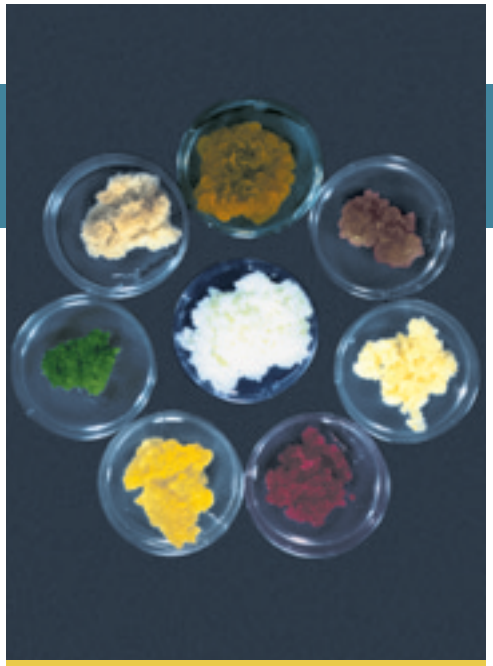
Biochemistry & Molecular Biology of Plants, B. Buchanan, W. Gruissem, R. Jones, Eds.



CHAPTER 24

Natural Products (Secondary Metabolites)

Rodney Croteau
Toni M. Kutchan
Norman G. Lewis

CHAPTER OUTLINE

Introduction

Natural products have primary ecological functions.

Plants produce a vast and diverse assortment of organic compounds, the great majority of which do not appear to participate directly in growth and development. These substances, traditionally referred to as **secondary metabolites,** often are differentially distributed among limited taxonomic groups within the plant kingdom. Their functions, many of which remain unknown, are being elucidated with increasing frequency. The **primary metabolites,** in contrast, such as phytosterols, acyl lipids, nucleotides, amino acids, and organic acids, are found in all plants and perform metabolic roles that are essential and usually evident.

Although noted for the complexity of their chemical structures and biosynthetic pathways, **natural products** have been widely perceived as biologically insignificant and have historically received little attention from most plant biologists. Organic chemists, however, have long been interested in these novel phytochemicals and have investigated their chemical properties extensively since the 1850s. Studies of natural products stimulated development of the separation techniques, spectroscopic approaches to structure elucidation, and synthetic methodologies that now constitute the foundation of contemporary organic chemistry. Interest in natural products was not purely academic but rather was prompted by their great utility as dyes, polymers, fibers, glues, oils, waxes, flavoring agents, perfumes, and drugs. Recognition of the biological properties of myriad natural products has fueled the current focus of this field, namely, the search for new drugs, antibiotics, insecticides, and herbicides. Importantly, this growing appreciation of the highly diverse biological effects produced by natural products has prompted a reevaluation of the possible roles these compounds play in plants, especially in the context of ecological interactions. As illustrated in this chapter, many of these compounds now have been shown to have important

adaptive significance in protection against herbivory and microbial infection, as attractants for pollinators and seed-dispersing animals, and as allelopathic agents (allelochemicals that influence competition among plant species). These ecological functions affect plant survival profoundly, and we think it reasonable to adopt the less pejorative term "plant natural products" to describe secondary plant metabolites that act primarily on other species.

The boundary between primary and secondary metabolism is blurred.

Based on their biosynthetic origins, plant natural products can be divided into three major groups: the terpenoids, the alkaloids, and the phenylpropanoids and allied phenolic compounds. All terpenoids, including both primary metabolites and more than 25,000 secondary compounds, are derived from the five-carbon precursor isopentenyl diphosphate (IPP). The 12,000 or so known alkaloids, which contain one or more nitrogen atoms, are biosynthesized principally from amino acids. The 8000 or so phenolic compounds are formed by way of either the shikimic acid pathway or the malonate/acetate pathway.

Primary and secondary metabolites cannot readily be distinguished on the basis of precursor molecules, chemical structures, or biosynthetic origins. For example, both primary and secondary metabolites are found among the diterpenes (C_{20}) and triterpenes (C_{30}). In the diterpene series, both kaurenoic acid and abietic acid are formed by a very similar sequence of related enzymatic reactions (Fig. 24.1); the former is an essential intermediate in the synthesis of gibberellins, i.e., growth hormones found in all plants (see Chapter 17), whereas the latter is a resin component largely restricted to members of the Fabaceae and Pinaceae. Similarly, the essential amino acid proline is classified as a primary metabolite, whereas the C_6 analog pipecolic acid is considered an alkaloid and thus a natural product (Fig. 24.1). Even lignin, the essential structural polymer of wood and second only to cellulose as the most abundant organic substance in plants, is considered a natural product rather than a primary metabolite.

In the absence of a valid distinction based on either structure or biochemistry, we return to a functional definition, with primary products participating in nutrition and essential metabolic processes inside the plant, and natural (secondary) products influencing ecological interactions between the plant and its environment. In this chapter, we provide an overview of the biosynthesis of the major classes of plant natural products, emphasizing the origins of their structural diversity, as well as their physiological functions, human uses, and potential biotechnological applications.

24.1 Terpenoids

Terpenoids perhaps are the most structurally varied class of plant natural products. The name terpenoid, or terpene, derives from the fact that the first members of the class were

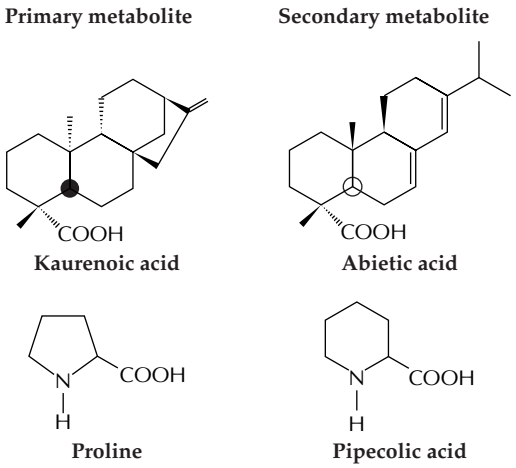


Figure 24.1
Kaurenoic acid and proline are primary metabolites, whereas the closely related compounds abietic acid and pipecolic acid are considered secondary metabolites.

isolated from turpentine ("terpentin" in German). All terpenoids are derived by repetitive fusion of branched five-carbon units based on isopentane skeleton. These monomers generally are referred to as isoprene units because thermal decomposition of many terpenoid substances yields the alkene gas isoprene as a product (Fig. 24.2, upper panel) and because suitable chemical conditions can induce isoprene to polymerize in multiples of five carbons, generating numerous terpenoid skeletons. For these reasons, the terpenoids are often called isoprenoids, although researchers have known for well over 100 years that isoprene itself is not the biological precursor of this family of metabolites.

24.1.1 Terpenoids are classified by the number of five-carbon units they contain.

The five-carbon (isoprene) units that make up the terpenoids are often joined in a "head-to-tail" fashion, but head-to-head fusions are also common, and some products are formed by head-to-middle fusions (Fig. 24.2, lower panel). Accordingly, and because extensive structural modifications with carbon–carbon bond rearrangements can occur, tracing the original pattern of isoprene units is sometimes difficult.

The smallest terpenes contain a single isoprene unit; as a group, they are named **hemiterpenes** (half-terpenes). The best known hemiterpene is isoprene itself, a volatile product released from photosynthetically active tissues. The enzyme isoprene synthase is present in the leaf plastids of numerous C_3 plant species, but the metabolic rationale for the light-dependent production of isoprene is unknown (acclimation to high temperatures has been suggested). Estimated annual foliar emissions of isoprene are quite substantial (5×10^8 metric tons of carbon), and the gas is a principal reactant in the NOx radical–induced formation of tropospheric ozone (see Chapter 22, Fig. 22.37).

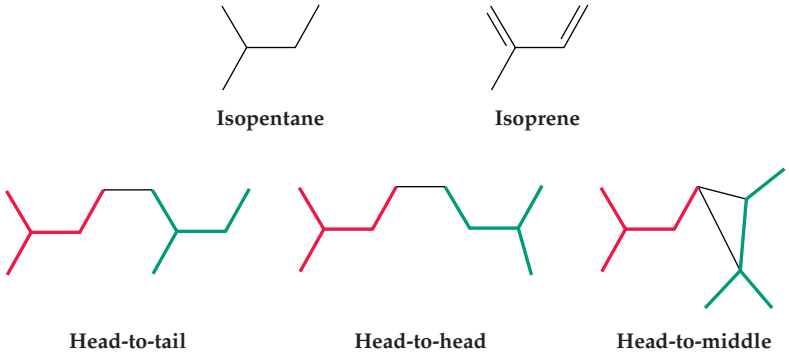


Figure 24.2
Terpenes are synthesized from C_5 units. The upper panel shows the structures of the isopentane skeleton and isoprene gas. The lower panel shows how different patterns of isoprene unit assembly yield a variety of different structures. For example, the triterpene squalene is formed by head-to-head fusion of two molecules of farnesyl diphosphate (FPP), which itself is the product of the head-to-tail fusion of isopentenyl diphosphate (IPP) and geranyl diphosphate (GPP) (see Fig. 24.7). The monoterpene pyrethrin I (see Fig. 24.10) results from a head-to-middle fusion of two C_5 units.

C_{10} terpenoids, although they consist of two isoprene units, are called **monoterpenes;** as the first terpenoids isolated from turpentine in the 1850s, they were considered to be the base unit from which the subsequent nomenclature is derived. The monoterpenes are best known as components of the volatile essences of flowers and of the essential oils of herbs and spices, in which they make up as much as 5% of plant dry weight. Monoterpenes are isolated by either distillation or extraction and find considerable industrial use in flavors and perfumes.

The terpenoids that derive from three isoprene units contain 15 carbon atoms and are known as **sesquiterpenes** (i.e., one and one-half terpenes). Like monoterpenes, many sesquiterpenes are found in essential oils. In addition, numerous sesquiterpenoids act as phytoalexins, antibiotic compounds produced by plants in response to microbial challenge, and as antifeedants that discourage opportunistic herbivory. Although the plant hormone abscisic acid is structurally a sesquiterpene, its C_{15} precursor, xanthoxin, is not synthesized directly from three isoprene units but rather is produced by asymmetric cleavage of a C_{40} carotenoid (see Chapter 17).

The **diterpenes,** which contain 20 carbons (four C_5 units), include phytol (the hydrophobic side chain of chlorophyll), the gibberellin hormones, the resin acids of conifer and legume species, phytoalexins, and a host of pharmacologically important metabolites, including taxol, an anticancer agent found at very low concentrations (0.01% dry weight) in yew bark, and forskolin, a compound used to treat glaucoma. Some gibberellins have only 19 carbon atoms and are considered norditerpenoids since they have lost 1 carbon through a metabolic cleavage reaction (see Chapter 17).

The **triterpenes,** which contain 30 carbon atoms, are generated by the head-to-head joining of two C_{15} chains, each of which constitutes three isoprene units joined head-to-tail. This large class of molecules includes the brassinosteroids (see Chapter 17), the phytosterol membrane components (see Chapter 1), certain phytoalexins, various toxins and feeding deterrents, and components of surface waxes, such as oleanolic acid of grapes.

The most prevalent **tetraterpenes** (40 carbons, eight isoprene units) are the carotenoid accessory pigments which perform essential functions in photosynthesis (see Chapter 12). The **polyterpenes,** those containing more than eight isoprene units, include the prenylated quinone electron carriers (plastoquinone and ubiquinone; see Chapters 12 and 14), long-chain polyprenols involved in sugar transfer reactions (e.g., dolichol; see Chapters 1 and 4), and enormously long polymers such as rubber (average molecular mass greater than 10^6 Da), often found in latex.

Natural products of mixed biosynthetic origins that are partially derived from terpenoids are often called **meroterpenes.** For example, both cytokinins (see Chapter 17) and numerous phenylpropanoid compounds contain C_5 isoprenoid side chains. Certain alkaloids, including the anticancer drugs vincristine and vinblastine, contain terpenoid fragments in their structures (see Fig. 24.34). Additionally, some modified proteins include a 15- or 20-carbon terpenoid side chain that anchors the protein in a membrane (see Chapter 1).

24.1.2 A diverse array of terpenoid compounds is synthesized by various conserved reaction mechanisms.

At the turn of the 20th century, structural investigations of many terpenoids led Otto Wallach to formulate the "isoprene rule," which postulated that most terpenoids could be constructed hypothetically by repetitively joining isoprene units. This principle provided the first conceptual framework for a common structural relationship among terpenoid natural products (Box 24.1). Wallach's idea was refined in the 1930s, when Leopold Ruzicka formulated the "biogenetic isoprene rule," emphasizing mechanistic considerations of terpenoid synthesis in terms of electrophilic elongations, cyclizations, and rearrangements. This hypothesis ignores the precise character of the biological precursors and assumes only that they are "isoprenoid" in structure. As a working model for terpenoid biosynthesis, the biogenetic isoprene rule has proved essentially correct.

Despite great diversity in form and function, the terpenoids are unified in their common biosynthetic origin. The biosynthesis of all terpenoids from simple, primary metabolites can be divided into four overall steps: *(a)* synthesis of the fundamental precursor IPP; *(b)* repetitive additions of IPP to form a series of prenyl diphosphate homologs, which serve as the immediate precursors of the different classes of terpenoids; *(c)* elaboration of these allylic prenyl diphosphates by specific terpenoid synthases to yield terpenoid skeletons; and *(d)* secondary enzymatic modifications to the skeletons (largely redox reactions) to give rise to the functional properties and great chemical diversity of this family of natural products.

24.2 Synthesis of IPP

24.2.1 Biosynthesis of terpenoids is compartmentalized, as is production of the terpenoid precursor IPP.

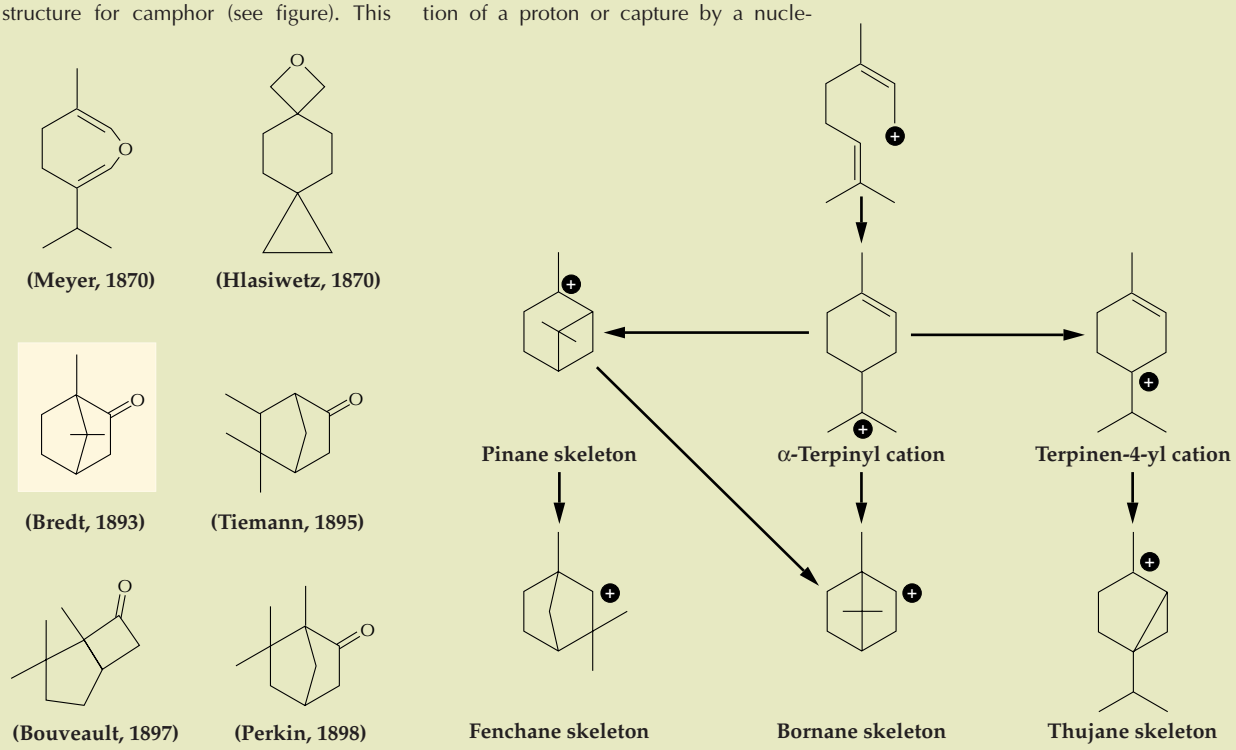
Although terpenoid biosynthesis in plants, animals, and microorganisms involves similar classes of enzymes, important differences exist among these processes. In particular, plants produce a much wider variety of terpenoids than do either animals or microbes, a difference reflected in the complex organization of plant terpenoid biosynthesis at the tissue, cellular, subcellular, and genetic levels. The production of large quantities of terpenoid natural products as well as their subsequent accumulation, emission, or secretion is almost always associated with the presence of anatomically highly specialized structures. The glandular trichomes (Fig. 24.3A, B) and secretory cavities of leaves (Fig. 24.3C) and the glandular epiderms of flower petals generate and store or emit terpenoid essential oils that are important because they encourage pollination by insects. The resin ducts and blisters of conifer species

Box 24.1

Early investigators formulated rules for identifying and naming isoprenoid structures.

In the late 1800s, chemists struggled to define the structures of the monoterpenes. The mixed results achieved by these efforts are illustrated by the numerous structures proposed for camphor ($C_{10}H_{16}O$; see structures at left of figure, which include the names of the proposers and the dates proposed). Chromatographic purification techniques and spectroscopic methods for structure elucidation were not available to these early chemists, who relied on the preparation of crystalline derivatives to assess purity and on chemical degradation studies to determine structures. Systematic study of the monoterpenes led the German chemist Otto Wallach to recognize that many terpenoid compounds might be constructed by joining isoprene units, generally in a repetitive head-to-tail fashion, as in Bredt's correct proposed structure for camphor (see figure). This concept, known as the **isoprene rule,** earned Wallach the Nobel Prize in Chemistry in 1910.

By the 1930s, faced with a bewildering array of terpenoid substances, Leopold Ruzicka and his contemporaries sought to develop a unifying principle that could rationalize the natural occurrence of all of the known terpenoids, even those that did not strictly fit Wallach's isoprene rule. Ruzicka's ingenious solution to the problem was to focus on reaction mechanisms and ignore the precise character of the biological precursor, assuming only that it had a terpenoid structure during reaction. He hypothesized the involvement of electrophilic reactions that generated carbocationic intermediates, which underwent subsequent C_5 addition, cyclization, and in some cases skeletal rearrangement before elimination of a proton or capture by a nucleophile to yield the observed terpenoid products. This proposal, which Ruzicka called the **biogenetic isoprene rule,** can be stated simply: A compound is "isoprenoid" if it is derived biologically from an "isoprenoid" precursor, with or without rearrangements. Ruzicka's concept differs from Wallach's in its emphasis on biochemical origin rather than structure. The great strength of the biogenetic isoprene rule lay in its use of mechanistic considerations to classify the bulk of known terpenoids, including structures that did not strictly follow Wallach's isoprene rule. Application of the biogenetic isoprene rule to the origin of several of the common monoterpene skeletons is illustrated in the right panel of the figure (note the bornane skeleton from which camphor is derived). Ruzicka was awarded the Nobel Prize in Chemistry in 1939.



(Fig. 24.3D) produce and accumulate a defensive resin consisting of turpentine (monoterpene olefins) and rosin (diterpenoid resin acids). Triterpenoid surface waxes are formed and excreted from specialized epidermis, and laticifers produce certain triterpenes and polyterpenes such as rubber. These specialized structures sequester natural products away from sensitive metabolic processes and thereby prevent autotoxicity. Most structures of this type are nonphotosynthetic and must therefore rely on adjacent cells to supply

the carbon and energy needed to drive terpenoid biosynthesis.

A more fundamental, and perhaps universal, feature of the organization of terpenoid metabolism exists at the subcellular level. The sesquiterpenes (C_{15}), triterpenes (C_{30}), and polyterpenes appear to be produced in the cytosolic and endoplasmic reticulum (ER) compartments, whereas isoprene, the monoterpenes (C_{10}), diterpenes (C_{20}), tetraterpenes (C_{40}), and certain prenylated quinones originate largely, if not exclusively, in the plastids. The evidence now indicates that the biosynthetic pathways for the formation of the fundamental precursor IPP differ markedly in these compartments, with the classical acetate/mevalonate pathway being active in the cytosol and ER and the glyceraldehyde phosphate/pyruvate pathway operating in the plastids. Regulation of these dual pathways may be difficult to assess, given that plastids may supply IPP to the cytosol for use in biosynthesis, and vice versa. Mitochondria, a third compartment, may generate the ubiquinone prenyl group by the acetate/mevalonate pathway, although little is known about the capability of these organelles for terpenoid biosynthesis.

24.2.2 Hydroxymethylglutaryl-CoA reductase, an enzyme in the acetate/mevalonate pathway, is highly regulated.

The basic enzymology of IPP biosynthesis by way of the acetate/mevalonate pathway is widely accepted (Fig. 24.4). This cytosolic IPP pathway involves the two-step condensation of three molecules of acetyl-CoA catalyzed by thiolase and hydroxymethylglutaryl-CoA synthase. The resulting product, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA), is subsequently reduced by

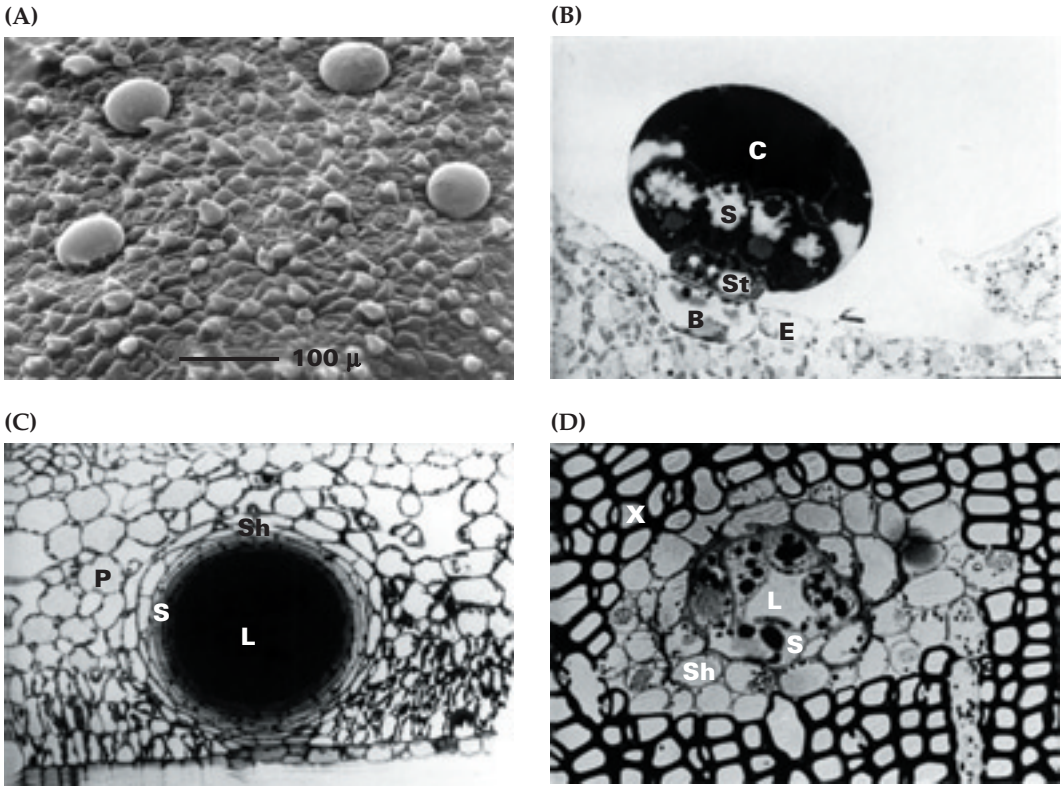


Figure 24.3
(A) Scanning electron micrograph of the leaf surface of thyme. The round structures are peltate glandular trichomes, in which monoterpenes and sesquiterpenes are synthesized. (B) Light micrograph of a glandular trichome from spearmint, shown in longitudinal section. C, subcuticular space; S, secretory cells; St, stalk; B, basal cell; E, epidermal cell. (C) Light micrograph of a secretory cavity in a lemon leaf, shown in cross-section. L, lumen; Sh, sheath cells; P, parenchyma cell. (D) Light micrograph of a resin duct in wood of Jeffrey pine, shown in cross-section. X, secondary xylem.

HMG-CoA reductase in two coupled reactions that form mevalonic acid. Two sequential ATP-dependent phosphorylations of mevalonic acid and a subsequent phosphorylation/elimination-assisted decarboxylation yield IPP.

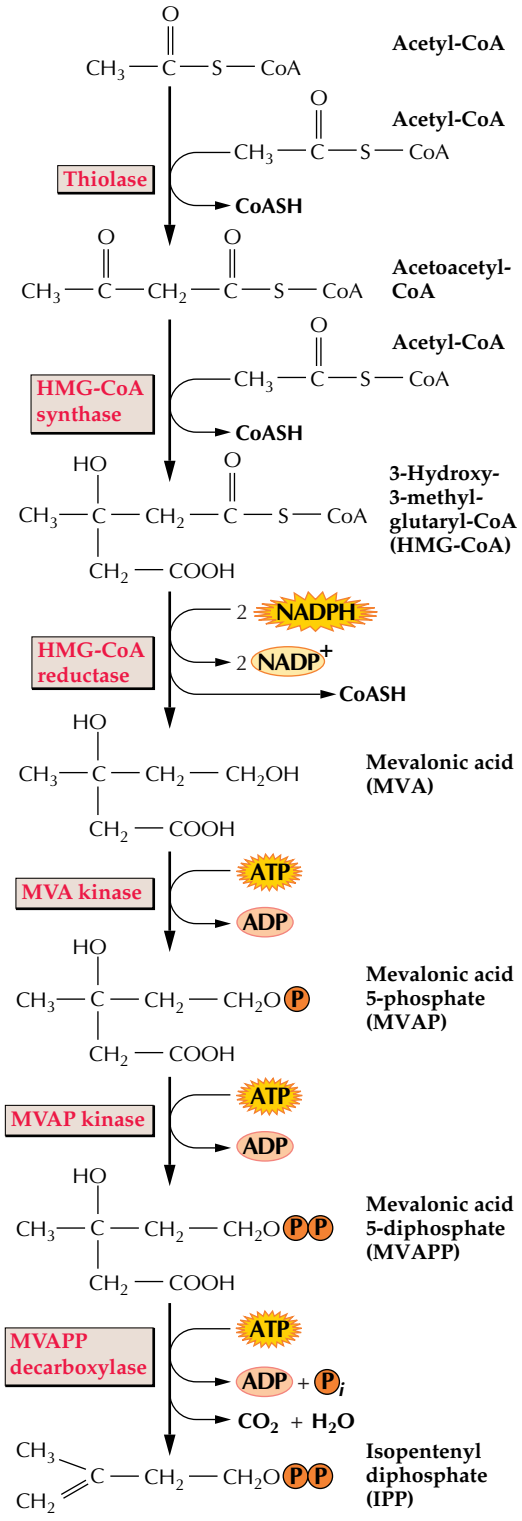


Figure 24.4
The acetate/mevalonate pathway for the formation of IPP, the basic five-carbon unit of terpenoid biosynthesis. Synthesis of each IPP unit requires three molecules of acetyl-CoA.

HMG-CoA reductase is one of the most highly regulated enzymes in animals, being largely responsible for the control of cholesterol biosynthesis. Accumulated evidence indicates that the plant enzyme, which is located in the ER membrane, is also highly regulated. In many cases, small gene families, each containing multiple members, encode this reductase. These gene families are expressed in complex patterns, with individual genes exhibiting constitutive, tissue- or development-specific, or hormone-inducible expression. Specific HMG-CoA reductase genes can be induced by wounding or pathogen infection. The activity of HMG-CoA reductase may be subject to posttranslational regulation, for example, by a protein kinase cascade that phosphorylates and thereby inactivates the enzyme. Allosteric modulation probably also plays a regulatory role. Proteolytic degradation of HMG-CoA reductase protein and the rate of turnover of the corresponding mRNA transcripts may also influence enzyme activity. Researchers have not arrived at a unified scheme that explains how the various mechanisms that regulate HMG-CoA reductase facilitate the production of different terpenoid families. The precise biochemical controls that influence activity have been difficult to assess in vitro because the enzyme is associated with the ER membrane. A model proposed to rationalize the selective participation of HMG-CoA reductase in the biosynthesis of different mevalonate-derived terpenoids is shown in Figure 24.5.

24.2.3 In plastids, IPP is synthesized from pyruvate and glyceraldehyde 3-phosphate.

The plastid-localized route to IPP involves a different pathway, demonstrated in green algae and many eubacteria as well as plants. In this pathway, pyruvate reacts with thiamine pyrophosphate (TPP) to yield a two-carbon fragment, hydroxyethyl-TPP, which condenses with glyceraldehyde 3-phosphate (see Chapter 12, Fig. 12.41, for similar TPP-mediated C_2 transfers catalyzed by transketolase). TPP is released to form a five-carbon intermediate, 1-deoxy-D-xylulose 5-phosphate, which is rearranged and reduced to form 2-*C*-methyl-D-erythritol 4-phosphate and subsequently transformed to yield IPP (Fig. 24.6, upper pathway). The

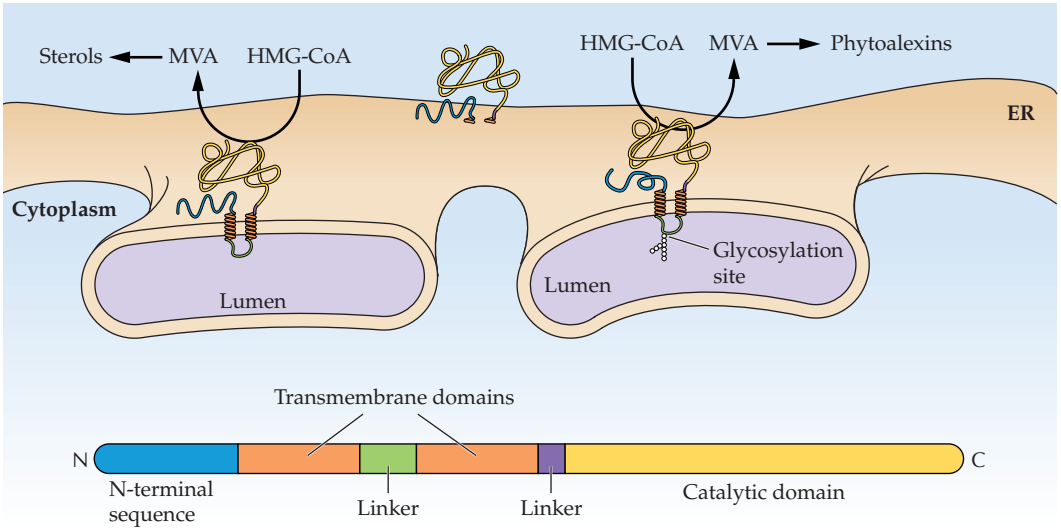


Figure 24.5
Model for the membrane topology of HMG-CoA reductase (HMGR). The protein includes a highly variable hydrophilic N-terminal sequence (blue), a conserved membrane anchor (orange), a highly variable linker sequence (green and purple), and a highly conserved, cytosol-exposed, C-terminal catalytic domain (yellow). Isoforms of HMGR that are associated with elicitor-induced synthesis of sesquiterpenoid phytoalexins contain an N-linked glycosylation site exposed to the ER lumen. Differences in N-terminal sequences and extent of glycosylation may affect targeting of HMGR to various ER domains and to other organelles of the endomembrane system (see Chapters 1 and 4). ER, endoplasmic reticulum; MVA, mevalonic acid.

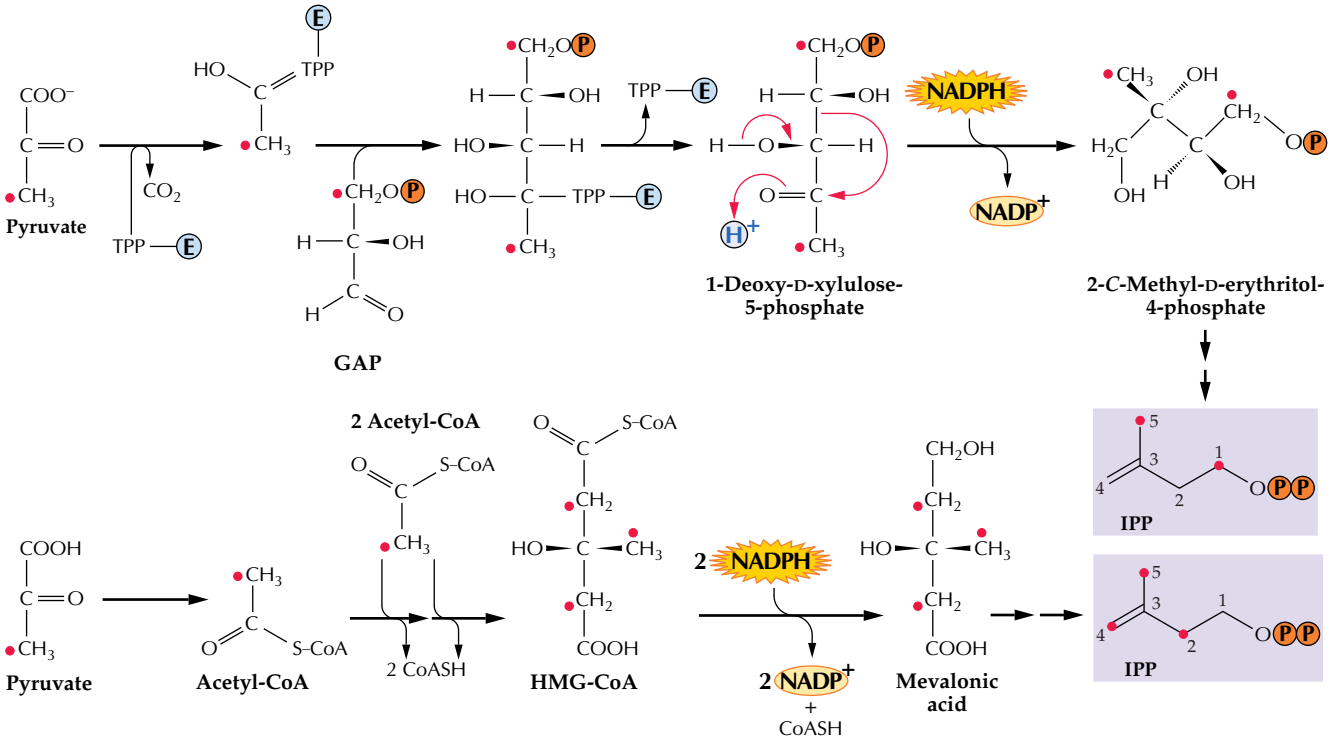


Figure 24.6
Feeding studies distinguish two pathways of isoprenoid biosynthesis. When glucose isotopically labeled at C-1 is transformed by glycolytic enzymes and pyruvate dehydrogenase, the label subsequently appears in the methyl groups of pyruvate and acetyl-CoA and in C-3 of glyceraldehyde 3-phosphate (GAP). IPP synthesized from labeled pyruvate and GAP by the plastid-localized pathway will be labeled at C-1 and C-5 (upper panel), whereas IPP formed from labeled acetyl-CoA by way of the cytosolic acetate/mevalonate pathway will be labeled at C-2, C-4, and C-5 (lower panel).

discovery of this new pathway for IPP formation in plastids suggests that these organelles, presumed to have originated as prokaryotic endosymbionts, have retained the bacterial machinery for the production of this key intermediate of terpenoid biosynthesis.

The details of the glyceraldehyde 3-phosphate/pyruvate pathway and the enzymes responsible have not yet been fully defined. However, products of the two IPP biosynthesis pathways can be easily distinguished in experiments that utilize [1-^{13}C]glucose as a precursor for terpenoid biosynthesis. Nuclear magnetic resonance (NMR) spectroscopy (see Chapter 2, Box 2.2) can be used to determine the ^{13}C-labeling pattern of each isoprene unit in a terpenoid compound, allowing researchers to infer the labeling pattern of the corresponding IPP units (Fig. 24.6).

24.3 Prenyltransferase and terpene synthase reactions

Prenyltransferase enzymes generate the allylic diphosphate esters geranyl diphosphate (GPP), farnesyl diphosphate (FPP), and geranylgeranyl diphosphate (GGPP). Reactions that these compounds undergo (often cyclizations), which are catalyzed by terpene synthases, yield a wide variety of terpenoid compounds. Both prenyltransferases and terpene synthases utilize electrophilic reaction mechanisms involving carbocationic intermediates, a feature of terpenoid biochemistry. Enzymes in both groups share similar properties and contain conserved sequence elements, such as an aspartate-rich DDxxD motif involved in substrate binding, which may participate in initiating divalent metal ion–dependent ionizations.

24.3.1 Repetitive addition of C_5 units is carried out by prenyltransferases.

IPP is utilized in a sequence of elongation reactions to produce a series of prenyl diphosphate homologs, which serve as the immediate precursors of the different families of terpenoids (Fig. 24.7). Isomerization of IPP by IPP isomerase produces the allylic isomer dimethylallyl diphosphate (DMAPP),

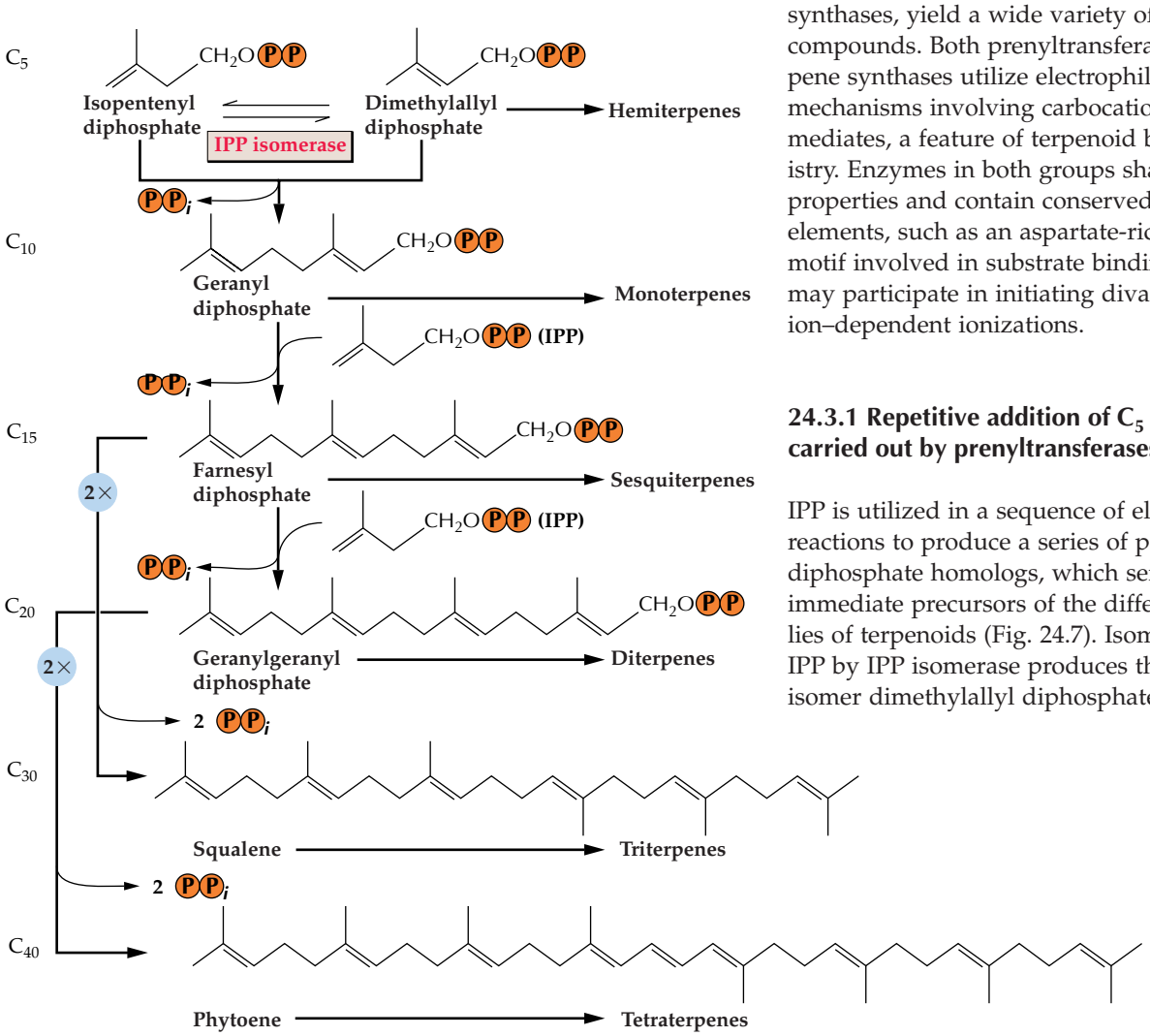


Figure 24.7
The major subclasses of terpenoids are biosynthesized from the basic five-carbon unit, IPP, and from the initial prenyl (allylic) diphosphate, dimethylallyl diphosphate, which is formed by isomerization of IPP. In reactions catalyzed by prenyltransferases, monoterpenes (C_{10}), sesquiterpenes (C_{15}), and diterpenes (C_{20}) are derived from the corresponding intermediates by sequential head-to-tail addition of C_5 units. Triterpenes (C_{30}) are formed from two C_{15} (farnesyl) units joined head-to-head, and tetraterpenes (C_{40}) are formed from two C_{20} (geranylgeranyl) units joined head-to-head.

which is considered the first prenyl diphosphate. Because DMAPP and related prenyl diphosphates contain an allylic double bond, these compounds can be ionized to generate resonance-stabilized carbocations. Once formed, a carbocation intermediate of *n* carbons can react with IPP to yield a prenyl diphosphate homolog containing *n* + 5 carbons. Thus, the reactive primer DMAPP undergoes condensation with IPP to yield the C_{10} intermediate GPP. Repetition of the reaction cycle by addition of one or two molecules of IPP provides FPP (C_{15}) or GGPP (C_{20}), respectively. Each prenyl homolog in the series arises as an allylic diphosphate ester that can ionize to form a resonance-stabilized carbocation and condense with IPP in another round of elongation (Fig. 24.8).

The electrophilic elongation reactions that yield C_{10}, C_{15}, and C_{20} prenyl diphosphates are catalyzed by enzymes known collectively as prenyltransferases. GPP, FPP, and GGPP are each formed by specific prenyltransferases named for their products (e.g., farnesyl diphosphate synthase). The new allylic double bond introduced in the course of the prenyltransferase reaction is commonly in the trans geometry, although this is not always the case: The transferase responsible for rubber biosynthesis introduces cis-double bonds, which are responsible for the elasticity of that polymer. Prenylation reactions are not limited to elongations involving IPP; the same basic carbocationic mechanism permits the attachment of prenyl side chains to atoms of carbon, oxygen, nitrogen, or sulfur in a wide range of nonterpenoid compounds, including proteins.

The most extensively studied prenyltransferase, farnesyl diphosphate synthase, plays an important role in cholesterol biosynthesis in humans. Farnesyl diphosphate synthases from microbes, plants, and animals exhibit high sequence conservation. The first enzyme of the terpenoid pathway to be structurally defined is recombinant avian farnesyl diphosphate synthase, the crystal structure of which has been determined.

24.3.2 The enzyme limonene synthase is a model for monoterpene synthase action.

The families of enzymes responsible for the formation of terpenoids from GPP, FPP, and GGPP are known as monoterpene, sesquiterpene, and diterpene synthases, respectively. These synthases use the corresponding prenyl diphosphates as substrates to form the enormous diversity of carbon skeletons characteristic of terpenoids. Most terpenoids are cyclic, and many contain multiple ring systems, the basic structures of which are determined by the highly specific synthases. Terpenoid synthases that produce cyclic products are also referred to as "cyclases," although examples of synthases producing acyclic products are also known.

A diverse array of monoterpene synthases has been isolated from essential oil-producing angiosperm species and resin-producing gymnosperms. These enzymes use a common mechanism in which ionization

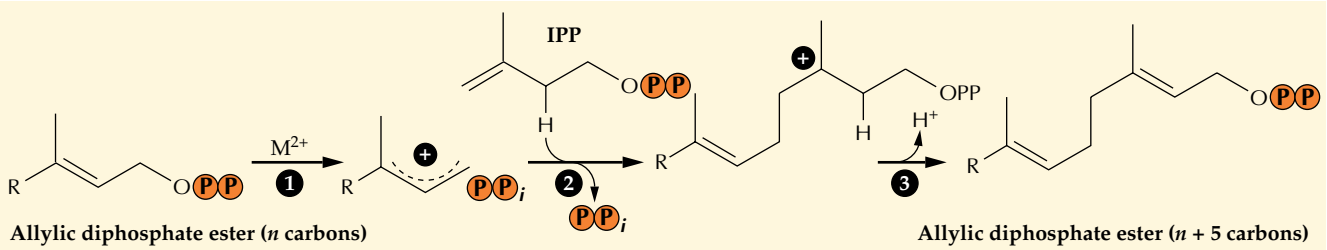
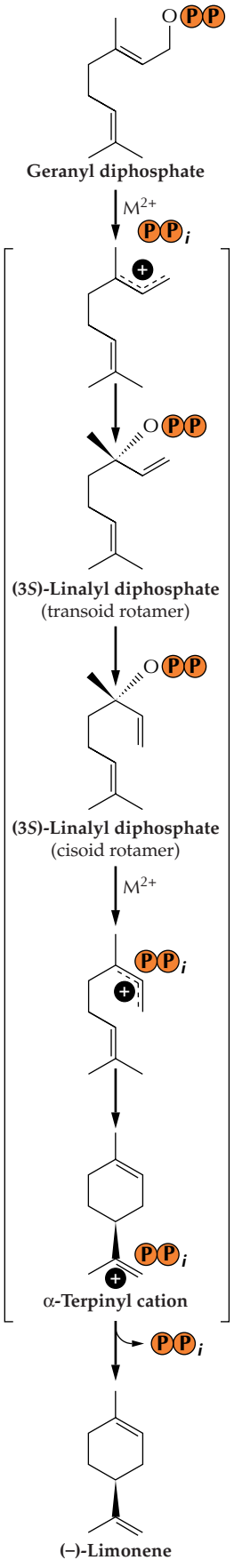


Figure 24.8
The prenyltransferase reaction.



of GPP leads initially to the tertiary allylic isomer linalyl diphosphate (LPP; Fig. 24.9). This isomerization step is required because GPP cannot cyclize directly, given the presence of the trans-double bond. Ionization of the enzyme-bound LPP intermediate promotes cyclization to a six-membered ring carbocation (the α-terpinyl cation), which may undergo additional electrophilic cyclizations, hydride shifts, or other rearrangements before the reaction is terminated by deprotonation of the carbocation or capture by a nucleophile (e.g., water). Variations on this simple mechanistic scheme, involving subsequent reactions of the α-terpinyl carbocation, are responsible for the enzymatic formation of most monoterpene skeletons (see Box 24.1).

The simplest monoterpene synthase reaction is catalyzed by limonene synthase, a useful model for all terpenoid cyclizations (Fig. 24.9). The electrophilic mechanism of action used by limonene synthase can be viewed as an intramolecular equivalent of the prenyltransferase reaction (see Fig. 24.8). Synthases that produce acyclic olefin products (e.g., myrcene) and bicyclic products (α- and β-pinene) from GPP are also known, as are enzymes that transform GPP to oxygenated derivatives such as 1,8-cineole and bornyl diphosphate (Fig. 24.10), the precursor of camphor (see Box 24.1).

An interesting feature of the monoterpene synthases is the ability of these enzymes to produce more than one product; for example, pinene synthase from several plant sources produces both α- and β-pinene. The pinenes are among the most common monoterpenes produced by plants and are principal components of turpentine of the pines, spruces, and firs. The compounds are toxic to bark beetles and their pathogenic fungal symbionts, which cause serious damage to conifer species worldwide. Many conifers respond to bark beetle infestation by up-regulating synthesis of monoterpenes, a process analogous to the production of antimicrobial phytoalexins, when under pathogen attack (Fig. 24.11). Other monoterpenes have quite different functions. Thus, linalool (see Fig. 24.10) and 1,8-cineole emitted by flowers serve as attractants for pollinators, including bees, moths, and bats. 1,8-Cineole and camphor act as foliar feeding deterrents to large herbivores such as hares and deer and also may provide a competitive advantage to several angiosperm species as allelopathic agents that inhibit germination of the seeds of other species.

Exceptions to the general pattern of head-to-tail joining of isoprene units seen in limonene, the pinenes, and most other monoterpenes derived from GPP are the "irregular" monoterpenes. An example of this type is the family of insecticidal monoterpene esters called pyrethrins, found in *Chrysanthemum* and *Tanacetum* species. These monoterpenoids, which exhibit a head-to-middle joining of C_5 units, have gained wide use as commercial insecticides because of their negligible toxicity to mammals and their limited persistence in the environment (see Fig. 24.10).

Figure 24.9
(–)-Limonene synthase catalyzes the simplest of all terpenoid cyclizations and serves as a model for this reaction type. Ionization of GPP, assisted by divalent metal ions, provides the delocalized carbocation–diphosphate anion pair, which collapses to form the enzyme-bound tertiary allylic intermediate linalyl diphosphate. This required isomerization step, followed by rotation about the C-2–C-3 single bond, overcomes the original stereochemical impediment to direct cyclization of the geranyl precursor. A subsequent assisted ionization of the linalyl diphosphate ester promotes an anti-endo-cyclization to the α-terpinyl cation, which undergoes deprotonation to form limonene, a compound now thought to be an important cancer preventive in humans.

24.3.3 Sesquiterpene synthases generate several compounds that function in plant defense.

The electrophilic mechanisms for the formation of the C_{15} sesquiterpenes from FPP closely resemble those used by monoterpene synthases, although the increased flexibility of the 15-carbon farnesyl chain eliminates the need for the preliminary isomerization step except in forming cyclohexanoid-type compounds. The additional C_5 unit and double bond of FPP also permit formation of a greater number of skeletal structures than in the monoterpene series. The best known sesquiterpene synthase of plant origin is *epi*-aristolochene synthase from tobacco, the crystal structure of which has been determined (Fig. 24.12). This enzyme cyclizes FPP

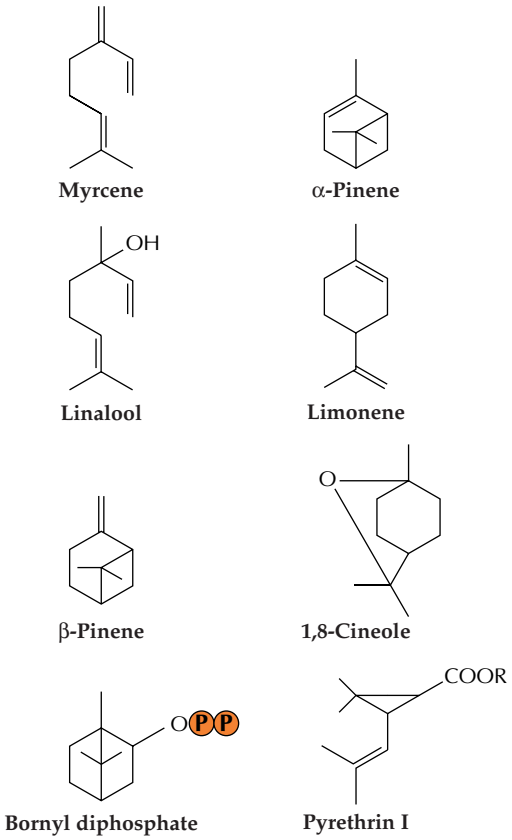


Figure 24.10
Structures of monoterpenes, including insecticidal compounds (α- and β-pinene, pyrethrin), pollinator attractants (linalool and 1,8-cineole), and antiherbivory agents (1,8-cineole).

Figure 24.11
Mass attack by mountain pine beetles on a lodgepole pine (*Pinus contorta*) bole. Each white spot on the trunk represents a beetle entry point at which resin has been secreted. This tree has survived the attack because turpentine production was sufficient to kill all of the bark beetles, which have been "pitched out" by resin outflow. On evaporation of the turpentine and exposure to air, the diterpenoid resin acids form a solid plug that seals the wound.

and catalyzes a methyl migration to yield the olefin precursor of the phytoalexin capsidiol, which is elicited by pathogen attack. Vetispiradiene synthase from potato provides the olefin precursor of the phytoalexin lubimin in this species, whereas δ-cadinene synthase from cotton yields the olefin precursor of the important defense compound gossypol, the latter being currently studied as a possible male contraceptive (Fig. 24.13). Some sesquiterpene synthases involved in the production of conifer resin are capable of individually producing more than 25 different olefins.

24.3.4 Diterpene synthases catalyze two distinct types of cyclization reactions.

Two fundamentally different types of enzymatic cyclization reactions occur in the transformation of GGPP to diterpenes (Fig. 24.14). The first resembles the reactions catalyzed by monoterpene and sesquiterpene synthases, in which the cyclization involves ionization of the diphosphate ester and attack of the resulting carbocation on an interior double bond of the geranylgeranyl substrate. An example of this type is casbene synthase, which is responsible for production of the phytoalexin casbene in castor bean. Taxadiene synthase from yew species uses a mechanistically similar, but more complex, cyclization to produce the tricyclic olefin precursor of taxol.

Abietadiene synthase from grand fir exemplifies the second type of cyclization, in which protonation of the terminal double bond to generate a carbonium ion initiates the first cyclization to a bicyclic intermediate (labdadienyl diphosphate, also known as copalyl diphosphate). Ionization of the diphosphate ester promotes the second cyclization step to give the tricyclic olefin product, abietadiene; a single enzyme catalyzes both

cyclization steps. Oxidation of a methyl group yields abietic acid (see Fig. 24.1), one of the most common diterpenoid resin acids of conifers and important for wound sealing in these species. Fossilization of this resin produces amber.

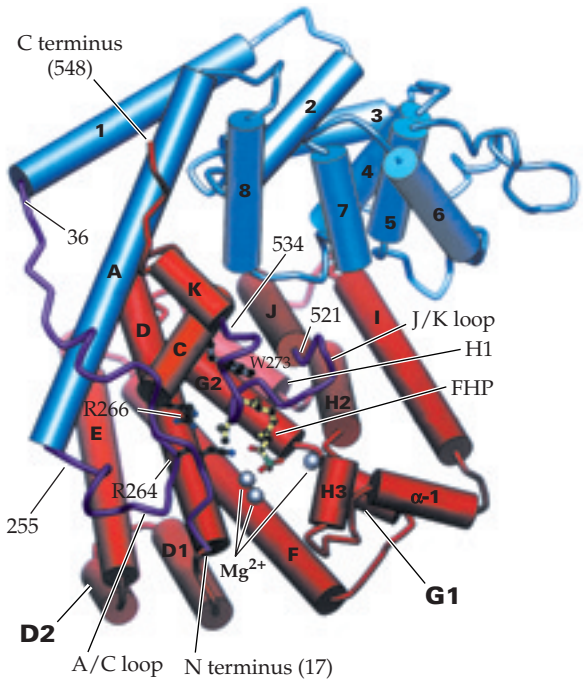


Figure 24.12
Schematic view of *epi*-aristolochene synthase complexed with the substrate analog farnesyl hydroxyphosphonate (FHP). Blue rods represent α-helices in the N-terminal domain; red rods represent α-helices in the C-terminal domain. Loop regions shown in purple are disordered in the native enzyme. Three Mg^{2+} ions and arginines 264 and 266 are involved in the initial steps of the reaction and are labeled near the entrance to the active site. Tryptophan 273, which serves as the general base in the final deprotonation step, is shown within the hydrophobic active-site pocket. The substrate analog FHP is shown in ball-and-stick representation, highlighted with yellow carbon–carbon bonds. Naming of helices in the C-terminal domain corresponds to the convention used for FPP synthase.

24.3.5 Triterpene synthesis proceeds from squalene, tetraterpene synthesis from phytoene.

Before cyclization can occur in the triterpene (C_{30}) series, two molecules of FPP (C_{15}) are first joined in a head-to-head condensation to produce squalene (see Fig. 24.7). The catalyst, squalene synthase, is a prenyltransferase that catalyzes a complex series of cationic rearrangements to accomplish the chemically difficult chore of joining the C-1 carbons of two farnesyl residues. Squalene is usually oxidized to form the 2,3-epoxide, oxidosqualene, and then cyclized in a protonation-initiated reaction to produce, for example,

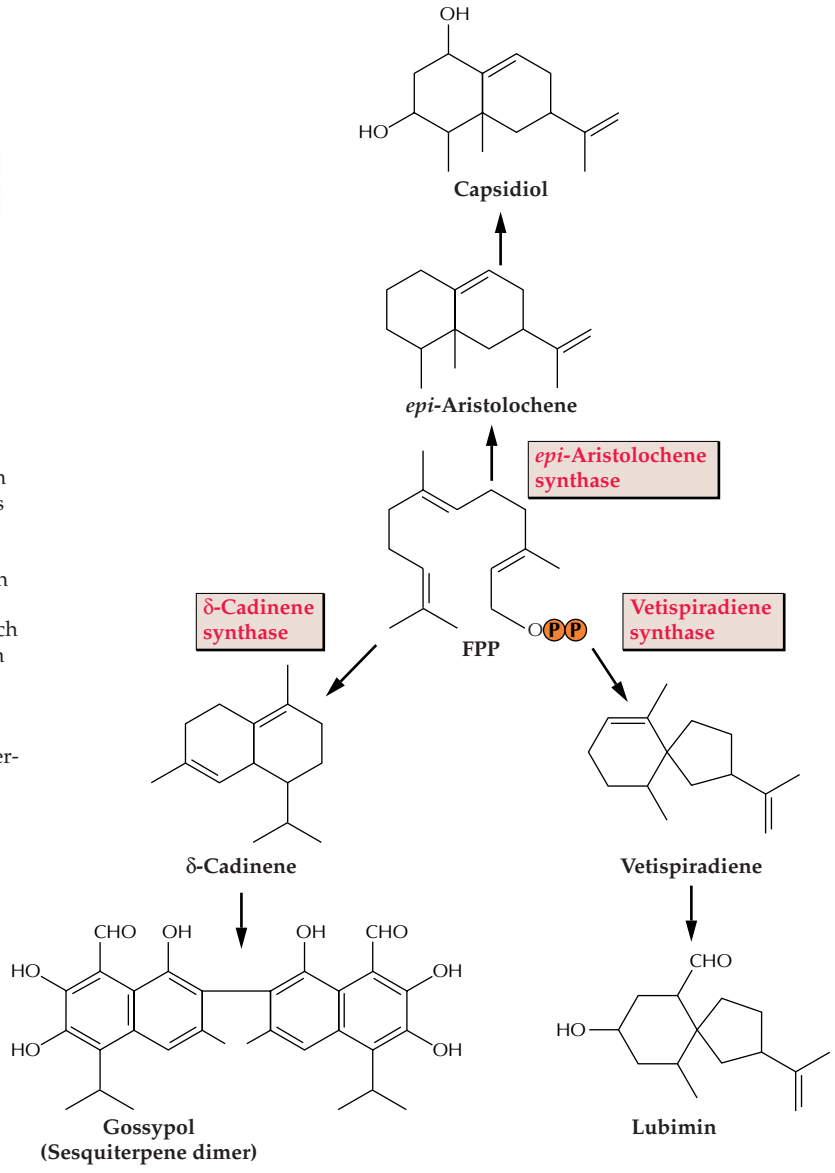


Figure 24.13
Structures of sesquiterpenes biosynthetically derived from FPP. The end products function in plant defense.

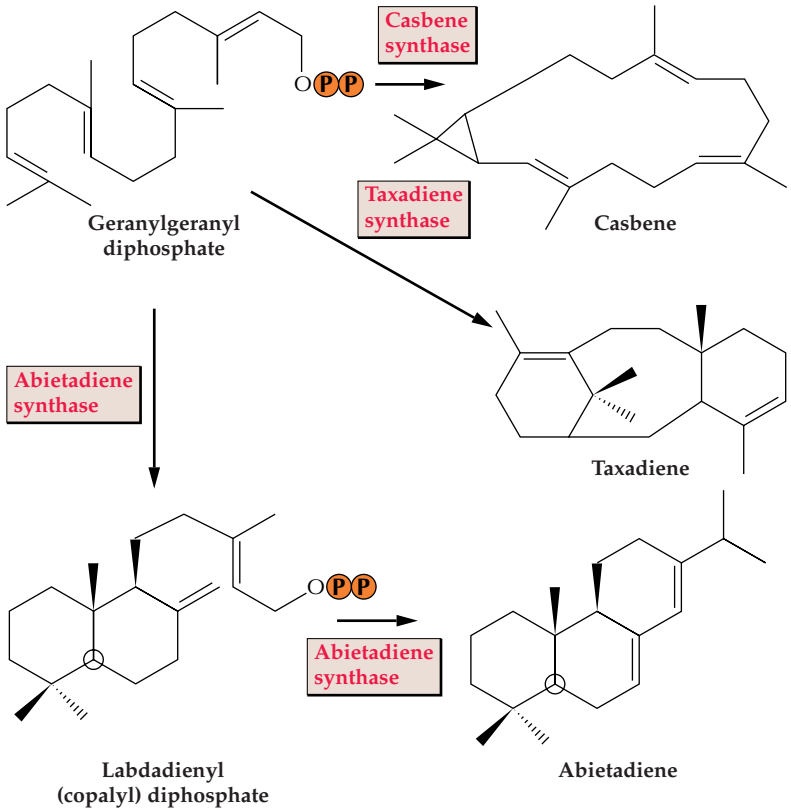


Figure 24.14
Cyclization of GGPP to form the diterpenes casbene, taxadiene, and abietadiene. Cyclization can proceed by one of two distinct mechanims, only one of which yields the intermediate labdadienyl (copalyl) diphosphate.

the common sterol cycloartenol (Fig. 24.15), a precursor of many other phytosterols and brassinosteroids (see Chapter 17). Several alternative modes of cyclization in the triterpene series are also known, such as that leading to the pentacyclic compound β-amyrin, the precursor of oleanolic acid found in the surface wax of several fruits (Fig. 24.15). Preliminary evidence suggests that sesquiterpene biosynthesis and triterpene biosynthesis (both of which utilize cytosolic FPP as a precursor) are reciprocally regulated during the induced defense responses, such that production of C_{15} defensive compounds is enhanced and C_{30} synthesis is repressed.

The tetraterpenes (C_{40}) are produced by joining two molecules of GGPP in head-to-head fashion to produce phytoene, in a manner analogous to the formation of squalene (see Fig. 24.7). The reaction is catalyzed by phytoene synthase, which deploys a mechanism very similar to that of squalene synthase. A series of desaturation steps precedes cyclization in the tetraterpene (carotenoid) series, usually involving formation of six-membered (ionone) rings at the chain termini to produce, for example, β-carotene from lycopene (see Chapter 12, Fig. 12.7).

24.4 Modification of terpenoid skeletons

Subsequent modifications of the basic parent skeletons produced by the terpenoid synthases are responsible for generating the myriad different terpenoids produced by plants. These secondary transformations most commonly involve oxidation, reduction, isomerization, and conjugation reactions, which impart functional properties to the terpenoid molecules. Several oxygenated derivatives of parent terpenoids have already been described in this chapter, including capsidiol, lubimin, gossypol, abietic acid, and oleanolic acid.

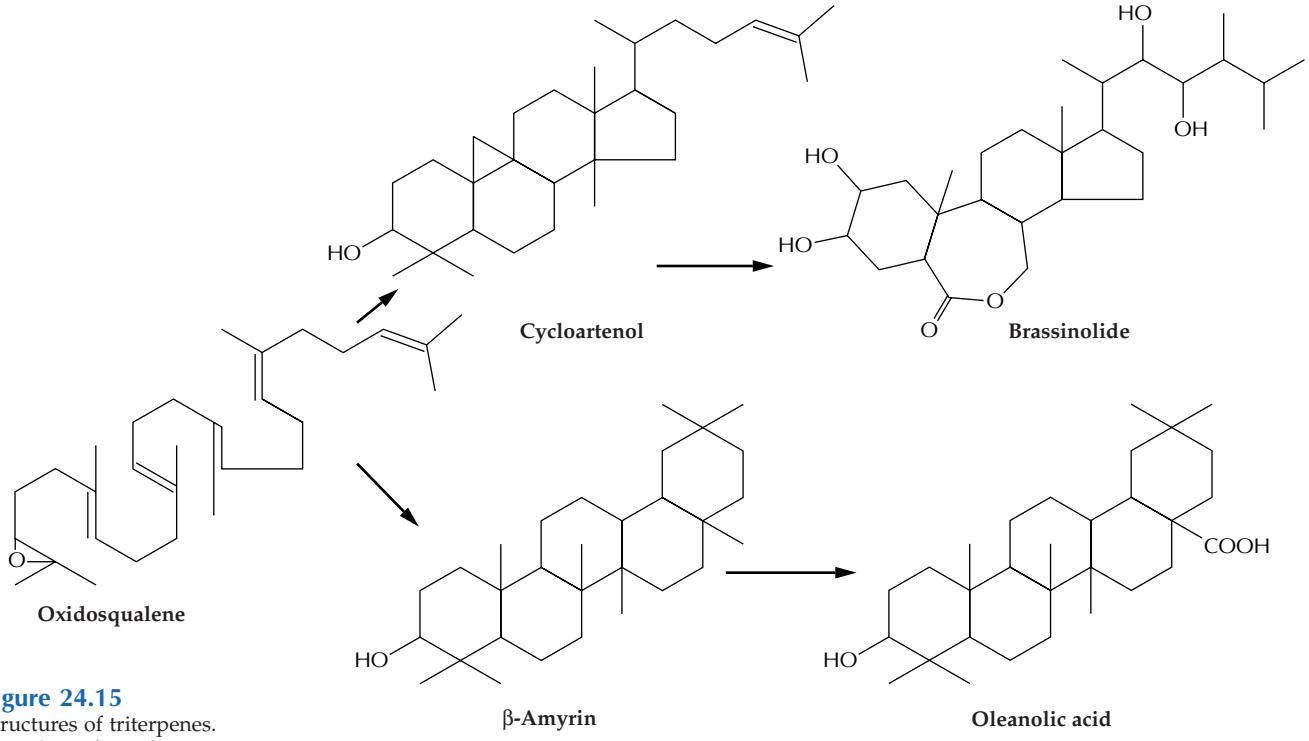


Figure 24.15
Structures of triterpenes. This class of squalene-derived products includes brassinosteroid regulators of plant growth and surface wax components.

Many of the hydroxylations or epoxidations involved in introducing oxygen atoms into the terpenoid skeletons are performed by cytochrome P450 mixed-function oxidases. Because these reactions are not unique to terpenoid biosynthesis, this section will not focus on specific enzyme types but rather on the general role of secondary transformations as the wellspring of diversity in terpenoid structure and function.

24.4.1 The conversion of (–)-limonene to (–)-menthol in peppermint and carvone in spearmint illustrates the biochemistry of terpenoid modification.

The principal and characteristic essential oil components of peppermint (*Mentha piperita*) and spearmint (*M. spicata*) are produced by secondary enzymatic transformations of (–)-limonene (Fig. 24.16). In peppermint, a microsomal cytochrome P450 limonene 3-hydroxylase introduces an oxygen atom at an allylic position to produce (–)-*trans*-isopiperitenol. A soluble $NADP^+$-dependent dehydrogenase oxidizes the alcohol to a ketone, (–)-isopiperitenone, thereby activating the adjacent double bond for reduction by a soluble, NADPH-dependent, regiospecific reductase to produce (+)-*cis*-isopulegone. An isomerase next moves the remaining double bond into conjugation with the carbonyl group, yielding (+)-pulegone. One regiospecific, NADPH-dependent, stereoselective reductase converts (–)-pulegone to either (+)-isomenthone or the predominant species, (–)-menthone. Similar reductases produce the menthol isomers from these ketones. (–)-Menthol greatly predominates among the menthol isomers (constituting as much as 40% of the essential oil) and is the component primarily responsible for the characteristic flavor and cooling sensation of peppermint. The menthol isomers are often found as acetate esters, formed by the action of an acetyl CoA-dependent acetyltransferase. The menthol and menthyl acetate content of peppermint oil glands increases with leaf maturity. Environmental factors greatly influence oil composition. Water stress and warm night growth conditions both promote the accumulation of the more-oxidized pathway intermediates such as (+)-pulegone.

The pathway in spearmint is much shorter. In this instance, a cytochrome P450 limonene 6-hydroxylase specifically introduces oxygen at the alternative allylic position to produce (–)-*trans*-carveol, which is oxidized to (–)-carvone by the soluble

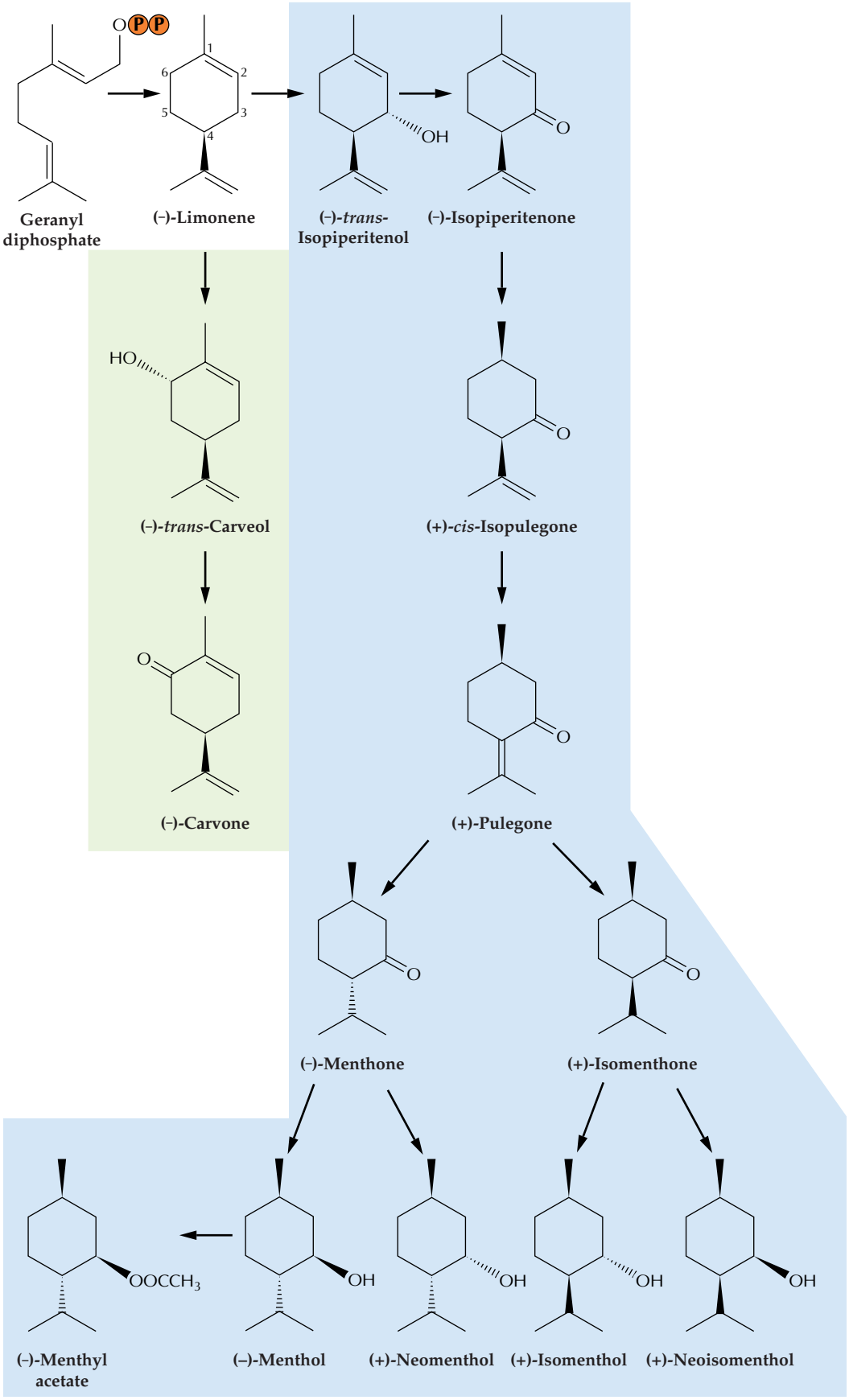


Figure 24.16
Essential oil synthesis in peppermint and spearmint. In peppermint, (–)-limonene is converted to (–)-isopiperitenone, which is modified to form (–)-menthol and related compounds. In spearmint, (–)-limonene is converted to (–)-carvone by a two-step pathway.

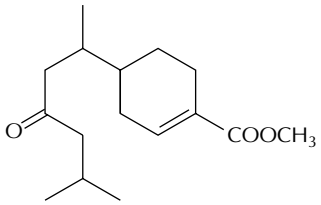
$NADP^+$-dependent dehydrogenase. Although most of the enzymatic machinery present in peppermint oil glands is also present in spearmint, the specificity of these enzymes is such that (–)-carvone is a very poor substrate. Consequently, carvone, the characteristic component of spearmint flavor, accumulates as the major essential oil component (about 70%). Similar reaction sequences initiated by allylic hydroxylations and subsequent redox metabolism and conjugations are very common in the monoterpene, sesquiterpene, and diterpene classes.

24.4.2 Some terpenoid skeletons are extensively decorated.

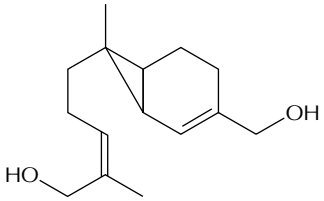
Reactions similar to those responsible for essential oil production in mints generate myriad terpenoid compounds of biological or pharmaceutical interest. Such reactions convert sesquiterpene olefin precursors to phytoalexins (see Fig. 24.13), allelopathic agents, and pollinator attractants. Additional sesquiterpenes generated by modifying olefin precursors include juvabione (Fig. 24.17), a compound from fir species that exhibits insect juvenile hormone activity; sirenin, a sperm attractant of the water mold allomyces; and artemisinin, a potent antimalarial drug from annual wormwood (*Artemisia annua,* also known as Qinghaosu, a plant used in traditional Chinese medicine since about 200 B.C.). A related enzymatic reaction sequence converts the parent diterpene olefin taxadiene to the anticancer drug taxol in yew species, in which the basic terpenoid nucleus is modified extensively by a complex pattern of hydroxylations and acylations. Esters of phorbol (another highly oxygenated diterpene) produced by species of the Euphorbiaceae are powerful irritants and cocarcinogens. After introduction of a hydroxyl group, subsequent oxidation can generate a carboxyl function such as that found in abietic acid (see Fig. 24.1) and oleanolic acid, and also provide the structural elements for lactone ring formation. Sesquiterpenes bearing such lactone rings, e.g., costunolide, are produced and accumulated in the glandular hairs on the leaf surfaces of members of the Asteraceae, where some of these compounds serve as feeding repellents to herbivorous insects and mammals. Monoterpene lactones include nepetalactone (the active principle of catnip as well as an aphid pheromone), a member of the iridoid family of monoterpenes, which are formed by a cyclization reaction quite different from that of other monoterpenes (Fig. 24.17).

The limonoids are a family of oxygenated nortriterpene antiherbivore compounds. Like the sesquiterpene lactones, these substances taste very bitter to humans and probably to other mammals as well. A powerful insect antifeedant compound is azadirachtin A, a highly modified limonoid from the neem tree (*Azadirachta indica*). Other oxygenated triterpenoid natural products with unusual biological properties include the phytoecdysones, a family of plant steroids that act as hormones and stimulate insect molting; the saponins, so named because of their soaplike, detergent properties; and the cardenolides, which, like the saponins, are glycosides, in that they bear one or more attached sugar residues. Ingestion of α-ecdysone by insects disrupts the molting cycle, usually with fatal consequences. The saponins and cardenolides are toxic to many vertebrate herbivores; this family of compounds includes well-known fish poisons and snail poisons of significance in the control of schistosomiasis. Many of these products are also cardioactive and anticholesterolemic agents of pharmacological significance. Digitoxin, the glycone (glycosylated form) of digitoxigenin (Fig. 24.17) extracted from foxglove (*Digitalis*), is used widely in carefully prescribed doses for treatment of congestive heart disease.

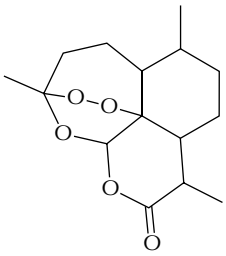
The broad range of insect and higher animal toxins and deterrents among the modified triterpenes leaves little doubt as to their role in plant defense. Interestingly, some herbivores have developed the means to circumvent the toxic effects of these terpenoids and adapt them to their own defense purposes. The classical example of this phenomenon is the monarch butterfly, a specialist feeder on milkweeds (*Asclepias*) which contain cardenolides that are toxic to most herbivores and are even associated with livestock poisoning. Monarch caterpillars, however, feed on milkweeds and accumulate the cardenolides without apparent ill effects. As a result, both caterpillars and the adult butterflies contain enough cardenolides to be toxic to their own predators such as birds.



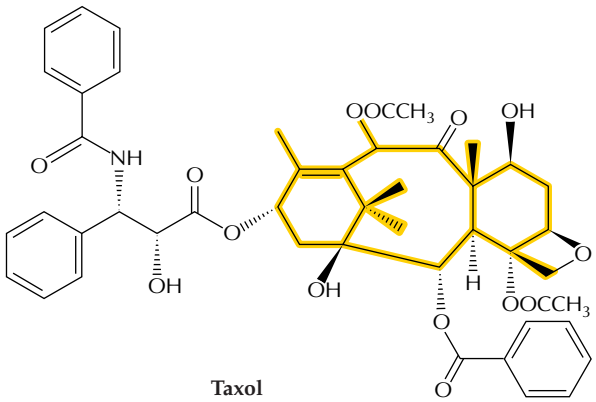
Juvabione
(insect juvenile hormone analog)



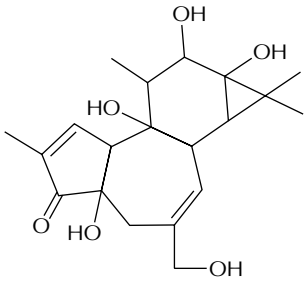
Sirenin
(sperm attractant in water molds)



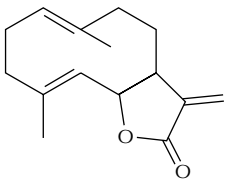
Artemisinin
(antimalarial drug)



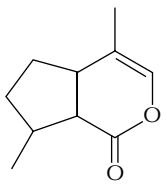
Taxol
(anticancer drug)



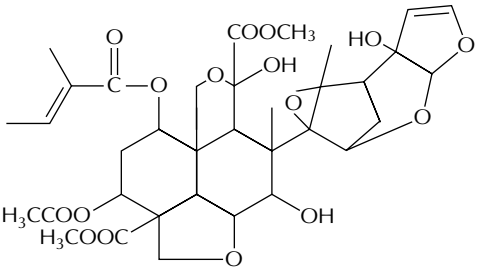
Phorbol
(irritant and cocarcinogen)



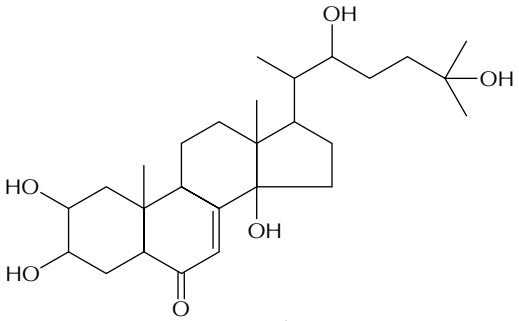
Costunolide
(insect repellent, mammal antifeedant)



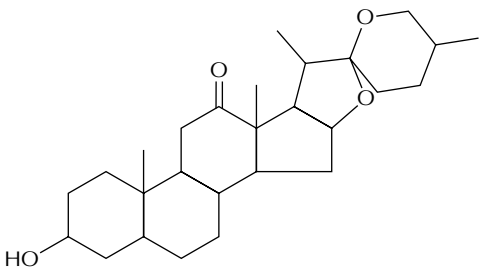
Nepetalactone
(active principle of catnip)



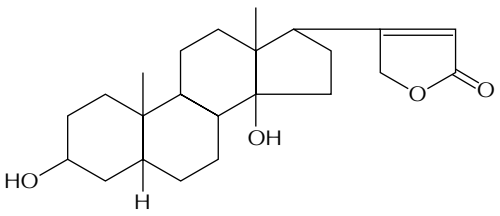
Azadirachtin A
(insect antifeedant)



α-Ecdysone
(disrupts insect molting cycle)



Hecogenin,
the aglycone of a saponin
(a detergent)



Digitoxigenin,
the aglycone of digitoxin, a cardenolide
(treatment of congestive heart disease)

Figure 24.17
Terpenoids formed by secondary transformations of parent cyclic compounds. The yellow highlighting delineates the terpenoid portion of the molecule taxol.

24.5 Toward transgenic terpenoid production

With recent success in the cloning of genes that encode enzymes of terpenoid synthesis, the transgenic manipulation of plant terpenoid metabolism may present a suitable avenue for achieving a number of goals. Several agriculturally important crop species have been bred selectively to produce relatively low amounts of unpalatable terpenoid defense compounds; in the process, these cultivars have lost not only defense capabilities but also, in some cases, quality attributes such as flavor and color. The selective reintroduction of terpenoid-based defense chemistry is certainly conceivable, as is the engineering of pathways into fruits and vegetables to impart desirable flavor properties. The aroma profiles of ornamental plant species might be modified by similar approaches. Likewise, transgene expression might accelerate the rate of slow biosynthetic steps and thereby increase the yields of essential oils used in flavors and perfumes, phytopharmaceuticals (e.g., artemisinin and taxol), insecticides (e.g., pyrethrins and azadirachtin), and a wide range of industrial intermediates that are economically inaccessible by traditional chemical synthesis.

The genetic engineering of terpenoid-based insect defenses is particularly appealing, given the array of available monoterpene, sesquiterpene, diterpene, and triterpene compounds that are toxic to insects not adapted to them. Attracting predators and parasitoids of the target insect or modifying host attractants, oviposition stimulators, and pheromone precursors offers even more sophisticated strategies for pest control. For effective transgenic manipulation of such terpenoid biosynthetic pathways, promoters for tissue-specific, developmentally controlled, and inducible expression are required, as are promoters for targeting production to secretory structures of essential oil plants and conifers. The latter are the most likely species for initial manipulation because they already are adapted for terpenoid accumulation, and the antecedent and subsequent metabolic steps are largely known.

The engineering of terpenoid biosynthetic pathways into plant species that do not ordinarily accumulate these natural products presents a greater opportunity but an even greater challenge, given that little metabolic context exists in these cases. In such species, issues of subcellular sites of synthesis, requirements for sufficiency of precursor flux, and the fate of the desired product might present additional difficulties. Clearly, targeting a terpenoid synthase to the cellular compartment containing the appropriate C_{10}, C_{15}, C_{20}, or C_{30} precursor will be an important consideration. Sufficient flux of IPP at the production site to drive the pathway also will be essential. Because constraints in precursor flow ultimately will limit the effectiveness of transgenes for subsequent pathway steps, information about the flux controls on IPP biosynthesis in both cytosol and plastid, and about the interactions of these controls, is sorely needed.

Very few published examples of the genetic engineering of terpenoid metabolism are currently available, although two notable successes have been achieved in the area of terpenoid vitamins. The ratio of beneficial tocopherol (vitamin E) isomers in oilseeds has been altered by this means, and an increased concentration of β-carotene (a vitamin A precursor) in both rice kernels and rapeseed has been obtained by manipulating the carotenoid pathway. In another, cautionary example, however, overexpression in a transgenic tomato of the enzyme that diverts GGPP to carotenoids resulted in a dwarf phenotype, an unintended consequence of depleting the precursor of the gibberellin plant hormones.

24.6 Alkaloids

24.6.1 Alkaloids have a 3000-year history of human use.

For much of human history, plant extracts have been used as ingredients in potions and poisons. In the eastern Mediterranean, use of the latex of the opium poppy (*Papaver somniferum;* Fig. 24.18) can be traced back at least to 1400 to 1200 B.C. The Sarpagandha root (*Rauwolfia serpentina*) has been used in India since approximately 1000 B.C. Ancient people used medicinal plant extracts as purgatives, antitussives, sedatives, and treatments for a wide range of ailments, including snakebite, fever, and insanity. As the use of medicinal plants spread westward across