary α-amino acid

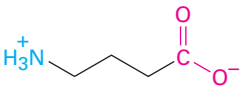
In addition to the 20 amino acids commonly found in proteins, 2 others—selenocysteine and pyrrolysine—are found in some organisms, and more than 700 nonprotein amino acids are also found in nature. γ-Aminobutyric acid (GABA), for instance, is found in the brain and acts as a neurotransmitter; homocysteine is found in blood and is linked to coronary heart disease; and thyroxine is found in the thyroid gland, where it acts as a hormone.



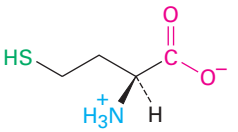
Selenocysteine



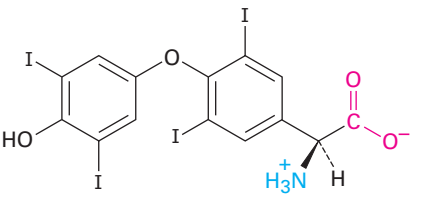
Pyrrolysine



γ-Aminobutyric acid



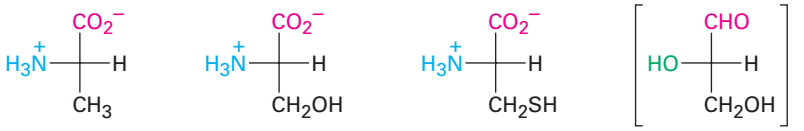
Homocysteine



Thyroxine

Except for glycine, $H_2NCH_2CO_2H$, the α carbons of amino acids are chirality centers. Two enantiomers of each are therefore possible, but nature uses only one to build proteins. In Fischer projections, naturally occurring amino acids are represented by placing the $-CO_2^-$ group at the top and the side chain down, as if drawing a carbohydrate **(Section 25.2)** and then placing the $-NH_3^+$ group on the left. Because of their stereochemical similarity to L sugars

(Section 25.3), the naturally occurring α-amino acids are often referred to as L amino acids.



L-Alanine
(*S*)-Alanine

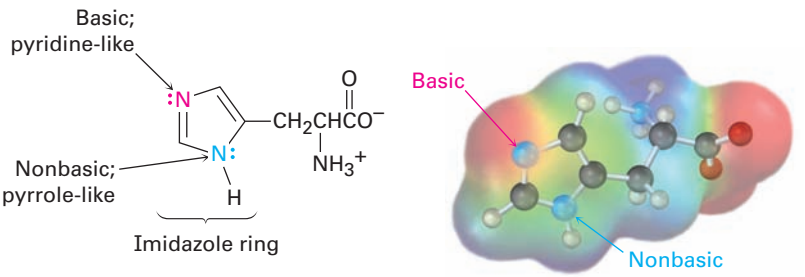
L-Serine
(*S*)-Serine

L-Cysteine
(*R*)-Cysteine

L-Glyceraldehyde

The 20 common amino acids can be further classified as neutral, acidic, or basic, depending on the structure of their side chains. Fifteen of the twenty have neutral side chains, two (aspartic acid and glutamic acid) have an extra carboxylic acid function in their side chains, and three (lysine, arginine, and histidine) have basic amino groups in their side chains. Note that both cysteine (a thiol) and tyrosine (a phenol), although usually classified as neutral amino acids, nevertheless have weakly acidic side chains that can be deprotonated in a sufficiently basic solution.

At the physiological pH of 7.3 within cells, the side-chain carboxyl groups of aspartic acid and glutamic acid are deprotonated and the basic side-chain nitrogens of lysine and arginine are protonated. Histidine, however, which contains a heterocyclic imidazole ring in its side chain, is not quite basic enough to be protonated at pH 7.3. Note that only the pyridine-like, doubly bonded nitrogen in histidine is basic. The pyrrole-like singly bonded nitrogen is nonbasic because its lone pair of electrons is part of the six-π-electron aromatic imidazole ring **(Section 24.9)**.



Histidine

Humans are able to synthesize only 11 of the 20 protein amino acids, called nonessential amino acids. The other 9, called essential amino acids, are biosynthesized only in plants and microorganisms and must be obtained in our diet. The division between essential and nonessential amino acids is not clear-cut, however. Tyrosine, for instance, is sometimes considered nonessential because humans can produce it from phenylalanine, but phenylalanine itself is essential and must be obtained in the diet. Arginine can be synthesized by humans, but much of the arginine we need also comes from our diet.

Problem 26.1

How many of the α-amino acids shown in Table 26.1 contain aromatic rings? How many contain sulfur? How many contain alcohols? How many contain hydrocarbon side chains?

Problem 26.2

Eighteen of the nineteen L amino acids have the *S* configuration at the α carbon. Cysteine is the only L amino acid that has an *R* configuration. Explain.

Problem 26.3

The amino acid threonine, (2*S*,3*R*)-2-amino-3-hydroxybutanoic acid, has two chirality centers.

(a) Draw a Fischer projection of threonine.

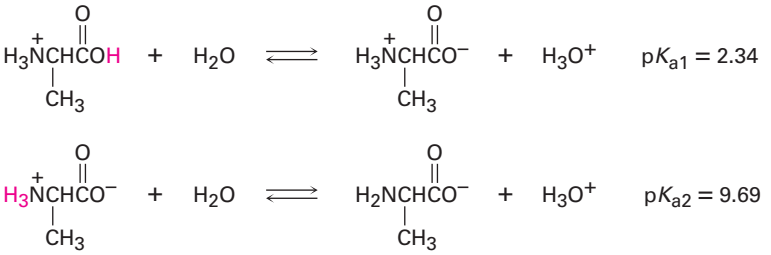
(b) Draw a Fischer projection of a threonine diastereomer, and label its chirality centers as *R* or *S*.

26.2 Amino Acids and the Henderson–Hasselbalch Equation: Isoelectric Points

According to the Henderson–Hasselbalch equation **(Sections 20.3 and 24.5)**, if we know both the pH of a solution and the pK_a of an acid HA, we can calculate the ratio of $[A^-]$ to [HA] in the solution. Furthermore, when pH = pK_a, the two forms A^- and HA are present in equal amounts because log 1 = 0.

$$\text{pH} = \text{p}K_a + \log\frac{[A^-]}{[HA]} \qquad \text{or} \qquad \log\frac{[A^-]}{[HA]} = \text{pH} - \text{p}K_a$$

To apply the Henderson–Hasselbalch equation to an amino acid, let's find out what species are present in a 1.00 M solution of alanine at pH = 9.00. According to Table 26.1, protonated alanine [$^+H_3NCH(CH_3)CO_2H$] has pK_{a1} = 2.34 and neutral zwitterionic alanine [$^+H_3NCH(CH_3)CO_2^-$] has pK_{a2} = 9.69:



Since the pH of the solution is much closer to pK_{a2} than to pK_{a1}, we need to use pK_{a2} for the calculation. From the Henderson–Hasselbalch equation, we have:

$$\log\frac{[A^-]}{[HA]} = \text{pH} - \text{p}K_a = 9.00 - 9.69 = -0.69$$

so

$$\frac{[A^-]}{[HA]} = \text{antilog}(-0.69) = 0.20 \qquad \text{and} \qquad [A^-] = 0.20[HA]$$

In addition, we know that

$$[A^-] + [HA] = 1.00\ M$$

Solving the two simultaneous equations gives [HA] = 0.83 and [A^-] = 0.17. In other words, at pH = 9.00, 83% of alanine molecules in a 1.00 M solution are neutral (zwitterionic) and 17% are deprotonated. Similar calculations can be done at any other pH and the results plotted to give the titration curve shown in **Figure 26.1**.

Each leg of the titration curve is calculated separately. The first leg, from pH 1 to 6, corresponds to the dissociation of protonated alanine, H_2A^+. The second leg, from pH 6 to 11, corresponds to the dissociation of zwitterionic alanine, HA. It's as if we started with H_2A^+ at low pH and then titrated with NaOH. When 0.5 equivalent of NaOH is added, the deprotonation of H_2A^+ is 50% done; when 1.0 equivalent of NaOH is added, the deprotonation of H_2A^+ is complete and HA predominates; when 1.5 equivalents of NaOH is added, the deprotonation of HA is 50% done; and when 2.0 equivalents of NaOH is added, the deprotonation of HA is complete.

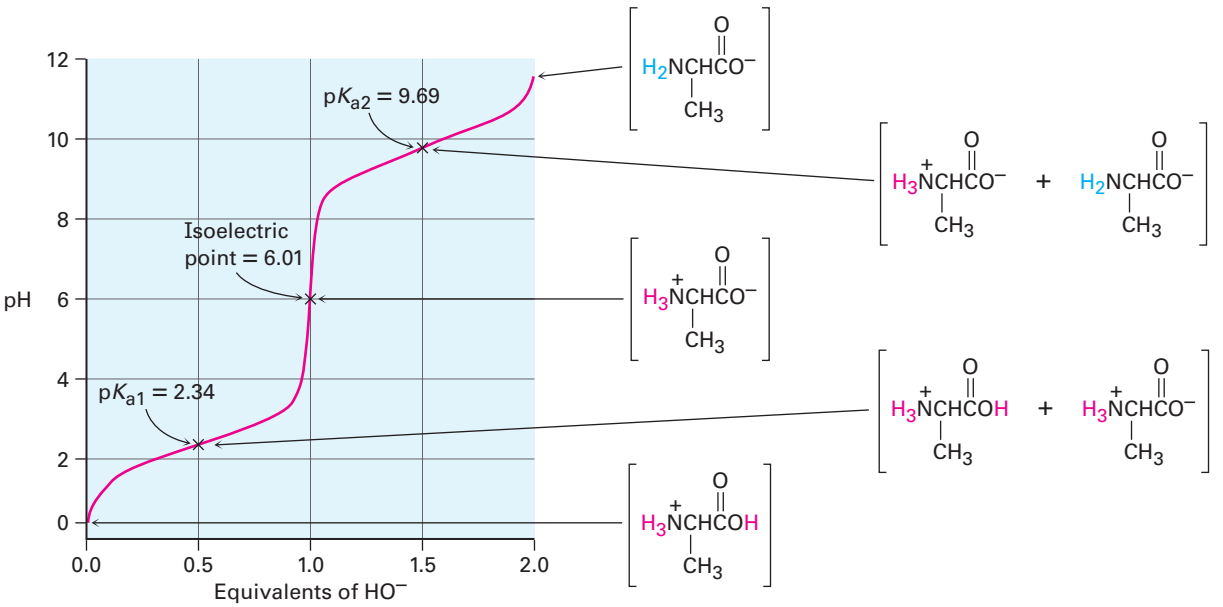
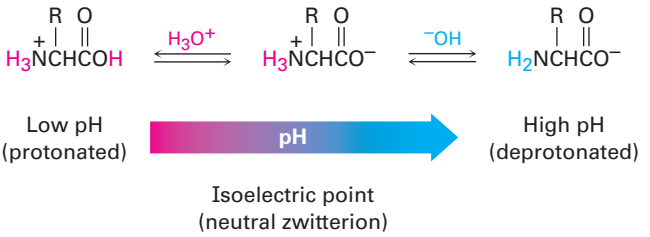


Figure 26.1 A titration curve for alanine, plotted using the Henderson–Hasselbalch equation. Each of the two legs is plotted separately. At pH < 1, alanine is entirely protonated; at pH = 2.34, alanine is a 50:50 mix of protonated and neutral forms; at pH = 6.01, alanine is entirely neutral; at pH = 9.69, alanine is a 50:50 mix of neutral and deprotonated forms; at pH > 11.5, alanine is entirely deprotonated.

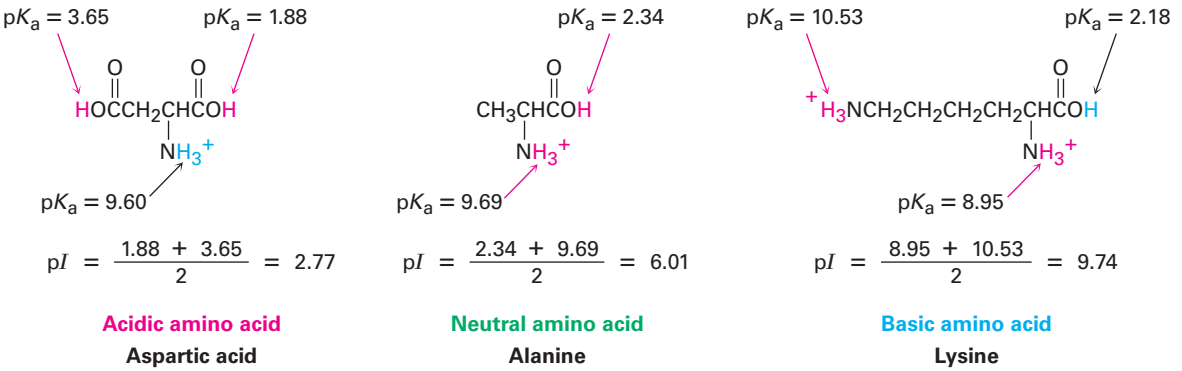
Look carefully at the titration curve in Figure 26.1. In acid solution, the amino acid is protonated and exists primarily as a cation. In basic solution, the amino acid is deprotonated and exists primarily as an anion. In between the two is an intermediate pH at which the amino acid is exactly balanced between anionic and cationic forms and exists primarily as the neutral, dipolar zwitterion. This

pH is called the amino acid's **isoelectric point (p*I*)** and has a value of 6.01 for alanine.



The isoelectric point of an amino acid depends on its structure, with values for the 20 common amino acids given in Table 26.1. The 15 neutral amino acids have isoelectric points near neutrality, in the pH range 5.0 to 6.5. The two acidic amino acids have isoelectric points at lower pH so that deprotonation of the side-chain $-CO_2H$ does not occur at their p*I*, and the three basic amino acids have isoelectric points at higher pH so that protonation of the side-chain amino group does not occur at their p*I*.

More specifically, the p*I* of any amino acid is the average of the two acid-dissociation constants that involve the neutral zwitterion. For the 13 amino acids with a neutral side chain, p*I* is the average of pK_{a1} and pK_{a2}. For the four amino acids with either a strongly or weakly acidic side chain, p*I* is the average of the two lowest pK_a values. For the three amino acids with a basic side chain, p*I* is the average of the two highest pK_a values.



Just as individual amino acids have isoelectric points, proteins have an overall p*I* because of the cumulative effect of all the acidic or basic amino acids they may contain. The enzyme lysozyme, for instance, has a preponderance of basic amino acids and thus has a high isoelectric point (p*I* = 11.0). Pepsin, however, has a preponderance of acidic amino acids and a low isoelectric point (p*I* ~ 1.0). Not surprisingly, the solubilities and properties of proteins with different p*I*'s are strongly affected by the pH of the medium. Solubility in water is usually lowest at the isoelectric point, where the protein has no net charge, and is higher both above and below the p*I*, where the protein is charged.

We can take advantage of the differences in isoelectric points to separate a mixture of proteins into its pure constituents. Using a technique known as

electrophoresis, a mixture of proteins is placed near the center of a strip of paper or gel. The paper or gel is moistened with an aqueous buffer of a given pH, and electrodes are connected to the ends of the strip. When an electric potential is applied, those proteins with negative charges (those that are deprotonated because the pH of the buffer is above their isoelectric point) migrate slowly toward the positive electrode. At the same time, those amino acids with positive charges (those that are protonated because the pH of the buffer is below their isoelectric point) migrate toward the negative electrode.

Different proteins migrate at different rates, depending on their isoelectric points and on the pH of the aqueous buffer, thereby effecting a separation of the mixture into its components. **Figure 26.2** illustrates this separation for a mixture containing basic, neutral, and acidic components.

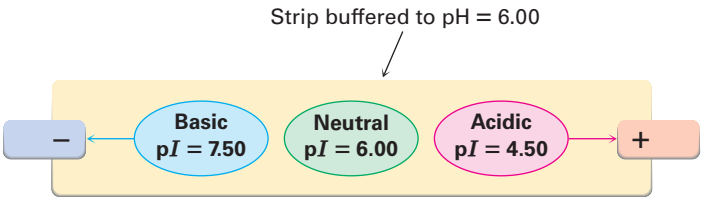


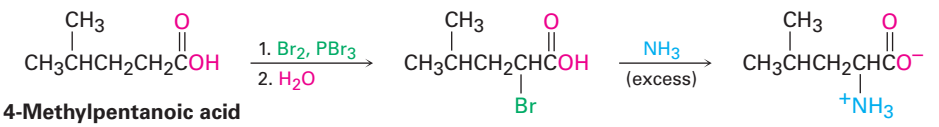
Figure 26.2 Separation of a protein mixture by electrophoresis. At pH = 6.00, a neutral protein does not migrate, a basic protein is protonated and migrates toward the negative electrode, and an acidic protein is deprotonated and migrates toward the positive electrode.

Problem 26.4

Hemoglobin has pI = 6.8. Does hemoglobin have a net negative charge or net positive charge at pH = 5.3? At pH = 7.3?

26.3 Synthesis of Amino Acids

α-Amino acids can be synthesized in the laboratory using some of the reactions discussed in previous chapters. One of the oldest methods of α-amino acid synthesis begins with α bromination of a carboxylic acid by treatment with Br_2 and PBr_3 (the Hell–Volhard–Zelinskii reaction; **Section 22.4**). S_N2 substitution of the α-bromo acid with ammonia then yields an α-amino acid.



4-Methylpentanoic acid

2-Bromo-4-methylpentanoic acid

(*R,S*)-Leucine (45%)

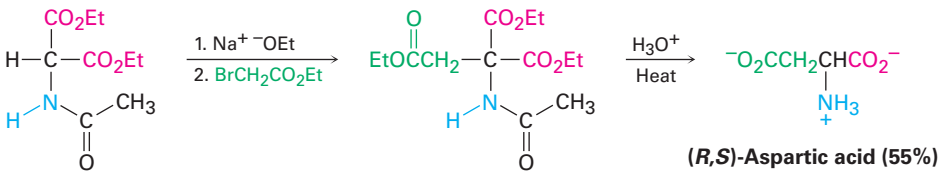
Problem 26.5

Show how you could prepare the following α-amino acids from the appropriate carboxylic acids:

(a) Phenylalanine **(b)** Valine

The Amidomalonate Synthesis

A more general method for preparation of α-amino acids is the *amidomalonate synthesis,* a straightforward extension of the malonic ester synthesis **(Section 22.7)**. The reaction begins with conversion of diethyl acetamidomalonate into an enolate ion by treatment with base, followed by S_N2 alkylation with a primary alkyl halide. Hydrolysis of both the amide protecting group and the esters occurs when the alkylated product is warmed with aqueous acid, and decarboxylation then takes place to yield an α-amino acid. For example, aspartic acid can be prepared from ethyl bromoacetate, $BrCH_2CO_2Et$:



Diethyl acetamidomalonate

(*R,S*)-Aspartic acid (55%)

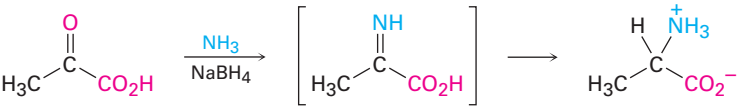
Problem 26.6

What alkyl halides would you use to prepare the following α-amino acids by the amidomalonate method?

(a) Leucine **(b)** Histidine **(c)** Tryptophan **(d)** Methionine

Reductive Amination of α-Keto Acids

Yet another method for the synthesis of α-amino acids is by reductive amination of an α-keto acid with ammonia and a reducing agent. Alanine, for instance, is prepared by treatment of pyruvic acid with ammonia in the presence of $NaBH_4$. As described in **Section 24.6**, the reaction proceeds through formation of an intermediate imine that is then reduced.



Pyruvic acid **Imine intermediate** **(*R,S*)-Alanine**

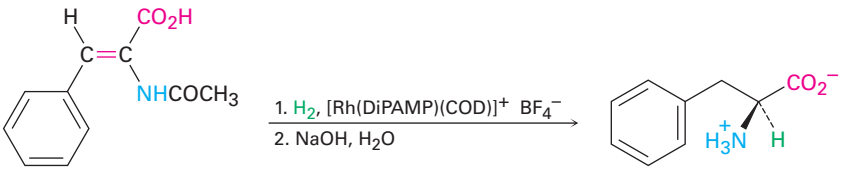
Enantioselective Synthesis

The synthesis of an α-amino acid from an achiral precursor by any of the methods just described yields a racemic mixture, with equal amounts of *S* and *R* enantiomers. To use an amino acid in the laboratory synthesis of a naturally occurring protein, however, the pure *S* enantiomer must be obtained.

Two methods are used in practice to obtain enantiomerically pure amino acids. One way is to resolve the racemic mixture into its pure enantiomers

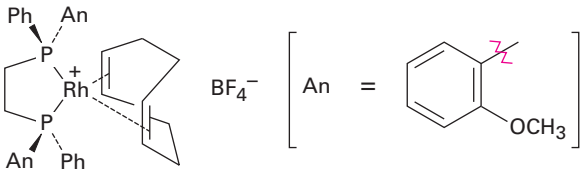
(Section 5.8). A more direct approach, however, is to use an enantioselective synthesis to prepare only the desired *S* enantiomer directly. As discussed in the Chapter 19 *A Deeper Look,* the idea behind enantioselective synthesis is to find a chiral reaction catalyst that will temporarily hold a substrate molecule in an unsymmetrical, chiral environment. While in that chiral environment, the substrate may be more open to reaction on one side than on another, leading to an excess of one enantiomeric product over another.

William Knowles at the Monsanto Company discovered some years ago that α-amino acids can be prepared enantioselectively by hydrogenation of a *Z* enamido acid with a chiral hydrogenation catalyst. (*S*)-Phenylalanine, for instance, is prepared in 98.7% purity contaminated by only 1.3% of the (*R*) enantiomer when a chiral rhodium catalyst is used. For this discovery, Knowles shared the 2001 Nobel Prize in Chemistry.



A (*Z*) enamido acid **(*S*)-Phenylalanine**

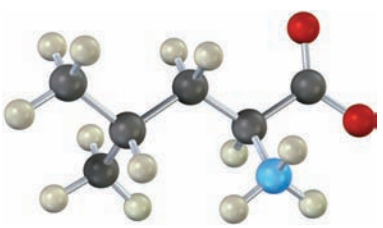
The most effective catalysts for enantioselective amino acid synthesis are coordination complexes of rhodium(I) with 1,5-cyclooctadiene (COD) and a chiral diphosphine such as (*R*,*R*)-1,2-bis(*o*-anisylphenylphosphino)ethane, the so-called DiPAMP ligand. The complex owes its chirality to the presence of the trisubstituted phosphorus atoms **(Section 5.10)**.



[Rh(*R*,*R*-DiPAMP)(COD)]+ BF_4^-

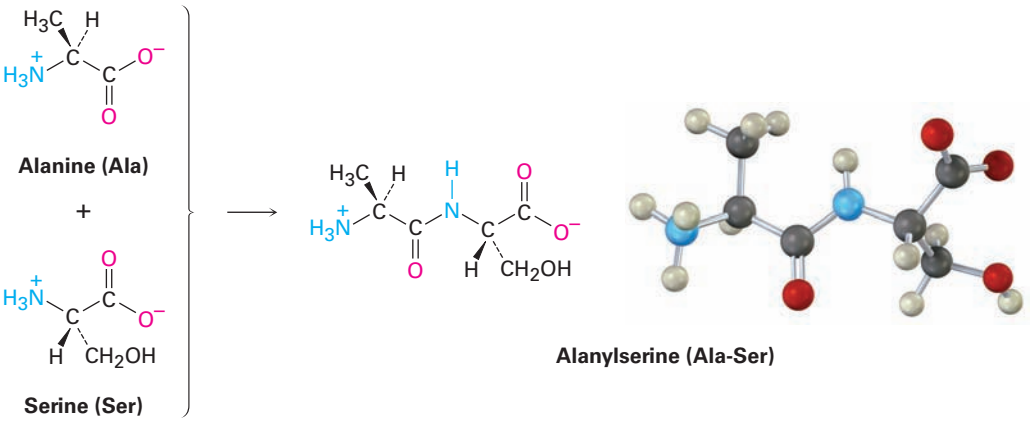
Problem 26.7

Show how you could prepare the following amino acid enantioselectively:

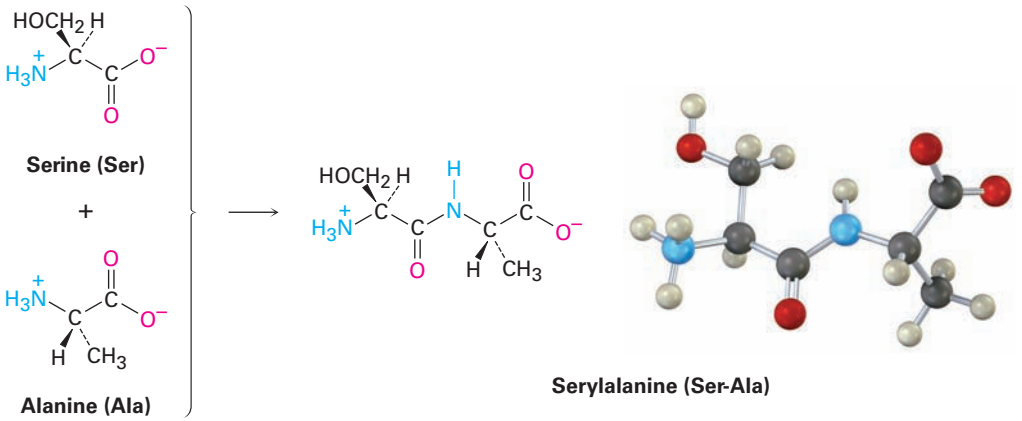


26.4 Peptides and Proteins

Proteins and peptides are amino acid polymers in which the individual amino acids, called **residues**, are linked together by amide bonds, or *peptide bonds*. An amino group from one residue forms an amide bond with the carboxyl of a second residue, the amino group of the second forms an amide bond with the carboxyl of a third, and so on. For example, alanylserine is the dipeptide that results when an amide bond forms between the alanine carboxyl and the serine amino group.

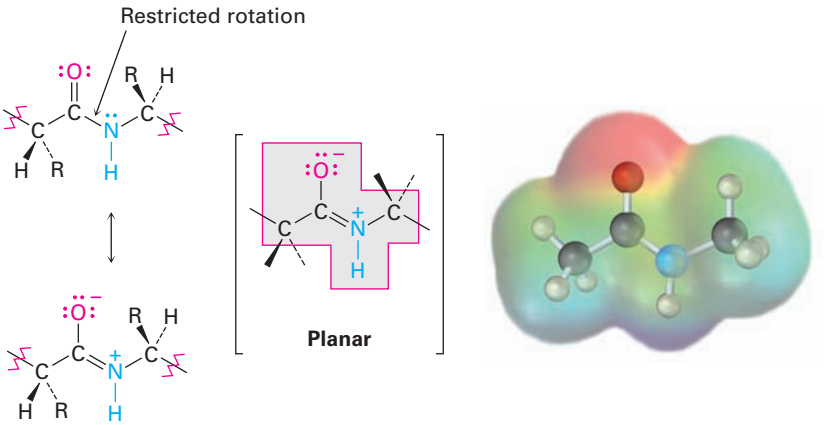


Note that two dipeptides can result from reaction between alanine and serine, depending on which carboxyl group reacts with which amino group. If the alanine amino group reacts with the serine carboxyl, serylalanine results.

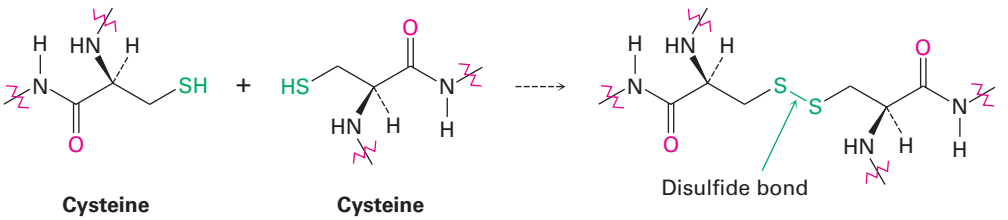


The long, repetitive sequence of –N–CH–CO– atoms that makes up a continuous chain is called the protein's **backbone**. By convention, peptides are written with the **N-terminal amino acid** (the one with the free $-NH_3^+$ group) on the left and the **C-terminal amino acid** (the one with the free $-CO_2^-$ group) on the right. The name of the peptide is indicated by using the abbreviations listed in Table 26.1 for each amino acid. Thus, alanylserine is abbreviated Ala-Ser or A-S, and serylalanine is abbreviated Ser-Ala or S-A. The one-letter abbreviations are more convenient, though less immediately recognizable, than the three-letter abbreviations.

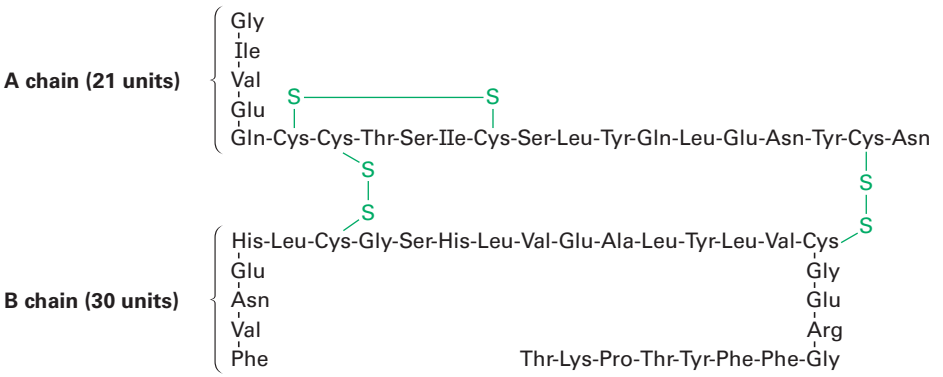
The amide bond that links different amino acids together in peptides is no different from any other amide bond **(Section 24.3)**. An amide nitrogen is nonbasic because its unshared electron pair is delocalized by interaction with the carbonyl group. This overlap of the nitrogen *p* orbital with the *p* orbitals of the carbonyl group imparts a certain amount of double-bond character to the C–N bond and restricts rotation around it. The amide bond is therefore planar, and the N–H is oriented 180° to the C=O.



A second kind of covalent bonding in peptides occurs when a disulfide linkage, RS–SR, is formed between two cysteine residues. As we saw in **Section 18.8**, a disulfide is formed by mild oxidation of a thiol, RSH, and is cleaved by mild reduction.



A disulfide bond between cysteine residues in different peptide chains links the otherwise separate chains together, while a disulfide bond between cysteine residues in the same chain forms a loop. Insulin, for instance, is composed of two chains that total 51 amino acids and are linked by two cysteine disulfide bridges.



Insulin

Problem 26.8

There are six isomeric tripeptides that contain valine, tyrosine, and glycine. Name them using both three- and one-letter abbreviations.

Problem 26.9

Draw the structure of M-P-V-G, and indicate the amide bonds.

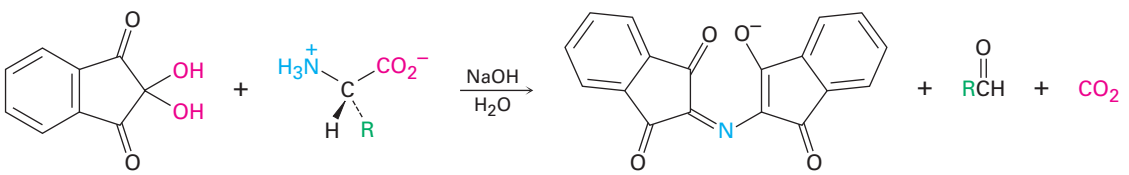
26.5 Amino Acid Analysis of Peptides

To determine the structure of a protein or peptide, we need to answer three questions: What amino acids are present? How much of each is present? In what sequence do the amino acids occur in the peptide chain? The answers to the first two questions are provided by an automated instrument called an amino acid analyzer.

An amino acid analyzer is based on analytical techniques worked out in the 1950s by William Stein and Stanford Moore, who shared the 1972 Nobel Prize in Chemistry for their work. In preparation for analysis, the peptide is broken into its constituent amino acids by reducing all disulfide bonds, capping the −SH groups of cysteine residues by S_N2 reaction with iodoacetic acid, and hydrolyzing the amide bonds by heating with aqueous 6 M HCl at 110 °C for 24 hours. The resultant amino acid mixture is then separated into its components by a technique called *chromatography,* either high-pressure liquid chromatography (HPLC) or ion-exchange chromatography.

In both HPLC and ion-exchange chromatography, the mixture to be separated is dissolved in a solvent, called the *mobile phase,* and passed through a metal tube or glass column that contains an adsorbent material, called the *stationary phase.* Because different compounds adsorb to the stationary phase to different extents, they migrate through the chromatography column at different rates and are separated as they emerge *(elute)* from the end.

In the ion-exchange technique, separated amino acids eluting from the chromatography column mix with a solution of a substance called *ninhydrin* and undergo a rapid reaction that produces an intense purple color. The color is detected by a spectrometer, and a plot of elution time versus spectrometer absorbance is obtained.



Ninhydrin **α-Amino acid** **(purple color)**

Because the time required for a given amino acid to elute from a standard column is reproducible, the identities of the amino acids in a peptide can be determined. The amount of each amino acid in the sample is determined by measuring the intensity of the purple color resulting from its reaction with ninhydrin. **Figure 26.3** shows the results of amino acid analysis of a standard equimolar mixture of 17 α-amino acids. Typically, amino acid analysis requires about 100 picomoles (2–3 μg) of sample for a protein containing about 200 residues.

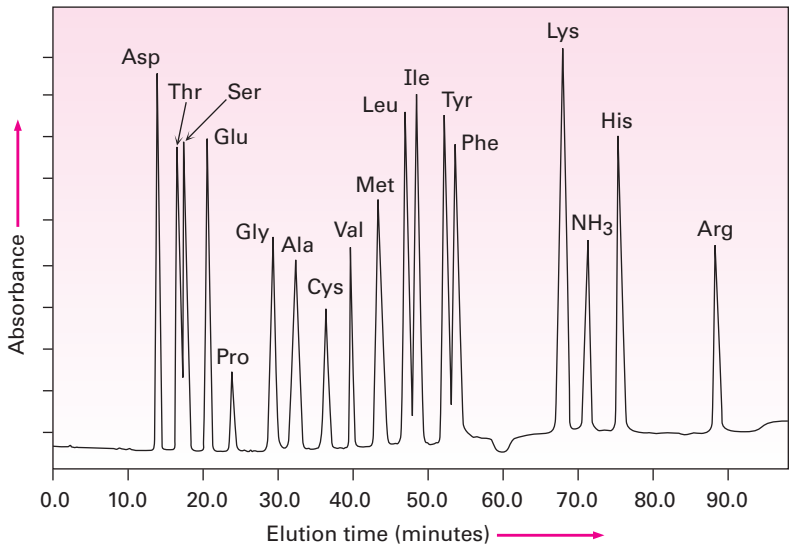


Figure 26.3 Amino acid analysis of an equimolar mixture of 17 amino acids.

Problem 26.10

Show the structure of the product you would expect to obtain by S_N2 reaction of a cysteine residue with iodoacetic acid.

Problem 26.11

Show the structures of the products obtained on reaction of valine with ninhydrin.

26.6 Peptide Sequencing: The Edman Degradation

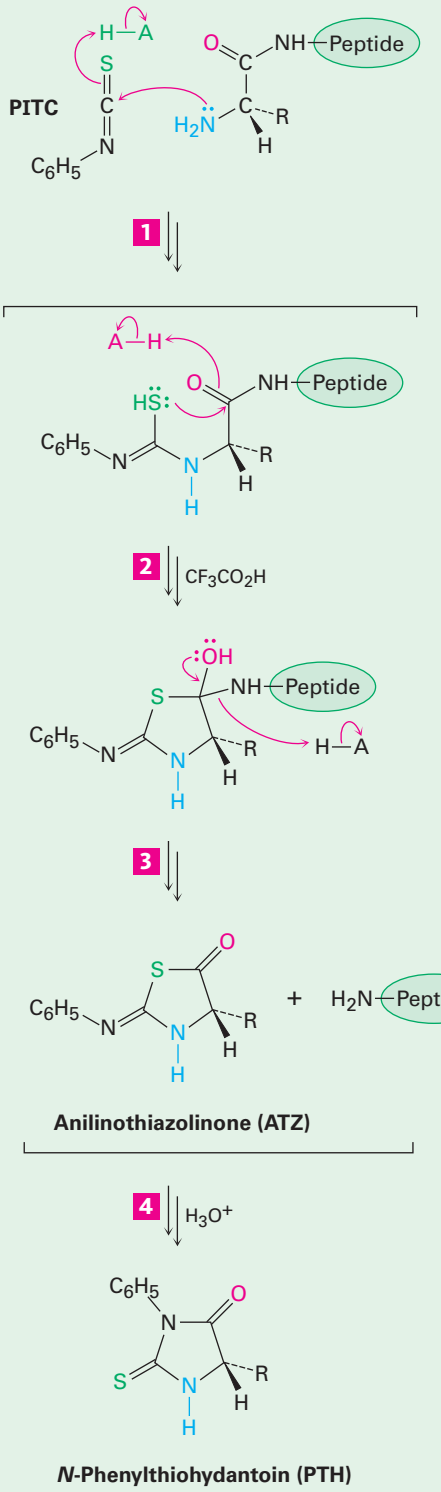
With the identities and relative amounts of amino acids known, the peptide is then sequenced to find out in what order the amino acids are linked together. Much peptide sequencing is now done by mass spectrometry, using either electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI) linked to a time-of-flight (TOF) mass analyzer, as described in **Section 12.4**. Also in common use is a chemical method of peptide sequencing called the Edman degradation.

The general idea of peptide sequencing by Edman degradation is to cleave one amino acid at a time from an end of the peptide chain. That terminal amino acid is then separated and identified, and the cleavage reactions are repeated on the chain-shortened peptide until the entire peptide sequence is known. Automated protein sequencers are available that allow as many as 50 repetitive sequencing cycles to be carried out before a buildup of unwanted by-products interferes with the results. So efficient are these instruments that sequence information can be obtained from as little as 1 to 5 picomoles of sample—less than 0.1 μg.

As shown in **Figure 26.4**, **Edman degradation** involves treatment of a peptide with phenyl isothiocyanate (PITC), $C_6H_5—N{=}C{=}S$, followed by reaction with trifluoroacetic acid. The first step attaches the PITC to the $-NH_2$ group of the N-terminal amino acid, and the second step splits the N-terminal residue from the peptide chain, yielding an anilinothiazolinone (ATZ) derivative plus the

Figure 26.4 | MECHANISM

Mechanism of the Edman degradation for N-terminal analysis of peptides.



1 Nucleophilic addition of the peptide terminal amino group to phenyl isothiocyanate (PITC) gives an *N*-phenylthiourea derivative.

2 Acid-catalyzed cyclization of the phenylthiourea yields a tetrahedral intermediate . . .

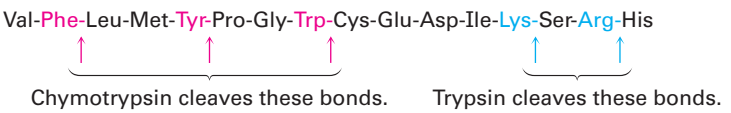
3 . . . which expels the chain-shortened peptide and forms an anilinothiazolinone (ATZ) derivative.

4 The ATZ rearranges in the presence of aqueous acid to an isomeric *N*-phenylthiohydantoin (PTH) as the final product.

chain-shortened peptide. Further acid-catalyzed rearrangement of the ATZ derivative with aqueous acid converts it into a phenylthiohydantoin (PTH), which is identified by comparison of its elution time with the known elution times of PTH derivatives of the 20 common amino acids. The chain-shortened peptide is then automatically resubmitted to another round of Edman degradation.

Complete sequencing of large proteins by Edman degradation is impractical because of the buildup of unwanted by-products. To get around the problem, a large peptide chain is first cleaved by partial hydrolysis into a number of smaller fragments, the sequence of each fragment is determined, and the individual fragments are fitted together by matching the overlapping ends. In this way, protein chains with more than 400 amino acids have been sequenced.

Partial hydrolysis of a peptide can be carried out either chemically with aqueous acid or enzymatically. Acid hydrolysis is unselective and gives a more-or-less random mixture of small fragments, but enzymatic hydrolysis is quite specific. The enzyme trypsin, for instance, catalyzes hydrolysis of peptides only at the carboxyl side of the basic amino acids arginine and lysine; chymotrypsin cleaves only at the carboxyl side of the aryl-substituted amino acids phenylalanine, tyrosine, and tryptophan.

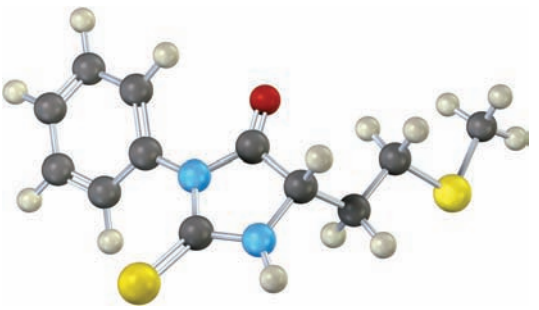


Problem 26.12

The octapeptide angiotensin II has the sequence Asp-Arg-Val-Tyr-Ile-His-Pro-Phe. What fragments would result if angiotensin II were cleaved with trypsin? With chymotrypsin?

Problem 26.13

What is the N-terminal residue on a peptide that gives the following PTH derivative on Edman degradation?



Problem 26.14

Draw the structure of the PTH derivative that would be formed on Edman degradation of angiotensin II (Problem 26.12).

Problem 26.15

Give the amino acid sequence of hexapeptides that produce the following sets of fragments on partial acid hydrolysis:

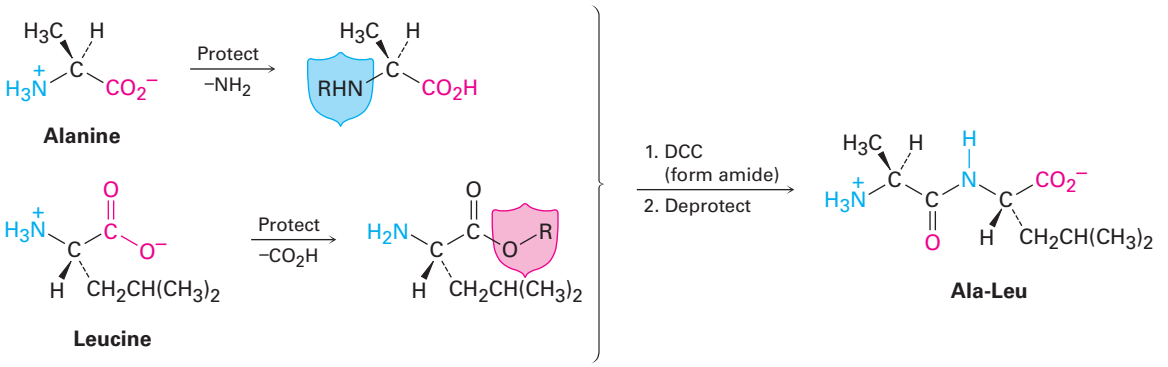
(a) Arg, Gly, Ile, Leu, Pro, Val gives Pro-Leu-Gly, Arg-Pro, Gly-Ile-Val

(b) N, L, M, W, V_2 gives V-L, V-M-W, W-N-V

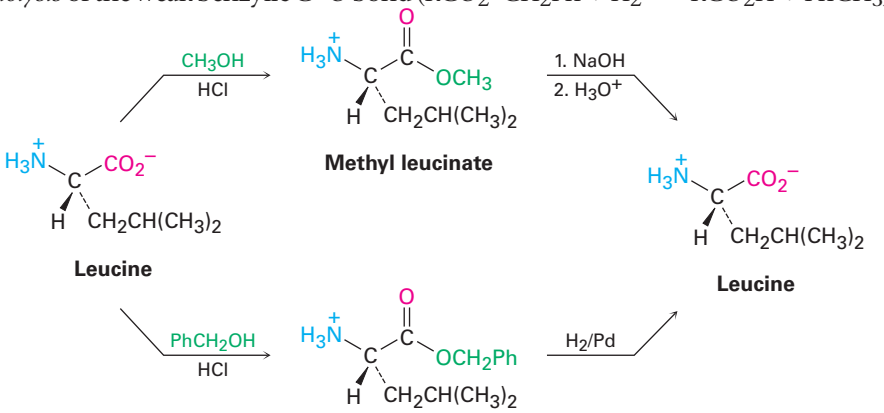
26.7 Peptide Synthesis

Once the structure of a peptide is known, its synthesis can then be undertaken—perhaps to obtain a larger amount for biological evaluation. A simple amide might be formed by treating an amine and a carboxylic acid with dicyclohexylcarbodiimide (DCC; **Section 21.7**), but peptide synthesis is a more difficult problem because many different amide bonds must be formed in a specific order rather than at random.

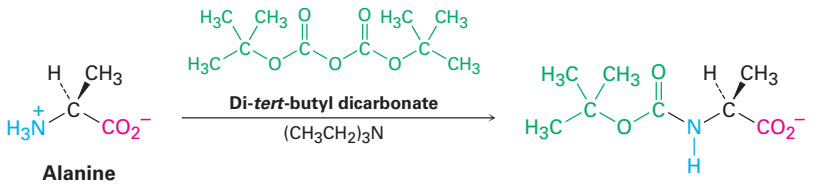
The solution to the specificity problem is protection **(Section 17.8)**. If we want to couple alanine with leucine to synthesize Ala-Leu, for instance, we could protect the $-NH_2$ group of alanine and the $-CO_2H$ group of leucine to shield them from reacting, then form the desired Ala-Leu amide bond by reaction with DCC, and then remove the protecting groups.

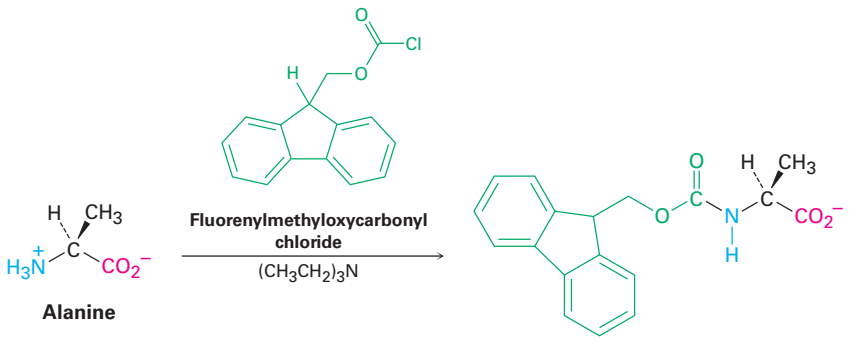


Many different amino- and carboxyl-protecting groups have been devised, but only a few are widely used. Carboxyl groups are often protected simply by converting them into methyl or benzyl esters. Both groups are easily introduced by standard methods of ester formation **(Section 21.6)** and are easily removed by mild hydrolysis with aqueous NaOH. Benzyl esters can also be cleaved by catalytic *hydrogenolysis* of the weak benzylic C–O bond ($RCO_2–CH_2Ph + H_2 \rightarrow RCO_2H + PhCH_3$).

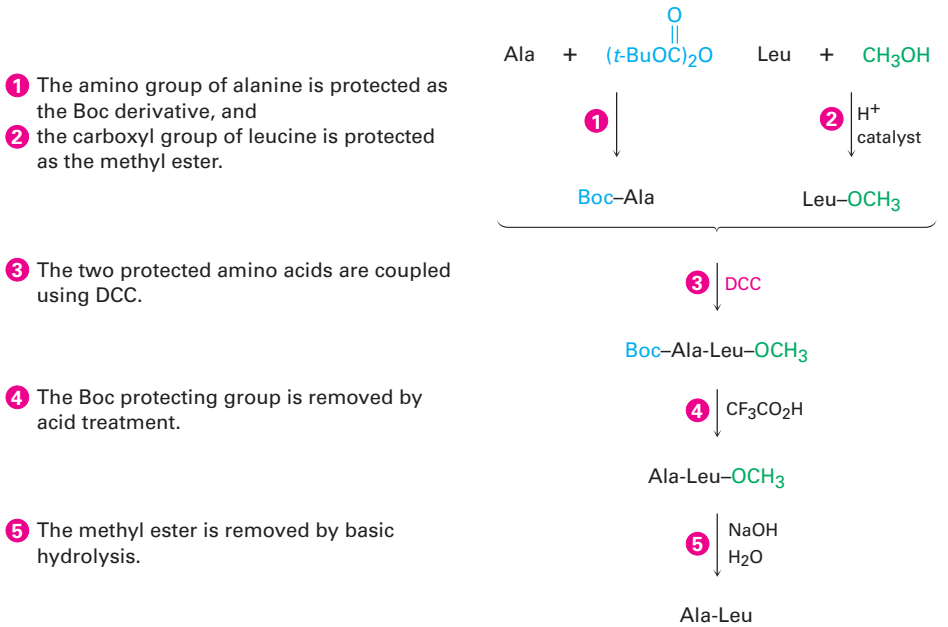


Amino groups are often protected as their *tert*-butyloxycarbonyl amide (Boc) or fluorenylmethyloxycarbonyl amide (Fmoc) derivatives. The Boc protecting group is introduced by reaction of the amino acid with di-*tert*-butyl dicarbonate in a nucleophilic acyl substitution reaction and is removed by brief treatment with a strong acid such as trifluoroacetic acid, CF_3CO_2H. The Fmoc protecting group is introduced by reaction with an acid chloride and is removed by treatment with base.





Thus, five steps are needed to synthesize a dipeptide such as Ala-Leu:



These steps can be repeated to add one amino acid at a time to the growing chain or to link two peptide chains together. Many remarkable achievements in peptide synthesis have been reported, including a complete synthesis of human insulin. Insulin is composed of two chains totaling 51 amino acids linked by two disulfide bridges. Its structure, shown previously on page 1057, was determined by Frederick Sanger, who received the 1958 Nobel Prize in Chemistry for his work.

Problem 26.16

Show the mechanism for formation of a Boc derivative by reaction of an amino acid with di-*tert*-butyl dicarbonate.

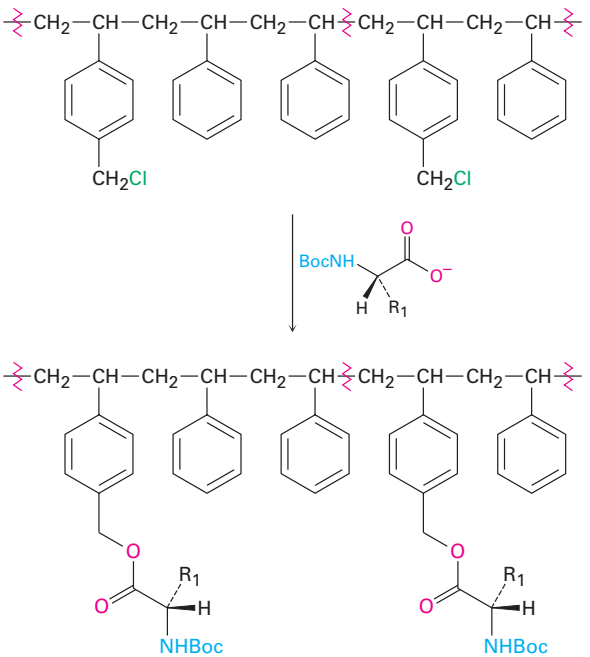
Problem 26.17

Write all five steps required for the synthesis of Leu-Ala from alanine and leucine.

26.8 Automated Peptide Synthesis: The Merrifield Solid-Phase Method

As you might imagine, the synthesis of a large peptide chain by sequential addition of one amino acid at a time is a long and arduous process. An immense simplification is possible, however, using methods introduced by R. Bruce Merrifield, who received the 1984 Nobel Prize in Chemistry for his work. In the Merrifield solid-phase method, peptide synthesis is carried out with the growing amino acid chain covalently bonded to small beads of a polymer resin rather than in solution.

In the original procedure, polystyrene resin was used, prepared so that 1 of every 100 or so benzene rings contained a chloromethyl ($-CH_2Cl$) group. A Boc-protected C-terminal amino acid was then attached to the resin through an ester bond formed by S_N2 reaction.

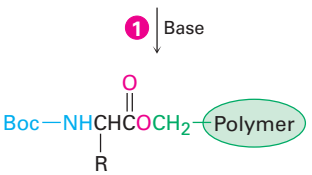


Chloromethylated polystyrene resin

Resin-bound amino acid

With the first amino acid bonded to the resin, a repeating series of four steps is then carried out to build a peptide.

1 A Boc-protected amino acid is covalently linked to the polystyrene polymer by formation of an ester bond (S_N2 reaction).

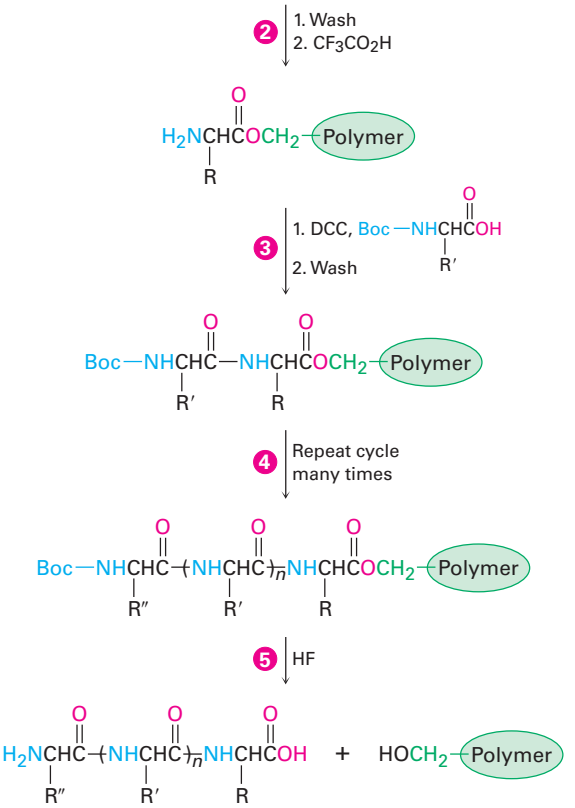


2 The polymer-bonded amino acid is washed free of excess reagent and then treated with trifluoroacetic acid to remove the Boc group.

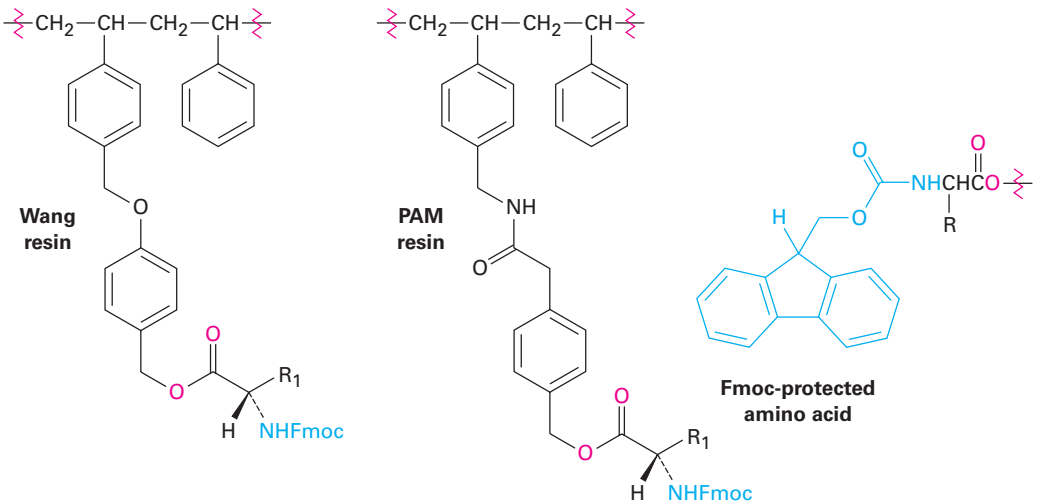
3 A second Boc-protected amino acid is coupled to the first by reaction with DCC. Excess reagents are removed by washing them from the insoluble polymer.

4 The cycle of deprotection, coupling, and washing is repeated as many times as desired to add amino acid units to the growing chain.

5 After the desired peptide has been made, treatment with anhydrous HF removes the final Boc group and cleaves the ester bond to the polymer, yielding the free peptide.



The steps in the solid-phase procedure have been improved substantially over the years, but the fundamental idea remains the same. The most commonly used resins at present are either the Wang resin or the PAM (phenylacetamidomethyl) resin, and the most commonly used N-protecting group is the Fmoc group rather than Boc.



Robotic peptide synthesizers are now used to automatically repeat the coupling, washing, and deprotection steps with different amino acids. Each step

occurs in high yield, and mechanical losses are minimized because the peptide intermediates are never removed from the insoluble polymer until the final step. Using this procedure, up to 25 to 30 mg of a peptide with 20 amino acids can be routinely prepared in a few hours.

26.9 Protein Structure

Proteins are usually classified as either fibrous or globular, according to their three-dimensional shape. **Fibrous proteins**, such as the collagen in tendons and connective tissue and the myosin in muscle tissue, consist of polypeptide chains arranged side by side in long filaments. Because these proteins are tough and insoluble in water, they are used in nature for structural materials. **Globular proteins**, by contrast, are usually coiled into compact, roughly spherical shapes. These proteins are generally soluble in water and are mobile within cells. Most of the 3000 or so enzymes that have been characterized to date are globular proteins.

Proteins are so large that the word *structure* takes on a broader meaning than it does with simpler organic compounds. In fact, chemists speak of four different levels of structure when describing proteins.

- The **primary structure** of a protein is simply the amino acid sequence.
- The **secondary structure** of a protein describes how *segments* of the peptide backbone orient into a regular pattern.
- The **tertiary structure** describes how the *entire* protein molecule coils into an overall three-dimensional shape.
- The **quaternary structure** describes how different protein molecules come together to yield large aggregate structures.

Primary structure is determined, as we've seen, by sequencing the protein. Secondary, tertiary, and quaternary structures are determined either by NMR or by X-ray crystallography (Chapter 12 *A Deeper Look*).

The most common secondary structures are the α helix and the β-pleated sheet. An ***α* helix** is a right-handed coil of the protein backbone, much like the coil of a spiral staircase **(Figure 26.5a)**. Each turn of the helix contains 3.6 amino acid residues, with a distance between coils of 540 pm, or 5.4 Å. The structure is stabilized by hydrogen bonds between amide N–H groups and C=O groups four residues away, with an N–H···O distance of 2.8 Å. The α helix is an extremely common secondary structure, and almost all globular proteins contain many helical segments. Myoglobin, a small globular protein containing 153 amino acid residues in a single chain, is an example (Figure 26.5b).

A ***β*-pleated sheet** differs from an α helix in that the peptide chain is fully extended rather than coiled and the hydrogen bonds occur between residues in adjacent chains **(Figure 26.6a)**. The neighboring chains can run either in the same direction (parallel) or in opposite directions (antiparallel), although the antiparallel arrangement is more common and energetically somewhat more favorable. Concanavalin A, for instance, consists of two identical chains of 237 residues, with extensive regions of antiparallel β sheets (Figure 26.6b).

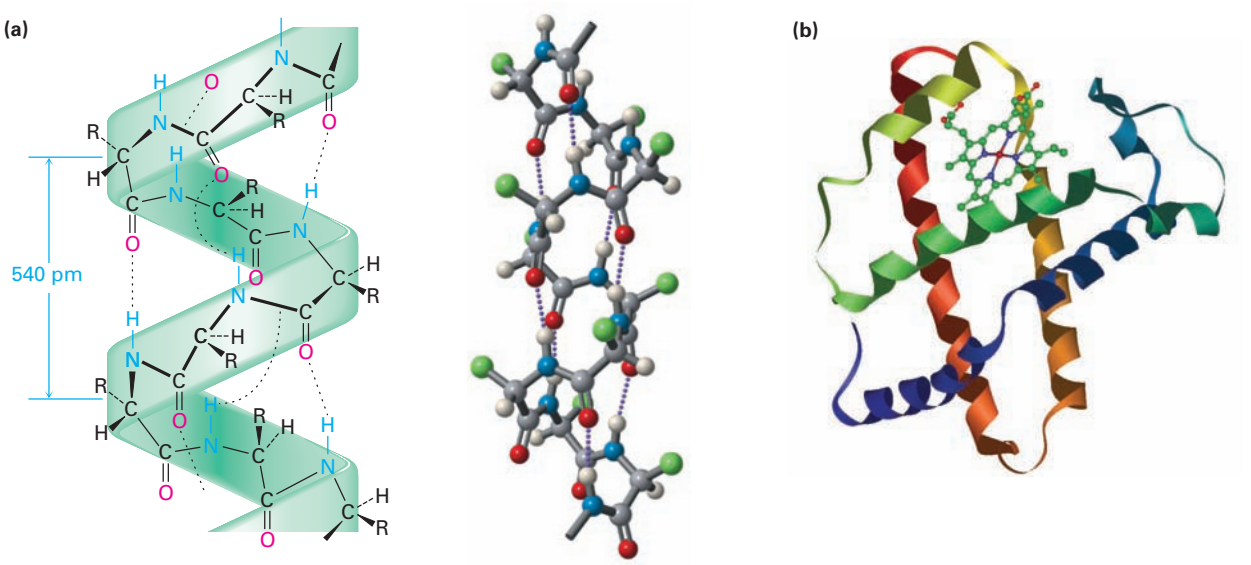


Figure 26.5 **(a)** The α-helical secondary structure of proteins is stabilized by hydrogen bonds between the N—H group of one residue and the C=O group four residues away. **(b)** The structure of myoglobin, a globular protein with extensive helical regions that are shown as coiled ribbons in this representation.

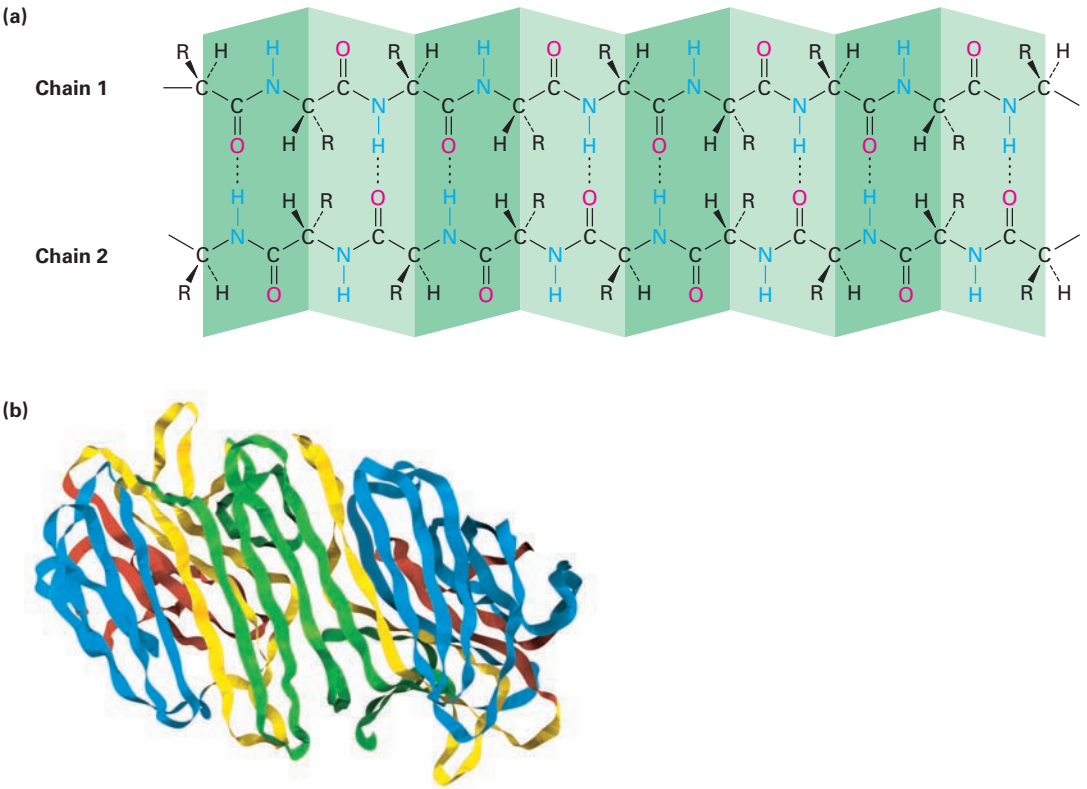


Figure 26.6 **(a)** The β-pleated sheet secondary structure of proteins is stabilized by hydrogen bonds between parallel or antiparallel chains. **(b)** The structure of concanavalin A, a protein with extensive regions of antiparallel β sheets, shown as flat ribbons.

What about tertiary structure? Why does any protein adopt the shape it does? The forces that determine the tertiary structure of a protein are the same forces that act on all molecules, regardless of size, to provide maximum stability. Particularly important are the hydrophilic (water-loving; **Section 2.12**) interactions of the polar side chains on acidic or basic amino acids and the hydrophobic (water-fearing) interactions of nonpolar side chains. Those acidic or basic amino acids with charged side chains tend to congregate on the exterior of the protein, where they can be solvated by water. Those amino acids with neutral, nonpolar side chains tend to congregate on the hydrocarbon-like interior of a protein molecule, away from the aqueous medium.

Also important for stabilizing a protein's tertiary structure are the formation of disulfide bridges between cysteine residues, the formation of hydrogen bonds between nearby amino acid residues, and the presence of ionic attractions, called *salt bridges,* between positively and negatively charged sites on various amino acid side chains within the protein.

Because the tertiary structure of a globular protein is delicately held together by weak intramolecular attractions, a modest change in temperature or pH is often enough to disrupt that structure and cause the protein to become **denatured**. Denaturation occurs under such mild conditions that the primary structure remains intact but the tertiary structure unfolds from a specific globular shape to a randomly looped chain **(Figure 26.7)**.

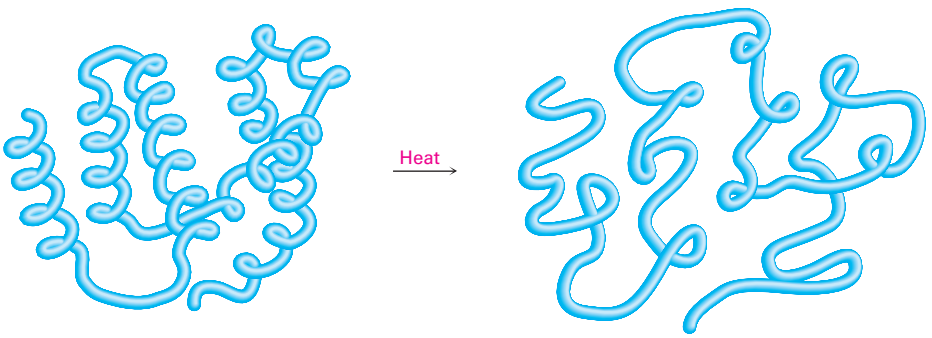


Figure 26.7 A representation of protein denaturation. A globular protein loses its specific three-dimensional shape and becomes randomly looped.

Denaturation is accompanied by changes in both physical and biological properties. Solubility is drastically decreased, as occurs when egg white is cooked and the albumins unfold and coagulate. Most enzymes also lose all catalytic activity when denatured, since a precisely defined tertiary structure is required for their action. Although most denaturation is irreversible, some cases are known where spontaneous renaturation of an unfolded protein to its stable tertiary structure occurs, accompanied by a full recovery of biological activity.

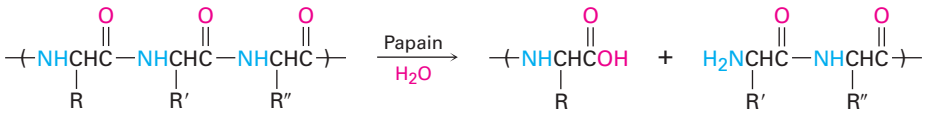
26.10 Enzymes and Coenzymes

An **enzyme** is a substance—usually a large protein—that acts as a catalyst for a biological reaction. Like all catalysts, an enzyme doesn't affect the equilibrium constant of a reaction and can't bring about a chemical change that is otherwise unfavorable. An enzyme acts only to lower the activation energy for a

reaction, thereby making the reaction take place more rapidly. Sometimes, in fact, the rate acceleration brought about by enzymes is extraordinary. Million-fold rate increases are common, and the glycosidase enzymes that hydrolyze polysaccharides increase the reaction rate by a factor of more than 10^{17}, changing the time required for the reaction from millions of years to milliseconds!

Unlike many of the catalysts that chemists use in the laboratory, enzymes are usually specific in their action. Often, in fact, an enzyme will catalyze only a single reaction of a single compound, called the enzyme's *substrate*. For example, the enzyme amylase, found in the human digestive tract, catalyzes only the hydrolysis of starch to yield glucose; cellulose and other polysaccharides are untouched by amylase.

Different enzymes have different specificities. Some, such as amylase, are specific for a single substrate, but others operate on a range of substrates. Papain, for instance, a globular protein of 212 amino acids isolated from papaya fruit, catalyzes the hydrolysis of many kinds of peptide bonds. In fact, it's this ability to hydrolyze peptide bonds that makes papain useful as a cleaner for contact lenses.



Enzymes function through a pathway that involves initial formation of an enzyme–substrate complex E · S, followed by a multistep chemical conversion of the enzyme-bound substrate into enzyme-bound product E · P and final release of product from the complex.

$$E + S \rightleftharpoons E \cdot S \rightleftharpoons E \cdot P \rightleftharpoons E + P$$

The overall rate constant for conversion of the E · S complex to products E + P is called the **turnover number** because it represents the number of substrate molecules a single enzyme molecule turns over into product per unit time. A value of about 10^3 per second is typical although carbonic anhydrase can reach a value of up to 600,000.

The extraordinary rate accelerations achieved by enzymes are due to a combination of several factors. One important factor is simple geometry: an enzyme will adjust its shape to hold the substrate, other reactants, and various catalytic sites on acidic or basic residues in the precise geometry needed for reaction. In addition, the wrapping of the enzyme around the substrate can create specialized microenvironments that protect the substrate from the aqueous medium and can dramatically change the behavior of acid–base catalytic residues in the active site. But perhaps most important is that the enzyme stabilizes and thus lowers the energy of the rate-limiting transition state for reaction. That is, it's not the ability of the enzyme to bind the *substrate* that matters but rather its ability to bind and stabilize the *transition state*. Often, in fact, the enzyme binds the transition structure as much as 10^{12} times more tightly than it binds the substrate or products. An energy diagram for an enzyme-catalyzed process might look like that in **Figure 26.8**.

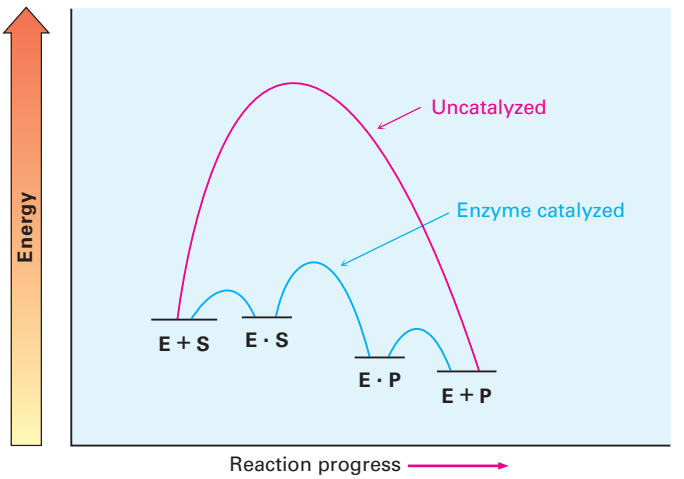


Figure 26.8 Energy diagrams for **uncatalyzed** and **enzyme-catalyzed** processes. The enzyme makes available an alternative, lower-energy pathway. Rate enhancement is due to the ability of the enzyme to bind to the transition state for product formation, thereby lowering its energy.

Enzymes are classified into six categories depending on the kind of reaction they catalyze, as shown in Table 26.2. *Oxidoreductases* catalyze oxidations and reductions; *transferases* catalyze the transfer of a group from one substrate to another; *hydrolases* catalyze hydrolysis reactions of esters, amides, and related substrates; *lyases* catalyze the elimination or addition of a small molecule such as H_2O from or to a substrate; *isomerases* catalyze isomerizations; and *ligases* catalyze the bonding together of two molecules, often coupled with the hydrolysis of ATP. The systematic name of an enzyme has two parts, ending with *-ase*. The first part identifies the enzyme's substrate, and the second part identifies its class. Hexose kinase, for example, is a transferase that catalyzes the transfer of a phosphate group from ATP to a hexose sugar.

Table 26.2 Classification of Enzymes

Class	Some subclasses	Function
Oxidoreductases	Dehydrogenases	Introduction of double bond
	Oxidases	Oxidation
	Reductases	Reduction
Transferases	Kinases	Transfer of phosphate group
	Transaminases	Transfer of amino group
Hydrolases	Lipases	Hydrolysis of ester
	Nucleases	Hydrolysis of phosphate
	Proteases	Hydrolysis of amide
Lyases	Decarboxylases	Loss of CO_2
	Dehydrases	Loss of H_2O
Isomerases	Epimerases	Isomerization of chirality center
Ligases	Carboxylases	Addition of CO_2
	Synthetases	Formation of new bond

In addition to their protein part, most enzymes also contain a small nonprotein part called a **cofactor**. A cofactor can be either an inorganic ion, such as Zn^{2+}, or a small organic molecule, called a **coenzyme**. A coenzyme is not a catalyst but is a reactant that undergoes chemical change during the reaction and requires an additional step to return to its initial state.

Many coenzymes are derived from vitamins—substances that an organism requires in small amounts for growth but is unable to synthesize and must receive in its diet (Chapter 20 *A Deeper Look*). Coenzyme A from pantothenate (vitamin B_3), NAD^+ from niacin, FAD from riboflavin (vitamin B_2), tetrahydrofolate from folic acid, pyridoxal phosphate from pyridoxine (vitamin B_6), and thiamin diphosphate from thiamin (vitamin B_1) are examples. Table 26.3 on the following two pages shows the structures of some common coenzymes.

Problem 26.18

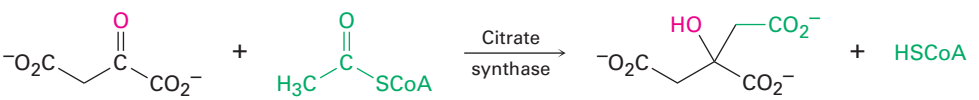
To what classes do the following enzymes belong?

(a) Pyruvate decarboxylase **(b)** Chymotrypsin

(c) Alcohol dehydrogenase

26.11 How Do Enzymes Work? Citrate Synthase

As we saw in the previous section, enzymes work by bringing substrate and other reactant molecules together, holding them in the orientation necessary for reaction, providing any necessary acidic or basic sites to catalyze specific steps, and stabilizing the transition state for reaction. As an example, let's look at citrate synthase, an enzyme that catalyzes the aldol-like addition of acetyl CoA to oxaloacetate to give citrate. The reaction is the first step in the citric acid cycle, in which acetyl groups produced by degradation of food molecules are metabolized to yield CO_2 and H_2O. We'll look at the details of the citric acid cycle in **Section 29.7**.

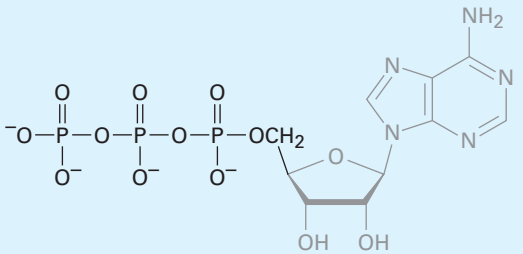


Oxaloacetate **Acetyl CoA** **Citrate**

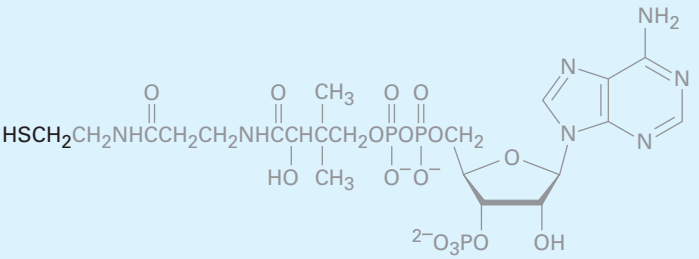
Citrate synthase is a globular protein of 433 amino acids with a deep cleft lined by an array of functional groups that can bind to the substrate, oxaloacetate. On binding oxaloacetate, the original cleft closes and another opens up nearby to bind acetyl CoA. This second cleft is also lined by appropriate functional groups, including a histidine at position 274 and an aspartic acid at position 375. The two reactants are now held by the enzyme in close proximity and with a suitable orientation for reaction. **Figure 26.9** shows the structure of citrate synthase as determined by X-ray crystallography, along with a close-up of the active site.

Table 26.3 Structures and Functions of Some Common Coenzymes

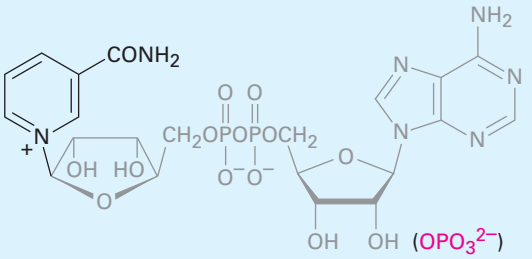
Adenosine triphosphate—ATP (phosphorylation)



Coenzyme A (acyl transfer)



Nicotinamide adenine dinucleotide—NAD^+ (oxidation/reduction)
($NADP^+$)



Flavin adenine dinucleotide—FAD (oxidation/reduction)

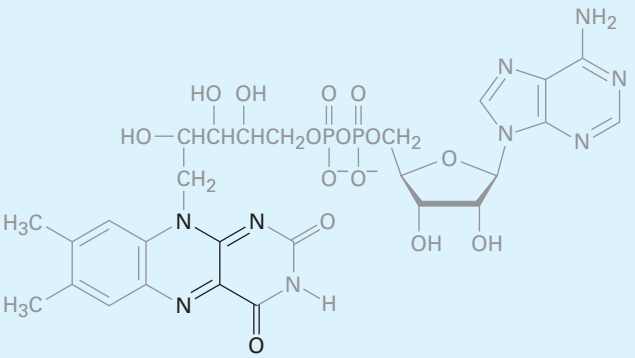
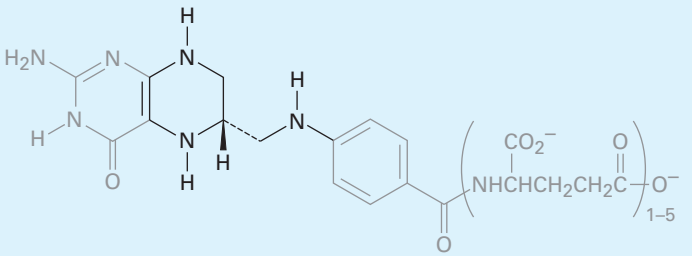
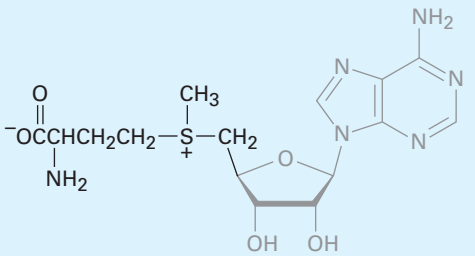


Table 26.3 Structures and Functions of Some Common Coenzymes *(continued)*

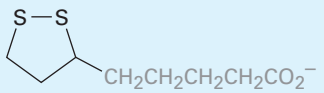
Tetrahydrofolate (transfer of C_1 units)



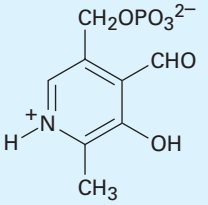
***S*-Adenosylmethionine (methyl transfer)**



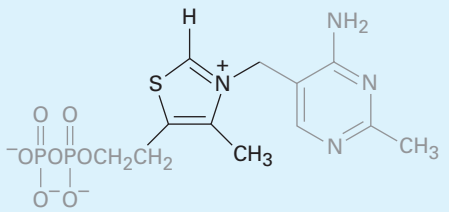
Lipoic acid (acyl transfer)



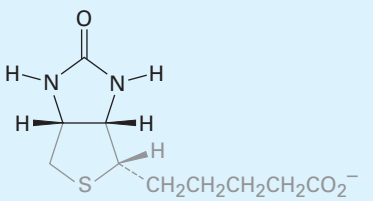
Pyridoxal phosphate (amino acid metabolism)



Thiamin diphosphate (decarboxylation)



Biotin (carboxylation)



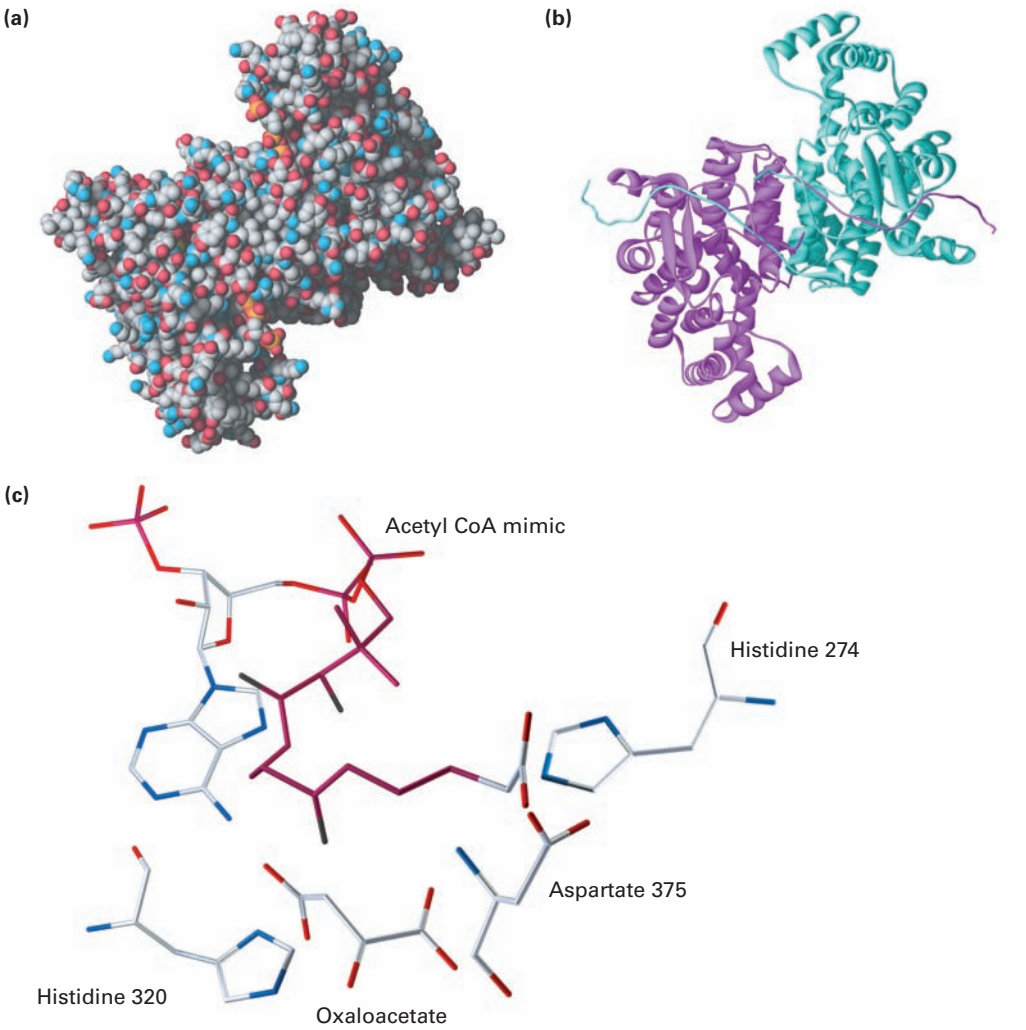


Figure 26.9 X-ray crystal structure of citrate synthase. Part **(a)** is a space-filling model and part **(b)** is a ribbon model, which emphasizes the α-helical segments of the protein chain and indicates that the enzyme is dimeric; that is, it consists of two identical chains held together by hydrogen bonds and other intermolecular attractions. Part **(c)** is a close-up of the active site in which oxaloacetate and an unreactive acetyl CoA mimic are bound.

As shown in **Figure 26.10**, the first step in the aldol reaction is generation of the enol of acetyl CoA. The side-chain carboxyl of an aspartate residue acts as base to abstract an acidic α proton, while at the same time the side-chain imidazole ring of a histidine donates H^+ to the carbonyl oxygen. The enol thus produced then does a nucleophilic addition to the ketone carbonyl group of oxaloacetate. The first histidine acts as a base to remove the −OH hydrogen from the enol, while a second histidine residue simultaneously donates a proton to the oxaloacetate carbonyl group, giving citryl CoA. Water then hydrolyzes the thiol ester group in citryl CoA in a nucleophilic acyl substitution reaction, releasing citrate and coenzyme A as the final products.

1 The side-chain carboxylate group of an aspartic acid acts as a base and removes an acidic α proton from acetyl CoA, while the N–H group on the side chain of a histidine acts as an acid and donates a proton to the carbonyl oxygen, giving an enol.

2 A histidine deprotonates the acetyl-CoA enol, which adds to the ketone carbonyl group of oxaloacetate in an aldol-like reaction. Simultaneously, an acid N–H proton of another histidine protonates the carbonyl oxygen, producing (*S*)-citryl CoA.

3 The thioester group of citryl CoA is hydrolyzed by a typical nucleophilic acyl substitution reaction to produce citrate plus coenzyme A.

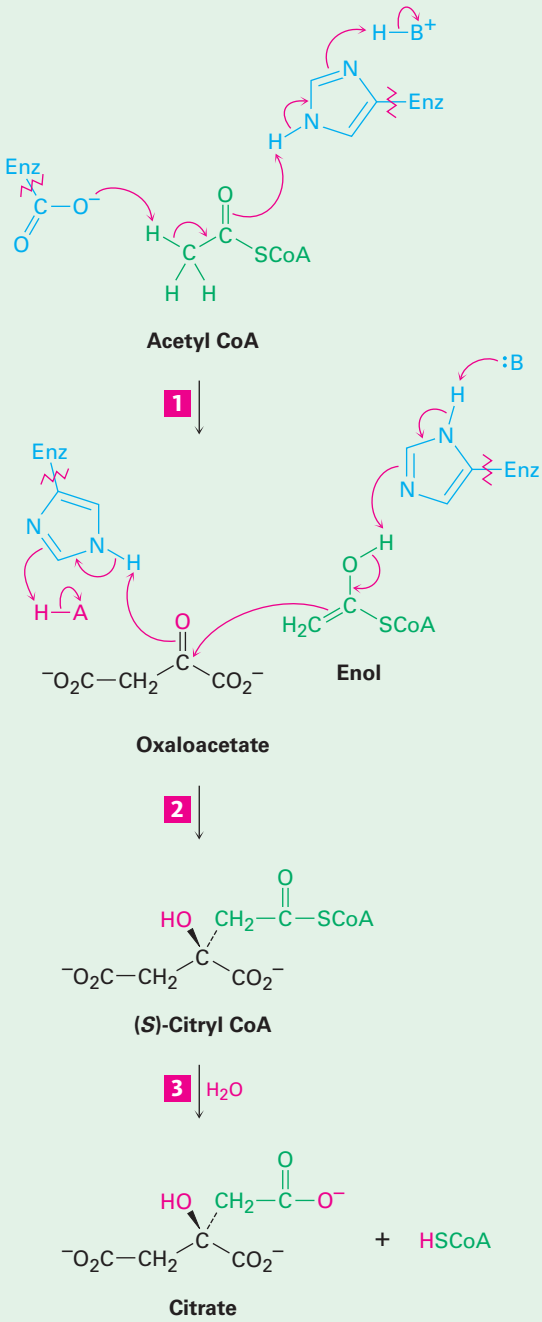


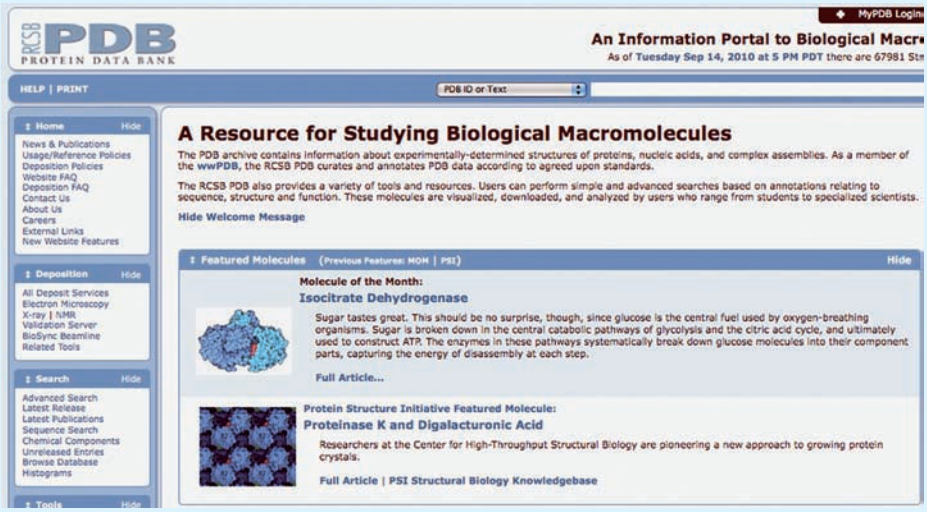
Figure 26.10 | MECHANISM

Mechanism of the addition of acetyl CoA to oxaloacetate to give (*S*)-citryl CoA, catalyzed by citrate synthase.

The Protein Data Bank | A DEEPER LOOK

Enzymes are so large, so structurally complex, and so numerous that the use of computer databases and molecular visualization programs has become an essential tool for studying biological chemistry. Of the various databases available online, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.ad.jp/kegg), maintained by the Kanehisa Laboratory of Kyoto University Bioinformatics Center, is useful for obtaining information on biosynthetic pathways of the sort we'll be describing in Chapter 29. For obtaining information on a specific enzyme, the BRENDA database (http://www.brenda-enzymes.org/), maintained by the Institute of Biochemistry at the University of Cologne, Germany, is particularly valuable.

Perhaps the most useful of all biological databases is the Protein Data Bank (PDB), operated by the Research Collaboratory for Structural Bioinformatics (RCSB). The PDB is a worldwide repository of X-ray and NMR structural data for biological macromolecules. In mid-2010, data for more than 66,000 structures were available, and more than 6000 new ones were being added yearly. To access the Protein Data Bank, go to http://www.rcsb.org/pdb/ and a home page like that shown in **Figure 26.11** will appear. As with much that is available online, however, the PDB site is changing rapidly, so you may not see quite the same thing.



Protein Data Bank

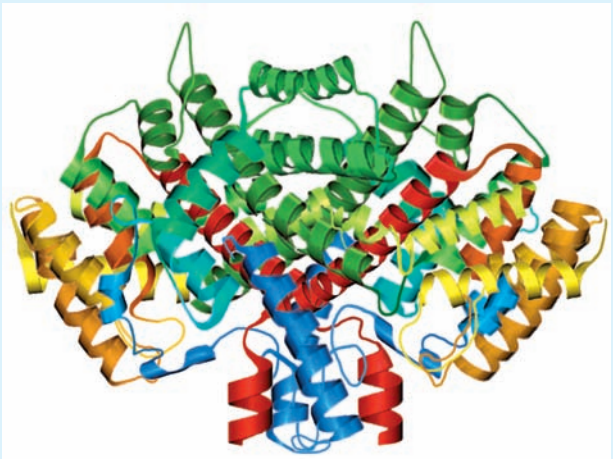
Figure 26.11 The Protein Data Bank home page.

To learn how to use the PDB, begin by running the short tutorial listed under Getting Started at the bottom of the page. After that introduction, start exploring. Let's say you want to view citrate synthase, the enzyme that catalyzes the addition of acetyl CoA to oxaloacetate to give citrate. Type "citrate synthase" (with quotation marks) into the small search box on the top line, click on "Search," and a list of 42 or so structures will appear. Scroll down near the end of the list until you find the entry with a PDB code of 5CTS and the title "Proposed Mechanism for the Condensation Reaction of Citrate Synthase: 1.9 Angstroms Structure of the Ternary Complex with Oxaloacetate and Carboxymethyl Coenzyme A." Alternatively, if you know the code of the enzyme you want, you can enter it directly into the search box. Click on the PDB code of entry 5CTS, and a new page containing information about the enzyme will open.

(continued)

(continued)

If you choose, you can download the structure file to your computer and open it with any of numerous molecular graphics programs to see an image like that in **Figure 26.12**. The biologically active molecule is a dimer of two identical subunits consisting primarily of α-helical regions displayed as coiled ribbons. For now, just click on "View in Jmol" under the enzyme image on the right side of the screen to see some of the options for visualizing and further exploring the enzyme.



PDB ID: 1AL6 B. Schwartz, et al. Mechanisms of Enzyme-Catalyzed Deprotonation of Acetyl-Coenzyme A

Figure 26.12 An image of citrate synthase, downloaded from the Protein Data Bank.

Summary

Proteins and **peptides** are large biomolecules made of ***α*-amino acid residues** linked together by amide, or peptide, bonds. Twenty amino acids are commonly found in proteins, and all except glycine have stereochemistry similar to that of L sugars. In neutral solution, amino acids exist as dipolar **zwitterions**.

Amino acids can be synthesized in racemic form by several methods, including ammonolysis of an α-bromo acid, alkylation of diethyl acetamidomalonate, and reductive amination of an α-keto acid. Alternatively, an enantioselective synthesis of amino acids can be carried out using a chiral hydrogenation catalyst.

Determining the structure of a peptide or protein begins with amino acid analysis. The peptide is hydrolyzed to its constituent α-amino acids, which are separated and identified. Next, the peptide is sequenced. **Edman degradation** by treatment with phenyl isothiocyanate (PITC) cleaves one residue from the N terminus of the peptide and forms an easily identifiable phenylthiohydantoin (PTH) derivative of the **N-terminal amino acid**. An automated series of Edman degradations can sequence peptide chains up to 50 residues in length.

Peptide synthesis involves the use of protecting groups. An N-protected amino acid with a free $-CO_2H$ group is coupled using DCC to an O-protected amino acid with a free $-NH_2$ group. Amide formation occurs, the protecting groups are removed, and the sequence is repeated. Amines are usually protected as their *tert*-butyloxycarbonyl (Boc) or fluorenylmethyloxycarbonyl (Fmoc) derivatives; acids are usually protected as esters. The synthesis is often carried out by the Merrifield solid-phase method, in which the peptide is bonded to insoluble polymer beads.

Key words

α-amino acid, 1045
α helix, 1066
backbone, 1056
β-pleated sheet, 1066
C-terminal amino acid, 1056
coenzyme, 1071
cofactor, 1071
denatured, 1068
Edman degradation, 1059
enzyme, 1068
fibrous protein, 1066
globular protein, 1066
isoelectric point, (p*I*), 1052
N-terminal amino acid, 1056
peptide, 1044
primary structure, 1066
protein, 1044
quaternary structure, 1066
residue, 1056
secondary structure, 1066
side chain, 1048
tertiary structure, 1066
turnover number, 1069
zwitterion, 1045

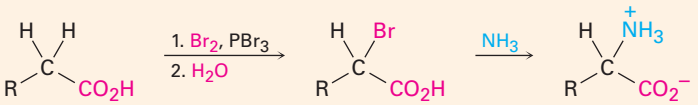
Proteins have four levels of structure. **Primary structure** describes a protein's amino acid sequence; **secondary structure** describes how segments of the protein chain orient into regular patterns—either ***α*-helix** or ***β*-pleated sheet**; **tertiary structure** describes how the entire protein molecule coils into an overall three-dimensional shape; and **quaternary structure** describes how individual protein molecules aggregate into larger structures.

Proteins are classified as either globular or fibrous. **Fibrous proteins** such as α-keratin are tough, rigid, and water-insoluble; **globular proteins** such as myoglobin are water-soluble and roughly spherical in shape. Many globular proteins are **enzymes**—substances that act as catalysts for biological reactions. Enzymes are grouped into six classes according to the kind of reaction they catalyze. In addition to their protein part, many enzymes contain **cofactors**, which can be either metal ions or small organic molecules called **coenzymes**.

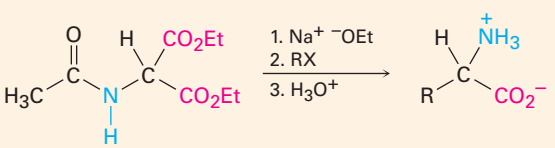
Summary of Reactions

1. Amino acid synthesis (Section 26.3)

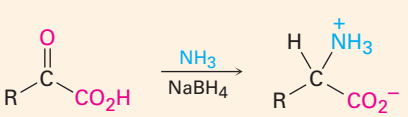
(a) From α-bromo acids



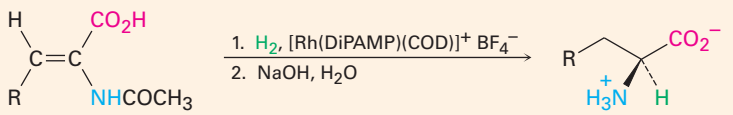
(b) Diethyl acetamidomalonate synthesis



(c) Reductive amination of an α-keto acid

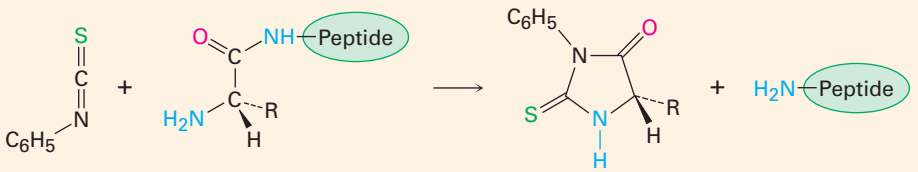


(d) Enantioselective synthesis

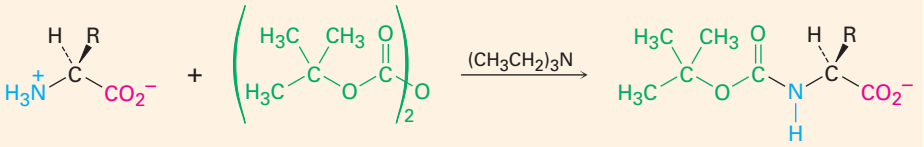


A (*Z*) enamido acid **An (*S*)-amino acid**

2. Peptide sequencing by Edman degradation (Section 26.6)

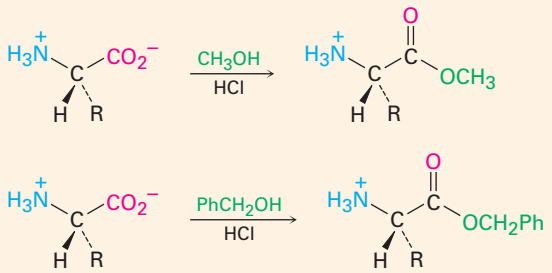


3. Peptide synthesis (Section 26.7)
   (a) Amine protection



Boc-protected amino acid

   (b) Carboxyl protection



Exercises

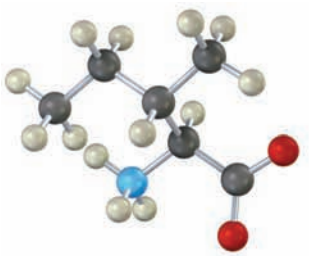
Visualizing Chemistry

OWL Interactive versions of these problems are assignable in OWL for Organic Chemistry.

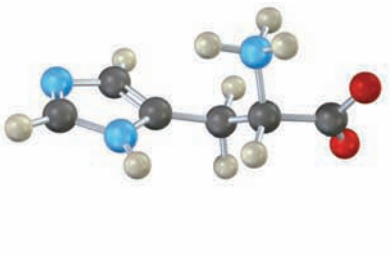
(Problems 26.1–26.18 appear within the chapter.)

26.19 Identify the following amino acids:

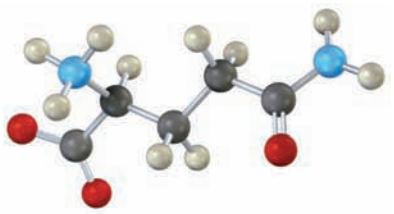
(a)



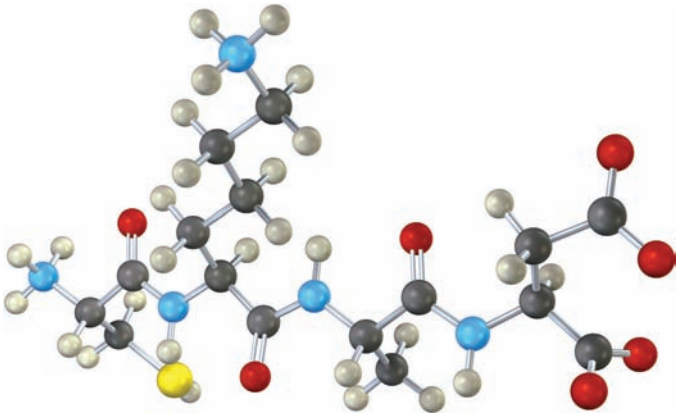
(b)



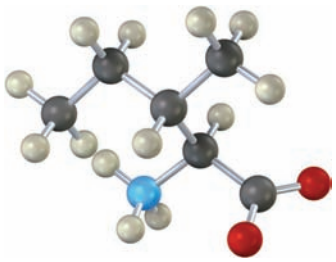
(c)



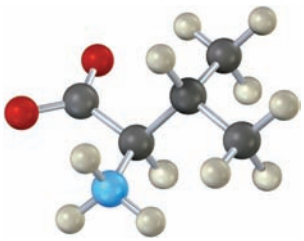
26.20 Give the sequence of the following tetrapeptide (yellow = S):



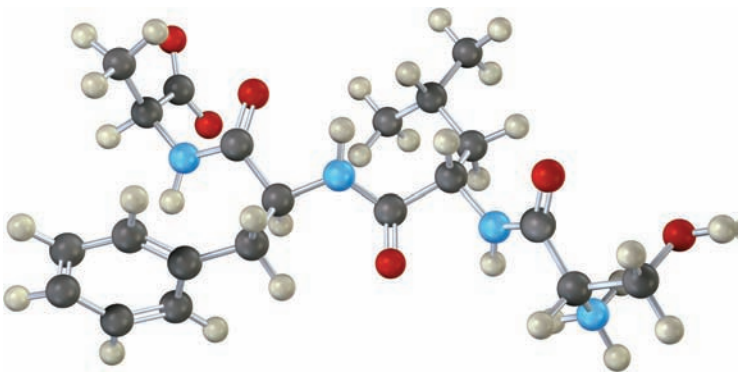
26.21 Isoleucine and threonine are the only two amino acids with two chirality centers. Assign *R* or *S* configuration to the methyl-bearing carbon atom of isoleucine.



26.22 Is the following structure a D amino acid or an L amino acid? Identify it.



26.23 Give the sequence of the following tetrapeptide:



Additional Problems

Amino Acid Structures and Chirality

26.24 Except for cysteine, only *S* amino acids occur in proteins. Several *R* amino acids are also found in nature, however. (*R*)-Serine is found in earthworms, and (*R*)-alanine is found in insect larvae. Draw Fischer projections of (*R*)-serine and (*R*)-alanine. Are these D or L amino acids?

26.25 Cysteine is the only amino acid that has L stereochemistry but an *R* configuration. Make up a structure for another L amino acid of your own creation that also has an *R* configuration.

26.26 Draw a Fischer projection of (*S*)-proline.

26.27 Show the structures of the following amino acids in their zwitterionic forms:
(a) Trp **(b)** Ile **(c)** Cys **(d)** His

26.28 Proline has $pK_{a1} = 1.99$ and $pK_{a2} = 10.60$. Use the Henderson–Hasselbalch equation to calculate the ratio of protonated and neutral forms at pH = 2.50. Calculate the ratio of neutral and deprotonated forms at pH = 9.70.

26.29 Using both three- and one-letter codes for amino acids, write the structures of all possible peptides containing the following amino acids:
(a) Val, Ser, Leu **(b)** Ser, Leu_2, Pro

26.30 Look at the side chains of the 20 amino acids in Table 26.1, and then think about what is *not* present. None of the 20 contain either an aldehyde or a ketone carbonyl group, for instance. Is this just one of nature's oversights, or is there a likely chemical reason? What complications might an aldehyde or ketone carbonyl group cause?

Amino Acid Synthesis and Reactions

26.31 Show how you could use the acetamidomalonate method to prepare the following amino acids:
(a) Leucine **(b)** Tryptophan

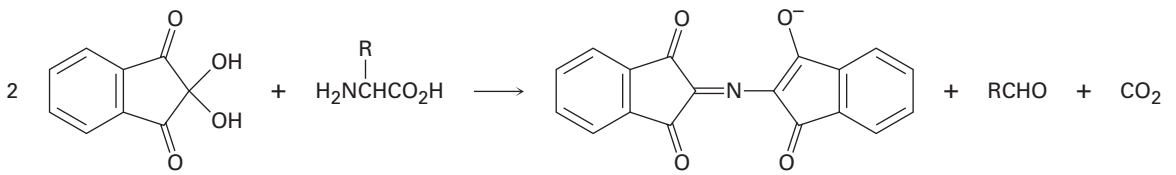
26.32 Show how you could prepare the following amino acids using a reductive amination:
(a) Methionine **(b)** Isoleucine

26.33 Show how you could prepare the following amino acids enantioselectively:
(a) Pro **(b)** Val

26.34 Serine can be synthesized by a simple variation of the amidomalonate method using formaldehyde rather than an alkyl halide. How might this be done?

26.35 Predict the product of the reaction of valine with the following reagents:
(a) CH_3CH_2OH, acid **(b)** Di-*tert*-butyl dicarbonate
(c) KOH, H_2O **(d)** CH_3COCl, pyridine; then H_2O

26.36 The reaction of ninhydrin with an α-amino acid occurs in several steps.
(a) The first step is formation of an imine by reaction of the amino acid with ninhydrin. Show its structure and the mechanism of its formation.
(b) The second step is a decarboxylation. Show the structure of the product and the mechanism of the decarboxylation reaction.
(c) The third step is hydrolysis of an imine to yield an amine and an aldehyde. Show the structures of both products and the mechanism of the hydrolysis reaction.
(d) The final step is formation of the purple anion. Show the mechanism of the reaction.



Ninhydrin

26.37 Draw resonance forms for the purple anion obtained by reaction of ninhydrin with an α-amino acid (Problem 26.36).

Peptides and Enzymes

26.38 Write full structures for the following peptides:
(a) C-H-E-M **(b)** P-E-P-T-I-D-E

26.39 Propose two structures for a tripeptide that gives Leu, Ala, and Phe on hydrolysis but does not react with phenyl isothiocyanate.

26.40 Show the steps involved in a synthesis of Phe-Ala-Val using the Merrifield procedure.

26.41 Draw the structure of the PTH derivative product you would obtain by Edman degradation of the following peptides:
(a) I-L-P-F **(b)** D-T-S-G-A

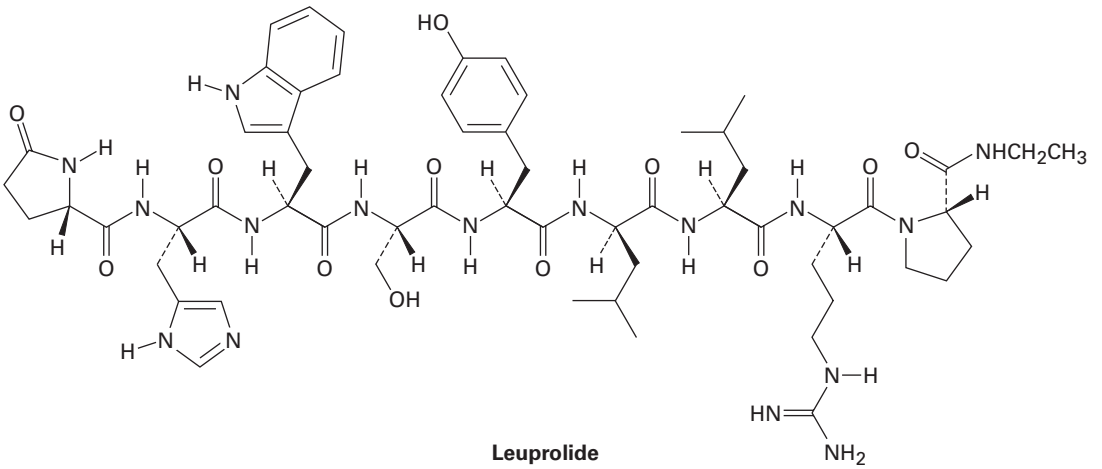
26.42 Which amide bonds in the following polypeptide are cleaved by trypsin? By chymotrypsin?

Phe-Leu-Met-Lys-Tyr-Asp-Gly-Gly-Arg-Val-Ile-Pro-Tyr

26.43 What kinds of reactions do the following classes of enzymes catalyze?
(a) Hydrolases **(b)** Lyases **(c)** Transferases

26.44 Which of the following amino acids are more likely to be found on the outside of a globular protein, and which on the inside? Explain.
(a) Valine **(b)** Aspartic acid **(c)** Phenylalanine **(d)** Lysine

26.45 Leuprolide is a synthetic nonapeptide used to treat both endometriosis in women and prostate cancer in men.



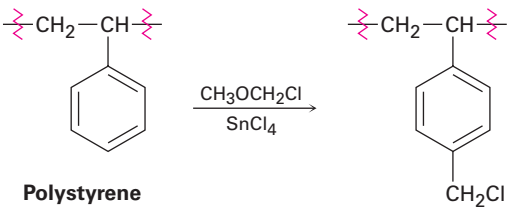
Leuprolide

(a) Both C-terminal and N-terminal amino acids in leuprolide have been structurally modified. Identify the modifications.
(b) One of the nine amino acids in leuprolide has D stereochemistry rather than the usual L. Which one?
(c) Write the structure of leuprolide using both one- and three-letter abbreviations.
(d) What charge would you expect leuprolide to have at neutral pH?

General Problems

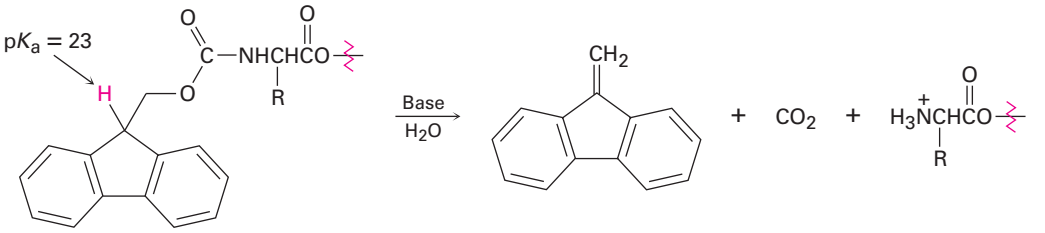
26.46 The α-helical parts of myoglobin and other proteins stop whenever a proline residue is encountered in the chain. Why is proline never present in a protein α-helix?

26.47 The chloromethylated polystyrene resin used for Merrifield solid-phase peptide synthesis is prepared by treatment of polystyrene with chloromethyl methyl ether and a Lewis acid catalyst. Propose a mechanism for the reaction.



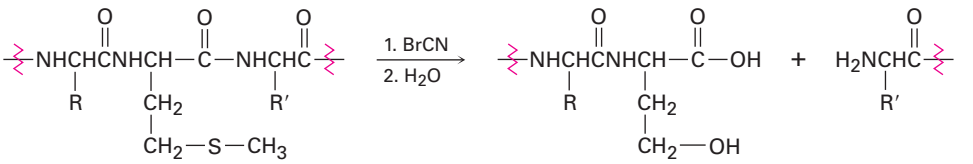
Polystyrene

26.48 An Fmoc protecting group can be removed from an amino acid by treatment with the amine base piperidine. Propose a mechanism.



Fmoc-protected amino acid

26.49 Proteins can be cleaved specifically at the amide bond on the carboxyl side of methionine residues by reaction with cyanogen bromide, BrC≡N.



The reaction occurs in several steps:

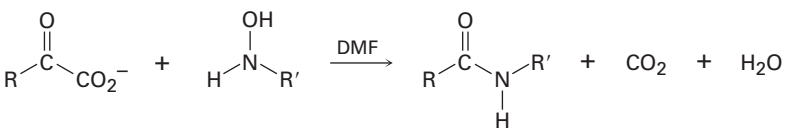
(a) The first step is a nucleophilic substitution reaction of the sulfur on the methionine side chain with BrCN to give a cyanosulfonium ion, $[R_2SCN]^+$. Show the structure of the product, and propose a mechanism for the reaction.

(b) The second step is an internal S_N2 reaction, with the carbonyl oxygen of the methionine residue displacing the positively charged sulfur leaving group and forming a five-membered ring product. Show the structure of the product and the mechanism of its formation.

(c) The third step is a hydrolysis reaction to split the peptide chain. The carboxyl group of the former methionine residue is now part of a lactone (cyclic ester) ring. Show the structure of the lactone product and the mechanism of its formation.

(d) The final step is a hydrolysis of the lactone to give the product shown. Show the mechanism of the reaction.

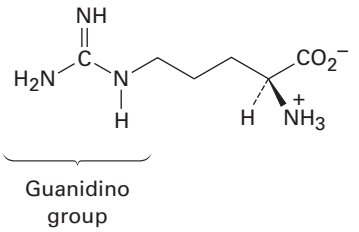
26.50 A clever new method of peptide synthesis involves formation of an amide bond by reaction of an α-keto acid with an *N*-alkylhydroxylamine:



An α-keto acid **A hydroxylamine** **An amide**

The reaction is thought to occur by nucleophilic addition of the *N*-alkylhydroxylamine to the keto acid as if forming an oxime (Section 19.8), followed by decarboxylation and elimination of water. Show the mechanism.

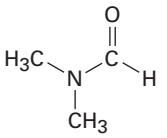
26.51 Arginine, the most basic of the 20 common amino acids, contains a *guanidino* functional group in its side chain. Explain, using resonance structures to show how the protonated guanidino group is stabilized.



Arginine

26.52 Cytochrome *c* is an enzyme found in the cells of all aerobic organisms. Elemental analysis of cytochrome *c* shows that it contains 0.43% iron. What is the minimum molecular weight of this enzyme?

26.53 Evidence for restricted rotation around amide CO–N bonds comes from NMR studies. At room temperature, the ^{1}H NMR spectrum of *N*,*N*-dimethylformamide shows three peaks: 2.9 δ (singlet, 3 H), 3.0 δ (singlet, 3 H), 8.0 δ (singlet, 1 H). As the temperature is raised, however, the two singlets at 2.9 δ and 3.0 δ slowly merge. At 180 °C, the ^{1}H NMR spectrum shows only two peaks: 2.95 δ (singlet, 6 H) and 8.0 δ (singlet, 1 H). Explain this temperature-dependent behavior.



***N,N*-Dimethylformamide**

26.54 Propose a structure for an octapeptide that shows the composition Asp, Gly_2, Leu, Phe, Pro_2, Val on amino acid analysis. Edman analysis shows a glycine N-terminal group, and leucine is the C-terminal group. Acidic hydrolysis gives the following fragments:

Val-Pro-Leu, Gly, Gly-Asp-Phe-Pro, Phe-Pro-Val

26.55 Look up the structure of human insulin (page 1057), and indicate where in each chain the molecule is cleaved by trypsin and chymotrypsin.

26.56 What is the structure of a nonapeptide that gives the following fragments when cleaved?

Trypsin cleavage: Val-Val-Pro-Tyr-Leu-Arg, Ser-Ile-Arg

Chymotrypsin cleavage: Leu-Arg, Ser-Ile-Arg-Val-Val-Pro-Tyr

26.57 Oxytocin, a nonapeptide hormone secreted by the pituitary gland, functions by stimulating uterine contraction and lactation during childbirth. Its sequence was determined from the following evidence:

1. Oxytocin is a cyclic compound containing a disulfide bridge between two cysteine residues.
2. When the disulfide bridge is reduced, oxytocin has the constitution Asn, Cys_2, Gln, Gly, Ile, Leu, Pro, Tyr.
3. Partial hydrolysis of reduced oxytocin yields seven fragments: Asp-Cys, Ile-Glu, Cys-Tyr, Leu-Gly, Tyr-Ile-Glu, Glu-Asp-Cys, Cys-Pro-Leu.
4. Gly is the C-terminal group.
5. Both Glu and Asp are present as their side-chain amides (Gln and Asn) rather than as free side-chain acids.

What is the amino acid sequence of reduced oxytocin? What is the structure of oxytocin itself?

26.58 *Aspartame,* a nonnutritive sweetener marketed under such trade names as Equal, NutraSweet, and Canderel, is the methyl ester of a simple dipeptide, Asp-Phe-OCH_3.

(a) Draw the structure of aspartame.

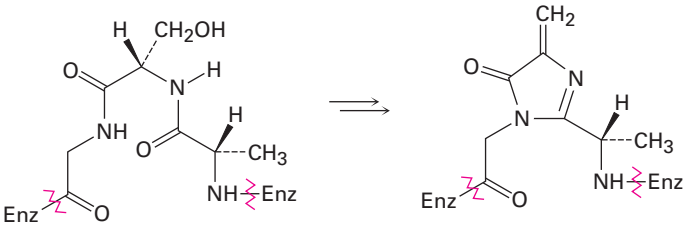
(b) The isoelectric point of aspartame is 5.9. Draw the principal structure present in aqueous solution at this pH.

(c) Draw the principal form of aspartame present at physiological pH = 7.3.

26.59 Refer to Figure 26.4 on page 1060 and propose a mechanism for the final step in the Edman degradation—the acid-catalyzed rearrangement of the ATZ derivative to the PTH derivative.

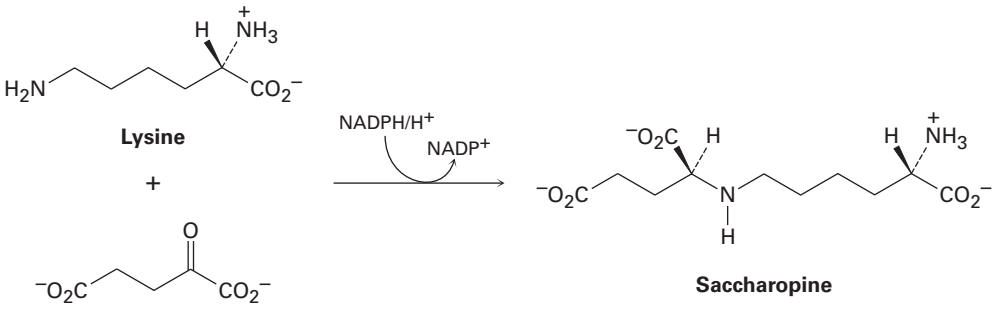
26.60 Amino acids are metabolized by a transamination reaction in which the $-NH_2$ group of the amino acid changes places with the keto group of an α-keto acid. The products are a new amino acid and a new α-keto acid. Show the product from transamination of isoleucine.

26.61 The first step in the biological degradation of histidine is formation of a 4-methylideneimidazol-5-one (MIO) by cyclization of a segment of the peptide chain in the histidine ammonia lyase enzyme. Propose a mechanism.



4-Methylidene-5-imidazolone (MIO)

26.62 The first step in the biological degradation of lysine is reductive amination with α-ketoglutarate to give saccharopine. Nicotinamide adenine dinucleotide phosphate (NADPH), a relative of NADH, is the reducing agent. Show the mechanism.



Lysine

+

α-Ketoglutarate

Saccharopine

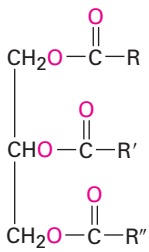
27

Soap bubbles, so common yet so beautiful, are made from animal fat, a lipid.

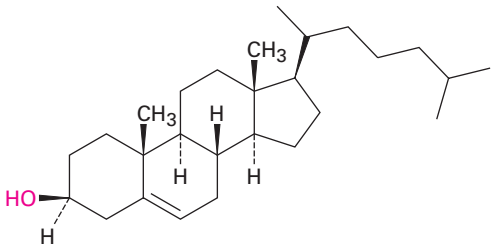
Biomolecules: Lipids

Lipids are naturally occurring organic molecules that have limited solubility in water and can be isolated from organisms by extraction with nonpolar organic solvents. Fats, oils, waxes, many vitamins and hormones, and most nonprotein cell-membrane components are examples. Note that this definition differs from the sort used for carbohydrates and proteins in that lipids are defined by a physical property (solubility) rather than by structure. Of the many kinds of lipids, we'll be concerned in this chapter only with a few: triacylglycerols, eicosanoids, terpenoids, and steroids.

Lipids are classified into two broad types: those like fats and waxes, which contain ester linkages and can be hydrolyzed, and those like cholesterol and other steroids, which don't have ester linkages and can't be hydrolyzed.



Animal fat—a triester (R, R′, R″ = C_{11}–C_{19} chains)



Cholesterol

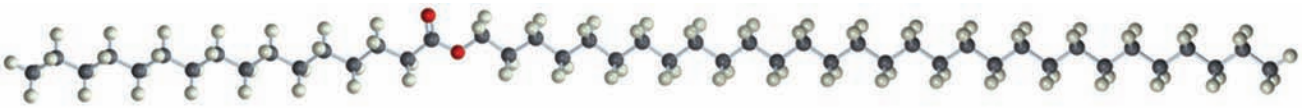
Why This Chapter? We've now covered two of the four major classes of biomolecules—proteins and carbohydrates—and have two remaining. We'll cover lipids, the largest and most diverse class of biomolecules, in this chapter, looking both at their structure and function and at their metabolism.

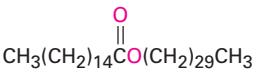
27.1 Waxes, Fats, and Oils

Waxes are mixtures of esters of long-chain carboxylic acids with long-chain alcohols. The carboxylic acid usually has an even number of carbons from 16 through 36, while the alcohol has an even number of carbons from 24 through 36. One of the major components of beeswax, for instance, is triacontyl hexadecanoate, the ester of the C_{30} alcohol 1-triacontanol and the

OWL Sign in to OWL for Organic Chemistry at **www.cengage.com/owl** to view tutorials and simulations, develop problem-solving skills, and complete online homework assigned by your professor.

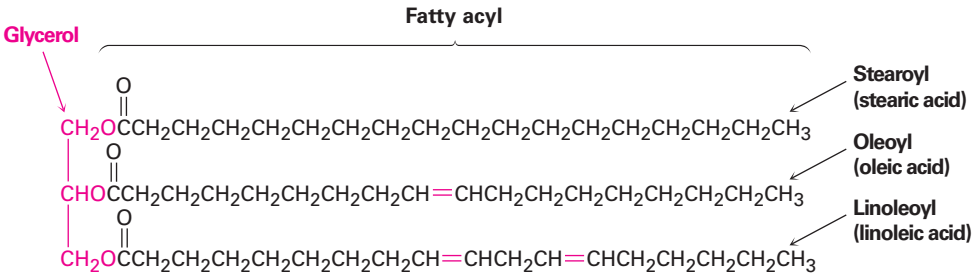
C_{16} acid hexadecanoic acid. The waxy protective coatings on most fruits, berries, leaves, and animal furs have similar structures.





Triacontyl hexadecanoate (from beeswax)

Animal fats and vegetable oils are the most widely occurring lipids. Although they appear different—animal fats like butter and lard are solids, whereas vegetable oils like corn and peanut oil are liquid—their structures are closely related. Chemically, fats and oils are *triglycerides,* or **triacylglycerols**—triesters of glycerol with three long-chain carboxylic acids called **fatty acids**. Animals use fats for long-term energy storage because they are much less highly oxidized than carbohydrates and provide about six times as much energy as an equal weight of stored, hydrated glycogen.



A triacylglycerol

Hydrolysis of a fat or oil with aqueous NaOH yields glycerol and three fatty acids. The fatty acids are generally unbranched and contain an even number of carbon atoms between 12 and 20. If double bonds are present, they have largely, although not entirely, *Z,* or cis, geometry. The three fatty acids of a specific triacylglycerol molecule need not be the same, and the fat or oil from a given source is likely to be a complex mixture of many different triacylglycerols. Table 27.1 lists some of the commonly occurring fatty acids, and Table 27.2 lists the approximate composition of fats and oils from different sources.

More than 100 different fatty acids are known, and about 40 occur widely. Palmitic acid (C_{16}) and stearic acid (C_{18}) are the most abundant saturated fatty acids; oleic and linoleic acids (both C_{18}) are the most abundant unsaturated ones. Oleic acid is monounsaturated because it has only one double bond, whereas linoleic, linolenic, and arachidonic acids are **polyunsaturated fatty acids** because they have more than one double bond. Linoleic and linolenic acids occur in cream and are essential in the human diet; infants grow poorly and develop skin lesions if fed a diet of nonfat milk for prolonged periods. Linolenic acid, in particular, is an example of an *omega-3* fatty acid, which has been found to lower blood triglyceride levels and reduce the risk of heart attack. The

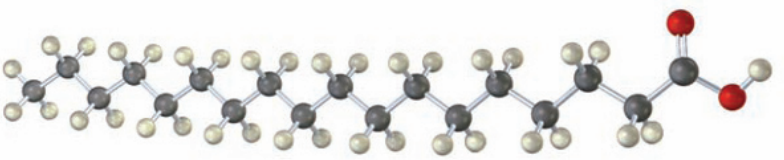
Table 27.1 Structures of Some Common Fatty Acids

Name	No. of carbons	Melting point (°C)	Structure
Saturated			
Lauric	12	43.2	$CH_3(CH_2)_{10}CO_2H$
Myristic	14	53.9	$CH_3(CH_2)_{12}CO_2H$
Palmitic	16	63.1	$CH_3(CH_2)_{14}CO_2H$
Stearic	18	68.8	$CH_3(CH_2)_{16}CO_2H$
Arachidic	20	76.5	$CH_3(CH_2)_{18}CO_2H$
Unsaturated			
Palmitoleic	16	−0.1	(Z)-$CH_3(CH_2)_5CH{=}CH(CH_2)_7CO_2H$
Oleic	18	13.4	(Z)-$CH_3(CH_2)_7CH{=}CH(CH_2)_7CO_2H$
Linoleic	18	−12	(Z,Z)-$CH_3(CH_2)_4(CH{=}CHCH_2)_2(CH_2)_6CO_2H$
Linolenic	18	−11	(all Z)-$CH_3CH_2(CH{=}CHCH_2)_3(CH_2)_6CO_2H$
Arachidonic	20	−49.5	(all Z)-$CH_3(CH_2)_4(CH{=}CHCH_2)_4CH_2CH_2CO_2H$

Table 27.2 Composition of Some Fats and Oils

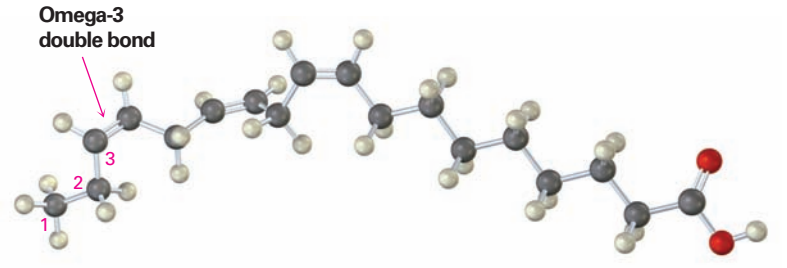
	Saturated fatty acids (%)				Unsaturated fatty acids (%)	
Source	C_{12} lauric	C_{14} myristic	C_{16} palmitic	C_{18} stearic	C_{18} oleic	C_{18} linoleic
Animal fat						
Lard	—	1	25	15	50	6
Butter	2	10	25	10	25	5
Human fat	1	3	25	8	46	10
Whale blubber	—	8	12	3	35	10
Vegetable oil						
Coconut	50	18	8	2	6	1
Corn	—	1	10	4	35	45
Olive	—	1	5	5	80	7
Peanut	—	—	7	5	60	20

name omega-3 means that there is a double bond three carbons in from the noncarboxyl end of the chain.



$CH_3CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2CH_2\overset{O}{\overset{\|}{C}}OH$

Stearic acid

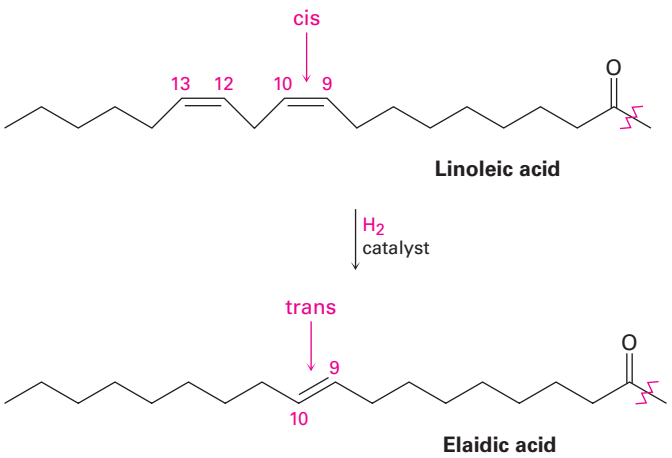


$$CH_3CH_2CH{=}CHCH_2CH{=}CHCH_2CH{=}CHCH_2CH_2CH_2CH_2CH_2CH_2\overset{\overset{O}{\|}}{C}OH$$

Linolenic acid, an omega-3 polyunsaturated fatty acid

The data in Table 27.1 show that unsaturated fatty acids generally have lower melting points than their saturated counterparts, a trend that is also true for triacylglycerols. Since vegetable oils generally have a higher proportion of unsaturated to saturated fatty acids than animal fats (Table 27.2), they have lower melting points. The difference is a consequence of structure. Saturated fats have a uniform shape that allows them to pack together efficiently in a crystal lattice. In unsaturated vegetable oils, however, the C=C bonds introduce bends and kinks into the hydrocarbon chains, making crystal formation more difficult. The more double bonds there are, the harder it is for the molecules to crystallize and the lower the melting point of the oil.

The C=C bonds in vegetable oils can be reduced by catalytic hydrogenation, typically carried out at high temperature using a nickel catalyst, to produce saturated solid or semisolid fats. Margarine and shortening are produced by hydrogenating soybean, peanut, or cottonseed oil until the proper consistency is obtained. Unfortunately, the hydrogenation reaction is accompanied by some cis–trans isomerization of the double bonds that remain, producing fats with about 10% to 15% trans unsaturated fatty acids. Dietary intake of trans fatty acids increases cholesterol levels in the blood, thereby increasing the risk of heart problems. The conversion of linoleic acid into elaidic acid is an example.

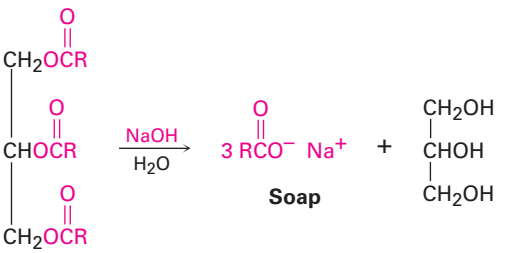


Problem 27.1
Carnauba wax, used in floor and furniture polishes, contains an ester of a C_{32} straight-chain alcohol with a C_{20} straight-chain carboxylic acid. Draw its structure.

Problem 27.2
Draw structures of glyceryl tripalmitate and glyceryl trioleate. Which would you expect to have a higher melting point?

27.2 Soap

Soap has been known since at least 600 BC, when the Phoenicians prepared a curdy material by boiling goat fat with extracts of wood ash. The cleansing properties of soap weren't generally recognized, however, and the use of soap did not become widespread until the 18th century. Chemically, soap is a mixture of the sodium or potassium salts of the long-chain fatty acids produced by hydrolysis *(saponification)* of animal fat with alkali. Wood ash was used as a source of alkali until the early 1800s, when the development of the LeBlanc process for making Na_2CO_3 by heating sodium sulfate with limestone became available.



A fat
(R = C_{11}–C_{19} aliphatic chains)

Glycerol

Crude soap curds contain glycerol and excess alkali as well as soap but can be purified by boiling with water and adding NaCl or KCl to precipitate the pure carboxylate salts. The smooth soap that precipitates is dried, perfumed, and pressed into bars for household use. Dyes are added to make colored soaps, antiseptics are added for medicated soaps, pumice is added for scouring soaps, and air is blown in for soaps that float. Regardless of these extra treatments and regardless of price, though, all soaps are basically the same.

Soaps act as cleansers because the two ends of a soap molecule are so different. The carboxylate end of the long-chain molecule is ionic and therefore hydrophilic **(Section 2.12)**, or attracted to water. The long hydrocarbon portion of the molecule, however, is nonpolar and hydrophobic, avoiding water and therefore more soluble in oils. The net effect of these two opposing tendencies is that soaps are attracted to both oils and water and are therefore useful as cleansers.

When soaps are dispersed in water, the long hydrocarbon tails cluster together on the inside of a tangled, hydrophobic ball, while the ionic heads on the surface of the cluster stick out into the water layer. These spherical clusters, called **micelles**, are shown schematically in **Figure 27.1**. Grease and oil droplets are solubilized in water when they are coated by the nonpolar, hydrophobic tails of soap molecules in the center of micelles. Once solubilized, the grease and dirt can be rinsed away.

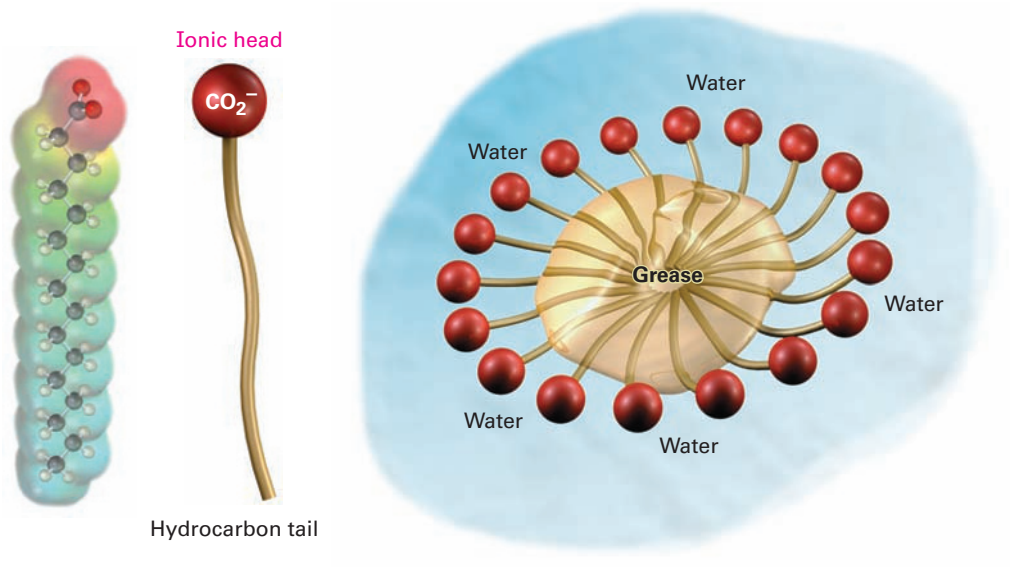
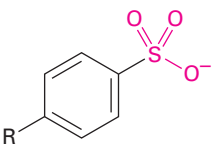


Figure 27.1 A soap micelle solubilizing a grease particle in water. An electrostatic potential map of a fatty acid carboxylate shows how the negative charge is located in the head group.

As useful as they are, soaps also have some drawbacks. In hard water, which contains metal ions, soluble sodium carboxylates are converted into insoluble magnesium and calcium salts, leaving the familiar ring of scum around bathtubs and the gray tinge on white clothes. Chemists have circumvented these problems by synthesizing a class of synthetic detergents based on salts of long-chain alkylbenzenesulfonic acids. The principle of synthetic detergents is the same as that of soaps: the alkylbenzene end of the molecule is attracted to grease, while the anionic sulfonate end is attracted to water. Unlike soaps, though, sulfonate detergents don't form insoluble metal salts in hard water and don't leave an unpleasant scum.



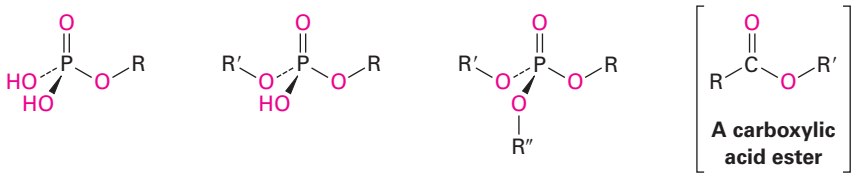
A synthetic detergent
(R = a mixture of C_{12} chains)

Problem 27.3
Draw the structure of magnesium oleate, a component of bathtub scum.

Problem 27.4
Write the saponification reaction of glyceryl dioleate monopalmitate with aqueous NaOH.

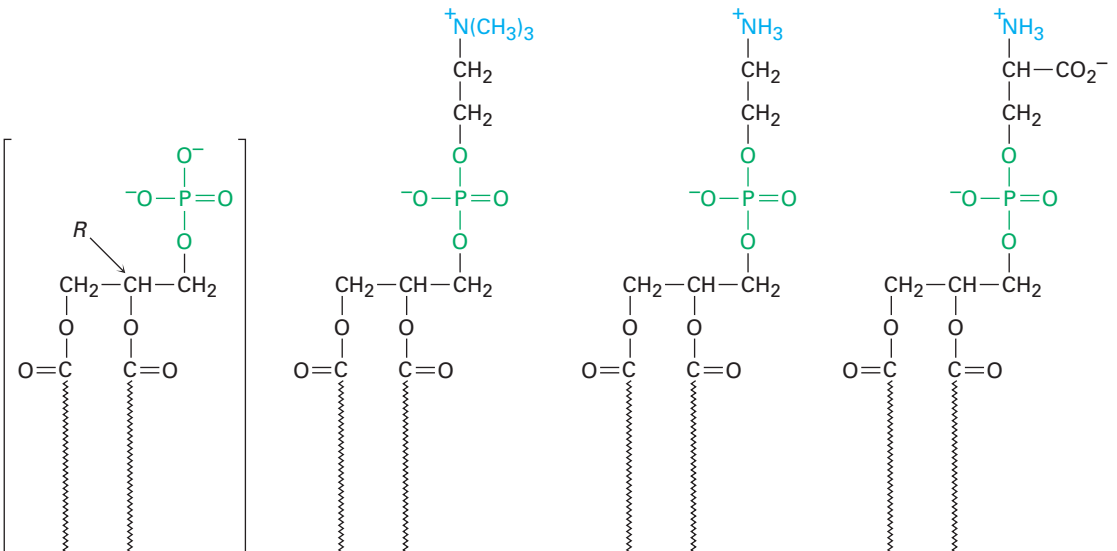
27.3 Phospholipids

Just as waxes, fats, and oils are esters of carboxylic acids, **phospholipids** are esters of phosphoric acid, H_3PO_4.



A phosphoric acid monoester | **A phosphoric acid diester** | **A phosphoric acid triester**

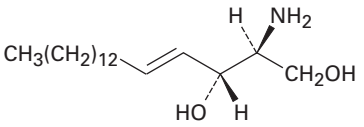
Phospholipids are of two general kinds: *glycerophospholipids* and *sphingomyelins*. Glycerophospholipids are based on phosphatidic acid, which contains a glycerol backbone linked by ester bonds to two fatty acids and one phosphoric acid. Although the fatty-acid residues can be any of the C_{12}–C_{20} units typically present in fats, the acyl group at C1 is usually saturated and the one at C2 is usually unsaturated. The phosphate group at C3 is also bonded to an amino alcohol such as choline $[HOCH_2CH_2N(CH_3)_3]^+$, ethanolamine ($HOCH_2CH_2NH_2$), or serine $[HOCH_2CH(NH_2)CO_2H]$. The compounds are chiral and have an L, or *R,* configuration at C2.



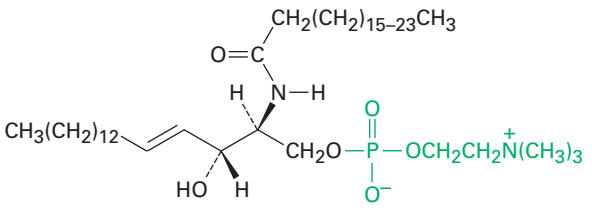
Phosphatidic acid | **Phosphatidylcholine** | **Phosphatidyl-ethanolamine** | **Phosphatidylserine**

Sphingomyelins are the second major group of phospholipids. These compounds have sphingosine or a related dihydroxyamine as their backbone and

are particularly abundant in brain and nerve tissue, where they are a major constituent of the coating around nerve fibers.



Sphingosine



A sphingomyelin

Phospholipids are found widely in both plant and animal tissues and make up approximately 50% to 60% of cell membranes. Because they are like soaps in having a long, nonpolar hydrocarbon tail bound to a polar ionic head, phospholipids in the cell membrane organize into a **lipid bilayer** about 5.0 nm (50 Å) thick. As shown in **Figure 27.2**, the nonpolar tails aggregate in the center of the bilayer in much the same way that soap tails aggregate in the center of a micelle. This bilayer serves as an effective barrier to the passage of water, ions, and other components into and out of cells.

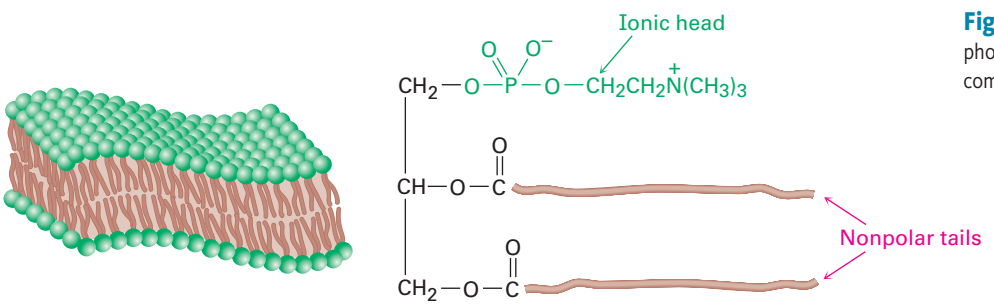


Figure 27.2 Aggregation of glycerophospholipids into the lipid bilayer that composes cell membranes.

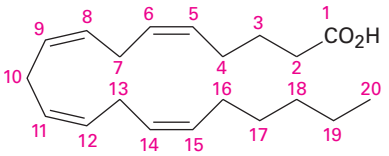
27.4 Prostaglandins and Other Eicosanoids

The **prostaglandins** are a group of C_{20} lipids that contain a five-membered ring with two long side chains. First isolated in the 1930s by Ulf von Euler at the Karolinska Institute in Sweden, much of the structural and chemical work on the prostaglandins was carried out by Sune Bergström and Bengt Samuelsson. All three received Nobel Prizes for their work. The name *prostaglandin* derives from the fact that the compounds were first isolated from sheep prostate glands, but they have subsequently been shown to be present in small amounts in all body tissues and fluids.

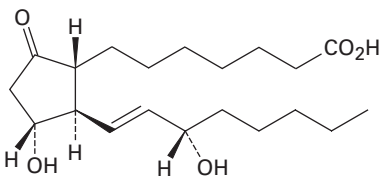
The several dozen known prostaglandins have an extraordinarily wide range of biological effects. Among their many properties, they can lower blood pressure, affect blood platelet aggregation during clotting, lower gastric secretions,

control inflammation, affect kidney function, affect reproductive systems, and stimulate uterine contractions during childbirth.

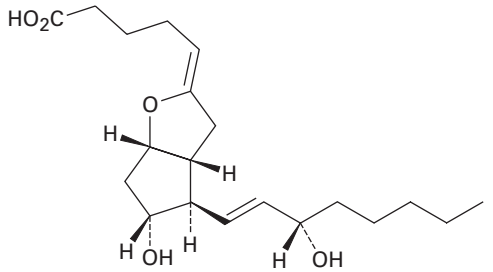
Prostaglandins, together with related compounds called thromboxanes and leukotrienes, make up a class of compounds called **eicosanoids** because they are derived biologically from 5,8,11,14-eicosatetraenoic acid, or arachidonic acid **(Figure 27.3)**. Prostaglandins (PG) have a cyclopentane ring with two long side chains; thromboxanes (TX) have a six-membered, oxygen-containing ring; and leukotrienes (LT) are acyclic.



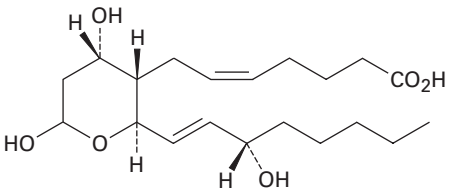
Arachidonic acid



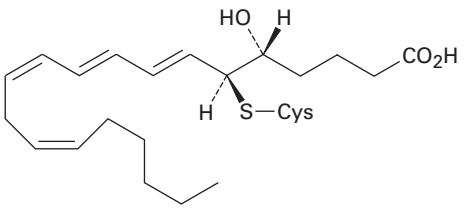
Prostaglandin E_1 (PGE_1)



Prostaglandin I_2 (PGI_2) (prostacyclin)



Thromboxane B_2 (TXB_2)



Leukotriene E_4 (LTE_4)

Figure 27.3 Structures of some representative eicosanoids. All are derived biologically from arachidonic acid.

Eicosanoids are named based on their ring system (PG, TX, or LT), substitution pattern, and number of double bonds. The various substitution patterns on the ring are indicated by letter as in **Figure 27.4**, and the number of double bonds is indicated by a subscript. Thus, PGE_1 is a prostaglandin with the "E" substitution pattern and one double bond. The numbering of the atoms in the various eicosanoids is the same as in arachidonic acid, starting with the $-CO_2H$ carbon as C1, continuing around the ring, and ending with the $-CH_3$ carbon at the other end of the chain as C20.

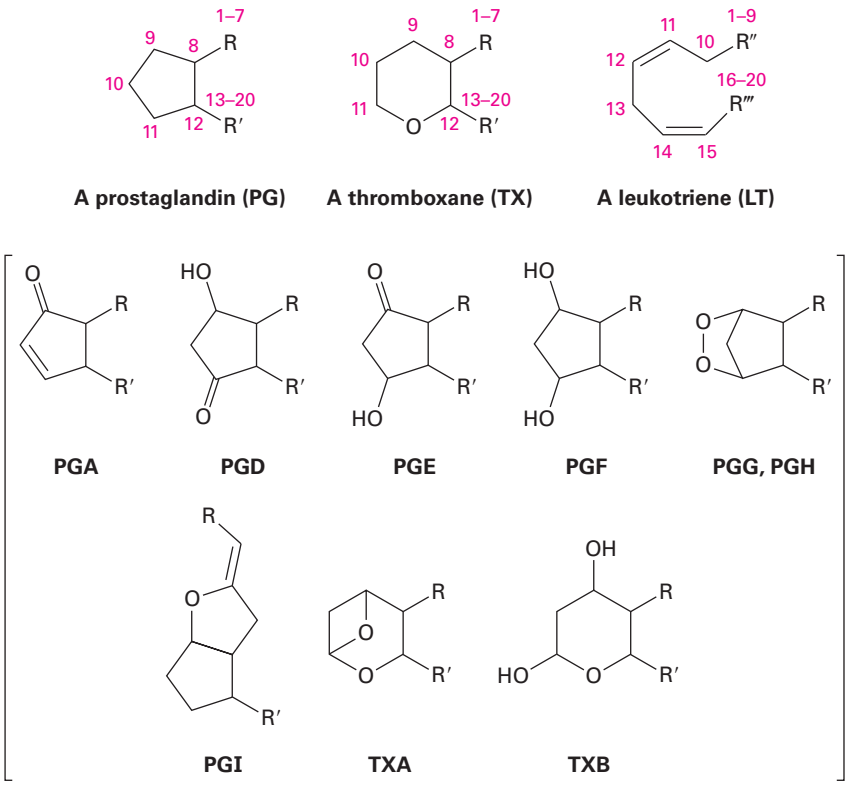


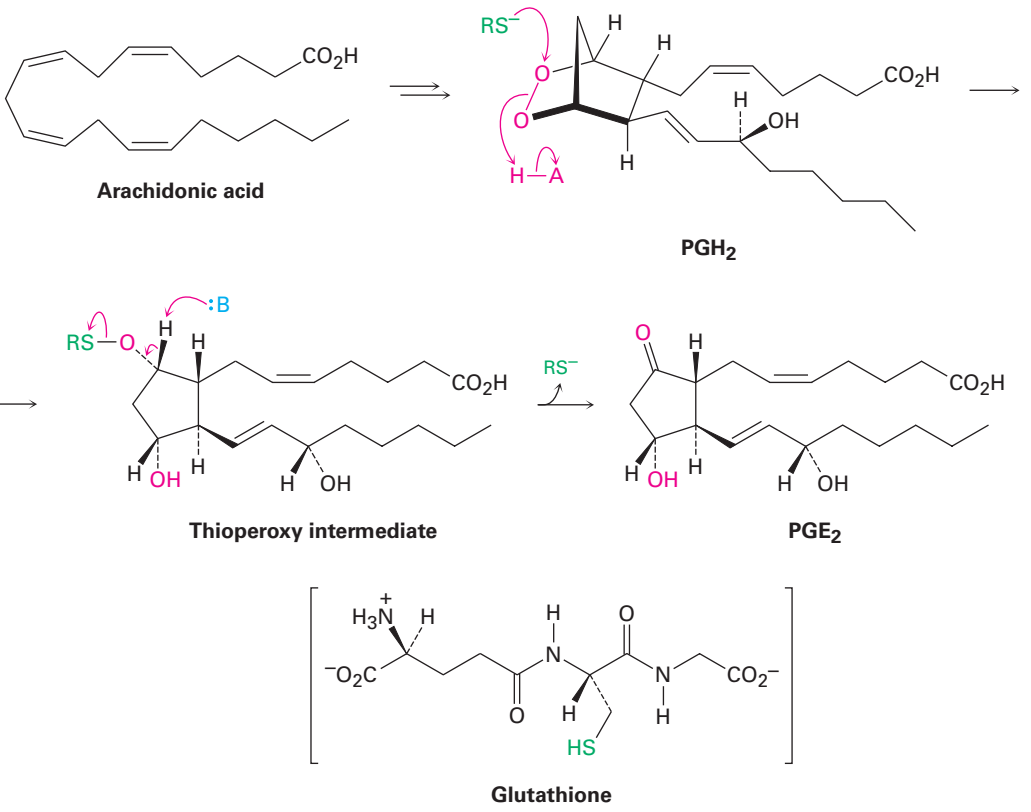
Figure 27.4 The nomenclature system for eicosanoids.

Eicosanoid biosynthesis begins with the conversion of arachidonic acid to PGH_2, catalyzed by the multifunctional PGH synthase (PGHS), also called cyclooxygenase (COX). There are two distinct enzymes, PGHS-1 and PGHS-2 (or COX-1 and COX-2), both of which accomplish the same reaction but appear to function independently. COX-1 carries out the normal physiological production of prostaglandins, and COX-2 produces additional prostaglandin in response to arthritis or other inflammatory conditions. Vioxx, Celebrex, Bextra, and several other drugs selectively inhibit the COX-2 enzyme but also appear to cause potentially serious heart problems in weakened patients. (See the Chapter 15 *A Deeper Look*.)

PGHS accomplishes two transformations, an initial reaction of arachidonic acid with O_2 to yield PGG_2 and a subsequent reduction of the hydroperoxide group (−OOH) to the alcohol PGH_2. The sequence of steps involved in these transformation was shown in Figure 8.10 on page 295.

Further processing of PGH_2 then leads to other eicosanoids. PGE_2, for instance, arises by an isomerization of PGH_2 catalyzed by PGE synthase (PGES). The coenzyme glutathione is needed for enzyme activity, although it is not chemically changed during the isomerization and its role is not fully understood. One possibility is that the glutathione thiolate anion breaks the O−O bond in PGH_2 by an S_N2-like attack on one of the oxygen atoms, giving a thioperoxy intermediate (R—S—O—R′) that eliminates glutathione to give the ketone **(Figure 27.5)**.

Figure 27.5 Mechanism of the conversion of PGH_2 into PGE_2.



Problem 27.5

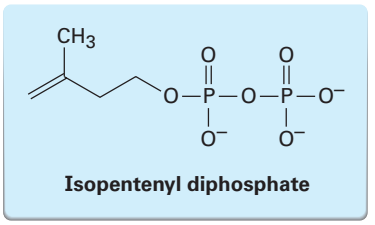
Assign *R* or *S* configuration to each chirality center in prostaglandin E_2 (Figure 27.5), the most abundant and biologically potent of mammalian prostaglandins.

27.5 Terpenoids

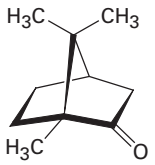
We saw in the Chapter 8 *A Deeper Look* that **terpenoids** are a vast and diverse group of lipids found in all living organisms. Despite their apparent structural differences, all terpenoids contain a multiple of five carbons and are derived biosynthetically from the five-carbon precursor isopentenyl diphosphate **(Figure 27.6)**. Although formally a *terpenoid* contains oxygen, while a hydrocarbon is called a *terpene,* we'll use the term *terpenoid* to refer to both for simplicity.

You might recall from Chapter 8 that terpenoids are classified according to the number of five-carbon multiples they contain. Monoterpenoids contain 10 carbons and are derived from two isopentenyl diphosphates, sesquiterpenoids contain 15 carbons and are derived from three isopentenyl diphosphates, diterpenoids contain 20 carbons and are derived from four isopentenyl diphosphates, and so on, up to triterpenoids (C_{30}) and tetraterpenoids (C_{40}). Lanosterol, for example, is a triterpenoid from which steroid hormones are made, and β-carotene is a tetraterpenoid that serves as a dietary source of vitamin A (Figure 27.6).

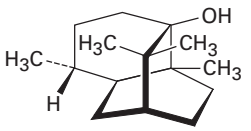
The terpenoid precursor isopentenyl diphosphate, formerly called isopentenyl pyrophosphate and thus abbreviated IPP, is biosynthesized by two different pathways, depending on the organism and the structure of the final product. In animals and higher plants, sesquiterpenoids and triterpenoids arise primarily from the *mevalonate* pathway, whereas monoterpenoids, diterpenoids, and



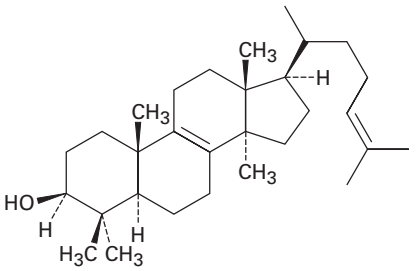
Isopentenyl diphosphate



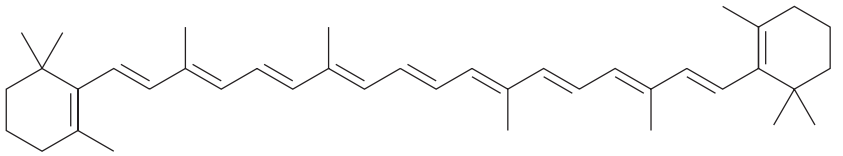
Camphor
(a monoterpenoid—C_{10})



Patchouli alcohol
(a sesquiterpenoid—C_{15})



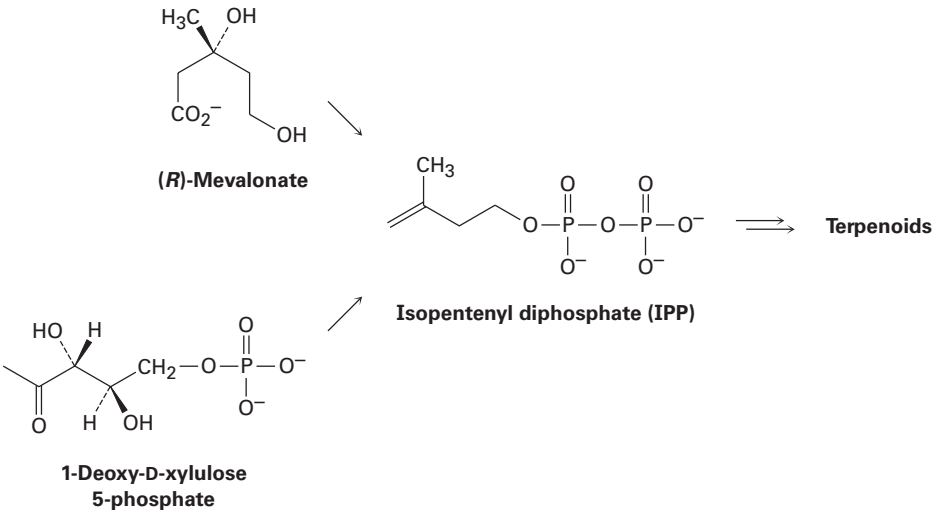
Lanosterol
(a triterpenoid—C_{30})



***β*-Carotene**
(a tetraterpenoid—C_{40})

Figure 27.6 Structures of some representative terpenoids.

tetraterpenoids are biosynthesized by the *1-deoxyxylulose 5-phosphate (DXP)* pathway, also called the methylerithritol phosphate, or MEP, pathway. In bacteria, both pathways are used. We'll look only at the mevalonate pathway, which is more common and better understood at present.



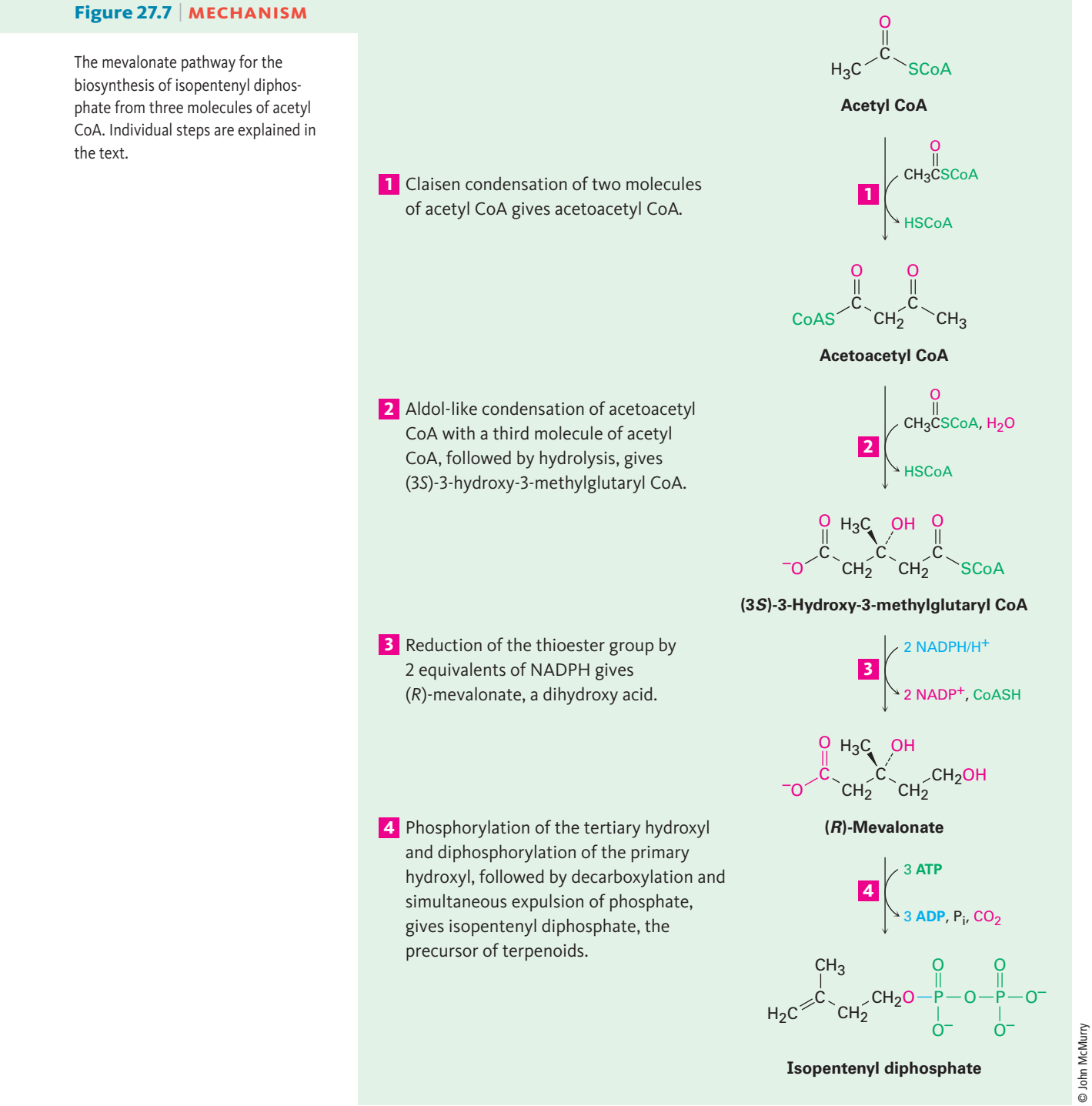
The Mevalonate Pathway to Isopentenyl Diphosphate

As shown in **Figure 27.7**, the mevalonate pathway begins with the conversion of acetate to acetyl CoA, followed by Claisen condensation to yield acetoacetyl CoA.

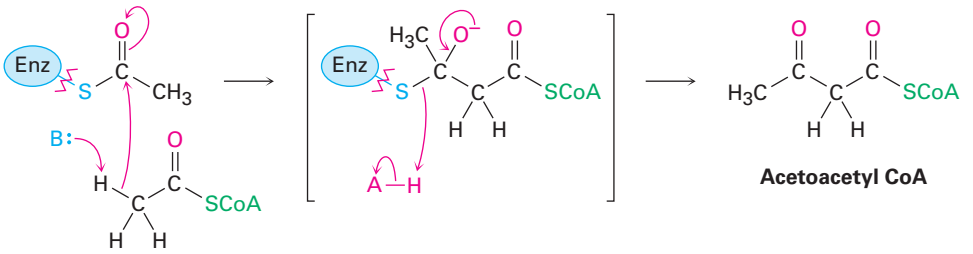
A second carbonyl condensation reaction with a third molecule of acetyl CoA, this one an aldol-like process, then yields the six-carbon compound 3-hydroxy-3-methylglutaryl CoA, which is reduced to give mevalonate. Phosphorylation, followed by loss of CO_2 and phosphate ion, completes the process.

Figure 27.7 | MECHANISM

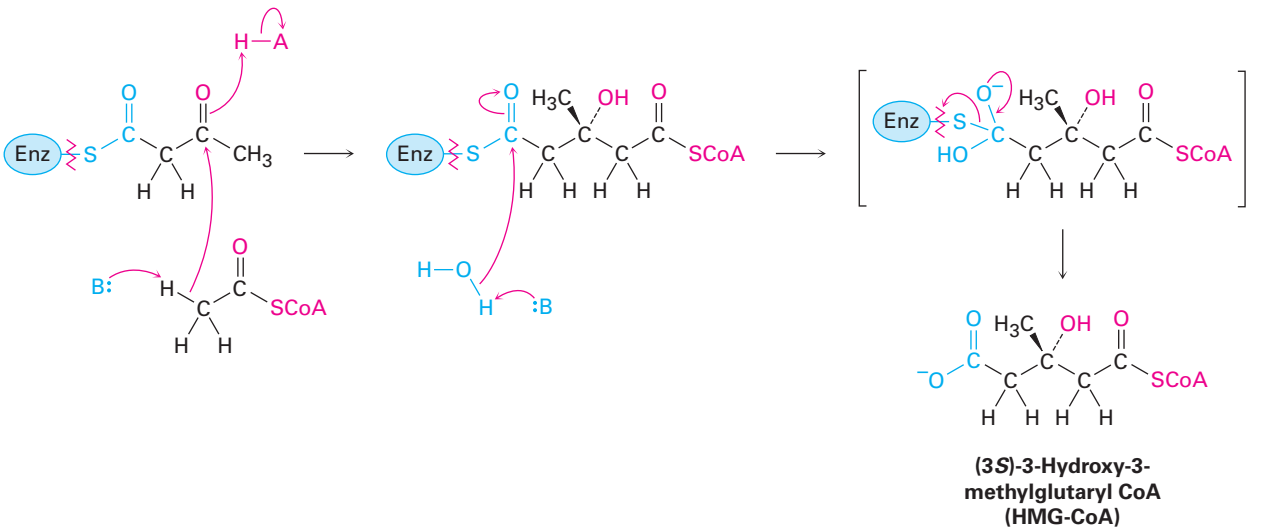
The mevalonate pathway for the biosynthesis of isopentenyl diphosphate from three molecules of acetyl CoA. Individual steps are explained in the text.



Step 1 of Figure 27.7: Claisen Condensation The first step in mevalonate biosynthesis is a Claisen condensation to yield acetoacetyl CoA, a reaction catalyzed by acetoacetyl-CoA acetyltransferase. An acetyl group is first bound to the enzyme by a nucleophilic acyl substitution reaction with a cysteine –SH group. Formation of an enolate ion from a second molecule of acetyl CoA, followed by Claisen condensation, then yields the product.

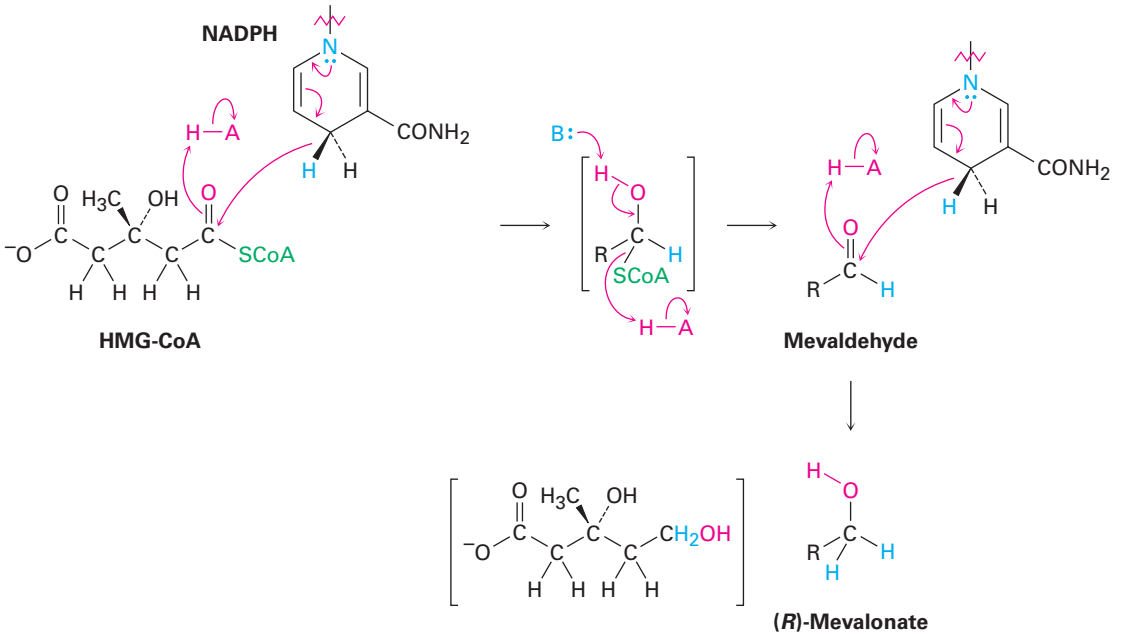


Step 2 of Figure 27.7: Aldol Condensation Acetoacetyl CoA next undergoes an aldol-like addition of an acetyl CoA enolate ion in a reaction catalyzed by 3-hydroxy-3-methylglutaryl-CoA synthase. The reaction occurs by initial binding of the substrate to a cysteine –SH group in the enzyme, followed by enolate-ion addition and subsequent hydrolysis to give (3*S*)-3-hydroxy-3-methylglutaryl CoA (HMG-CoA).

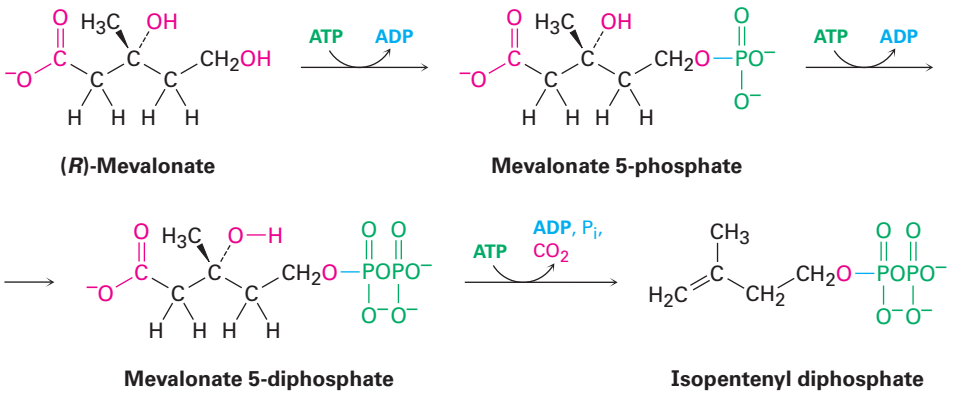


Step 3 of Figure 27.7: Reduction Reduction of HMG-CoA to give (*R*)-mevalonate is catalyzed by 3-hydroxy-3-methylglutaryl-CoA reductase and requires 2 equivalents of reduced nicotinamide adenine dinucleotide phosphate (NADPH), a close relative of NADH **(Section 19.12)**. The reaction occurs in two steps and proceeds through an aldehyde intermediate. The first step is a nucleophilic acyl substitution reaction involving hydride transfer from NADPH to the thioester carbonyl group of HMG-CoA. Following expulsion of HSCoA as

leaving group, the aldehyde intermediate undergoes a second hydride addition to give mevalonate.



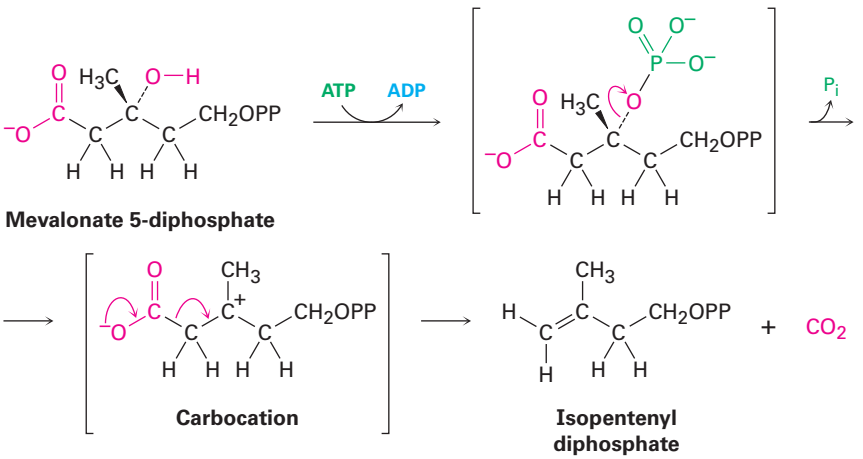
Step 4 of Figure 27.7: Phosphorylation and Decarboxylation Three additional reactions are needed to convert mevalonate to isopentenyl diphosphate. The first two are straightforward phosphorylations by ATP that occur through nucleophilic substitution reactions on the terminal phosphorus. Mevalonate is first converted to mevalonate 5-phosphate (phosphomevalonate) by reaction with ATP, and mevalonate 5-phosphate then reacts with a second ATP to give mevalonate 5-diphosphate (diphosphomevalonate). The third reaction results in phosphorylation of the tertiary hydroxyl group, followed by decarboxylation and loss of phosphate ion.



The final decarboxylation of mevalonate 5-diphosphate appears unusual because decarboxylations of acids do not typically occur except in β-keto acids

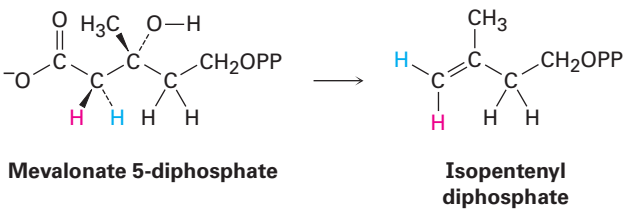
and malonic acids, in which the carboxylate group is two atoms away from an additional carbonyl group. As discussed in **Section 22.7**, the function of this second carbonyl group is to act as an electron acceptor and stabilize the charge resulting from loss of CO_2. In fact, though, the decarboxylation of a β-keto acid and the decarboxylation of mevalonate 5-diphosphate are closely related.

Catalyzed by mevalonate-5-diphosphate decarboxylase, the substrate is first phosphorylated on the free −OH group by reaction with ATP to give a tertiary phosphate, which undergoes spontaneous S_N1-like dissociation to give a tertiary carbocation. The positive charge then acts as an electron acceptor to facilitate decarboxylation in exactly the same way a β carbonyl group does, giving isopentenyl diphosphate. (In the following structures, the diphosphate group is abbreviated OPP.)



Problem 27.6

The conversion of mevalonate 5-phosphate to isopentenyl diphosphate occurs with the following result. Which hydrogen, *pro-R* or *pro-S*, ends up cis to the methyl group, and which ends up trans?



Conversion of Isopentenyl Diphosphate to Terpenoids

The conversion of isopentenyl diphosphate (IPP) to terpenoids begins with its isomerization to dimethylallyl diphosphate, abbreviated DMAPP and formerly called dimethylallyl pyrophosphate. These two C_5 building blocks

then combine to give the C_{10} unit geranyl diphosphate (GPP). The corresponding alcohol, geraniol, is itself a fragrant terpenoid that occurs in rose oil.

Further combination of GPP with another IPP gives the C_{15} unit farnesyl diphosphate (FPP), and so on, up to C_{25}. Terpenoids with more than 25 carbons—that is, triterpenoids (C_{30}) and tetraterpenoids (C_{40})—are synthesized by dimerization of C_{15} and C_{20} units, respectively **(Figure 27.8)**. Triterpenoids and steroids, in particular, arise from dimerization of farnesyl diphosphate to give squalene.

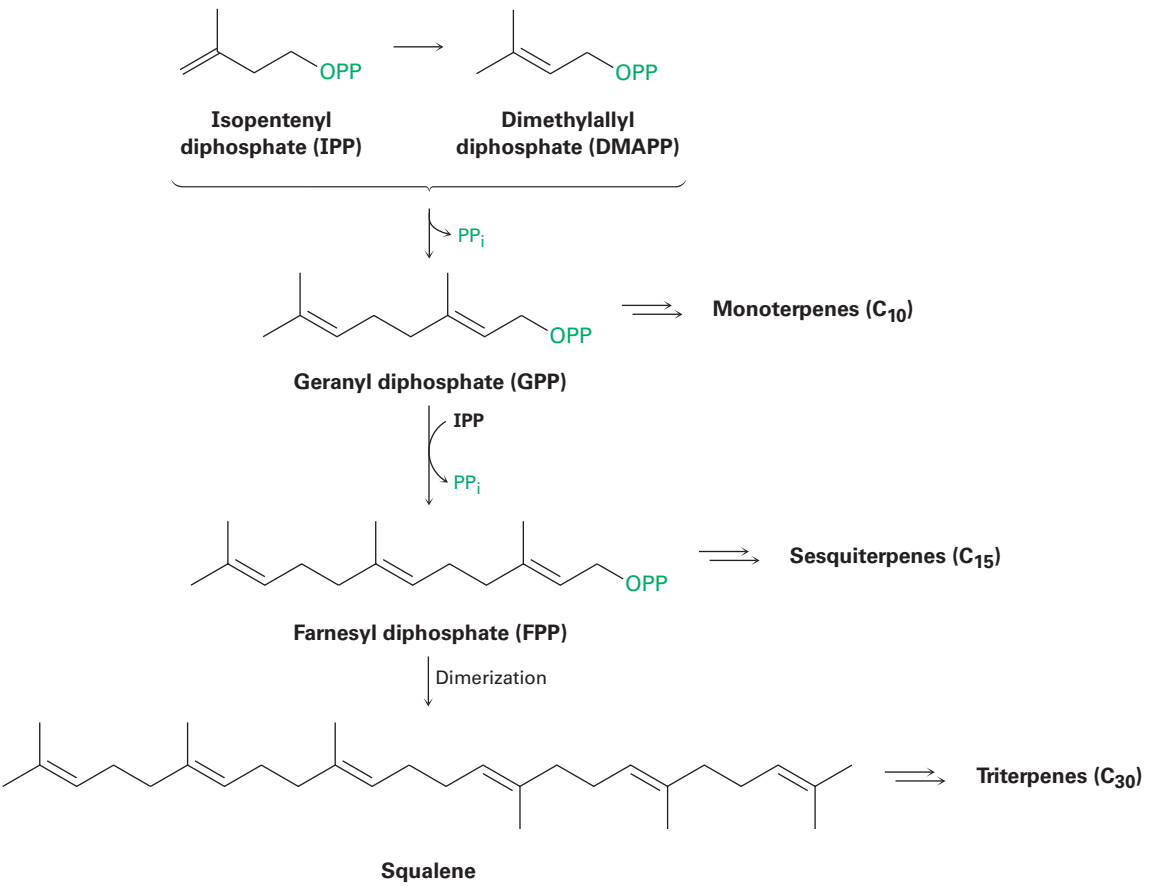
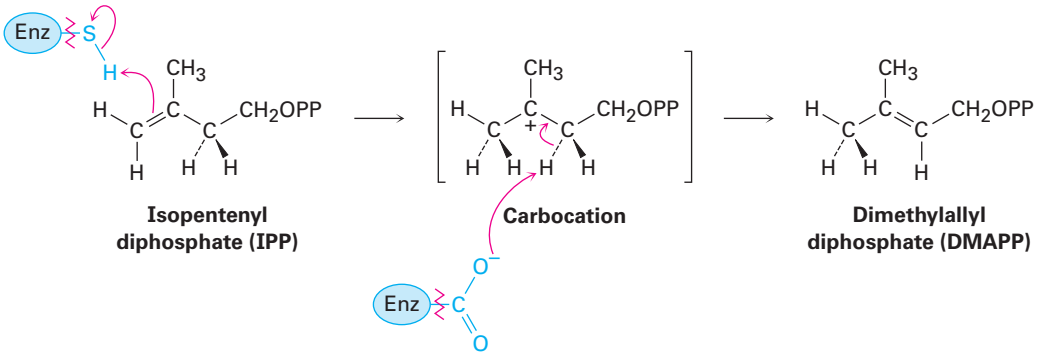


Figure 27.8 An overview of terpenoid biosynthesis from isopentenyl diphosphate.

The isomerization of isopentenyl diphosphate to dimethylallyl diphosphate is catalyzed by IPP isomerase and occurs through a carbocation pathway. Protonation of the IPP double bond by a hydrogen-bonded cysteine residue in the enzyme gives a tertiary carbocation intermediate, which is deprotonated by a glutamate residue as base to yield DMAPP. X-ray structural studies on the enzyme show that it holds the substrate in an unusually deep, well-protected

pocket to shield the highly reactive carbocation from reaction with solvent or other external substances.



Both the initial coupling of DMAPP with IPP to give geranyl diphosphate and the subsequent coupling of GPP with a second molecule of IPP to give farnesyl diphosphate are catalyzed by farnesyl diphosphate synthase. The process requires Mg^{2+} ion, and the key step is a nucleophilic substitution reaction in which the double bond of IPP behaves as a nucleophile in displacing diphosphate ion leaving group (PP_i) on DMAPP. Evidence suggests that the DMAPP develops considerable cationic character and that spontaneous dissociation of the allylic diphosphate ion in an S_N1-like pathway probably occurs **(Figure 27.9)**.

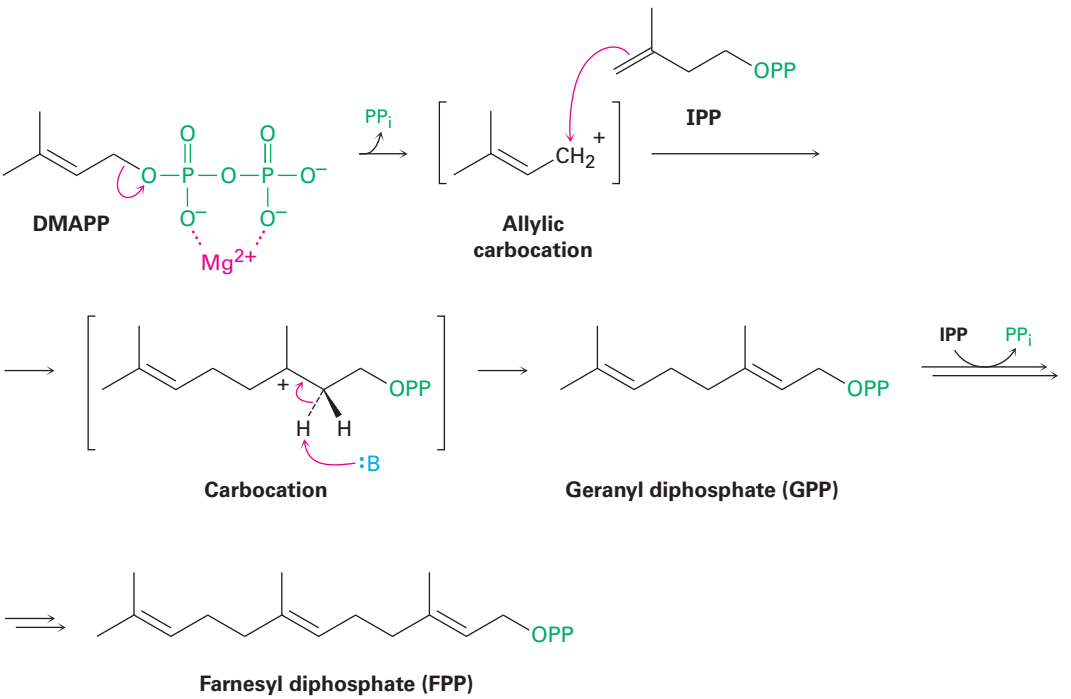


Figure 27.9 Mechanism of the coupling reaction of dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP), to give geranyl diphosphate (GPP).

The further conversion of geranyl diphosphate into monoterpenoids typically involves carbocation intermediates and multistep reaction pathways that are catalyzed by terpene cyclases. Monoterpene cyclases function by first isomerizing geranyl diphosphate to its allylic isomer linalyl diphosphate (LPP), a process that occurs by spontaneous S_N1-like dissociation to an allylic carbocation, followed by recombination. The effect of this isomerization is to convert the C2–C3 double bond of GPP into a single bond, thereby making cyclization possible and allowing *E*/*Z* isomerization of the double bond.

Further dissociation and cyclization by electrophilic addition of the cationic carbon to the terminal double bond then gives a cyclic cation, which might either rearrange, undergo a hydride shift, be captured by a nucleophile, or be deprotonated to give any of the several hundred known monoterpenoids. As just one example, limonene, a monoterpenoid found in many citrus oils, arises by the biosynthetic pathway shown in **Figure 27.10**.

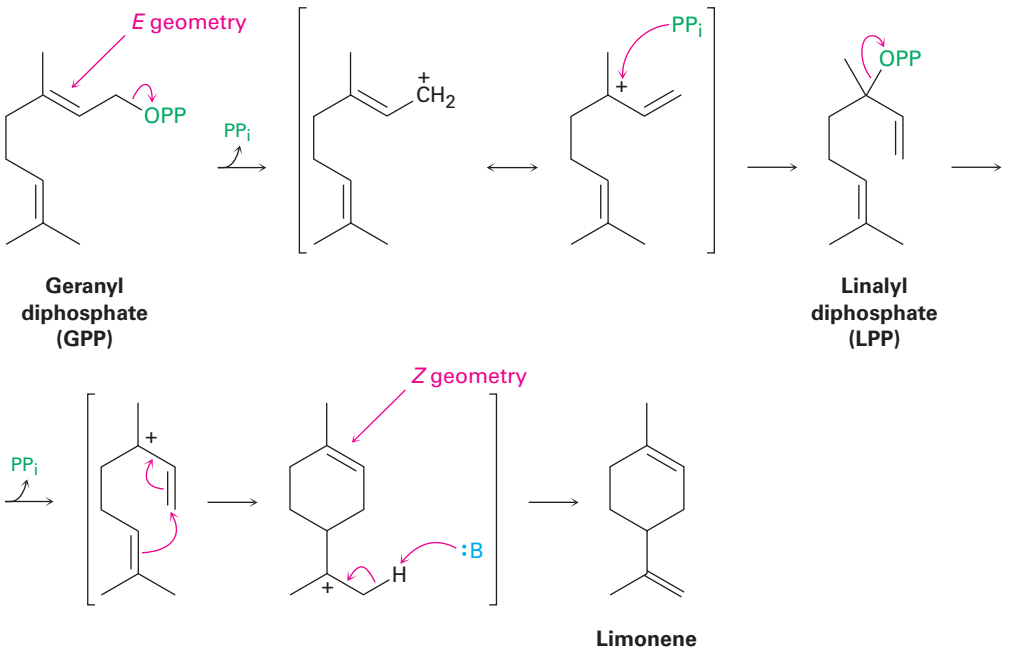
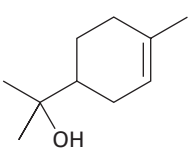


Figure 27.10 Mechanism of the formation of the monoterpenoid limonene from geranyl diphosphate.

Worked Example 27.1 Proposing a Terpenoid Biosynthesis Pathway

Propose a mechanistic pathway for the biosynthesis of α-terpineol from geranyl diphosphate.

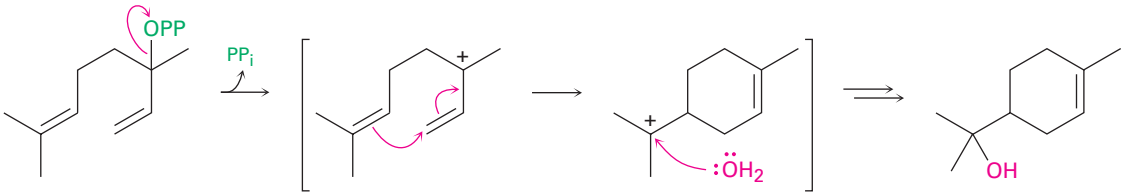


α-Terpineol

Strategy

α-Terpineol, a monoterpenoid, must be derived biologically from geranyl diphosphate through its isomer linalyl diphosphate. Draw the precursor in a conformation that approximates the structure of the target molecule, and then carry out a cationic cyclization, using the appropriate double bond to displace the diphosphate leaving group. Since the target is an alcohol, the carbocation resulting from cyclization evidently reacts with water.

Solution

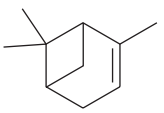


Linalyl diphosphate **α-Terpineol**

Problem 27.7

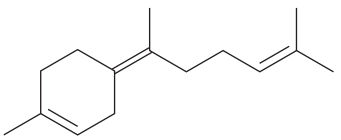
Propose mechanistic pathways for the biosynthetic formation of the following terpenoids:

(a)



α-Pinene

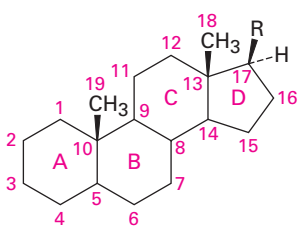
(b)



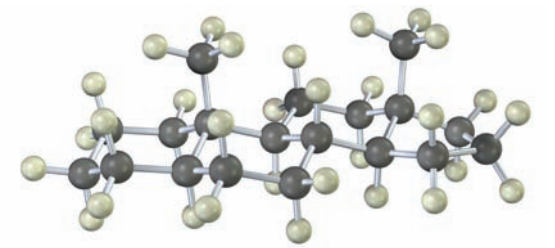
γ-Bisabolene

27.6 Steroids

In addition to fats, phospholipids, eicosanoids, and terpenoids, the lipid extracts of plants and animals also contain **steroids**, molecules that are derived from the triterpenoid lanosterol (Figure 27.6) and whose structures are based on a tetracyclic ring system. The four rings are designated A, B, C, and D, beginning at the lower left, and the carbon atoms are numbered beginning in the A ring. The three 6-membered rings (A, B, and C) adopt chair conformations but are prevented by their rigid geometry from undergoing the usual cyclohexane ring-flips **(Section 4.6)**.

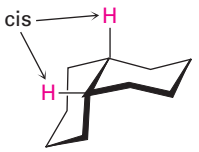


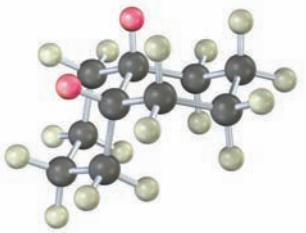
A steroid (R = various side chains)



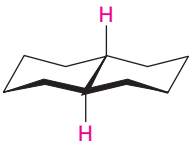
Two cyclohexane rings can be joined in either a cis or a trans manner. With cis fusion to give *cis*-decalin, both groups at the ring-junction positions (the *angular* groups) are on the same side of the two rings. With trans fusion to give *trans*-decalin, the groups at the ring junctions are on opposite sides.

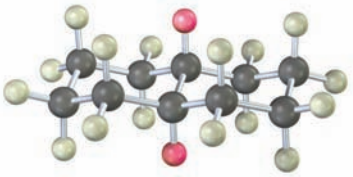
***cis*-Decalin**





***trans*-Decalin**

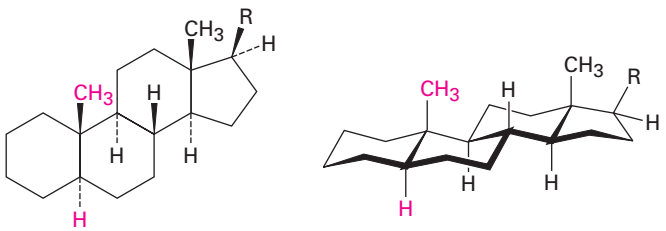




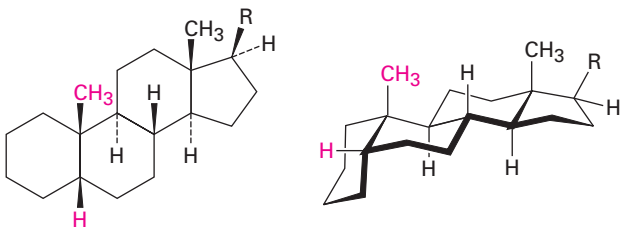
As shown in **Figure 27.11**, steroids can have either a cis or a trans fusion of the A and B rings, but the other ring fusions (B–C and C–D) are usually trans. An A–B trans steroid has the C19 angular methyl group up, denoted β, and the hydrogen atom at C5 down, denoted α, on opposite sides of the molecule. An A–B cis steroid, by contrast, has both the C19 angular methyl

Figure 27.11 Steroid conformations. The three 6-membered rings have chair conformations but are unable to undergo ring-flips. The A and B rings can be either cis-fused or trans-fused.

An A–B trans steroid

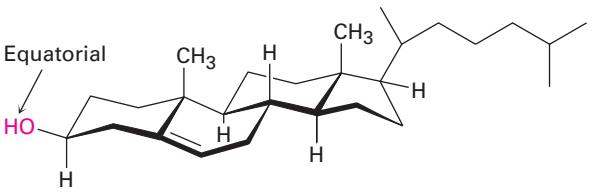


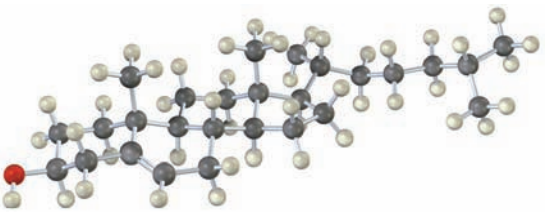
An A–B cis steroid



group and the C5 hydrogen atom on the same side (β) of the molecule. Both kinds of steroids are relatively long, flat molecules that have their two methyl groups (C18 and C19) protruding axially above the ring system. The A–B trans steroids are the more common, although A–B cis steroids are found in liver bile.

Substituent groups on the steroid ring system can be either axial or equatorial. As with simple cyclohexanes **(Section 4.7)**, equatorial substitution is generally more favorable than axial substitution for steric reasons. The hydroxyl group at C3 of cholesterol, for example, has the more stable equatorial orientation. Unlike what happens with simple cyclohexanes, however, steroids are rigid molecules whose geometry prevents cyclohexane ring-flips.



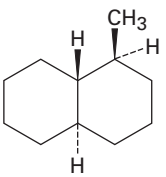


Cholesterol

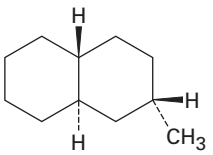
Problem 27.8

Draw the following molecules in chair conformations, and tell whether the ring substituents are axial or equatorial:

(a)

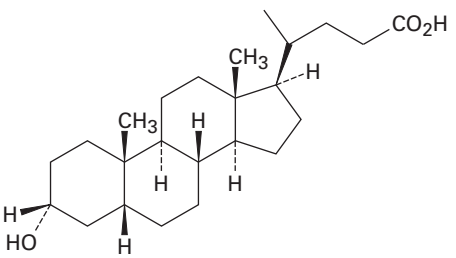


(b)



Problem 27.9

Lithocholic acid is an A–B cis steroid found in human bile. Draw lithocholic acid showing chair conformations, as in Figure 27.11, and tell whether the hydroxyl group at C3 is axial or equatorial.



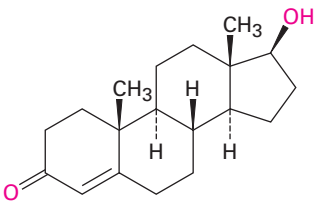
Lithocholic acid

Steroid Hormones

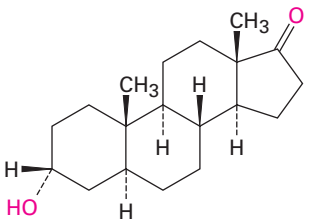
In humans, most steroids function as **hormones**, chemical messengers that are secreted by endocrine glands and carried through the bloodstream to target tissues. There are two main classes of steroid hormones: the *sex hormones,* which control maturation, tissue growth, and reproduction, and the *adrenocortical hormones,* which regulate a variety of metabolic processes.

Sex Hormones

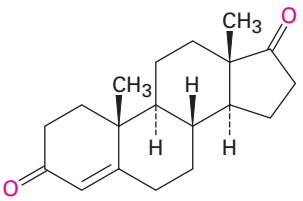
Testosterone and androsterone are the two most important male sex hormones, or *androgens*. Androgens are responsible for the development of male secondary sex characteristics during puberty and for promoting tissue and muscle growth. Both are synthesized in the testes from cholesterol. Androstenedione is another minor hormone that has received particular attention because of its use by prominent athletes.



Testosterone



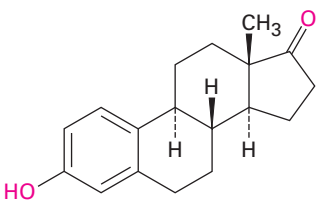
Androsterone



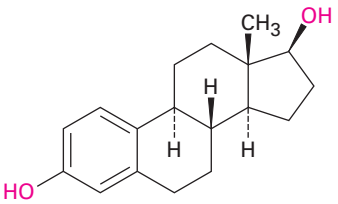
Androstenedione

(Androgens)

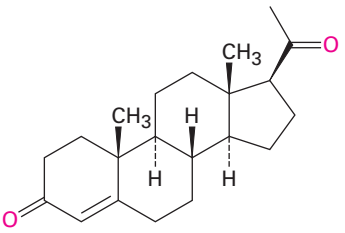
Estrone and estradiol are the two most important female sex hormones, or *estrogens*. Synthesized in the ovaries from testosterone, estrogenic hormones are responsible for the development of female secondary sex characteristics and for regulation of the menstrual cycle. Note that both have a benzene-like aromatic A ring. In addition, another kind of sex hormone called a *progestin* is essential for preparing the uterus for implantation of a fertilized ovum during pregnancy. Progesterone is the most important progestin.



Estrone



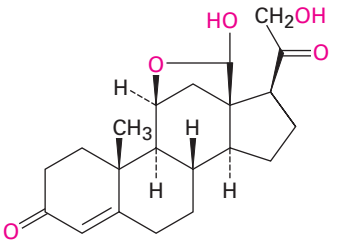
Estradiol



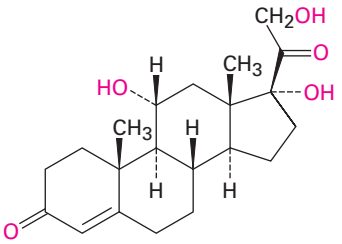
Progesterone (a progestin)

(Estrogens)

Adrenocortical Hormones Adrenocortical steroids are secreted by the adrenal glands, small organs located near the upper end of each kidney. There are two types of adrenocortical steroids, called *mineralocorticoids* and *glucocorticoids*. Mineralocorticoids, such as aldosterone, control tissue swelling by regulating cellular salt balance between Na^+ and K^+. Glucocorticoids, such as hydrocortisone, are involved in the regulation of glucose metabolism and in the control of inflammation. Glucocorticoid ointments are widely used to bring down the swelling from exposure to poison oak or poison ivy.

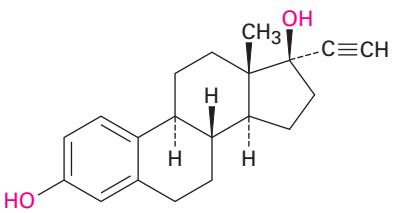


Aldosterone
(a mineralocorticoid)

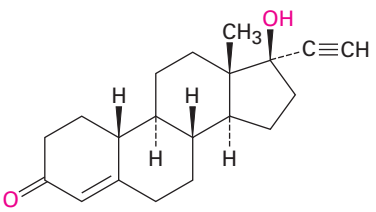


Hydrocortisone
(a glucocorticoid)

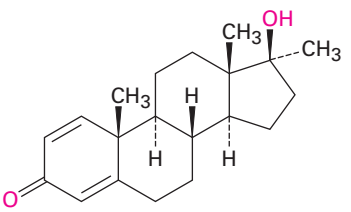
Synthetic Steroids In addition to the many hundreds of steroids isolated from plants and animals, thousands more have been synthesized in pharmaceutical laboratories in a search for new drugs. Among the best-known synthetic steroids are the oral contraceptives and anabolic agents. Most birth control pills are a mixture of two compounds, a synthetic estrogen, such as ethynylestradiol, and a synthetic progestin, such as norethindrone. Anabolic steroids, such as methandrostenolone (Dianabol), are synthetic androgens that mimic the tissue-building effects of natural testosterone.



Ethynylestradiol
(a synthetic estrogen)



Norethindrone
(a synthetic progestin)



Methandrostenolone
(Dianabol)

27.7 Biosynthesis of Steroids

Steroids are heavily modified triterpenoids that are biosynthesized in living organisms from farnesyl diphosphate (C_{15}). A reductive dimerization first converts farnesyl diphosphate to the acyclic hydrocarbon squalene (C_{30}), which is converted into lanosterol **(Figure 27.12)**. Further rearrangements and degradations then take place to yield various steroids. The conversion of squalene to lanosterol is among the most intensively studied of all biosynthetic transformations. Starting from an achiral, open-chain polyene, the entire process requires only two enzymes and results in the formation of six carbon–carbon bonds, four rings, and seven chirality centers.

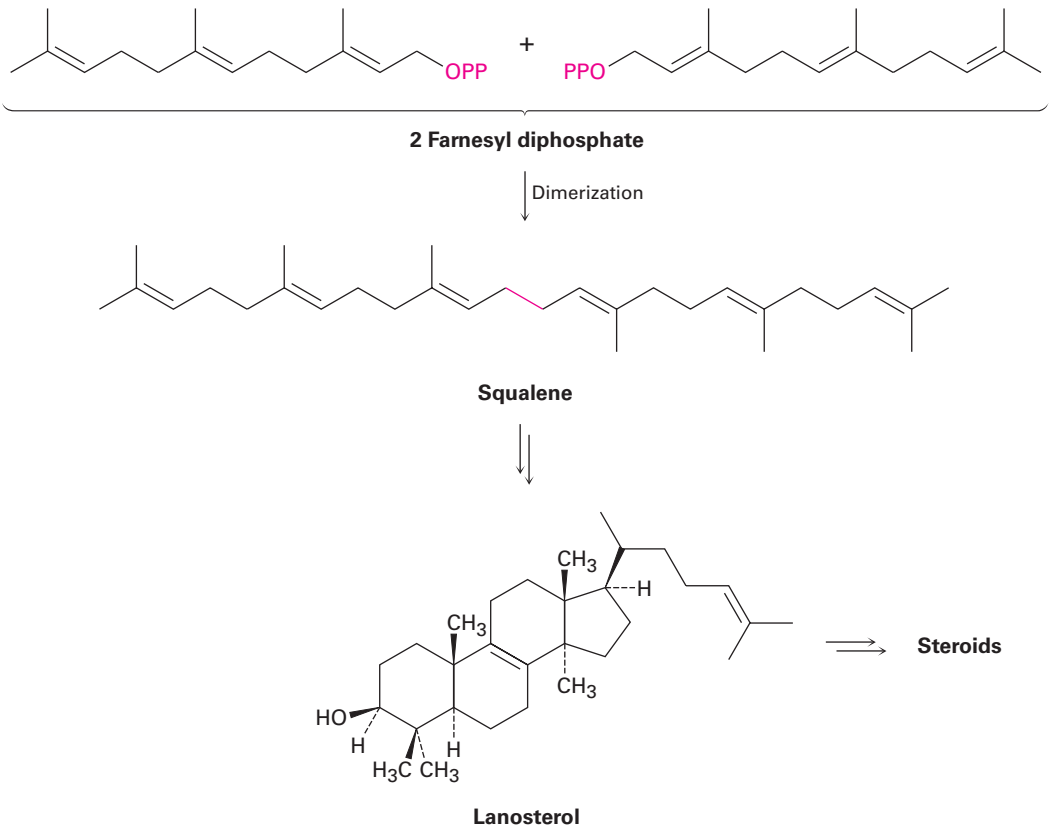


Figure 27.12 An overview of steroid biosynthesis from farnesyl diphosphate.

Lanosterol biosynthesis begins with the selective epoxidation of squalene to give (3*S*)-2,3-oxidosqualene, catalyzed by squalene epoxidase. Molecular O_2 provides the source of the epoxide oxygen atom, and NADPH is required, along with a flavin coenzyme. The proposed mechanism involves reaction of $FADH_2$ with O_2 to produce a flavin hydroperoxide intermediate (ROOH), which transfers an oxygen to squalene in a pathway initiated by nucleophilic

attack of the squalene double bond on the terminal hydroperoxide oxygen **(Figure 27.13)**. The flavin alcohol formed as a by-product loses H_2O to give FAD, which is reduced back to $FADH_2$ by NADPH. As noted in **Section 8.7**, this biological epoxidation mechanism is closely analogous to the mechanism by which peroxyacids (RCO_3H) react with alkenes to give epoxides in the laboratory.

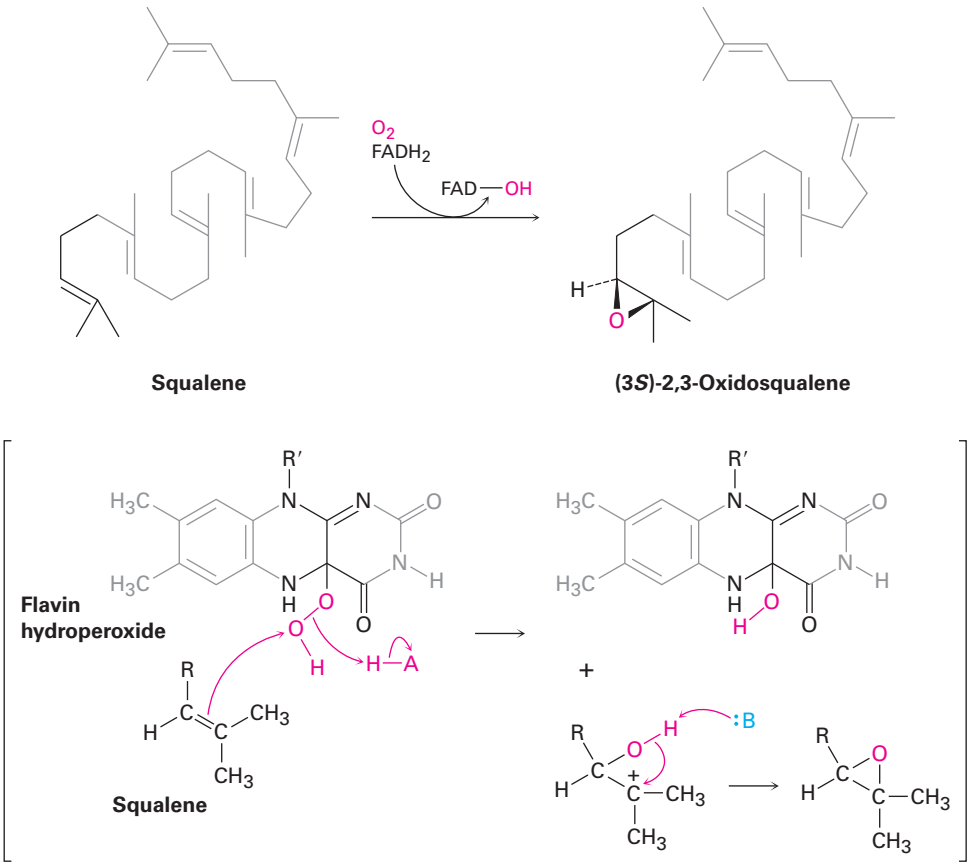
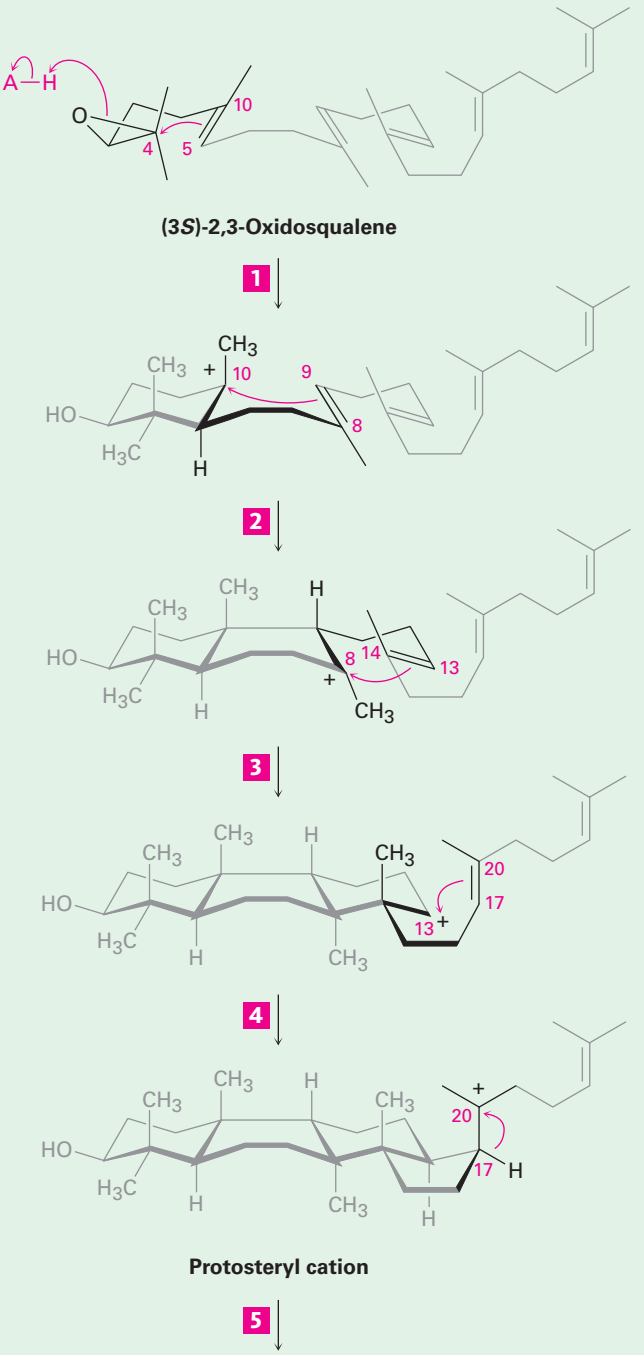


Figure 27.13 Proposed mechanism of the oxidation of squalene by flavin hydroperoxide.

The second part of lanosterol biosynthesis is catalyzed by oxidosqualene: lanosterol cyclase and occurs as shown in **Figure 27.14**. Squalene is folded by the enzyme into a conformation that aligns the various double bonds for undergoing a cascade of successive intramolecular electrophilic additions, followed by a series of hydride and methyl migrations. Except for the initial epoxide protonation/cyclization, the process is probably stepwise and appears to involve discrete carbocation intermediates that are stabilized by electrostatic interactions with electron-rich aromatic amino acids in the enzyme.



1 Protonation on oxygen opens the epoxide ring and gives a tertiary carbocation at C4. Intramolecular electrophilic addition of C4 to the 5,10 double bond then yields a tertiary monocyclic carbocation at C10.

2 The C10 carbocation adds to the 8,9 double bond, giving a C8 tertiary bicyclic carbocation.

3 Further intramolecular addition of the C8 carbocation to the 13,14 double bond occurs with non-Markovnikov regiochemistry and gives a tricyclic *secondary* carbocation at C13.

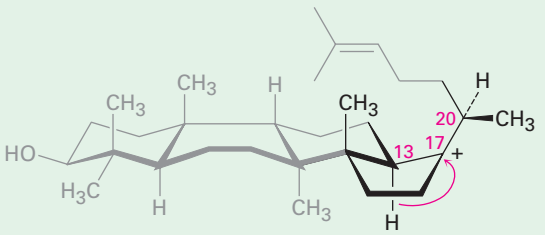
4 The fourth and final cyclization occurs by addition of the C13 cation to the 17,20 double bond, giving the protosteryl cation with 17β stereochemistry.

Figure 27.14 | **MECHANISM**

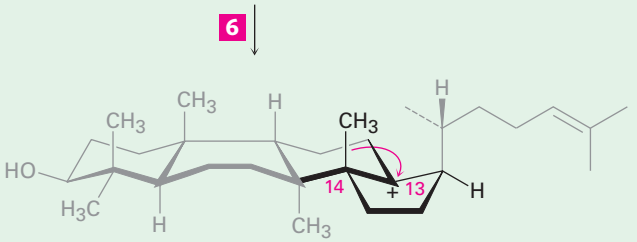
Mechanism of the conversion of 2,3-oxidosqualene to lanosterol. Four cationic cyclizations are followed by four rearrangements and a final loss of H^+ from C9. The steroid numbering system is used for referring to specific positions in the intermediates (Section 27.6). Individual steps are explained in the text

Protosteryl cation

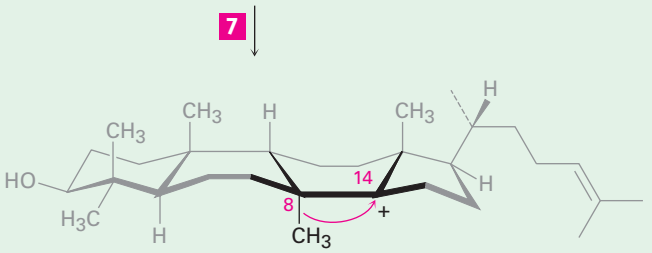
5 Hydride migration from C17 to C20 occurs, establishing *R* stereochemistry at C20.



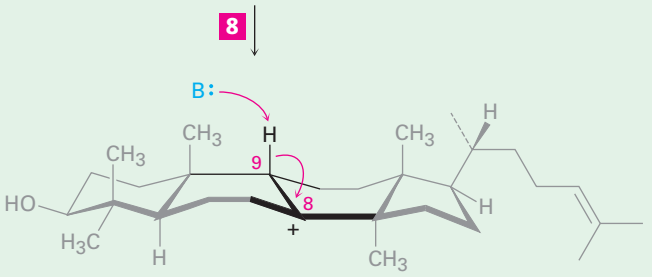
6 A second hydride migration takes place, from C13 to C17, establishing the final 17β stereochemistry of the side chain.



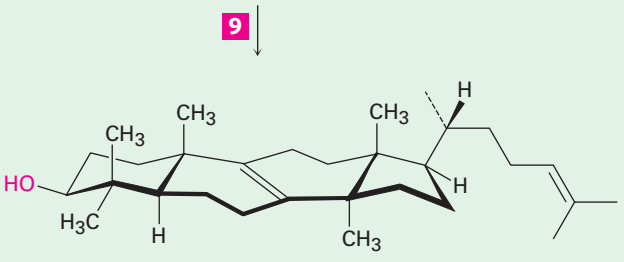
7 Methyl migration from C14 to C13 occurs.



8 A second methyl migration occurs, from C8 to C14.



9 Loss of a proton from C9 forms an 8,9 double bond and gives lanosterol.

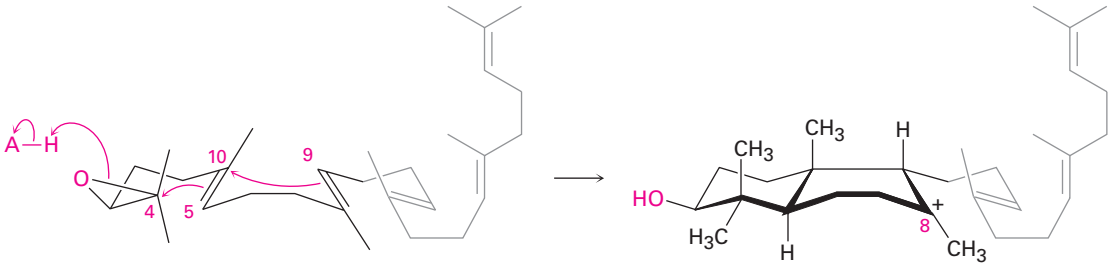


Lanosterol

© John McMurry

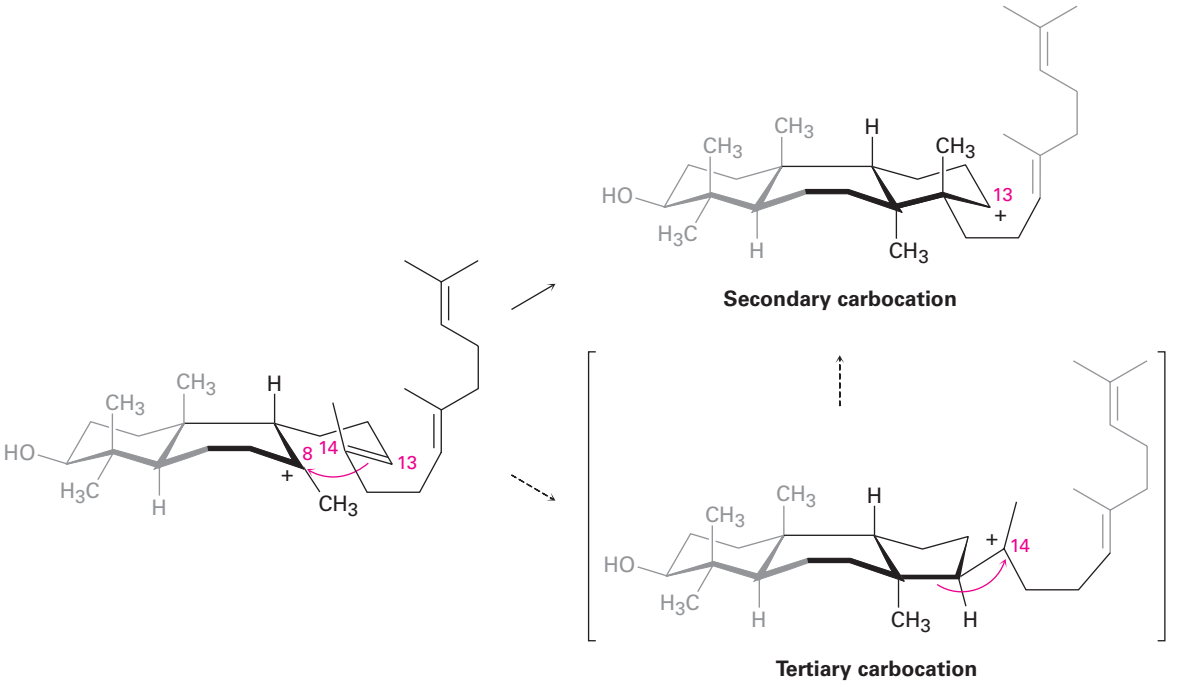
Figure 27.14 (Continued)

Steps 1, 2 of Figure 27.14: Epoxide Opening and Initial Cyclizations Cyclization begins in step 1 with protonation of the epoxide ring by an aspartic acid residue in the enzyme. Nucleophilic opening of the protonated epoxide by the nearby 5,10 double bond (steroid numbering; **Section 27.6**) then yields a tertiary carbocation at C10. Further addition of C10 to the 8,9 double bond in step 2 next gives a bicyclic tertiary cation at C8.

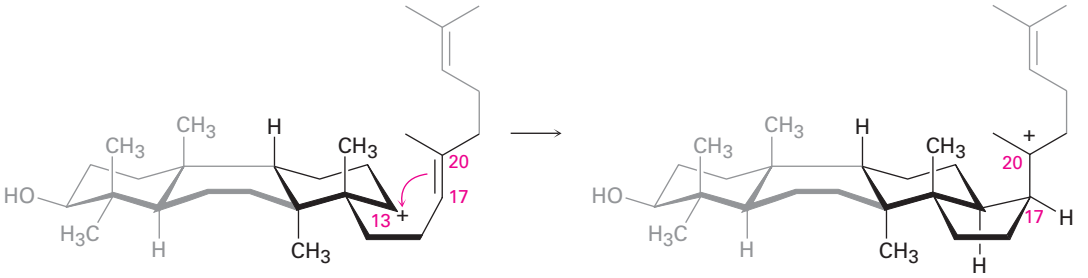


(3*S*)-2,3-Oxidosqualene

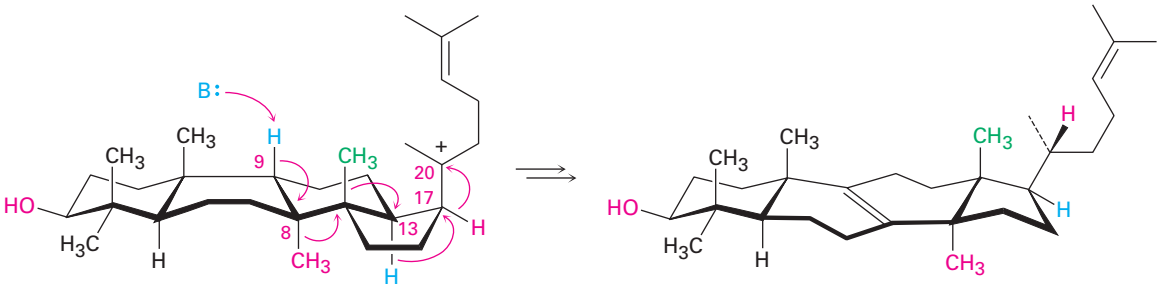
Step 3 of Figure 27.14: Third Cyclization The third cationic cyclization is somewhat unusual because it occurs with non-Markovnikov regiochemistry and gives a secondary cation at C13 rather than the alternative tertiary cation at C14. There is growing evidence, however, that the tertiary carbocation may in fact be formed initially and that the secondary cation arises by subsequent rearrangement. The secondary cation is probably stabilized in the enzyme pocket by the proximity of an electron-rich aromatic ring.



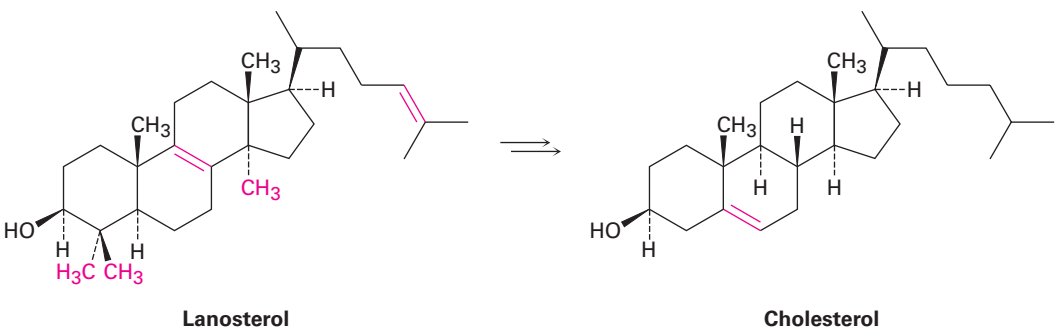
Step 4 of Figure 27.14: Final Cyclization The fourth and last cyclization occurs in step 4 by addition of the cationic center at C13 to the 17,20 double bond, giving what is known as the *protosteryl* cation. The side-chain alkyl group at C17 has β (up) stereochemistry, although this stereochemistry is lost in step 5 and then reset in step 6.



Steps 5–9 of Figure 27.14: Carbocation Rearrangements Once the tetracyclic carbon skeleton of lanosterol has been formed, a series of carbocation rearrangements occur **(Section 7.11)**. The first rearrangement, hydride migration from C17 to C20, occurs in step 5 and results in establishment of *R* stereochemistry at C20 in the side chain. A second hydride migration then occurs from C13 to C17 on the α (bottom) face of the ring in step 6 and reestablishes the 17β orientation of the side chain. Finally, two methyl migrations, the first from C14 to C13 on the top (β) face and the second from C8 to C14 on the bottom (α) face, place the positive charge at C8. A basic histidine residue in the enzyme then removes the neighboring β proton from C9 to give lanosterol.



From lanosterol, the pathway for steroid biosynthesis continues on to yield cholesterol. Cholesterol then becomes a branch point, serving as the common precursor from which all other steroids are derived.



Problem 27.10

Compare the structures of lanosterol and cholesterol, and catalog the changes needed for the transformation.

iStockphoto.com/Rob Friedman

It's hard to resist, but a high intake of saturated animal fat doesn't do much for your cholesterol level.

Saturated Fats, Cholesterol, and Heart Disease

A DEEPER LOOK

We hear a lot these days about the relationships between saturated fats, cholesterol, and heart disease. What are the facts? It's well established that a diet rich in saturated animal fats often leads to an increase in blood serum cholesterol, particularly in sedentary, overweight people. Conversely, a diet lower in saturated fats and higher in polyunsaturated fats leads to a lower serum cholesterol level. Studies have shown that a serum cholesterol level greater than 240 mg/dL (a desirable value is <200 mg/dL) is correlated with an increased incidence of coronary artery disease, in which cholesterol deposits build up on the inner walls of coronary arteries, blocking the flow of blood to the heart muscles.

A better indication of a person's risk of heart disease comes from a measurement of blood lipoprotein levels. *Lipoproteins* are complex molecules with both lipid and protein parts that transport lipids through the body. They can be divided into three types according to density, as shown in Table 27.3. Very-low-density lipoproteins (VLDLs) act primarily as carriers of triglycerides from the intestines to peripheral tissues, whereas low-density lipoproteins (LDLs) and high-density lipoproteins (HDLs) act as carriers of cholesterol to and from the liver.

Evidence suggests that LDLs transport cholesterol as its fatty-acid ester to peripheral tissues, whereas HDLs remove cholesterol as its stearate ester from dying cells. If LDLs deliver more cholesterol than is needed, and if insufficient HDLs are present to remove it, the excess is deposited in arteries. Thus, a *low* level of *low*-density lipoproteins is good because it means that less cholesterol is being transported, and a *high* level of *high*-density lipoproteins is good because it means that more cholesterol is being removed. In addition, HDL contains an enzyme that has antioxidant properties, offering further protection against heart disease.

As a rule of thumb, a person's risk drops about 25% for each increase of 5 mg/dL in HDL concentration. Normal values are about 45 mg/dL for men and 55 mg/dL for women, perhaps explaining why premenopausal women appear to be somewhat less susceptible than men to heart disease.

Table 27.3 Serum Lipoproteins

Name	Density (g/mL)	% Lipid	% Protein	Optimal (mg/dL)	Poor (mg/dL)
VLDL	0.940–1.006	90	10	—	—
LDL	1.006–1.063	75	25	<100	>130
HDL	1.063–1.210	60	40	>60	<40

Not surprisingly, the most important factor in gaining high HDL levels is a generally healthful lifestyle. Obesity, smoking, and lack of exercise lead to low HDL levels, whereas regular exercise and a sensible diet lead to high HDL levels. Distance runners and other endurance athletes have HDL levels nearly 50% higher than the general population. Failing that—not everyone wants to run 30 miles or bike 100 miles per week—diet is also important. Diets high in cold-water fish like salmon and whitefish, raise HDL and lower blood cholesterol because these fish contain almost entirely polyunsaturated fat, including a large percentage of omega-3 fatty acids. Animal fat from red meat and cooking fats should be minimized because saturated fats and monounsaturated trans fats raise blood cholesterol.

Summary

Lipids are the naturally occurring materials isolated from plants and animals by extraction with nonpolar organic solvents. Animal fats and vegetable oils are the most widely occurring lipids. Both are **triacylglycerols**—triesters of glycerol with long-chain **fatty acids**. Animal fats are usually saturated, whereas vegetable oils usually have unsaturated fatty acid residues.

Phospholipids are important constituents of cell membranes and are of two kinds. *Glycerophospholipids,* such as phosphatidylcholine and phosphatidylethanolamine, are closely related to fats in that they have a glycerol backbone esterified to two fatty acids (one saturated and one unsaturated) and to one phosphate ester. *Sphingomyelins* have the amino alcohol sphingosine for their backbone.

Eicosanoids and **terpenoids** are still other classes of lipids. Eicosanoids, of which prostaglandins are the most abundant kind, are derived biosynthetically from arachidonic acid, are found in all body tissues, and have a wide range of physiological activity. Terpenoids are often isolated from the essential oils of plants, have an immense diversity of structure, and are produced biosynthetically from the five-carbon precursor isopentenyl diphosphate (IPP). Isopentenyl diphosphate is itself biosynthesized from 3 equivalents of acetate in the mevalonate pathway.

Steroids are plant and animal lipids with a characteristic tetracyclic carbon skeleton. Like the eicosanoids, steroids occur widely in body tissues and have a large variety of physiological activities. Steroids are closely related to terpenoids and arise biosynthetically from the triterpenoid lanosterol. Lanosterol, in turn, arises from cationic cyclization of the acyclic hydrocarbon squalene.

Key words

eicosanoid, 1096
fatty acid, 1089
hormone, 1110
lipid, 1088
lipid bilayer, 1095
micelle, 1092
phospholipid, 1094
polyunsaturated fatty acid, 1089
prostaglandin, 1095
steroid, 1107
terpenoid, 1098
triacylglycerol, 1089
wax, 1088

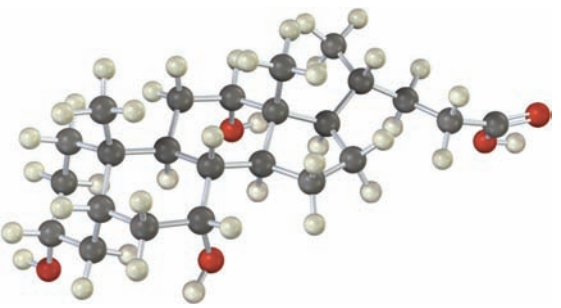
Exercises

Visualizing Chemistry

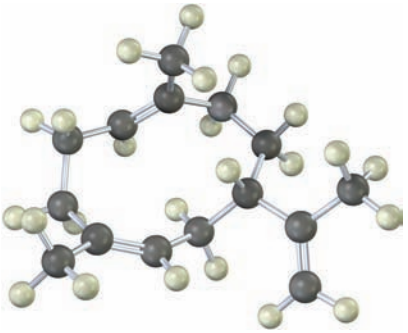
OWL Interactive versions of these problems are assignable in OWL for Organic Chemistry.

(Problems 27.1–27.10 appear within the chapter.)

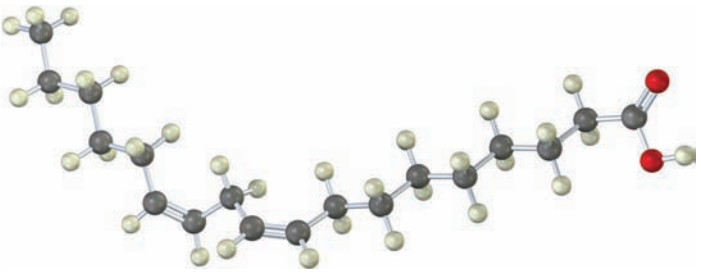
27.11 The following model is that of cholic acid, a constituent of human bile. Locate the three hydroxyl groups, and identify each as axial or equatorial. Is cholic acid an A–B trans steroid or an A–B cis steroid?



27.12 Propose a biosynthetic pathway for the sesquiterpenoid helminthogermacrene from farnesyl diphosphate.



27.13 Identify the following fatty acid, and tell whether it is more likely to be found in peanut oil or in red meat:



Additional Problems

Fats, Oils, and Related Lipids

27.14 Fatty fish like salmon and albacore are rich in *omega-3* fatty acids, which have a double bond three carbons in from the noncarboxyl end of the chain and have been shown to lower blood cholesterol levels. Draw the structure of 5,8,11,14,17-eicosapentaenoic acid, a common example. (Eicosane = $C_{20}H_{42}$.)

27.15 Fats can be either optically active or optically inactive, depending on their structure. Draw the structure of an optically active fat that yields 2 equivalents of stearic acid and 1 equivalent of oleic acid on hydrolysis. Draw the structure of an optically inactive fat that yields the same products.

27.16 Spermaceti, a fragrant substance from sperm whales, was much used in cosmetics until it was banned in 1976 to protect the whales from extinction. Chemically, spermaceti is cetyl palmitate, the ester of cetyl alcohol (n-$C_{16}H_{33}OH$) with palmitic acid. Draw its structure.

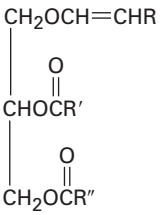
27.17 Show the products you would expect to obtain from reaction of glyceryl trioleate with the following reagents:

(a) Excess Br_2 in CH_2Cl_2 **(b)** H_2/Pd
(c) $NaOH/H_2O$ **(d)** O_3, then Zn/CH_3CO_2H
(e) $LiAlH_4$, then H_3O^+ **(f)** CH_3MgBr, then H_3O^+

27.18 How would you convert oleic acid into the following substances?

(a) Methyl oleate
(b) Methyl stearate
(c) Nonanal
(d) Nonanedioic acid
(e) 9-Octadecynoic acid (stearolic acid)
(f) 2-Bromostearic acid
(g) 18-Pentatriacontanone, $CH_3(CH_2)_{16}CO(CH_2)_{16}CH_3$

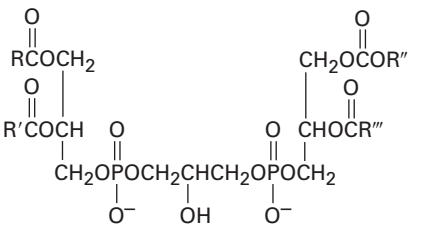
27.19 The *plasmalogens* are a group of lipids found in nerve and muscle cells. How do plasmalogens differ from fats?



A plasmalogen

27.20 What products would you obtain from hydrolysis of a plasmalogen (Problem 27.19) with aqueous NaOH? With H_3O^+?

27.21 *Cardiolipins* are a group of lipids found in heart muscles. What products would be formed if all ester bonds, including phosphates, were saponified by treatment with aqueous NaOH?



A cardiolipin

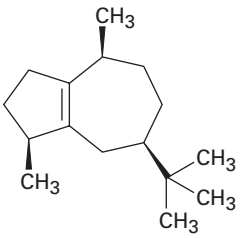
27.22 Stearolic acid, $C_{18}H_{32}O_2$, yields stearic acid on catalytic hydrogenation and undergoes oxidative cleavage with ozone to yield nonanoic acid and nonanedioic acid. What is the structure of stearolic acid?

27.23 How would you synthesize stearolic acid (Problem 27.22) from 1-decyne and 1-chloro-7-iodoheptane?

Terpenoids and Steroids

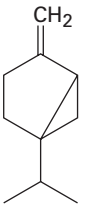
27.24 Without proposing an entire biosynthetic pathway, draw the appropriate precursor, either geranyl diphosphate or farnesyl diphosphate, in a conformation that shows a likeness to each of the following terpenoids:

(a)



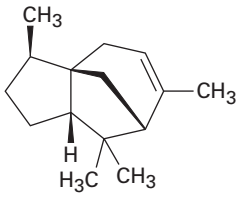
Guaiol

(b)



Sabinene

(c)



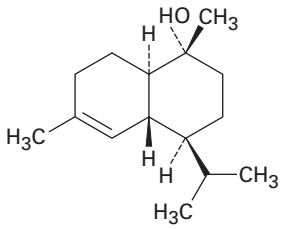
Cedrene

27.25 Indicate by asterisks the chirality centers present in each of the terpenoids shown in Problem 27.24. What is the maximum possible number of stereoisomers for each?

27.26 Assume that the three terpenoids in Problem 27.24 are derived biosynthetically from isopentenyl diphosphate and dimethylallyl diphosphate, each of which was isotopically labeled at the diphosphate-bearing carbon atom (C1). At what positions would the terpenoids be isotopically labeled?

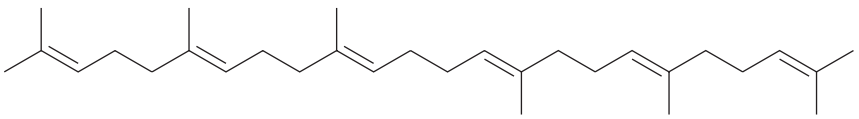
27.27 Assume that acetyl CoA containing a ^{14}C isotopic label in the carboxyl carbon atom is used as starting material for the biosynthesis of mevalonate, as shown in Figure 27.7. At what positions in mevalonate would the isotopic label appear?

27.28 Assume that acetyl CoA containing a ^{14}C isotopic label in the carboxyl carbon atom is used as starting material and that the mevalonate pathway is followed. Identify the positions in α-cadinol where the label would appear.



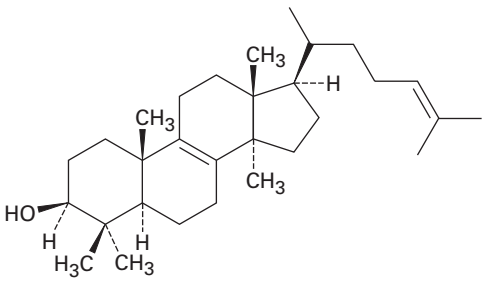
α-Cadinol

27.29 Assume that acetyl CoA containing a ^{14}C isotopic label in the carboxyl carbon atom is used as starting material and that the mevalonate pathway is followed. Identify the positions in squalene where the label would appear.



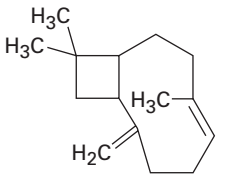
Squalene

27.30 Assume that acetyl CoA containing a ^{14}C isotopic label in the carboxyl carbon atom is used as starting material and that the mevalonate pathway is followed. Identify the positions in lanosterol where the label would appear.



Lanosterol

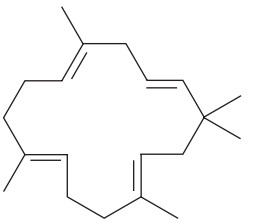
27.31 Propose a mechanistic pathway for the biosynthesis of caryophyllene, a substance found in clove oil.



Caryophyllene

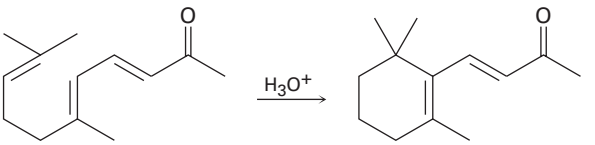
General Problems

27.32 Flexibilene, a compound isolated from marine coral, is the first known terpenoid to contain a 15-membered ring. What is the structure of the acyclic biosynthetic precursor of flexibilene? Show the mechanistic pathway for the biosynthesis.



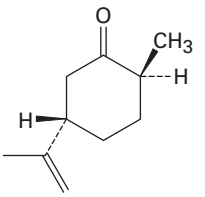
Flexibilene

27.33 Suggest a mechanism by which ψ-ionone is transformed into β-ionone on treatment with acid.



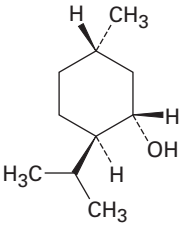
ψ-Ionone **β-Ionone**

27.34 Draw the most stable chair conformation of dihydrocarvone.



Dihydrocarvone

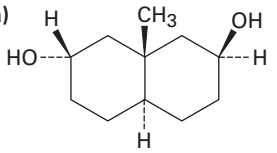
27.35 Draw the most stable chair conformation of menthol, and label each substituent as axial or equatorial.



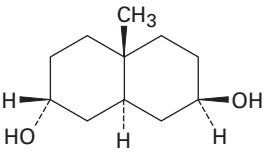
Menthol (from peppermint oil)

27.36 As a general rule, equatorial alcohols are esterified more readily than axial alcohols. What product would you expect to obtain from reaction of the following two compounds with 1 equivalent of acetic anhydride?

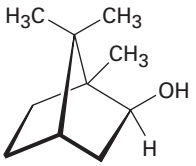
(a)



(b)

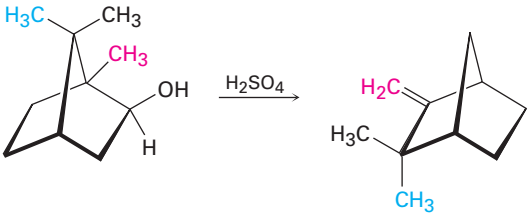


27.37 Propose a mechanistic pathway for the biosynthesis of isoborneol. A carbocation rearrangement is needed at one point in the scheme.



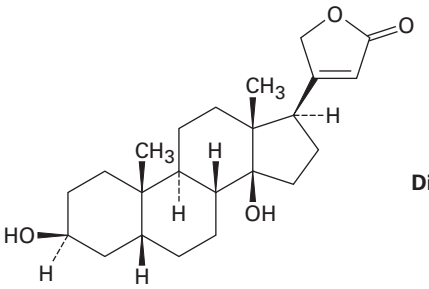
Isoborneol

27.38 Isoborneol (Problem 27.37) is converted into camphene on treatment with dilute sulfuric acid. Propose a mechanism for the reaction, which involves a carbocation rearrangement.



Isoborneol **Camphene**

27.39 Digitoxigenin is a heart stimulant obtained from the purple foxglove *Digitalis purpurea* and used in the treatment of heart disease. Draw the three-dimensional conformation of digitoxigenin, and identify the two –OH groups as axial or equatorial.



Digitoxigenin

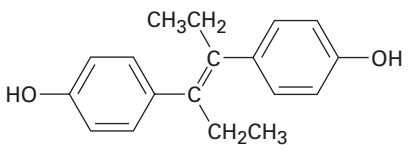
27.40 What product would you obtain by reduction of digitoxigenin (Problem 27.39) with $LiAlH_4$? By oxidation with the Dess–Martin periodinane?

27.41 Vaccenic acid, $C_{18}H_{34}O_2$, is a rare fatty acid that gives heptanal and 11-oxoundecanoic acid [$OHC(CH_2)_9CO_2H$] on ozonolysis followed by zinc treatment. When allowed to react with $CH_2I_2/Zn(Cu)$, vaccenic acid is converted into lactobacillic acid. What are the structures of vaccenic and lactobacillic acids?

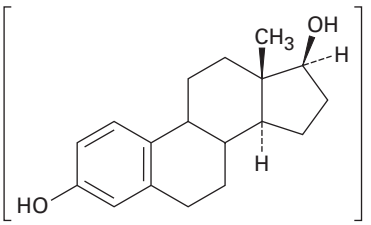
27.42 Eleostearic acid, $C_{18}H_{30}O_2$, is a rare fatty acid found in the tung oil used for finishing furniture. On ozonolysis followed by treatment with zinc, eleostearic acid furnishes one part pentanal, two parts glyoxal (OHC—CHO), and one part 9-oxononanoic acid [$OHC(CH_2)_7CO_2H$]. What is the structure of eleostearic acid?

27.43 Diterpenoids are derived biosynthetically from geranylgeranyl diphosphate (GGPP), which is itself biosynthesized by reaction of farnesyl diphosphate with isopentenyl diphosphate. Show the structure of GGPP, and propose a mechanism for its biosynthesis from FPP and IPP.

27.44 Diethylstilbestrol (DES) has estrogenic activity even though it is structurally unrelated to steroids. Once used as an additive in animal feed, DES has been implicated as a causative agent in several types of cancer. Show how DES can be drawn so that it is sterically similar to estradiol.



Diethylstilbestrol



Estradiol

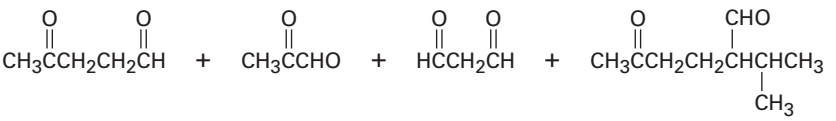
27.45 Propose a synthesis of diethylstilbestrol (Problem 27.44) from phenol and any other organic compound required.

27.46 What products would you expect from reaction of estradiol (Problem 27.44) with the following reagents?

(a) NaH, then CH_3I **(b)** CH_3COCl, pyridine

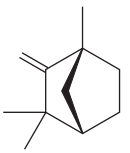
(c) Br_2, $FeBr_3$ **(d)** Dess–Martin periodinane

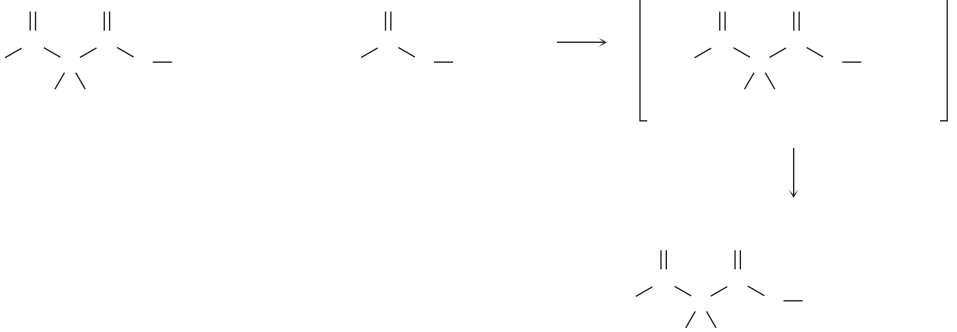
27.47 Cembrene, $C_{20}H_{32}$, is a diterpenoid hydrocarbon isolated from pine resin. Cembrene has a UV absorption at 245 nm, but dihydrocembrene ($C_{20}H_{34}$), the product of hydrogenation with 1 equivalent of H_2, has no UV absorption. On exhaustive hydrogenation, 4 equivalents of H_2 react, and octahydrocembrene, $C_{20}H_{40}$, is produced. On ozonolysis of cembrene, followed by treatment of the ozonide with zinc, four carbonyl-containing products are obtained:

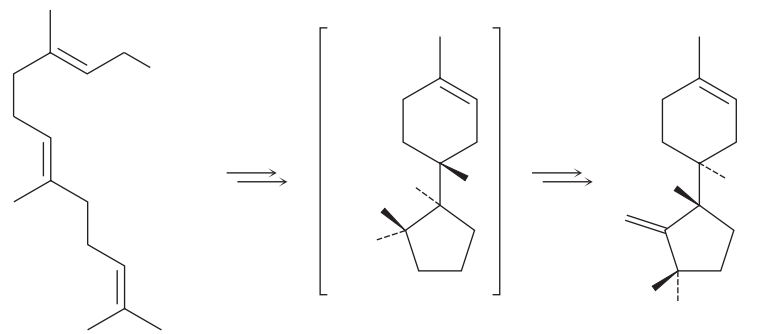


Propose a structure for cembrene that is consistent with its formation from geranylgeranyl diphosphate.

27.48 α-Fenchone is a pleasant-smelling terpenoid isolated from oil of lavender. Propose a pathway for the formation of α-fenchone from geranyl diphosphate. A carbocation rearrangement is required.







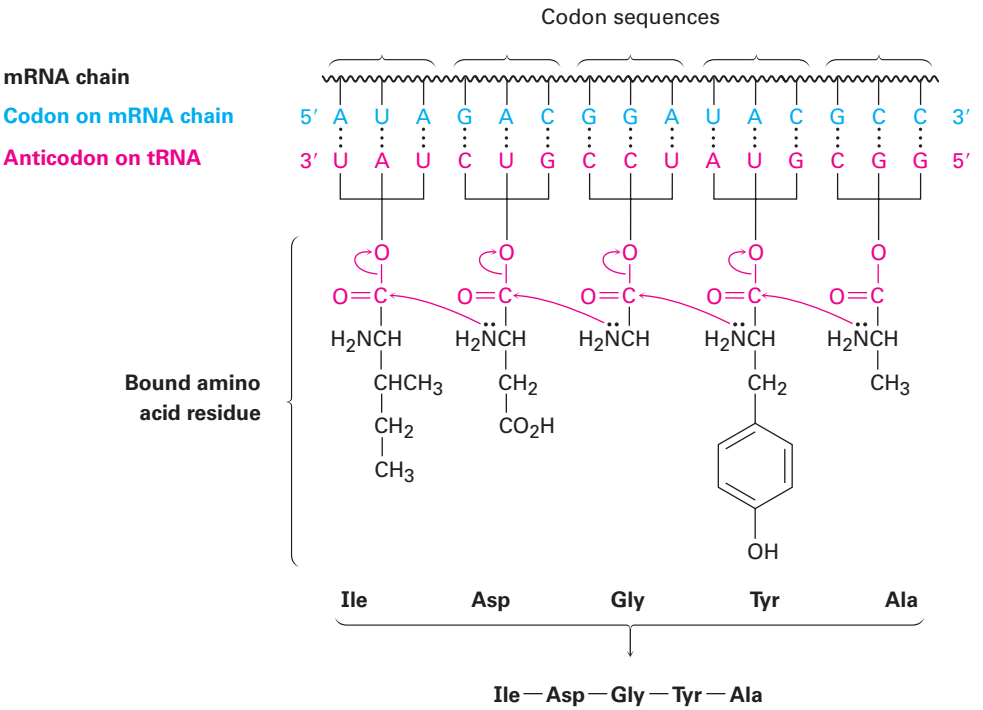


Figure 28.7 A representation of protein biosynthesis. The codon base sequences on mRNA are read by tRNAs containing complementary anticodon base sequences. Transfer RNAs assemble the proper amino acids into position for incorporation into the growing peptide.

Predicting the Amino Acid Sequence Transcribed from DNA

Worked Example 28.2

What amino acid sequence is coded by the following segment of a DNA coding strand (sense strand)?

(5′) CTA-ACT-AGC-GGG-TCG-CCG (3′)

Strategy

The mRNA produced during translation is a copy of the DNA coding strand, with each T replaced by U. Thus, the mRNA has the sequence

(5′) CUA-ACU-AGC-GGG-UCG-CCG (3′)

Each set of three bases forms a codon, whose meaning can be found in Table 28.1.

Solution

Leu-Thr-Ser-Gly-Ser-Pro.

Problem 28.7

List codon sequences for the following amino acids:

(a) Ala **(b)** Phe **(c)** Leu **(d)** Tyr

Problem 28.8

List anticodon sequences on the tRNAs carrying the amino acids shown in Problem 28.7.

Problem 28.9
What amino acid sequence is coded by the following mRNA base sequence?

CUU-AUG-GCU-UGG-CCC-UAA

Problem 28.10
What is the base sequence in the original DNA strand on which the mRNA sequence in Problem 28.9 was made?

28.6 DNA Sequencing

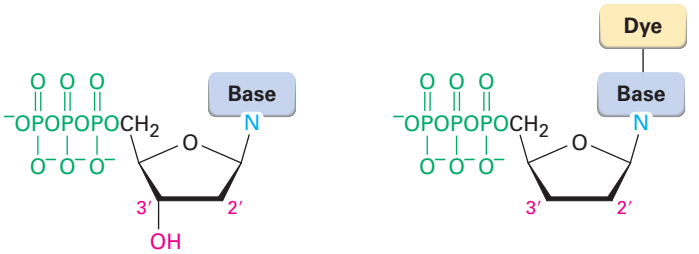
One of the greatest scientific revolutions in history is now under way in molecular biology, as scientists are learning how to manipulate and harness the genetic machinery of organisms. None of the extraordinary advances of the past two decades would have been possible, however, were it not for the discovery in 1977 of methods for sequencing immense DNA chains.

The first step in DNA sequencing is to cleave the enormous chain at known points to produce smaller, more manageable pieces, a task accomplished by the use of *restriction endonucleases*. Each different restriction enzyme, of which more than 3800 are known and approximately 375 are commercially available, cleaves a DNA molecule at a point in the chain where a specific base sequence occurs. For example, the restriction enzyme *Alu*I cleaves between G and C in the four-base sequence AG-CT. Note that the sequence is a *palindrome,* meaning that the sequence (5′)-AGCT-(3′) is the same as its complement (3′)-TCGA-(5′) when both are read in the same 5′ → 3′ direction. The same is true for other restriction endonucleases.

If the original DNA molecule is cut with another restriction enzyme having a different specificity for cleavage, still other segments are produced whose sequences partially overlap those produced by the first enzyme. Sequencing of all the segments, followed by identification of the overlapping regions, allows complete DNA sequencing.

Two methods of DNA sequencing are commonly available, and several others are being developed. The *Maxam–Gilbert method* uses chemical techniques, while the **Sanger dideoxy method** uses enzymatic reactions. The Sanger method is the more frequently used of the two and was the method responsible for sequencing the entire human genome of 3.0 billion base pairs. In commercial sequencing instruments, the dideoxy method begins with a mixture of the following:

- The restriction fragment to be sequenced
- A small piece of DNA called a *primer,* whose sequence is complementary to that on the 3′ end of the restriction fragment
- The four 2′-deoxyribonucleoside triphosphates (dNTPs)
- Very small amounts of the four 2′,3′-dideoxyribonucleoside triphosphates (ddNTPs), each of which is labeled with a fluorescent dye of a different color (A 2′,3′-*dideoxy*ribonucleoside triphosphate is one in which both 2′ and 3′ −OH groups are missing from ribose.)



A 2′-deoxyribonucleoside triphosphate (dNTP)

A 2′,3′-dideoxyribonucleoside triphosphate (ddNTP)

DNA polymerase is added to the mixture, and a strand of DNA complementary to the restriction fragment begins to grow from the end of the primer. Most of the time, only normal deoxyribonucleotides are incorporated into the growing chain because of their much higher concentration in the mixture, but every so often, a dideoxyribonucleotide is incorporated. When that happens, DNA synthesis stops because the chain end no longer has a 3′-hydroxyl group for adding further nucleotides.

When reaction is complete, the product consists of a mixture of DNA fragments of all possible lengths, each terminated by one of the four dye-labeled dideoxyribonucleotides. This product mixture is then separated according to the size of the pieces by gel electrophoresis **(Section 26.2)**, and the identity of the terminal dideoxyribonucleotide in each piece—and thus the sequence of the restriction fragment—is determined by noting the color with which the attached dye fluoresces. **Figure 28.8** shows a typical result.

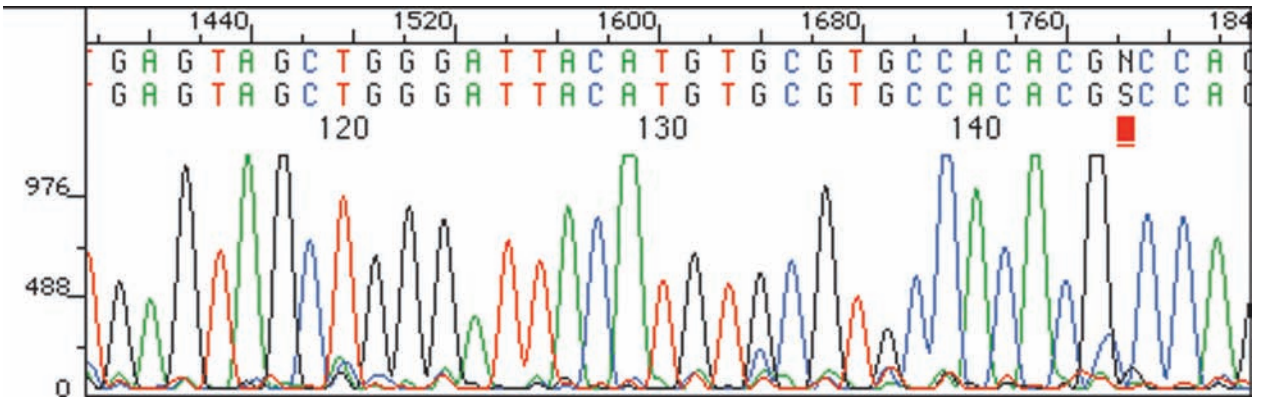


Figure 28.8 The sequence of a restriction fragment determined by the Sanger dideoxy method can be read simply by noting the colors of the dye attached to each of the various terminal nucleotides.

So efficient is the automated dideoxy method that sequences up to 1100 nucleotides in length, with a throughput of up to 19,000 bases per hour, can be sequenced with 98% accuracy. After a decade of work and a cost of about $500 million, preliminary sequence information for the entire human genome of 3.0 billion base pairs was announced early in 2001 and complete information was released in 2003. More recently, the genome sequencing of specific individuals, including that of James Watson, discoverer of the double helix, has been accomplished. The sequencing price per genome is dropping rapidly and is currently approaching $10,000, meaning that the routine sequencing of individuals is within reach.

Remarkably, our genome appears to contain only about 21,000 genes, less than one-fourth the previously predicted number and only about twice the number found in the common roundworm. It's also interesting to note that the number of genes in a human (21,000) is much smaller than the number of kinds of proteins (perhaps 150,000). The discrepancy arises because most proteins are modified in various ways after translation *(posttranslational modifications)*, so a single gene can ultimately give many different proteins.

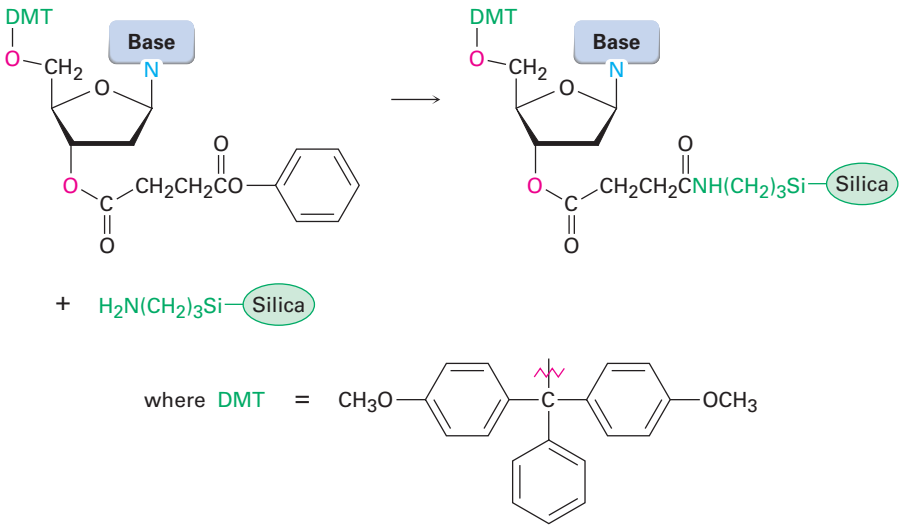
28.7 DNA Synthesis

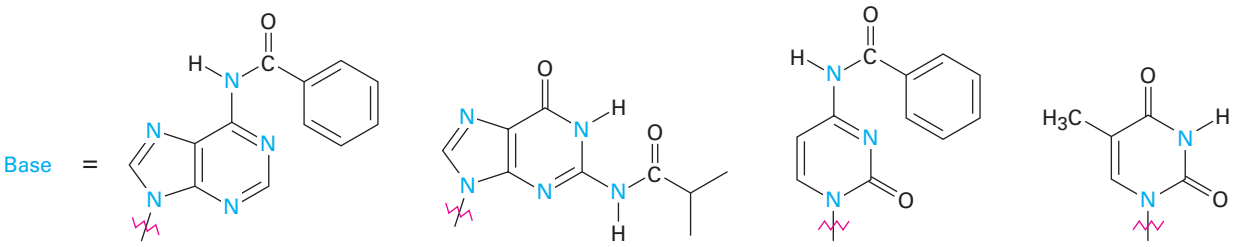
The ongoing revolution in molecular biology has brought with it an increased demand for the efficient chemical synthesis of short DNA segments, called *oligonucleotides*, or simply *oligos*. The problems of DNA synthesis are similar to those of peptide synthesis **(Section 26.7)** but are more difficult because of the complexity of the nucleotide monomers. Each nucleotide has multiple reactive sites that must be selectively protected and deprotected at the proper times, and coupling of the four nucleotides must be carried out in the proper sequence. Automated DNA synthesizers are available, however, that allow the fast and reliable synthesis of DNA segments up to 200 nucleotides in length.

DNA synthesizers operate on a principle similar to that of the Merrifield solid-phase peptide synthesizer **(Section 26.8)**. In essence, a protected nucleotide is covalently bonded to a solid support, and one nucleotide at a time is added to the growing chain by the use of a coupling reagent. After the final nucleotide has been added, all the protecting groups are removed and the synthetic DNA is cleaved from the solid support. Five steps are needed:

STEP 1

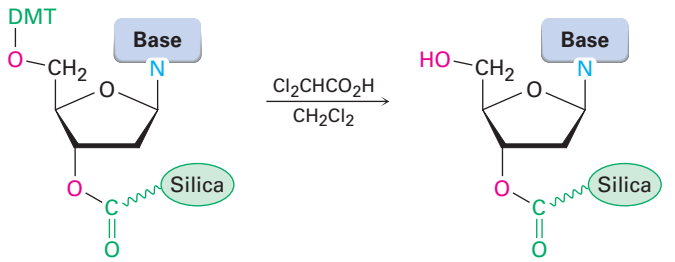
The first step in DNA synthesis is to attach a protected deoxynucleoside to a silica (SiO_2) support by an ester linkage to the 3′ −OH group of the deoxynucleoside. Both the 5′ −OH group on the sugar and free $-NH_2$ groups on the heterocyclic bases must be protected. Adenine and cytosine bases are protected by benzoyl groups, guanine is protected by an isobutyryl group, and thymine requires no protection. The deoxyribose 5′ −OH is protected as its *p*-dimethoxytrityl (DMT) ether.





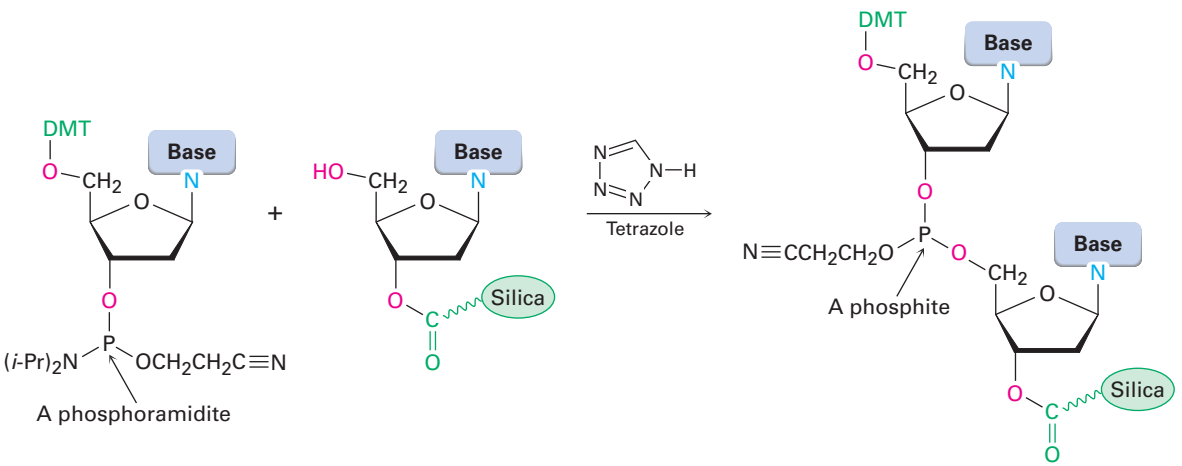
STEP 2

The second step is removal of the DMT protecting group by treatment with dichloroacetic acid in CH_2Cl_2. The reaction occurs by an S_N1 mechanism and proceeds rapidly because of the stability of the tertiary, benzylic dimethoxytrityl cation.



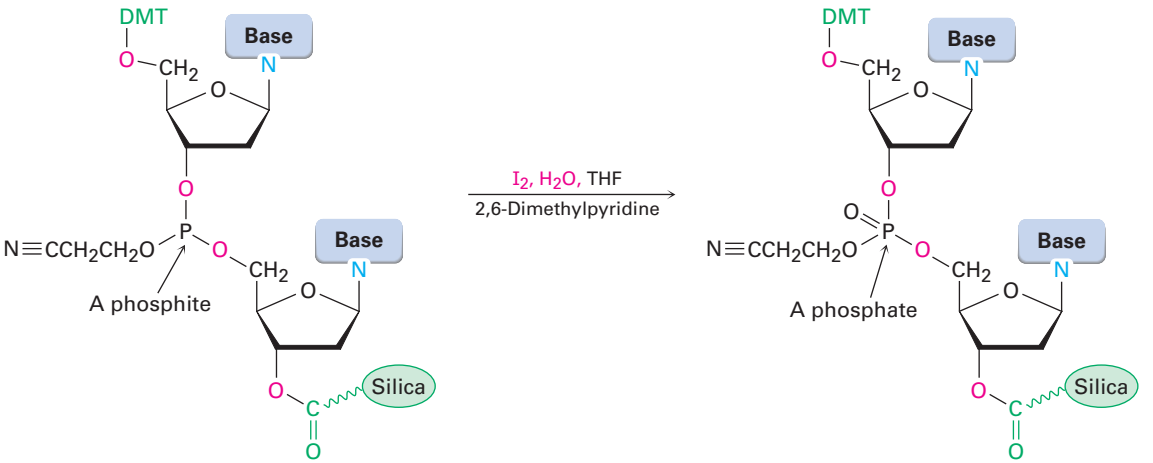
STEP 3

The third step is the coupling of the polymer-bonded deoxynucleoside with a protected deoxynucleoside containing a *phosphoramidite* group [$R_2NP(OR)_2$] at its 3′ position. The coupling reaction takes place in the polar aprotic solvent acetonitrile, requires catalysis by the heterocyclic amine tetrazole, and yields a *phosphite,* $P(OR)_3$, as product. Note that one of the phosphorus oxygen atoms is protected by a β-cyanoethyl group, $-OCH_2CH_2C{\equiv}N$. The coupling step takes place in better than 99% yield.



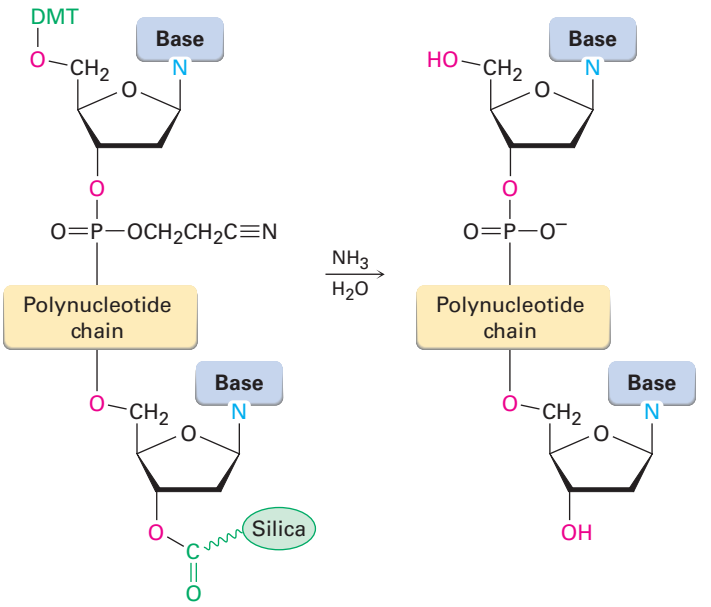
STEP 4

With the coupling accomplished, the phosphite product is oxidized to a phosphate by treatment with iodine in aqueous tetrahydrofuran in the presence of 2,6-dimethylpyridine. The cycle (1) deprotection, (2) coupling, and (3) oxidation is then repeated until an oligonucleotide chain of the desired sequence has been built.



STEP 5

The final step is removal of all protecting groups and cleavage of the ester bond holding the DNA to the silica. All these reactions are done at the same time by treatment with aqueous NH_3. Purification by electrophoresis then yields the synthetic DNA.



Problem 28.11

p-Dimethoxytrityl (DMT) ethers are easily cleaved by mild acid treatment. Show the mechanism of the cleavage reaction.

Problem 28.12

Propose a mechanism to account for cleavage of the β-cyanoethyl protecting group from the phosphate groups on treatment with aqueous ammonia. (Acrylonitrile, $H_2C{=}CHCN$, is a by-product.) What kind of reaction is occurring?

28.8 The Polymerase Chain Reaction

It often happens that only a tiny amount of DNA can be obtained directly, as might occur at a crime scene, so methods for obtaining larger amounts are sometimes needed to carry out the sequencing and characterization. The invention of the **polymerase chain reaction (PCR)** by Kary Mullis in 1986 has been described as being to genes what Gutenberg's invention of the printing press was to the written word. Just as the printing press produces multiple copies of a book, PCR produces multiple copies of a given DNA sequence. Starting from less than 1 picogram of DNA with a chain length of 10,000 nucleotides ($1\ pg = 10^{-12}$ g; about 10^5 molecules), PCR makes it possible to obtain several micrograms ($1\ \mu g = 10^{-6}$ g; about 10^{11} molecules) in just a few hours.

The key to the polymerase chain reaction is *Taq* DNA polymerase, a heat-stable enzyme isolated from the thermophilic bacterium *Thermus aquaticus* found in a hot spring in Yellowstone National Park. *Taq* polymerase is able to take a single strand of DNA that has a short, primer segment of complementary chain at one end and then finish constructing the entire complementary strand. The overall process takes three steps, as shown in **Figure 28.9**. More recently, improved heat-stable DNA polymerases have become available, including Vent polymerase and *Pfu* polymerase, both isolated from bacteria growing near geothermal vents in the ocean floor. The error rate of both enzymes is substantially less than that of *Taq*.

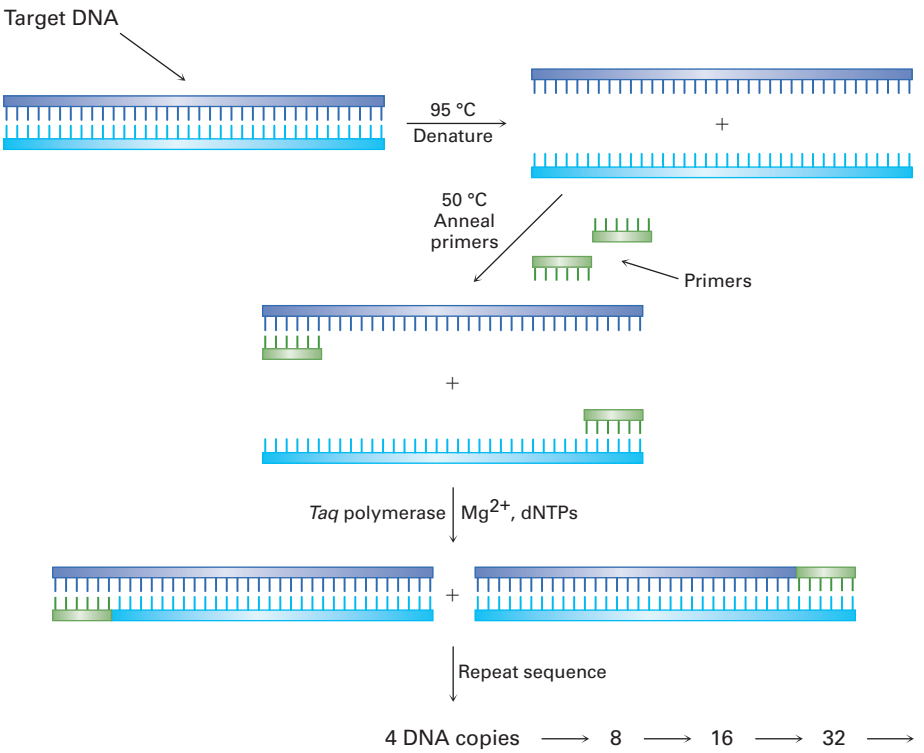


Figure 28.9 The polymerase chain reaction. Details are explained in the text.

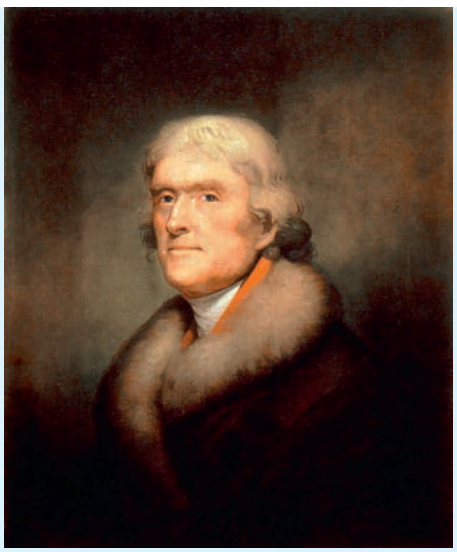
STEP 1
The double-stranded DNA to be amplified is heated in the presence of *Taq* polymerase, Mg^{2+} ion, the four deoxynucleotide triphosphate monomers (dNTPs), and a large excess of two short oligonucleotide primers of about 20 bases each. Each primer is complementary to the sequence at the end of one of the target DNA segments. At a temperature of 95 °C, double-stranded DNA denatures, spontaneously breaking apart into two single strands.

STEP 2
The temperature is lowered to between 37 and 50 °C, allowing the primers, because of their relatively high concentration, to anneal by hydrogen-bonding to their complementary sequence at the end of each target strand.

STEP 3
The temperature is then raised to 72 °C, and *Taq* polymerase catalyzes the addition of further nucleotides to the two primed DNA strands. When replication of each strand is finished, *two* copies of the original DNA now exist. Repeating the denature–anneal–synthesize cycle a second time yields four DNA copies, repeating a third time yields eight copies, and so on, in an exponential series.

PCR has been automated, and 30 or so cycles can be carried out in an hour, resulting in a theoretical amplification factor of 2^{30} ($\sim 10^9$). In practice, however, the efficiency of each cycle is less than 100%, and an experimental amplification of about 10^6 to 10^8 is routinely achieved for 30 cycles.

DNA Fingerprinting | A DEEPER LOOK



Rembrandt Peale (1805)/New York Historical Society

Historians have wondered for many years whether Thomas Jefferson fathered a child by Sally Hemings. DNA fingerprinting evidence obtained in 1998 strongly suggests that he did.

The invention of DNA sequencing has affected society in many ways, few more dramatic than those stemming from the development of *DNA fingerprinting*. DNA fingerprinting arose from the discovery in 1984 that human genes contain short, repeating sequences of noncoding DNA, called *short tandem repeat* (STR) loci. Furthermore, the STR loci are slightly different for every individual, except identical twins. By sequencing these loci, a pattern unique to each person can be obtained.

Perhaps the most common and well-publicized use of DNA fingerprinting is that carried out by crime laboratories to link suspects to biological evidence—blood, hair follicles, skin, or semen—found at a crime scene. Many thousands of court cases have now been decided based on DNA evidence.

For use in criminal cases, forensic laboratories in the United States have agreed on 13 core STR loci that are most accurate for identification of an individual. Based on these 13 loci, a Combined DNA Index System (CODIS) has been established to serve as a registry of convicted offenders. When a DNA sample is obtained from a crime scene, the sample is subjected to cleavage with restriction endonucleases to cut out fragments containing the STR loci, the fragments are amplified using the polymerase chain reaction, and the sequences of the fragments are determined.

(continued)

(continued)

If the profile of sequences from a known individual and the profile from DNA obtained at a crime scene match, the probability is approximately 82 billion to 1 that the DNA is from the same individual. In paternity cases, where the DNA of father and offspring are related but not fully identical, the identity of the father can be established with a probability of around 100,000 to 1. Even after several generations have passed, paternity can still be inferred from DNA analysis of the Y chromosome of direct male-line descendants. The most well-known such case is that of Thomas Jefferson, who likely fathered a child by his slave Sally Hemings. Although Jefferson himself has no male-line descendants, DNA analysis of the male-line descendants of Jefferson's paternal uncle contained the same Y chromosome as a male-line descendant of Eston Hemings, the youngest son of Sally Hemings. Thus, a mixing of the two genomes is clear, although the male individual responsible for that mixing can't be conclusively identified.

Among its many other applications, DNA fingerprinting is widely used for the diagnosis of genetic disorders, both prenatally and in newborns. Cystic fibrosis, hemophilia, Huntington's disease, Tay–Sachs disease, sickle cell anemia, and thalassemia are among the many diseases that can be detected, enabling early treatment of an affected child. Furthermore, by studying the DNA fingerprints of relatives with a history of a particular disorder, it's possible to identify DNA patterns associated with the disease and perhaps obtain clues for eventual cure. In addition, the U.S. Department of Defense now requires blood and saliva samples from all military personnel. The samples are stored, and DNA is extracted if the need for identification of a casualty arises.

Summary

DNA (deoxyribonucleic acid) and **RNA (ribonucleic acid)** are biological polymers that act as chemical carriers of an organism's genetic information. Enzyme-catalyzed hydrolysis of nucleic acids yields **nucleotides**, the monomer units from which RNA and DNA are constructed. Further enzyme-catalyzed hydrolysis of the nucleotides yields **nucleosides** plus phosphate. Nucleosides, in turn, consist of a purine or pyrimidine base linked to C1 of an aldopentose sugar—ribose in RNA and 2-deoxyribose in DNA. The nucleotides are joined by phosphate links between the 5′ phosphate of one nucleotide and the 3′ hydroxyl on the sugar of another nucleotide.

Molecules of DNA consist of two complementary polynucleotide strands held together by hydrogen bonds between heterocyclic bases on the different strands and coiled into a **double helix**. Adenine and thymine form hydrogen bonds to each other, as do cytosine and guanine.

Three processes take place in deciphering the genetic information of DNA:

- **Replication** of DNA is the process by which identical DNA copies are made. The DNA double helix unwinds, complementary deoxyribonucleotides line up in order, and two new DNA molecules are produced.
- **Transcription** is the process by which RNA is produced to carry genetic information from the nucleus to the ribosomes. A short segment of the DNA double helix unwinds, and complementary ribonucleotides line up to produce **messenger RNA (mRNA)**.

Key words

anticodon, 1138
antisense strand, 1136
codon, 1137
deoxyribonucleic acid (DNA), 1128
double helix, 1131
3′ end, 1131
5′ end, 1131
messenger RNA (mRNA), 1135
nucleoside, 1128
nucleotide, 1128
polymerase chain reaction (PCR), 1145
replication, 1133
ribonucleic acid (RNA), 1128
ribosomal RNA (rRNA), 1135
Sanger dideoxy method, 1140

Key words—cont'd

sense strand, 1136
small RNAs, 1135
transcription, 1135
transfer RNA (tRNA), 1135
translation, 1138

- **Translation** is the process by which mRNA directs protein synthesis. Each mRNA is divided into **codons**, ribonucleotide triplets that are recognized by small amino acid–carrying molecules of **transfer RNA (tRNA)**, which deliver the appropriate amino acids needed for protein synthesis.

Sequencing of DNA is carried out by the **Sanger dideoxy method**, and small DNA segments can be synthesized in the laboratory by automated instruments. Small amounts of DNA can be amplified by factors of 10^6 using the **polymerase chain reaction (PCR)**.

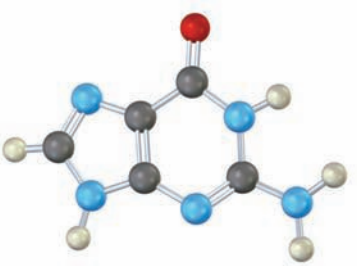
Exercises

OWL Interactive versions of these problems are assignable in OWL for Organic Chemistry.

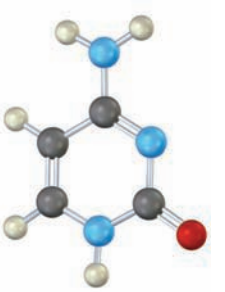
Visualizing Chemistry

(Problems 28.1–28.12 appear within the chapter.)

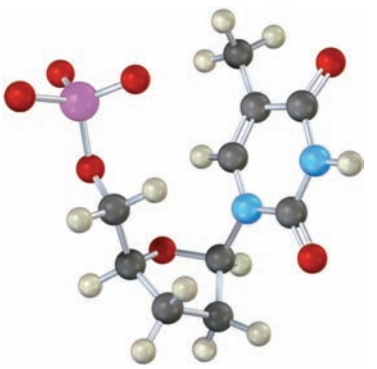
28.13 Identify the following bases, and tell whether each is found in DNA, RNA, or both:

(a) 

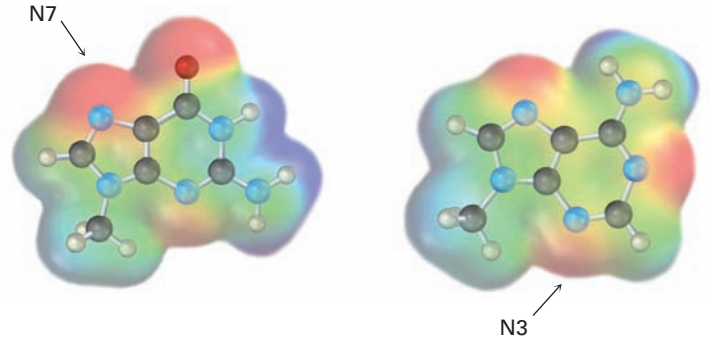
(b) 

(c) 

28.14 Identify the following nucleotide, and tell how it is used:



28.15 Amine bases in nucleic acids can react with alkylating agents in typical S_N2 reactions. Look at the following electrostatic potential maps, and tell which is the better nucleophile, guanine or adenine. The reactive positions in each are indicated.



9-Methylguanine **9-Methyladenine**

Additional Problems

28.16 Human brain natriuretic peptide (BNP) is a small peptide of 32 amino acids used in the treatment of congestive heart failure. How many nitrogen bases are present in the DNA that codes for BNP?

28.17 Human and horse insulin both have two polypeptide chains, with one chain containing 21 amino acids and the other containing 30 amino acids. They differ in primary structure at two places. At position 9 in one chain, human insulin has Ser and horse insulin has Gly; at position 30 in the other chain, human insulin has Thr and horse insulin has Ala. How must the DNA for the two insulins differ?

28.18 The DNA of sea urchins contains about 32% A. What percentages of the other three bases would you expect in sea urchin DNA? Explain.

28.19 The codon UAA stops protein synthesis. Why does the sequence UAA in the following stretch of mRNA not cause any problems?

-GCA-UUC-GAG-GUA-ACG-CCC-

28.20 Which of the following base sequences would most likely be recognized by a restriction endonuclease? Explain.
(a) GAATTC **(b)** GATTACA **(c)** CTCGAG

28.21 For what amino acids do the following ribonucleotide triplets code?
(a) AAU **(b)** GAG **(c)** UCC **(d)** CAU

28.22 From what DNA sequences were each of the mRNA codons in Problem 28.21 transcribed?

28.23 What anticodon sequences of tRNAs are coded for by the codons in Problem 28.21?

28.24 Draw the complete structure of the ribonucleotide codon UAC. For what amino acid does this sequence code?

28.25 Draw the complete structure of the deoxyribonucleotide sequence from which the mRNA codon in Problem 28.24 was transcribed.

28.26 Give an mRNA sequence that will code for synthesis of metenkephalin.

Tyr-Gly-Gly-Phe-Met

28.27 Give an mRNA sequence that will code for the synthesis of angiotensin II.

Asp-Arg-Val-Tyr-Ile-His-Pro-Phe

28.28 What amino acid sequence is coded for by the following DNA coding strand (sense strand)?

(5′) CTT-CGA-CCA-GAC-AGC-TTT (3′)

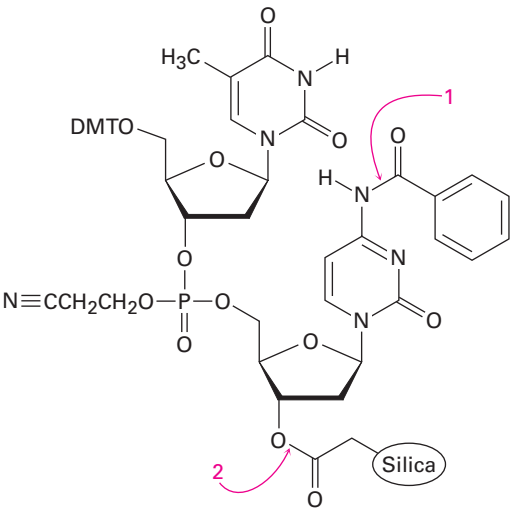
28.29 What amino acid sequence is coded for by the following mRNA base sequence?

(5′) CUA-GAC-CGU-UCC-AAG-UGA (3′)

28.30 If the DNA coding sequence -CAA-CCG-GAT- were miscopied during replication and became -CGA-CCG-GAT-, what effect would there be on the sequence of the protein produced?

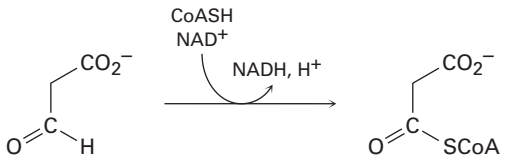
28.31 Show the steps involved in a laboratory synthesis of the DNA fragment with the sequence CTAG.

28.32 The final step in DNA synthesis is deprotection by treatment with aqueous ammonia. Show the mechanisms by which deprotection occurs at the points indicated in the following structure:



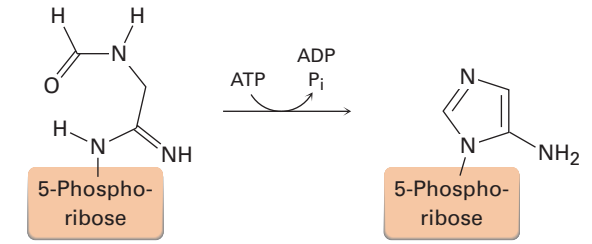
28.33 Draw the structure of cyclic adenosine monophosphate (cAMP), a messenger involved in the regulation of glucose production in the body. Cyclic AMP has a phosphate ring connecting the 3′ and 5′ hydroxyl groups on adenosine.

28.34 The final step in the metabolic degradation of uracil is the oxidation of malonic semialdehyde to give malonyl CoA. Propose a mechanism.



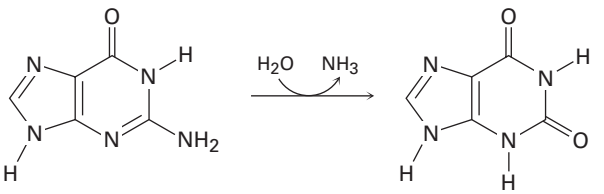
Malonic semialdehyde **Malonyl CoA**

28.35 One of the steps in the biosynthesis of a nucleotide called inosine monophosphate is the formation of aminoimidazole ribonucleotide from formylglycinamidine ribonucleotide. Propose a mechanism.



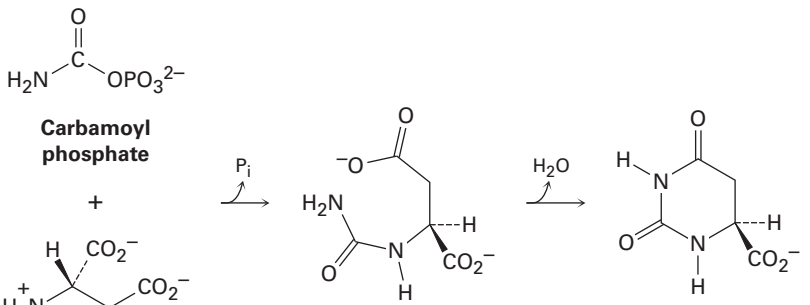
Formylglycinamidine ribonucleotide **Aminoimidazole ribonucleotide**

28.36 One of the steps in the metabolic degradation of guanine is hydrolysis to give xanthine. Propose a mechanism.



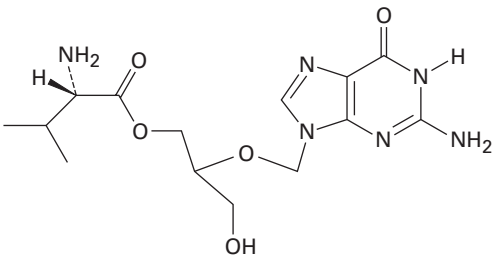
Guanine **Xanthine**

28.37 One of the steps in the biosynthesis of uridine monophosphate is the reaction of aspartate with carbamoyl phosphate to give carbamoyl aspartate followed by cyclization to form dihydroorotate. Propose mechanisms for both steps.



Aspartate **Carbamoyl aspartate** **Dihydroorotate**

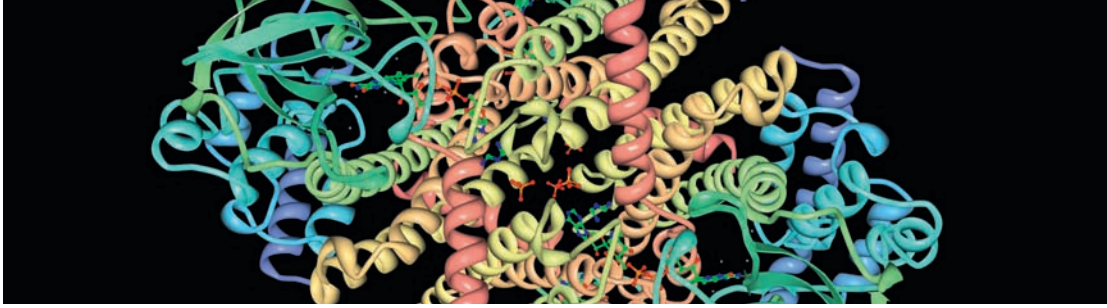
28.38 Valganciclovir, marketed as Valcyte, is an antiviral agent used for the treatment of cytomegalovirus. Called a *prodrug,* valganciclovir is inactive by itself but is rapidly converted in the intestine by hydrolysis of its ester bond to produce an active drug, called ganciclovir, along with an amino acid.



Valganciclovir

(a) What amino acid is produced by hydrolysis of the ester bond in valganciclovir?
(b) What is the structure of ganciclovir?
(c) What atoms present in the nucleotide deoxyguanine are missing from ganciclovir?
(d) What role do the atoms missing from deoxyguanine play in DNA replication?
(e) How might valganciclovir interfere with DNA synthesis?

29



Acyl CoA dehydrogenase is an enzyme that catalyzes the introduction of a C=C double bond into fatty acids during their metabolism. PDB ID: 2WBI. Muniz, J.R.C., Guo, K., Savitsky, P., Roos, A., Yue, W., Pilka, E., Vondelft, F., Edwards, A.M., Bountra, C., Arrowsmith, C.H., Weigelt, J., Oppermann, U. CRYSTAL STRUCTURE OF HUMAN ACYL-COA DEHYDROGENASE 11

The Organic Chemistry of Metabolic Pathways

Anyone who wants to understand or contribute to the revolution now taking place in the biological sciences must first understand life processes at the molecular level. This understanding, in turn, must be based on a detailed knowledge of the chemical reactions and pathways used by living organisms. Just knowing *what* occurs is not enough; it's also necessary to understand *how* and *why* organisms use the chemistry they do.

Biochemical reactions are not mysterious. Even though the biological reactions that take place in living organisms often appear complicated, they follow the same rules of reactivity as laboratory reactions and they take place by the same mechanisms.

A word of caution: some of the molecules we'll be encountering are substantially larger and more complex than those we've been dealing with thus far. But don't be intimidated; keep your focus on the parts of the molecules where changes occur, and ignore the parts where nothing changes. The reactions themselves are exactly the same additions, eliminations, substitutions, carbonyl condensations, and so forth, that we've been dealing with all along. By the end of this chapter, it should be clear that the chemistry of living organisms *is* organic chemistry.

Why This Chapter? In this chapter, we'll look at some of the pathways by which organisms carry out their chemistry, focusing primarily on how they metabolize fats and carbohydrates. The treatment will be far from complete, but it should give you an idea of the kinds of processes that occur.

29.1 An Overview of Metabolism and Biochemical Energy

The many reactions that go on in the cells of living organisms are collectively called **metabolism**. The pathways that break down larger molecules into smaller ones are called **catabolism**, and the pathways that synthesize larger biomolecules from smaller ones are known as **anabolism**. Catabolic reaction pathways are usually exergonic and release energy, while anabolic pathways are

often endergonic and absorb energy. Catabolism can be divided into the four stages shown in **Figure 29.1**.

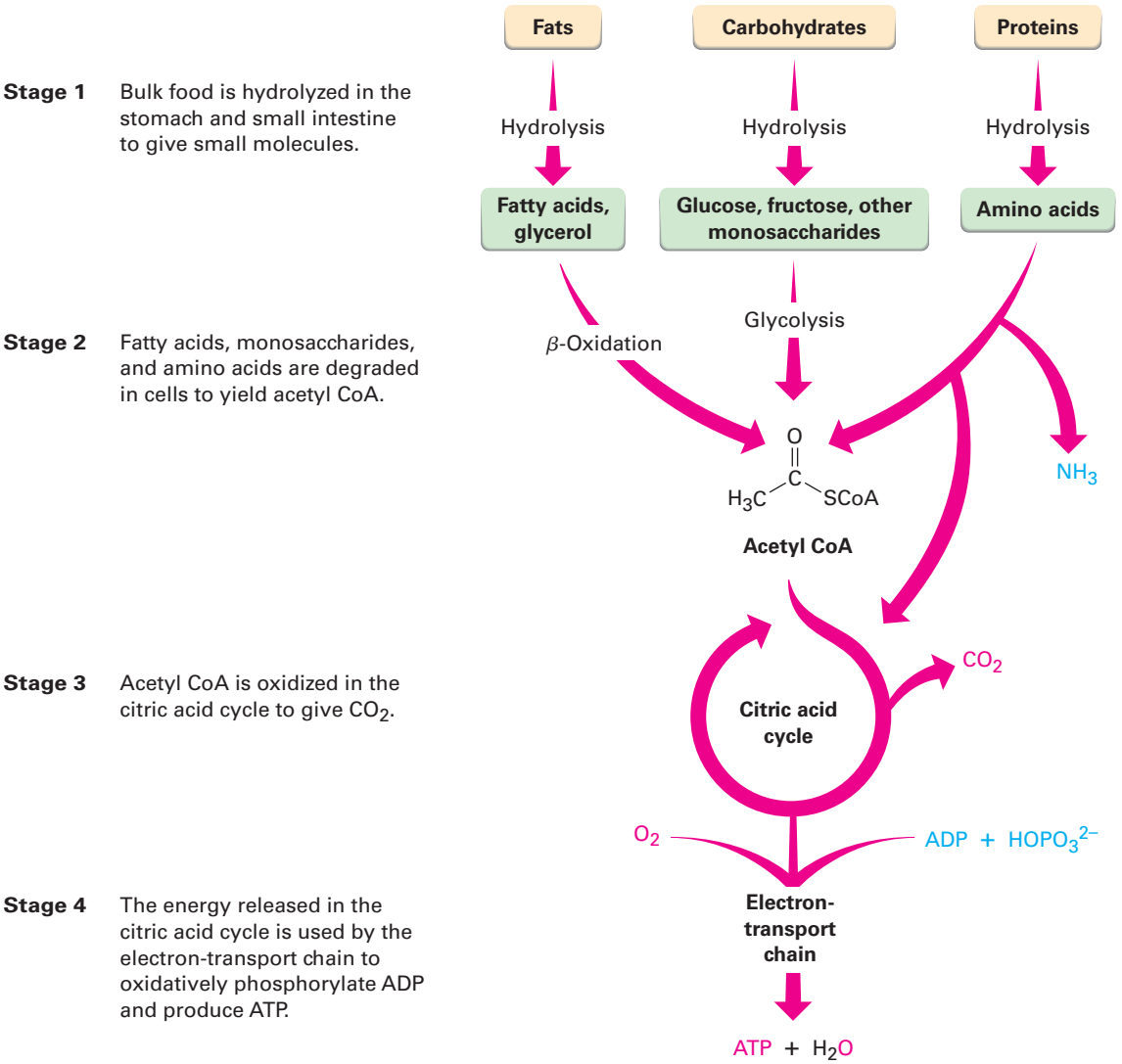
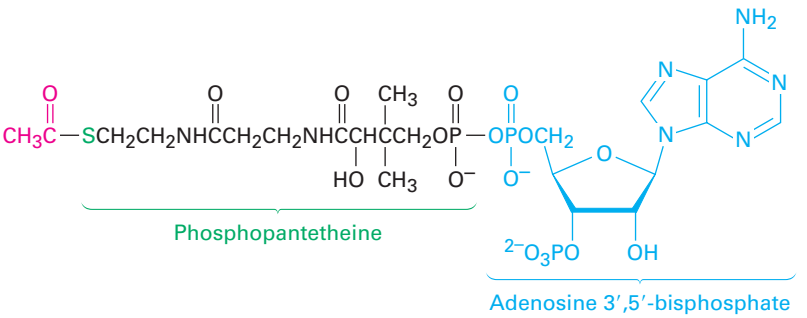


Figure 29.1 An overview of catabolic pathways for the degradation of food and the production of biochemical energy. The ultimate products of food catabolism are CO_2 and H_2O, with the energy released in the citric acid cycle used to drive the endergonic synthesis of adenosine triphosphate (ATP) from adenosine diphosphate (ADP) plus phosphate ion, $HOPO_3^{2-}$.

In the first catabolic stage, commonly called digestion, food is broken down in the mouth, stomach, and small intestine by hydrolysis of ester, acetal (glycoside), and amide (peptide) bonds to yield fatty acids, simple sugars, and amino acids. These smaller molecules are then absorbed and further degraded in the second stage of catabolism to yield acetyl groups attached by a thioester bond to the large carrier molecule coenzyme A. The resultant compound, acetyl coenzyme A (acetyl CoA), is a key substance in the metabolism of food

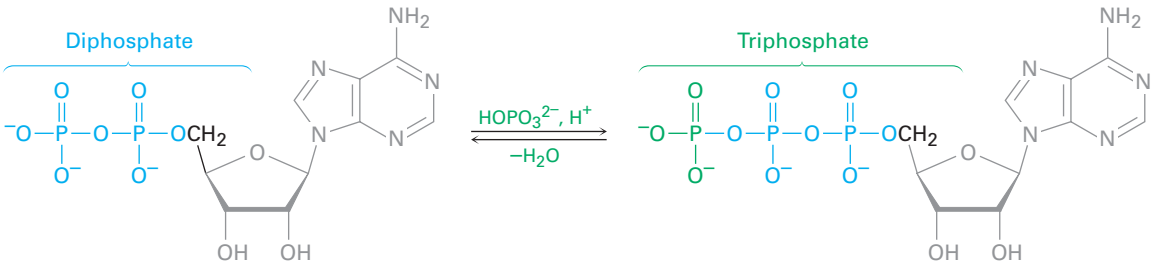
molecules and in many other biological pathways. As noted in **Section 21.8**, the acetyl group in acetyl CoA is linked to the sulfur atom of phosphopantetheine, which is itself linked to adenosine 3′,5′-bisphosphate.



Acetyl CoA—a thioester

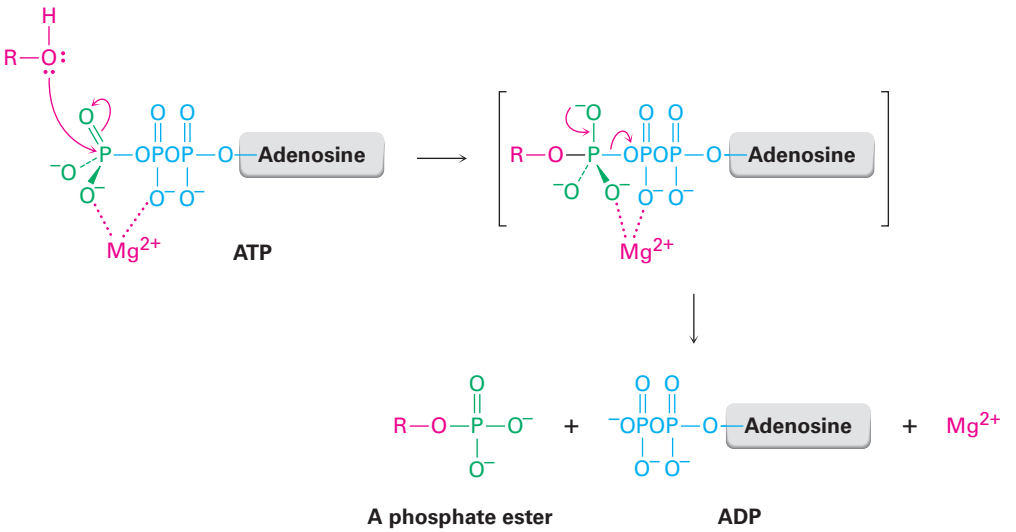
Acetyl groups are oxidized inside cellular mitochondria in the third stage of catabolism, the *citric acid cycle,* to yield CO_2. (We'll see the details of the process in **Section 29.7**.) Like most oxidations, this stage releases a large amount of energy, which is used in the fourth stage, the *electron-transport chain,* to accomplish the endergonic phosphorylation of adenosine diphosphate (ADP) with hydrogen phosphate ion ($HOPO_3^{2-}$, abbreviated P_i) to give adenosine triphosphate (ATP).

As the final result of food catabolism, ATP has been called the "energy currency" of the cell. Catabolic reactions "buy" ATP by using the energy they release to synthesize it from ADP plus hydrogen phosphate ion. Anabolic reactions then spend the ATP by transferring a phosphate group to another molecule, thereby regenerating ADP. Energy production and use in living organisms thus revolves around the ATP ⇄ ADP interconversion.



Adenosine diphosphate (ADP) **Adenosine triphosphate (ATP)**

ADP and ATP are both phosphoric acid anhydrides, which contain —P—O—P— linkages analogous to the —C—O—C— linkage in carboxylic acid anhydrides. Just as carboxylic acid anhydrides react with alcohols by breaking a C–O bond and forming a carboxylic ester, ROCOR′, **(Section 21.5)**, phosphoric acid anhydrides react with alcohols by breaking a P–O bond and forming a phosphate ester, $ROPO_3^{2-}$. The reaction is, in effect, a nucleophilic acyl substitution at phosphorus. Note that phosphorylation reactions with ATP generally require the presence of a divalent metal cation in the enzyme, usually Mg^{2+}, to form a Lewis acid–base complex with the phosphate oxygen atoms and neutralize negative charge.



How does the body use ATP? Recall from **Section 6.7** that the free-energy change ΔG must be negative and energy must be released for a reaction to be favorable and occur spontaneously. If ΔG is positive, the reaction is energetically unfavorable and the process can't occur spontaneously.

For an energetically unfavorable reaction to occur, it must be "coupled" to an energetically favorable reaction so that the overall free-energy change for the two reactions together is favorable. To understand what it means for reactions to be coupled, imagine that reaction 1 does not occur to any reasonable extent because it has a small equilibrium constant and is energetically unfavorable; that is, the reaction has $\Delta G > 0$.

$$(1)\ \mathbf{A} + m \rightleftharpoons \mathbf{B} + n \qquad \Delta G > 0$$

where **A** and **B** are the biochemically "important" substances while m and n are enzyme cofactors, H_2O, or other small molecules.

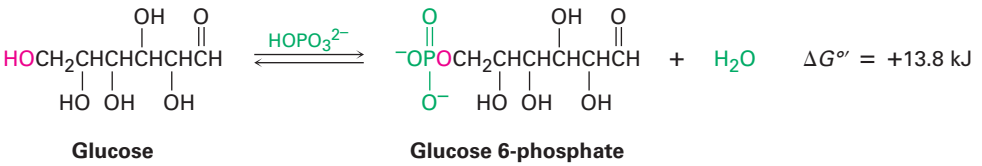
Imagine also that product n can react with substance o to yield p and q in a second, strongly favorable reaction that has a large equilibrium constant and $\Delta G << 0$.

$$(2)\ n + o \rightleftharpoons p + q \qquad \Delta G << 0$$

Taking the two reactions together, they share, or are coupled through, the common intermediate n, which is a product in the first reaction and a reactant in the second. When even a tiny amount of n is formed in reaction 1, it undergoes essentially complete conversion in reaction 2, thereby removing it from the first equilibrium and forcing reaction 1 to continually replenish n until the reactant **A** is gone. That is, the two reactions added together have a favorable $\Delta G < 0$, and we say that the favorable reaction 2 "drives" the unfavorable reaction 1. Because the two reactions are coupled through n, the transformation of **A** to **B** becomes favorable.

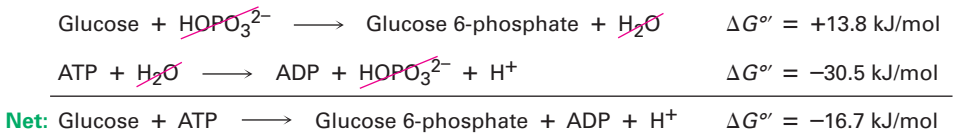
(1)	$\mathbf{A} + m \rightleftharpoons \mathbf{B} + \cancel{n}$	$\Delta G > 0$
(2)	$\cancel{n} + o \rightleftharpoons p + q$	$\Delta G << 0$
Net:	$\mathbf{A} + m + o \rightleftharpoons \mathbf{B} + p + q$	$\Delta G < 0$

As an example of two reactions that are coupled, look at the phosphorylation reaction of glucose to yield glucose 6-phosphate plus water, an important step in the breakdown of dietary carbohydrates.



Glucose **Glucose 6-phosphate**

The reaction of glucose with $HOPO_3^{2-}$ does not occur spontaneously because it is energetically unfavorable, with $\Delta G^{\circ\prime} = +13.8$ kJ/mol. (The standard free-energy change for a biological reaction is denoted $\Delta G^{\circ\prime}$ and refers to a process in which reactants and products have a concentration of 1.0 M in a solution with pH = 7.) At the same time, however, the reaction of water with ATP to yield ADP plus $HOPO_3^{2-}$ is strongly favorable, with $\Delta G^{\circ} = -30.5$ kJ/mol. When the two reactions are coupled, glucose reacts with ATP to yield glucose 6-phosphate plus ADP in a reaction that is favorable by about 16.7 kJ/mol (4.0 kcal/mol). That is, ATP drives the phosphorylation reaction of glucose.



It's this ability to drive otherwise unfavorable phosphorylation reactions that makes ATP so useful. The resultant phosphates are much more reactive as leaving groups in nucleophilic substitutions and eliminations than the alcohols they're derived from and are therefore more chemically useful.

Problem 29.1

One of the steps in fat metabolism is the reaction of glycerol (1,2,3-propanetriol) with ATP to yield glycerol 1-phosphate. Write the reaction, and draw the structure of glycerol 1-phosphate.

29.2 Catabolism of Triacylglycerols: The Fate of Glycerol

The metabolic breakdown of triacylglycerols begins with their hydrolysis in the stomach and small intestine to yield glycerol plus fatty acids. The reaction is catalyzed by a lipase, whose mechanism is shown in **Figure 29.2**. The active site of the enzyme contains a catalytic triad of aspartic acid, histidine, and serine residues, which act cooperatively to provide the necessary acid and base catalysis for the individual steps. Hydrolysis is accomplished by two sequential nucleophilic acyl substitution reactions, one that covalently binds an acyl group to the side chain −OH of a serine residue on the enzyme and a second that frees the fatty acid from the enzyme.

1 The enzyme active site contains an aspartic acid, a histidine, and a serine. First, histidine acts as a base to deprotonate the –OH group of serine, with the negatively charged carboxylate of aspartic acid stabilizing the nearby histidine cation that results. Serine then adds to the carbonyl group of the triacylglycerol, yielding a tetrahedral intermediate.

2 This intermediate expels a diacylglycerol as leaving group in a nucleophilic acyl substitution reaction, giving an acyl enzyme. The diacylglycerol is protonated by the histidine cation.

3 Histidine deprotonates a water molecule, which adds to the acyl group. A tetrahedral intermediate is again formed, and the histidine cation is again stabilized by the nearby carboxylate.

4 The tetrahedral intermediate expels the serine as leaving group in a second nucleophilic acyl substitution reaction, yielding a free fatty acid. The serine accepts a proton from histidine, and the enzyme has now returned to its starting structure.

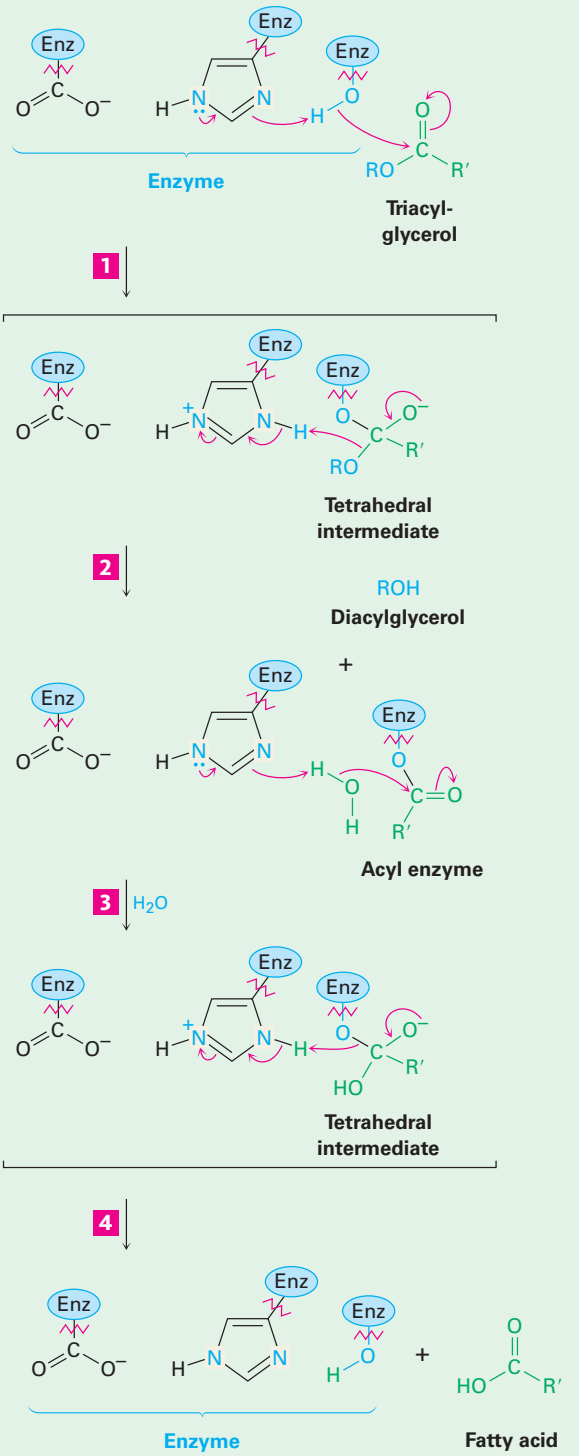
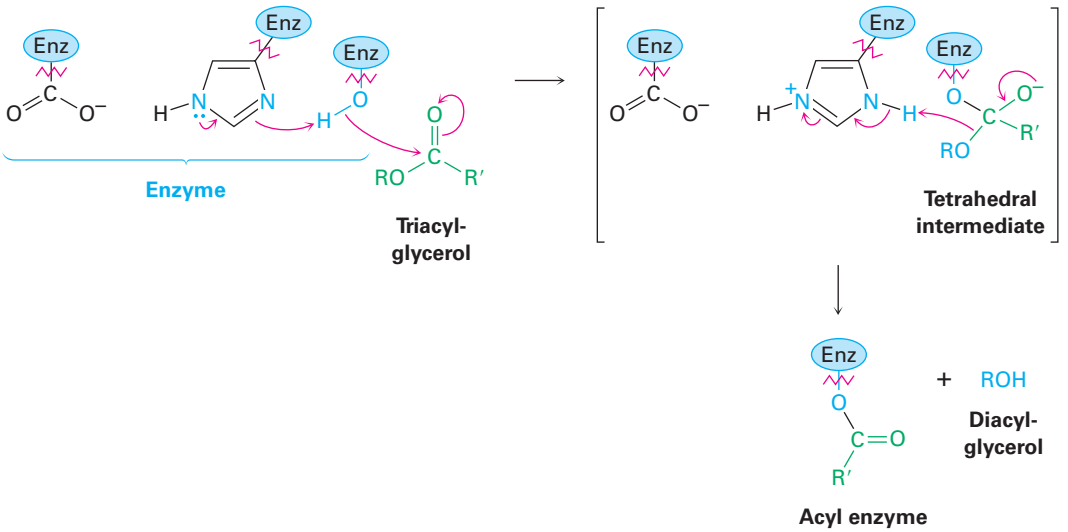


Figure 29.2 | MECHANISM

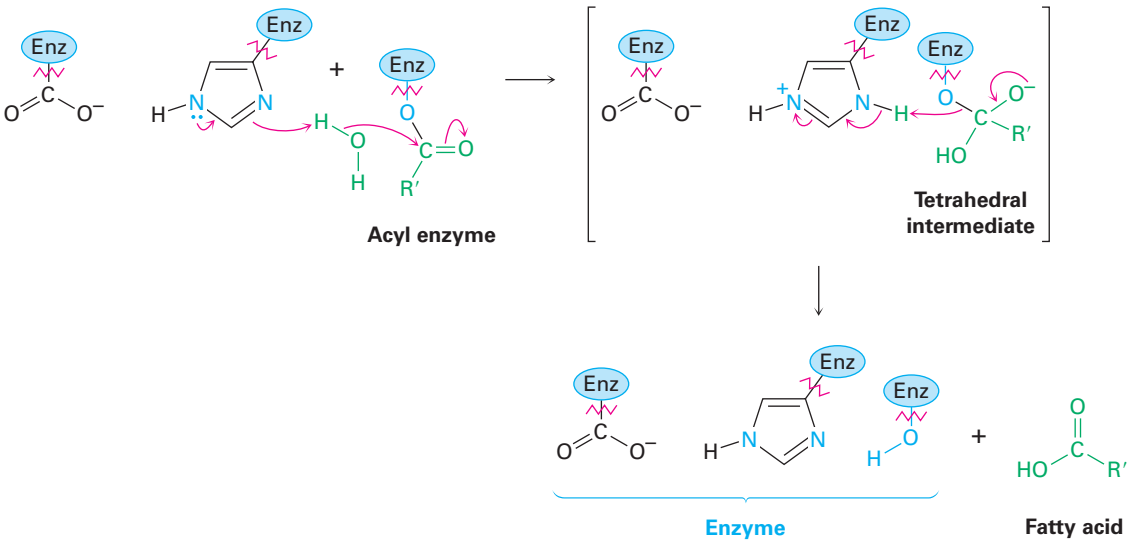
Mechanism of action of lipase. The active site of the enzyme contains a catalytic triad of aspartic acid, histidine, and serine, which react cooperatively to carry out two nucleophilic acyl substitution reactions. Individual steps are explained in the text.

Steps 1–2 of Figure 29.2: Acyl Enzyme Formation The first nucleophilic acyl substitution step—reaction of the triacylglycerol with the active-site serine to give an acyl enzyme—begins with deprotonation of the serine alcohol by histidine to form the more strongly nucleophilic alkoxide ion. This proton transfer is facilitated by a nearby side-chain carboxylate anion of aspartic acid, which makes the histidine more basic and stabilizes the resultant histidine cation by electrostatic interactions. The deprotonated serine adds to a carbonyl group of a triacylglycerol to give a tetrahedral intermediate.

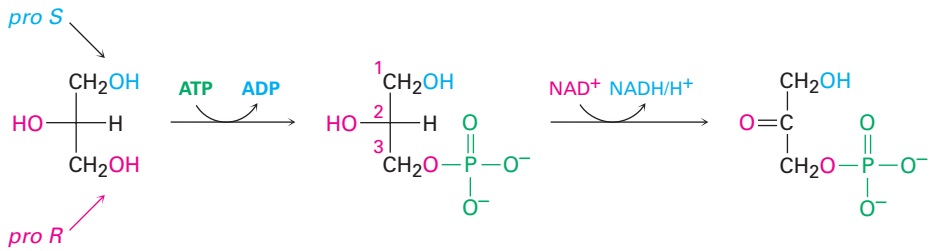
The tetrahedral intermediate expels a diacylglycerol as the leaving group and produces an acyl enzyme. The step is catalyzed by a proton transfer from histidine to make the leaving group a neutral alcohol.



Steps 3–4 of Figure 29.2: Hydrolysis The second nucleophilic acyl substitution step hydrolyzes the acyl enzyme and gives the free fatty acid by a mechanism analogous to that of the first two steps. Water is deprotonated by histidine to give hydroxide ion, which adds to the enzyme-bound acyl group. The tetrahedral intermediate then expels the neutral serine residue as the leaving group, freeing the fatty acid and returning the enzyme to its active form.



The fatty acids released on triacylglycerol hydrolysis are transported to mitochondria and degraded to acetyl CoA, while the glycerol is carried to the liver for further metabolism. In the liver, glycerol is first phosphorylated by reaction with ATP and then oxidized by NAD^+. The dihydroxyacetone phosphate (DHAP) that results enters the carbohydrate glycolysis pathway, which we'll discuss in **Section 29.5**.



You might note that C2 of glycerol is a prochiral center **(Section 5.11)** with two identical "arms." As is typical for enzyme-catalyzed reactions, the phosphorylation of glycerol is selective. Only the *pro-R* arm undergoes reaction, although this can't be predicted in advance.

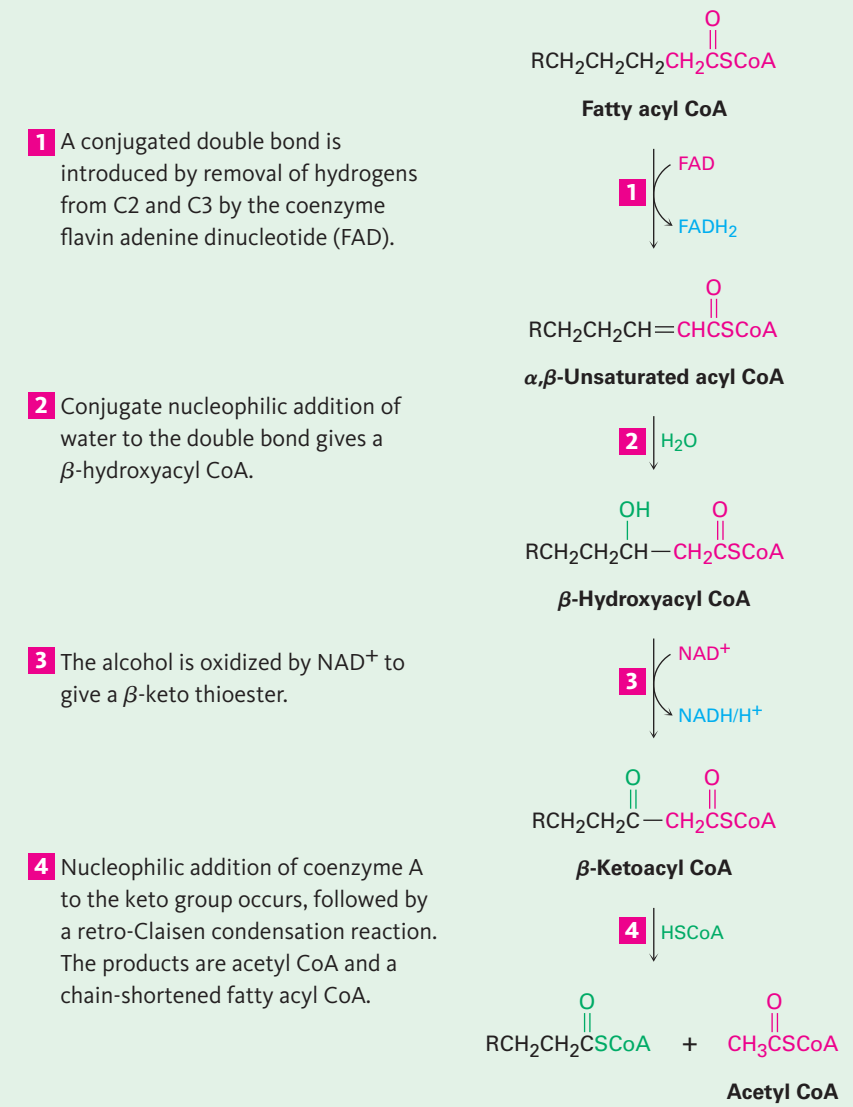
Note also that the phosphorylation product is named *sn*-glycerol 3-phosphate, where the *sn-* prefix means "stereospecific numbering." In this convention, the molecule is drawn in Fischer projection with the –OH group at C2 pointing to the left and the glycerol carbon atoms numbered beginning at the top.

29.3 Catabolism of Triacylglycerols: β-Oxidation

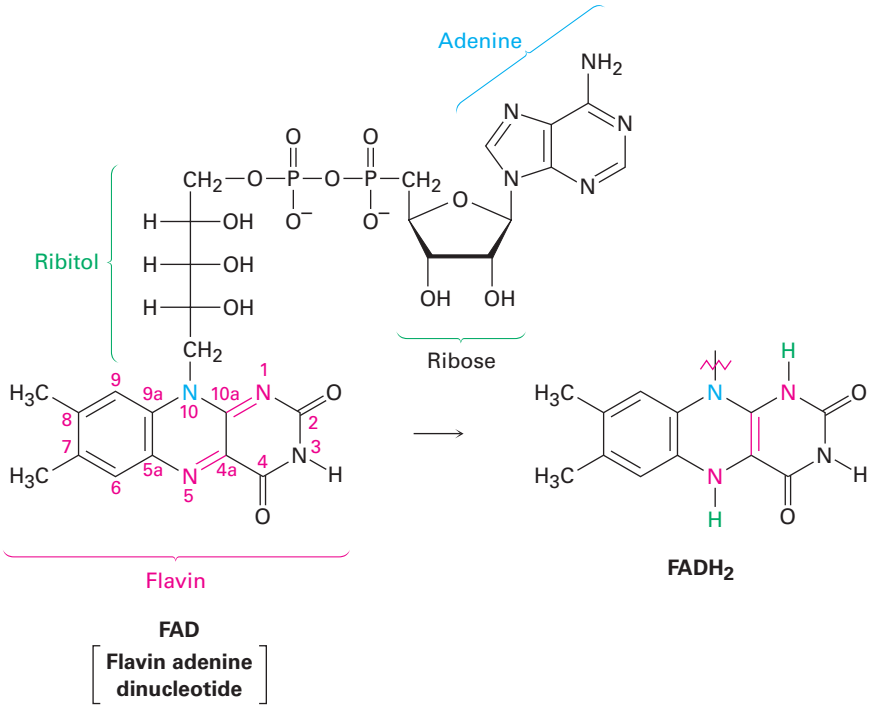
The fatty acids that result from triacylglycerol hydrolysis are converted into thioesters with coenzyme A and then catabolized by a repetitive four-step sequence of reactions called the **β-oxidation pathway**, shown in **Figure 29.3**. Each passage along the pathway results in the cleavage of an acetyl group from the end of the fatty-acid chain, until the entire molecule is ultimately degraded. As each acetyl group is produced, it enters the citric acid cycle and is further degraded to CO_2, as we'll see in **Section 29.7**.

Figure 29.3 | MECHANISM

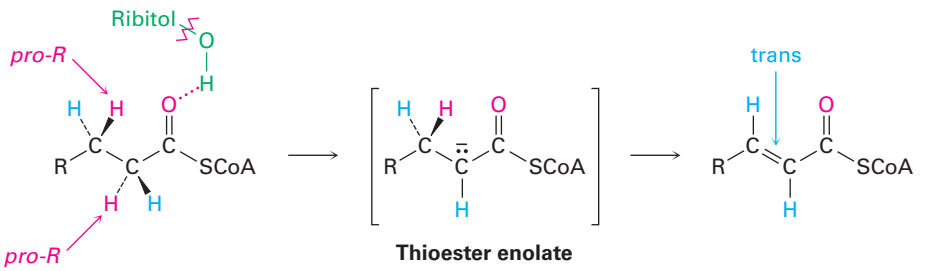
The four steps of the β-oxidation pathway, resulting in the cleavage of an acetyl group from the end of the fatty-acid chain. The key chain-shortening step is a retro-Claisen reaction of a β-keto thioester. Individual steps are explained in the text.



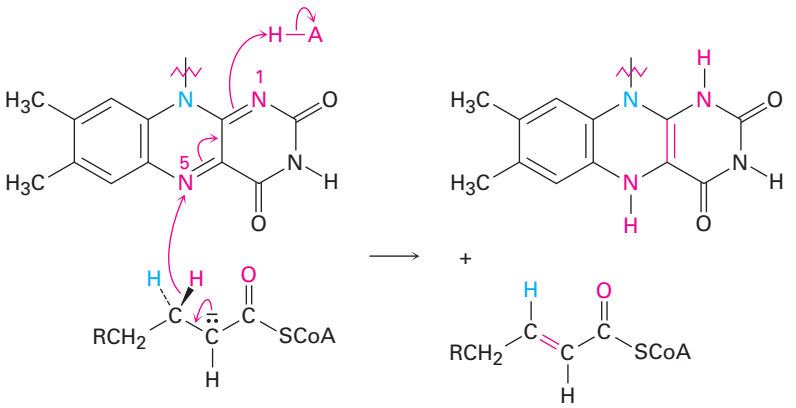
Step 1 of Figure 29.3: Introduction of a Double Bond The β-oxidation pathway begins when two hydrogen atoms are removed from C2 and C3 of the fatty acyl CoA by one of a family of acyl-CoA dehydrogenases to yield an α,β-unsaturated acyl CoA. This kind of oxidation—the introduction of a conjugated double bond into a carbonyl compound—occurs frequently in biochemical pathways and usually involves the coenzyme flavin adenine dinucleotide (FAD). Reduced $FADH_2$ is the by-product.



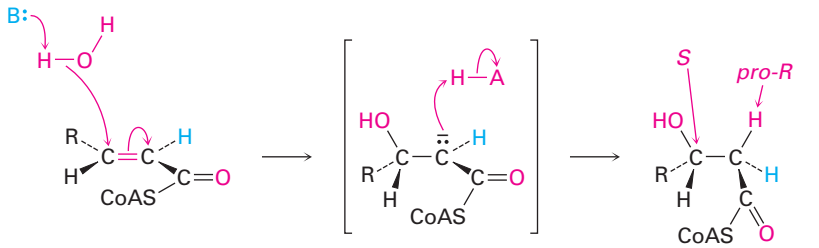
The mechanisms of FAD-catalyzed reactions are often difficult to establish because flavin coenzymes can operate by both two-electron (polar) and one-electron (radical) pathways. As a result, extensive studies of the family of acyl-CoA dehydrogenases have not yet provided a clear picture of how these enzymes function. What is known is that: (1) The first step is abstraction of the *pro-R* hydrogen from the acidic α position of the acyl CoA to give a thioester enolate ion. Hydrogen-bonding between the acyl carbonyl group and the ribitol hydroxyls of FAD increases the acidity of the acyl group. (2) The *pro-R* hydrogen at the β position is transferred to FAD. (3) The α,β-unsaturated acyl CoA that results has a trans double bond.



One suggested mechanism is that the reaction may take place by a conjugate nucleophilic addition of hydride, analogous to what occurs during alcohol oxidations with NAD^+. Electrons on the enolate ion might expel a β hydride ion, which could add to the doubly bonded N5 nitrogen on FAD. Protonation of the intermediate at N1 would give the product.



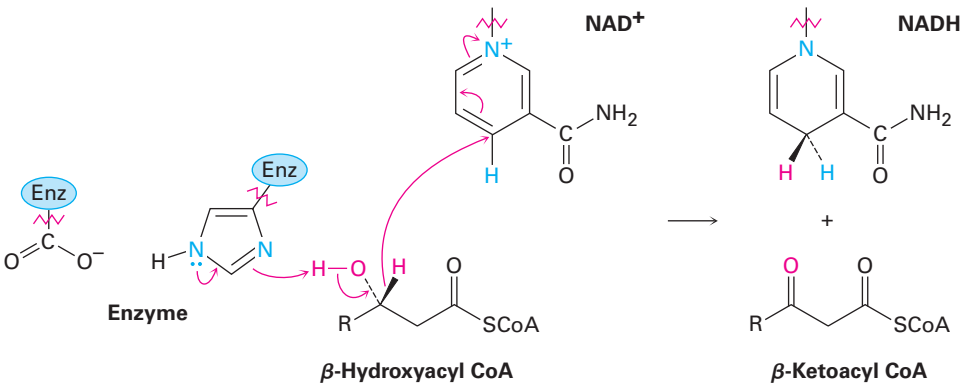
Step 2 of Figure 29.3: Conjugate Addition of Water The α,β-unsaturated acyl CoA produced in step 1 reacts with water by a conjugate addition pathway **(Section 19.13)** to yield a β-hydroxyacyl CoA in a process catalyzed by enoyl CoA hydratase. Water as nucleophile adds to the β carbon of the double bond, yielding an intermediate thioester enolate ion that is protonated on the α position.



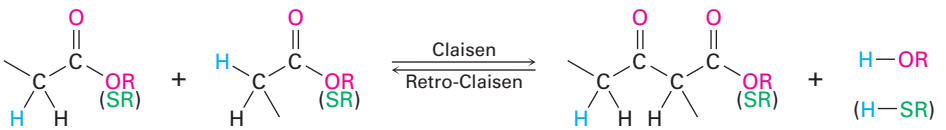
(3*S*)-Hydroxyacyl CoA

Step 3 of Figure 29.3: Alcohol Oxidation The β-hydroxyacyl CoA from step 2 is oxidized to a β-ketoacyl CoA in a reaction catalyzed by one of a family of L-3-hydroxyacyl-CoA dehydrogenases, which differ in substrate specificity

according to the chain length of the acyl group. As in the oxidation of *sn*-glycerol 3-phosphate to dihydroxyacetone phosphate mentioned at the end of **Section 29.2**, this alcohol oxidation requires NAD^+ as a coenzyme and yields reduced $NADH/H^+$ as by-product. Deprotonation of the hydroxyl group is carried out by a histidine residue at the active site.

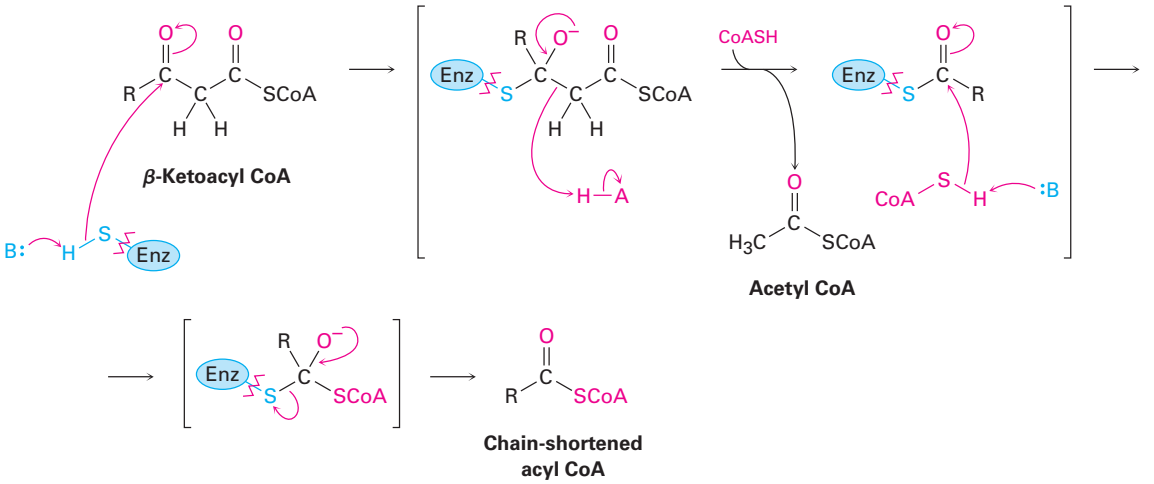


Step 4 of Figure 29.3: Chain Cleavage Acetyl CoA is split off from the chain in the final step of β-oxidation, leaving an acyl CoA that is two carbon atoms shorter than the original. The reaction is catalyzed by β-ketoacyl-CoA thiolase and is mechanistically the reverse of a Claisen condensation reaction **(Section 23.7)**. In the forward direction, a Claisen condensation joins two esters together to form a β-keto ester product. In the reverse direction, a retro-Claisen reaction splits apart a β-keto ester (or β-keto thioester in this case) to form two esters (or two thioesters).



The retro-Claisen reaction occurs by nucleophilic addition of a cysteine –SH group on the enzyme to the keto group of the β-ketoacyl CoA to yield an alkoxide ion intermediate. Cleavage of the C2–C3 bond then follows, with expulsion of an acetyl CoA enolate ion that is immediately protonated. The enzyme-bound acyl group then undergoes nucleophilic acyl substitution by reaction with a molecule of coenzyme A, and the chain-shortened acyl CoA

that results enters another round of the β-oxidation pathway for further degradation.



Look at the catabolism of myristic acid shown in **Figure 29.4** to see the overall results of the β-oxidation pathway. The first passage converts the 14-carbon myristoyl CoA into the 12-carbon lauroyl CoA plus acetyl CoA, the second passage converts lauroyl CoA into the 10-carbon caproyl CoA plus acetyl CoA, the third passage converts caproyl CoA into the 8-carbon capryloyl CoA, and so on. Note that the final passage produces *two* molecules of acetyl CoA because the precursor has four carbons.

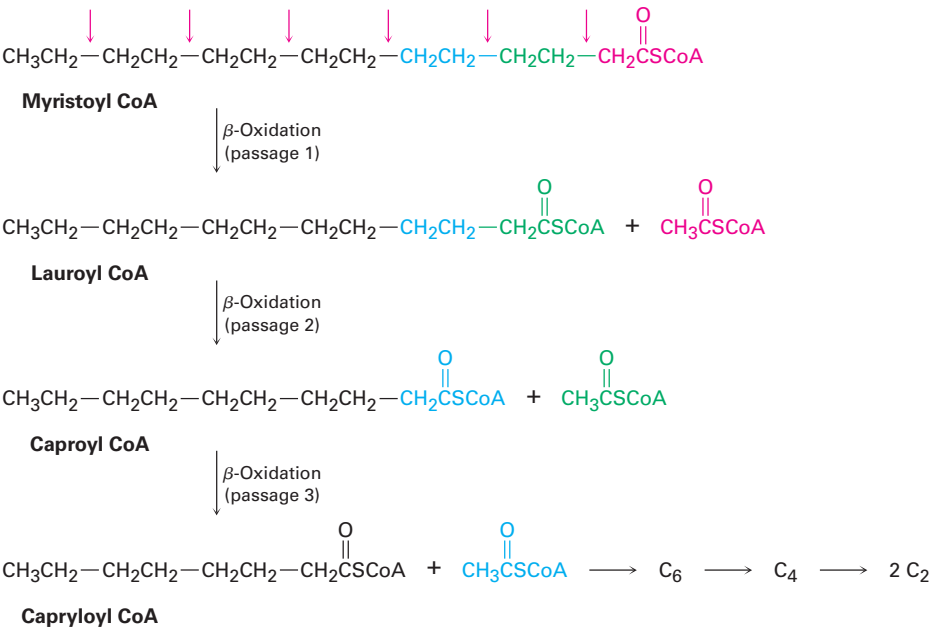


Figure 29.4 Catabolism of the 14-carbon myristic acid by the β-oxidation pathway yields seven molecules of acetyl CoA after six passages.

Most fatty acids have an even number of carbon atoms, so none are left over after β-oxidation. Those fatty acids with an odd number of carbon atoms yield the three-carbon propionyl CoA in the final β-oxidation. Propionyl CoA is then converted to succinate by a multistep radical pathway, and succinate enters the citric acid cycle **(Section 29.7)**. Note that the three-carbon propionyl group should properly be called *propanoyl,* but biochemists generally use the non-systematic name.

Problem 29.2

Write the equations for the remaining passages of the β-oxidation pathway following those shown in Figure 29.4.

Problem 29.3

How many molecules of acetyl CoA are produced by catabolism of the following fatty acids, and how many passages of the β-oxidation pathway are needed?

(a) Palmitic acid, $CH_3(CH_2)_{14}CO_2H$

(b) Arachidic acid, $CH_3(CH_2)_{18}CO_2H$

29.4 Biosynthesis of Fatty Acids

One of the most striking features of the common fatty acids is that they have an even number of carbon atoms (Table 27.1, page 1090).This even number results because all fatty acids are derived biosynthetically from acetyl CoA by sequential addition of two-carbon units to a growing chain. The acetyl CoA, in turn, arises primarily from the metabolic breakdown of carbohydrates in the glycolysis pathway, which we'll see in **Section 29.5**. Thus, dietary carbohydrates consumed in excess of immediate energy needs are turned into fats for storage.

As a general rule in biological chemistry, the anabolic pathway by which a substance is made is not the reverse of the catabolic pathway by which the same substance is degraded. The two paths must differ in some respects for both to be energetically favorable. Thus, the β-oxidation pathway for converting fatty acids into acetyl CoA and the biosynthesis of fatty acids from acetyl CoA are related but are not exact opposites. Differences include the identity of the acyl-group carrier, the stereochemistry of the β-hydroxyacyl reaction intermediate, and the identity of the redox coenzyme. FAD is used to introduce a double bond in β-oxidation, while NADPH is used to reduce the double bond in fatty-acid biosynthesis.

In bacteria, each step in fatty-acid synthesis is catalyzed by a separate enzyme. In vertebrates, however, fatty-acid synthesis is catalyzed by an immense, multienzyme complex called a *synthase* that contains two identical subunits of 2505 amino acids each and catalyzes all steps in the pathway. In fact, for an 18-carbon fatty acid, the synthase catalyzes 42 separate steps! An overview of fatty-acid biosynthesis is shown in **Figure 29.5**.

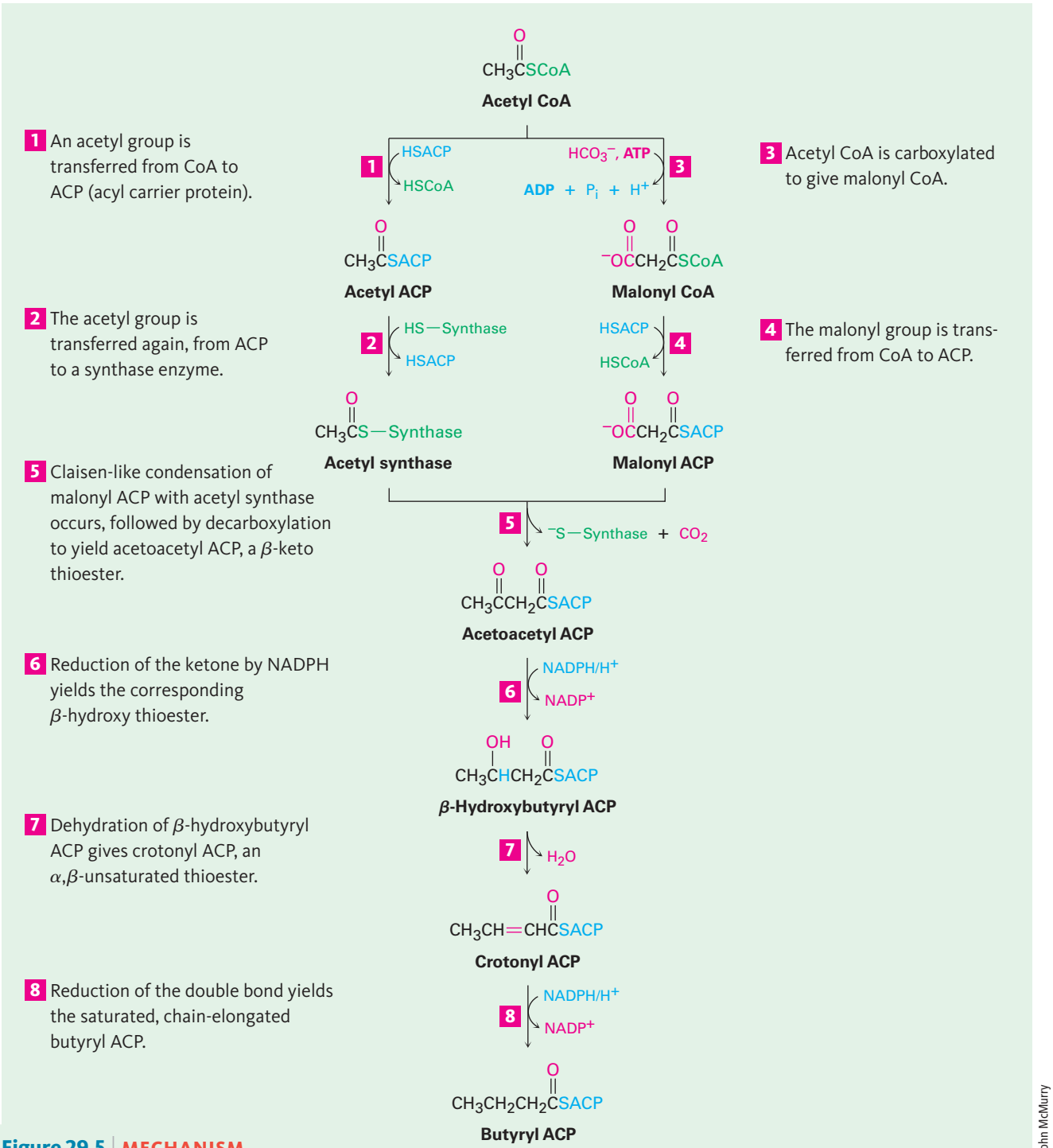
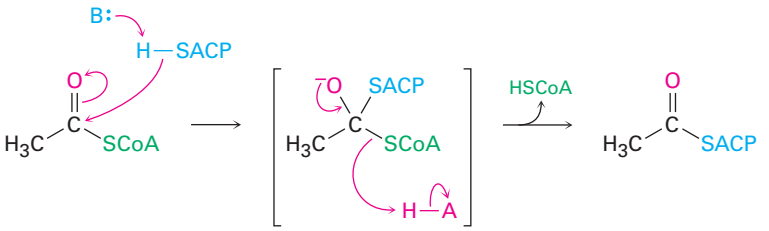


Figure 29.5 | **MECHANISM**

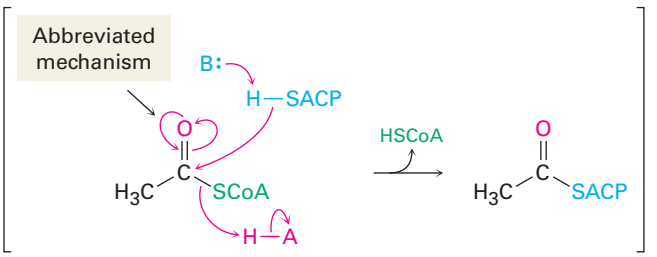
The pathway for fatty-acid biosynthesis from the two-carbon precursor, acetyl CoA. Individual steps are explained in the text.

Steps 1–2 of Figure 29.5: Acyl Transfers The starting material for fatty-acid biosynthesis is the thioester acetyl CoA, the final product of carbohydrate breakdown, as we'll see in **Section 29.6**. The pathway begins with several *priming reactions,* which transport acetyl CoA and convert it into more reactive species. The first priming reaction is a nucleophilic acyl substitution reaction that converts acetyl CoA into acetyl ACP (acyl carrier protein).

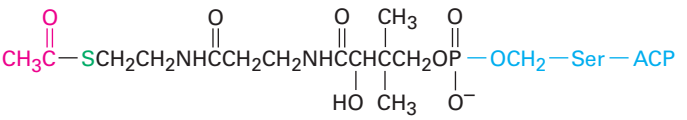
Notice that the mechanism of the nucleophilic acyl substitution step can be given in an abbreviated form that saves space by not explicitly showing the tetrahedral reaction intermediate. Instead, electron movement is shown as a heart-shaped path around the carbonyl oxygen to imply the two steps of the full mechanism. Biochemists use this kind of format commonly, and we'll also use it on occasion in the remainder of this chapter.



Tetrahedral intermediate



In bacteria, ACP is a small protein of 77 residues that transports an acyl group from one enzyme to another. In vertebrates, however, ACP appears to be a long arm on a multienzyme synthase complex, whose apparent function is to shepherd an acyl group from site to site within the complex. As in acetyl CoA, the acyl group in acetyl ACP is linked by a thioester bond to the sulfur atom of phosphopantetheine. The phosphopantetheine is in turn linked to ACP through the side-chain –OH group of a serine residue in the enzyme.



Acetyl ACP

Step 2, another priming reaction, involves a further exchange of thioester linkages by another nucleophilic acyl substitution and results in covalent bonding of the acetyl group to a cysteine residue in the synthase complex that will catalyze the upcoming condensation step.

Steps 3–4 of Figure 29.5: Carboxylation and Acyl Transfer Step 3 is a *loading* reaction in which acetyl CoA is carboxylated by reaction with HCO_3^- and ATP to yield malonyl CoA plus ADP. This step requires the coenzyme biotin, which is bonded to the lysine residue of acetyl CoA carboxylase and acts as a carrier of CO_2. Biotin first reacts with bicarbonate ion to give *N*-carboxybiotin, which then reacts with the enolate ion of acetyl CoA and transfers the CO_2 group. Thus, biotin acts as a carrier of CO_2, binding it in one step and releasing it in another.

The mechanism of the CO_2 transfer reaction with acetyl CoA to give malonyl CoA is thought to involve CO_2 as the reactive species. One proposal is that loss of CO_2 is favored by hydrogen-bond formation between the *N*-carboxybiotin carbonyl group and a nearby acidic site in the enzyme. Simultaneous deprotonation of acetyl CoA by a basic site in the enzyme gives a thioester enolate ion that can react with CO_2 as it is formed **(Figure 29.6)**.

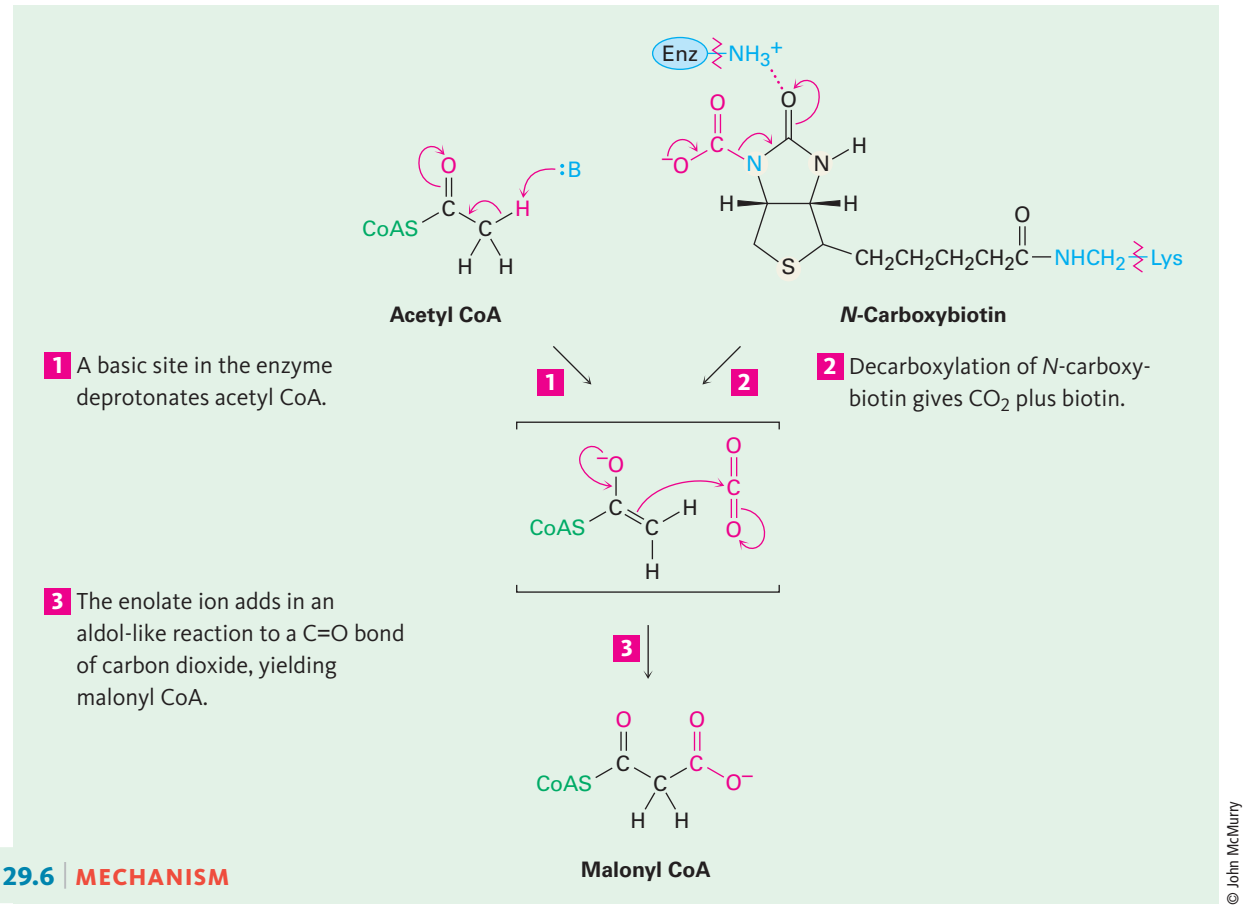
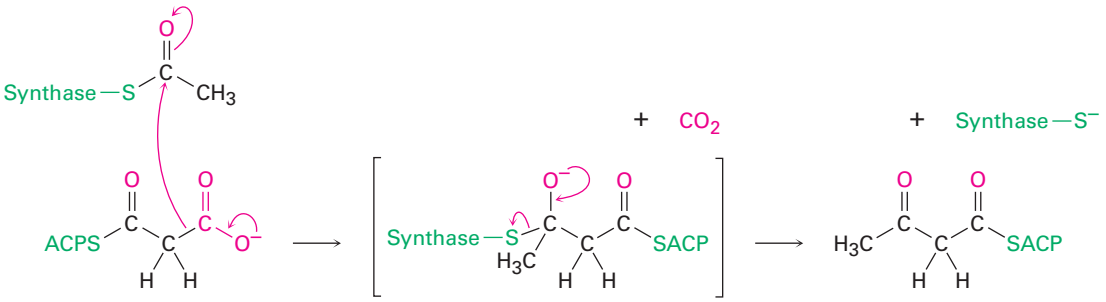


Figure 29.6 | MECHANISM

Mechanism of step 3 in Figure 29.5, the biotin-dependent carboxylation of acetyl CoA to yield malonyl CoA.

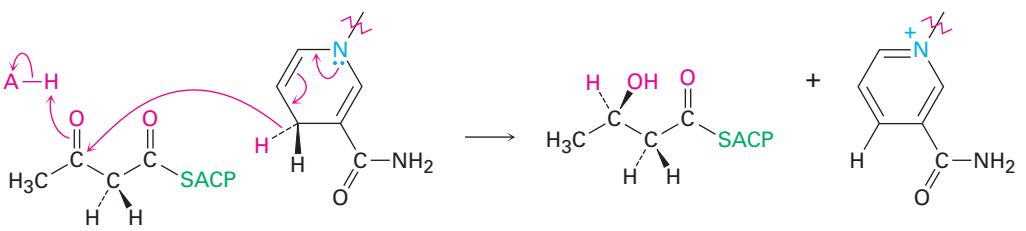
Following the formation of malonyl CoA, another nucleophilic acyl substitution reaction occurs in step 4 to form the more reactive malonyl ACP, thereby binding the malonyl group to an ACP arm of the multienzyme synthase. At this point, both acetyl and malonyl groups are bound to the enzyme, and the stage is set for their condensation.

Step 5 of Figure 29.5: Condensation The key carbon–carbon bond-forming reaction that builds the fatty-acid chain occurs in step 5. This step is simply a Claisen condensation between acetyl synthase as the electrophilic acceptor and malonyl ACP as the nucleophilic donor. The mechanism of the condensation is thought to involve decarboxylation of malonyl ACP to give an enolate ion, followed by immediate nucleophilic addition of the enolate ion to the carbonyl group of acetyl synthase. Breakdown of the tetrahedral intermediate then gives the four-carbon condensation product acetoacetyl ACP and frees the synthase binding site for attachment of the chain-elongated acyl group at the end of the sequence.



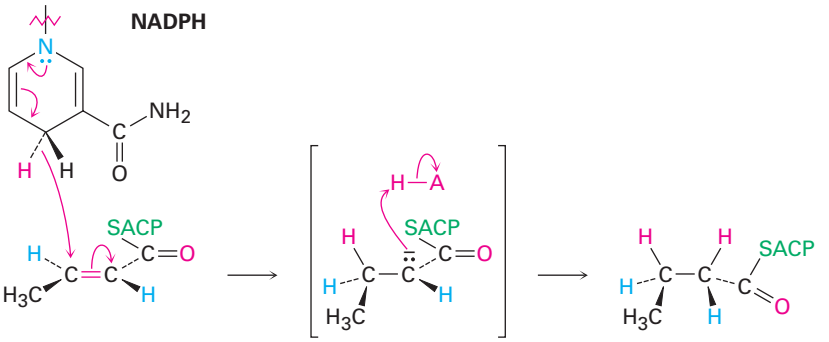
Malonyl ACP **Acetoacetyl ACP**

Steps 6–8 of Figure 29.5: Reduction and Dehydration The ketone carbonyl group in acetoacetyl ACP is next reduced to the alcohol β-hydroxybutyryl ACP by β-keto thioester reductase and NADPH, a reducing coenzyme closely related to NADH. *R* Stereochemistry results at the newly formed chirality center in the β-hydroxy thioester product. (Note that the systematic name of a butyryl group is *butanoyl*.)



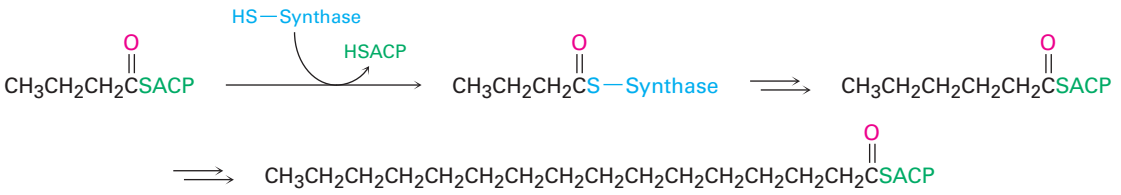
Acetoacetyl ACP **NADPH** **β-Hydroxybutyryl ACP** **$NADP^+$**

Subsequent dehydration of β-hydroxybutyryl ACP by an E1cB reaction in step 7 yields *trans*-crotonyl ACP, and the carbon–carbon double bond of crotonyl ACP is reduced by NADPH in step 8 to yield butyryl ACP. The double-bond reduction occurs by conjugate nucleophilic addition of a hydride ion from NADPH to the β carbon of *trans*-crotonyl ACP. In vertebrates, the reduction occurs by an overall *syn* addition, but other organisms carry out similar chemistry with different stereochemistry.



Crotonyl ACP **Butyryl ACP**

The net effect of the eight steps in the fatty-acid biosynthesis pathway is to take two 2-carbon acetyl groups and combine them into a 4-carbon butyryl group. Further condensation of the butyryl group with another malonyl ACP yields a 6-carbon unit, and still further repetitions of the pathway add two more carbon atoms to the chain each time until the 16-carbon palmitoyl ACP is reached.



Palmitoyl ACP

Further chain elongation of palmitic acid occurs by reactions similar to those just described, but CoA rather than ACP is the carrier group, and separate enzymes are needed for each step rather than a multienzyme synthase complex.

Problem 29.4

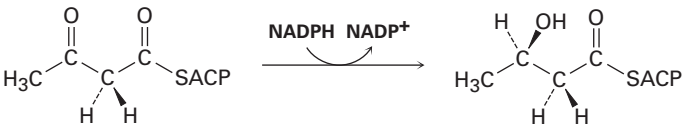
Write a mechanism for the dehydration reaction of β-hydroxybutyryl ACP to yield crotonyl ACP in step 7 of fatty-acid synthesis.

Problem 29.5

Evidence for the role of acetate in fatty-acid biosynthesis comes from isotope-labeling experiments. If acetate labeled with ^{13}C in the methyl group ($^{13}CH_3CO_2H$) were incorporated into fatty acids, at what positions in the fatty-acid chain would you expect the ^{13}C label to appear?

Problem 29.6

Does the reduction of acetoacetyl ACP in step 6 occur on the *Re* face or the *Si* face of the molecule?



Acetoacetyl ACP **β-Hydroxybutyryl ACP**

29.5 Catabolism of Carbohydrates: Glycolysis

Glucose is the body's primary short-term energy source. Its catabolism begins with **glycolysis**, a series of ten enzyme-catalyzed reactions that break down glucose into 2 equivalents of pyruvate, $CH_3COCO_2^-$. The steps of glycolysis, also called the *Embden–Meyerhoff pathway* after its discoverers, are summarized in **Figure 29.7**.

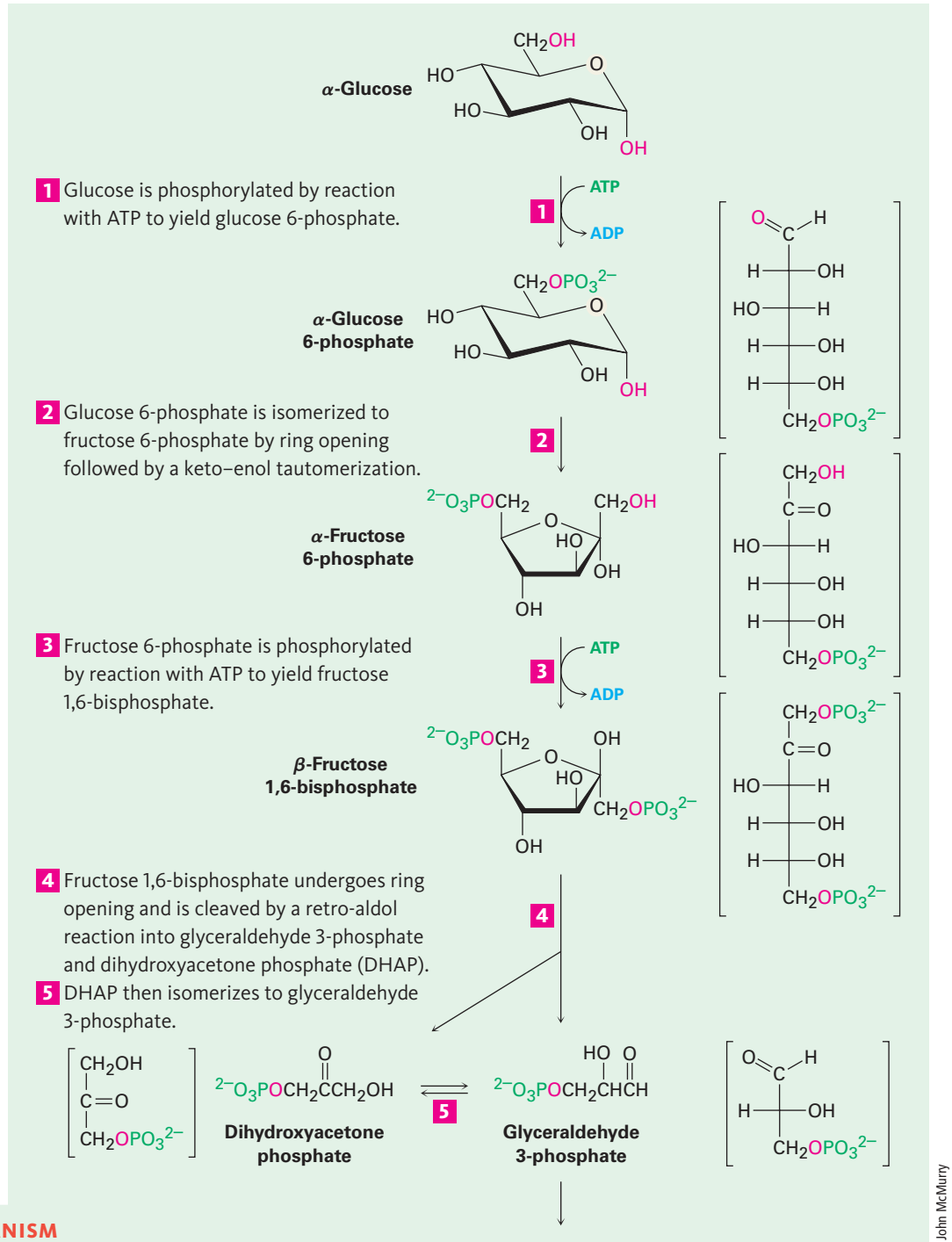


Figure 29.7 | MECHANISM

The ten-step glycolysis pathway for catabolizing glucose to two molecules of pyruvate. Individual steps are described in the text.

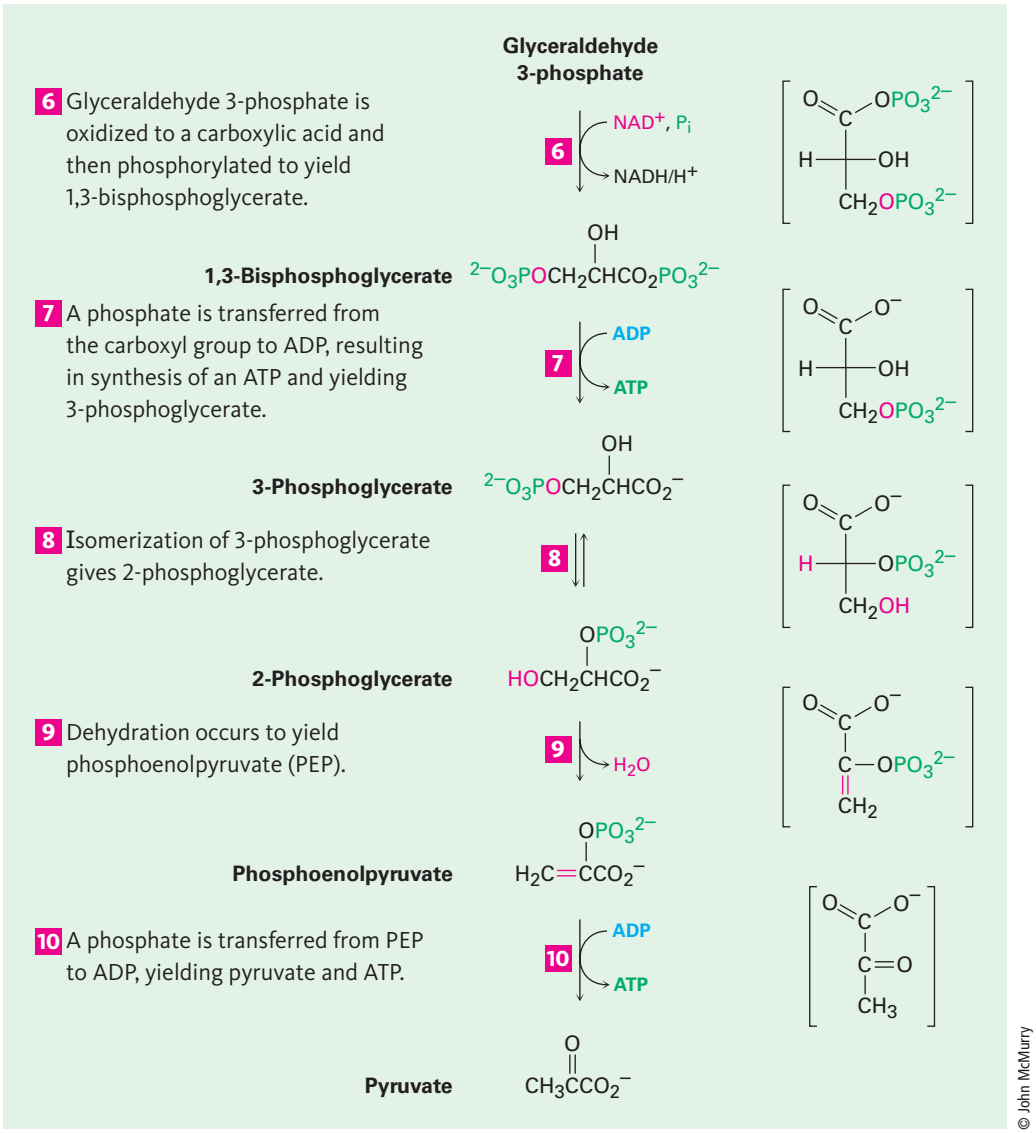


Figure 29.7 (Continued)

Steps 1–2 of Figure 29.7: Phosphorylation and Isomerization Glucose, produced by the digestion of dietary carbohydrates, is phosphorylated at the C6 hydroxyl group by reaction with ATP in a process catalyzed by hexokinase. As noted in **Section 29.1**, the reaction requires Mg^{2+} as a cofactor to complex with the negatively charged phosphate oxygens. The glucose 6-phosphate that results is then isomerized by glucose 6-phosphate isomerase to give fructose 6-phosphate. The isomerization takes place by initial opening of the

glucose hemiacetal ring to the open-chain form, followed by keto–enol tautomerization to a cis enediol, HO—C=C—OH. But because glucose and fructose share a common enediol, further tautomerization to a different keto form produces open-chain fructose, and cyclization completes the process **(Figure 29.8)**.

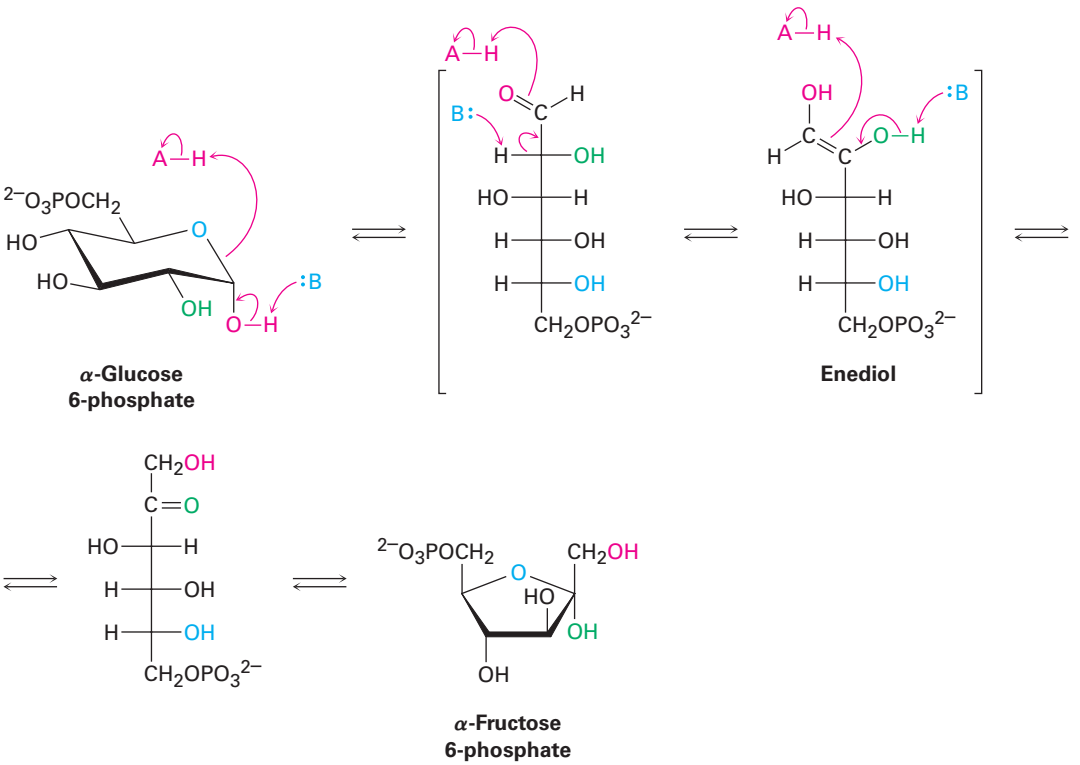
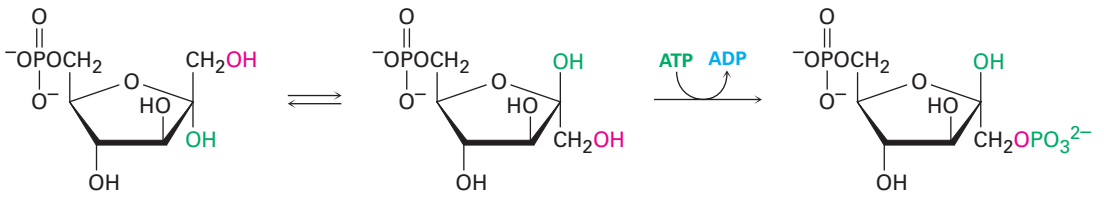


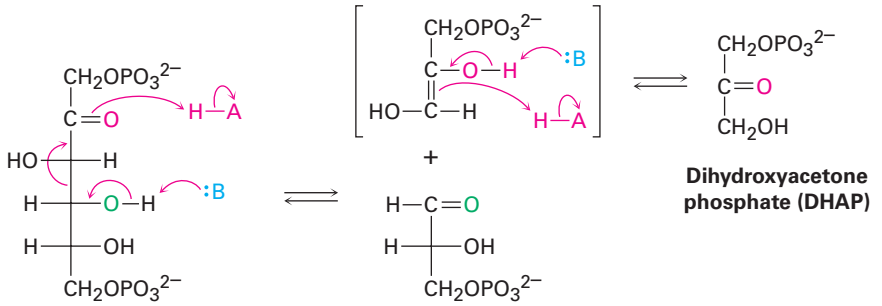
Figure 29.8 Mechanism of step **2** in glycolysis, the isomerization of glucose 6-phosphate to fructose 6-phosphate.

Step 3 of Figure 29.7: Phosphorylation Fructose 6-phosphate is converted in step 3 to fructose 1,6-bisphosphate (FBP) by a phosphofructokinase-catalyzed reaction with ATP (recall that the prefix *bis-* means two). The mechanism is similar to that in step 1, with Mg^{2+} ion again required as cofactor. Interestingly, the product of step 2 is the α anomer of fructose 6-phosphate, but it is the β anomer that is phosphorylated in step 3, implying that the two anomers equilibrate rapidly through the open-chain form. The result is a molecule ready to be split into the two three-carbon intermediates that will ultimately become two molecules of pyruvate.



α-Fructose 6-phosphate **β-Fructose 6-phosphate** **β-Fructose 1,6-bisphosphate (FBP)**

Step 4 of Figure 29.7: Cleavage Fructose 1,6-bisphosphate is cleaved in step 4 into two 3-carbon pieces, dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP). The bond between C3 and C4 of fructose 1,6-bisphosphate breaks, and a C=O group is formed at C4. Mechanistically, the cleavage is the reverse of an aldol reaction **(Section 23.1)** and is catalyzed by an aldolase. A forward aldol reaction joins two aldehydes or ketones to give a β-hydroxy carbonyl compound, while a retro-aldol reaction such as that occurring here cleaves a β-hydroxy carbonyl compound into two aldehydes or ketones.



Fructose 1,6-bisphosphate **Glyceraldehyde 3-phosphate (GAP)** **Dihydroxyacetone phosphate (DHAP)**

Two classes of aldolases are used by organisms for catalysis of the retro-aldol reaction. In fungi, algae, and some bacteria, the retro-aldol reaction is catalyzed by class II aldolases, which function by coordination of the fructose carbonyl group with Zn^{2+} as Lewis acid. In plants and animals, the reaction is catalyzed by class I aldolases and does not take place on the free ketone. Instead, fructose 1,6-bisphosphate undergoes reaction with the side-chain $-NH_2$ group of a lysine residue on the aldolase to yield a protonated enzyme-bound imine **(Section 19.8)**, often called a **Schiff base** in biochemistry.

Because of its positive charge, the iminium ion is a better electron acceptor than a ketone carbonyl group. Retro-aldol reaction ensues, giving glyceraldehyde 3-phosphate and an enamine, which is protonated to give another iminium ion that is hydrolyzed to yield dihydroxyacetone phosphate **(Figure 29.9)**.

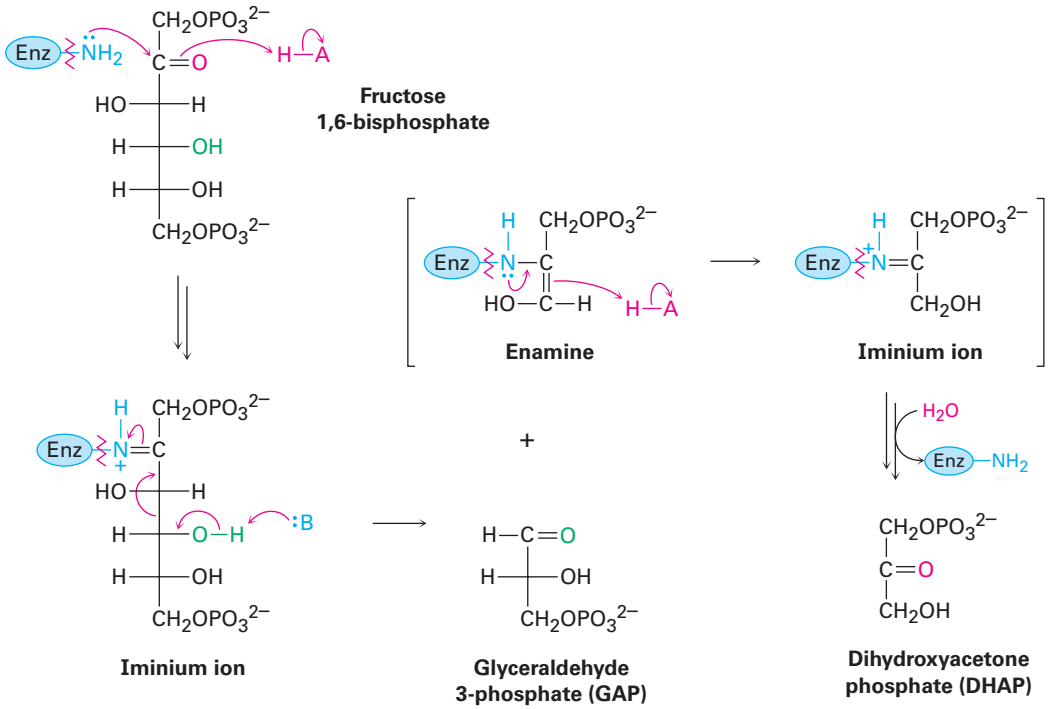
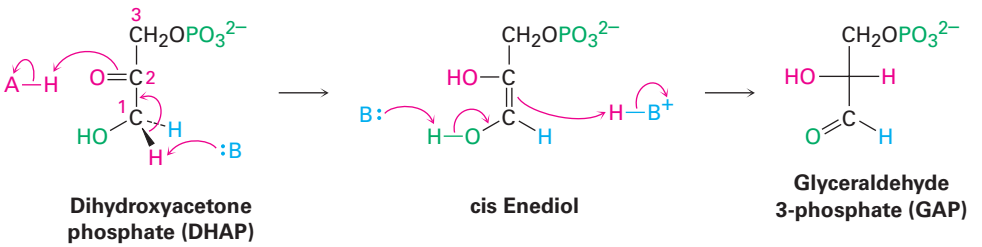


Figure 29.9 Mechanism of step **4** in Figure 29.7, the cleavage of fructose 1,6-bisphosphate to yield glyceraldehyde 3-phosphate and dihydroxyacetone phosphate. The reaction occurs through an iminium ion formed by reaction with a lysine residue in the enzyme.

Step 5 of Figure 29.7: Isomerization Dihydroxyacetone phosphate is isomerized in step 5 by triose phosphate isomerase to form a second equivalent of glyceraldehyde 3-phosphate. As in the conversion of glucose 6-phosphate to fructose 6-phosphate in step 2, the isomerization takes place by keto–enol tautomerization through a common enediol intermediate. A base deprotonates C1 and then reprotonates C2 using the same hydrogen. The net result of steps 4 and 5 together is the production of two glyceraldehyde 3-phosphate molecules, both of which pass down the rest of the pathway. Thus, each of the remaining five steps of glycolysis takes place twice for every glucose molecule that enters at step 1.



Steps 6–7 of Figure 29.7: Oxidation, Phosphorylation, and Dephosphorylation Glyceraldehyde 3-phosphate is oxidized and phosphorylated in step 6 to give 1,3-bisphosphoglycerate **(Figure 29.10)**. The reaction is catalyzed by

glyceraldehyde 3-phosphate dehydrogenase and begins by nucleophilic addition of the –SH group of a cysteine residue in the enzyme to the aldehyde carbonyl group to yield a *hemithioacetal,* the sulfur analog of a hemiacetal. Oxidation of the hemithioacetal –OH group by NAD^+ then yields a thioester, which reacts with phosphate ion in a nucleophilic acyl substitution step to yield 1,3-bisphosphoglycerate, a mixed anhydride between a carboxylic acid and phosphoric acid.

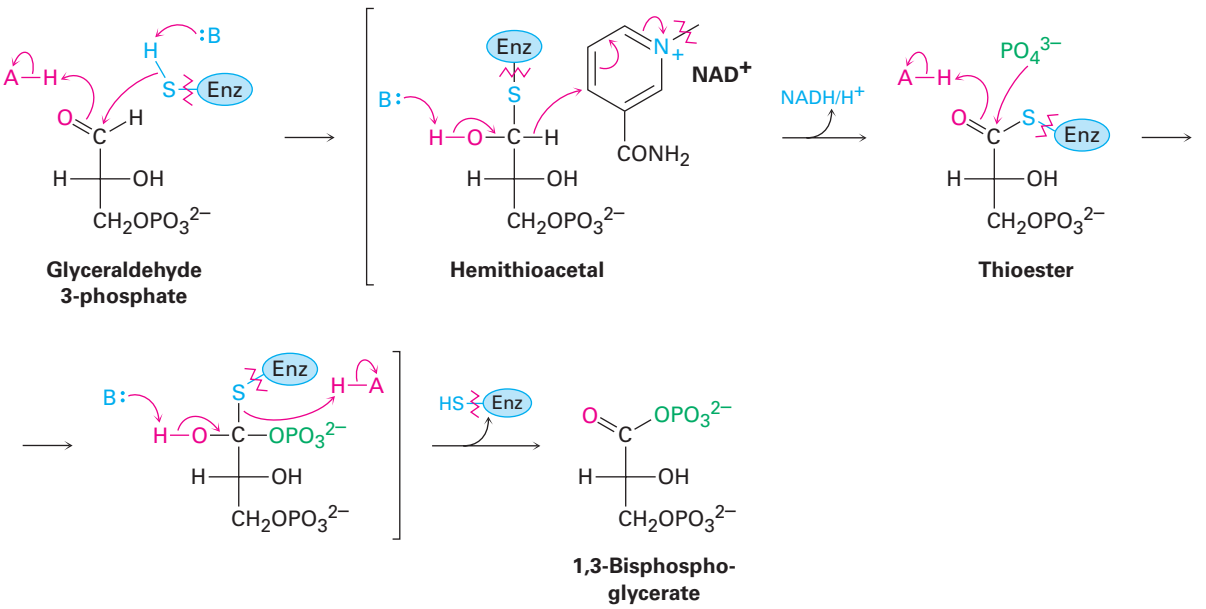
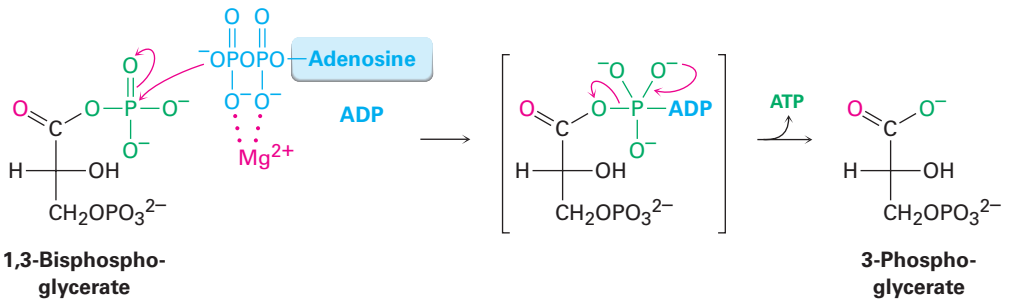
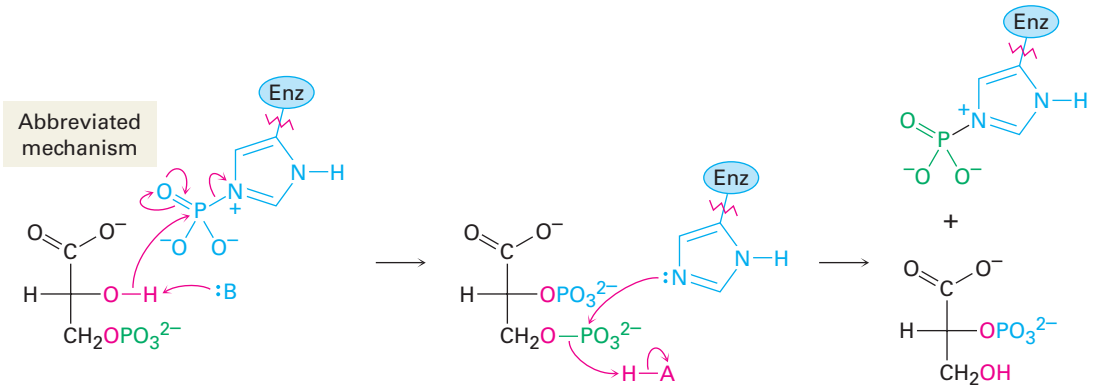


Figure 29.10 Mechanism of step **6** in Figure 29.7, the oxidation and phosphorylation of glyceraldehyde 3-phosphate to give 1,3-bisphosphoglycerate.The process occurs through initial formation of a hemiacetal that is oxidized to a thioester and converted into an acyl phosphate.

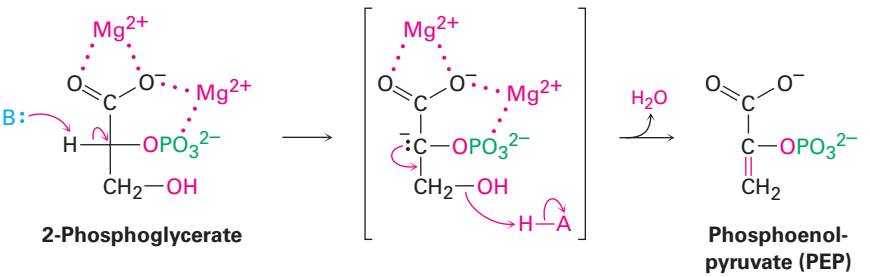
Like all anhydrides **(Section 21.5)**, the mixed carboxylic–phosphoric anhydride is a reactive substrate in nucleophilic acyl (or phosphoryl) substitution reactions. Reaction of 1,3-bisphosphoglycerate with ADP occurs in step 7 by substitution on phosphorus, resulting in transfer of a phosphate group to ADP and giving ATP plus 3-phosphoglycerate. The process is catalyzed by phosphoglycerate kinase and requires Mg^{2+} as cofactor. Together, steps 6 and 7 accomplish the oxidation of an aldehyde to a carboxylic acid.



Step 8 of Figure 29.7: Isomerization 3-Phosphoglycerate isomerizes to 2-phosphoglycerate in a step catalyzed by phosphoglycerate mutase. In plants, 3-phosphoglycerate transfers its phosphoryl group from its C3 oxygen to a histidine residue on the enzyme in one step and then accepts the same phosphoryl group back onto the C2 oxygen in a second step. In animals and yeast, however, the enzyme contains a phosphorylated histidine, which transfers its phosphoryl group to the C2 oxygen of 3-phosphoglycerate and forms 2,3-bisphosphoglycerate as intermediate. The same histidine then accepts a phosphoryl group from the C3 oxygen to yield the isomerized product plus regenerated enzyme. As explained in **Section 29.4**, we'll occasionally use an abbreviated mechanism for nucleophilic acyl substitution reactions to save space.

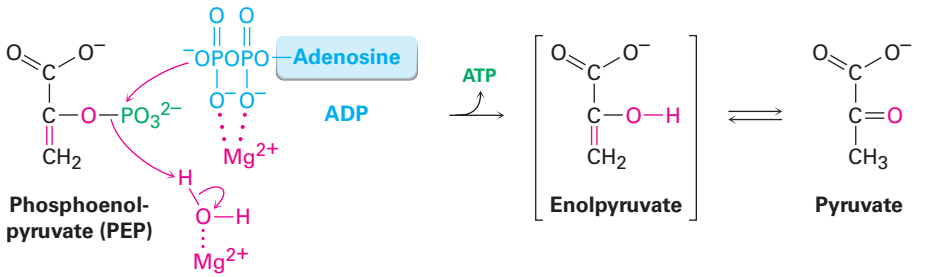


Steps 9–10 of Figure 29.7: Dehydration and Dephosphorylation Like most β-hydroxy carbonyl compounds, 2-phosphoglycerate undergoes a ready dehydration in step 9 by an E1cB mechanism **(Section 23.3)**. The process is catalyzed by enolase, and the product is phosphoenolpyruvate, abbreviated PEP. Two Mg^{2+} ions are associated with the 2-phosphoglycerate to neutralize the negative charges.

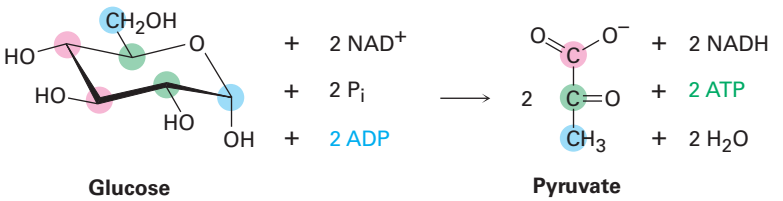


Transfer of the phosphoryl group to ADP in step 10 then generates ATP and gives enolpyruvate, which tautomerizes to pyruvate. The reaction is catalyzed

by pyruvate kinase and requires that a molecule of fructose 1,6-bisphosphate also be present, as well as 2 equivalents of Mg^{2+}. One Mg^{2+} ion coordinates to ADP, and the other increases the acidity of a water molecule necessary for protonation of the enolate ion.



The overall result of glycolysis can be summarized by the following equation:



Problem 29.7

Identify the two steps in glycolysis in which ATP is produced.

Problem 29.8

Look at the entire glycolysis pathway, and make a list of the kinds of organic reactions that take place—nucleophilic acyl substitutions, aldol reactions, E1cB reactions, and so forth.

29.6 Conversion of Pyruvate to Acetyl CoA

Pyruvate, produced by catabolism of glucose (and by degradation of several amino acids), can undergo several further transformations depending on the conditions and on the organism. In the absence of oxygen, pyruvate can be either reduced by NADH to yield lactate [$CH_3CH(OH)CO_2^-$] or, in yeast, fermented to give ethanol. Under typical aerobic conditions in mammals, however, pyruvate is converted by a process called *oxidative decarboxylation* to give acetyl CoA plus CO_2. (*Oxidative* because the oxidation state of the carbonyl carbon rises from that of a ketone to that of a thioester.)

The conversion occurs through a multistep sequence of reactions catalyzed by a complex of enzymes and cofactors called the *pyruvate dehydrogenase complex*. The process occurs in three stages, each catalyzed by one of the enzymes in the complex, as outlined in **Figure 29.11**. Acetyl CoA, the ultimate product, then acts as fuel for the final stage of catabolism, the citric acid cycle.

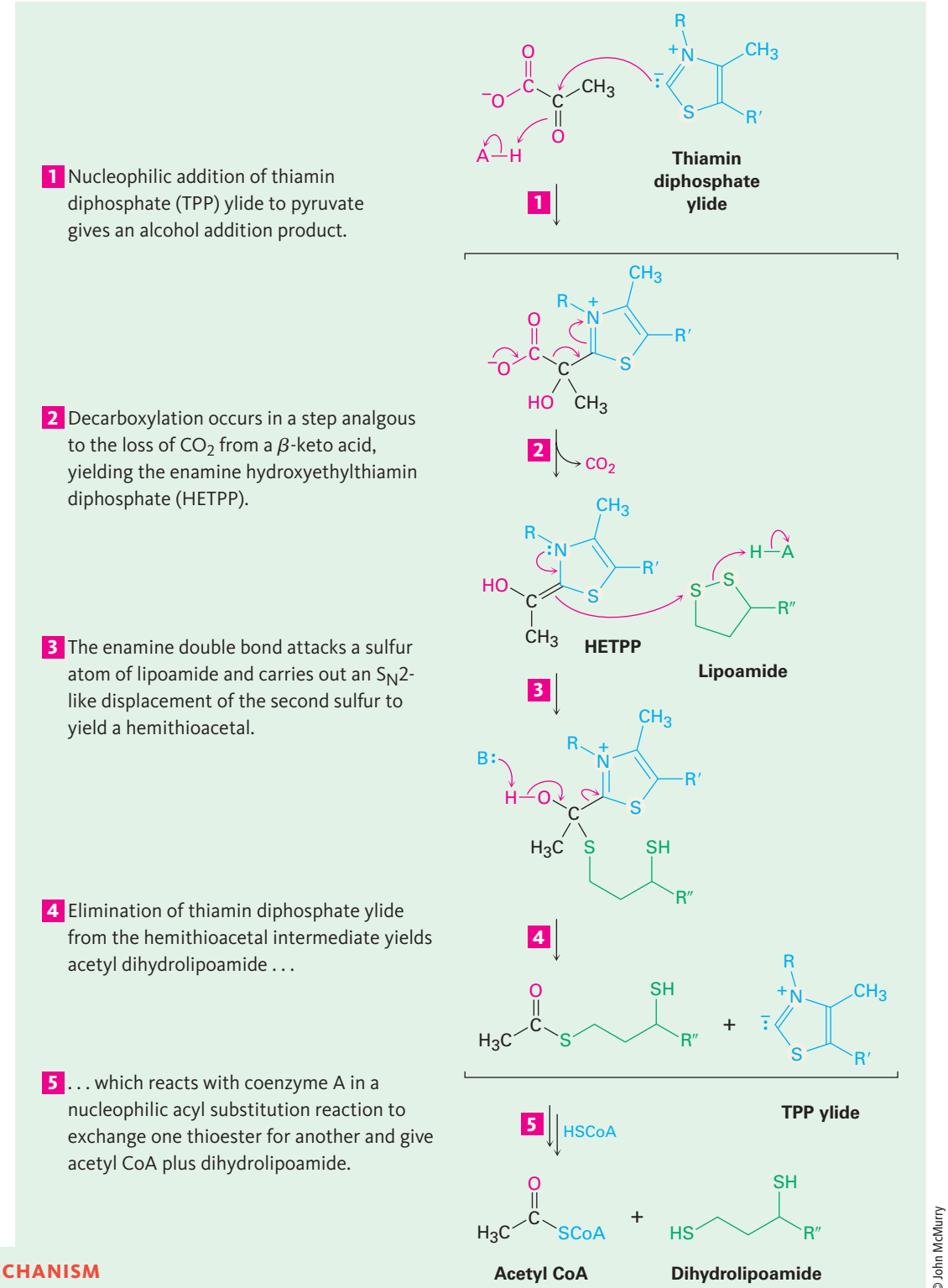
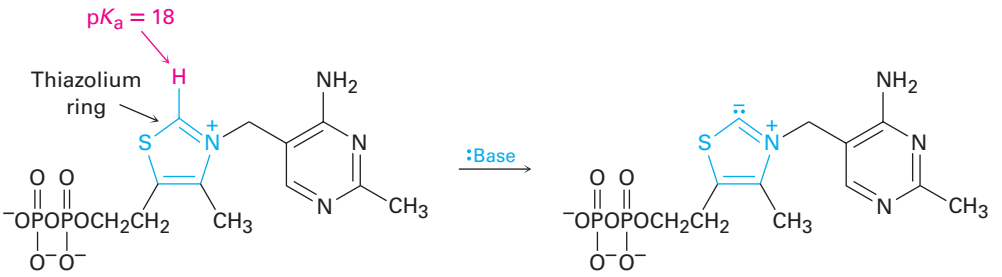


Figure 29.11 | MECHANISM

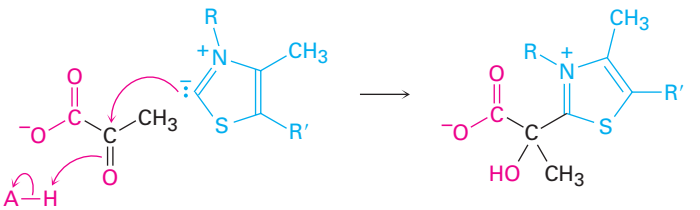
Mechanism of the conversion of pyruvate to acetyl CoA through a multistep sequence of reactions that requires three different enzymes and four different coenzymes. The individual steps are explained in the text.

Step 1 of Figure 29.11: Addition of Thiamin Diphosphate The conversion of pyruvate to acetyl CoA begins by reaction of pyruvate with thiamin diphosphate, a derivative of vitamin B_1. Formerly called thiamin *pyro*phosphate, thiamin diphosphate is usually abbreviated as TPP. The spelling *thiamine* is also correct and frequently used.

The key structural element in thiamin diphosphate is the presence of a thiazolium ring—a five-membered, unsaturated heterocycle containing a sulfur atom and a positively charged nitrogen atom. The thiazolium ring is weakly acidic, with a pK_a of approximately 18 for the ring hydrogen between N and S. Bases can therefore deprotonate thiamin diphosphate, leading to formation of an ylide much like the phosphonium ylides used in Wittig reactions **(Section 19.11)**. As in the Wittig reaction, the TPP ylide is a nucleophile and adds to the ketone carbonyl group of pyruvate to yield an alcohol addition product.



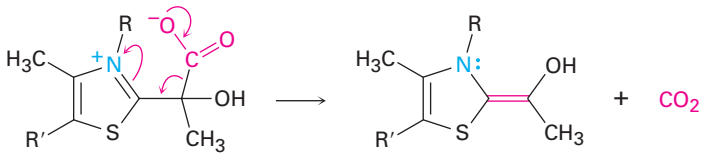
Thiamin diphosphate (TPP) **Thiamin diphosphate ylide (adjacent + and – charges)**



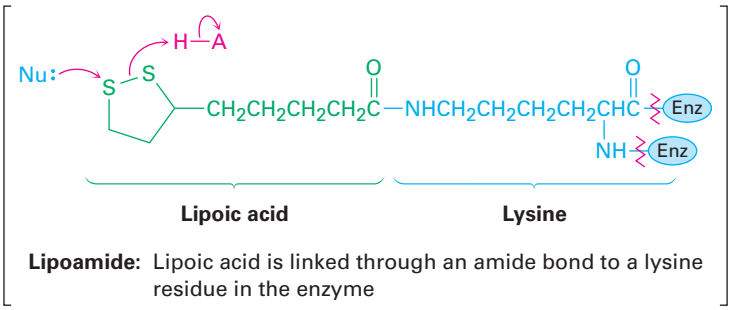
Pyruvate **Thiamin diphosphate ylide**

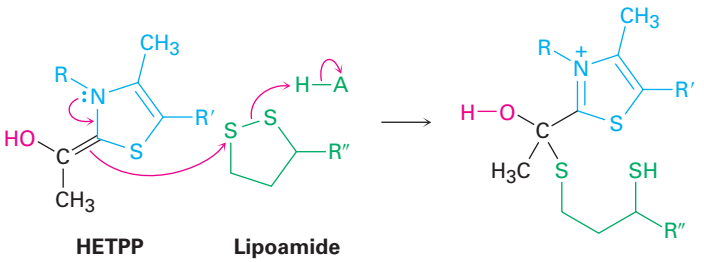
Step 2 of Figure 29.11: Decarboxylation The TPP addition product, which contains an iminium ion β to a carboxylate anion, undergoes decarboxylation in much the same way that a β-keto acid decarboxylates in the acetoacetic ester synthesis **(Section 22.7)**. The C=N$^+$ bond of the pyruvate addition product

acts like the C=O bond of a β-keto acid to accept electrons as CO_2 leaves, giving hydroxyethylthiamin diphosphate (HETPP).

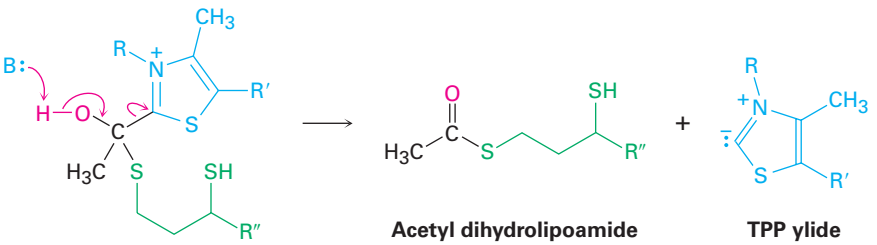


Step 3 of Figure 29.11: Reaction with Lipoamide Hydroxyethylthiamin diphosphate is an enamine ($R_2N—C=C$), which, like all enamines, is nucleophilic **(Section 23.11)**. It therefore reacts with the enzyme-bound disulfide lipoamide by nucleophilic attack on a sulfur atom, displacing the second sulfur in an S_N2-like process.

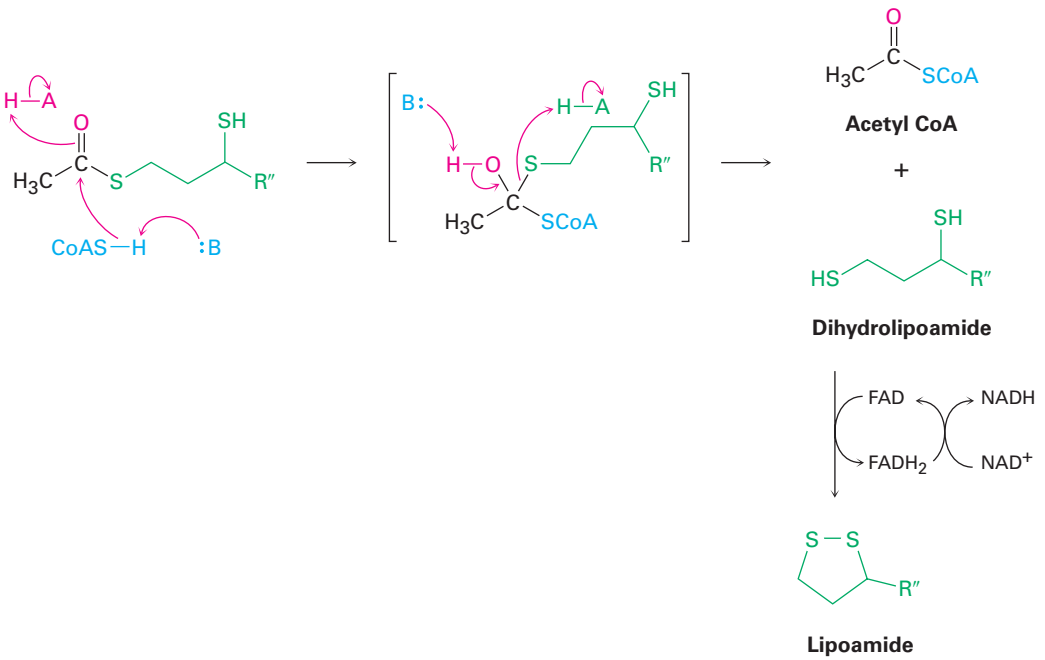




Step 4 of Figure 29.11: Elimination of Thiamin Diphosphate The product of the HETPP reaction with lipoamide is a hemithioacetal, which eliminates thiamin diphosphate ylide. This elimination is the reverse of the ketone addition in step 1 and generates acetyl dihydrolipoamide.



Step 5 of Figure 29.11: Acyl Transfer Acetyl dihydrolipoamide, a thioester, undergoes a nucleophilic acyl substitution reaction with coenzyme A to yield acetyl CoA plus dihydrolipoamide. The dihydrolipoamide is then oxidized back to lipoamide by FAD **(Section 29.3)**, and the $FADH_2$ that results is in turn oxidized back to FAD by NAD^+, completing the catalytic cycle.



Problem 29.9

Which carbon atoms in glucose end up as $-CH_3$ carbons in acetyl CoA? Which carbons end up as CO_2?

29.7 The Citric Acid Cycle

The initial stages of catabolism result in the conversion of both fats and carbohydrates into acetyl groups that are bonded through a thioester link to coenzyme A. Acetyl CoA then enters the next stage of catabolism—the **citric acid cycle**, also called the *tricarboxylic acid (TCA) cycle,* or *Krebs cycle,* after Hans Krebs, who unraveled its complexities in 1937. The overall result of the cycle is the conversion of an acetyl group into two molecules of CO_2 plus reduced coenzymes by the eight-step sequence of reactions shown in **Figure 29.12**.

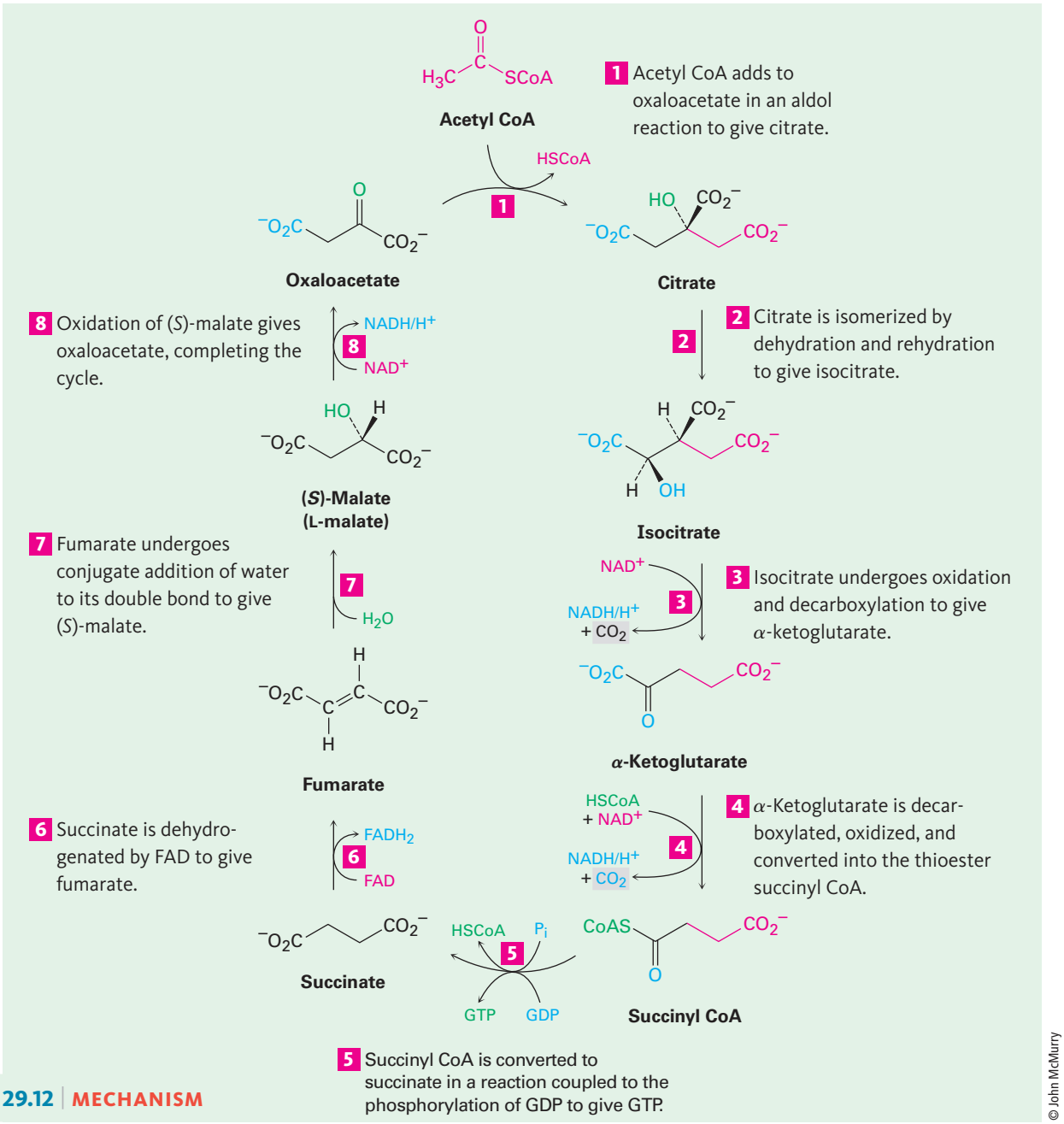


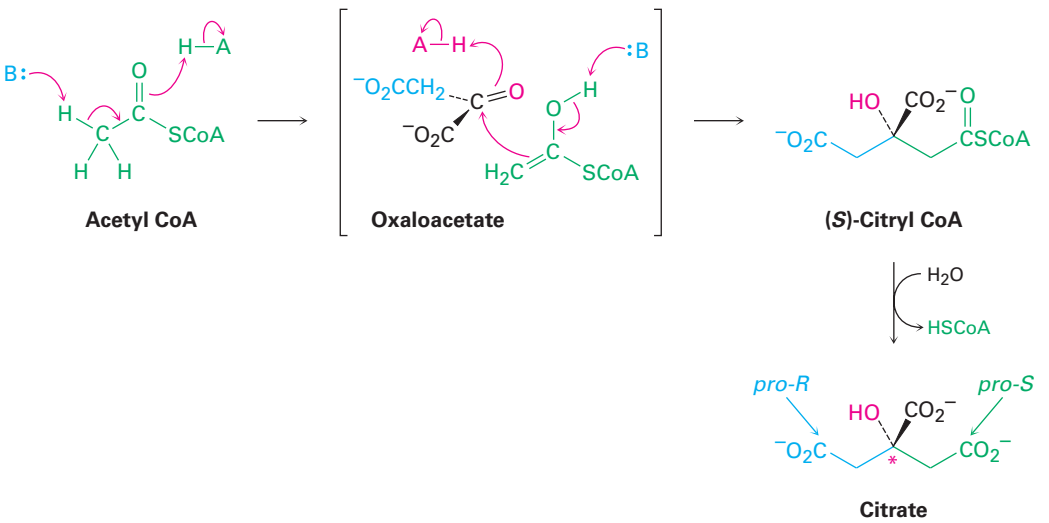
Figure 29.12 | **MECHANISM**

The citric acid cycle is an eight-step series of reactions that results in the conversion of an acetyl group into two molecules of CO_2 plus reduced coenzymes. Individual steps are explained in the text.

As its name implies, the citric acid *cycle* is a closed loop of reactions in which the product of the final step (oxaloacetate) is a reactant in the first step. The intermediates are constantly regenerated and flow continuously through the cycle, which operates as long as the oxidizing coenzymes NAD^+ and FAD are available. To meet this condition, the reduced coenzymes NADH and $FADH_2$ must be reoxidized via the electron-transport chain, which in turn relies on oxygen as the ultimate electron acceptor. Thus, the cycle is dependent on the availability of oxygen and on the operation of the electron-transport chain.

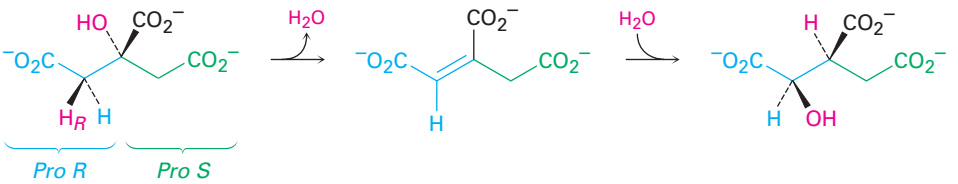
Step 1 of Figure 29.12: Addition to Oxaloacetate Acetyl CoA enters the citric acid cycle in step 1 by nucleophilic addition to the oxaloacetate carbonyl group, to give (*S*)-citryl CoA. The addition is an aldol reaction and is catalyzed by citrate synthase, as discussed in **Section 26.11**. (*S*)-Citryl CoA is then hydrolyzed to citrate by a typical nucleophilic acyl substitution reaction with water, catalyzed by the same citrate synthase enzyme.

Note that the hydroxyl-bearing carbon of citrate is a prochirality center and contains two identical arms. Because the initial aldol reaction of acetyl CoA to oxaloacetate occurs specifically from the *Si* face of the ketone carbonyl group, the *pro-S* arm of citrate is derived from acetyl CoA and the *pro-R* arm is derived from oxaloacetate.



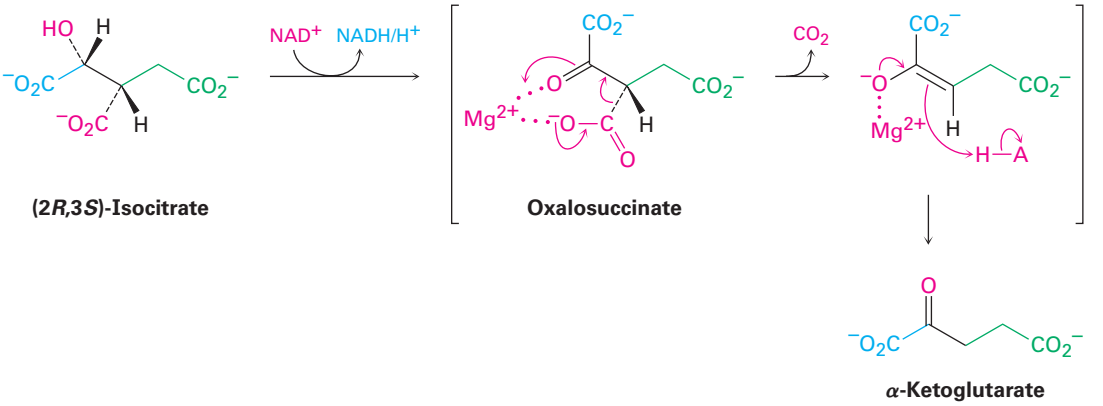
Step 2 of Figure 29.12: Isomerization Citrate, a prochiral tertiary alcohol, is next converted into its isomer, (2*R*,3*S*)-isocitrate, a chiral secondary alcohol. The isomerization occurs in two steps, both of which are catalyzed by the same aconitase enzyme. The initial step is an E1cB dehydration of a β-hydroxy acid to give *cis*-aconitate, the same sort of reaction that occurs in step 9 of glycolysis (Figure 29.7 on page 1175). The second step is a conjugate nucleophilic addition of water to the C=C bond **(Section 19.13)**. The dehydration of citrate takes

place specifically on the *pro-R* arm—the one derived from oxaloacetate—rather than on the *pro-S* arm derived from acetyl CoA.



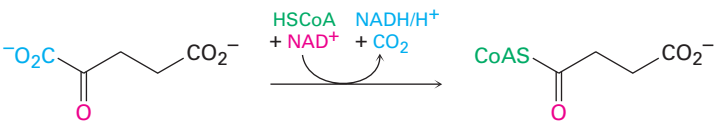
Citrate ***cis*-Aconitate** **(2*R*,3*S*)-Isocitrate**

Step 3 of Figure 29.12: Oxidation and Decarboxylation (2*R*,3*S*)-Isocitrate, a secondary alcohol, is oxidized by NAD^+ in step 3 to give the ketone oxalosuccinate, which loses CO_2 to give α-ketoglutarate. Catalyzed by isocitrate dehydrogenase, the decarboxylation is a typical reaction of a β-keto acid, just like that in the acetoacetic ester synthesis **(Section 22.7)**. The enzyme requires a divalent cation as cofactor to polarize the ketone carbonyl group and make it a better electron acceptor.



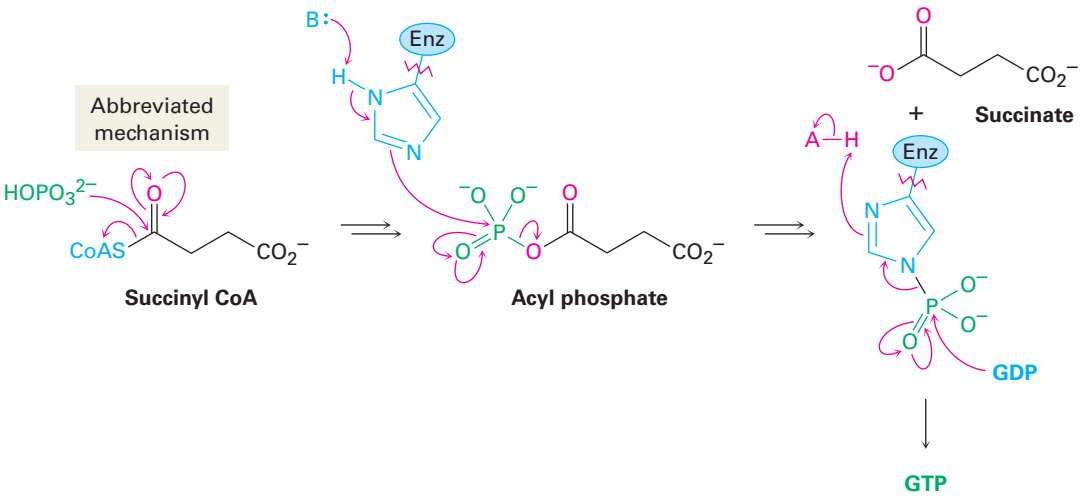
(2*R*,3*S*)-Isocitrate **Oxalosuccinate** **α-Ketoglutarate**

Step 4 of Figure 29.12: Oxidative Decarboxylation The transformation of α-ketoglutarate to succinyl CoA in step 4 is a multistep process just like the transformation of pyruvate to acetyl CoA that we saw in Figure 29.11 on page 1182. In both cases, an α-keto acid loses CO_2 and is oxidized to a thioester in a series of steps catalyzed by a multienzyme dehydrogenase complex. As in the conversion of pyruvate to acetyl CoA, the reaction involves an initial nucleophilic addition reaction of thiamin diphosphate ylide to α-ketoglutarate, followed by decarboxylation. Reaction with lipoamide, elimination of TPP ylide, and finally a transesterification of the dihydrolipoamide thioester with coenzyme A yields succinyl CoA.

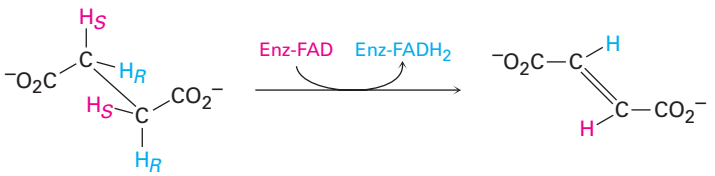


α-Ketoglutarate **Succinyl CoA**

Step 5 of Figure 29.12: Acyl CoA Cleavage Succinyl CoA is converted to succinate in step 5. The reaction is catalyzed by succinyl CoA synthetase and is coupled with phosphorylation of guanosine diphosphate (GDP) to give guanosine triphosphate (GTP). The overall transformation is similar to that of steps 6 through 8 in glycolysis (Figure 29.7), in which a thioester is converted into an acyl phosphate and a phosphate group is then transferred to ADP. The overall result is a "hydrolysis" of the thioester group without involvement of water.

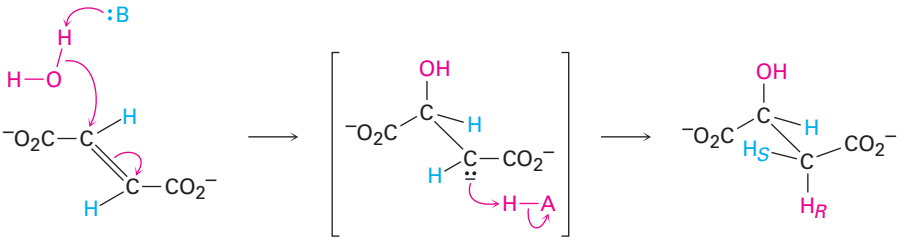


Step 6 of Figure 29.12: Dehydrogenation Succinate is dehydrogenated in step 6 by the FAD-dependent succinate dehydrogenase to give fumarate. The process is analogous to what occurs during the β-oxidation pathway of fatty-acid catabolism **(Section 29.3)**. The reaction is stereospecific, removing the *pro-S* hydrogen from one carbon and the *pro-R* hydrogen from the other.



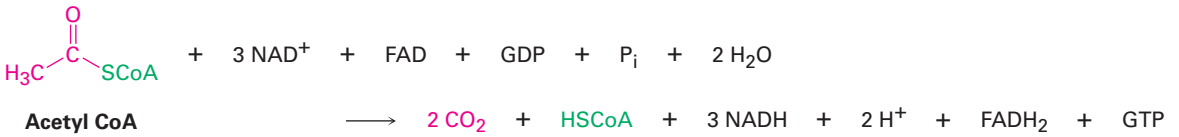
Steps 7–8 of Figure 29.12: Hydration and Oxidation The final two steps in the citric acid cycle are the conjugate nucleophilic addition of water to fumarate to yield (*S*)-malate and the oxidation of (*S*)-malate by NAD^+ to give oxaloacetate. The addition is catalyzed by fumarase and is mechanistically similar to the addition of water to *cis*-aconitate in step 2. The reaction occurs through an

enolate-ion intermediate, which is protonated on the side opposite the OH, leading to a net anti addition.



Fumarate **(*S*)-Malate**

The final step is the oxidation of (*S*)-malate by NAD^+ to give oxaloacetate, a reaction catalyzed by malate dehydrogenase. The citric acid cycle has now returned to its starting point, ready to revolve again. The overall result of the cycle is



Acetyl CoA

Problem 29.10

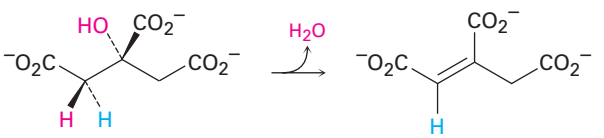
Which of the substances in the citric acid cycle are tricarboxylic acids, thus giving the cycle its alternative name?

Problem 29.11

Write mechanisms for step 2 of the citric acid cycle, the dehydration of citrate and the addition of water to aconitate.

Problem 29.12

Is the *pro-R* or *pro-S* hydrogen removed from citrate during the dehydration in step 2 of the citric acid cycle? Does the elimination reaction occur with syn or anti geometry?

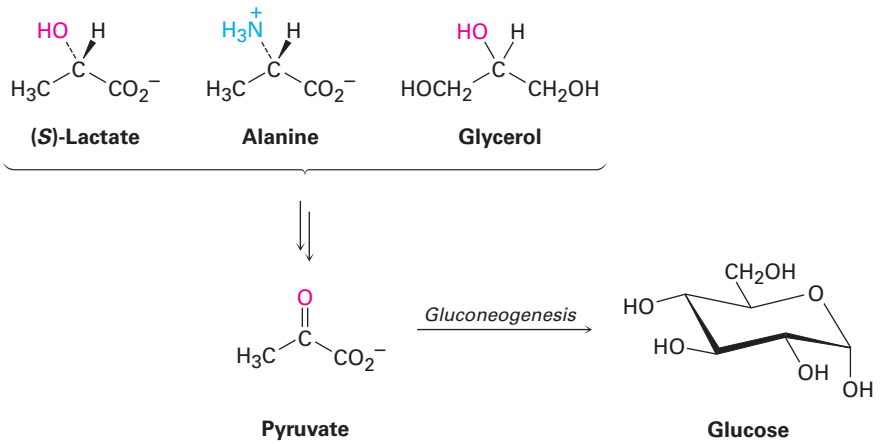


Citrate ***cis*-Aconitate**

29.8 Carbohydrate Biosynthesis: Gluconeogenesis

Glucose is the body's primary fuel when food is plentiful, but in times of fasting or prolonged exercise, glucose stores can become depleted. Most tissues then begin metabolizing fats as their source of acetyl CoA, but the brain is different. The brain relies almost entirely on glucose for fuel and is dependent on receiving a continuous supply in the blood. When the supply of glucose fails, even for a brief time, irreversible damage can occur. Thus, a pathway for synthesizing glucose from simple precursors is crucial.

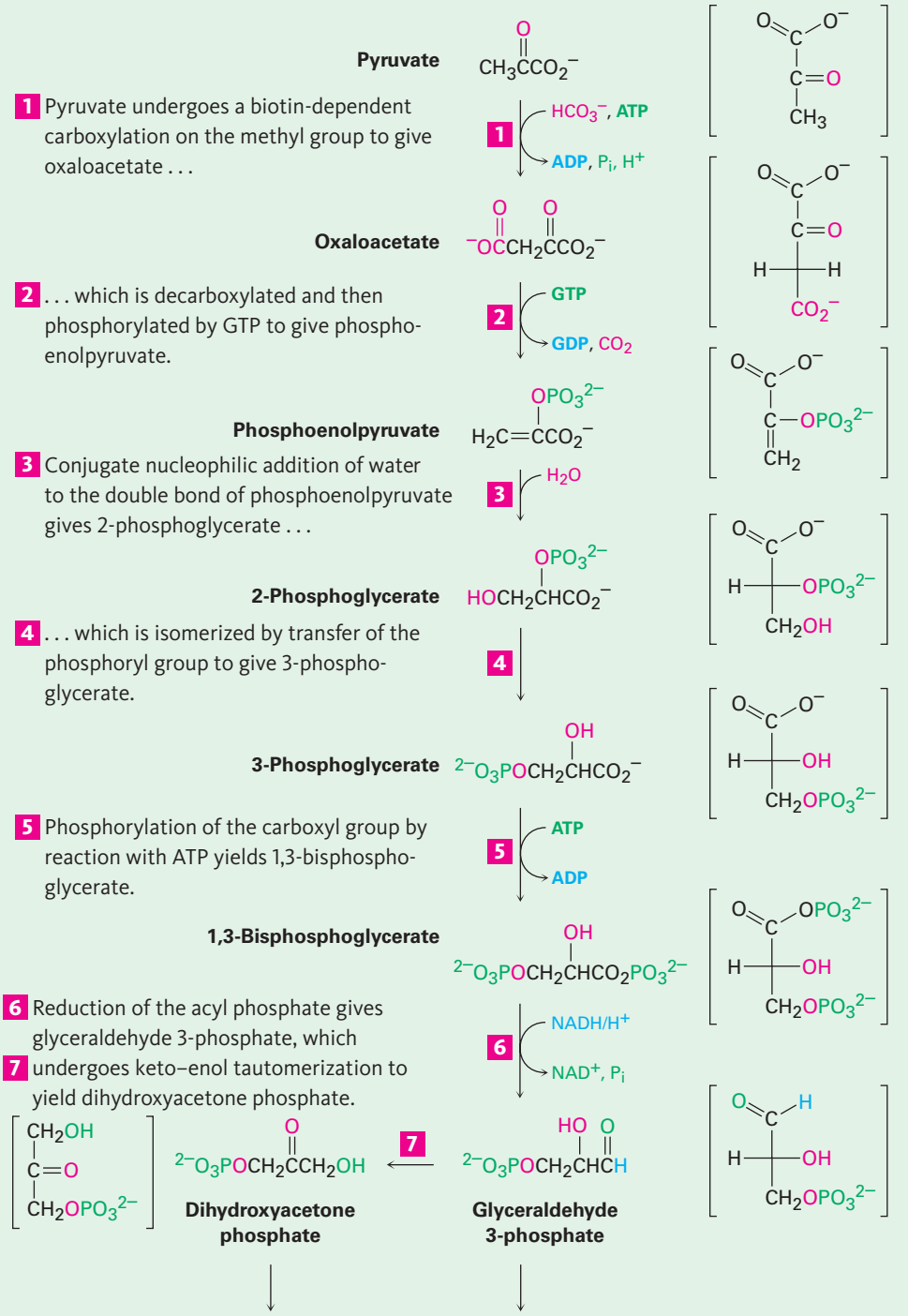
Higher organisms are not able to synthesize glucose from acetyl CoA but must instead use one of the three-carbon precursors lactate, glycerol, or alanine, all of which are readily converted into pyruvate.



Pyruvate then becomes the starting point for **gluconeogenesis**, the 11-step biosynthetic pathway by which organisms make glucose **(Figure 29.13)**. The gluconeogenesis pathway by which glucose is made, however, is not the reverse of the glycolysis pathway by which it is degraded. As with the catabolic and anabolic pathways for fatty acids **(Sections 29.3 and 29.4)**, the catabolic and anabolic pathways for carbohydrates differ in some details so that both are energetically favorable.

Figure 29.13 | MECHANISM

The gluconeogenesis pathway for the biosynthesis of glucose from pyruvate. Individual steps are explained in the text.



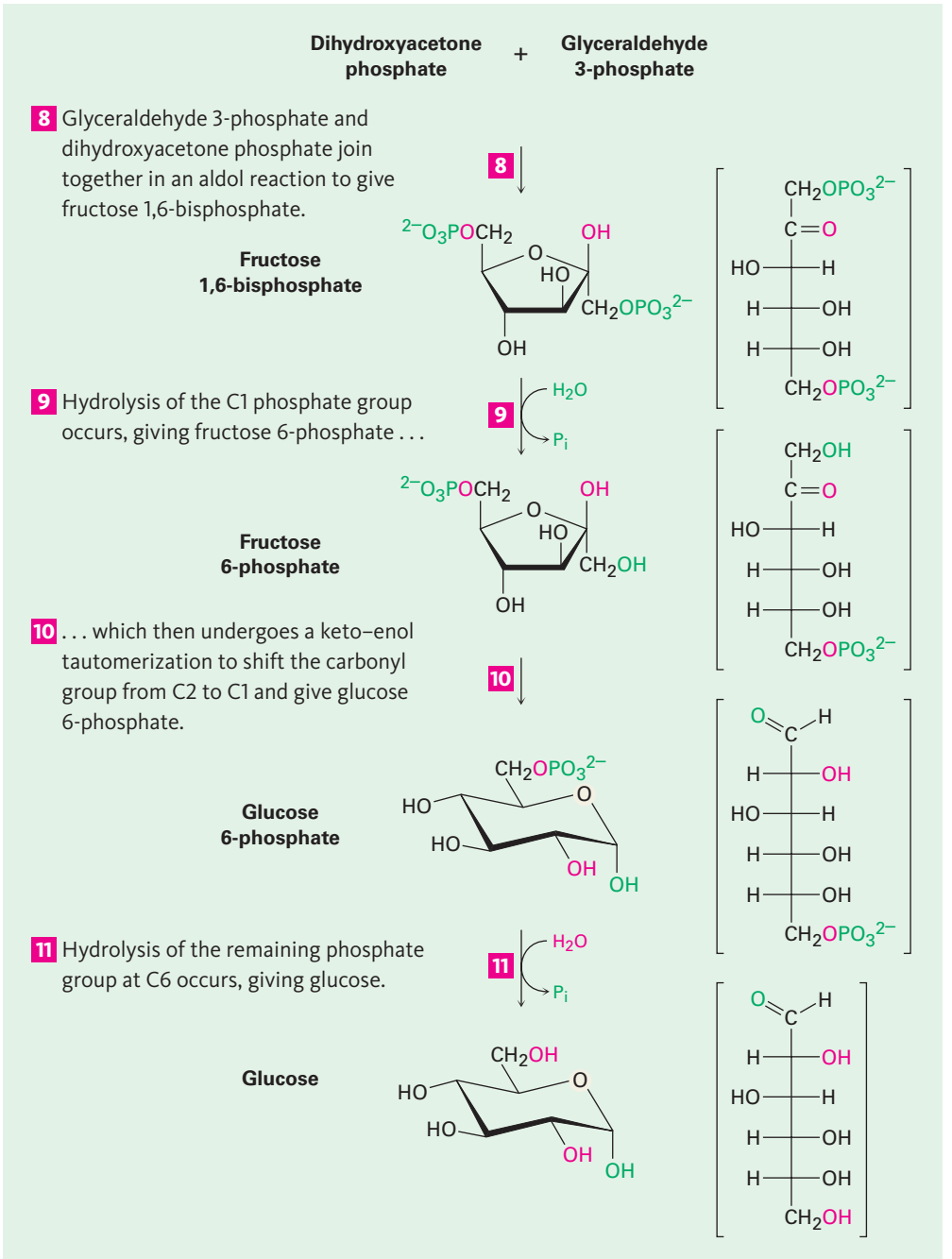
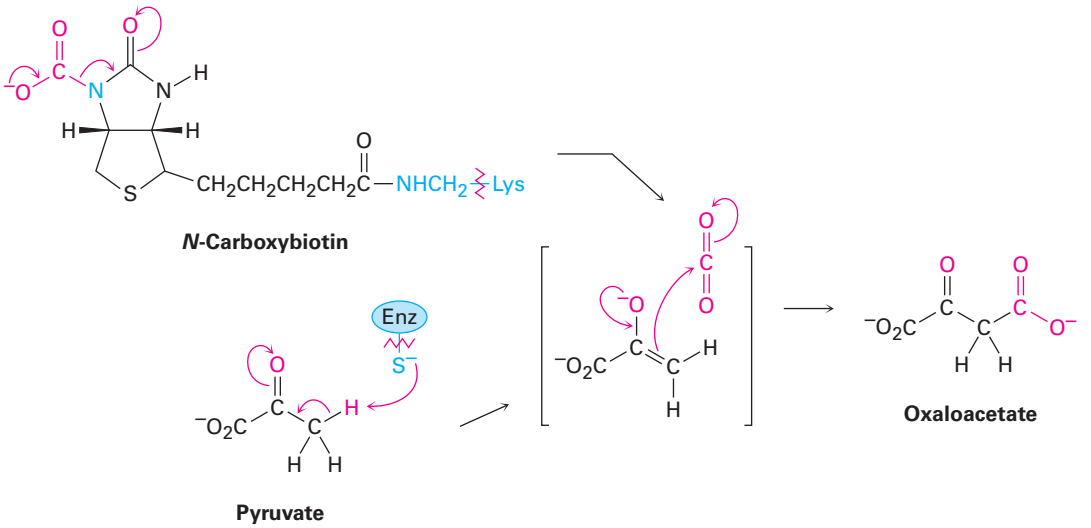
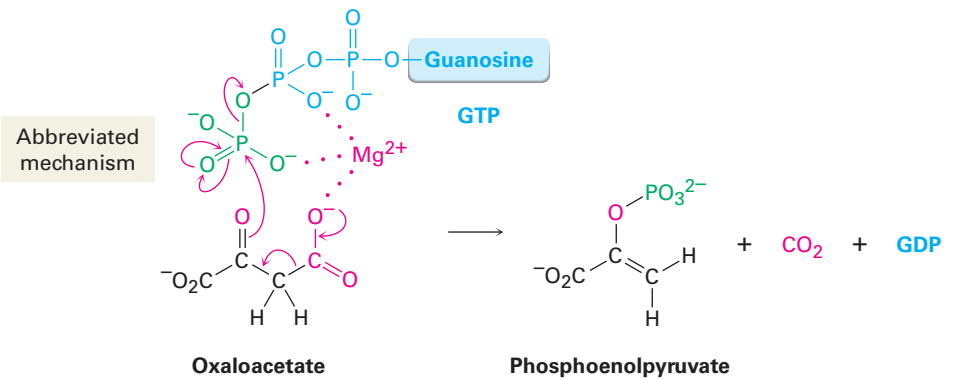


Figure 29.13 (Continued)

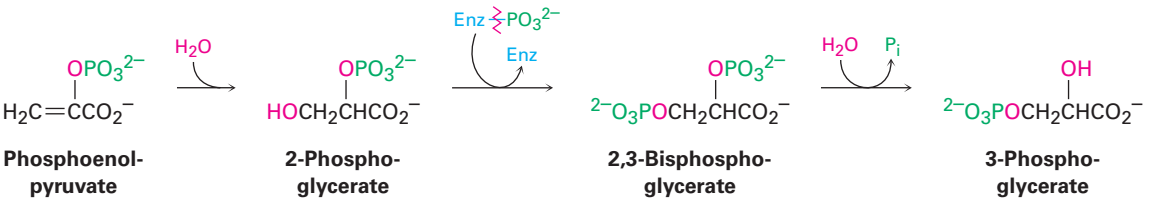
Step 1 of Figure 29.13: Carboxylation Gluconeogenesis begins with the carboxylation of pyruvate to yield oxaloacetate. The reaction is catalyzed by pyruvate carboxylase and requires ATP, bicarbonate ion, and the coenzyme biotin, which acts as a carrier to transport CO_2 to the enzyme active site. The mechanism is analogous to that of step 3 in fatty-acid biosynthesis (Figure 29.5 on page 1168), in which acetyl CoA is carboxylated to yield malonyl CoA.



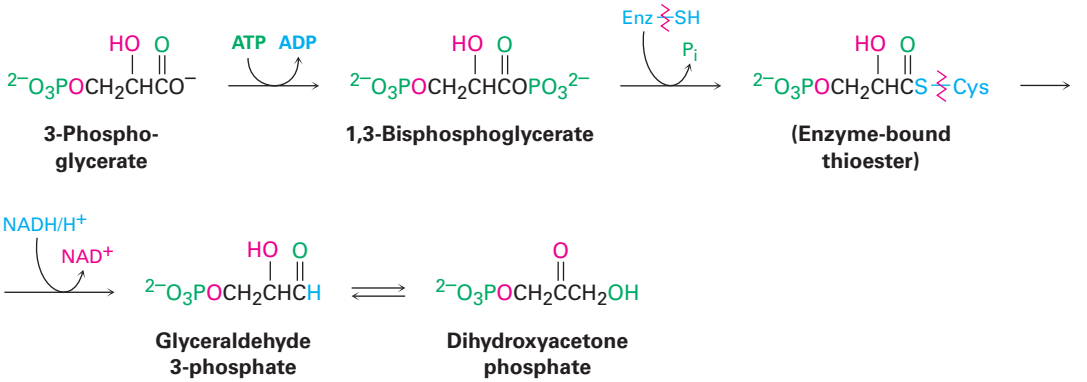
Step 2 of Figure 29.13: Decarboxylation and Phosphorylation Decarboxylation of oxaloacetate, a β-keto acid, occurs by the typical retro-aldol mechanism like that in step 3 in the citric acid cycle (Figure 29.12 on page 1186), and phosphorylation of the resultant pyruvate enolate ion by GTP occurs concurrently to give phosphoenolpyruvate. The reaction is catalyzed by phosphoenolpyruvate carboxykinase.



Steps 3–4 of Figure 29.13: Hydration and Isomerization Conjugate nucleophilic addition of water to the double bond of phosphoenolpyruvate gives 2-phosphoglycerate by a process similar to that of step 7 in the citric acid cycle. Phosphorylation of C3 and dephosphorylation of C2 then yields 3-phosphoglycerate. Mechanistically, these steps are the reverse of steps 9 and 8 in glycolysis (Figure 29.7), which have equilibrium constants near 1 so that substantial amounts of reactant and product are both present.

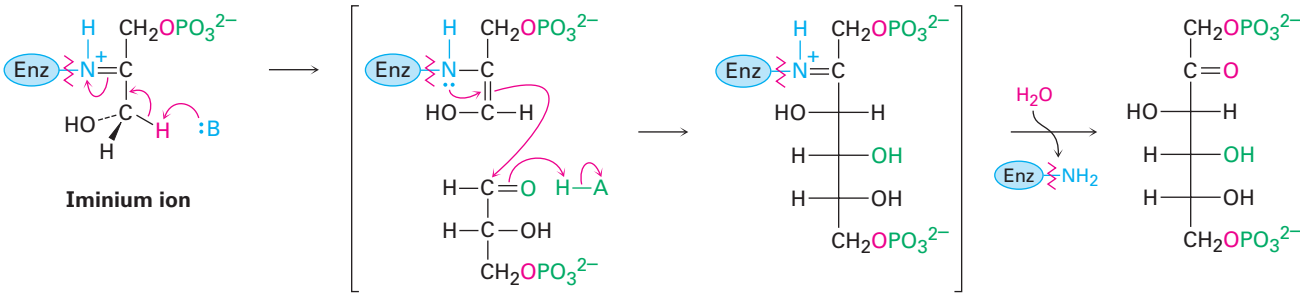


Steps 5–7 of Figure 29.13: Phosphorylation, Reduction, and Tautomerization Reaction of 3-phosphoglycerate with ATP generates the corresponding acyl phosphate, 1,3-bisphosphoglycerate, which binds to the glyceraldehyde 3-phosphate dehydrogenase by a thioester bond to a cysteine residue. Reduction of the thioester by NADH/H^+ yields the corresponding aldehyde, and keto–enol tautomerization of the aldehyde gives dihydroxyacetone phosphate. All three steps are mechanistically the reverse of the corresponding steps 7, 6, and 5 of glycolysis and have equilibrium constants near 1.



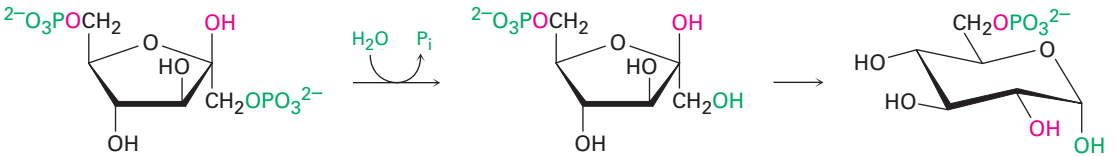
Step 8 of Figure 29.13: Aldol Reaction Dihydroxyacetone phosphate and glyceraldehyde 3-phosphate, the two 3-carbon units produced in step 7, join by an aldol reaction to give fructose 1,6-bisphosphate, the reverse of step 4 in glycolysis (Figure 29.9 on page 1178). As in glycolysis, the reaction is catalyzed in plants and animals by a class I aldolase and takes place on an iminium ion formed by reaction of dihydroxyacetone phosphate with a side-chain

lysine $-NH_2$ group on the enzyme. Loss of a proton from the neighboring carbon then generates an enamine, an aldol-like reaction ensues, and the product is hydrolyzed.



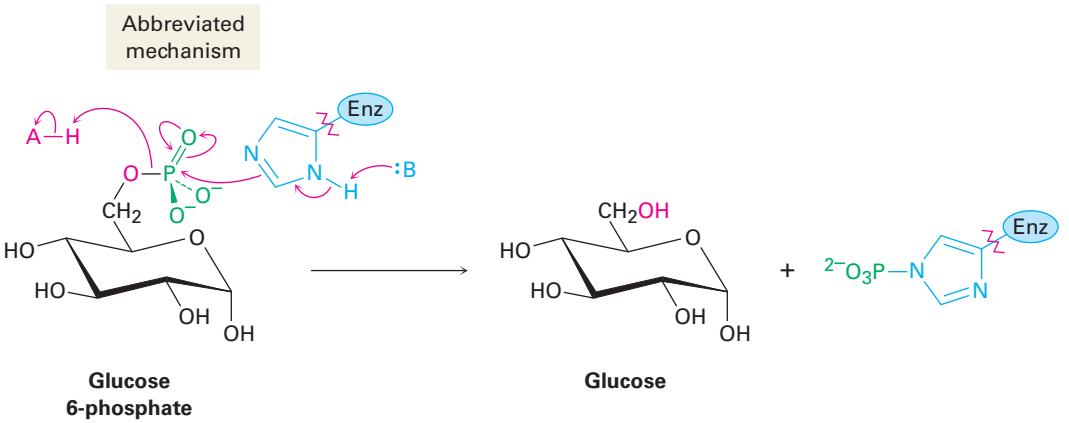
Steps 9–10 of Figure 29.13: Hydrolysis and Isomerization Hydrolysis of the phosphate group at C1 of fructose 1,6-bisphosphate gives fructose 6-phosphate. Although the result of the reaction is the exact opposite of step 3 in glycolysis, the mechanism is not. In glycolysis, the phosphorylation is accomplished by reaction of fructose with ATP, with formation of ADP as by-product. The reverse of that process, however—the reaction of fructose 1,6-bisphosphate with ADP to give fructose 6-phosphate and ATP—is energetically unfavorable because ATP is too high in energy. Thus, an alternative pathway is used in which the C1 phosphate group is removed by a direct hydrolysis reaction, catalyzed by fructose 1,6-bisphosphatase.

Following hydrolysis, keto–enol tautomerization of the carbonyl group from C2 to C1 gives glucose 6-phosphate. The isomerization is the reverse of step 2 in glycolysis.

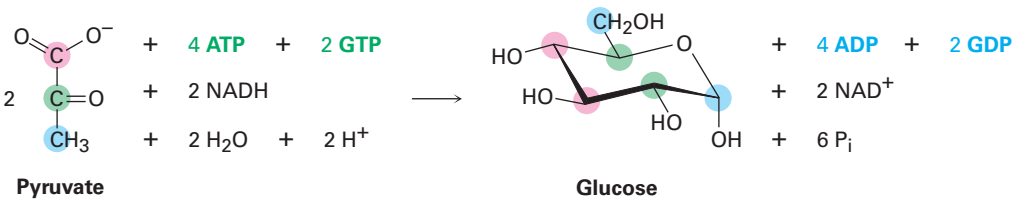


Step 11 of Figure 29.13: Hydrolysis The final step in gluconeogenesis is the conversion of glucose 6-phosphate to glucose by a second phosphatase-catalyzed hydrolysis reaction. As just discussed for the hydrolysis of fructose 1,6-bisphosphate in step 9, and for the same energetic reasons, the mechanism of the glucose 6-phosphate hydrolysis is not the exact opposite of the corresponding step 1 in glycolysis.

Interestingly, however, the mechanisms of the two phosphate hydrolysis reactions in steps 9 and 11 are not the same. In step 9, water is the nucleophile, but in the glucose 6-phosphate reaction of step 11, a histidine residue on the enzyme attacks phosphorus, giving a phosphoryl enzyme intermediate that subsequently reacts with water.



The overall result of gluconeogenesis is summarized by the following equation:



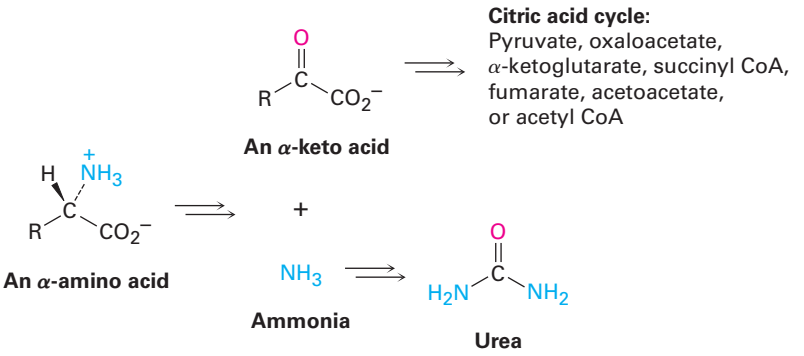
Problem 29.13

Write a mechanism for step 6 of gluconeogenesis, the reduction of 3-phosphoglyceryl phosphate with $NADH/H^+$ to yield glyceraldehyde 3-phosphate.

29.9 Catabolism of Proteins: Deamination

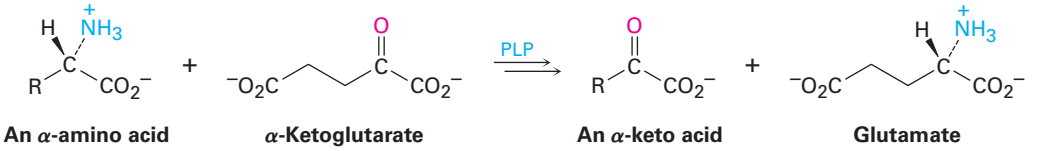
The catabolism of proteins is much more complex than that of fats and carbohydrates because each of the 20 α-amino acids is degraded through its own unique pathway. The general idea, however, is that (1) the α amino group is first removed as ammonia by a *deamination* process, (2) the ammonia is converted

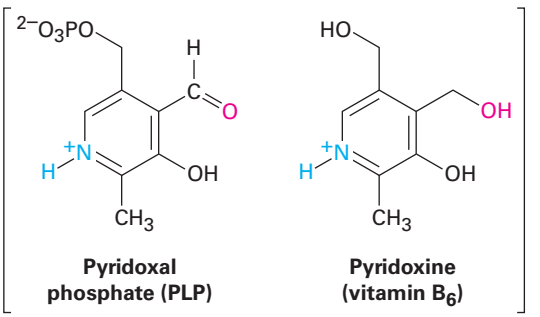
into urea, and (3) the remaining amino acid carbon skeleton (usually an α-keto acid) is converted into a compound that enters the citric acid cycle.



Transamination

Deamination is usually accomplished by a **transamination** reaction in which the $-NH_2$ group of the amino acid is exchanged with the keto group of α-ketoglutarate, forming a new α-keto acid plus glutamate. The overall process occurs in two parts, is catalyzed by aminotransferases, and involves participation of the coenzyme pyridoxal phosphate, abbreviated PLP, a derivative of pyridoxine (vitamin B_6). Different aminotransferases differ in their specificity for amino acids, but the mechanism remains the same.





The mechanism of the first part of transamination is shown in **Figure 29.14**. The process begins with reaction between the α-amino acid and pyridoxal phosphate, which is covalently bonded to the aminotransferase by an imine linkage between the side-chain $-NH_2$ group of a lysine residue in the enzyme and the PLP aldehyde group. Deprotonation/reprotonation of the PLP–amino acid imine effects tautomerization of the imine C=N bond, and hydrolysis of the tautomerized imine gives an α-keto acid plus pyridoxamine phosphate (PMP).

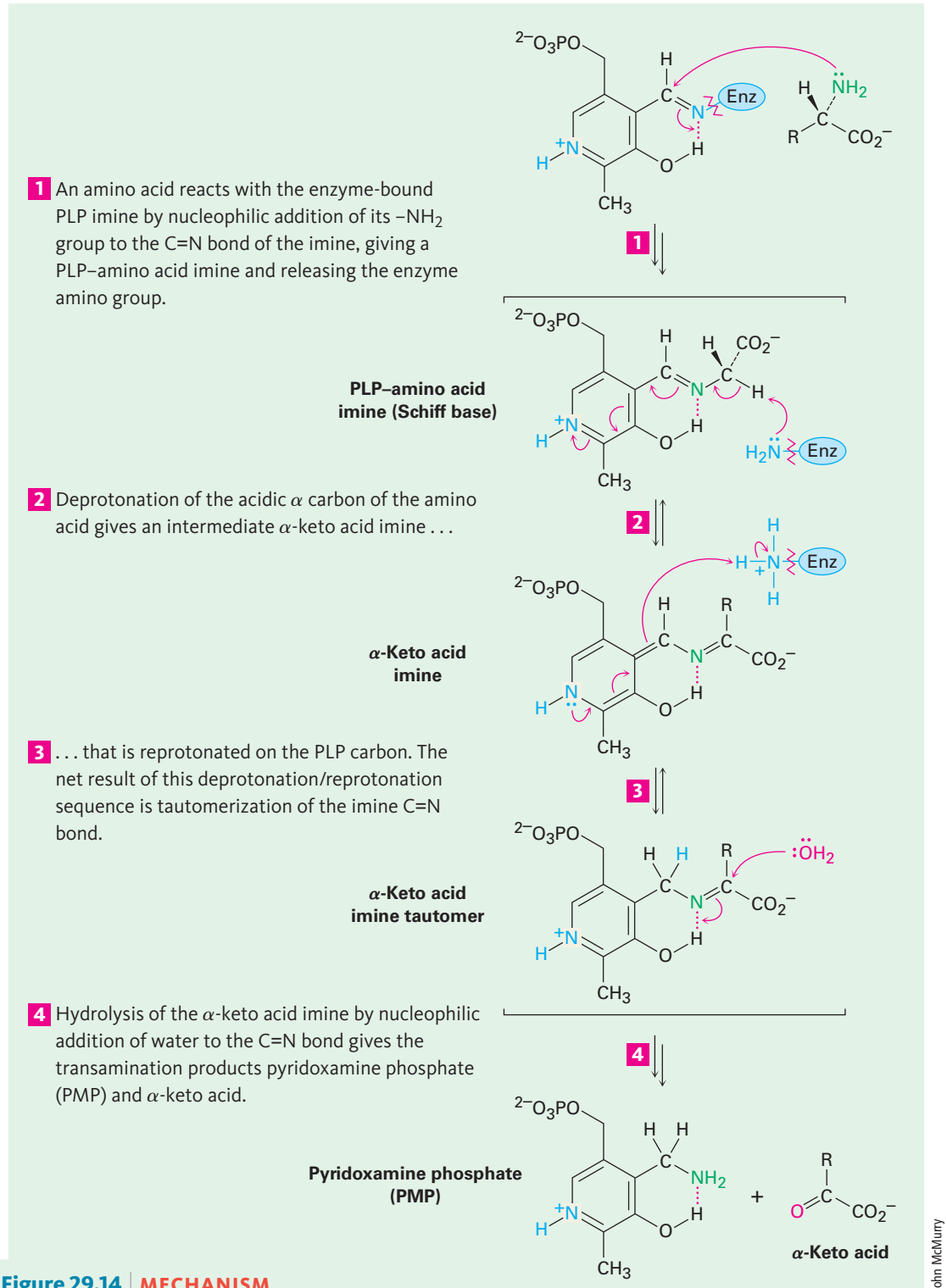
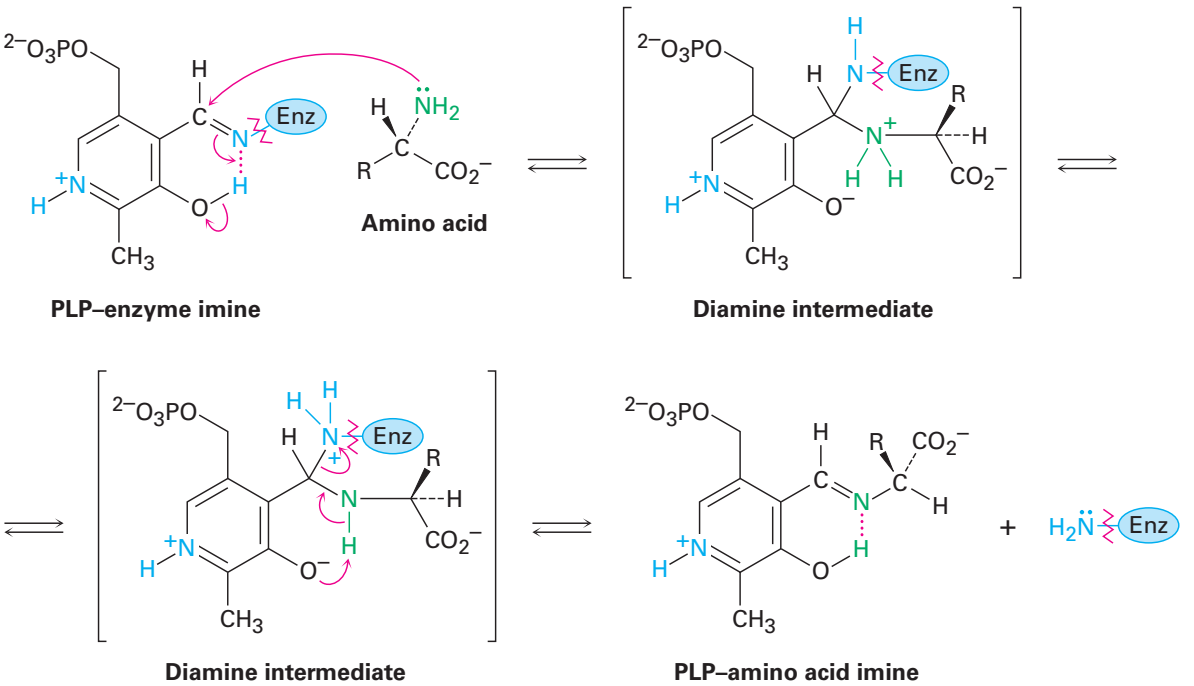


Figure 29.14 | MECHANISM

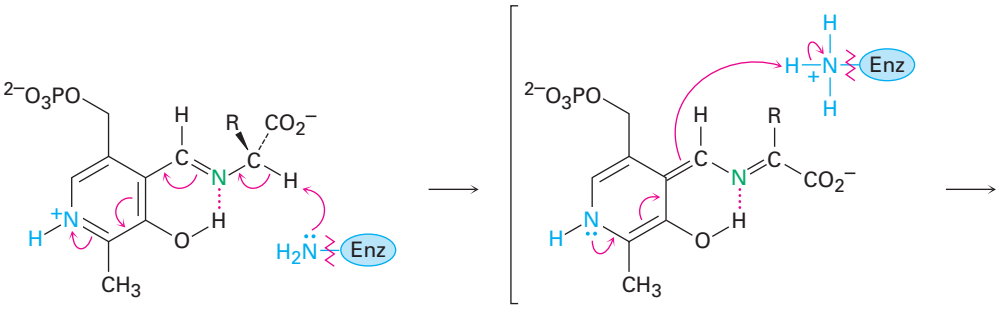
Mechanism of the enzyme-catalyzed, PLP-dependent transamination of an α-amino acid to give an α-keto acid. Individual steps are explained in the text.

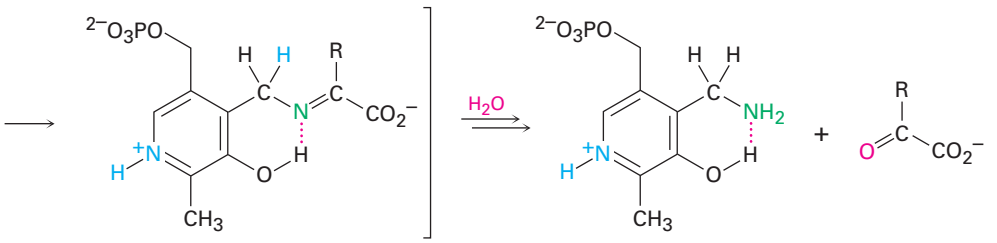
Step 1 of Figure 29.14: Transimination The first step in trans*amination* is trans*imination*—the reaction of the PLP–enzyme imine with an α-amino acid to give a PLP–amino acid imine plus expelled enzyme as the leaving group. The reaction occurs by nucleophilic addition of the amino acid $-NH_2$ group to the C=N bond of the PLP imine, much as an amine adds to the C=O bond of a ketone or aldehyde in a nucleophilic addition reaction **(Section 19.8)**. The protonated diamine intermediate undergoes a proton transfer and expels the lysine amino group in the enzyme to complete the step.



Steps 2–4 of Figure 29.14: Tautomerization and Hydrolysis Following formation of the PLP–amino acid imine in step 1, a tautomerization of the C=N bond occurs in step 2. The basic lysine residue in the enzyme that was expelled as a leaving group during transimination deprotonates the acidic α position of the amino acid, with the protonated pyridine ring of PLP acting as the electron acceptor. Reprotonation occurs on the carbon atom next to the ring, generating a tautomeric product that is the imine of an α-keto acid with pyridoxamine phosphate, abbreviated PMP.

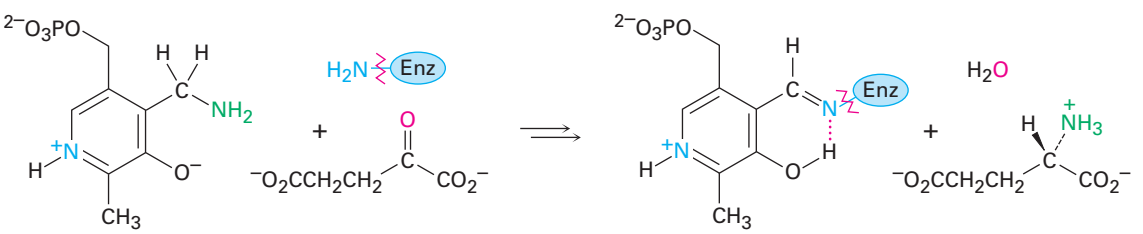
Hydrolysis of this PMP–α-keto acid imine then completes the first part of the transamination reaction. The hydrolysis is the mechanistic reverse of imine formation and occurs by nucleophilic addition of water to the imine, followed by proton transfer and expulsion of PMP as leaving group.





Regeneration of PLP from PMP

With PLP plus the α-amino acid now converted into PMP plus an α-keto acid, PMP must be transformed back into PLP to complete the catalytic cycle. The conversion occurs by another transamination reaction, this one between PMP and an α-keto acid, usually α-ketoglutarate. The products are PLP plus glutamate, and the mechanism of the process is the exact reverse of that shown in Figure 29.14. That is, PMP and α-ketoglutarate give an imine; the PMP–α-ketoglutarate imine undergoes tautomerization of the C=N bond to give a PLP–glutamate imine; and the PLP–glutamate imine reacts with a lysine residue on the enzyme in a transimination process to yield PLP–enzyme imine plus glutamate.



Problem 29.14

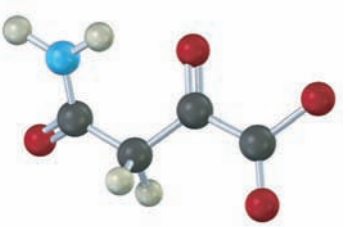
Write all the steps in the transamination reaction of PMP with α-ketoglutarate plus a lysine residue in the enzyme to give the PLP–enzyme imine plus glutamate.

Problem 29.15

What α-keto acid is formed on transamination of leucine?

Problem 29.16

From what amino acid is the following α-keto acid derived?



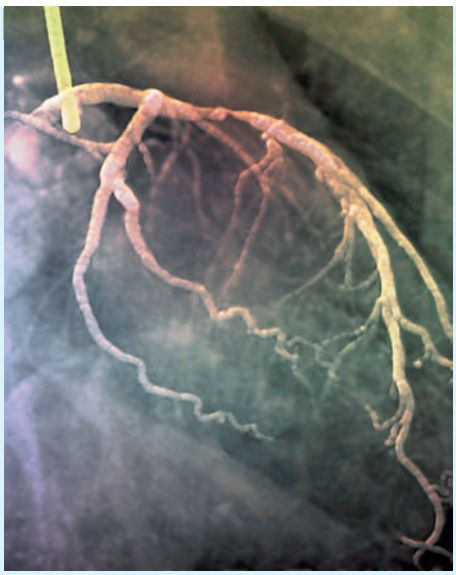
29.10 Some Conclusions about Biological Chemistry

As promised in the chapter introduction, the past few sections have been a fast-paced tour of a large number of reactions. Following it all undoubtedly required a lot of work and a lot of page-turning to look at earlier sections.

After examining the various metabolic pathways, perhaps the main conclusion about biological chemistry is the remarkable similarity between the mechanisms of biological reactions and the mechanisms of laboratory reactions. In all the pathways described in this chapter, terms like *imine formation, aldol reaction, nucleophilic acyl substitution reaction, E1cB reaction,* and *Claisen reaction* appear constantly. Biological reactions aren't mysterious; there are clear, understandable explanations for the reactions carried out by living organisms. Biological chemistry *is* organic chemistry.

But other than satisfying a sense of curiosity about how life works at the molecular level, what is there to be gained from studying biological chemistry and learning about metabolism? One good answer is given in the following *A Deeper Look,* where the story is told of how knowledge of a biosynthetic pathway led to the design of new drugs that have saved many millions of lives.

A DEEPER LOOK Statin Drugs



Zephyr/Photo Researchers, Inc.

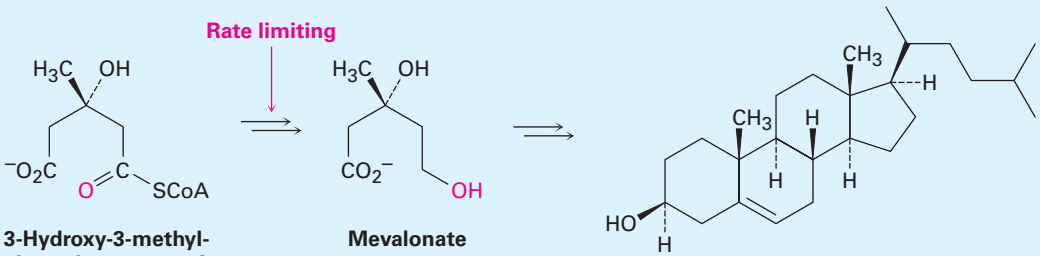
The buildup of cholesterol deposits inside arteries can cause coronary heart disease, a leading cause of death for both men and women.

Coronary heart disease—the buildup of cholesterol-containing plaques on the walls of heart arteries—is the leading cause of death for both men and women older than age 20 in industrialized countries. It's estimated that up to one-third of women and one-half of men will develop the disease at some point in their lives.

The onset of coronary heart disease is directly correlated with blood cholesterol levels (see the Chapter 27 *A Deeper Look*), and the first step in disease prevention is to lower those levels. It turns out that only about 25% of your blood cholesterol comes from what you eat; the remaining 75%—about 1000 mg each day—is biosynthesized in your liver from dietary fats and carbohydrates. Thus, any effective plan for lowering your cholesterol level means limiting the amount that your body makes, which is where a detailed chemical knowledge of cholesterol biosynthesis comes in.

We saw in **Sections 27.5 and 27.7** that all steroids, including cholesterol, are biosynthesized from the triterpenoid lanosterol, which in turn comes from acetyl CoA through isopentenyl diphosphate. If you knew all the mechanisms for all the chemical steps in cholesterol biosynthesis, you might be able to devise a drug that would block one of those steps, thereby short-circuiting the biosynthetic process and controlling the amount of cholesterol produced.

But we do know those mechanisms! Look back at the pathway for the biosynthesis of isopentenyl diphosphate from acetyl CoA, shown in Figure 27.7 on page 1100. It turns out that the rate-limiting step in the pathway is the reduction of 3-hydroxy-3-methylglutaryl CoA (abbreviated HMG-CoA) to mevalonate, brought about by the enzyme HMG-CoA reductase. If that enzyme could be stopped from functioning, cholesterol biosynthesis would also be stopped.



3-Hydroxy-3-methylglutaryl coenzyme A (HMG-CoA)

Mevalonate

Cholesterol

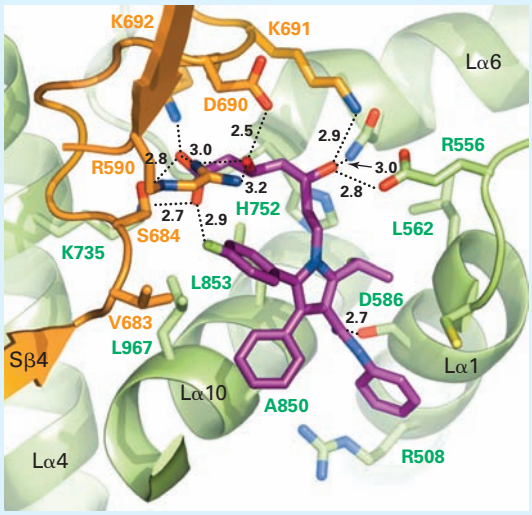
To find a drug that blocks HMG-CoA reductase, chemists did two simultaneous experiments on a large number of potential drug candidates isolated from soil microbes. In one experiment, the drug candidate and mevalonate were added to liver extract; in the second experiment, only the drug candidate was added without mevalonate. If cholesterol was produced only in the presence of added mevalonate but not in the absence of mevalonate, the drug candidate must have blocked the enzyme for mevalonate synthesis.

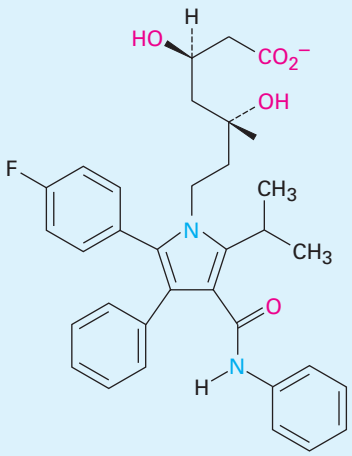
The drugs that block HMG-CoA reductase, and thus control cholesterol synthesis in the body, are called *statins*. They are the most widely prescribed drugs in the world, with an estimated $15 billion in annual sales. So effective are they that in the 10-year period following their introduction in 1994, the death rate from coronary heart disease decreased by

(continued)

(continued)

33% in the United States. Atorvastatin (Lipitor), simvastatin (Zocor), rosuvastatin (Crestor), pravastatin (Pravachol), and lovastatin (Mevacor) are examples. An X-ray crystal structure of the active site in the HMG-CoA reductase enzyme is shown in the accompanying graphic, along with a molecule of atorvastatin (blue) that is tightly bound in the active site and stops the enzyme from functioning. A good understanding of organic chemistry certainly paid off in this instance.





Atorvastatin (Lipitor)

Summary

Key words

anabolism, 1154
***β*-oxidation pathway, 1162**
catabolism, 1154
citric acid cycle, 1185
gluconeogenesis, 1191
glycolysis, 1173
metabolism, 1154
Schiff base, 1177
transamination, 1198

Metabolism is the sum of all chemical reactions in the body. Reactions that break down large molecules into smaller fragments are called **catabolism**, and those that build up large molecules from small pieces are called **anabolism**. Although the details of specific biochemical pathways are sometimes complex, all the reactions that occur follow the normal rules of organic chemical reactivity.

The catabolism of fats begins with digestion, in which ester bonds are hydrolyzed to give glycerol and fatty acids. The fatty acids are degraded in the four-step ***β*-oxidation pathway** by removal of two carbons at a time, yielding acetyl CoA. Catabolism of carbohydrates begins with the hydrolysis of glycoside bonds to give glucose, which is degraded in the ten-step **glycolysis** pathway. Pyruvate, the initial product of glycolysis, is then converted into acetyl CoA. Acetyl CoA next enters the eight-step **citric acid cycle**, where it is further degraded into CO_2. The cycle is a closed loop of reactions in which the product of the final step (oxaloacetate) is a reactant in the first step.

Catabolism of proteins is more complex than that of fats or carbohydrates because each of the 20 different amino acids is degraded by its own unique pathway. In general, though, the amino nitrogen atoms are removed and the substances that remain are converted into compounds that enter the citric acid cycle. Most amino acids lose their nitrogen atom by **transamination**, a

reaction in which the $-NH_2$ group of the amino acid changes places with the keto group of an α-keto acid such as α-ketoglutarate. The products are a new α-keto acid and glutamate.

The energy released in catabolic pathways is used in the *electron-transport chain* to make molecules of adenosine triphosphate, ATP. ATP, the final result of food catabolism, couples to and drives many otherwise unfavorable reactions.

Biomolecules are synthesized as well as degraded, but the pathways for anabolism and catabolism are not the exact reverse of one another. Fatty acids are biosynthesized from acetate by an 8-step pathway, and carbohydrates are made from pyruvate by the 11-step **gluconeogenesis** pathway.

Exercises

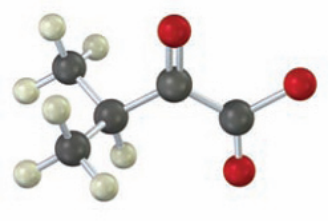
Visualizing Chemistry

OWL Interactive versions of these problems are assignable in OWL for Organic Chemistry.

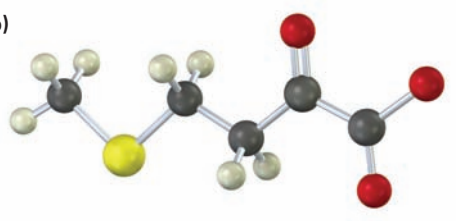
(Problems 29.1–29.16 appear within the chapter.)

29.17 Identify the amino acid that is a catabolic precursor of each of the following α-keto acids:

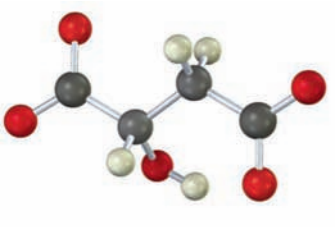
(a)



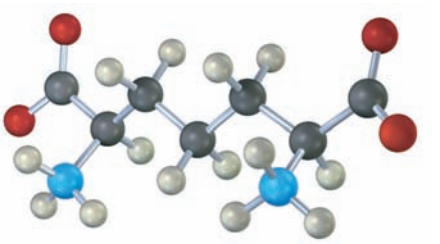
(b)



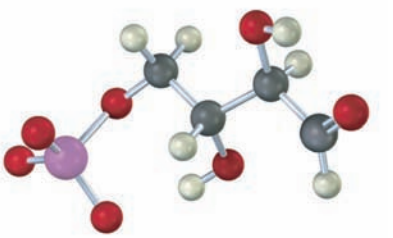
29.18 Identify the following intermediate in the citric acid cycle, and tell whether it has *R* or *S* stereochemistry:



29.19 The following compound is an intermediate in the biosynthesis of one of the 20 common α-amino acids. Which one is it likely to be, and what kind of chemical change must take place to complete the biosynthesis?



29.20 The following compound is an intermediate in the pentose phosphate pathway, an alternative route for glucose metabolism. Identify the sugar it is derived from.



Additional Problems

Enzymes and Coenzymes

29.21 What chemical events occur during the digestion of food?

29.22 What is the difference between digestion and metabolism?

29.23 What is the difference between anabolism and catabolism?

29.24 Draw the structure of adenosine 5′-monophosphate (AMP), an intermediate in some biochemical pathways.

29.25 Cyclic adenosine monophosphate (cyclic AMP), a modulator of hormone action, is related to AMP (Problem 29.24) but has its phosphate group linked to *two* hydroxyl groups at C3′ and C5′ of the sugar. Draw the structure of cyclic AMP.

29.26 What general kind of reaction does ATP carry out?

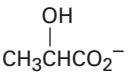
29.27 What general kind of reaction does NAD^+ carry out?

29.28 What general kind of reaction does FAD carry out?

29.29 What enzyme cofactor is associated with each of the following kinds of reactions?

(a) Transamination
(b) Carboxylation of a ketone
(c) Decarboxylation of an α-keto acid

29.30 Lactate, a product of glucose catabolism in oxygen-starved muscles, can be converted into pyruvate by oxidation. What coenzyme do you think is needed? Write the equation in the normal biochemical format using a curved arrow.



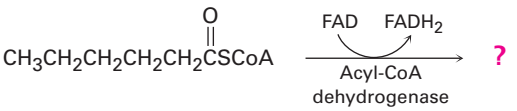
Lactate

Metabolism

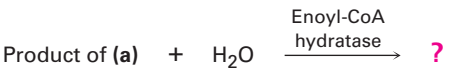
29.31 Write the equation for the final step in the β-oxidation pathway of any fatty acid with an even number of carbon atoms.

29.32 Show the products of each of the following reactions:

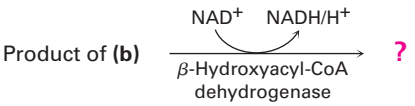
(a)



(b)



(c)



29.33 Why aren't the glycolysis and gluconeogenesis pathways the exact reverse of each other?

29.34 How many moles of acetyl CoA are produced by catabolism of the following substances?

(a) 1.0 mol of glucose **(b)** 1.0 mol of palmitic acid **(c)** 1.0 mol of maltose

29.35 How many grams of acetyl CoA (MW = 809.6 amu) are produced by catabolism of the following substances? Which substance is the most efficient precursor of acetyl CoA on a weight basis?

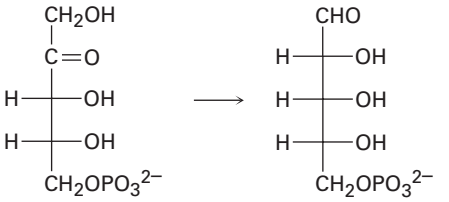
(a) 100.0 g of glucose
(b) 100.0 g of palmitic acid
(c) 100.0 g of maltose

29.36 What is the structure of the α-keto acid formed by transamination of each of the following amino acids?

(a) Threonine **(b)** Phenylalanine **(c)** Asparagine

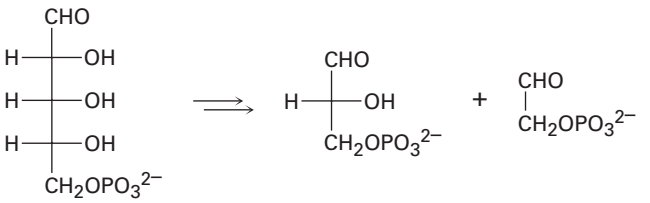
29.37 The glycolysis pathway shown in Figure 29.7 has a number of intermediates that contain phosphate groups. Why can 3-phosphoglyceryl phosphate and phosphoenolpyruvate transfer a phosphate group to ADP while glucose 6-phosphate cannot?

29.38 In the *pentose phosphate* pathway for degrading sugars, ribulose 5-phosphate is converted to ribose 5-phosphate. Propose a mechanism for the isomerization.



Ribulose 5-phosphate **Ribose 5-phosphate**

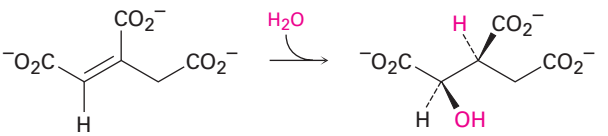
29.39 Another step in the pentose phosphate pathway for degrading sugars (see Problem 29.38) is the conversion of ribose 5-phosphate to glyceraldehyde 3-phosphate. What kind of organic process is occurring? Propose a mechanism for the conversion.



Ribose 5-phosphate **Glyceraldehyde 3-phosphate**

29.40 Write a mechanism for the conversion of α-ketoglutarate to succinyl CoA in step 4 of the citric acid cycle (Figure 29.12).

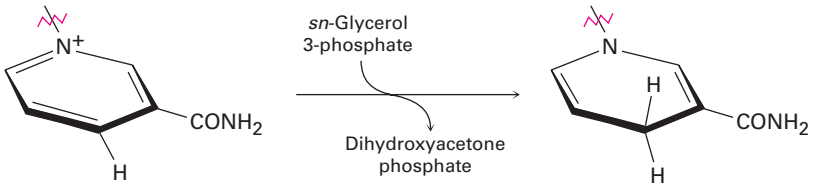
29.41 In step 2 of the citric acid cycle (Figure 29.12), *cis*-aconitate reacts with water to give (2*R*,3*S*)-isocitrate. Does −OH add from the *Re* face of the double bond or from the *Si* face? What about −H? Does the addition of water occur with syn or anti geometry?



***cis*-Aconitate** **(2*R*,3*S*)-Isocitrate**

General Problems

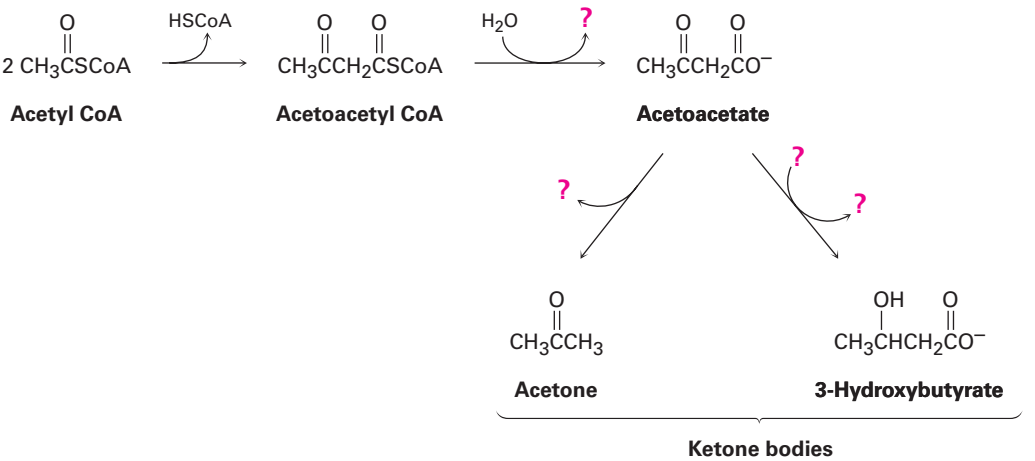
29.42 In glycerol metabolism, the oxidation of *sn*-glycerol 3-phosphate to give dihydroxyacetone phosphate is catalyzed by *sn*-glycerol-3-phosphate dehydrogenase, with NAD^+ as cofactor. The reaction is stereospecific, occurring exclusively on the *Re* face of the nicotinamide ring.



NAD^+ **NADH**

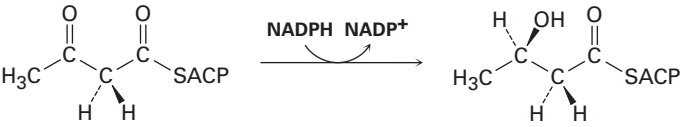
Which hydrogen in the NADH product comes from *sn*-glycerol 3-phosphate? Does it have *pro-R* or *pro-S* stereochemistry?

29.43 The primary fate of acetyl CoA under normal metabolic conditions is degradation in the citric acid cycle to yield CO_2. When the body is stressed by prolonged starvation, however, acetyl CoA is converted into compounds called *ketone bodies,* which can be used by the brain as a temporary fuel. Fill in the missing information indicated by the four question marks in the following biochemical pathway for the synthesis of ketone bodies from acetyl CoA:



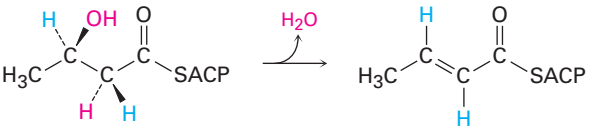
29.44 The initial reaction in Problem 29.43, conversion of two molecules of acetyl CoA to one molecule of acetoacetyl CoA, is a Claisen reaction. Assuming that there is a base present, show the mechanism of the reaction.

29.45 In step 6 of fatty-acid biosynthesis (Figure 29.5), acetoacetyl ACP is reduced stereospecifically by NADPH to yield an alcohol. Does hydride ion add to the *Si* face or the *Re* face of acetoacetyl ACP?



Acetoacetyl ACP **β-Hydroxybutyryl ACP**

29.46 In step 7 of fatty-acid biosynthesis (Figure 29.5), dehydration of a β-hydroxy thioester occurs to give *trans*-crotonyl ACP. Is the dehydration a syn elimination or an anti elimination?

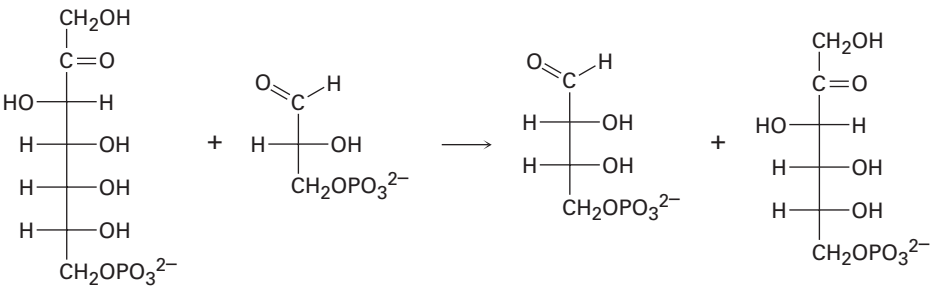


***trans*-Crotonyl ACP**

29.47 In step 8 of fatty-acid biosynthesis (Figure 29.5), reduction of *trans*-crotonyl ACP gives butyryl ACP. A hydride from NADPH adds to C3 of the crotonyl group from the *Re* face, and protonation on C2 occurs on the *Si* face. Is the reduction a syn addition or an anti addition?

Crotonyl ACP **Butyryl ACP**

29.48 One of the steps in the pentose phosphate pathway for glucose catabolism is the reaction of sedoheptulose 7-phosphate with glyceraldehyde 3-phosphate in the presence of a transaldolase to yield erythrose 4-phosphate and fructose 6-phosphate.

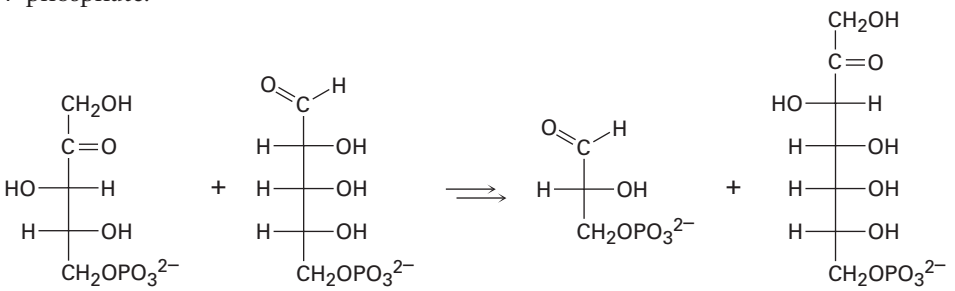


Sedoheptulose 7-phosphate **Glyceraldehyde 3-phosphate** **Erythrose 4-phosphate** **Fructose 6-phosphate**

(a) The first part of the reaction is formation of a protonated Schiff base of sedoheptulose 7-phosphate with a lysine residue in the enzyme followed by a retro-aldol cleavage to give an enamine plus erythrose 4-phosphate. Show the structure of the enamine and the mechanism by which it is formed.

(b) The second part of the reaction is nucleophilic addition of the enamine to glyceraldehyde 3-phosphate followed by hydrolysis of the Schiff base to give fructose 6-phosphate. Show the mechanism.

29.49 One of the steps in the pentose phosphate pathway for glucose catabolism is the reaction of xylulose 5-phosphate with ribose 5-phosphate in the presence of a transketolase to give glyceraldehyde 3-phosphate and sedoheptulose 7-phosphate.

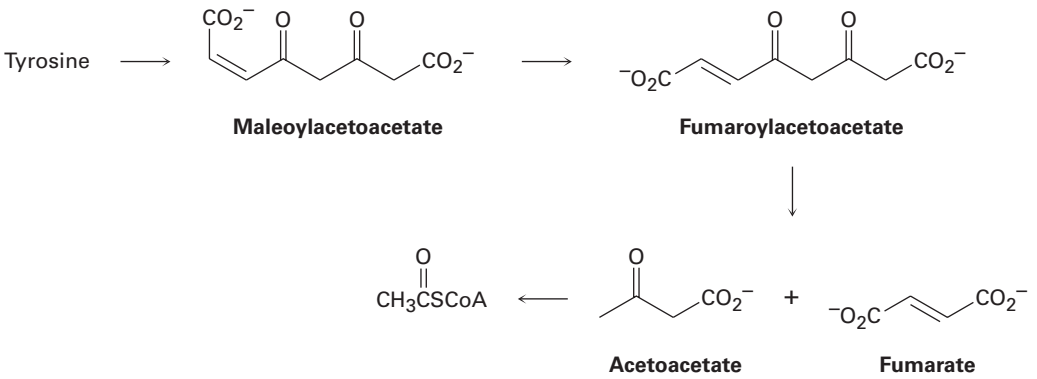


Xylulose 5-phosphate + **Ribose 5-phosphate** ⟹ **Glyceraldehyde 3-phosphate** + **Sedoheptulose 7-phosphate**

(a) The first part of the reaction is nucleophilic addition of thiamin diphosphate (TPP) ylide to xylulose 5-phosphate, followed by a retro-aldol cleavage to give glyceraldehyde 3-phosphate and a TPP-containing enamine. Show the structure of the enamine and the mechanism by which it is formed.

(b) The second part of the reaction is addition of the enamine to ribose 5-phosphate followed by loss of TPP ylide to give sedoheptulose 7-phosphate. Show the mechanism.

29.50 The amino acid tyrosine is biologically degraded by a series of steps that include the following transformations:



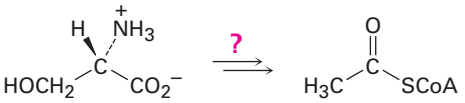
Maleoylacetoacetate **Fumaroylacetoacetate** **Acetoacetate** **Fumarate**

The double-bond isomerization of maleoylacetoacetate to fumaroylacetoacetate is catalyzed by practically any nucleophile, :Nu$^-$. Propose a mechanism.

29.51 Propose a mechanism for the conversion of fumaroylacetoacetate to fumarate plus acetoacetate (Problem 29.50).

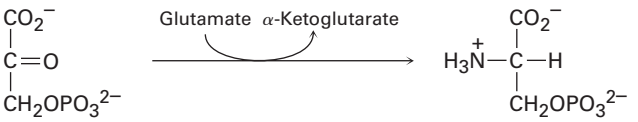
29.52 Propose a mechanism for the conversion of acetoacetate to acetyl CoA (Problem 29.50).

29.53 Design your own degradative pathway. You know the rules (organic mechanisms), and you've seen the kinds of reactions that occur in the biological degradation of fats and carbohydrates into acetyl CoA. If you were Mother Nature, what series of steps would you use to degrade the amino acid serine into acetyl CoA?



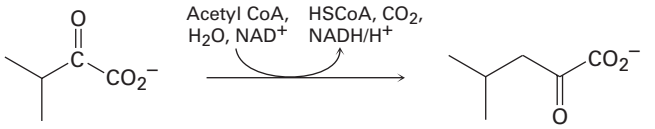
Serine **Acetyl CoA**

29.54 The amino acid serine is biosynthesized by a route that involves reaction of 3-phosphohydroxypyruvate with glutamate to give 3-phosphoserine. Propose a mechanism.



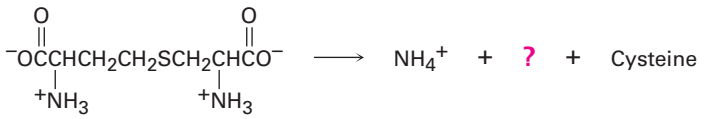
3-Phosphohydroxypyruvate **3-Phosphoserine**

29.55 The amino acid leucine is biosynthesized from α-ketoisocaproate, which is itself prepared from α-ketoisovalerate by a multistep route that involves (1) reaction with acetyl CoA, (2) hydrolysis, (3) dehydration, (4) hydration, (5) oxidation, and (6) decarboxylation. Show the steps in the transformation, and propose a mechanism for each.



α-Ketoisovalerate **α-Ketoisocaproate**

29.56 The amino acid cysteine, $C_3H_7NO_2S$, is biosynthesized from a substance called cystathionine by a multistep pathway.



Cystathionine

(a) The first step is a transamination. What is the product?
(b) The second step is an E1cB reaction. Show the products and the mechanism of the reaction.
(c) The final step is a double-bond reduction. What organic cofactor is required for this reaction, and what is the product represented by the question mark in the equation?

30

All vertebrates need vitamin D, which is synthesized by a pericyclic reaction when skin oils are exposed to sunlight. If the animal has no exposed skin, however, vitamin D is made from oily skin secretions that are deposited onto fur and then ingested during grooming.

Orbitals and Organic Chemistry: Pericyclic Reactions

Most organic reactions take place by polar mechanisms, in which a nucleophile donates two electrons to an electrophile in forming a new bond. Other reactions take place by radical mechanisms, in which each of two reactants donates one electron in forming a new bond. Both kinds of reactions occur frequently in the laboratory and in living organisms. Less common, however, is the third major class of organic reactions—*pericyclic reactions*.

A **pericyclic reaction** is one that occurs by a concerted process through a cyclic transition state. The word *concerted* means that all bonding changes occur simultaneously; no intermediates are involved. Rather than try to expand this definition now, we'll begin by briefly reviewing some of the ideas of molecular orbital theory introduced in Chapters 1 and 14 and then looking individually at the three main classes of pericyclic reactions: *electrocyclic reactions, cycloadditions,* and *sigmatropic rearrangements*.

Why This Chapter? The broad outlines of both polar and radical reactions have been known for more than a century, but our understanding of pericyclic reactions has emerged more recently. Prior to the mid-1960s, in fact, they were even referred to on occasion as "no-mechanism reactions." They occur largely in the laboratory rather than in biological processes, but a knowledge of them is necessary, both for completeness in studying organic chemistry and in understanding those biological pathways where they do occur.

30.1 Molecular Orbitals of Conjugated Pi Systems

A conjugated polyene, as we saw in **Section 14.1**, is one with alternating double and single bonds. According to molecular orbital (MO) theory, the *p* orbitals on the sp^2-hybridized carbons of a conjugated polyene interact to form a set of π molecular orbitals whose energies depend on the number of nodes they have

between nuclei. Those molecular orbitals with fewer nodes are lower in energy than the isolated *p* atomic orbitals and are *bonding MOs;* those molecular orbitals with more nodes are higher in energy than the isolated *p* orbitals and are *antibonding MOs*. Pi molecular orbitals of ethylene and 1,3-butadiene are shown in **Figure 30.1**.

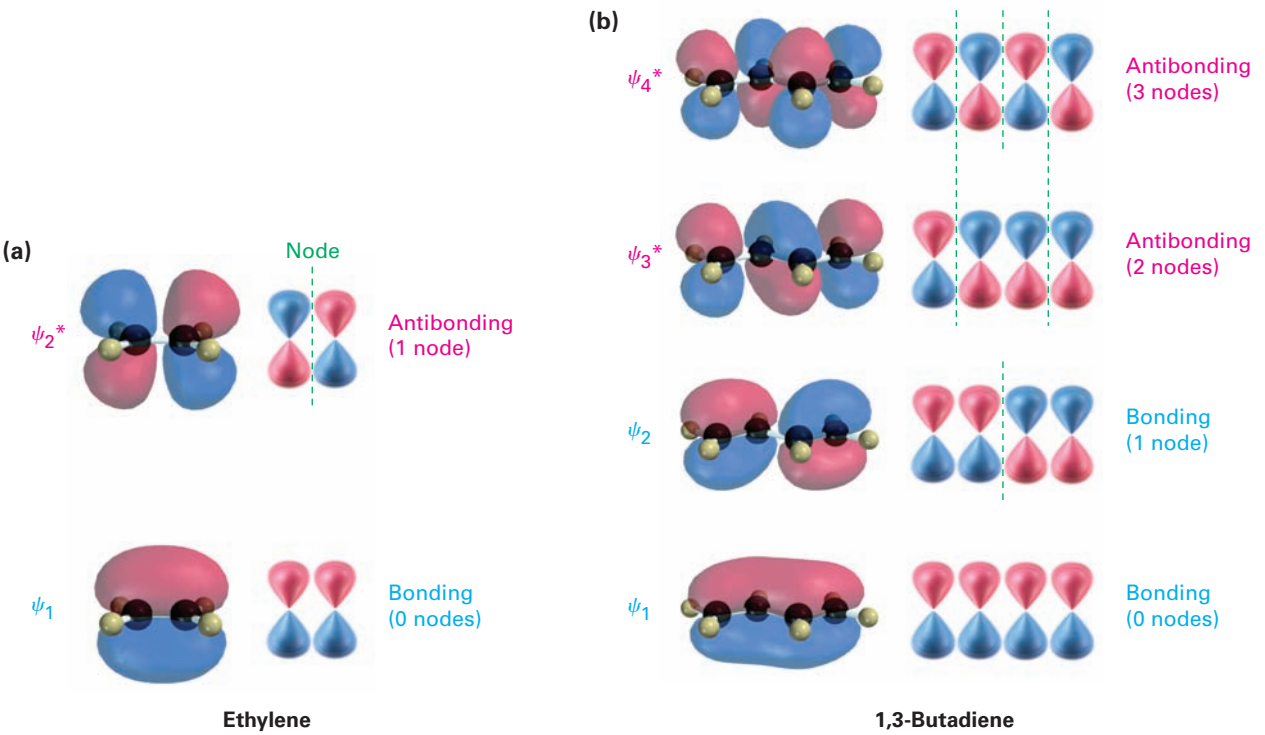


Figure 30.1 Pi molecular orbitals of **(a)** ethylene and **(b)** 1,3-butadiene.

A similar sort of molecular orbital description can be derived for any conjugated π electron system. 1,3,5-Hexatriene, for example, has three double bonds and six π MOs, as shown in **Figure 30.2**. In the ground state, only the three bonding orbitals, ψ_1, ψ_2, and ψ_3, are filled. On irradiation with ultraviolet light, however, an electron is promoted from the highest-energy filled orbital (ψ_3) to the lowest-energy unfilled orbital (ψ_4*) to give an excited state **(Section 14.7)**, in which ψ_3 and ψ_4* are each half-filled. (An asterisk denotes an antibonding orbital.)

What do molecular orbitals and their nodes have to do with pericyclic reactions? The answer is, *everything*. According to a series of rules formulated in the mid-1960s by R. B. Woodward and Roald Hoffmann, a pericyclic reaction can take place only if the symmetries of the reactant MOs are the same as the symmetries of the product MOs. In other words, the lobes of reactant MOs must be of the correct algebraic sign for bonding to occur in the transition state leading to product.

If the symmetries of reactant and product orbitals match up, or correlate, the reaction is said to be **symmetry-allowed**. If the symmetries of reactant and product orbitals don't correlate, the reaction is **symmetry-disallowed**.

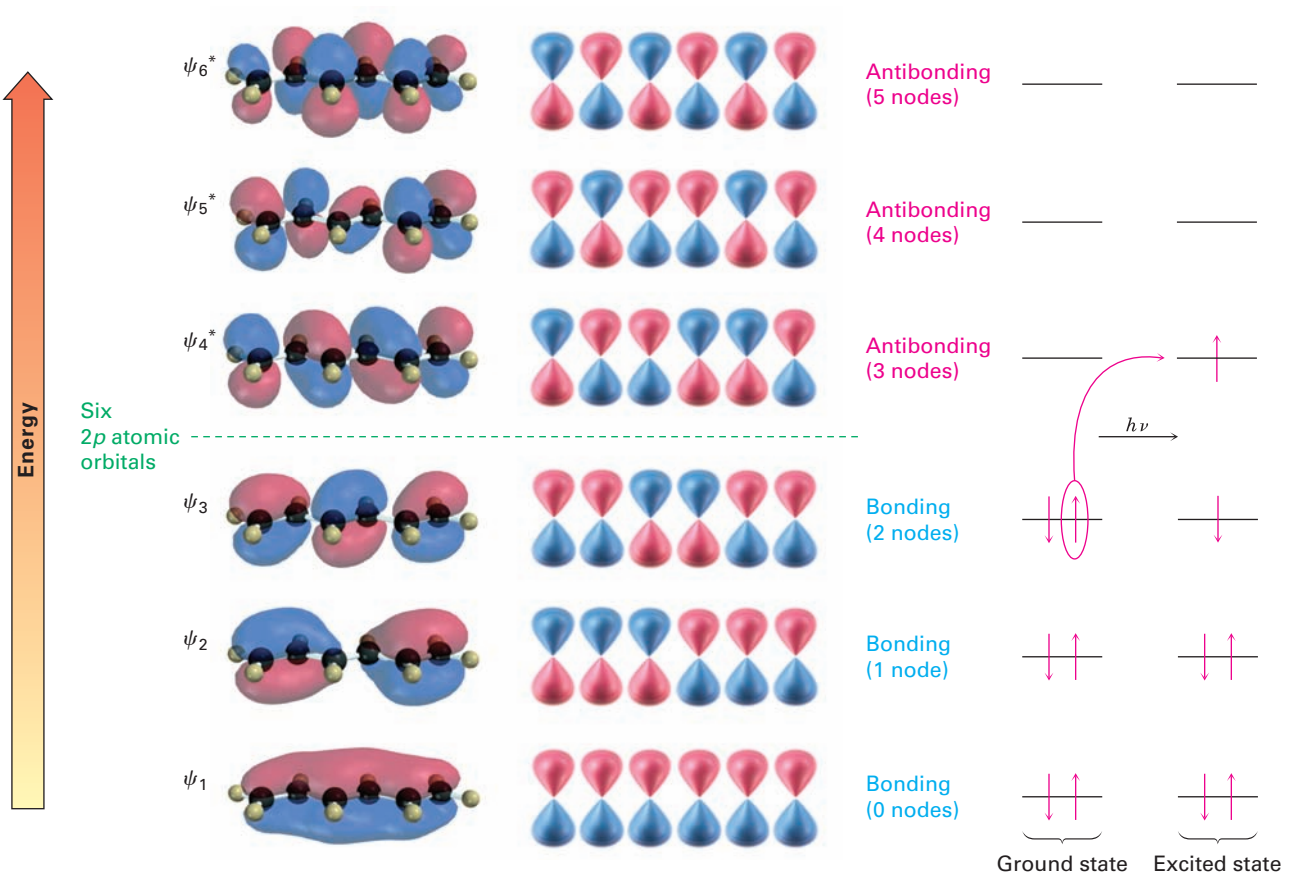


Figure 30.2 The six π molecular orbitals of 1,3,5-hexatriene. In the ground state, the three bonding MOs, ψ_1, ψ_2, and ψ_3, are filled. In the excited state, ψ_3 and ψ_4* are both half-filled.

Symmetry-allowed reactions often occur under relatively mild conditions, but symmetry-disallowed reactions can't occur by concerted paths. Either they take place by nonconcerted, higher-energy pathways, or they don't take place at all.

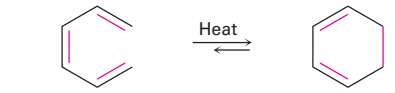
The Woodward–Hoffmann rules for pericyclic reactions require an analysis of all reactant and product molecular orbitals, but Kenichi Fukui at Kyoto Imperial University in Japan introduced a simplified version. According to Fukui, we need to consider only two molecular orbitals, called the **frontier orbitals**. These frontier orbitals are the **highest occupied molecular orbital (HOMO)** and the **lowest unoccupied molecular orbital (LUMO)**. In ground-state 1,3,5-hexatriene, for example, ψ_3 is the HOMO and ψ_4* is the LUMO (Figure 30.2). In excited-state 1,3,5-hexatriene, however, ψ_4* is the HOMO and ψ_5* is the LUMO.

Problem 30.1

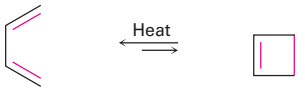
Look at Figure 30.1, and tell which molecular orbital is the HOMO and which is the LUMO for both ground and excited states of ethylene and 1,3-butadiene.

30.2 Electrocyclic Reactions

The best way to understand how orbital symmetry affects pericyclic reactions is to look at some examples. Let's look first at a group of polyene rearrangements called *electrocyclic reactions*. An **electrocyclic reaction** is a pericyclic process that involves the cyclization of a conjugated acyclic polyene. One π bond is broken, the other π bonds change position, a new σ bond is formed, and a cyclic compound results. For example, a conjugated triene can be converted into a cyclohexadiene, and a conjugated diene can be converted into a cyclobutene.



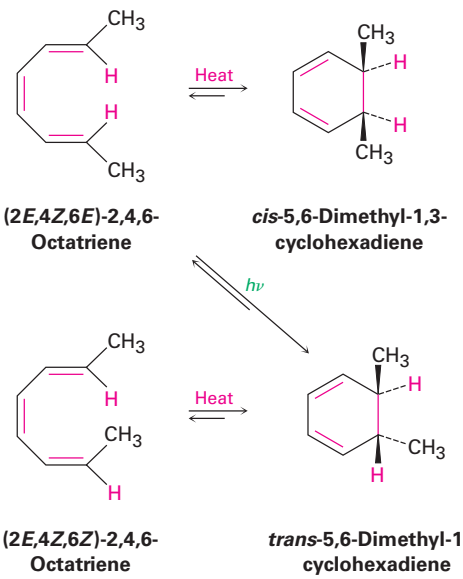
A conjugated triene **A cyclohexadiene**



A conjugated diene **A cyclobutene**

Pericyclic reactions are reversible, and the position of the equilibrium depends on the specific case. In general, the triene ⇄ cyclohexadiene equilibrium favors the cyclic product, whereas the diene ⇄ cyclobutene equilibrium favors less strained open-chain product.

The most striking feature of electrocyclic reactions is their stereochemistry. For example, (2*E*,4*Z*,6*E*)-2,4,6-octatriene yields only *cis*-5,6-dimethyl-1,3-cyclohexadiene when heated, and (2*E*,4*Z*,6*Z*)-2,4,6-octatriene yields only *trans*-5,6-dimethyl-1,3-cyclohexadiene. Remarkably, however, the stereochemical results change completely when the reactions are carried out under what are called **photochemical**, rather than thermal, conditions. Irradiation, or *photolysis,* of (2*E*,4*Z*,6*E*)-2,4,6-octatriene with ultraviolet light yields *trans*-5,6-dimethyl-1,3-cyclohexadiene **(Figure 30.3)**.



(2*E*,4*Z*,6*E*)-2,4,6-Octatriene **_cis_-5,6-Dimethyl-1,3-cyclohexadiene**

(2*E*,4*Z*,6*Z*)-2,4,6-Octatriene **_trans_-5,6-Dimethyl-1,3-cyclohexadiene**

Figure 30.3 Electrocyclic interconversions of 2,4,6-octatriene isomers and 5,6-dimethyl-1,3-cyclohexadiene isomers.

A similar result is obtained for the thermal electrocyclic ring-opening of 3,4-dimethylcyclobutene. The trans isomer yields only (2*E*,4*E*)-2,4-hexadiene when heated, and the cis isomer yields only (2*E*,4*Z*)-2,4-hexadiene. On UV irradiation, however, the results are opposite. Cyclization of the 2*E*,4*E* isomer under photochemical conditions yields cis product **(Figure 30.4)**.

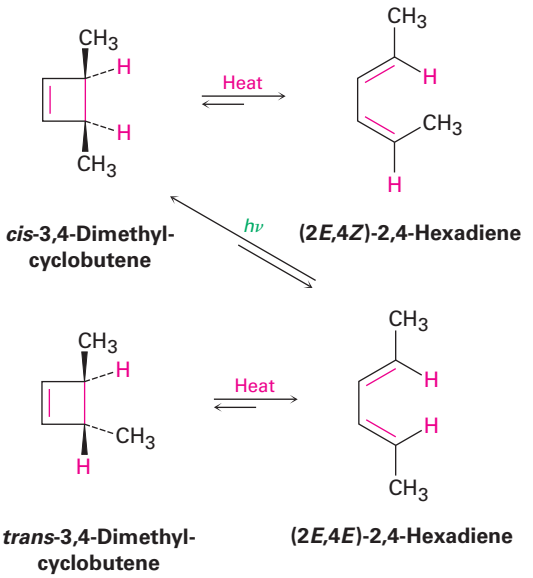
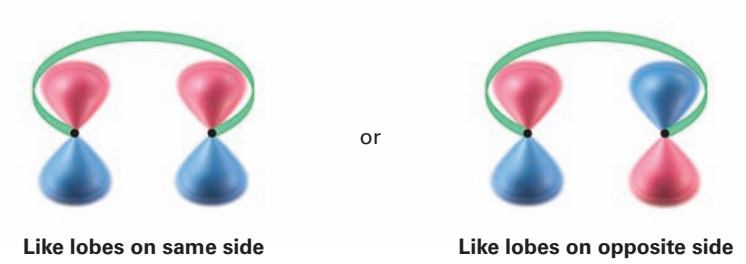


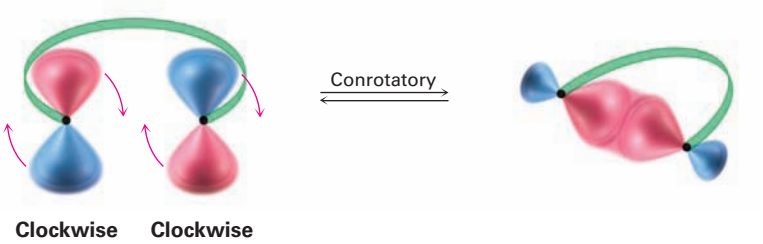
Figure 30.4 Electrocyclic interconversions of 2,4-hexadiene isomers and 3,4-dimethylcyclobutene isomers.

To account for these results, we need to look at the two outermost lobes of the polyene MOs—the lobes that interact when cyclization occurs. There are two possibilities: the lobes of like sign can be either on the same side or on opposite sides of the molecule.



For a bond to form, the outermost π lobes must rotate so that favorable bonding interaction is achieved—a positive lobe with a positive lobe or a negative lobe with a negative lobe. If two lobes of like sign are on the *same* side of the molecule, the two orbitals must rotate in *opposite* directions—one clockwise and one counterclockwise. This kind of motion is referred to as **disrotatory**.

Conversely, if lobes of like sign are on *opposite* sides of the molecule, both orbitals must rotate in the *same* direction, either both clockwise or both counterclockwise. This kind of motion is called **conrotatory**.



30.3 Stereochemistry of Thermal Electrocyclic Reactions

How can we predict whether conrotatory or disrotatory motion will occur in a given case? According to frontier orbital theory, *the stereochemistry of an electrocyclic reaction is determined by the symmetry of the polyene HOMO*. The electrons in the HOMO are the highest-energy, most loosely held electrons and are therefore most easily moved during reaction. For thermal reactions, the ground-state electron configuration is used to identify the HOMO; for photochemical reactions, the excited-state electron configuration is used.

Let's look again at the thermal ring-closure of conjugated *trienes*. According to Figure 30.2, the HOMO of a conjugated triene in its ground state has lobes of like sign on the same side of the molecule, a symmetry that predicts disrotatory ring-closure. This disrotatory cyclization is exactly what is observed in the thermal cyclization of 2,4,6-octatriene. The 2*E*,4*Z*,6*E* isomer yields cis product; the 2*E*,4*Z*,6*Z* isomer yields trans product **(Figure 30.5)**.

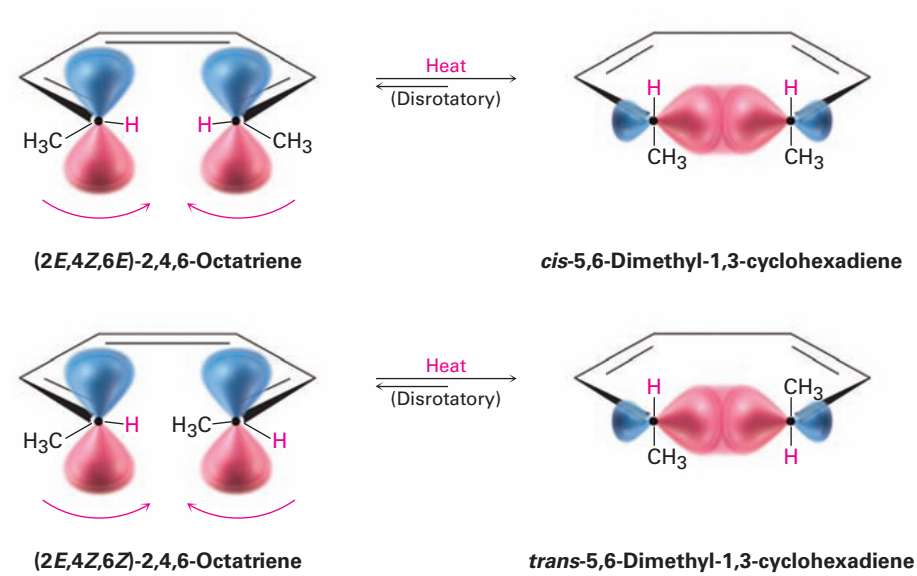


Figure 30.5 Thermal cyclizations of 2,4,6-octatrienes occur by disrotatory ring-closures.

In the same way, the ground-state HOMO of conjugated *dienes* (Figure 30.1) has a symmetry that predicts conrotatory ring-closure. In practice, however, the conjugated diene reaction can be observed only in the reverse direction (cyclobutene → diene) because of the position of the equilibrium. We therefore find that the 3,4-dimethylcyclobutene ring *opens* in a conrotatory fashion. *cis*-3,4-Dimethylcyclobutene yields (2*E*,4*Z*)-2,4-hexadiene, and *trans*-3,4-dimethylcyclobutene yields (2*E*,4*E*)-2,4-hexadiene by conrotatory opening **(Figure 30.6)**.

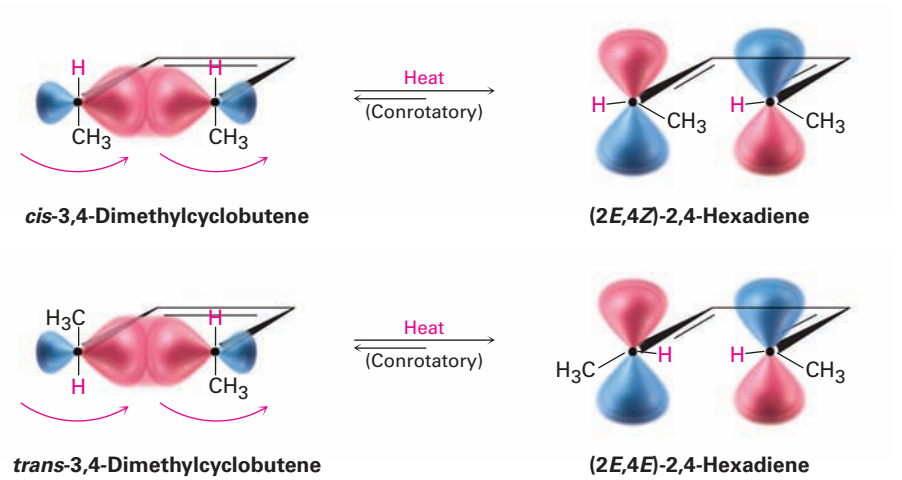
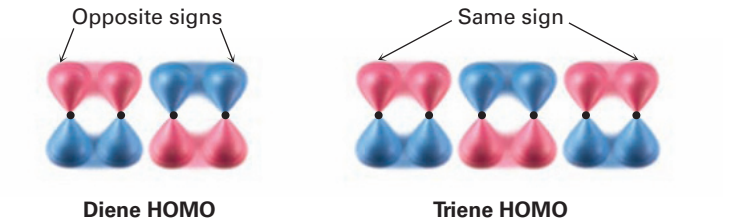


Figure 30.6 Thermal ring-openings of *cis*- and *trans*-dimethylcyclobutene occur by conrotatory paths.

Note that a conjugated diene and a conjugated triene react with opposite stereochemistry. The diene opens and closes by a conrotatory path, whereas the triene opens and closes by a disrotatory path. The difference is due to the different symmetries of the diene and triene HOMOs.



It turns out that there is an alternating relationship between the number of electron pairs (double bonds) undergoing bond reorganization and the stereochemistry of ring-opening or -closure. Polyenes with an even number of electron pairs undergo thermal electrocyclic reactions in a conrotatory sense, whereas polyenes with an odd number of electron pairs undergo the same reactions in a disrotatory sense.

Problem 30.2

Draw the products you would expect from conrotatory and disrotatory cyclizations of (2*Z*,4*Z*,6*Z*)-2,4,6-octatriene. Which of the two paths would you expect the thermal reaction to follow?

Problem 30.3

trans-3,4-Dimethylcyclobutene can open by two conrotatory paths to give either (2*E*,4*E*)-2,4-hexadiene or (2*Z*,4*Z*)-2,4-hexadiene. Explain why both products are symmetry-allowed, and then account for the fact that only the 2*E*,4*E* isomer is obtained in practice.

30.4 Photochemical Electrocyclic Reactions

We noted previously that photochemical electrocyclic reactions take a different stereochemical course than their thermal counterparts, and we can now explain this difference. Ultraviolet irradiation of a polyene causes an excitation of one electron from the ground-state HOMO to the ground-state LUMO, thus changing their symmetries. But because electronic excitation changes the symmetries of HOMO and LUMO, it also changes the reaction stereochemistry. (2*E*,4*E*)-2,4-Hexadiene, for instance, undergoes photochemical cyclization by a disrotatory path, whereas the thermal reaction is conrotatory. Similarly, (2*E*,4*Z*,6*E*)-2,4,6-octatriene undergoes photochemical cyclization by a conrotatory path, whereas the thermal reaction is disrotatory **(Figure 30.7)**.

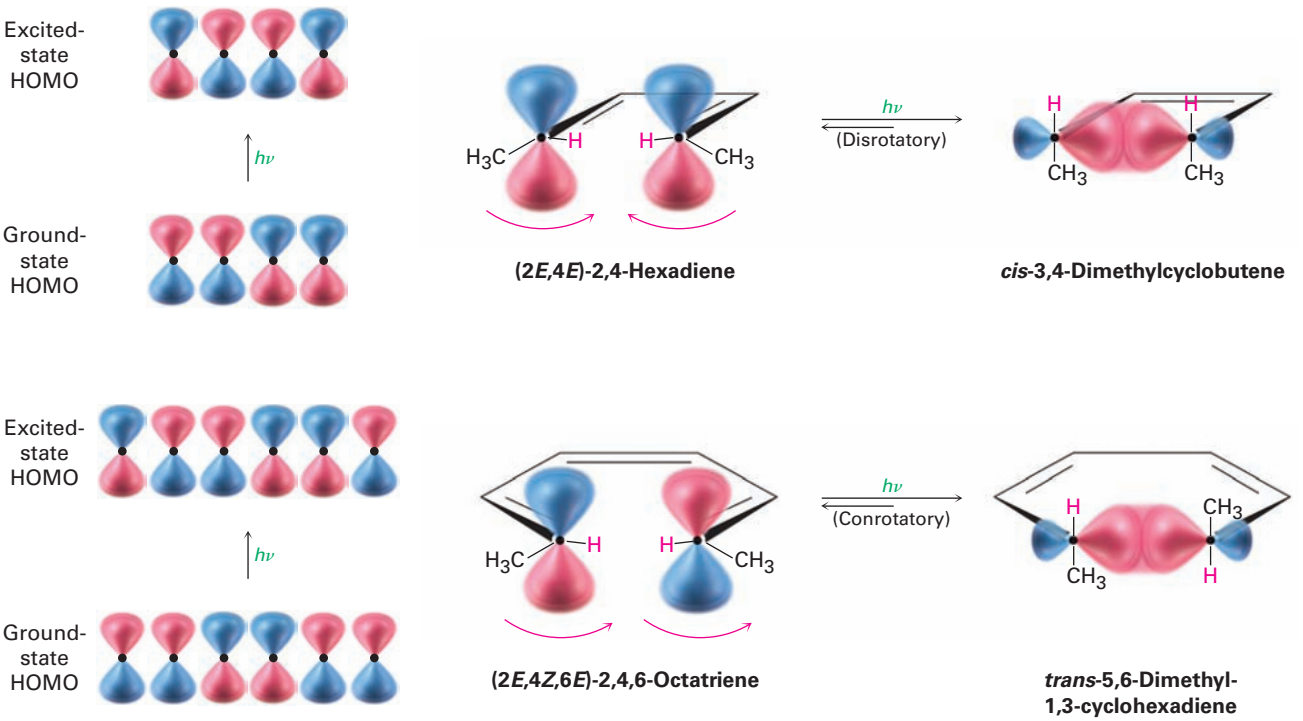


Figure 30.7 Photochemical cyclizations of conjugated dienes and trienes. The two processes occur with different stereochemistry because of their different orbital symmetries.

Thermal and photochemical electrocyclic reactions always take place with opposite stereochemistry because the symmetries of the frontier orbitals are always different. Table 30.1 gives some simple rules that make it possible to predict the stereochemistry of electrocyclic reactions.

Table 30.1 Stereochemical Rules for Electrocyclic Reactions

Electron pairs (double bonds)	Thermal reaction	Photochemical reaction
Even number	Conrotatory	Disrotatory
Odd number	Disrotatory	Conrotatory

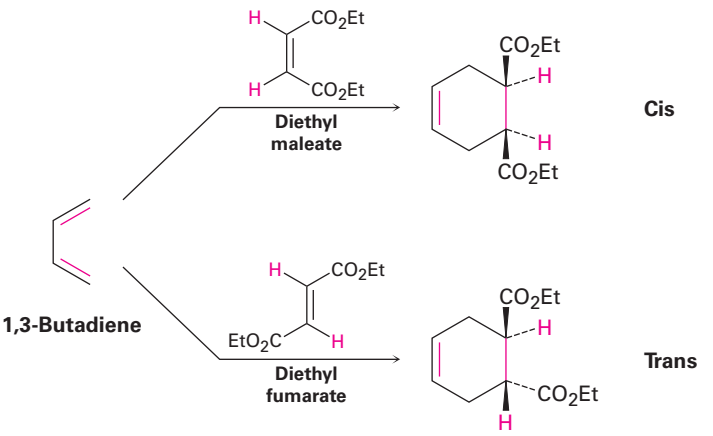
Problem 30.4

What product would you expect to obtain from the photochemical cyclization of (2*E*,4*Z*,6*E*)-2,4,6-octatriene? Of (2*E*,4*Z*,6*Z*)-2,4,6-octatriene?

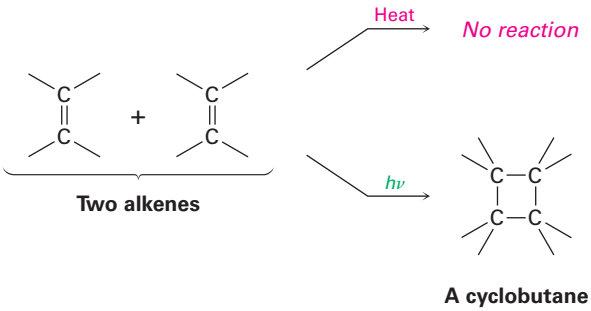
30.5 Cycloaddition Reactions

A **cycloaddition reaction** is one in which two unsaturated molecules add to one another to yield a cyclic product. As with electrocyclic reactions, cycloadditions are controlled by the orbital symmetry of the reactants. Symmetry-allowed processes often take place readily, but symmetry-disallowed processes take place with difficulty, if at all, and then only by nonconcerted pathways. Let's look at two examples to see how they differ.

The Diels–Alder cycloaddition reaction **(Section 14.4)** is a pericyclic process that takes place between a diene (four π electrons) and a dienophile (two π electrons) to yield a cyclohexene product. Many thousands of examples of Diels–Alder reactions are known. They often take place easily at room temperature or slightly above, and they are stereospecific with respect to substituents. For example, room-temperature reaction between 1,3-butadiene and diethyl maleate (cis) yields exclusively the cis-disubstituted cyclohexene product. A similar reaction between 1,3-butadiene and diethyl fumarate (trans) yields exclusively the trans-disubstituted product.



In contrast to the [4 + 2]-π-electron Diels–Alder reaction, the [2 + 2]-π-electron cycloaddition between two alkenes does not occur thermally. The [2 + 2] cycloaddition takes place only on irradiation, yielding cyclobutane products.



For a successful cycloaddition to take place, the terminal π lobes of the two reactants must have the correct symmetry for bonding to occur. This can happen in either of two ways, called *suprafacial* and *antarafacial*. **Suprafacial** cycloadditions take place when a bonding interaction occurs between lobes on the same face of one reactant and lobes on the same face of the other reactant. **Antarafacial** cycloadditions take place when a bonding interaction occurs between lobes on the same face of one reactant and lobes on *opposite* faces of the other reactant **(Figure 30.8)**.

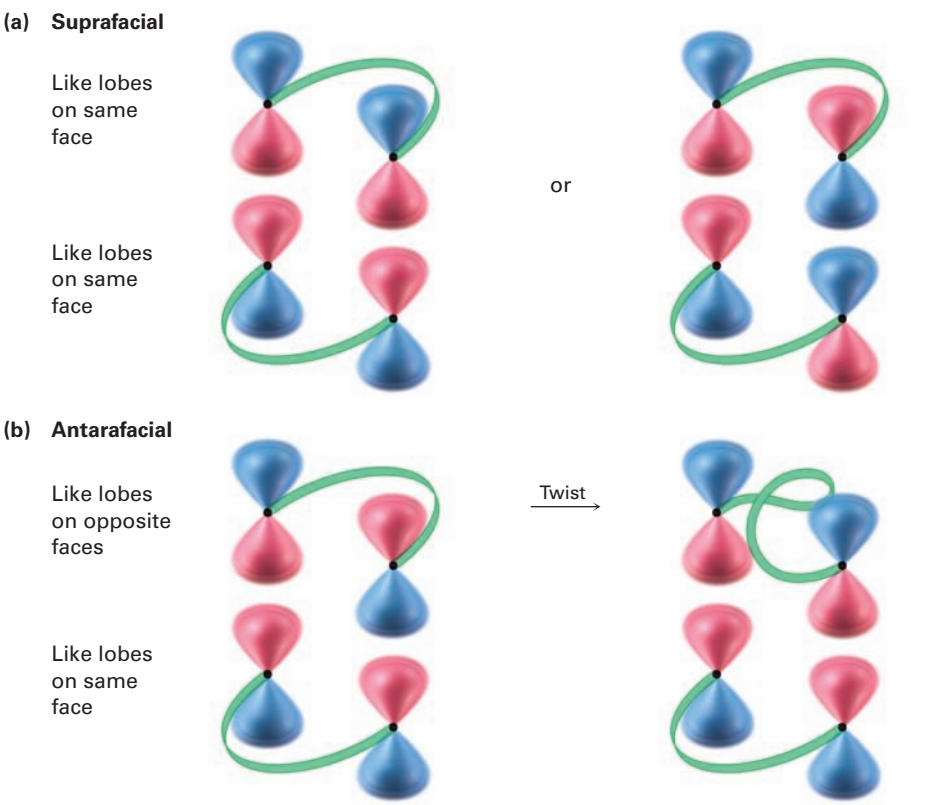


Figure 30.8 **(a)** Suprafacial cycloaddition occurs when there is bonding between lobes on the same face of one reactant and lobes on the same face of the other reactant. **(b)** Antarafacial cycloaddition occurs when there is bonding between lobes on the same face of one reactant and lobes on opposite faces of the other, which requires a twist in one π system.

Note that both suprafacial and antarafacial cycloadditions are symmetry-allowed. Geometric constraints often make antarafacial reactions difficult, however, because there must be a twisting of the π orbital system in one of the reactants. Thus, suprafacial cycloadditions are much more common for small π systems.

30.6 Stereochemistry of Cycloadditions

How can we predict whether a given cycloaddition reaction will occur with suprafacial or with antarafacial geometry? According to frontier orbital theory, a cycloaddition reaction takes place when a bonding interaction occurs between the HOMO of one reactant and the LUMO of the other. An intuitive explanation of this rule is to imagine that one reactant donates electrons to the other. As with electrocyclic reactions, it's the electrons in the HOMO of the first reactant that are least tightly held and most likely to be donated. But of course when the second reactant accepts those electrons, they must go into a vacant, unoccupied orbital—the LUMO.

For a [4 + 2] cycloaddition (Diels–Alder reaction), let's arbitrarily select the diene LUMO and the alkene HOMO. The symmetries of the two ground-state orbitals are such that bonding of the terminal lobes can occur with suprafacial geometry **(Figure 30.9)**, so the Diels–Alder reaction takes place readily under thermal conditions. Note that, as with electrocyclic reactions, we need be concerned only with the terminal lobes. For purposes of prediction, interactions among the interior lobes need not be considered.

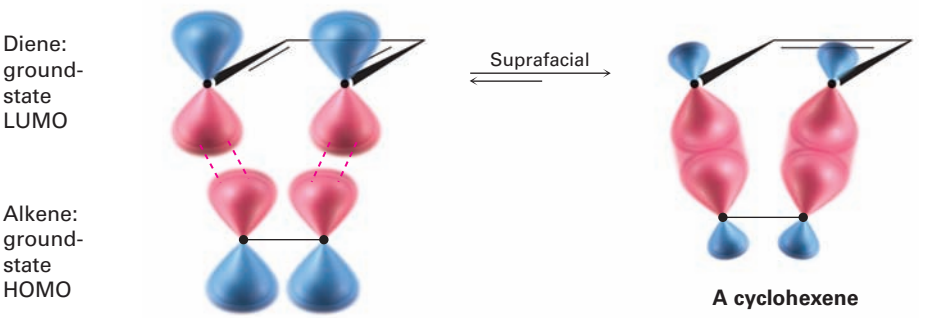


Figure 30.9 Interaction of diene LUMO and alkene HOMO in a suprafacial [4 + 2] cycloaddition reaction (Diels–Alder reaction).

In contrast with the thermal [4 + 2] Diels–Alder reaction, the [2 + 2] cycloaddition of two alkenes to yield a cyclobutane can only be observed photochemically. The explanation follows from orbital-symmetry arguments. Looking at the ground-state HOMO of one alkene and the LUMO of the second alkene, it's apparent that a thermal [2 + 2] cycloaddition must take place by an antarafacial pathway **(Figure 30.10a)**. Geometric constraints make the antarafacial transition state difficult, however, and so concerted thermal [2 + 2] cycloadditions are not observed.

In contrast with the thermal process, photochemical [2 + 2] cycloadditions *are* observed. Irradiation of an alkene with UV light excites an electron from ψ_1, the ground-state HOMO, to ψ_2^*, which becomes the excited-state HOMO. Interaction between the excited-state HOMO of one alkene and the LUMO of the second alkene allows a photochemical [2 + 2] cycloaddition reaction to occur by a suprafacial pathway (Figure 30.10b).

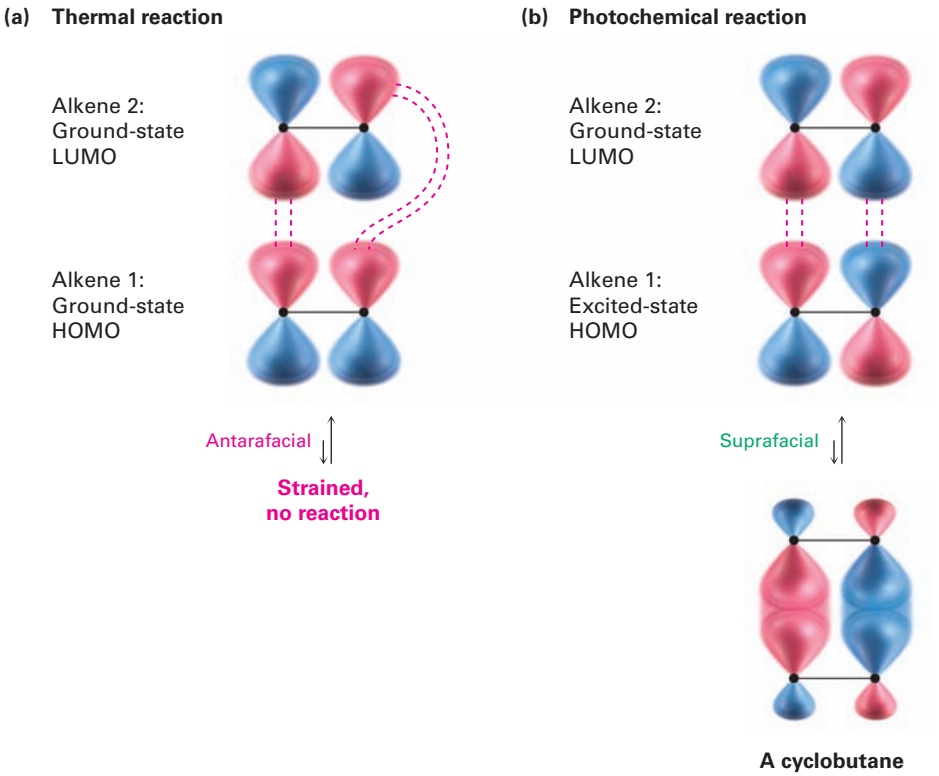
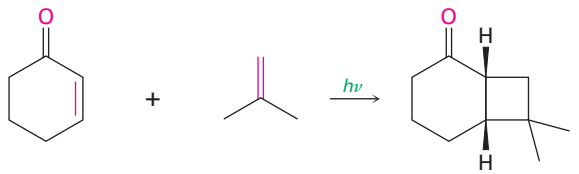


Figure 30.10 **(a)** Interaction of a ground-state HOMO and a ground-state LUMO in a potential [2 + 2] cycloaddition does not occur thermally because the antarafacial geometry is too strained. **(b)** Interaction of an excited-state HOMO and a ground-state LUMO in a photochemical [2 + 2] cycloaddition reaction is less strained, however, and occurs with suprafacial geometry.

The photochemical [2 + 2] cycloaddition reaction occurs smoothly, particularly with α,β-unsaturated carbonyl compounds, and represents one of the best methods known for synthesizing cyclobutane rings. For example:



2-Cyclohexenone **2-Methylpropene** **(40%)**

Thermal and photochemical cycloaddition reactions always take place with opposite stereochemistry. As with electrocyclic reactions, we can categorize cycloadditions according to the total number of electron pairs (double bonds) involved in the rearrangement. Thus, a thermal [4 + 2] Diels–Alder reaction between a diene and a dienophile involves an odd number (three) of electron pairs and takes place by a suprafacial pathway. A thermal [2 + 2] reaction between two alkenes involves an even number (two) of electron pairs and must take place by an antarafacial pathway. For photochemical cyclizations, these selectivities are reversed. The general rules are given in Table 30.2.

Table 30.2 Stereochemical Rules for Cycloaddition Reactions

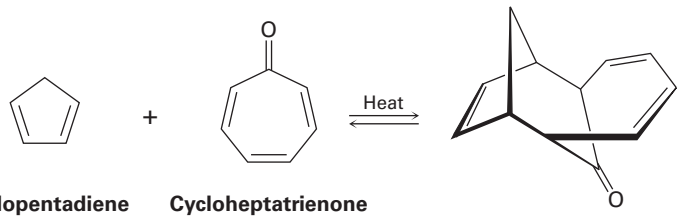
Electron pairs (double bonds)	Thermal reaction	Photochemical reaction
Even number	Antarafacial	Suprafacial
Odd number	Suprafacial	Antarafacial

Problem 30.5

What stereochemistry would you expect for the product of the Diels–Alder reaction between (2*E*,4*E*)-2,4-hexadiene and ethylene? What stereochemistry would you expect if (2*E*,4*Z*)-2,4-hexadiene were used instead?

Problem 30.6

1,3-Cyclopentadiene reacts with cycloheptatrienone to give the product shown. Tell what kind of reaction is involved, and explain the observed result. Is the reaction suprafacial or antarafacial?

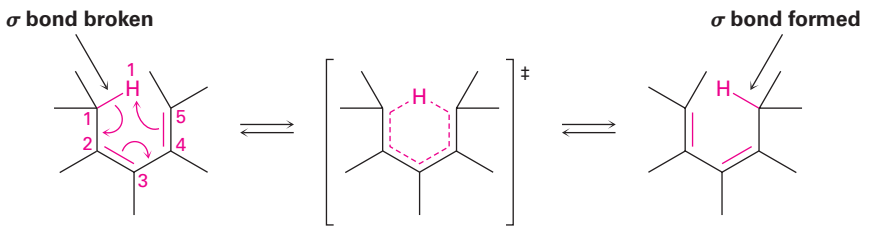


Cyclopentadiene **Cycloheptatrienone**

30.7 Sigmatropic Rearrangements

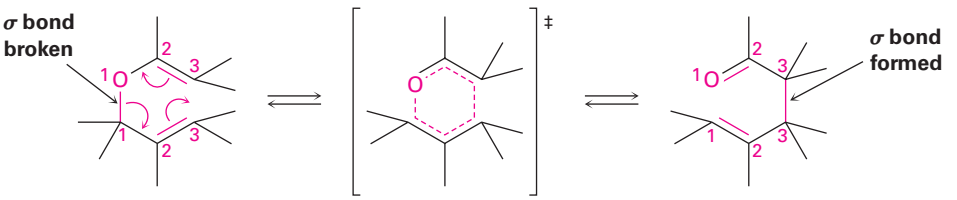
A **sigmatropic rearrangement**, the third general kind of pericyclic reaction, is a process in which a σ-bonded substituent atom or group migrates across a π electron system from one position to another. A σ bond is broken in the reactant, the π bonds move, and a new σ bond is formed in the product. The σ-bonded group can be either at the end or in the middle of the π system, as the following [1,5] and [3,3] rearrangements illustrate:

A [1,5] sigmatropic rearrangement



A 1,3-diene **Cyclic transition state** **A 1,3-diene**

A [3,3] sigmatropic rearrangement



An allylic vinylic ether **Cyclic transition state** **An unsaturated ketone**

The notations [1,5] and [3,3] describe the kind of rearrangement that is occurring. The numbers refer to the two groups connected by the σ bond in the reactant and designate the positions in those groups to which migration occurs. For example, in the [1,5] sigmatropic rearrangement of a 1,3-diene, the two groups connected by the σ bond are a hydrogen atom and a pentadienyl group. Migration occurs to position 1 of the H group (the only possibility) and to position 5 of the pentadienyl group. In the [3,3] Claisen rearrangement of an allylic vinylic ether **(Section 18.4)**, the two groups connected by the σ bond are an allylic group and a vinylic ether group. Migration occurs to position 3 of the allylic group and also to position 3 of the vinylic ether.

Like electrocyclic reactions and cycloadditions, sigmatropic rearrangements are controlled by orbital symmetries. There are two possible modes of reaction: migration of a group across the same face of the π system is suprafacial, and migration of a group from one face of the π system to the other face is antarafacial **(Figure 30.11)**.

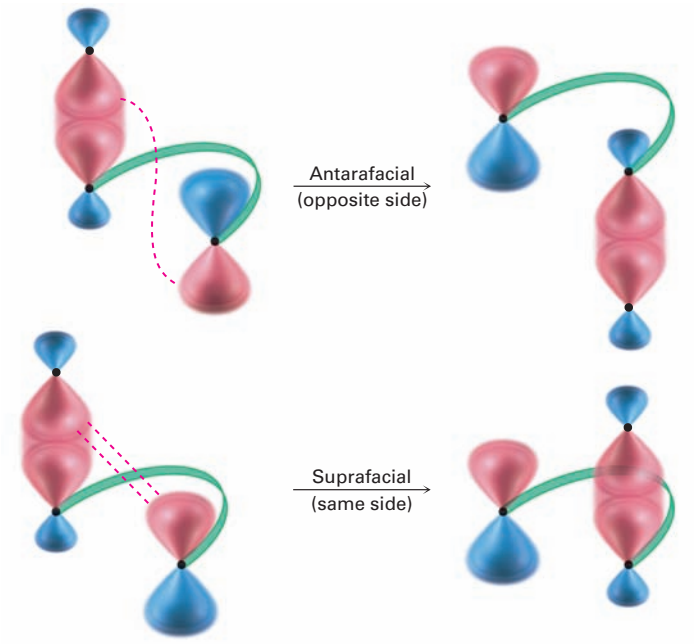


Figure 30.11 Suprafacial and antarafacial sigmatropic rearrangements.

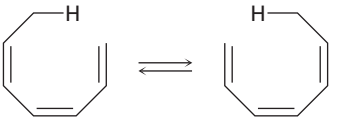
Both suprafacial and antarafacial sigmatropic rearrangements are symmetry-allowed, but suprafacial rearrangements are often easier for geometric reasons. The rules for sigmatropic rearrangements are identical to those for cycloaddition reactions (Table 30.3).

Table 30.3 Stereochemical Rules for Sigmatropic Rearrangements

Electron pairs (double bonds)	Thermal reaction	Photochemical reaction
Even number	Antarafacial	Suprafacial
Odd number	Suprafacial	Antarafacial

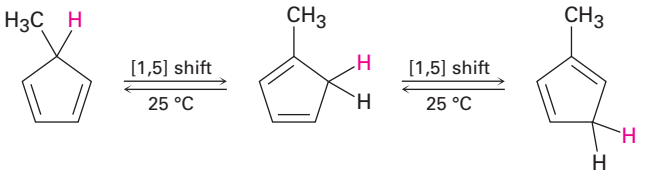
Problem 30.7

Classify the following sigmatropic reaction by order [*x*,*y*], and tell whether it will proceed with suprafacial or antarafacial stereochemistry:

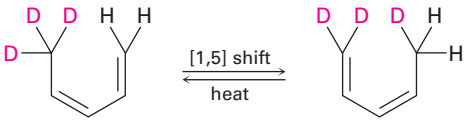


30.8 Some Examples of Sigmatropic Rearrangements

Because a [1,5] sigmatropic rearrangement involves three electron pairs (two π bonds and one σ bond), the orbital-symmetry rules in Table 30.3 predict a suprafacial reaction. In fact, the [1,5] suprafacial shift of a hydrogen atom across two double bonds of a π system is one of the most commonly observed of all sigmatropic rearrangements. For example, 5-methyl-1,3-cyclopentadiene rapidly rearranges at room temperature to yield a mixture of 1-methyl-, 2-methyl-, and 5-methyl-substituted products.



As another example, heating 5,5,5-trideuterio-(3*Z*)-1,3-pentadiene causes scrambling of deuterium between positions 1 and 5.



Both these [1,5] hydrogen shifts occur by a symmetry-allowed suprafacial pathway, as illustrated in **Figure 30.12**. In contrast with these thermal [1,*5*] sigmatropic hydrogen shifts, however, thermal [1,*3*] hydrogen shifts are unknown. Were they to occur, they would have to proceed by a strained antarafacial reaction pathway.

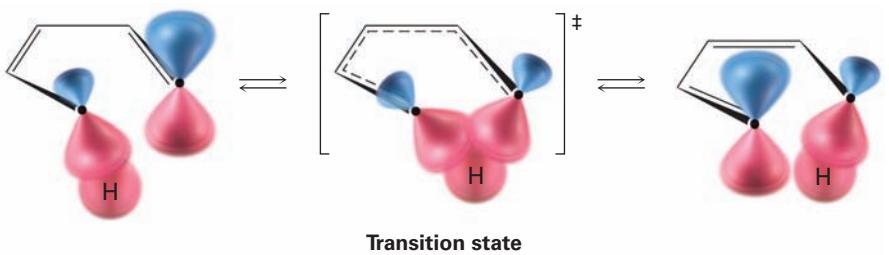
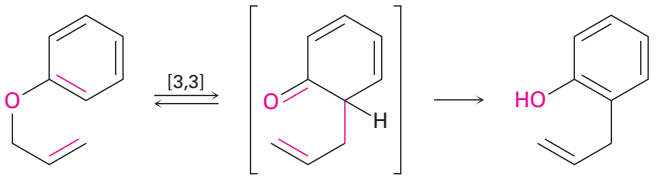


Figure 30.12 An orbital view of a suprafacial [1,5] hydrogen shift.

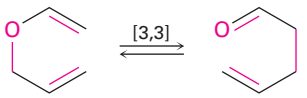
Two other important sigmatropic reactions are the *Cope rearrangement* of a 1,5-hexadiene and the *Claisen rearrangement* of an allyl aryl ether or an allyl vinyl ether discussed in **Section 18.4**. These two, along with the Diels–Alder reaction, are the most useful pericyclic reactions for organic synthesis; many thousands of examples of all three are known.

Claisen rearrangement



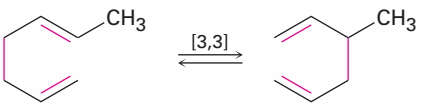
An allylic aryl ether — **An *o*-allylphenol**

Claisen rearrangement



An allylic vinylic ether — **An unsaturated ketone**

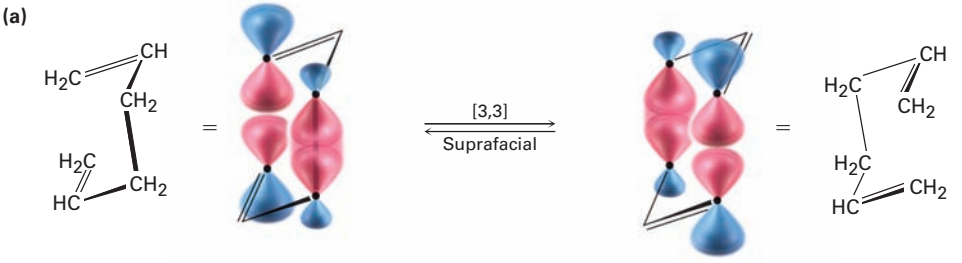
Cope rearrangement



A 1,5-diene — **An isomeric 1,5-diene**

Both Cope and Claisen rearrangements involve reorganization of an odd number of electron pairs (two π bonds and one σ bond), and both react by suprafacial pathways **(Figure 30.13)**.

Biological examples of pericyclic reactions are relatively rare, although one much-studied example occurs in bacteria during biosynthesis of the essential amino acid phenylalanine. Phenylalanine arises from the precursor chorismate through a Claisen rearrangement to prephenate, followed by decarboxylation to phenylpyruvate and reductive amination **(Figure 30.14)**. You might note that the reductive amination of phenylpyruvate is the exact reverse of the transamination process shown in Figure 29.14 on page 1199, by which amino acids are deaminated. In addition, reductive amination of ketones is a standard method for preparing amines in the laboratory, as we saw in **Section 24.6**.



Cope rearrangement of a 1,5-hexadiene

Claisen rearrangement of an allylic vinylic ether

Figure 30.13 Suprafacial [3,3] **(a)** Cope and **(b)** Claisen rearrangements.

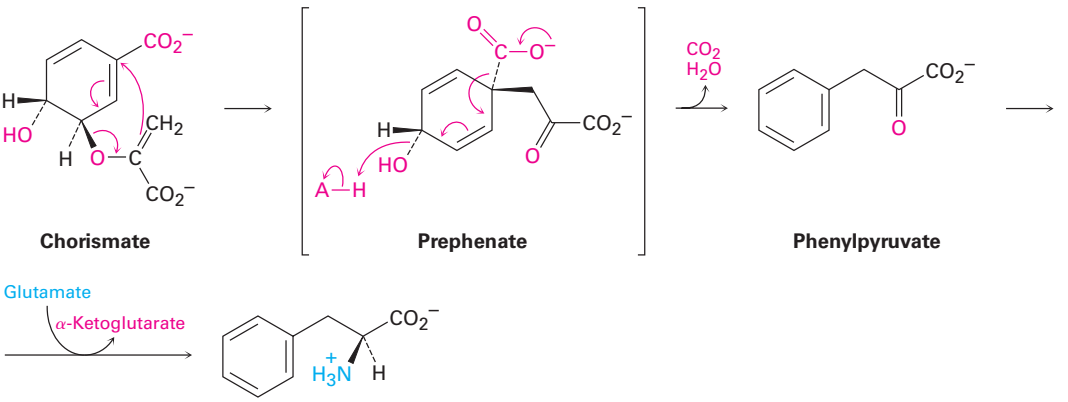
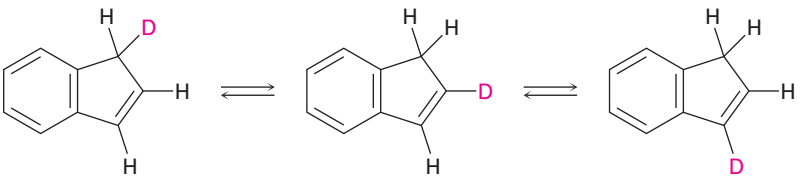


Figure 30.14 Pathway for the bacterial biosynthesis of phenylalanine from chorismate, involving a Claisen rearrangement.

Problem 30.8

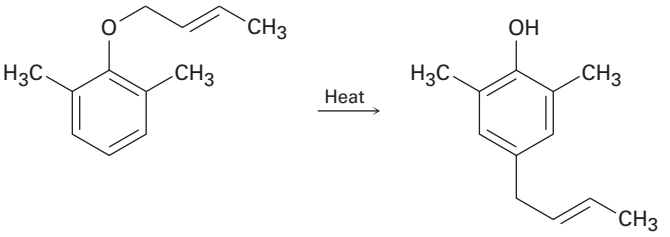
Propose a mechanism to account for the fact that heating 1-deuterioindene scrambles the isotope label to all three positions on the five-membered ring.



1-Deuterioindene

Problem 30.9

When a 2,6-disubstituted allyl phenyl ether is heated in an attempted Claisen rearrangement, migration occurs to give the *p*-allyl product as the result of two sequential pericyclic reactions. Explain.



30.9 A Summary of Rules for Pericyclic Reactions

How can you keep straight all the rules about pericyclic reactions? The summary information in Tables 30.1 to 30.3 can be distilled into one mnemonic phrase that provides an easy way to predict the stereochemical outcome of any pericyclic reaction:

The Electrons Circle Around (TECA)

*T*hermal reactions with an *E*ven number of electron pairs are *C*onrotatory or *A*ntarafacial.

A change either from thermal to photochemical or from an even to an odd number of electron pairs changes the outcome from conrotatory/antarafacial to disrotatory/suprafacial. A change from both thermal and even to photochemical and odd causes no change because two negatives make a positive.

These selection rules are summarized in Table 30.4; knowing them gives you the ability to predict the stereochemistry of literally thousands of pericyclic reactions.

Table 30.4 Stereochemical Rules for Pericyclic Reactions

Electronic state	Electron pairs	Stereochemistry
Ground state (thermal)	Even number	Antara–con
	Odd number	Supra–dis
Excited state (photochemical)	Even number	Supra–dis
	Odd number	Antara–con

Problem 30.10

Predict the stereochemistry of the following pericyclic reactions:

(a) The thermal cyclization of a conjugated tetraene
(b) The photochemical cyclization of a conjugated tetraene
(c) A photochemical [4 + 4] cycloaddition
(d) A thermal [2 + 6] cycloaddition
(e) A photochemical [3,5] sigmatropic rearrangement

Synthesizing vitamin D takes dedication and hard work.

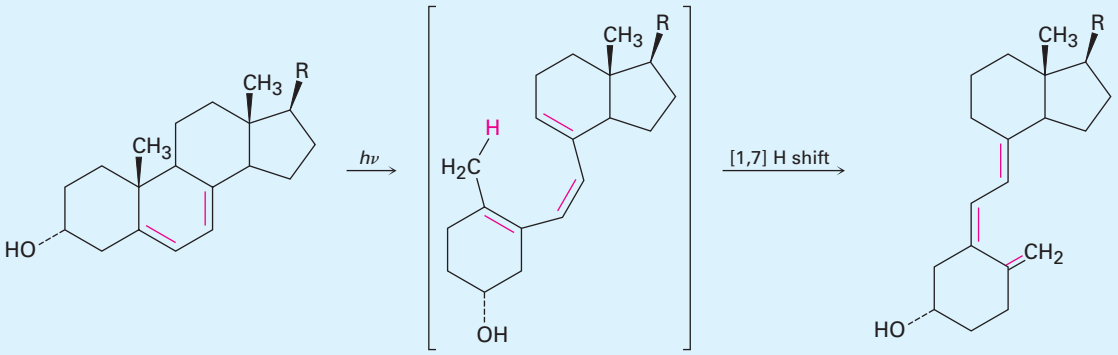
Vitamin D, the Sunshine Vitamin

A DEEPER LOOK

Vitamin D, discovered in 1918, is a general name for two related compounds, *cholecalciferol* (vitamin D_3) and *ergocalciferol* (vitamin D_2). Both are derived from steroids **(Section 27.6)** and differ only in the nature of the hydrocarbon side chain attached to the five-membered ring. Cholecalciferol comes primarily from dairy products and fish; ergocalciferol comes from some vegetables.

The function of vitamin D in the body is to control the calcification of bones by increasing intestinal absorption of calcium. When sufficient vitamin D is present, approximately 30% of ingested calcium is absorbed, but in the absence of vitamin D, calcium absorption falls to about 10%. A deficiency of vitamin D thus leads to poor bone growth and to the diseases *rickets* in children and *osteoporosis* in adults.

Actually, neither vitamin D_2 nor D_3 is present in foods. Rather, foods contain the precursor molecules 7-dehydrocholesterol and ergosterol. In the presence of sunlight, both precursors are converted in the outer, epidermal layer of skin to the active vitamins, hence the nickname for vitamin D, the "sunshine vitamin."



7-Dehydrocholesterol
Ergosterol

R = $CH(CH_3)CH_2CH_2CH_2CH(CH_3)_2$
R = $CH(CH_3)CH{=}CHCH(CH_3)CH(CH_3)_2$

Cholecalciferol
Ergocalciferol

Pericyclic reactions are unusual in living organisms, and the photochemical synthesis of vitamin D is one of only a few well-studied examples. The reaction takes place in two steps, an electrocyclic ring-opening of a cyclohexadiene to yield an open-chain hexatriene, followed by a sigmatropic [1,7] H shift to yield an isomeric hexatriene. Only the initial, electrocyclic ring-opening requires irradiation, with so-called UVB light of 295 to 300 nm wavelength required. The subsequent sigmatropic [1,7] H shift occurs spontaneously by a thermal isomerization.

Following synthesis under the skin, further metabolic processing of cholecalciferol and ergocalciferol in the liver and kidney introduces two additional —OH groups to give the active forms of the vitamin, calcitriol and ergocalcitriol.

Summary

A **pericyclic reaction** takes place in a single step through a cyclic transition state without intermediates. There are three major classes of pericyclic processes: electrocyclic reactions, cycloaddition reactions, and sigmatropic rearrangements. The stereochemistry of these reactions is controlled by the symmetry of the orbitals involved in bond reorganization.

Electrocyclic reactions involve the cyclization of conjugated acyclic polyenes. For example, 1,3,5-hexatriene cyclizes to 1,3-cyclohexadiene on heating. Electrocyclic reactions can occur by either **conrotatory** or **disrotatory** pathways, depending on the symmetry of the terminal lobes of the π system. Conrotatory cyclization requires that both lobes rotate in the same direction, whereas disrotatory cyclization requires that the lobes rotate in opposite directions. The reaction course in a specific case can be found by looking at the symmetry of the **highest occupied molecular orbital (HOMO)**.

Cycloaddition reactions are those in which two unsaturated molecules add together to yield a cyclic product. For example, Diels–Alder reaction between a diene (four π electrons) and a dienophile (two π electrons) yields a cyclohexene. Cycloadditions can take place either by **suprafacial** or **antarafacial** pathways. Suprafacial cycloaddition involves interaction between lobes on the same face of one component and on the same face of the second component. Antarafacial cycloaddition involves interaction between lobes on the same face of one component and on opposite faces of the other component. The reaction course in a specific case can be found by looking at the symmetry of the HOMO of one component and the **lowest unoccupied molecular orbital (LUMO)** of the other component.

Sigmatropic rearrangements involve the migration of a σ-bonded group across a π electron system. For example, Claisen rearrangement of an allylic vinylic ether yields an unsaturated carbonyl compound, and Cope rearrangement of a 1,5-hexadiene yields an isomeric 1,5-hexadiene. Sigmatropic rearrangements can occur with either suprafacial or antarafacial stereochemistry; the selection rules for a given case are the same as those for cycloaddition reactions.

The stereochemistry of any pericyclic reaction can be predicted by counting the total number of electron pairs (bonds) involved in bond reorganization and then applying the mnemonic "The Electrons Circle Around." That is, **thermal** (ground-state) reactions involving an even number of electron pairs occur with either conrotatory or antarafacial stereochemistry. Exactly the opposite rules apply to **photochemical** (excited-state) reactions.

Key words

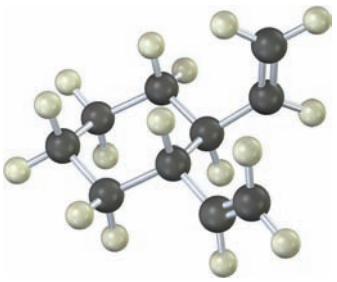
Exercises

OWL Interactive versions of these problems are assignable in OWL for Organic Chemistry.

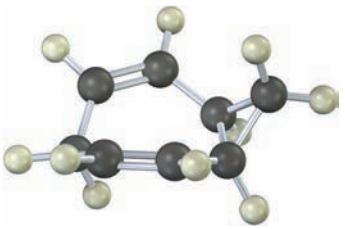
Visualizing Chemistry

(Problems 30.1–30.10 appear within the chapter.)

30.11 Predict the product obtained when the following substance is heated:



30.12 The ^{13}C NMR spectrum of homotropilidene taken at room temperature shows only three peaks. Explain.

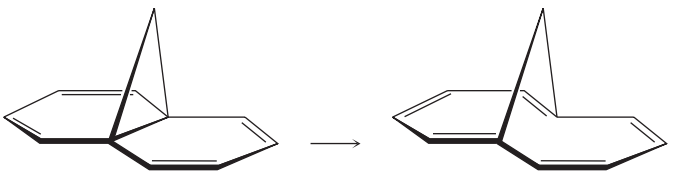


Additional Problems

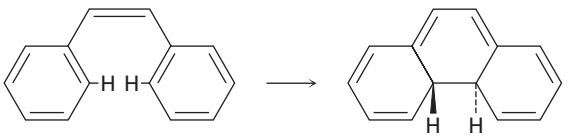
Electrocyclic Reactions

30.13 Have the following electrocyclic reactions taken place in a conrotatory or disrotatory manner? Under what conditions, thermal or photochemical, would you carry out each reaction?

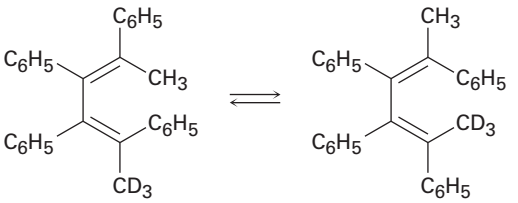
(a)



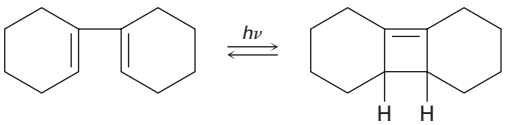
(b)



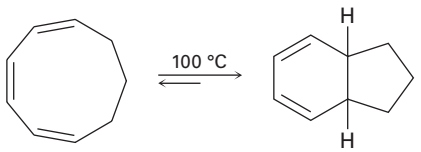
30.14 The following thermal isomerization occurs under relatively mild conditions. Identify the pericyclic reactions involved, and show how the rearrangement occurs.



30.15 Would you expect the following reaction to proceed in a conrotatory or disrotatory manner? Show the stereochemistry of the cyclobutene product, and explain your answer.



30.16 Heating (1*Z*,3*Z*,5*Z*)-1,3,5-cyclononatriene to 100 °C causes cyclization and formation of a bicyclic product. Is the reaction conrotatory or disrotatory? What is the stereochemical relationship of the two hydrogens at the ring junctions, cis or trans?

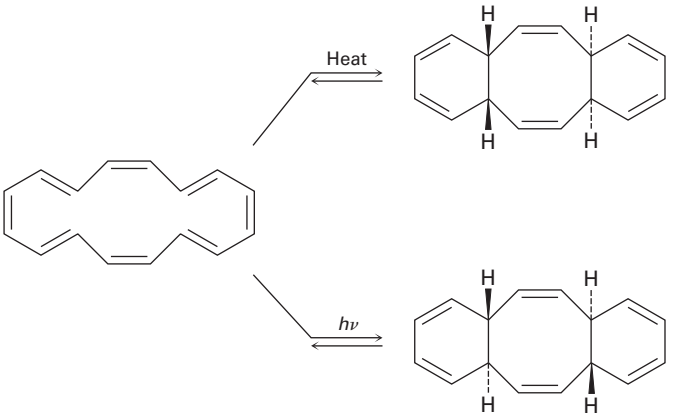


(1*Z*,3*Z*,5*Z*)-1,3,5-Cyclononatriene

30.17 (2*E*,4*Z*,6*Z*,8*E*)-2,4,6,8-Decatetraene has been cyclized to give 7,8-dimethyl-1,3,5-cyclooctatriene. Predict the manner of ring-closure—conrotatory or disrotatory—for both thermal and photochemical reactions, and predict the stereochemistry of the product in each case.

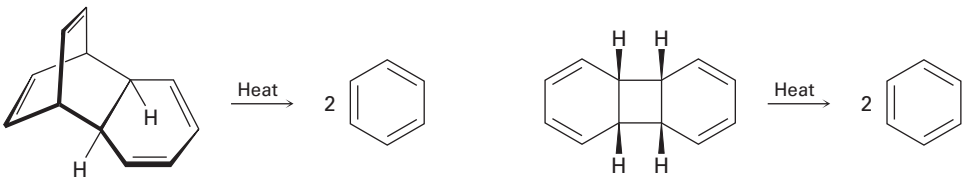
30.18 Answer Problem 30.17 for the thermal and photochemical cyclizations of (2*E*,4*Z*,6*Z*,8*Z*)-2,4,6,8-decatetraene.

30.19 The cyclohexadecaoctaene shown isomerizes to two different isomers, depending on reaction conditions. Explain the observed results, and indicate whether each reaction is conrotatory or disrotatory.

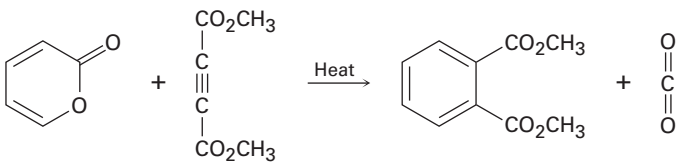


Cycloaddition Reactions

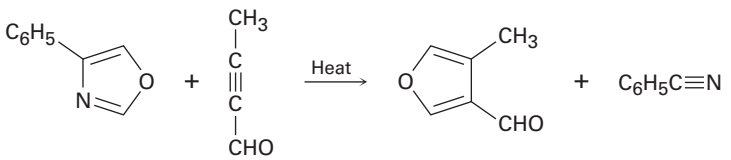
30.20 Which of the following reactions is more likely to occur? Explain.



30.21 The following reaction takes place in two steps, one of which is a cycloaddition and the other of which is a *reverse* cycloaddition. Identify the two pericyclic reactions, and show how they occur.

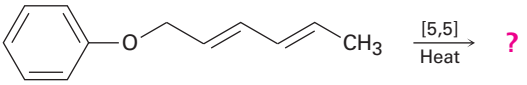


30.22 Two sequential pericyclic reactions are involved in the following furan synthesis. Identify them, and propose a mechanism for the transformation.

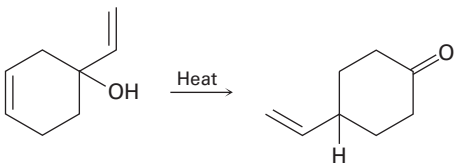


Sigmatropic Rearrangements

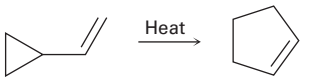
30.23 Predict the product of the following pericyclic reaction. Is this [5,5] shift a suprafacial or an antarafacial process?



30.24 Propose a pericyclic mechanism to account for the following transformation:

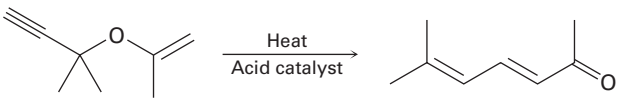


30.25 Vinyl-substituted cyclopropanes undergo thermal rearrangement to yield cyclopentenes. Propose a mechanism for the reaction, and identify the pericyclic process involved.

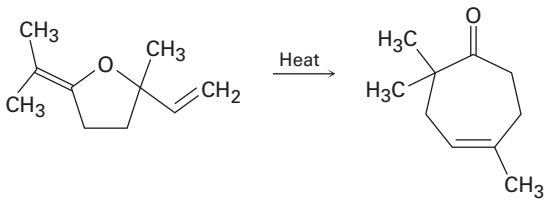


Vinylcyclopropane **Cyclopentene**

30.26 The following synthesis of dienones occurs readily. Propose a mechanism to account for the results, and identify the kind of pericyclic reaction involved.



30.27 Karahanaenone, a terpenoid isolated from oil of hops, has been synthesized by the thermal reaction shown. Identify the kind of pericyclic reaction, and explain how karahanaenone is formed.



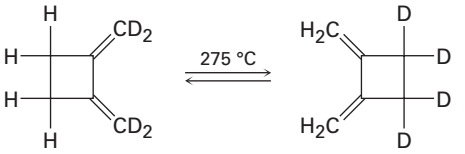
Karahanaenone

General Problems

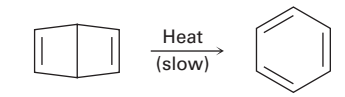
30.28 What stereochemistry—antarafacial or suprafacial—would you expect to observe in the following reactions?

(a) A photochemical [1,5] sigmatropic rearrangement
(b) A thermal [4 + 6] cycloaddition
(c) A thermal [1,7] sigmatropic rearrangement
(d) A photochemical [2 + 6] cycloaddition

30.29 The following thermal rearrangement involves two pericyclic reactions in sequence. Identify them, and propose a mechanism to account for the observed result.

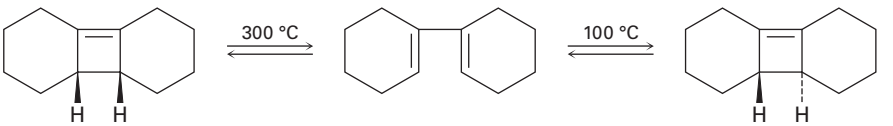


30.30 Bicyclohexadiene, also known as *Dewar benzene,* is extremely stable despite the fact that its rearrangement to benzene is energetically favored. Explain why the rearrangement is so slow.

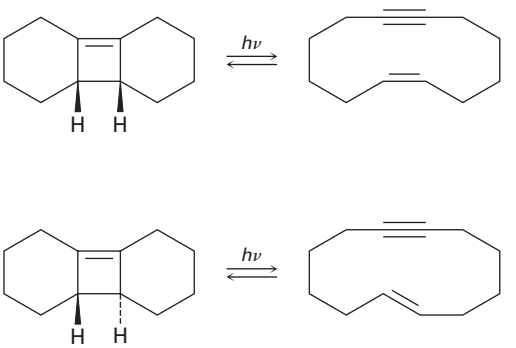


Dewar benzene **Benzene**

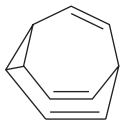
30.31 Ring-opening of the *trans*-cyclobutene isomer shown takes place at much lower temperature than a similar ring-opening of the *cis*-cyclobutene isomer. Explain the temperature effect, and identify the stereochemistry of each reaction as either conrotatory or disrotatory.



30.32 Photolysis of the *cis*-cyclobutene isomer in Problem 30.31 yields *cis*-cyclododecaen-7-yne, but photolysis of the trans isomer yields *trans*-cyclododecaen-7-yne. Explain these results, and identify the type and stereochemistry of the pericyclic reaction.

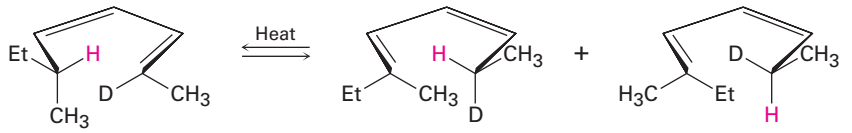


30.33 The ^{1}H NMR spectrum of bullvalene at 100 °C consists only of a single peak at 4.22 δ. Explain.

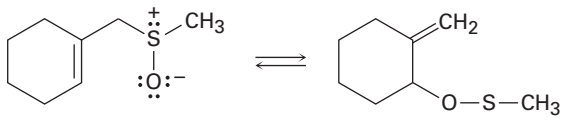


Bullvalene

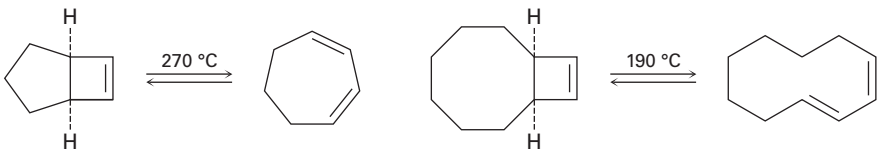
30.34 The following rearrangement was devised and carried out to prove the stereochemistry of [1,5] sigmatropic hydrogen shifts. Explain how the observed result confirms the predictions of orbital symmetry.



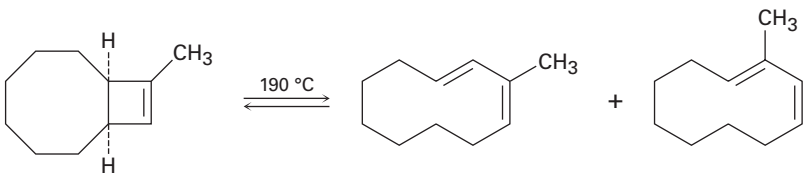
30.35 The following reaction is an example of a [2,3] sigmatropic rearrangement. Would you expect the reaction to be suprafacial or antarafacial? Explain.



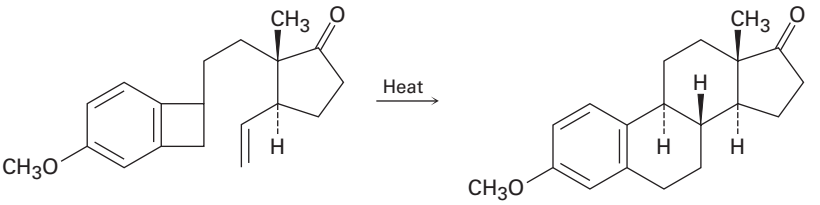
30.36 When the compound having a cyclobutene fused to a five-membered ring is heated, (1*Z*,3*Z*)-1,3-cycloheptadiene is formed. When the related compound having a cyclobutene fused to an eight-membered ring is heated, however, (1*E*,3*Z*)-1,3-cyclodecadiene is formed. Explain these results, and suggest a reason why opening of the eight-membered ring occurs at a lower temperature.



30.37 In light of your answer to Problem 30.36, explain why a mixture of products occurs in the following reaction:

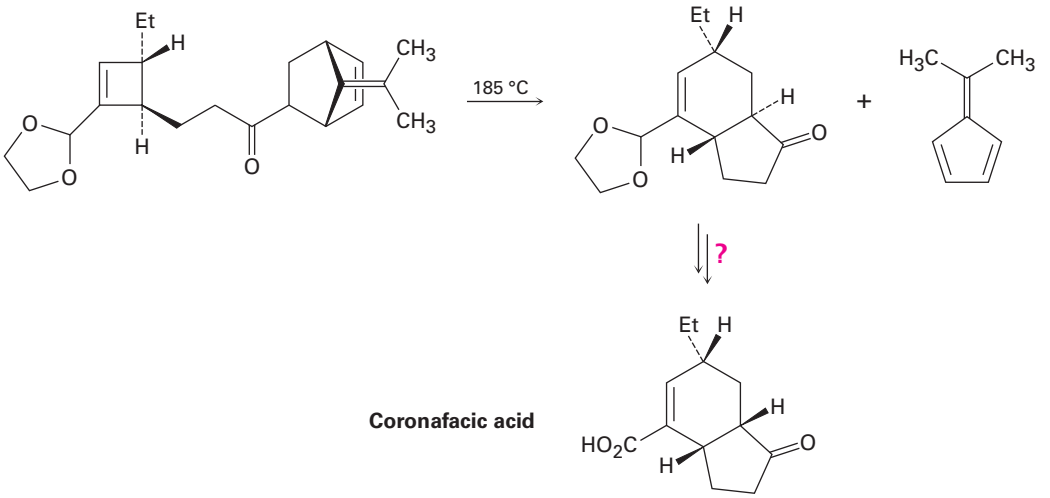


30.38 The sex hormone estrone has been synthesized by a route that involves the following step. Identify the pericyclic reactions involved, and propose a mechanism.



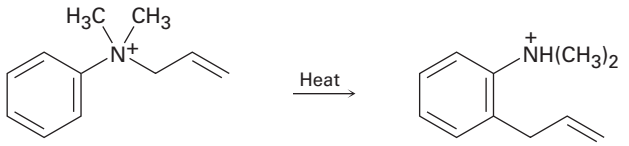
Estrone methyl ether

30.39 Coronafacic acid, a bacterial toxin, was synthesized using a key step that involves three sequential pericyclic reactions. Identify them, and propose a mechanism for the overall transformation. How would you complete the synthesis?



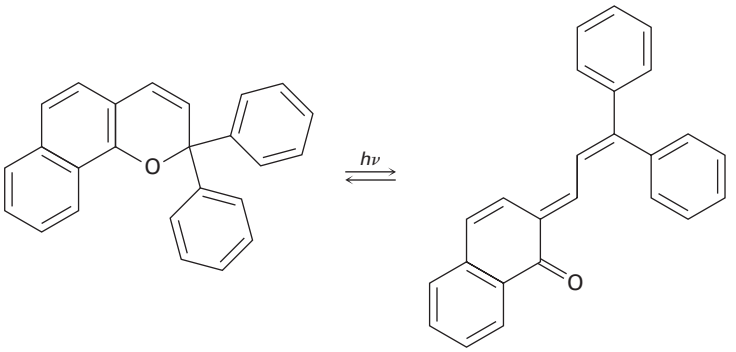
Coronafacic acid

30.40 The following rearrangement of *N*-allyl-*N*,*N*-dimethylanilinium ion has been observed. Propose a mechanism.



***N*-Allyl-*N*,*N*-dimethylanilinium ion** ***o*-Allyl-*N*,*N*-dimethylanilinium ion**

30.41 Plastic photochromic sunglasses are based on the following reversible rearrangement of a dye inside the lenses that occurs when the lenses are exposed to sunlight. The original dye absorbs UV light but not visible light and is thus colorless, while the rearrangement product absorbs visible light and is thus darkened.



(Colorless) **(Dark)**

(a) Show the mechanism of the rearrangement.
(b) Why does the rearrangement product absorb at a longer wavelength (visible light) than the original dye (UV)?

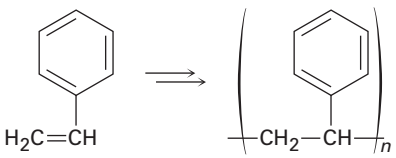
31

If you ride a bike, wear your helmet! Most bike helmets are made of two different polymers, a hard polycarbonate shell and an inner layer of polystyrene. © Dattatreya/Alamy

Synthetic Polymers

Polymers are a fundamental part of the modern world, used in everything from coffee cups to cars to clothing. In medicine, too, their importance is growing for purposes as diverse as cardiac pacemakers, artificial heart valves, and biodegradable sutures.

We've seen on several occasions in previous chapters that a **polymer**, whether synthetic or biological, is a large molecule built up by repetitive bonding together of many smaller units, or **monomers**. Polyethylene, for instance, is a synthetic polymer made from ethylene **(Section 8.10)**, nylon is a synthetic polyamide made from a diacid and a diamine **(Section 21.9)**, and proteins are biological polyamides made from amino acids. Note that polymers are often drawn by indicating their repeating unit in parentheses. The repeat unit in polystyrene, for example, comes from the monomer styrene.



Styrene **Polystyrene**

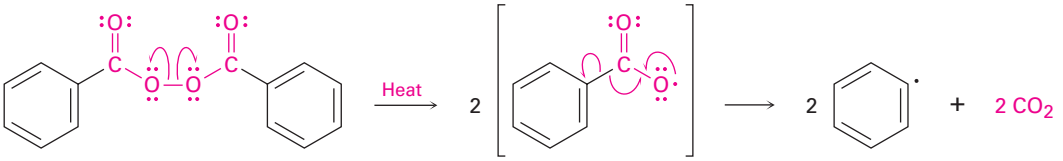
Why This Chapter? Our treatment of polymers has thus far been dispersed over several chapters, but it's also important to take a more comprehensive view. In the present chapter, we'll look further at how polymers are made, and we'll see how polymer structure correlates with physical properties. No course in organic chemistry would be complete without a look at polymers.

31.1 Chain-Growth Polymers

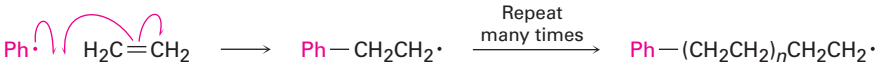
Synthetic polymers are classified by their method of synthesis as either *chain-growth* or *step-growth*. The categories are somewhat imprecise but nevertheless provide a useful distinction. Chain-growth polymers are produced by chain-reaction polymerization in which an initiator adds to a carbon–carbon double bond of an unsaturated substrate (a *vinyl monomer*) to yield a reactive intermediate. This intermediate reacts with a second molecule of monomer to yield a new intermediate, which reacts with a third monomer unit, and so on.

OWL Sign in to OWL for Organic Chemistry at **www.cengage.com/owl** to view tutorials and simulations, develop problem-solving skills, and complete online homework assigned by your professor.

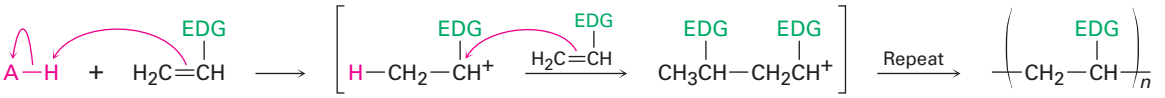
The initiator can be a radical, an acid, or a base. Historically, as we saw in **Section 8.10**, radical polymerization was the most common method because it can be carried out with practically any vinyl monomer.



Benzoyl peroxide | **Benzoyloxy radical** | **Phenyl radical (Ph·)**

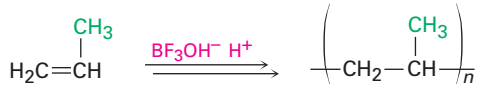


Acid-catalyzed (cationic) polymerization, by contrast, is effective only with vinyl monomers that contain an electron-donating group (EDG) capable of stabilizing the chain-carrying carbocation intermediate.



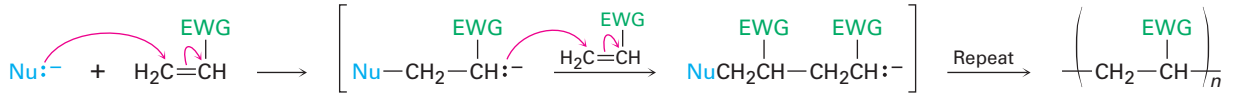
where EDG = an electron-donating group

Isobutylene (2-methylpropene) is a good example of a monomer that polymerizes rapidly under cationic conditions. The reaction is carried out commercially at −80 °C, using BF_3 and a small amount of water to generate BF_3OH^- H^+ catalyst. The product is used in the manufacture of truck and bicycle inner tubes.



Isobutylene | **Polyisobutylene**

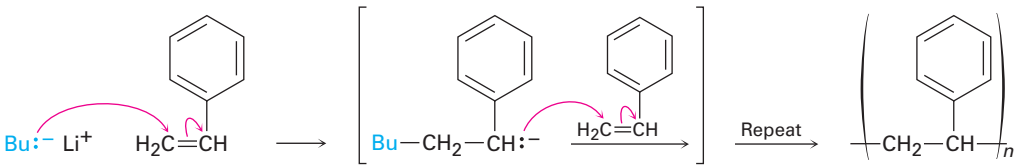
Vinyl monomers with electron-withdrawing groups (EWG) can be polymerized by basic (anionic) catalysts. The chain-carrying step is conjugate nucleophilic addition of an anion to the unsaturated monomer **(Section 19.13)**.



where EWG = an electron-withdrawing group

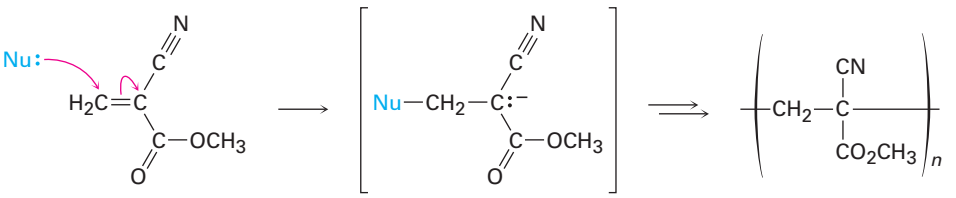
Acrylonitrile ($H_2C{=}CHCN$), methyl methacrylate [$H_2C{=}C(CH_3)CO_2CH_3$], and styrene ($H_2C{=}CHC_6H_5$) can all be polymerized anionically. The polystyrene

used in foam coffee cups, for example, is prepared by anionic polymerization of styrene using butyllithium as catalyst.

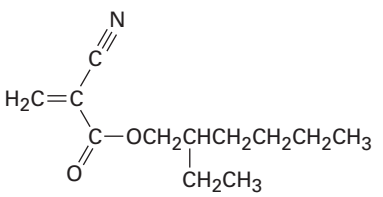


Styrene **Polystyrene**

An interesting example of anionic polymerization accounts for the remarkable properties of "super glue," one drop of which can support up to 2000 lb. Super glue is simply a solution of pure methyl α-cyanoacrylate, which has two electron-withdrawing groups that make anionic addition particularly easy. Trace amounts of water or bases on the surface of an object are sufficient to initiate polymerization of the cyanoacrylate and bind articles together. Skin is a good source of the necessary basic initiators, and many people have found their fingers stuck together after inadvertently touching super glue. So good is super glue at binding tissues together that related cyanoacrylate esters such as Dermabond are often used in place of sutures to close wounds.



Methyl α-cyanoacrylate



Dermabond
(2-ethylhexyl α-cyanoacrylate)

Problem 31.1

Order the following monomers with respect to their expected reactivity toward cationic polymerization, and explain your answer:

$$H_2C{=}CHCH_3,\ H_2C{=}CHCl,\ H_2C{=}CH{-}C_6H_5,\ H_2C{=}CHCO_2CH_3$$

Problem 31.2

Order the following monomers with respect to their expected reactivity toward anionic polymerization, and explain your answer:

$$H_2C{=}CHCH_3,\ H_2C{=}CHC{\equiv}N,\ H_2C{=}CHC_6H_5$$

Problem 31.3

Polystyrene is produced commercially by reaction of styrene with butyllithium as an anionic initiator. Using resonance structures, explain how the chain-carrying intermediate is stabilized.

31.2 Stereochemistry of Polymerization: Ziegler–Natta Catalysts

Although we didn't point it out when discussing chain-growth polymers in **Section 8.10**, the polymerization of a substituted vinyl monomer can lead to a polymer with numerous chirality centers in its chain. Propylene, for example, might polymerize with any of the three stereochemical outcomes shown in **Figure 31.1**. The polymer having all methyl groups on the same side of the zigzag backbone is called **isotactic**, the one in which the methyl groups alternate regularly on opposite sides of the backbone is called **syndiotactic**, and the one having the methyl groups randomly oriented is called **atactic**.

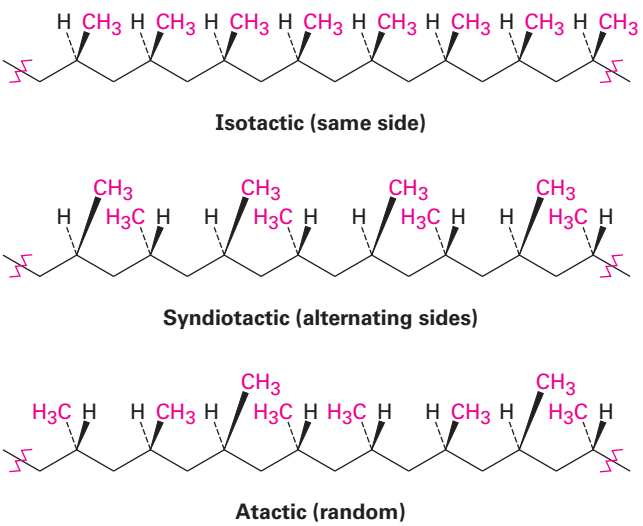


Figure 31.1 Isotactic, syndiotactic, and atactic forms of polypropylene.

The three different stereochemical forms of polypropylene all have somewhat different properties, and all can be made by using the right polymerization catalyst. Propylene polymerization using radical initiators does not work well, but polymerization using *Ziegler–Natta catalysts* allows preparation of isotactic, syndiotactic, and atactic polypropylene.

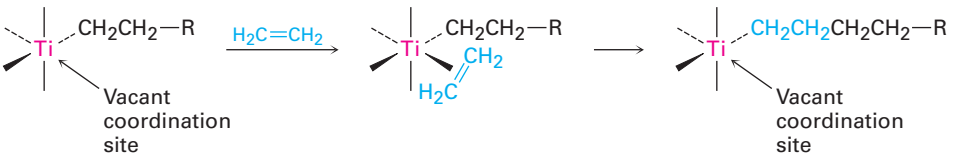
Ziegler–Natta catalysts—there are many different formulations—are organometallic transition-metal complexes prepared by treatment of an alkylaluminum with a titanium compound. Triethylaluminum and titanium tetrachloride form a typical preparation.

$$(CH_3CH_2)_3Al \ + \ TiCl_4 \rightarrow \text{A Ziegler–Natta catalyst}$$

Following their introduction in 1953, Ziegler–Natta catalysts revolutionized the field of polymer chemistry because of two advantages: first, the resultant

polymers are linear, with practically no chain branching, and second, they are stereochemically controllable. Isotactic, syndiotactic, and atactic forms can all be produced, depending on the catalyst system used.

The active form of a Ziegler–Natta catalyst is an alkyltitanium intermediate with a vacant coordination site on the metal. Coordination of alkene monomer to the titanium occurs, and the coordinated alkene then inserts into the carbon–titanium bond to extend the alkyl chain. A new coordination site opens up during the insertion step, so the process repeats indefinitely.



The linear polyethylene produced by the Ziegler–Natta process, called *high-density polyethylene,* is a highly crystalline polymer with 4000 to 7000 ethylene units per chain and molecular weights in the range 100,000 to 200,000 amu. High-density polyethylene has greater strength and heat resistance than the branched product of radical-induced polymerization, called *low-density polyethylene,* and is used to produce plastic squeeze bottles and molded housewares.

Polyethylenes of even higher molecular weights are produced for specialty applications. So-called high-molecular-weight (HMW) polyethylene contains 10,000 to 18,000 monomer units per chain (MW = 300,000–500,000 amu) and is used for underground pipes and large containers. Ultrahigh-molecular-weight (UHMW) polyethylene contains more than 100,000 monomer units per chain and has molecular weights ranging from 3,000,000 to 6,000,000 amu. It is used in bearings, conveyor belts, and bulletproof vests, among other applications requiring unusual wear resistance.

Problem 31.4

Vinylidene chloride, $H_2C{=}CCl_2$, does not polymerize in isotactic, syndiotactic, and atactic forms. Explain.

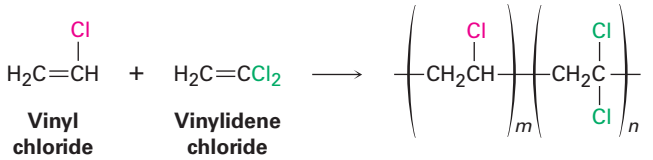
Problem 31.5

Polymers such as polypropylene contain a large number of chirality centers. Would you therefore expect samples of isotactic, syndiotactic, or atactic polypropylene to rotate plane-polarized light? Explain.

31.3 Copolymers

Up to this point we've discussed only **homopolymers**—polymers that are made up of identical repeating units. In practice, however, *copolymers* are more important commercially. **Copolymers** are obtained when two or more different monomers are allowed to polymerize together. For example, copolymerization

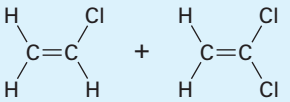
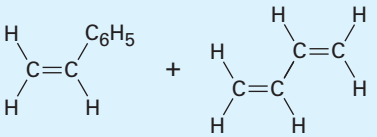
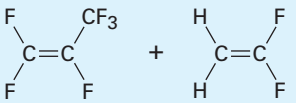
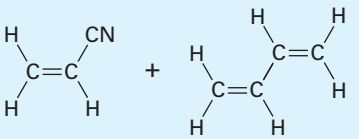
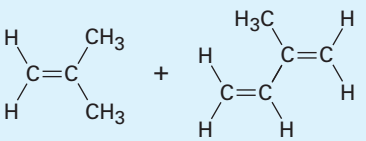
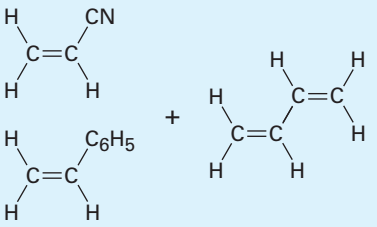
of vinyl chloride with vinylidene chloride (1,1-dichloroethylene) in a 1:4 ratio leads to the polymer Saran.



Vinyl chloride **Vinylidene chloride** **Saran**

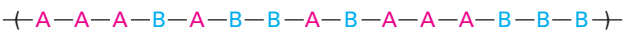
Copolymerization of monomer mixtures often leads to materials with properties quite different from those of either corresponding homopolymer, giving the polymer chemist a vast amount of flexibility for devising new materials. Table 31.1 lists some common copolymers and their commercial applications.

Table 31.1 Some Common Copolymers and Their Uses

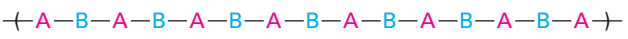
Monomers	Structures	Trade name	Uses
Vinyl chloride Vinylidene chloride		Saran	Fibers, food packaging
Styrene 1,3-Butadiene		SBR (styrene–butadiene rubber)	Tires, rubber articles
Hexafluoropropene Vinylidene fluoride		Viton	Gaskets, seals
Acrylonitrile 1,3-Butadiene		Nitrile rubber	Adhesives, hoses
Isobutylene Isoprene		Butyl rubber	Inner tubes
Acrylonitrile 1,3-Butadiene Styrene		ABS (monomer initials)	Pipes, high-impact applications

Several different types of copolymers can be defined, depending on the distribution of monomer units in the chain. If monomer A is copolymerized with

monomer B, for instance, the resultant product might have a random distribution of the two units throughout the chain, or it might have an alternating distribution.



Random copolymer



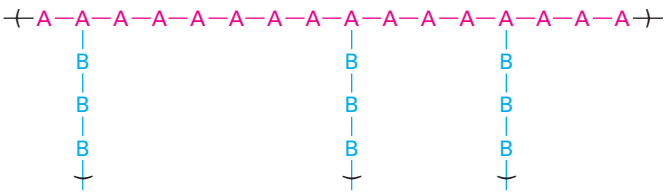
Alternating copolymer

The exact distribution of monomer units depends on the initial proportions of the two reactant monomers and their relative reactivities. In practice, neither perfectly random nor perfectly alternating copolymers are usually found. Most copolymers have many random imperfections.

Two other forms of copolymers that can be prepared under certain conditions are called *block copolymers* and *graft copolymers*. **Block copolymers** are those in which different blocks of identical monomer units alternate with each other; **graft copolymers** are those in which homopolymer branches of one monomer unit are "grafted" onto a homopolymer chain of another monomer unit.

A block copolymer

A graft copolymer



Block copolymers are prepared by initiating the polymerization of one monomer as if growing a homopolymer chain and then adding an excess of the second monomer to the still-active reaction mix. Graft copolymers are made by gamma irradiation of a completed homopolymer chain in the presence of the second monomer. The high-energy irradiation knocks hydrogen atoms off the homopolymer chain at random points, thus generating new radical sites that can initiate polymerization of the added monomer.

Problem 31.6

Draw the structure of an alternating segment of butyl rubber, a copolymer of isoprene (2-methyl-1,3-butadiene) and isobutylene (2-methylpropene) prepared using a cationic initiator.

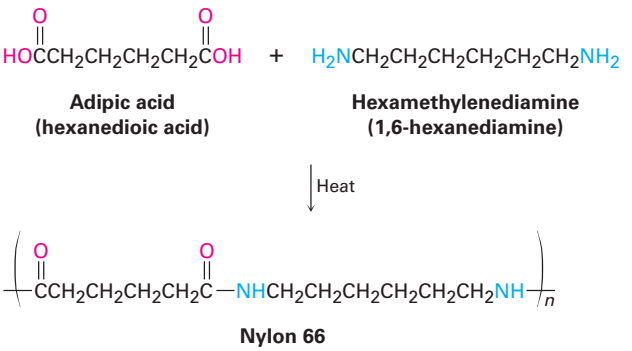
Problem 31.7

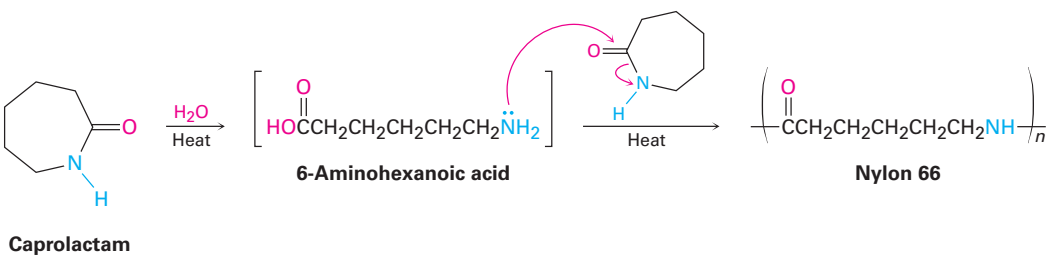
Irradiation of poly(1,3-butadiene), followed by addition of styrene, yields a graft copolymer that is used to make rubber soles for shoes. Draw the structure of a representative segment of this styrene–butadiene graft copolymer.

31.4 Step-Growth Polymers

Step-growth polymers are produced by reactions in which each bond in the polymer is formed stepwise, independently of the others. Like the polyamides (nylons) and polyesters that we saw in **Section 21.9**, most step-growth polymers

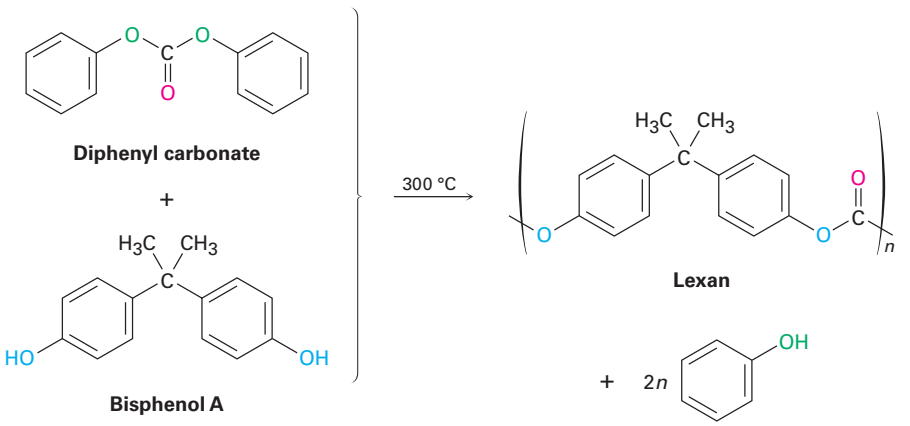
are produced by reaction between two difunctional reactants. Nylon 66, for instance, is made by reaction between the six-carbon adipic acid and the six-carbon hexamethylenediamine (1,6-hexanediamine). Alternatively, a single reactant with two different functional groups can polymerize. Nylon 6 is made by polymerization of the six-carbon caprolactam. The reaction is initiated by addition of a small amount of water, which hydrolyzes some caprolactam to 6-aminohexanoic acid. Nucleophilic addition of the amino group to caprolactam then propagates the polymerization.





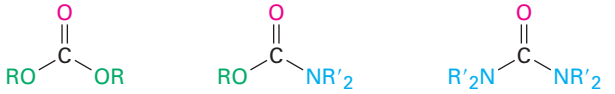
Polycarbonates

Polycarbonates are like polyesters, but their carbonyl group is linked to two −OR groups, [O═C(OR)$_2$]. Lexan, for instance, is a polycarbonate prepared from diphenyl carbonate and a diphenol called bisphenol A. Lexan has unusually high impact strength, making it valuable for use in machinery housings, telephones, bicycle safety helmets, and bulletproof glass.



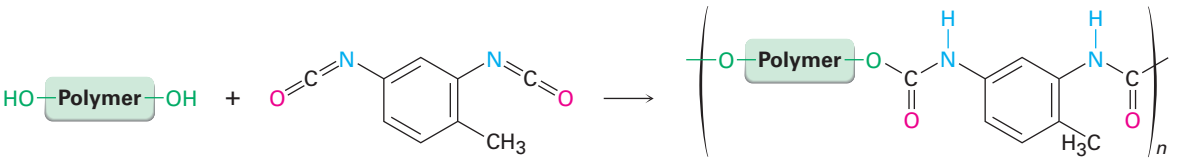
Polyurethanes

A *urethane* is a carbonyl-containing functional group in which the carbonyl carbon is bonded to both an −OR group and an $-NR_2$ group. As such, a urethane is halfway between a carbonate and a urea.



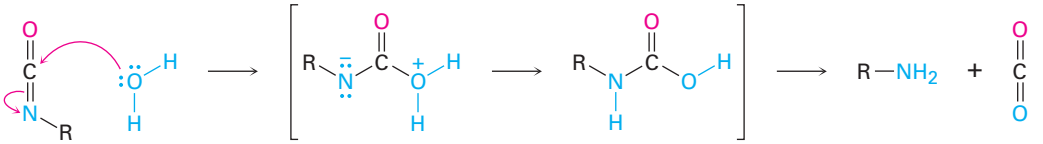
A carbonate **A urethane** **A urea**

A urethane is typically prepared by nucleophilic addition reaction between an alcohol and an isocyanate (R—N=C=O), so a **polyurethane** is prepared by reaction between a diol and a diisocyanate. The diol is usually a low-molecular-weight polymer (MW ≈ 1000 amu) with hydroxyl end-groups; the diisocyanate is often toluene-2,4-diisocyanate.



Toluene-2,4-diisocyanate **A polyurethane**

Several different kinds of polyurethanes are produced, depending on the nature of the polymeric alcohol used. One major use of polyurethane is in the stretchable spandex fibers used for bathing suits and athletic gear. These polyurethanes have a fairly low degree of cross-linking so that the resultant polymer is soft and elastic. A second major use of polyurethanes is in the foams used for insulation. Foaming occurs when a small amount of water is added during polymerization, giving a carbamic acid intermediate that spontaneously loses bubbles of CO_2.



A carbamic acid

Polyurethane foams are generally made using a *poly*alcohol rather than a diol as the monomer so that the polymer has a high amount of three-dimensional cross-linking. The result is a rigid but very light foam suitable for use as thermal insulation in building construction and portable ice chests.

Problem 31.8

Poly(ethylene terephthalate), or PET, is a polyester used to make soft-drink bottles. It is prepared by reaction of ethylene glycol with 1,4-benzenedicarboxylic acid (terephthalic acid). Draw the structure of PET.

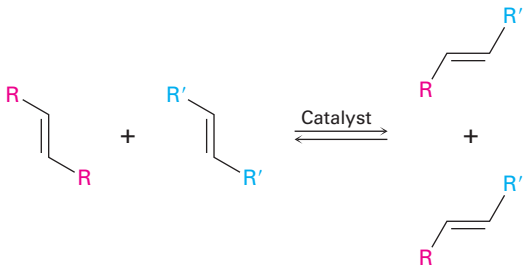
Problem 31.9

Show the mechanism of the nucleophilic addition reaction of an alcohol with an isocyanate to yield a urethane.

31.5 Olefin Metathesis Polymerization

Perhaps the most important advance in polymer synthesis in recent years has been the development of *olefin metathesis polymerization*. At its simplest, an olefin metathesis reaction is one in which two olefins (alkenes) exchange substituents on their double bonds.

An olefin metathesis reaction



Olefin metathesis catalysts, such as the Grubbs catalyst now in common use, contain a carbon–metal double bond (usually to ruthenium, Ru) and have the general structure M=CHR. They function by reacting reversibly with an alkene to form a four-membered, metal-containing intermediate called a *metallacycle,* which immediately opens to give a different catalyst and a different alkene. The mechanism is shown in **Figure 31.2**.

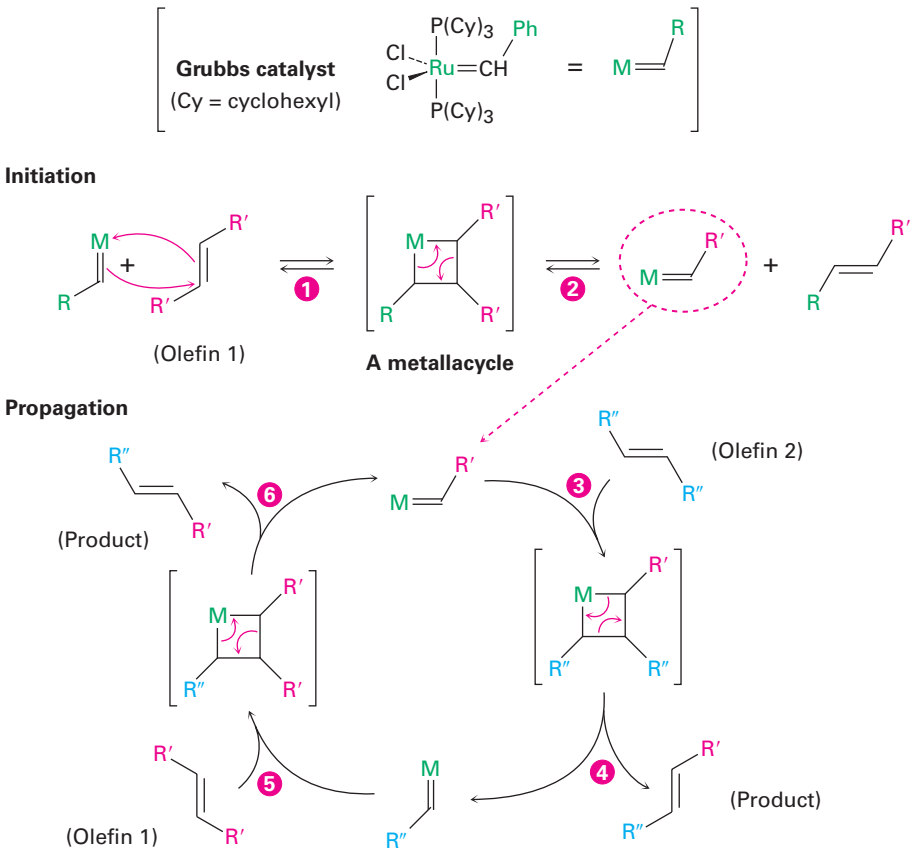
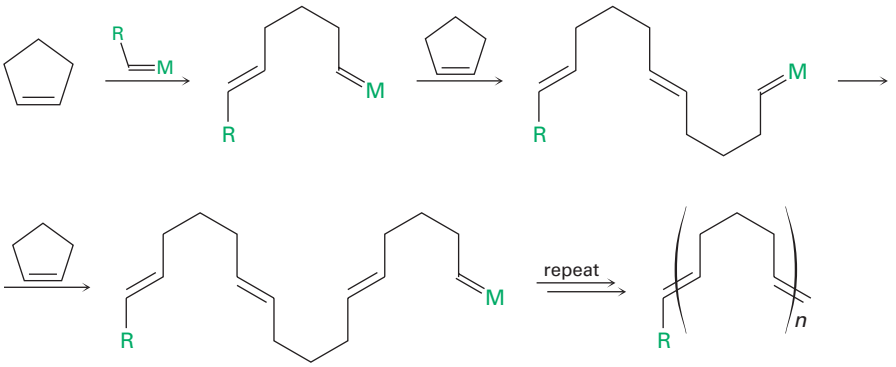


Figure 31.2 Mechanism of the olefin metathesis reaction. The process is initiated by a two-step sequence that involves (1) reaction of the catalyst and olefin 1 to give a four-membered metallacycle intermediate, followed by (2) ring-opening to give a different form of catalyst that contains part of olefin 1. (3) Reaction of this new catalyst with olefin 2 gives another metallacycle intermediate, (4) which opens to give metathesis product and another form of catalyst. (5, 6) The repeating ring-forming and ring-opening steps then continue.

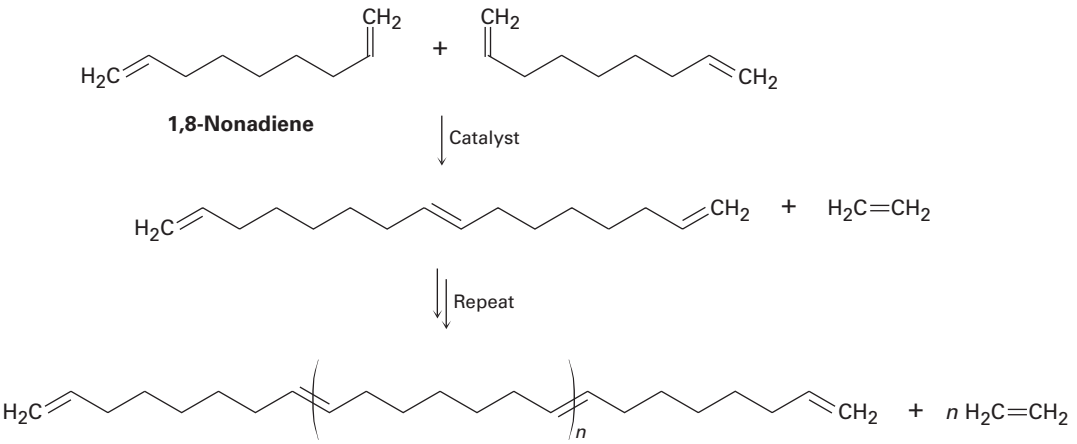
There are several methods for implementing the olefin metathesis reaction to prepare polymers. One method, called *ring-opening metathesis polymerization,* or ROMP, involves use of a moderately strained cycloalkene, such as cyclopentene. The strain of the ring favors ring-opening, thereby driving formation of the open-chain product. The polymer that results has double bonds spaced regularly along the chain, allowing for either hydrogenation or further functionalization if desired.

Ring-opening metathesis polymerization (ROMP)



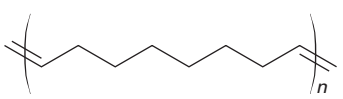
A second method of using olefin metathesis to prepare polymers is by *acyclic diene metathesis,* or ADMET. As the name suggests, ADMET involves olefin metathesis of an open-chain substrate with two double bonds at the ends of a long chain, such as 1,8-nonadiene. As the reaction proceeds, the gaseous ethylene by-product escapes, thereby driving the equilibrium toward polymer product. So efficient is the reaction that polymers with molecular weights as high as 80,000 amu have been prepared.

Acyclic diene metathesis (ADMET)

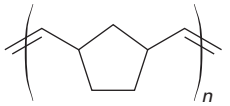


The ROMP and ADMET procedures are particularly valuable because the metathesis reaction is compatible with the presence in the olefin monomer of many different functional groups. In addition, the double bonds in the polymers allow still more flexibility for further manipulations. Among the commercial polymers produced by olefin metathesis are Vestenamer, used in the

manufacture of tires and other molded rubber objects, and Norsorex, used in the automobile industry as a sealing material.



Vestenamer



Norsorex

Problem 31.10

Look at the structures of Vestenamer and Norsorex, and show how they might be made by olefin metathesis polymerization.

31.6 Polymer Structure and Physical Properties

Polymers aren't really that different from other organic molecules. They're much larger, of course, but their chemistry is similar to that of analogous small molecules. Thus, the alkane chains of polyethylene undergo radical-initiated halogenation, the aromatic rings of polystyrene undergo typical electrophilic aromatic substitution reactions, and the amide linkages of nylon are hydrolyzed by aqueous base.

The major difference between small and large organic molecules is in their physical properties. For instance, their large size means that polymers experience substantially larger van der Waals forces than do small molecules **(Section 2.12)**. But because van der Waals forces operate only at close distances, they are strongest in polymers like high-density polyethylene, in which chains can pack together closely in a regular way. Many polymers, in fact, have regions that are essentially crystalline. These regions, called **crystallites**, consist of highly ordered portions in which the zigzag polymer chains are held together by van der Waals forces **(Figure 31.3)**.

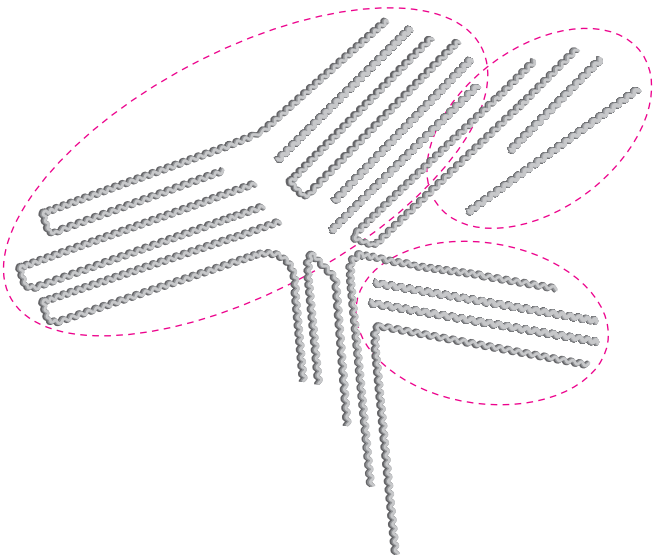
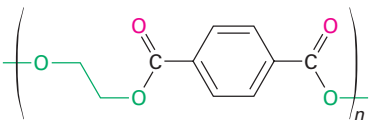


Figure 31.3 Crystallites in linear polyethylene. The long polymer chains are arranged in parallel lines in the crystallite regions.

As you might expect, polymer crystallinity is strongly affected by the steric requirements of substituent groups on the chains. Linear polyethylene is highly crystalline, but poly(methyl methacrylate) is noncrystalline because the chains can't pack closely together in a regular way. Polymers with a high degree of crystallinity are generally hard and durable. When heated, the crystalline regions melt at the **melt transition temperature**, T_m, to give an amorphous material.

Noncrystalline, amorphous polymers like poly(methyl methacrylate), sold under the trade name Plexiglas, have little or no long-range ordering among chains but can nevertheless be very hard at room temperature. When heated, the hard amorphous polymer becomes soft and flexible at a point called the **glass transition temperature**, T_g. Much of the art in polymer synthesis lies in finding methods for controlling the degree of crystallinity and the glass transition temperature, thereby imparting useful properties to the polymer.

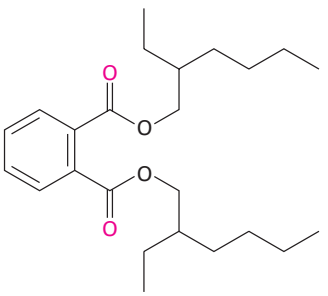
In general, polymers can be divided into four major categories, depending on their physical behavior: *thermoplastics, fibers, elastomers,* and *thermosetting resins*. **Thermoplastics** are the polymers most people think of when the word *plastic* is mentioned. These polymers have a high T_g and are therefore hard at room temperature but become soft and viscous when heated. As a result, they can be molded into toys, beads, telephone housings, or any of a thousand other items. Because thermoplastics have little or no cross-linking, the individual chains can slip past one another in the melt. Some thermoplastic polymers, such as poly(methyl methacrylate) and polystyrene, are amorphous and noncrystalline; others, such as polyethylene and nylon, are partially crystalline. Among the better-known thermoplastics is poly(ethylene terephthalate), or PET, used for making plastic soft-drink bottles.



Poly(ethylene terephthalate)

Plasticizers—small organic molecules that act as lubricants between chains—are usually added to thermoplastics to keep them from becoming brittle at room temperature. An example is poly(vinyl chloride), which is brittle when pure but becomes supple and pliable when a plasticizer is added. In fact, most drip bags used in hospitals to deliver intravenous saline solutions are made of poly(vinyl chloride), although replacements are appearing.

Dialkyl phthalates such as di(2-ethylhexyl) phthalate (generally called dioctyl phthalate) are commonly used as plasticizers although questions about their safety have been raised. The U.S. Food and Drug Administration (FDA) has advised the use of alternative materials in compromised patients and infants but has found no evidence of toxicity for healthy individuals. In addition, children's toys that contain phthalates have been banned in the United States.



Di(2-ethylhexyl) phthalate (or dioctyl phthalate), a plasticizer

Fibers are thin threads produced by extruding a molten polymer through small holes in a die, or spinneret. The fibers are then cooled and drawn out, which orients the crystallite regions along the axis of the fiber and adds considerable tensile strength **(Figure 31.4)**. Nylon, Dacron, and polyethylene all have the semicrystalline structure necessary for drawing into oriented fibers.

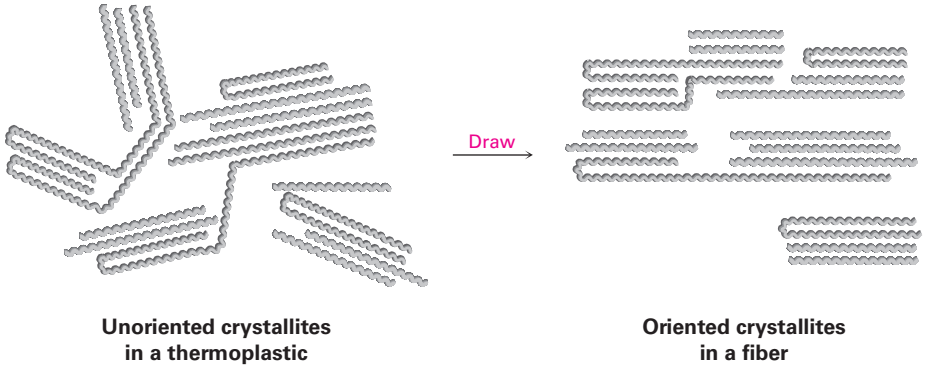


Figure 31.4 Oriented crystallite regions in a polymer fiber.

Elastomers are amorphous polymers that have the ability to stretch out and spring back to their original shapes. These polymers must have low T_g values and a small amount of cross-linking to prevent the chains from slipping over one another. In addition, the chains must have an irregular shape to prevent crystallite formation. When stretched, the randomly coiled chains straighten out and orient along the direction of the pull. Van der Waals forces are too weak and too few to maintain this orientation, however, and the elastomer therefore reverts to its random coiled state when the stretching force is released **(Figure 31.5)**.

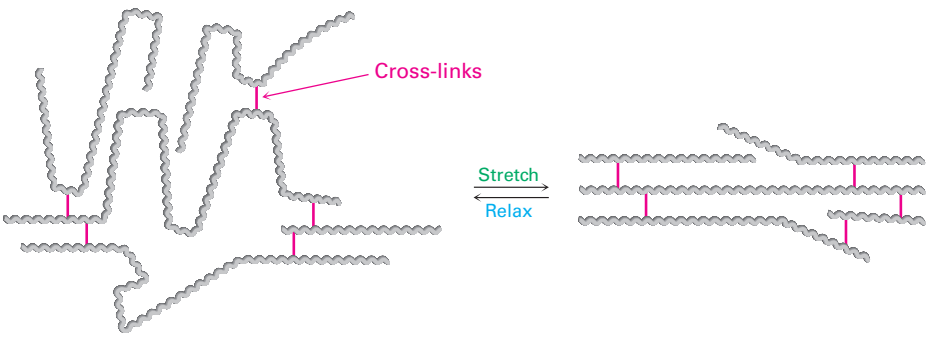


Figure 31.5 Unstretched and stretched forms of an elastomer.

Natural rubber **(Section 14.6)** is the most common example of an elastomer. Rubber has the long chains and occasional cross-links needed for elasticity, but its irregular geometry prevents close packing of the chains into crystallites. Gutta-percha, by contrast, is highly crystalline and is not an elastomer **(Figure 31.6)**.

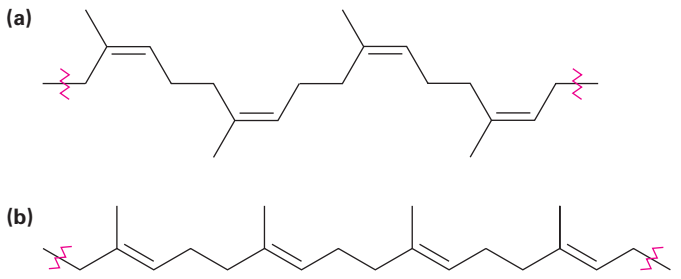
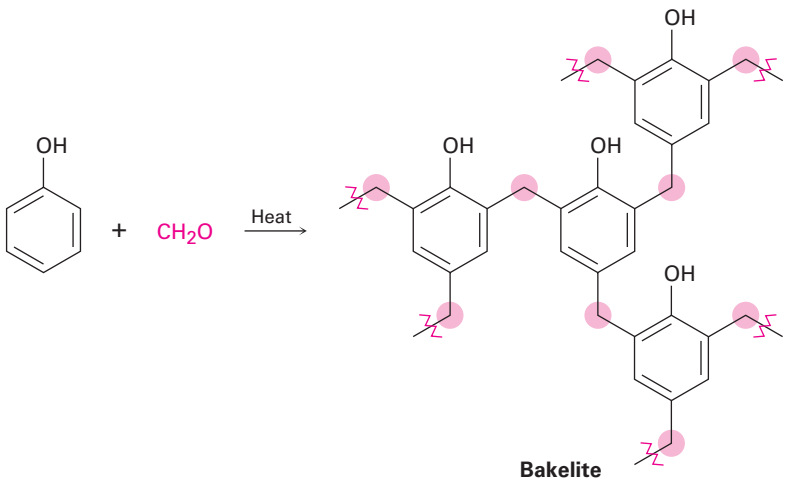


Figure 31.6 **(a)** Natural rubber is elastic and noncrystalline because of its cis double-bond geometry, but **(b)** gutta-percha is nonelastic and crystalline because its geometry allows for better packing together of chains.

Thermosetting resins are polymers that become highly cross-linked and solidify into a hard, insoluble mass when heated. *Bakelite,* a thermosetting resin first produced in 1907, has been in commercial use longer than any other synthetic polymer. It is widely used for molded parts, adhesives, coatings, and even high-temperature applications such as missile nose cones.

Chemically, Bakelite is a *phenolic resin,* produced by reaction of phenol and formaldehyde. On heating, water is eliminated, many cross-links form, and the polymer sets into a rocklike mass. The cross-linking in Bakelite and other thermosetting resins is three-dimensional and is so extensive that we can't really speak of polymer "chains." A piece of Bakelite is essentially one large molecule.



Bakelite

Problem 31.11

What product would you expect to obtain from catalytic hydrogenation of natural rubber? Would the product be syndiotactic, atactic, or isotactic?

Problem 31.12

Propose a mechanism to account for the formation of Bakelite from acid-catalyzed polymerization of phenol and formaldehyde.

Biodegradable Polymers A DEEPER LOOK

The high chemical stability of many polymers is both a blessing and a curse. Heat resistance, wear resistance, and long life are valuable characteristics of clothing fibers, bicycle helmets, underground pipes, food wrappers, and many other items. Yet when those items outlive their usefulness, disposal becomes a problem.

Recycling of unwanted polymers is the best solution, and six types of plastics in common use are frequently stamped with identifying codes assigned by the Society of the Plastics Industry (Table 31.2). After being sorted by type, the items to be recycled are shredded into

(continued)

(continued)

small chips, washed, dried, and melted for reuse. Soft-drink bottles, for instance, are made from recycled poly(ethylene terephthalate), trash

iStockphoto.com/Joy Fera

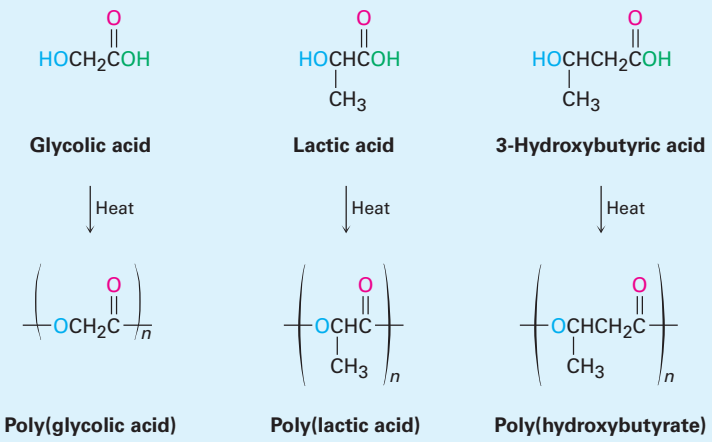
What happens to the plastics that end up here?

Table 31.2 Recyclable Plastics

Polymer	Recycling code	Use
Poly(ethylene terephthalate)	1—PET	Soft-drink bottles
High-density polyethylene	2—HDPE	Bottles
Poly(vinyl chloride)	3—V	Floor mats
Low-density polyethylene	4—LDPE	Grocery bags
Polypropylene	5—PP	Furniture
Polystyrene	6—PS	Molded articles
Mixed plastics	7	Benches, plastic lumber

bags are made from recycled low-density polyethylene, and garden furniture is made from recycled polypropylene and mixed plastics.

Frequently, however, plastics are simply thrown away rather than recycled, and much work has therefore been carried out on developing *biodegradable* polymers, which can be broken down rapidly by soil microorganisms. Among the most common biodegradable polymers are polyglycolic acid (PGA), polylactic acid (PLA), and polyhydroxybutyrate (PHB). All are polyesters and are therefore susceptible to hydrolysis of their ester links. Copolymers of PGA with PLA have found a particularly wide range of uses. A 90/10 copolymer of polyglycolic acid with polylactic acid is used to make absorbable sutures that are degraded and absorbed by the body within 90 days after surgery.



Summary

Key words

atactic, 1245
block copolymer, 1248
copolymer, 1246
crystallite, 1253
elastomer, 1255
fiber, 1255
glass transition temperature (T_g), 1254
graft copolymer, 1248
homopolymer, 1246
isotactic, 1245
melt transition temperature (T_m), 1254
monomer, 1242
plasticizer, 1254
polycarbonate, 1249
polymer, 1242
polyurethane, 1250
syndiotactic, 1245
thermoplastic, 1254
thermosetting resin, 1256
Ziegler–Natta catalyst, 1245

Synthetic polymers can be classified as either chain-growth or step-growth. Chain-growth polymers are prepared by chain-reaction polymerization of vinyl monomers in the presence of a radical, an anion, or a cation initiator. Radical polymerization is sometimes used, but alkenes such as 2-methylpropene that have electron-donating substituents on the double bond polymerize easily by a cationic route through carbocation intermediates. Similarly, monomers such as methyl α-cyanoacrylate that have electron-withdrawing substituents on the double bond polymerize by an anionic, conjugate addition pathway.

Copolymerization of two monomers gives a product with properties different from those of either homopolymer. **Graft copolymers** and **block copolymers** are two examples.

Alkene polymerization can be carried out in a controlled manner using a **Ziegler–Natta catalyst**. Ziegler–Natta polymerization minimizes the amount of chain branching in the polymer and leads to stereoregular chains—either **isotactic** (substituents on the same side of the chain) or **syndiotactic** (substituents on alternate sides of the chain), rather than **atactic** (substituents randomly disposed).

Step-growth polymers, the second major class of polymers, are prepared by reactions between difunctional molecules, with the individual bonds in the polymer formed independently of one another. **Polycarbonates** are formed from a diester and a diol, and **polyurethanes** are formed from a diisocyanate and a diol.

The chemistry of synthetic polymers is similar to the chemistry of small molecules with the same functional groups, but the physical properties of polymers are greatly affected by size. Polymers can be classified by physical property into four groups: **thermoplastics**, **fibers**, **elastomers**, and **thermosetting resins**. The properties of each group can be accounted for by the structure, the degree of crystallinity, and the amount of cross-linking they contain.

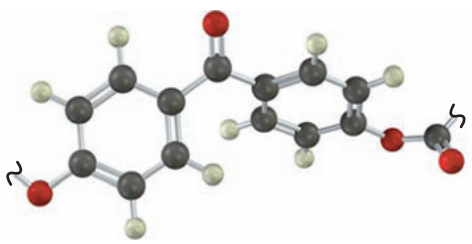
Exercises

OWL Interactive versions of these problems are assignable in OWL for Organic Chemistry.

Visualizing Chemistry

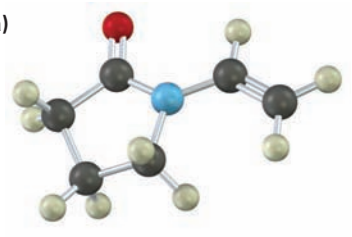
(Problems 31.1–31.12 appear within the chapter.)

31.13 Identify the structural class to which the following polymer belongs, and show the structure of the monomer units used to make it:

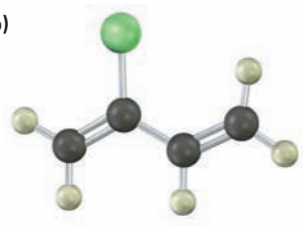


31.14 Show the structures of the polymers that could be made from the following monomers (green = Cl):

(a)



(b)



Additional Problems

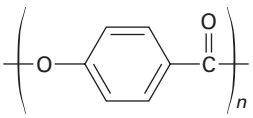
31.15 Identify the monomer units from which each of the following polymers is made, and tell whether each is a chain-growth or a step-growth polymer:

(a) $-(-CH_2-O-)_n-$

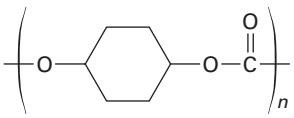
(b) $-(-CF_2-CFCl-)_n-$

(c)

(d)



(e)



31.16 Draw a three-dimensional representation of segments of the following polymers:

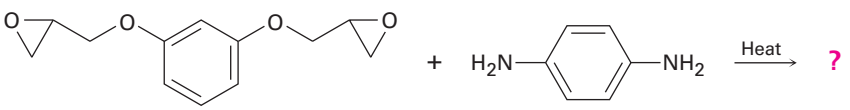
(a) Syndiotactic polyacrylonitrile **(b)** Atactic poly(methyl methacrylate)

(c) Isotactic poly(vinyl chloride)

31.17 Draw the structure of Kodel, a polyester prepared by heating dimethyl 1,4-benzenedicarboxylate with 1,4-bis(hydroxymethyl)cyclohexane.

1,4-Bis(hydroxymethyl)cyclohexane

31.18 Show the structure of the polymer that results from heating the following diepoxide and diamine:



31.19 Nomex, a polyamide used in such applications as fire-retardant clothing, is prepared by reaction of 1,3-benzenediamine with 1,3-benzenedicarbonyl chloride. Show the structure of Nomex.

31.20 Nylon 10,10 is an extremely tough, strong polymer used to make reinforcing rods for concrete. Draw a segment of nylon 10,10, and show its monomer units.

31.21 1,3-Cyclopentadiene undergoes thermal polymerization to yield a polymer that has no double bonds in the chain. On strong heating, the polymer breaks down to regenerate cyclopentadiene. Propose a structure for the polymer.

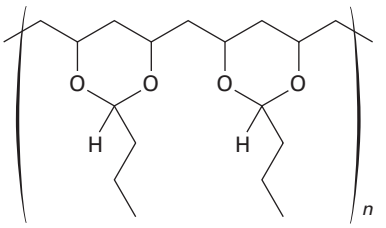
31.22 When styrene, $C_6H_5CH{=}CH_2$, is copolymerized in the presence of a few percent *p*-divinylbenzene, a hard, insoluble, cross-linked polymer is obtained. Show how this cross-linking of polystyrene chains occurs.

31.23 Poly(ethylene glycol), or Carbowax, is made by anionic polymerization of ethylene oxide using NaOH as catalyst. Propose a mechanism.

$$-\!\left(\!-O-CH_2CH_2-\right)_{\overline{n}}$$ **Poly(ethylene glycol)**

31.24 Nitroethylene, $H_2C{=}CHNO_2$, is a sensitive compound that must be prepared with great care. Attempted purification of nitroethylene by distillation often results in low recovery of product and a white coating on the inner walls of the distillation apparatus. Explain.

31.25 Poly(vinyl butyral) is used as the plastic laminate in the preparation of automobile windshield safety glass. How would you synthesize this polymer?



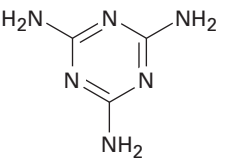
Poly(vinyl butyral)

31.26 What is the structure of the polymer produced by anionic polymerization of *β*-propiolactone using NaOH as catalyst?

***β*-Propiolactone**

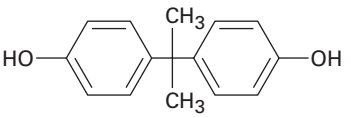
31.27 Glyptal is a highly cross-linked thermosetting resin produced by heating glycerol and phthalic anhydride (1,2-benzenedicarboxylic acid anhydride). Show the structure of a representative segment of glyptal.

31.28 Melmac, a thermosetting resin often used to make plastic dishes, is prepared by heating melamine with formaldehyde. Look at the structure of Bakelite shown in Section 31.6, and then propose a structure for Melmac.

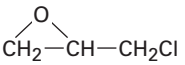


Melamine

31.29 Epoxy adhesives are cross-linked resins prepared in two steps. The first step involves S_N2 reaction of the disodium salt of bisphenol A with epichlorohydrin to form a low-molecular-weight prepolymer. This prepolymer is then "cured" into a cross-linked resin by treatment with a triamine such as $H_2NCH_2CH_2NHCH_2CH_2NH_2$.



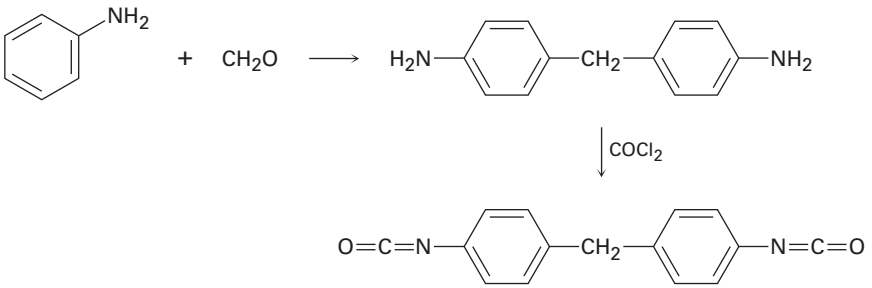
Bisphenol A



Epichlorohydrin

(a) What is the structure of the prepolymer?
(b) How does addition of the triamine to the prepolymer result in cross-linking?

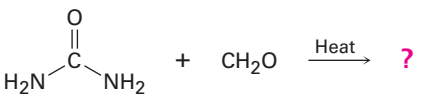
31.30 The polyurethane foam used for home insulation uses methanediphenyldiisocyanate (MDI) as monomer. The MDI is prepared by acid-catalyzed reaction of aniline with formaldehyde, followed by treatment with phosgene, $COCl_2$. Propose mechanisms for both steps.



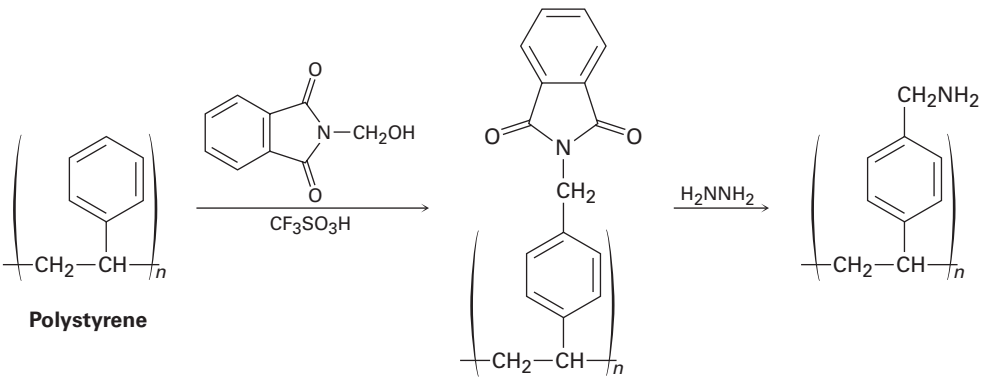
MDI

31.31 Write the structure of a representative segment of polyurethane prepared by reaction of ethylene glycol with MDI (Problem 31.30).

31.32 The smoking salons of the Hindenburg and other hydrogen-filled dirigibles of the 1930s were insulated with urea–formaldehyde polymer foams. The structure of this polymer is highly cross-linked, like that of Bakelite (Section 31.6). Propose a structure.



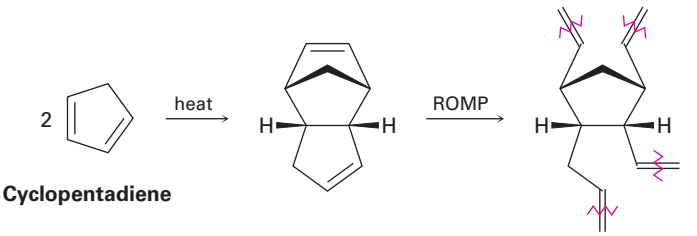
31.33 The polymeric resin used for Merrifield solid-phase peptide synthesis (Section 26.8) is prepared by treating polystyrene with *N*-(hydroxymethyl) phthalimide and trifluoromethanesulfonic acid, followed by reaction with hydrazine. Propose a mechanism for both steps.



Polystyrene

31.34 2-Ethyl-1-hexanol, used in the synthesis of di(2-ethylhexyl) phthalate plasticizer, is made commercially from butanal. Show the likely synthesis route.

31.35 Polydicyclopentadiene (PDCPD), marketed as Telene and Metton, is a highly cross-linked thermosetting resin used for molding such impact-resistant parts as cabs for large trucks and earth-moving equipment. PDCPD is prepared by ring-opening metathesis polymerization of dicyclopentadiene, which is itself prepared from 1,3-cyclopentadiene. The polymerization occurs by initial metathesis of the more highly strained double bond in the bicyclo[2.2.1] heptane part of the molecule (Section 4.9) to give a linear polymer, followed by cross-linking of different chains in a second metathesis of the remaining cyclopentene double bond.



Cyclopentadiene **Dicyclopentadiene** **Polydicyclopentadiene**

(a) Show the mechanism of the formation of dicyclopentadiene from cyclopentadiene.
(b) Draw the structure of a representative sample of the initially formed linear polymer containing three monomer units.
(c) Draw the structure of a representative sample of PDCPD that shows how cross-linking of the linear chains takes place.