5

THE MEVALONATE AND METHYLERYTHRITOL PHOSPHATE PATHWAYS: TERPENOIDS AND STEROIDS

Terpenoids form a large and structurally diverse family of natural products derived from C_5 **isoprene units** (Figure 5.1) joined in a head-to-tail fashion. Typical structures contain carbon skeletons represented by $(C_5)_n$, and are classified as **hemiterpenes** (C_5), **monoterpenes** (C_{10}), **sesquiterpenes** (C_{15}), **diterpenes** (C_{20}), **sesterterpenes** (C_{25}), **triterpenes** (C_{30}), and **tetraterpenes** (C_{40}) (Figure 5.2). Higher polymers are encountered in materials such as rubber. Isoprene itself (Figure 5.1) was known as a decomposition product from various natural cyclic hydrocarbons, and had been suggested as the fundamental building block for these compounds, also referred to as 'isoprenoids'. Isoprene is produced naturally but is not involved in the formation of these compounds; the biochemically active isoprene units were subsequently identified as the diphosphate (pyrophosphate) esters **dimethylallyl diphosphate** (**DMAPP**) and **isopentenyl diphosphate** (**IPP**) (Figure 5.2). Relatively few of the natural terpenoids conform exactly to the simple concept of a linear head-to-tail combination of isoprene units as seen with **geraniol** (C_{10}), **farnesol** (C_{15}), and **geranylgeraniol** (C_{20}) (Figure 5.3). **Squalene** (C_{30}) and **phytoene** (C_{40}), although formed entirely of isoprene units, display a tail-to-tail linkage at the centre of the molecules. Most terpenoids are modified further by cyclization reactions, though the head-to-tail arrangement of the units can usually still be recognized, e.g. **menthol, bisabolene**, and **taxadiene**. The linear arrangement of isoprene units can be much more difficult to appreciate in many other structures when rearrangement reactions have taken place, e.g. steroids, where, in addition, several carbon atoms have been lost. Nevertheless, such compounds are formed by way of regular terpenoid precursors. Terpenoids comprise the largest group of natural products, with over 35 000 known members.

Many other natural products contain terpenoid elements in their molecules, in combination with carbon skeletons derived from other sources, such as the acetate and shikimate pathways. Many alkaloids, phenolics, and vitamins discussed in other chapters are examples of this. A particularly common terpenoid fragment in such cases is a single C_5 unit, usually a dimethylallyl substituent, and molecules containing these isolated isoprene units are sometimes referred to as '**meroterpenoids**'. Some examples include furocoumarins (see page 162), rotenoids (see page 175), and ergot alkaloids (see page 387). One should also note that the term '**prenyl**' is in general use to indicate the dimethylallyl substituent. Even macromolecules like proteins can be modified by attaching terpenoid chains. Cysteine residues in proteins are alkylated with farnesyl (C_{15}) or geranylgeranyl (C_{20}) groups, thereby increasing the lipophilicity of the protein and its ability to associate with membranes.

Medicinal Natural Products: A Biosynthetic Approach, 3rd Edition. Paul M. Dewick
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-74168-9

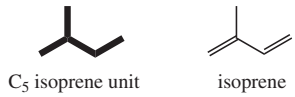


Figure 5.1

MEVALONIC ACID AND METHYLERYTHRITOL PHOSPHATE

The biochemical isoprene units may be derived by two pathways: by way of intermediates **mevalonic acid** (**MVA**) (Figure 5.4) or 2-*C*-methyl-D-erythritol 4-phosphate (**methylerythritol phosphate; MEP**; see Figure 5.6). MVA, itself a product of acetate metabolism, had been established as a precursor of the animal sterol cholesterol, and the steps leading to and from MVA were gradually detailed in a series of painstakingly executed experiments. For many years, the early parts of the mevalonate pathway were believed to be common to the whole range of natural terpenoid derivatives in all organisms. However, after detailed investigation of inconsistencies in labelling patterns, it has since been proven that an alternative pathway to IPP and DMAPP exists, via MEP, and that this pathway is probably more widely utilized in nature than is the mevalonate pathway. This pathway is also referred to as the **mevalonate-independent pathway** or the **deoxyxylulose phosphate pathway**; the terminology **MEP pathway** is preferred, in that MEP is the first committed terpenoid precursor, whilst deoxyxylulose phosphate is also used for the biosynthesis of pyridoxal phosphate (vitamin B_6, page 32) and thiamine (vitamin B_1, page 31).

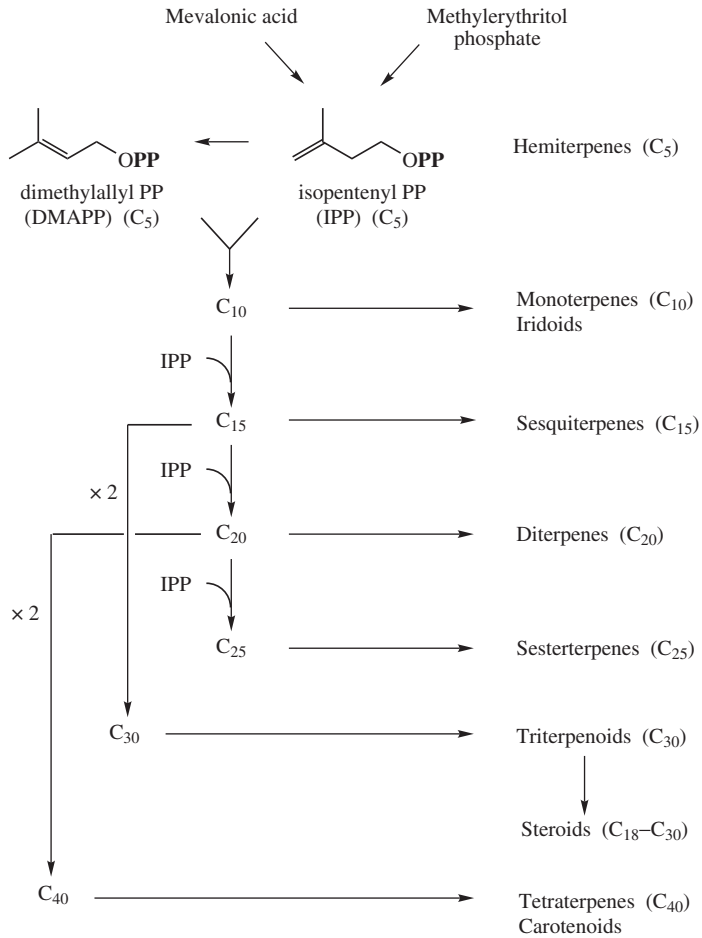


Figure 5.2

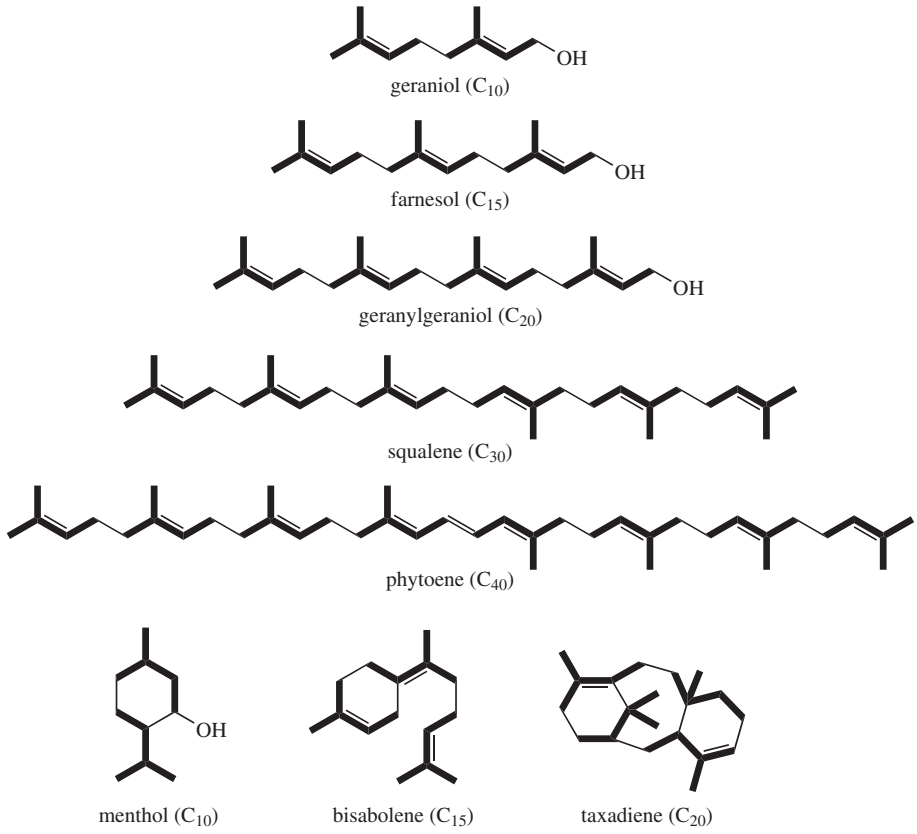
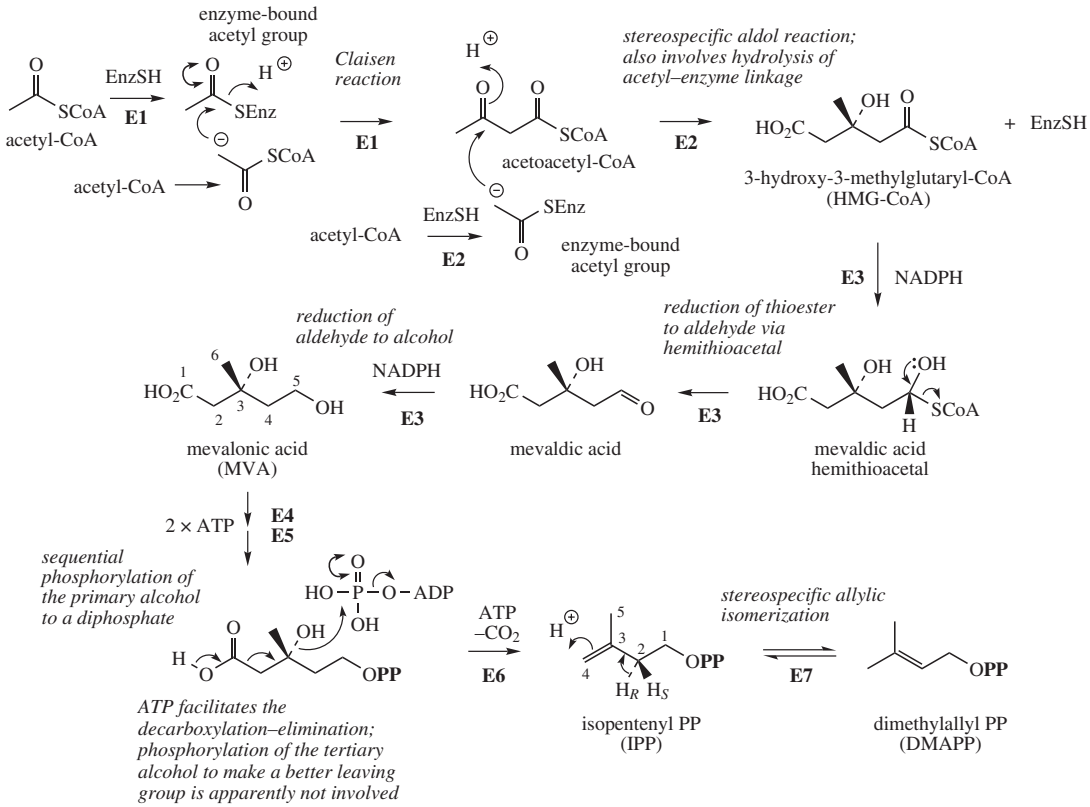


Figure 5.3

Three molecules of acetyl-coenzyme A are used to form **MVA**. Two molecules combine initially in a Claisen condensation to give acetoacetyl-CoA, and a third is incorporated via a stereospecific aldol addition giving the branched-chain ester **3-hydroxy-3-methylglutaryl-CoA** (**HMG-CoA**) (Figure 5.4). Two of the acetyl-CoA molecules appear to be bound to the enzyme via a thiol group. One linkage is broken during the Claisen reaction and the second is subsequently hydrolysed to form the free-acid group of HMG-CoA. Note that the mevalonate pathway does not use malonyl-CoA and it thus diverges from the acetate pathway at the very first step. In the acetate pathway, an equivalent acetoacetyl thioester (bound to the acyl carrier protein, see page 67) would be formed using the thioester of malonic acid as a more nucleophilic species. In the second step of the mevalonate pathway, it should also be noted that, on purely chemical grounds, acetoacetyl-CoA is the more acidic substrate and might be expected to act as the nucleophile rather than the third acetyl-CoA molecule. The enzyme thus achieves what is a less favourable reaction. The conversion of HMG-CoA into (3*R*)-MVA involves a two-step reduction of the thioester group to a primary alcohol via the aldehyde, and provides an essentially irreversible and rate-limiting transformation. Drug-mediated inhibition of this enzyme (**HMG-CoA reductase**) is an important means of regulating the biosynthesis of mevalonate and ultimately of the steroid cholesterol (see statins, page 98).

The six-carbon compound MVA is transformed into the five-carbon phosphorylated isoprene units in a series of reactions, beginning with phosphorylation of the primary alcohol group. Two different ATP-dependent enzymes are involved, resulting in mevalonic acid diphosphate, and decarboxylation–dehydration then follows to give **IPP**. Whilst a third molecule of ATP is required for this last transformation, there is no evidence for phosphorylation of the tertiary hydroxyl, though this would convert the hydroxyl into a better leaving group. Hydrolysis of ATP may assist decarboxylation, as shown in Figure 5.4. IPP



E1: acetoacetyl-CoA synthase
E2: 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase)
E3: 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase)
E4: mevalonate kinase
E5: phosphomevalonate kinase
E6: mevalonate 5-diphosphate decarboxylase
E7: isopentenyl diphosphate isomerase (IPP isomerase)

Figure 5.4

is isomerized to the other isoprene unit, **DMAPP**, by an isomerase enzyme which incorporates a proton from water onto C-4 and stereospecifically removes the *pro-R* proton (H_R) from C-2. This reaction is used to provide the two compounds in the amounts required for further metabolism. Two different types of isomerase enzyme have been distinguished: a type I enzyme requiring a divalent metal ion and a type II enzyme that requires a divalent metal ion together with FMN for activity. Both enzymes appear to employ a protonation–deprotonation mechanism. The conversion of IPP into DMAPP generates a reactive electrophile and, therefore, a good alkylating agent. DMAPP possesses a good leaving group, the diphosphate, and can ionize readily to yield an allylic carbocation which is stabilized by charge delocalization (Figure 5.5). In contrast, IPP with its terminal double bond is more likely to act as a nucleophile, especially towards

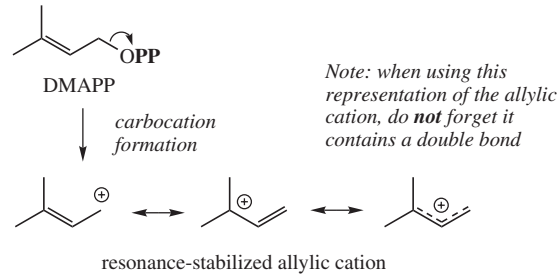
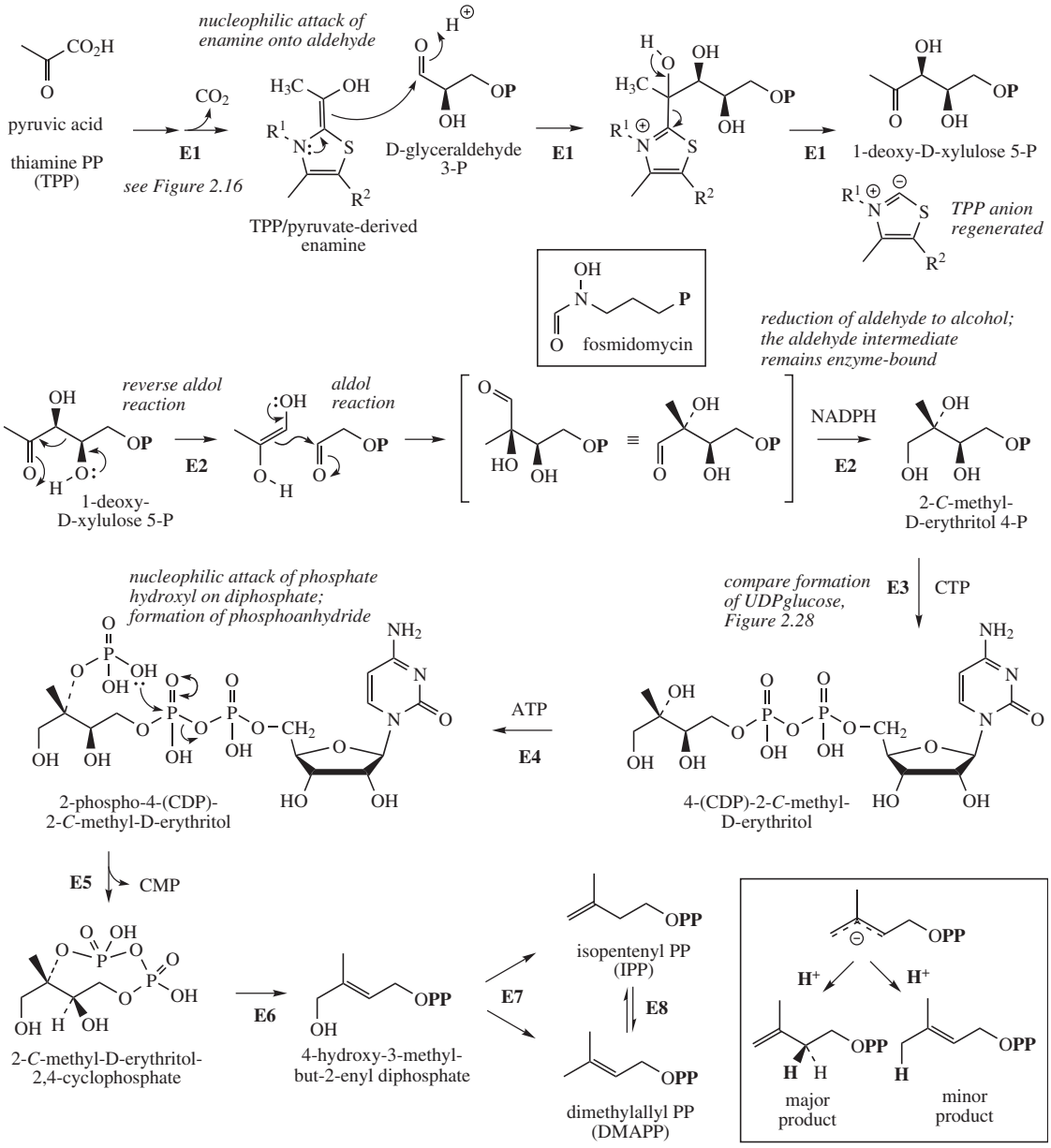


Figure 5.5

the electrophilic DMAPP. These differing reactivities are the basis of terpenoid biosynthesis, and carbocations feature strongly in mechanistic rationalizations of the pathways.



E1: 1-deoxy-D-xylulose 5-phosphate synthase (DXP synthase)
E2: 2-*C*-methyl-D-erythritol 4-phosphate synthase; 1-deoxy-D-xylulose 5-phosphate reductoisomerase (IspC)
E3: 4-diphosphocytidyl-2-*C*-methyl-D-erythritol synthase (IspD)
E4: 4-diphosphocytidyl-2-*C*-methyl-D-erythritol kinase (IspE)
E5: 2-*C*-methyl-D-erythritol-2,4-cyclodiphosphate synthase (IspF)
E6: 4-hydroxy-3-methylbut-2-enyl diphosphate synthase (IspG)
E7: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase (IspH)
E8: isopentenyl diphosphate isomerase (IPP isomerase)

Figure 5.6

Glycolytic pathway intermediates pyruvic acid and glyceraldehyde 3-phosphate are used in the production of **MEP**; the pyruvate carboxyl is lost in this process (Figure 5.6). Thiamine diphosphate-mediated decarboxylation of pyruvate (compare page 23) produces an acetaldehyde-equivalent bound in the form of an enamine. This reacts as a nucleophile in an addition reaction with the glyceraldehyde 3-phosphate. Subsequent release from the TPP carrier generates 1-deoxy-D-xylulose 5-phosphate (**deoxyxylulose phosphate**), which is transformed into MEP by a rearrangement process. This has been shown to involve a reverse aldol–aldol sequence (Figure 5.6), coupled with a reduction. A single enzyme catalyses these skeletal rearrangement and reduction reactions without release of any intermediate; the product now contains the branched-chain system equivalent to the isoprene unit. Reaction of MEP with cytidine triphosphate (CTP) produces a cytidine diphospho derivative (compare uridine diphosphoglucose in glucosylation, page 31), which is then phosphorylated via ATP. The resultant 2-phosphate is then converted into a cyclic phosphoanhydride with loss of cytidine phosphate. The subsequent steps leading to **IPP** and **DMAPP** are the least understood part of the pathway. Gene methodology has shown that two enzymes are involved, the first producing 4-hydroxy-3-methylbut-2-enyl diphosphate and the second converting this into predominantly IPP, but also DMAPP. Both steps are reductive in nature, but mechanisms are yet to be elucidated. The formation of both IPP and DMAPP (ratios are typically in the region 5:1 to 4:1) is suggested to involve a delocalized allylic system (radical or anion, shown in Figure 5.6 as an anion), with protons being supplied by water. Although this pathway coproduces IPP and DMAPP, isomerism of IPP to DMAPP as in the mevalonate pathway is also possible to balance the pool sizes of these intermediates.

Whether the mevalonate pathway or the MEP pathway supplies isoprene units for the biosynthesis of a particular terpenoid must be established experimentally. This can be determined from the results of feeding [1-^{13}C]-D-glucose as precursor; this leads to different labelling patterns in the isoprene unit according to the pathway operating (Figure 5.7). Animals and fungi appear to lack the MEP pathway, so utilize the mevalonate pathway exclusively. The MEP pathway is present in plants, algae, and most bacteria. Plants and some bacteria are equipped with and employ both pathways, often concurrently. In plants, the two pathways appear to be compartmentalized, so that the mevalonate pathway enzymes are localized in the cytosol, whereas the MEP pathway enzymes are found in chloroplasts. The cytosolic pool of IPP serves as a precursor of C_{15} derivatives (farnesyl PP (FPP) and sesquiterpenes; see Figure 5.2), and in due course triterpenoids and steroids ($2 \times$ FPP). Accordingly, triterpenoids and steroids, and some sesquiterpenoids (cytosolic products) are formed by the mevalonate pathway, whilst most other terpenoids (C_{10}, C_{20}, C_{40}) are formed in the chloroplasts and are MEP derived. Of course there are exceptions. There are also examples where the two pathways can supply different portions of a molecule, or where there is exchange of late-stage common intermediates between the two pathways (cross-talk), resulting in a contribution of isoprene units from each pathway. In the following part of this chapter, these complications will not be considered further, and in most cases there is no need to consider the precise source of the isoprene units.

An area of special pharmacological interest where the early pathway is of particular concern is steroid biosynthesis, which appears to be from mevalonate in the vast majority of organisms. Thus, inhibitors of mevalonate pathway enzymes will reduce steroid production in plants, but will not affect the formation of terpenoids derived via MEP. Equally, it is possible to inhibit terpenoid production without affecting steroid formation by the use of MEP pathway inhibitors, such as the antibiotic **fosmidomycin** from *Streptomyces lavendulae*. This acts as an analogue of the rearrangement intermediate in the reaction catalysed by MEP synthase (Figure 5.6). Enzymes of the MEP pathway are attractive targets for development of drugs against microbial diseases such as malaria or tuberculosis, since the MEP pathway is utilized by the pathogen but is not present in humans. Regulation of cholesterol production in humans is an important health concern (see page 251); the widely used statin drugs are specific inhibitors of the mevalonate pathway enzyme HMG-CoA reductase (see page 98).

HEMITERPENES (C_5)

IPP and DMAPP are reactive hemiterpene intermediates in the pathways leading to more complex terpenoid structures. They are also used as alkylating agents in the formation of meroterpenoids, as indicated above, but examples of these structures are discussed elsewhere under the section appropriate to the major substructure, e.g. alkaloids, shikimate, acetate. Relatively few true hemiterpenes are produced in nature. **Isoprene**, a volatile compound which is released in huge amounts by many species of plants, especially woody trees such as oaks, willows, poplars, and spruce, is the notable example. Isoprene is formed by loss of a proton from the allylic cation. Alternatively, quenching the allylic cation with water leads to methylbutenol, produced by several species of pine (Figure 5.8).

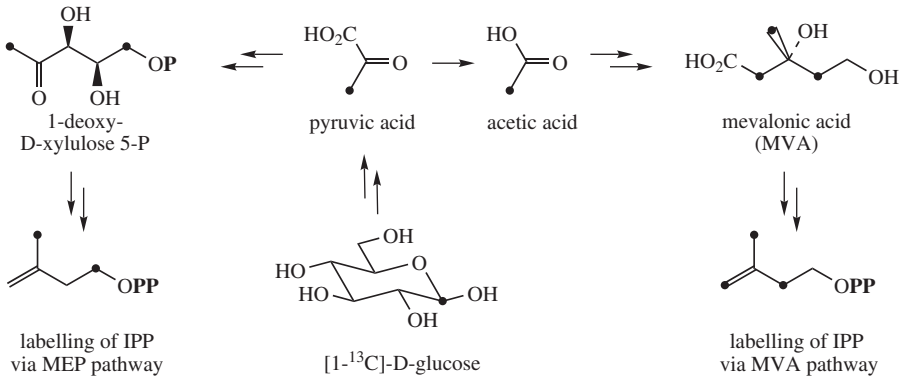


Figure 5.7

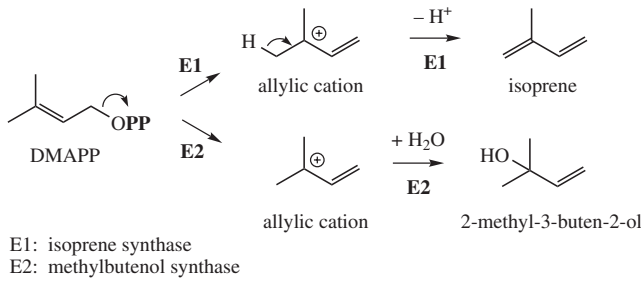


Figure 5.8

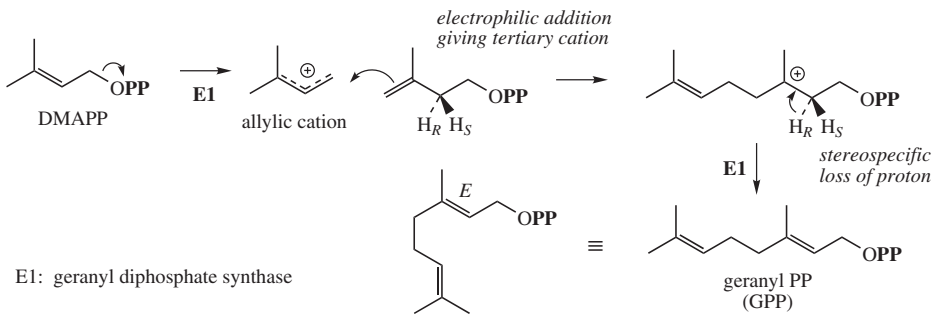


Figure 5.9

MONOTERPENES (C_{10})

Enzyme-catalysed combination of DMAPP and IPP yields **geranyl diphosphate** (**GPP**; Figure 5.9). This is believed to involve ionization of DMAPP to the allylic cation, addition to the double bond of IPP, followed by loss of a proton. The proton lost (H_R) is stereochemically analogous to that lost on the isomerization of IPP to DMAPP; indeed, the two reaction mechanisms are essentially the same, one involving a proton and the other a carbocation. This produces a monoterpene diphosphate, geranyl PP, in which the new double bond is *trans* (*E*).

Linalyl PP and **neryl PP** are isomers of GPP, and are likely to be formed from GPP by ionization to the allylic cation, allowing a change in attachment of the diphosphate group (to the tertiary carbon in linalyl PP) or a

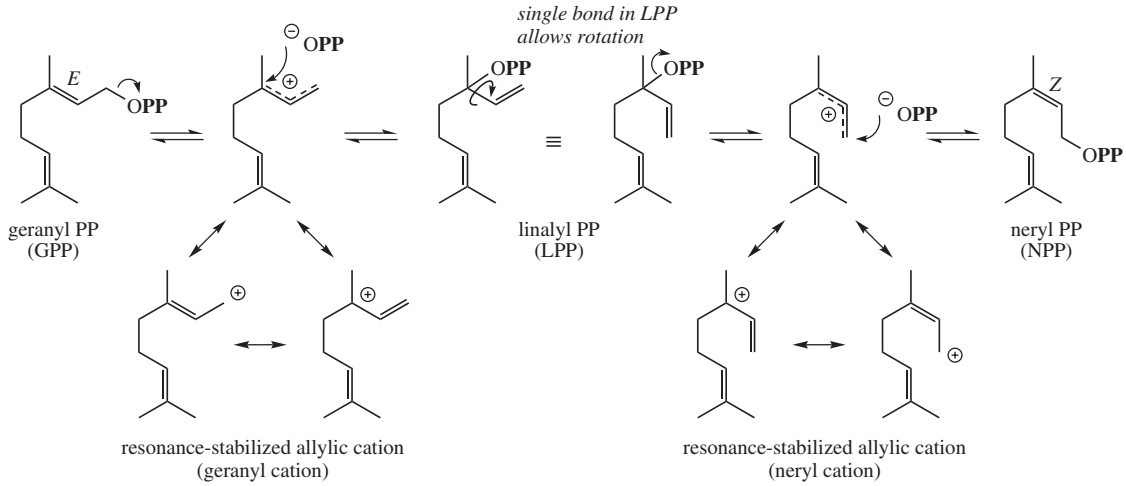


Figure 5.10

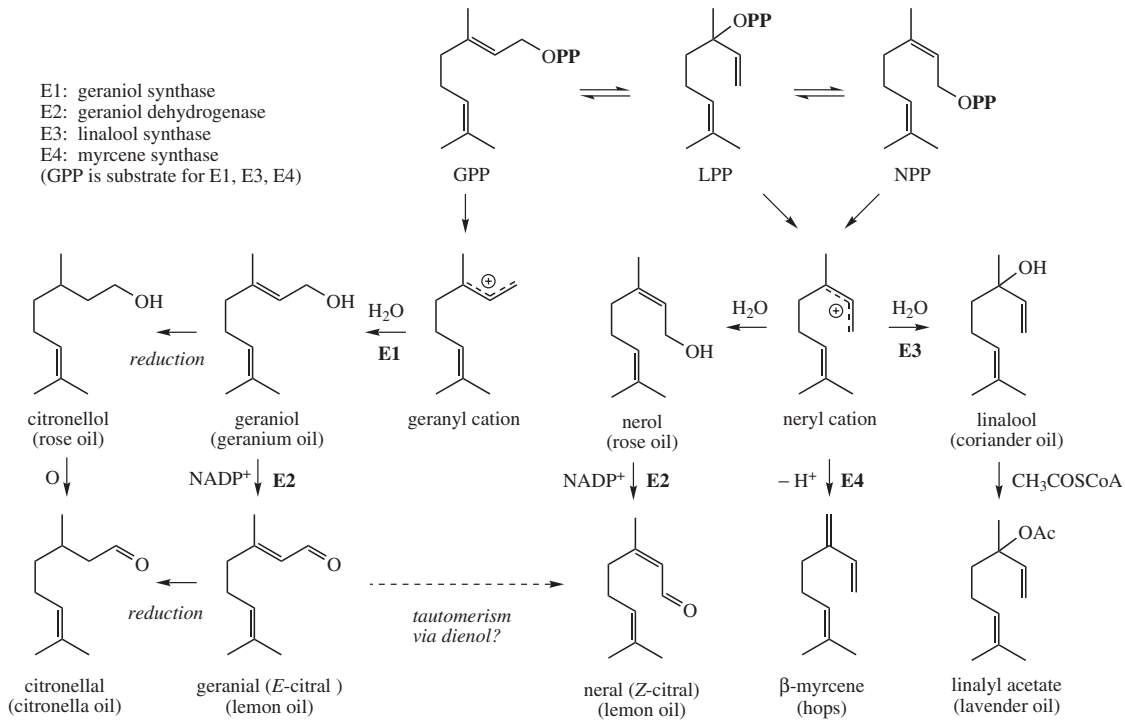


Figure 5.11

change in stereochemistry at the double bond (to *Z* in neryl PP) (Figure 5.10). These three compounds, by relatively modest changes, can give rise to a range of linear monoterpenes found as components of volatile oils used in flavouring and perfumery (Figure 5.11). The resulting compounds may be hydrocarbons, alcohols, aldehydes, or perhaps esters, especially acetates by reaction with acetyl-CoA. Where enzymes have been characterized, it

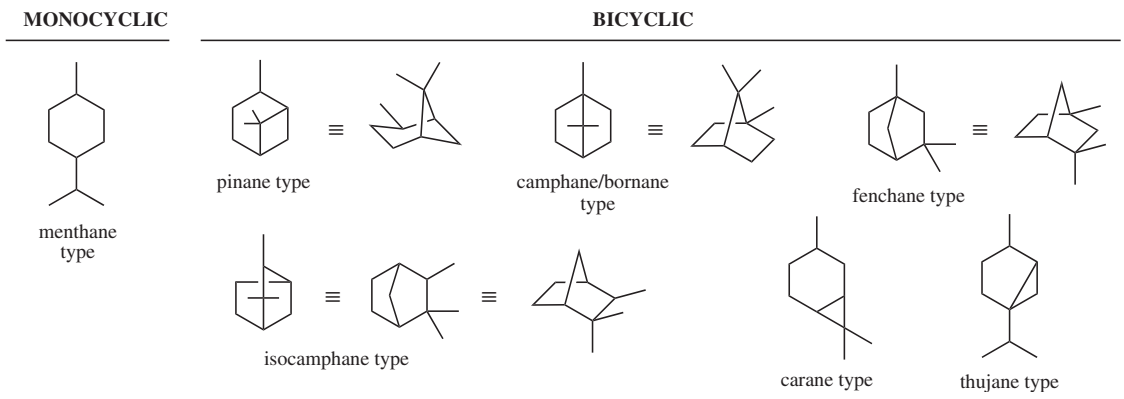


Figure 5.12

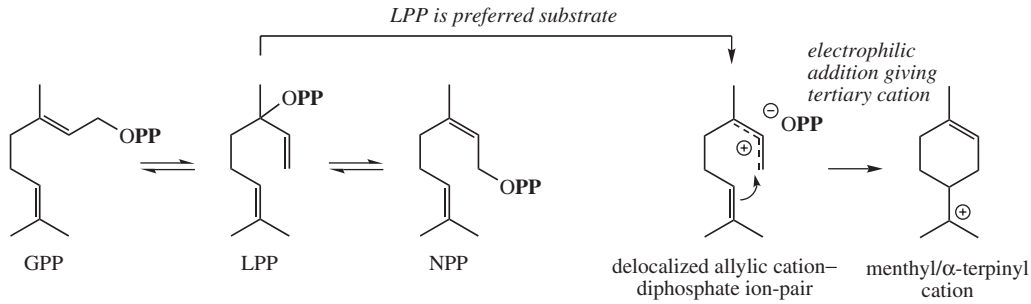


Figure 5.13

has been demonstrated that the reactions proceed through the carbocation intermediates. Thus, geraniol is the result of addition of water to the geranyl cation, and is not formed by hydrolysis of GPP.

The range of monoterpenes encountered is extended considerably by cyclization reactions, and monocyclic or bicyclic systems can be created. Some of the more important examples of these ring systems are shown in Figure 5.12. A considerable amount of information about the enzymes (terpene cyclases), genes, and cyclization mechanisms is now available, providing a satisfactory and detailed picture of these natural products. Cyclizations would not be expected to occur with the precursor GPP, the *E* stereochemistry of the double bond being unfavourable for ring formation (Figure 5.13). Neryl PP or linalyl PP, however, do have favourable stereochemistry, and either or both of these would be more immediate precursors of the monocyclic menthane system, formation of which could be represented as shown, generating a carbocation (termed menthyl or α-terpinyl) that has the menthane skeleton. It has been found that monoterpene cyclase enzymes are able to accept all three diphosphates, with linalyl PP being the best substrate, and it appears they have the ability to isomerize the substrates initially as well as to cyclize them. It is convenient, therefore, to consider the species involved in the cyclization as the delocalized allylic cation tightly bound to the diphosphate anion, and bond formation follows due to the proximity of the π-electrons of the double bond (Figure 5.13).

In Chapter 2, the possible fates of carbocations were discussed. These include quenching with nucleophiles (especially water), loss of a proton, cyclization, and the possibility that Wagner–Meerwein rearrangements might occur (see page 15). All of these feature strongly in terpenoid biosynthesis, and examples are shown in Figures 5.14 and 5.15. The newly generated menthyl cation could be quenched by attack of water, in which case the alcohol **α-terpineol** would be formed, or it could lose a proton to give **limonene** (Figure 5.14). Alternatively, folding the cationic side-chain towards the double bond (via the surface characteristics of the enzyme) would allow a repeat of the cyclization mechanism and produce

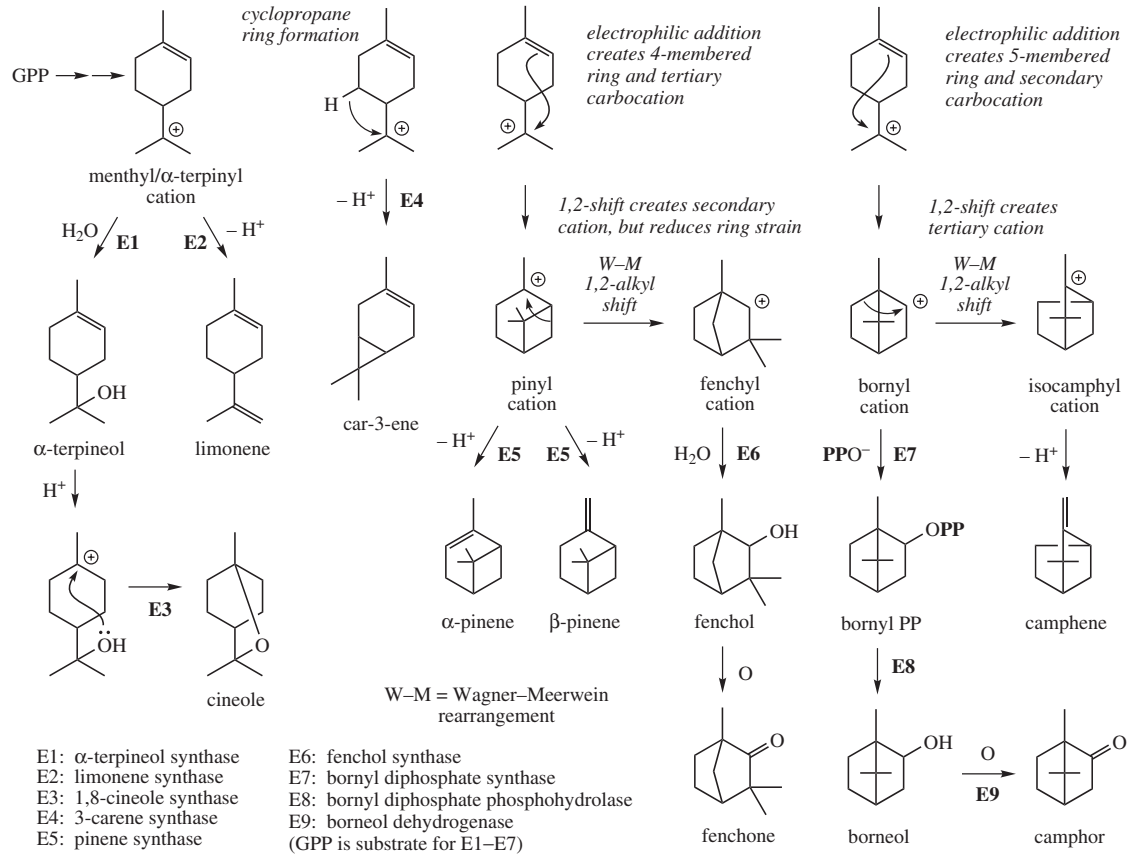


Figure 5.14

bicyclic pinyl and bornyl cations, according to which end of the double bond was involved in forming the new bonds. Thus **α-pinene** and **β-pinene** arise by loss of different protons from the pinyl cation, producing the double bonds as cyclic or exocyclic respectively. **Borneol** could potentially result from quenching of the bornyl cation with water; unusually, though, this alcohol is actually derived by hydrolysis of bornyl diphosphate. Oxidation of the secondary alcohol in borneol then generates the ketone **camphor**. A less common termination step involving loss of a proton is also shown in Figure 5.14. This is the formation of a cyclopropane ring as exemplified by **3-carene** and generation of the carane skeleton.

The chemistry of terpenoid formation is essentially based on the reactivity of carbocations, even though in nature these cations may not exist as discrete species, but rather as tightly bound ion pairs with a counter anion, e.g. diphosphate. The analogy with carbocation chemistry is justified, however, since a high proportion of natural terpenoids have skeletons which have suffered rearrangement processes. Rearrangements of the Wagner–Meerwein type (see page 15), in which carbon atoms or hydride migrate to achieve enhanced stability for the cation via tertiary against secondary character, or by reduction of ring strain, give a mechanistic rationalization for the biosynthetic pathway. The bicyclic pinyl cation, with a strained four-membered ring, rearranges to the less-strained five-membered fenchyl cation, a change which presumably more than makes up for the unfavourable tertiary to secondary carbocation transformation. This produces the fenchane skeleton, exemplified by **fenchol** and **fenchone** (Figure 5.14). The isocamphyl tertiary carbocation is formed from the bornyl secondary carbocation by a Wagner–Meerwein rearrangement, and so leads to **camphene**.

Examples of Wagner–Meerwein rearrangements involving hydride migrations are featured in Figure 5.15.

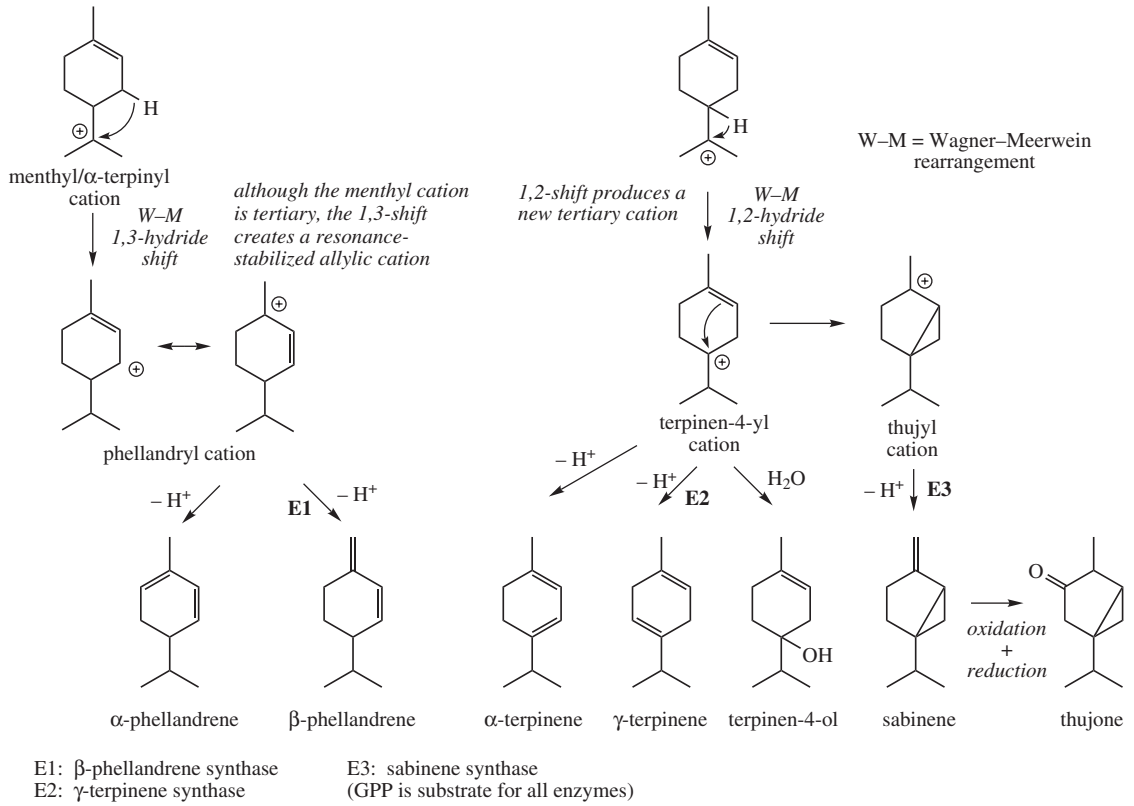


Figure 5.15

The menthyl cation, although it is a tertiary, may be converted by a 1,3-hydride shift into a favourable resonance-stabilized allylic cation. This allows the formation of **α-** and **β-phellandrenes** by loss of a proton from the phellandryl carbocation. A 1,2-hydride shift converting the menthyl cation into the terpinen-4-yl cation only changes one tertiary carbocation system for another, but allows formation of **α-terpinene**, **γ-terpinene**, and the α-terpineol isomer **terpinen-4-ol**. A further cyclization reaction on the terpinen-4-yl cation generates the thujane skeleton, e.g. **sabinene** and **thujone**. Terpinen-4-ol is the primary antibacterial component of tea tree oil from *Melaleuca alternifolia* (Myrtaceae); thujone has achieved notoriety as the neurotoxic agent in wormwood oil from *Artemisia absinthium* (Compositae/Asteraceae) used in preparation of the drink absinthe, now banned in most countries.

So far, little attention has been given to the stereochemical features of the resultant monoterpene. Individual enzyme systems present in a particular organism will, of course, control the folding of the substrate molecule and, thus, define the stereochemistry of the final product. Most monoterpenes are optically active, and there are many examples known where enantiomeric forms of the same compound can be isolated from different sources, e.g. (+)-**camphor** in sage (*Salvia officinalis*; Labiatae/Lamiaceae) and (−)-camphor in tansy (*Tanacetum vulgare*; Compositae/Asteraceae), or (+)-**carvone** in caraway (*Carum carvi*: Umbelliferae/Apiaceae) and (−)-carvone in spearmint (*Mentha spicata*; Labiatae/Lamiaceae). There are also examples of compounds found in both enantiomeric forms in the same organism, examples being (+)- and (−)-**limonene** in peppermint (*Mentha* × *piperita*; Labiatae/Lamiaceae) and (+)- and (−)-**α-pinene** in pine (*Pinus* species; Pinaceae). The individual enantiomers can produce different biological responses, especially towards olfactory receptors in the nose. Thus, the characteristic odour of caraway is due to (+)-carvone, whereas (−)-carvone smells of mint. (+)-Limonene smells of oranges, whilst (−)-limonene resembles the smell of lemons. The origins of the different enantiomeric forms of limonene and

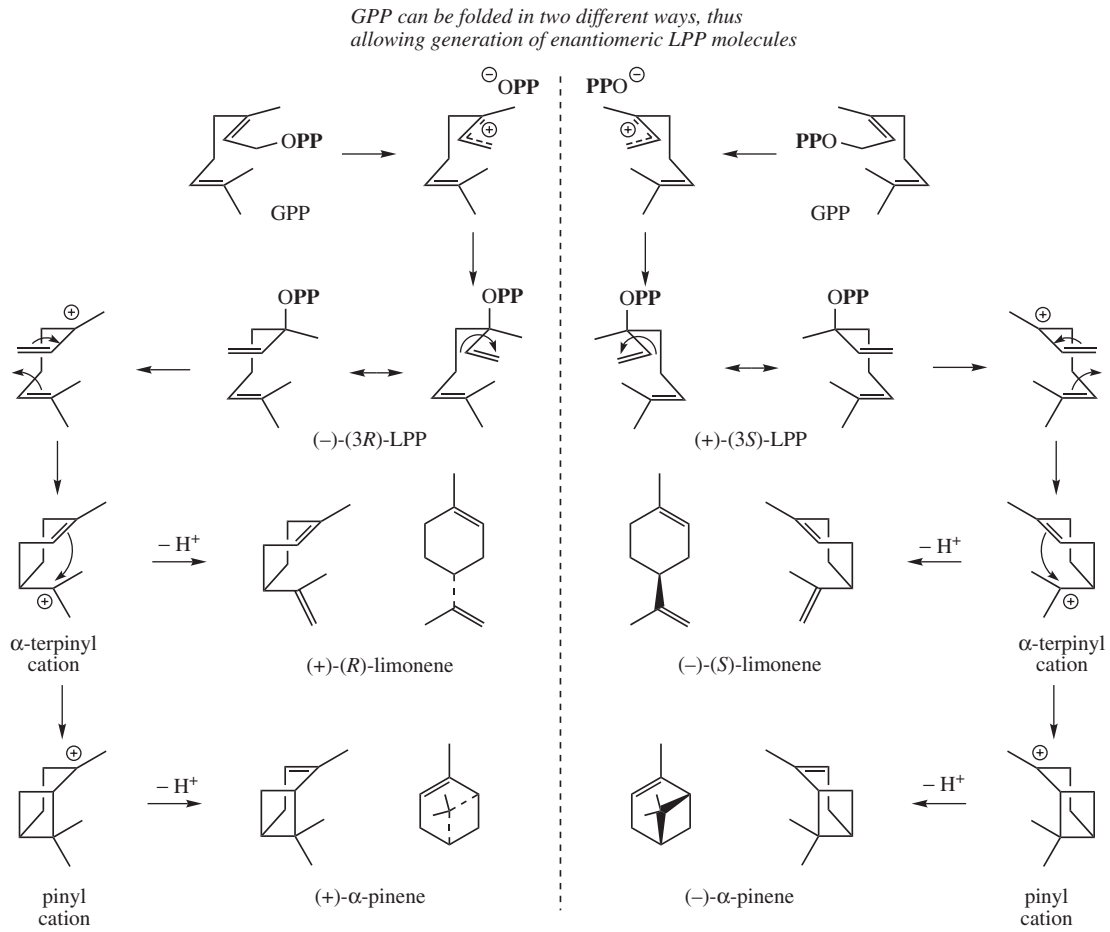


Figure 5.16

α-pinene are illustrated in Figure 5.16. This shows the precursor GPP being folded in two mirror-image conformations, leading to formation of the separate enantiomers of linalyl PP. Analogous carbocation reactions will then explain production of the optically active monoterpenes. Where a single plant produces both enantiomers, it appears to contain two separate enzyme systems each capable of elaborating a single enantiomer. Furthermore, a single enzyme typically accepts GPP as substrate, catalyses the isomerization to linalyl PP, and converts this into a final product without the release of free intermediates.

Terpenoid cyclase enzymes, even highly pure proteins obtained by gene expression, rarely convert their substrate into a single product. Sometimes, multiple products in varying amounts, e.g. limonene, myrcene, α-pinene, and β-pinene, are synthesized by a single enzyme, reflecting the common carbocation chemistry involved in these biosyntheses. This suggests that the enzyme is predominantly providing a suitable environment for the folding and cyclization of the substrate, whilst carbocation chemistry is responsible for product formation. Many of the transformations included in Figures 5.14 and 5.15 are based on analysis of these alternative products and how carbocation chemistry interrelates them. Subsequent reactions, such as oxidation of an alcohol to a ketone, e.g. borneol to **camphor** (Figure 5.14), or heterocyclic ring formation in the conversion of α-terpineol into **cineole** (Figure 5.14), require additional enzyme systems.

In other systems, a particular structure may be found as a mixture of diastereoisomers. Peppermint (*Mentha* × *piperita*; Labiatae/Lamiaceae) typically produces (−)-**menthol**, with smaller amounts of the stereoisomers (+)-**neomenthol**, (+)-**isomenthol**, and

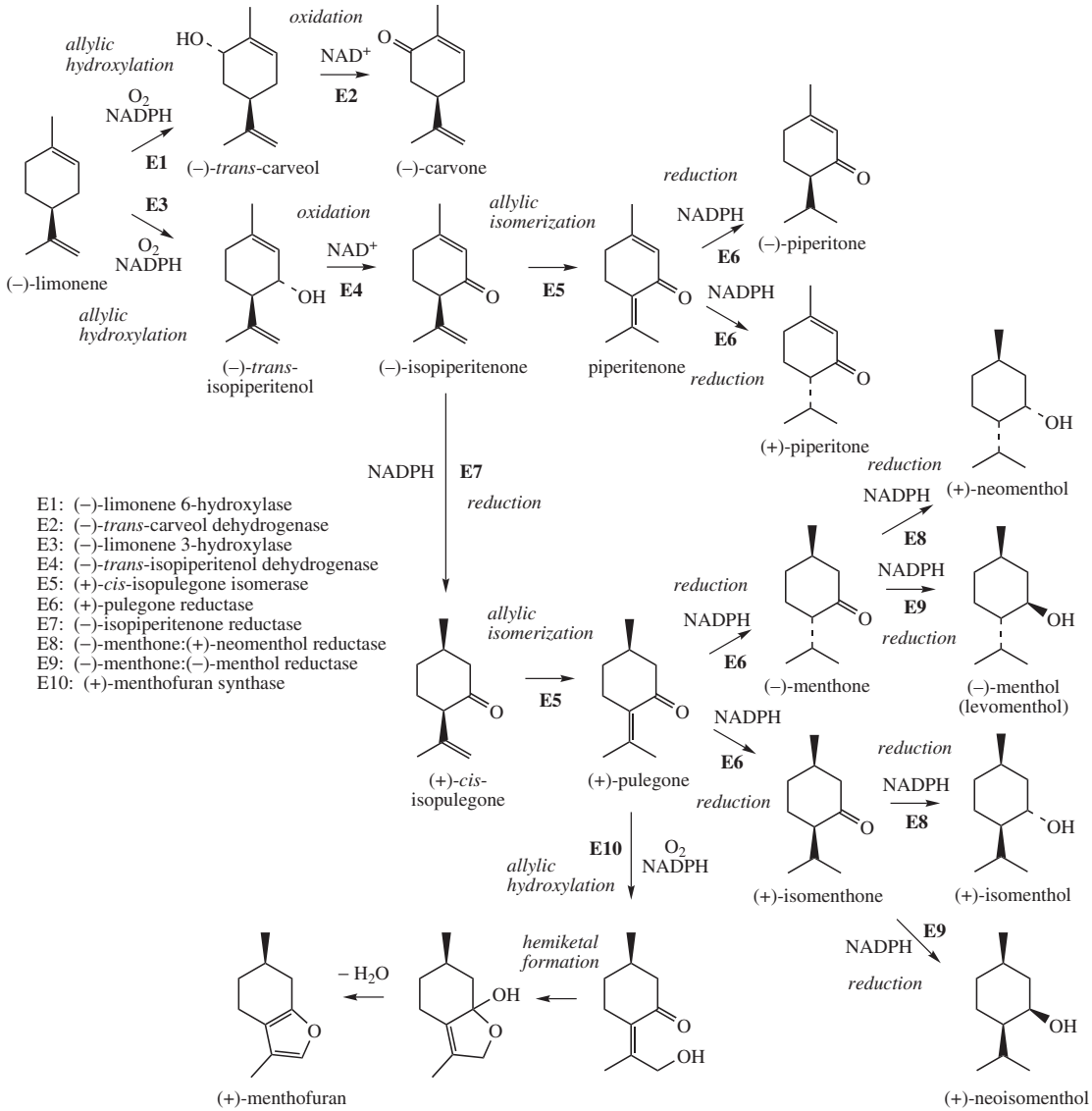


Figure 5.17

(+)-**neoisomenthol**, covering four of the possible eight stereoisomers (Figure 5.17). Oils from various *Mentha* species also contain significant amounts of ketones, e.g. (−)-**menthone**, (+)-**isomenthone**, (−)-**piperitone**, or (+)-**pulegone**. The metabolic relationship of these various compounds has been established as in Figure 5.17, which illustrates how the stereochemistry at each centre can be determined by stereospecific reduction processes on double bonds or carbonyl groups. Note that some of the enzymes involved have rather broad substrate specificity. The pathway also exemplifies that oxygen functions can be introduced into the molecule at positions activated by adjacent double bonds (allylic oxidation), as well as being introduced by quenching of carbocations with water. These reactions are catalysed by cytochrome P-450-dependent monooxygenases. Thus, limonene is a precursor of **carvone** (the main constituent of spearmint oil from *Mentha spicata*) as well as menthone and piperitone, where initial hydroxylation occurs at an alternative allylic site on the ring. **Menthofuran** exemplifies a

Table 5.1 Volatile oils (ii): containing principally terpenoid compounds

Oil	Plant source	Plant part used	Oil content (%)	Major constituents with typical (%) composition	Uses, notes
Bergamot	*Citrus aurantium* ssp. *bergamia* (Rutaceae)	fresh fruit peel (expression)	0.5	limonene (42) linalyl acetate (27) γ-terpinene (8) linalool (7)	flavouring, aromatherapy, perfumery; also contains the furocoumarin bergapten (up to 5%) and may cause severe photosensitization (see page 165)
Camphor oil	*Cinnamomum camphora* (Lauraceae)	wood	1–3	camphor (27–45) cineole (4–21) safrole (1–18)	soaps
Caraway	*Carum carvi* (Umbelliferae /Apiaceae)	ripe fruit	3–7	(+)-carvone (50–70) limonene (47)	flavour, carminative, aromatherapy
Cardamom	*Elettaria cardamomum* (Zingiberaceae)	ripe fruit	3–7	α-terpinyl acetate (25–35) cineole (25–45) linalool (5)	flavour, carminative; ingredient of curries, pickles
Chamomile (Roman chamomile)	*Chamaemelum nobile* (*Anthemis nobilis*) (Compositae/ Asteraceae)	dried flowers	0.4–1.5	aliphatic esters of angelic, tiglic, isovaleric, and isobutyric acids (75–85) small amounts of monoterpenes	flavouring, aromatherapy; blue colour of oil is due to chamazulene
Citronella	*Cymbopogon winterianus, C. nardus* (Graminae/Poaceae)	fresh leaves	0.5–1.2	(+)-citronellal (25–55) geraniol (20–40) (+)-citronellol (10–15) geranyl acetate (8)	perfumery, aromatherapy, insect repellent
Coriander	*Coriandrum sativum* (Umbelliferae /Apiaceae)	ripe fruit	0.3–1.8	(+)-linalool (60–75) γ-terpinene (5) α-pinene (5) camphor (5)	flavour, carminative
Dill	*Anethum graveolens* (Umbelliferae /Apiaceae)	ripe fruit	3–4	(+)-carvone (40–65)	flavour, carminative
Eucalyptus	*Eucalyptus globulus, E. smithii, E. polybractea* (Myrtaceae)	fresh leaves	1–3	cineole (= eucalyptol) (70–85) α-pinene (14)	flavour, antiseptic, aromatherapy
Eucalyptus (lemon-scented)	*Eucalyptus citriodora* (Myrtaceae)	fresh leaves	0.8	citronellal (65–85)	perfumery

Ginger	*Zingiber officinale* (Zingiberaceae)	dried rhizome	1.5–3	zingiberene (34) β-sesquiphellandrene (12) β-phellandrene (8) β-bisabolene (6)	flavouring; the main pungent principals (gingerols) in ginger are not volatile (see page 168)
Juniper	*Juniperus communis* (Cupressaceae)	dried ripe berries	0.5–2	α-pinene (45–80) myrcene (10–25) limonene (1–10) sabinene (0–15)	flavouring, antiseptic, diuretic, aromatherapy; juniper berries provide the flavouring for gin
Lavender	*Lavandula angustifolia, L. officinalis* (Labiatae/Lamiaceae)	fresh flowering tops	0.3–1	linalyl acetate (25–45) linalool (25–38)	perfumery, aromatherapy; inhalation produces mild sedation and facilitates sleep
Lemon	*Citrus limon* (Rutaceae)	dried fruit peel (expression)	0.1–3	(+)-limonene (60–80) β-pinene (8–12) γ-terpinene (8–10) citral (= geranial + neral) (2–3)	flavouring, perfumery, aromatherapy; terpeneless lemon oil is obtained by removing much of the terpenes under reduced pressure; this oil is more stable and contains 40–50% citral
Lemon-grass	*Cymbopogon citratus* (Graminae/Poaceae)	fresh leaves	0.1–0.3	citral (= geranial + neral) (50–85)	perfumery, aromatherapy
Matricaria (German chamomile)	*Matricaria chamomilla* (*Chamomilla recutica*) (Compositae /Asteraceae)	dried flowers	0.3–1.5	(−)-α-bisabolol (10–25) bisabolol oxides A and B (10–25) chamazulene (1–15)	flavouring; dark blue colour of oil is due to chamazulene
Orange (bitter)	*Citrus aurantium* ssp. *amara* (Rutaceae)	dried fruit peel (expression)	0.5–2.5	(+)-limonene (92–94) myrcene (2)	flavouring, aromatherapy; the main flavour and odour come from the minor oxygenated components; terpeneless orange oil is obtained by removing much of the terpenes under reduced pressure; this oil contains about 20% aldehydes, mainly decanal
Orange (sweet)	*Citrus sinensis* (Rutaceae)	dried fruit peel (expression)	0.3	(+)-limonene (90–95) myrcene (2)	flavouring, aromatherapy; the main flavour and odour come from the minor oxygenated components; terpeneless orange oil is obtained by removing much of the terpenes under reduced pressure; this oil contains about 20% aldehydes, mainly octanal and decanal

(*continued overleaf*)

Table 5.1 (*continued*)

Oil	Plant source	Plant part used	Oil content (%)	Major constituents with typical (%) composition	Uses, notes
Orange flower (Neroli)	*Citrus aurantium* ssp. *amara* (Rutaceae)	fresh flowers	0.1	linalool (36) β-pinene (16) limonene (12) linalyl acetate (6)	flavour, perfumery, aromatherapy
Peppermint	*Mentha* × *piperita* (Labiatae/Lamiaceae)	fresh leaf	1–3	menthol (30–50) menthone (15–32) menthyl acetate (2–10) menthofuran (1–9)	flavouring, carminative, aromatherapy
Pine	*Pinus palustris* or other *Pinus* species (Pinaceae)	needles, twigs		α-terpineol (65)	antiseptic, disinfectant, aromatherapy
Pumilio pine	*Pinus mugo* ssp. *pumilio* (Pinaceae)	needles	0.3–0.4	α- and β-phellandrene (60) α- and β-pinene (10–20) bornyl acetate (3–10)	inhalant; the minor components bornyl acetate and borneol are mainly responsible for the aroma
Rose (attar of rose, otto of rose)	*Rosa damascena, R. gallic*a, *R. alba*, and *R. centifolia* (Rosaceae)	fresh flowers	0.02–0.03	citronellol (36) geraniol (17) 2-phenylethanol (3) C_{14}–C_{23} straight-chain hydrocarbons (25)	perfumery, aromatherapy
Rosemary	*Rosmarinus officinalis* (Labiatae/Lamiaceae)	fresh flowering tops	1–2	cineole (15–45) α-pinene (10–25) camphor (10–25) β-pinene (8)	perfumery, aromatherapy
Sage	*Salvia officinalis* (Labiatae/Lamiaceae)	fresh flowering tops	0.7–2.5	thujone (40–60) camphor (5–22) cineole (5–14) β-caryophyllene (10) limonene (6)	aromatherapy, food flavouring
Sandalwood	*Santalum album* (Santalaceae)	heartwood	4.5–6.3	sesquiterpenes: α-santalol (50) β-santalol (21)	perfumery, aromatherapy
Spearmint	*Mentha spicata* (Labiatae/Lamiaceae)	fresh leaf	1–2	(−)-carvone (50–70) (−)-limonene (2–25)	flavouring, carminative, aromatherapy

Tea tree	*Melaleuca alternifolia* (Myrtaceae)	fresh leaf	1.8	terpinen-4-ol (30–45) γ-terpinene (10–28) α-terpinene (5–13) *p*-cymene (0.5–12) cineole (0.5–10) α-terpineol (1.5–8)	antiseptic, aromatherapy; an effective broad-spectrum antiseptic widely used in creams, cosmetics, toiletries
Thyme	*Thymus vulgaris* (Labiatae/Lamiaceae)	fresh flowering tops	0.5–2.5	thymol (40) *p*-cymene (30) linalool (7) carvacrol (1)	antiseptic, aromatherapy, food flavouring
Turpentine oil	*Pinus palustris* and other *Pinus* spp. (Pinaceae)	distillation of the resin (turpentine) secreted from bark		(+)- and (−)-α-pinene (35:65) (60–70) β-pinene (20–25)	counter-irritant, important source of industrial chemicals; the residue from distillation is colophony (rosin), composed chiefly of diterpene acids (abietic acids, see page 230)

Note: Major volatile oils have been divided into two groups. Those oils containing principally chemicals which are terpenoid in nature and which are derived by the methylerythritol phosphate or mevalonate pathways are given in this table. Oils which are composed predominantly of aromatic compounds which are derived via the shikimate pathway are listed in Table 4.1 on page 158. The remarks in the footnotes to Table 4.1 are also applicable here. Structures of terpenoids are shown in Figures 5.11–5.18.

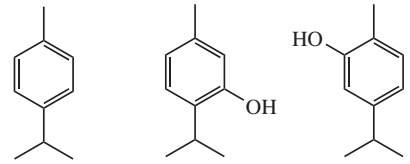


Figure 5.18

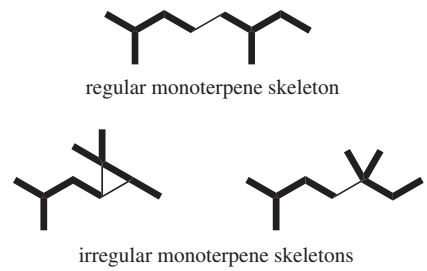


Figure 5.19

further cytochrome P-450-dependent oxidative modification that leads to generation of a heterocyclic furan ring via the hemiketal. Both pulegone and menthofuran are considered hepatotoxic. Pulegone is a major constituent of oil of pennyroyal from *Mentha pulegium*, which has a folklore history as an abortifacient. Pulegone is metabolized in humans first to menthofuran, by the same mechanism as it is in the plant pathway, and then to electrophilic metabolites that form adducts with cellular proteins (compare pyrrolizidine alkaloids, page 325). High menthofuran levels in peppermint oils are regarded as undesirable. Peppermint plants transformed with an antisense version of the menthofuran synthase gene produce less than half the normal amounts of this metabolite.

***p*-Cymene** and the phenol derivatives **thymol** and **carvacrol** (Figure 5.18), found in thyme (*Thymus vulgaris*; Labiatae/Lamiaceae), are representatives of a small group of aromatic compounds that are produced in nature from isoprene units, rather than by the much more common routes to aromatics involving acetate or shikimate (see also cannabinol, page 119, and gossypol, page 220). These compounds all possess the carbon skeleton typical of monocyclic monoterpenes, and their structural relationship to limonene and the more common oxygenated monoterpenes, such as menthone or carvone, suggests pathways in which additional dehydrogenation reactions are involved.

Data on volatile oils containing terpenoid constituents isolated from these and other plant materials are given in Table 5.1. Volatile oils in which the main components are aromatic and derived from the shikimate pathway are listed in Table 4.1, page 158.

IRREGULAR MONOTERPENES

A number of natural monoterpene structures contain carbon skeletons which, although obviously derived from two isoprene C_5 units, do not seem to fit the regular head-to-tail coupling mechanism, e.g. those in Figure 5.19. These structures are termed irregular monoterpenes and seem to be limited almost exclusively to members of the Compositae/Asteraceae plant family. Allowing for possible rearrangements, the two isoprene units appear to have coupled in another manner, and this is borne out by information available on their biosynthesis, though this is far from complete. Thus, although DMAPP and IPP are utilized in their biosynthesis, GPP and neryl PP do not appear to be involved. Pre-eminent amongst these structures are **chrysanthemic acid** and **pyrethric acid** (Figure 5.20), found in ester form as the **pyrethrins** (pyrethrins, cinerins, and jasmolins, Figure 5.20), which are valuable insecticidal components in pyrethrum flowers, the flower heads of *Chrysanthemum cinerariaefolium* (Compositae/Asteraceae) [Box 5.1]. These cyclopropane structures are readily recognizable as derived from two isoprene units, and a mechanism for the derivation of chrysanthemic acid via chrysanthemyl diphosphate is given in Figure 5.21. This invokes two DMAPP units joining by a modification of the standard mechanism, with termination achieved by cyclopropane ring formation (compare carene, page 196). This mechanism is seen again later; it is identical to that involved in the formation of presqualene PP during steroid biosynthesis (see page 235) and of prephytoene PP in carotenoid formation (see page 300). Relatively little is known about the origins of **pyrethrolone, cinerolone**, and **jasmolone** (Figure 5.20), the alcohol portions of the pyrethrins, though it is likely that these are cyclized and modified fatty acid derivatives, the cyclization process resembling the biosynthetic pathway to prostaglandins (see page 60). Thus, α-linolenic acid via 12-oxophytodienoic acid could be the precursor of jasmolone, with β-oxidation and then decarboxylation accounting for the chain shortening (Figure 5.22). Certainly, this type of pathway operates in the formation of **jasmonic acid** (Figure 5.22), which forms part of a general signalling system in plants, particularly the synthesis of secondary metabolites in response to wounding or microbial infection.

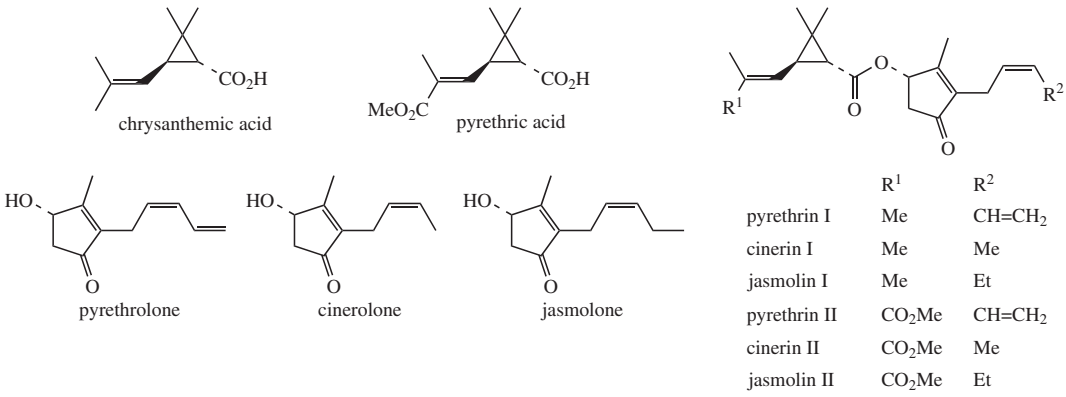


Figure 5.20

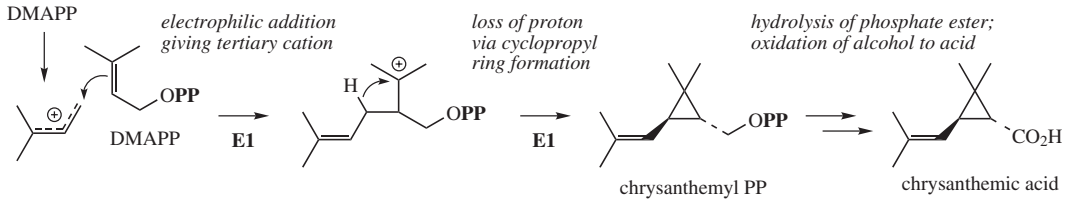


Figure 5.21

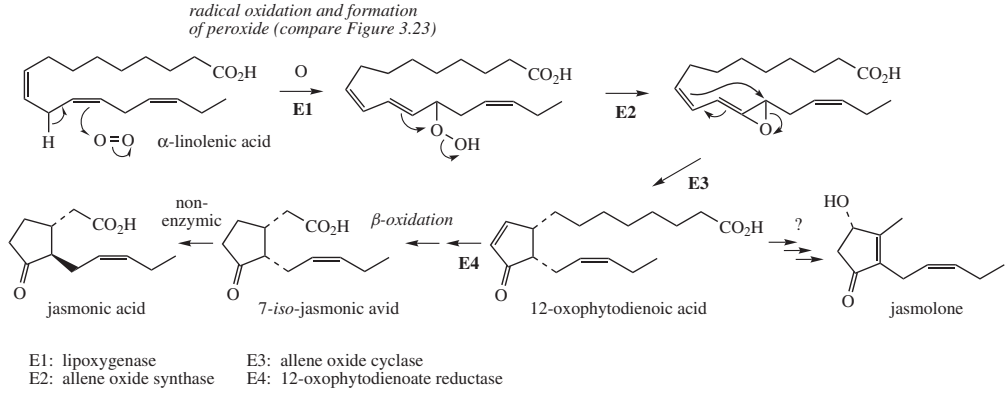


Figure 5.22

Box 5.1

Pyrethrins

The **pyrethrins** are valuable insecticidal components of pyrethrum flowers, *Chrysanthemum cinerariaefolium* (=*Tanacetum cinerariifolium*) (Compositae/Asteraceae). The flowers are harvested just before they are fully expanded, and usually processed to an extract. Pyrethrum cultivation is conducted in East Africa, especially Kenya, and more recently in Ecuador and Australia. The natural pyrethrins are used as a constituent of insect sprays for household use and as a post-harvest insecticides, having a

Box 5.1 (continued)

rapid action on the nervous system of insects, whilst being biodegradable and non-toxic to mammals, though they are toxic to fish and amphibians. This biodegradation, initiated by air and light, means few insects develop resistance to the pyrethrins, but it does limit the lifetime of the insecticide under normal conditions to just a few hours.

The flowers may contain 0.7–2% of pyrethrins, representing about 25–50% of the extract. A typical pyrethrin extract contains pyrethrin I (35%), pyrethrin II (32%), cinerin I (10%), cinerin II (14%), jasmolin I (5%), and jasmolin II (4%), which structures represent esters of chrysanthemic acid or pyrethric acid with the alcohols pyrethrolone, cinerolone, and jasmolone (Figure 5.20). Pyrethrin I is the most insecticidal component, with pyrethrin II providing much of the rapid knock-down (paralysing) effect.

A wide range of synthetic pyrethroid analogues, e.g. **bioresmethrin, tetramethrin, phenothrin, permethrin, cypermethrin**, and **deltamethrin** (Figure 5.23), have been developed which have increased lifetimes up to several days and greater toxicity towards insects. These materials have become widely used household and agricultural insecticides; the commercial insecticides are often a mixture of stereoisomers. Tetramethrin, bioresmethrin, and phenothrin are all esters of chrysanthemic acid, but with a modified alcohol portion, providing improvements in knock-down effect and in insecticidal activity. Replacement of the terminal methyl groups of chrysanthemic acid with halogen atoms, e.g. permethrin, conferred greater stability towards air and light and opened up the use of pyrethroids in agriculture. Inclusion of a cyano group in the alcohol portion, as in cypermethrin and deltamethrin, improved insecticidal activity several-fold. Modern pyrethroids now have insecticidal activities over a thousand times that of pyrethrin I, whilst maintaining extremely low mammalian toxicity. Permethrin and phenothrin are employed against skin parasites such as head lice in humans.

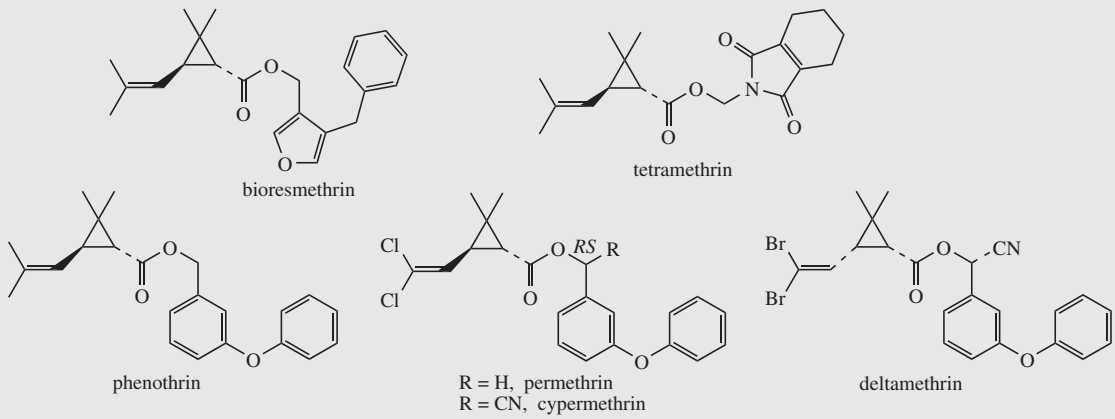


Figure 5.23

IRIDOIDS (C_{10})

The **iridane** skeleton (Figure 5.24) found in **iridoids** is monoterpenoid in origin and contains a cyclopentane ring which is usually fused to a six-membered oxygen heterocycle. The iridoid system arises from geraniol by a type of folding (Figure 5.25) which is different from that already encountered with monoterpenoids; also markedly different is the lack of phosphorylated intermediates and subsequent carbocation mechanism in its formation. The fundamental cyclization to **iridodial** is formulated as attack of hydride on the dialdehyde, produced by a series of hydroxylation and oxidation reactions on geraniol. Further oxidation gives **iridotrial**, in which hemiacetal formation then leads to production of the heterocyclic ring. In

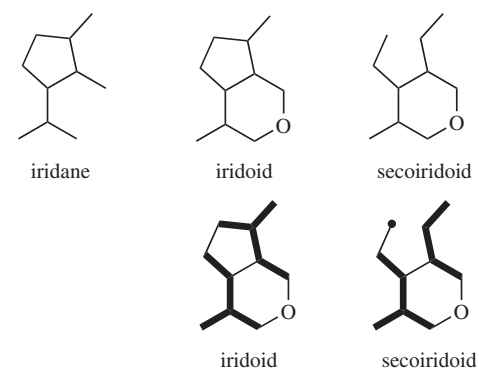


Figure 5.24

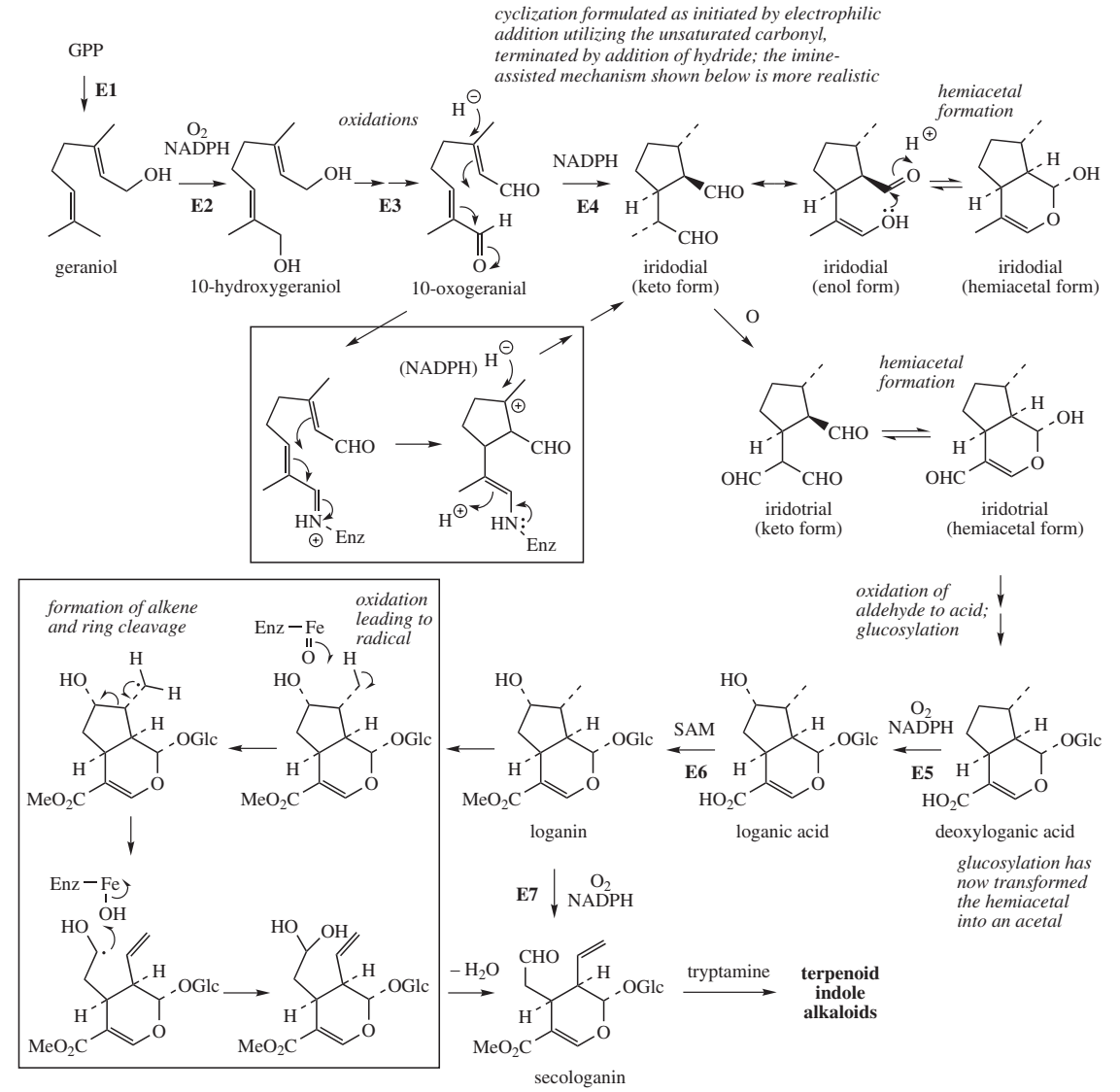


Figure 5.25

iridotrial, there is an equal chance that the original methyl groups from the head of geraniol end up as the aldehyde or in the heterocyclic ring. A large number of iridoids are found as glycosides, e.g. **loganin**; glycosylation effectively transforms the hemiacetal linkage into an acetal. The pathway to loganin involves, in addition, a sequence of reactions in which the remaining aldehyde group is oxidized to the acid and methylated, and the cyclopentane ring is hydroxylated. Loganin is a key intermediate in the pathway to a range of complex terpenoid indole alkaloids (see page 369) and tetrahydroisoquinoline alkaloids (see page 363). Fundamental in this further metabolism is cleavage of the simple monoterpene skeleton, which is still recognizable in

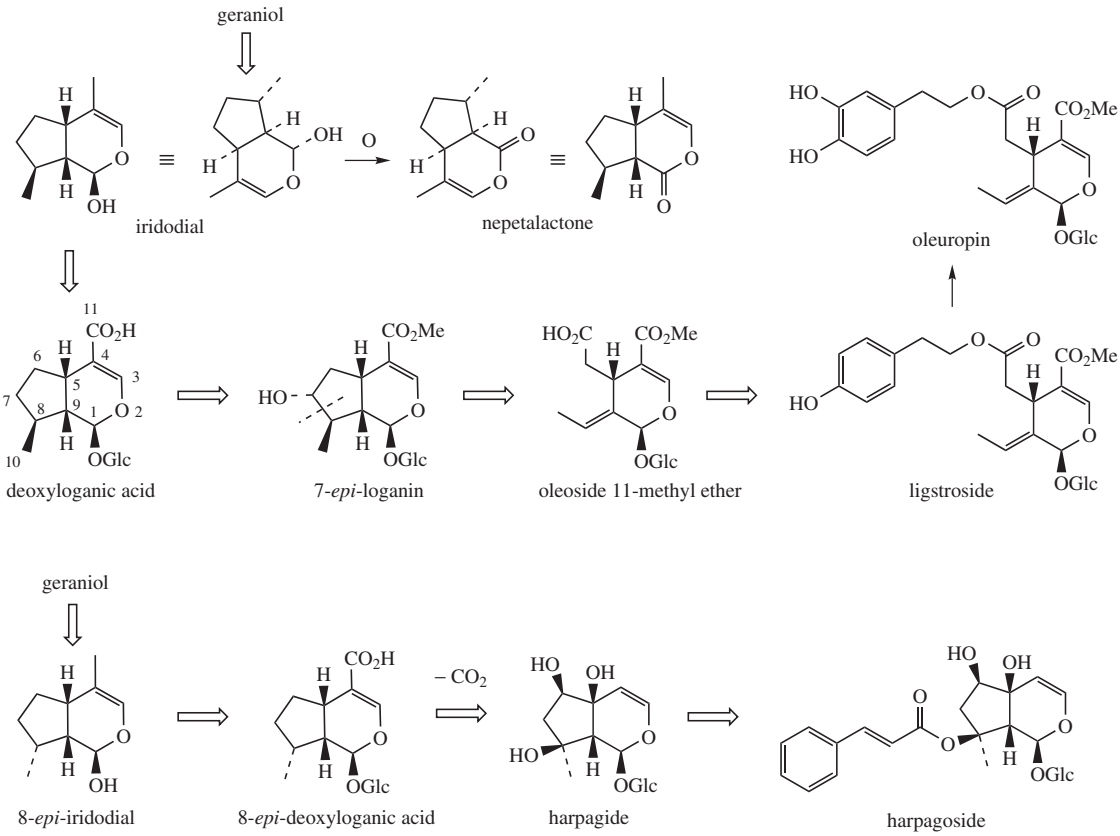


Figure 5.26

loganin, to give **secologanin**, representative of the **secoiridoids** (Figure 5.24). This is catalysed by a cytochrome P-450-dependent monooxygenase, and a radical mechanism is proposed in Figure 5.25. Secologanin now contains a free aldehyde group, together with further aldehyde and enol groups, with these latter two fixed as an acetal by the presence of the glucose. As we shall see with some of the complex alkaloids, these functionalities can be released again by hydrolysing off the glucose and reopening the hemiacetal linkage.

Well over a thousand different natural iridoids and secoiridoids are known. Structural variation arises predominantly from hydroxylations, esterifications, and changes in stereochemistry. Another major change is the loss of a carbon atom by a decarboxylation mechanism. A few examples showing their relationship to the intermediates of Figure 5.25 are presented in Figure 5.26.

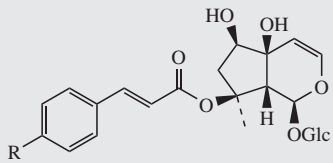
Nepetalactone from catmint *Nepeta cataria* (Labiatae/Lamiaceae), a powerful attractant and stimulant for cats, is produced from iridodial by simple oxidation of the hemiacetal to a lactone (Figure 5.26). **Ligstroside** and **oleuropin** are phenolic esters found in olive oil (see page 47) from *Olea europea* (Oleaceae), and are believed to contribute to the health benefits of this oil. These compounds are secoiridoids, but are formed from the 7-epimer of loganin; in this case, the stereochemistry of 7-hydroxylation is different from the normal loganin pathway. **Harpagide** and **harpagoside** are examples of decarboxylated iridoids; these particular examples are formed via 8-*epi*-iridodial, in which geraniol is cyclized by a stereochemically different mechanism. Harpagoside is the cinnamoyl ester of harpagide. These compounds are found in devil's claw (*Harpagophytum procumbens*; Pedaliaceae) and contribute to the anti-inflammatory and analgesic properties associated with this plant drug [Box 5.2].

Box 5.2

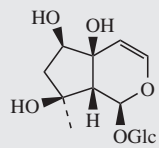
Devil's Claw

Devil's claw is the common name for *Harpagophytum procumbens* (Pedaliaceae) from the appearance of its fruit, which have curved, sharp hooks. The plant is a weedy, perennial, tuberous plant with long creeping stems found in southern Africa (South Africa, Namibia, Botswana); commercial material is collected from the wild, but is often a mixture from two species: *H. procumbens* and *H. zeyheri*. Preparations of the secondary roots have gained a reputation as an anti-inflammatory and antirheumatic agent to relieve pain and inflammation in people with arthritis and similar disorders. Clinical studies appear to support its medicinal value as an anti-inflammatory and analgesic, though some findings are less positive.

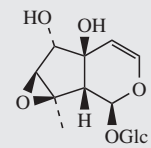
The main constituents of devil's claw root are a group of decarboxylated iridoid glycosides (about 3%), including harpagoside (at least 1.2%) as the main component and smaller amounts of procumbide, harpagide, and 8-(4-coumaroyl)harpagide (Figure 5.27). The latter compound appears representative of *H. zeyheri* only, so indicates if this species is present in the sample. The secondary roots contain significantly higher levels of iridoids than the primary tubers. Harpagoside and related iridoids have been shown to inhibit thromboxane biosynthesis (see page 65), which may relate to the observed anti-inflammatory activity of devil's claw.



R = H, harpagoside
R = OH, 8-(4-coumaroyl)harpagide



harpagide



procumbide

Figure 5.27

A range of epoxyiridoid esters has been identified in the drug valerian (*Valeriana officinalis*; Valerianaceae) [Box 5.3]. These materials, responsible for the sedative activity of the crude drug, are termed **valepotriates. Valtrate** (Figure 5.28) is a typical example, and illustrates the structural relationship to loganin, though these compounds contain additional ester functions, frequently isovaleryl. The hemiacetal is now fixed as an ester, rather than as a glycoside.

Box 5.3

Valerian

Valerian root consists of the dried underground parts of *Valeriana officinalis* (Valerianaceae), a perennial herb found throughout Europe. Drug material comes from wild and cultivated plants, and is carefully dried at low temperature (less than 40°C) to minimize decomposition of constituents. Valerian preparations are widely used as herbal tranquillizers to relieve nervous tension, anxiety, and as a mild sedative to promote sleep; the drug was especially popular during the First World War, when it was used to treat shell-shock. The drug does possess mild sedative and anxiolytic properties, but the roots need to be freshly harvested and carefully dried for maximum activity. The major active principles are generally held to be a number of epoxyiridoid esters called valepotriates (0.5–1.6%), the principal component of which is valtrate (about 80%) (Figure 5.28). Minor valepotriates have the same parent iridoid alcohol as valtrate, but differ with respect to esterifying acids, e.g. isovaltrate and acevaltrate (Figure 5.28), or are based on the reduced iridoid seen in didrovaltrate, again with various ester functionalities. Acid entities characterized in this group of compounds are mainly isovaleric (3-methylbutyric) and acetic (as in valtrate/isovaltrate/didrovaltrate), though more complex diester groups involving 3-acetoxyisovaleric and isovaleroxyisovaleric acids are encountered. During drying and storage, some of the valepotriate content may decompose by hydrolysis to liberate quantities of isovaleric acid, giving a characteristic odour, and structures such as baldrinal (Figure 5.28) (from valtrate) and homobaldrinal (from isovaltrate). Samples of old or poorly prepared valerian may contain negligible amounts of valepotriates. Standardized mixtures of valepotriates are available in some countries. These materials are usually extracted from the roots of other species of *Valeriana* which produce higher amounts

Box 5.3 (continued)

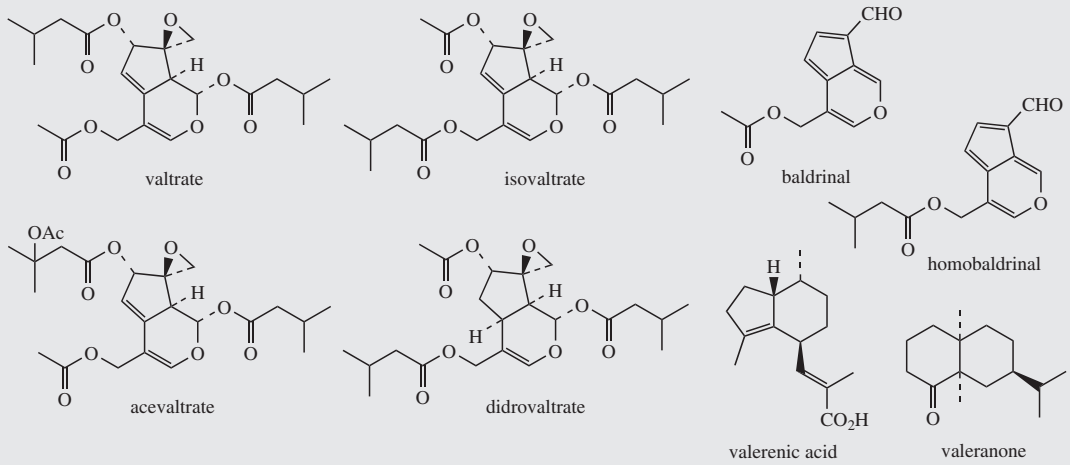


Figure 5.28

of valepotriates than *V. officinalis*, e.g. *V. mexicana* contains up to about 8%. Some other species of *Valeriana* which contain similar valepotriate constituents are used medicinally, including *V. wallichi* (Indian valerian) and *V. edulis* (Mexican valerian).

Despite the information given above, many workers believe the sedative activity of valerian cannot be due to the valepotriates, which are rather unstable and not water soluble. Some of the sedative activity is said to arise from sesquiterpene derivatives such as valerenic acid (about 0.3%) and those found in the volatile oil content (0.5–1.3%), e.g. valeranone (Figure 5.28), which have been shown to be physiologically active. GABA and glutamine have also been identified in aqueous extracts of valerian, and these have been suggested to contribute to the sedative properties. The valepotriates valtrate and didrovaltrate are reported to be cytotoxic *in vitro*, and this may restrict future use of valerian. The reactive epoxide group is likely to be responsible for these cytotoxic properties.

SESQUITERPENES (C_{15})

Sesquiterpenes are formed from three C_5 units; the terminology comes from the Latin prefix sesqui: 'one and a half times'. Though the mechanisms involved in their formation closely parallel those seen for the monoterpenes, sesquiterpenes are generally synthesized via the mevalonate pathway (see page 192) rather than from MEP. Addition of a further C_5 IPP unit to GPP in an extension of the GPP synthase reaction (Figure 5.9) leads to the fundamental sesquiterpene precursor **farnesyl diphosphate (FPP)** (Figure 5.29). Again, an initial ionization of GPP seems likely, and the proton lost from C-2 of IPP is stereochemically analogous to that lost in the previous isoprenylation step.

FPP can then give rise to linear and cyclic sesquiterpenes. Because of the increased chain length and an additional double bond, the number of possible cyclization modes is increased, and a huge range of mono-, bi-, and tri-cyclic structures can result; the number of known natural sesquiterpenes greatly exceeds that of known natural

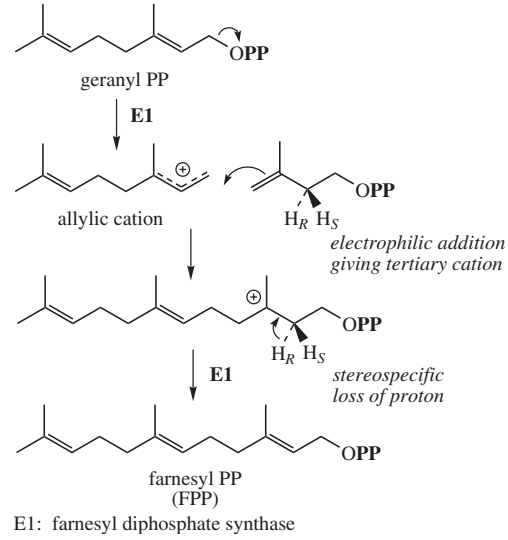


Figure 5.29

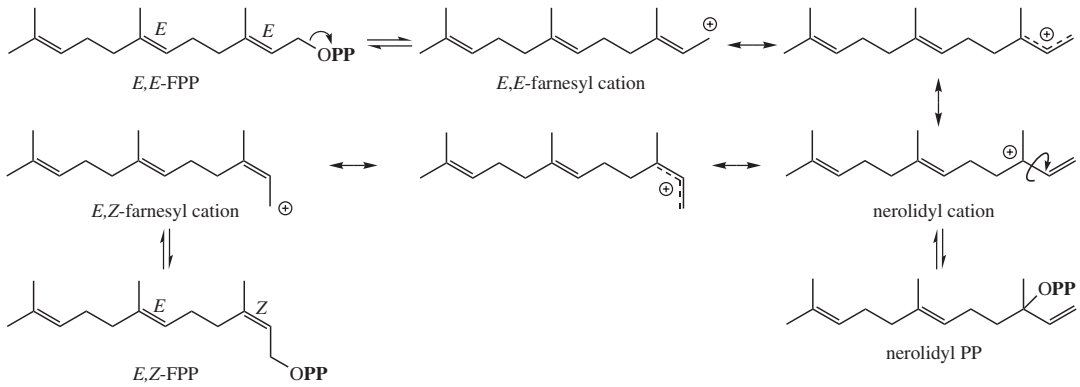


Figure 5.30

monoterpenes. The stereochemistry of the double bond nearest the diphosphate can adopt an *E* configuration (as in FPP) or a *Z* configuration via ionization as found with GPP/neryl PP (Figure 5.30). In some systems, the tertiary diphosphate **nerolidyl PP** (compare linalyl PP, page 194) has been implicated as a more immediate precursor than FPP (Figure 5.30). This allows different possibilities for folding the carbon chain, dictated of course by the enzyme involved, and cyclization by electrophilic attack onto an appropriate double bond.

As with the monoterpenes, standard reactions of carbocations rationally explain most of the common structural skeletons encountered, and a small but representative selection of these is given in Figure 5.31; sesquiterpene cyclase enzymes typically synthesize a major product accompanied by a range of related structures. One of these cyclized systems, the bisabolyl cation, is analogous to the monoterpene menthane system, and further modifications in the six-membered ring can take place to give essentially monoterpene variants with an extended hydrocarbon substituent, e.g. **γ-bisabolene** (Figure 5.32), which contributes to the aroma of ginger (*Zingiber officinale*; Zingiberaceae) along with obviously related structures such as **zingiberene** and **β-sesquiphellandrene** (Figure 5.33). Sesquiterpenes will, in general, be less volatile than monoterpenes. Simple quenching of the bisabolyl cation with water leads to **α-bisabolol** (Figure 5.32), a major component of matricaria (German chamomile) flowers (*Matricaria chamomilla*; Compositae/Asteraceae) [Box 5.5]. So-called **bisabolol oxides** A and B are also present, compounds probably derived from bisabolol by cyclization reactions (Figure 5.32) on an intermediate epoxide (compare Figure 4.34, page 164).

Other cyclizations in Figure 5.31 lead to ring systems larger than six carbon atoms, and 7-, 10-, and 11-membered rings can be formed as shown. The two 10-membered ring systems (germacryl and *cis*-germacryl cations), or the two 11-membered systems (humulyl and *cis*-humulyl cations), differ only in the stereochemistry associated with the double bonds. However, this affects further cyclization processes and is responsible for extending the variety of natural sesquiterpene derivatives. The germacryl cation, without further cyclization, is a precursor of the germacrane class of sesquiterpenes, as exemplified by **costunolide** (Figure 5.34), a bitter principle found in the roots of chicory (*Cichorium intybus*; Compositae/Asteraceae). Costunolide is actually classified as a germacranolide, the suffix 'olide' referring to the lactone group (compare macrolide, page 68). The antimigraine agent in feverfew (*Tanacetum parthenium*; Compositae/Asteraceae) is **parthenolide** [Box 5.4], an epoxide derivative of costunolide (Figure 5.34). At present it is not known whether epoxidation of the double bond in costunolide is involved, or whether an earlier intermediate in the pathway is oxidized.

The α,β-unsaturated carbonyl functionality seen in costunolide and parthenolide is a common feature of many of the biologically active terpenoids. The activity frequently manifests itself as a toxicity, especially cytotoxicity, as seen with the germacranolide **elephantopin** (Figure 5.33) from *Elephantopus elatus* (Compositae/Asteraceae), or skin allergies, as caused by the pseudoguaianolide (a rearranged guaianolide) **parthenin** (Figure 5.33) from *Parthenium hysterophorus* (Compositae/Asteraceae), a highly troublesome weed in India. These compounds can be considered as powerful alkylating agents by a Michael-type addition of a suitable nucleophile, e.g. thiols, onto the α,β-unsaturated carbonyl system. Such alkylation reactions are believed to explain

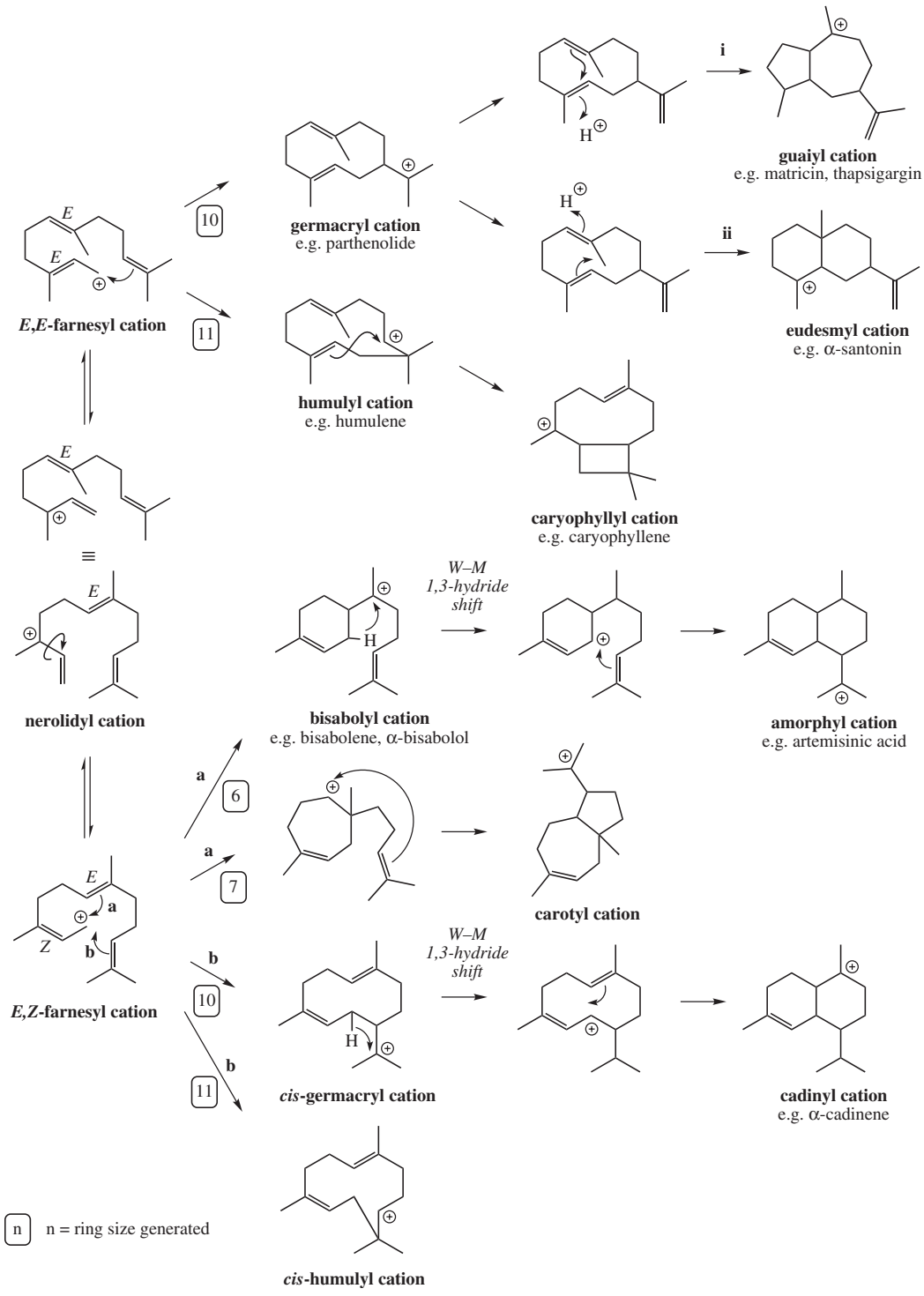

E
E
E,E-farnesyl cation
10
11
germacryl cation
e.g. parthenolide
humulyl cation
e.g. humulene
i
guaiyl cation
e.g. matricin, thapsigargin
ii
eudesmyl cation
e.g. α-santonin
caryophyllyl cation
e.g. caryophyllene
nerolidyl cation
W–M
1,3-hydride
shift
bisabolyl cation
e.g. bisabolene, α-bisabolol
amorphyl cation
e.g. artemisinic acid
a
6
a
7
carotyl cation
E,Z-farnesyl cation
b
10
b
11
cis-germacryl cation
cadinyl cation
e.g. α-cadinene
cis-humulyl cation
n = ring size generated

Figure 5.31

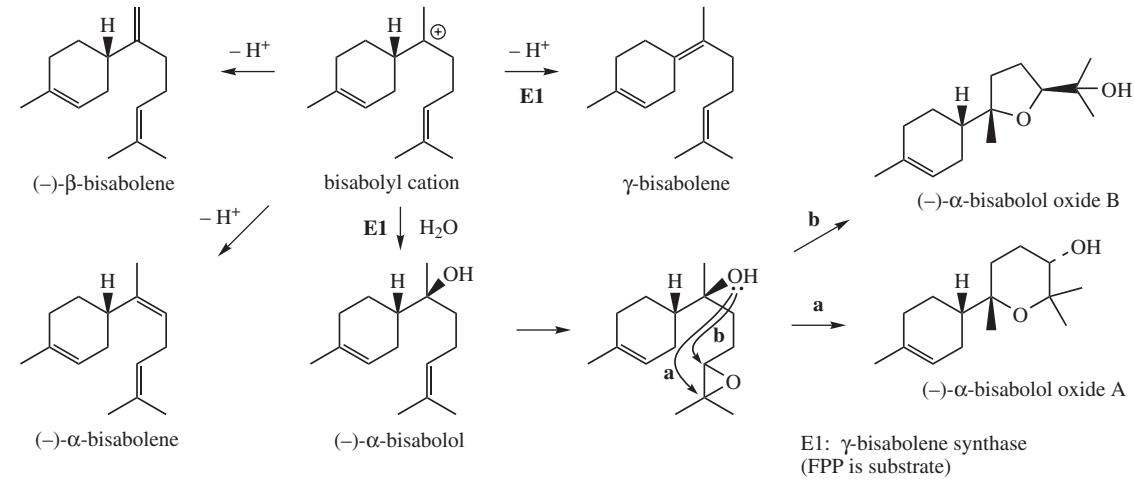


Figure 5.32

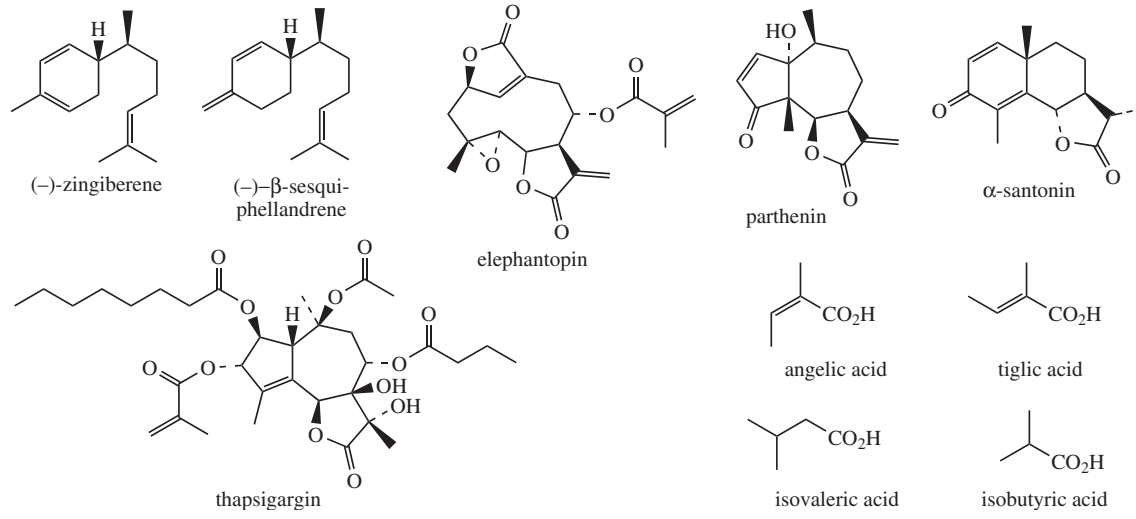


Figure 5.33

biological activity; indeed, activity is typically lost if either the double bond or the carbonyl group is chemically reduced. In some structures, additional electrophilic centres offer further scope for alkylation reactions. In **parthenolide** (Figure 5.35), an electrophilic epoxide group is also present, allowing transannular cyclization and generation of a second alkylation site. Cytotoxic agents may irreversibly alkylate critical enzymes that control cell division, whilst allergenic compounds may conjugate with proteins to form antigens which trigger the allergic response. The beneficial effects of parthenolide and structurally related compounds in feverfew have been demonstrated to relate to alkylation of thiol groups.

Box 5.4

Feverfew

Feverfew is a traditional herbal remedy for the relief of arthritis, migraine, toothache, and menstrual difficulties. The plant is a perennial, strongly aromatic herb of the Compositae/Asteraceae family, and has been classified variously as *Tanacetum parthenium* (which is currently favoured), *Chrysanthemum parthenium, Leucanthemum parthenium*, or *Pyrethrum parthenium*. Studies have confirmed that feverfew is an effective prophylactic treatment for migraine in about 70% of sufferers. It reduces the frequency and severity of attacks and the vomiting associated with them. The herb has been shown to inhibit blood platelet aggregation, the release of 5-hydroxytryptamine (5-HT, serotonin) from platelets, the release of histamine from mast cells, and the production of prostaglandins, thromboxanes, and leukotrienes. Of a range of sesquiterpene lactones of the germacrane and guianane groups characterized in the leaf material, the principal constituent and major active component is parthenolide (Figure 5.34) (up to about 1% in dried leaves). The powerful pungent odour of the plant arises from the volatile oil constituents, of which the monoterpene camphor (Figure 5.14) is a major constituent. Feverfew may be taken as the fresh leaf, often eaten with bread in the form of a sandwich to minimize the bitter taste, or it can be obtained in commercial dosage forms as tablets or capsules of the dried powdered leaf. The parthenolide content of dried leaf deteriorates on storage, and many commercial preparations of feverfew have been shown to contain little parthenolide, or to be well below the stated content. This may be a consequence of complexation with plant thiols via Michael addition. Consumers of fresh leaf can be troubled by sore mouth or mouth ulcers, caused by the sesquiterpenes. Parthenolide is also known to be capable of causing some allergic effects, e.g. contact dermatitis. The proposed mechanism of action of parthenolide via alkylation of thiol groups in proteins is shown in Figure 5.35.

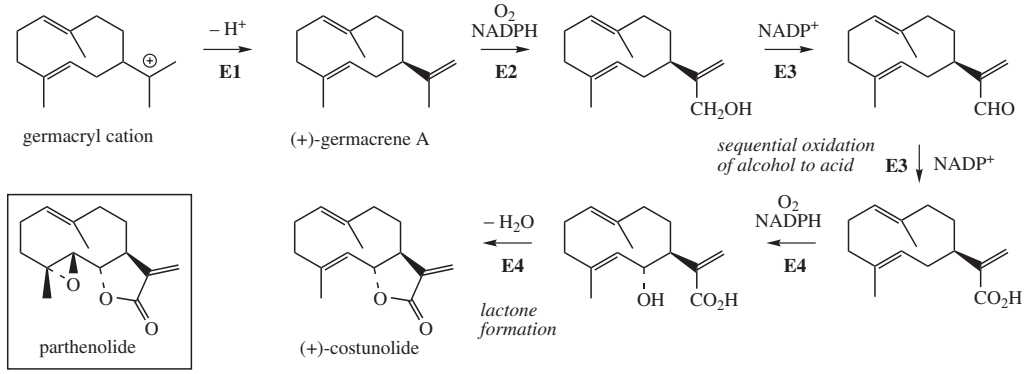


Figure 5.34

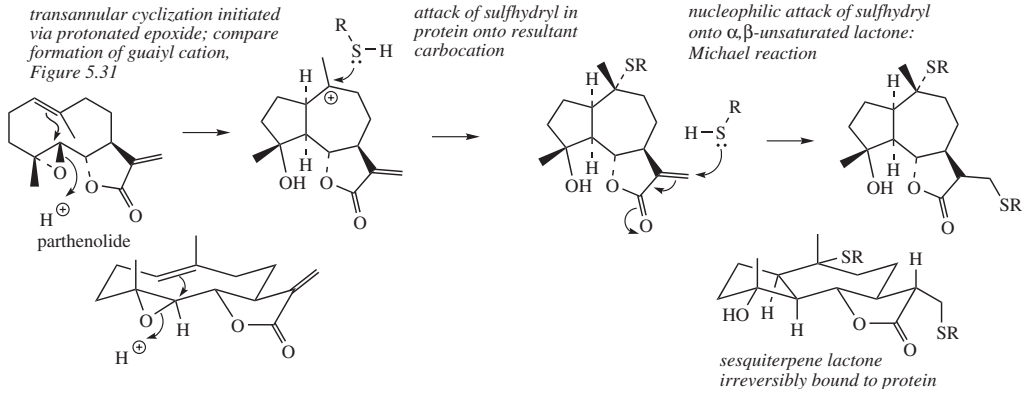


Figure 5.35

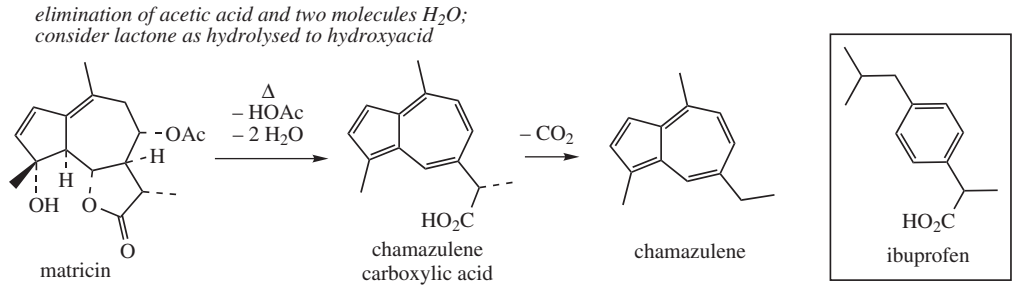


Figure 5.36

α-Santonin (Figure 5.33) has been identified as the principal anthelmintic component of various *Artemisia* species, e.g. wormseed (*A. cinia*; Compositae/Asteraceae), and has found considerable use for removal of roundworms, although potential toxicity limits its application. Structurally, α-santonin bears much similarity to parthenolide, and the most marked difference lies in the presence of the bicyclic decalin ring system. This basic skeleton, the eudesmane system, is formed from the germacryl cation by protonation and cyclization via the eudesmyl cation (Figure 5.31, route ii), whereas protonation at the more substituted end of a double bond (anti-Markovnikov addition, route i), could generate the guaiyl cation and guaiane skeleton. This latter skeleton is found in **matricin** (Figure 5.36), again from matricaria flowers [Box 5.5]. This compound degrades on heating, via hydrolysis then elimination of acetic acid and water, followed by decarboxylation to the azulene derivative **chamazulene**, which is responsible for the blue coloration of oil distilled from the flowers. Chamazulene carboxylic acid (Figure 5.36) is also found in yarrow (*Achillea millefolium*; Compositae/Asteraceae) and has been shown to be an anti-inflammatory agent by inhibiting cyclooxygenase COX-2, though not COX-1 (see page 62). The structural resemblance to the synthetic analgesic ibuprofen is striking. **Thapsigargin** (Figure 5.33) from seeds and roots of the Mediterranean plant *Thapsia garganica* (Compositae/Asteraceae) provides a further example of a guaianolide, highly oxygenated and esterified with a variety of acid groups. This compound is of considerable pharmacological interest in the study of Ca^{2+} signalling pathways. Thapsigargin is a potent inhibitor of Ca^{2+}-ATPases, and is capable of severely unbalancing cellular Ca^{2+} concentrations, often leading to disrupted cell function and growth, and apoptosis. A thapsigargin pro-drug has been tested in the treatment of prostate cancer.

Box 5.5

Chamomile and Matricaria

Two types of **chamomile** (**camomile**) are commonly employed in herbal medicine: Roman chamomile *Chamaemelum nobile* (formerly *Anthemis nobilis*) (Compositae/Asteraceae) and German chamomile *Matricaria chamomilla* (*Chamomilla recutica*) (Compositae/Asteraceae). German chamomile, an annual plant, is the more important commercially and is often called **matricaria** to distinguish it from the perennial Roman chamomile. Both plants are cultivated in various European countries to produce the flower-heads which are then dried for drug use. Volatile oils obtained by steam distillation or solvent extraction are also available.

Matricaria is also used as a digestive aid, but is mainly employed for its anti-inflammatory and spasmolytic properties. Extracts or the volatile oil find use in creams and ointments to treat inflammatory skin conditions, and as an antibacterial and antifungal agent. Taken internally, matricaria may help in the control of gastric ulcers. The flowers yield 0.5–1.5% volatile oil containing the sesquiterpenes α-bisabolol (10–25%), bisabolol oxides A and B (10–25%) (Figure 5.32), and chamazulene (0–15%) (Figure 5.36). Chamazulene is a thermal decomposition product from matricin, and is responsible for the dark blue coloration of the oil (Roman chamomile oil contains only trace amounts of chamazulene). α-Bisabolol has some anti-inflammatory, antibacterial, and ulcer-protective properties, but chamazulene probably contributes to the anti-inflammatory activity of matricaria preparations. It has been found to block the cyclooxygenase enzyme COX-2, though not COX-1, in prostaglandin biosynthesis (see page 62); the anti-inflammatory activity may result from the subsequent inhibition of leukotriene formation. Matricin itself produces rather greater anti-inflammatory activity than chamazulene and appears to be metabolized in the body to chamazulene carboxylic acid, a natural analogue of the synthetic analgesic ibuprofen (Figure 5.36).

Box 5.5 (continued)

Roman chamomile is usually taken as an aqueous infusion (chamomile tea) to aid digestion, curb flatulence, etc., but extracts also feature in mouthwashes, shampoos, and many pharmaceutical preparations. It has mild antiseptic and anti-inflammatory properties. The flower-heads yield 0.4–1.5% of volatile oil, which contains over 75% of aliphatic esters of angelic, tiglic, isovaleric, and isobutyric acids (Figure 5.33), products from metabolism of the amino acids isoleucine, leucine, and valine (see pages 56, 79, 316), with small amounts of monoterpenes and sesquiterpenes.

The 11-carbon ring of the humulyl carbocation (Figure 5.31) may be retained, as in the formation of **α-humulene** (Figure 5.37), or modified to give the caryophyllyl cation containing a nine-membered ring fused to a four-membered ring, as in **β-caryophyllene** (Figure 5.37). Humulene is found in hops (*Humulus lupulus*; Cannabaceae), and β-caryophyllene is found in a number of plants, e.g. in the oils from cloves (*Syzygium aromaticum*; Myrtaceae) and cinnamon (*Cinnamomum zeylanicum*; Lauraceae).

Another type of decalin-containing sesquiterpene is seen in the structures of cadinenes and amorpha-4,11-diene. **α-Cadinene** (Figure 5.38) is one of the many terpenoids found in juniper berries (*Juniperus communis*; Cupressaceae) used in making gin, and this compound is derived from the 10-carbon ring-containing *cis*-germacryl cation. The double bonds in the *cis*-germacryl cation are unfavourably placed for a cyclization reaction, as observed with the germacryl cation, and available evidence points to an initial 1,3-shift of hydride to the

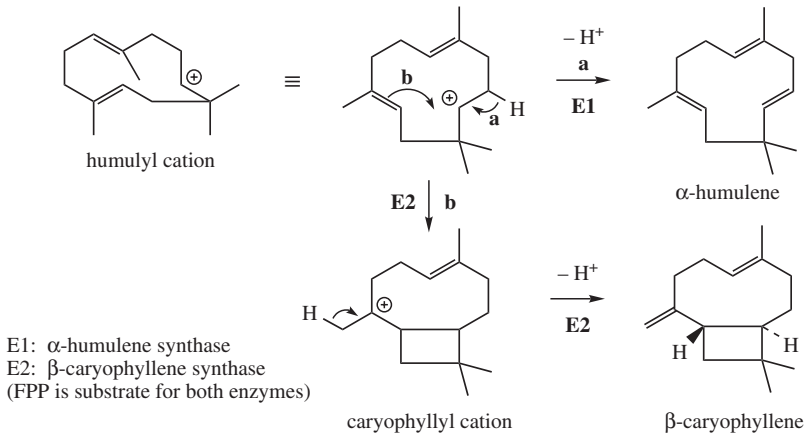


Figure 5.37

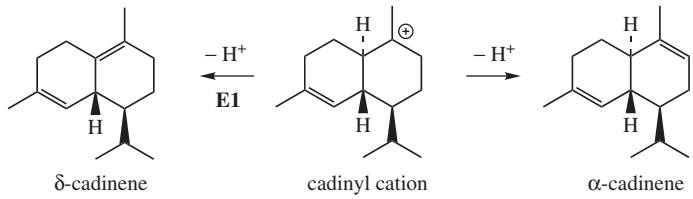


Figure 5.38

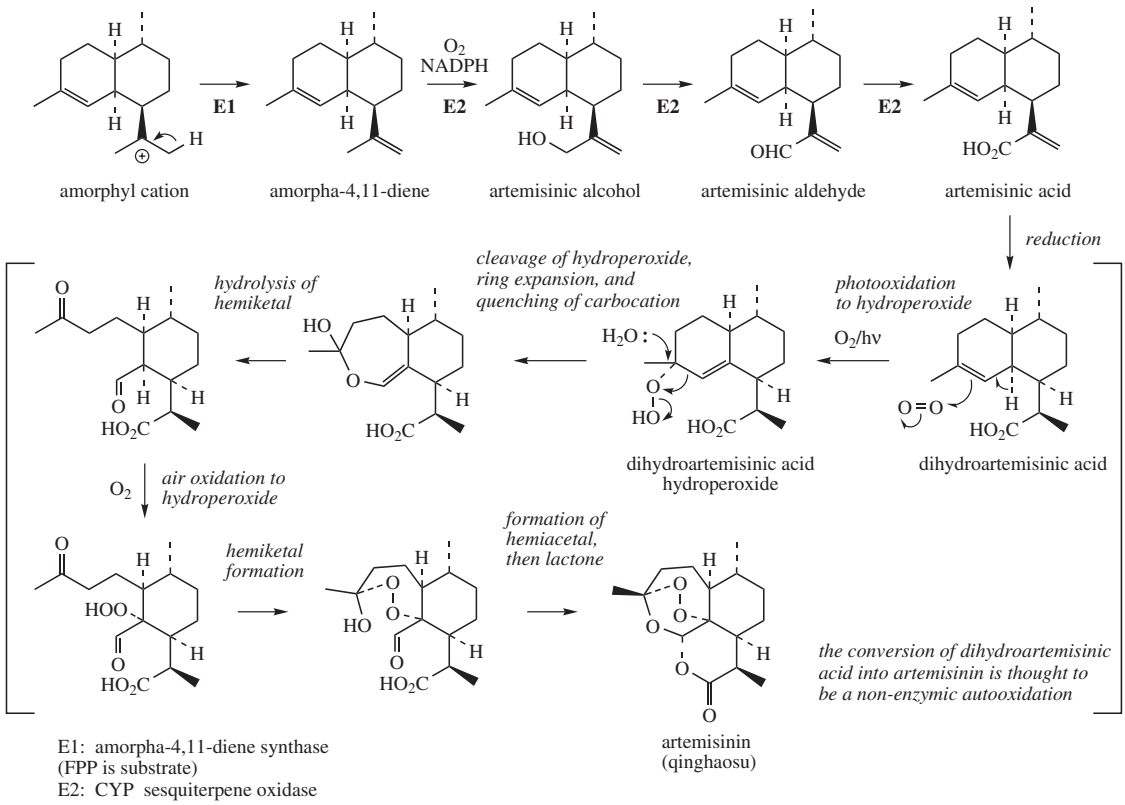


Figure 5.39

isopropyl side-chain to generate a new cation, and thus allowing cyclization to the cadinyl cation (Figure 5.31). **δ-Cadinene** is an alternative deprotonation product; it is further elaborated to gossypol in cotton (see page 220).

Amorpha-4,11-diene (Figure 5.39) is structurally related to the cadinenes, but the different stereochemistry of ring fusion and the position of the second double bond is a consequence of a different cyclization mechanism operating to produce the initial decalin ring system. In this case, a six-membered ring is formed first to give the bisabolyl cation; again, a 1,3-hydride shift is implicated prior to forming the decalin system of the amorphyl cation (Figure 5.31).

Amorpha-4,11-diene is an intermediate in the pathway leading to **artemisinin** in *Artemisia annua* (Compositae/Asteraceae) (Figure 5.39). **Artemisinic acid** is formed from amorphadiene via modest oxidation processes all catalysed by a single cytochrome P-450-dependent enzyme system. Reduction to **dihydroartemisinic acid** is followed by transformation into artemisinin by a sequence of reactions that is currently regarded as non-enzymic, and brought about by an oxygen-mediated photochemical oxidation under conditions that might normally be present in the plant. An intermediate in this process, also found naturally in *A. annua*, is the hydroperoxide of dihydroartemisinic acid. The further modifications postulated in Figure 5.39 include ring expansion by cleavage of this hydroperoxide and a second oxygen-mediated hydroperoxidation. The 1,2,4-trioxane system in artemisinin can be viewed more simply as a combination of hemiketal, hemiacetal, and lactone functions, and the later stages of the pathway merely reflect their construction. Artemisinin is an important antimalarial component in *Artemisia annua*, a Chinese herbal drug [Box 5.6]. There is currently strong research effort to produce artemisinin or analogues as new antimalarial drugs, since the malarial parasite has developed resistance to many of the current drugs (see quinine, page 382).

Box 5.6

Artemisia annua and Artemisinin

Artemisia annua (Compositae/Asteraceae) is known as qinghao in Chinese traditional medicine, where it has been used for centuries in the treatment of fevers and malaria. The plant is sometimes called annual or sweet wormwood, and is quite widespread, being found in Europe, North and South America, as well as China. Artemisinin (qinghaosu; Figure 5.40) was subsequently extracted and shown to be responsible for the antimalarial properties, being an effective blood schizontocide in humans infected with malaria, and showing virtually no toxicity. Malaria is caused by protozoa of the genus *Plasmodium*, especially *P. falciparum*, entering the blood system from the salivary glands of mosquitoes, and worldwide is responsible for 2–3 million deaths each year. Established antimalarial drugs, such as chloroquine (see page 382), are proving less effective in the treatment of malaria due to the appearance of drug-resistant strains of *P. falciparum*. Artemisinin is currently effective against these drug-resistant strains.

Artemisinin is a sesquiterpene lactone containing a rare peroxide linkage which appears to be essential for activity. Some plants of *A. annua* have been found to produce as much as 1.4% artemisinin, but the yield is normally very much less, typically 0.05–0.2%. Apart from one or two low-yielding species, the compound has not been found in any other species of the genus *Artemisia* (about 400 species). *A. annua* is grown for drug use in China, Vietnam, East Africa, the United States, Russia, India, and Brazil. Small amounts (about 0.01%) of the related peroxide structure artemisitene (Figure 5.40) are also present in *A. annua*, though this has a lower antimalarial activity. The most abundant sesquiterpenes in the plant are artemisinic acid (arteannuic acid, qinghao acid; typically 0.2–0.8%) (Figure 5.39), and lesser amounts (0.1%) of arteannuin B (qinghaosu-II; Figure 5.40). Fortunately, the artemisinic acid content may be converted chemically into artemisinin by a relatively straightforward and efficient process. Artemisinin may also be reduced to the lactol (hemiacetal) dihydroartemisinin (Figure 5.40), and this has been used for the semi-synthesis of a range of analogues, of which the acetals artemether and arteether (Figure 5.40), and the water-soluble sodium salts of artelinic acid and artesunic acid (Figure 5.40) appear very promising antimalarial agents. These materials have increased activity compared with artemisinin, and the chances of infection recurring are also reduced. **Artemether** has rapid action against chloroquinine-resistant *P. falciparum* malaria, and is currently being used in combination with the

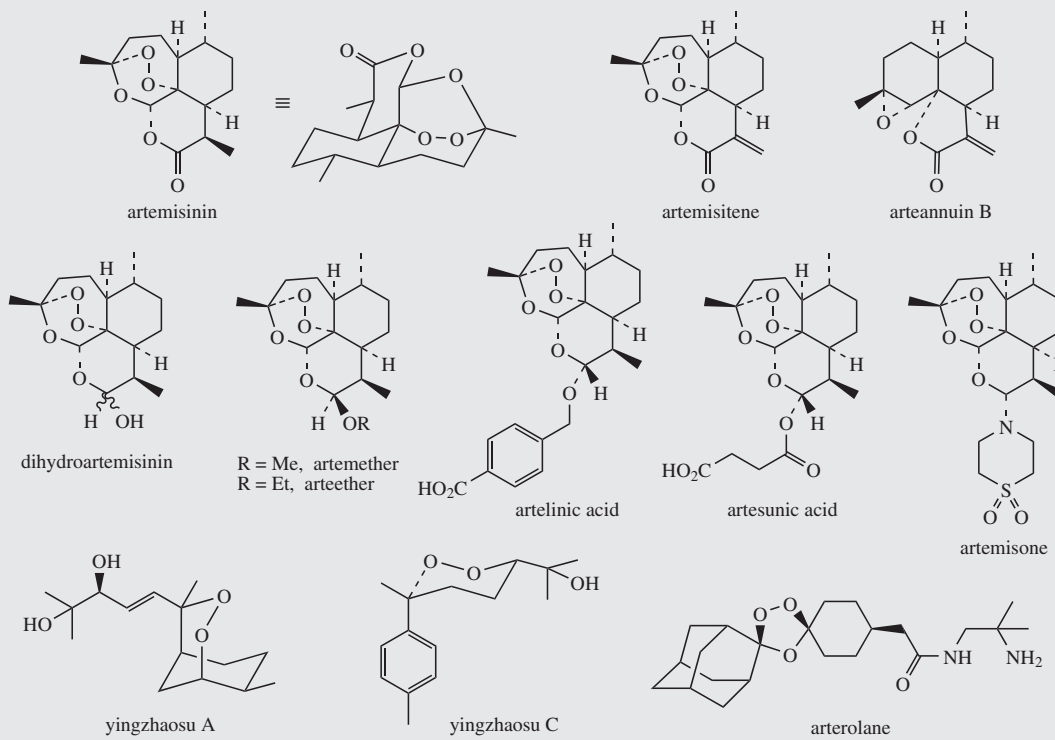


Figure 5.40

Box 5.6 (continued)

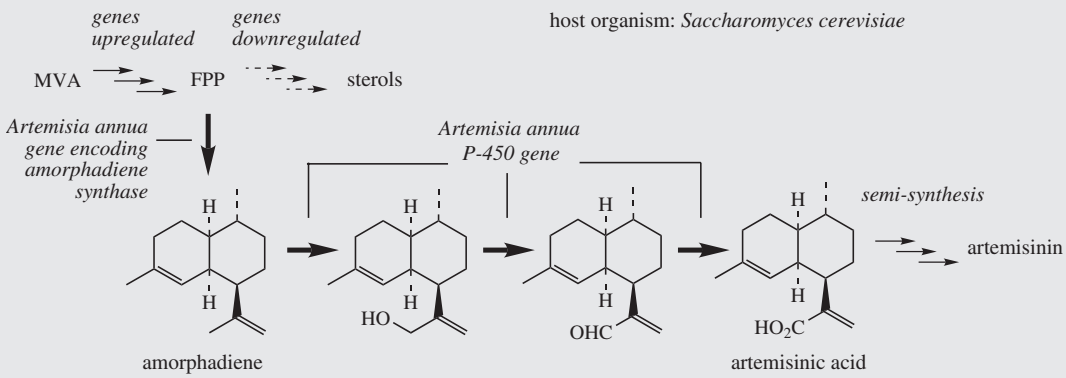


Figure 5.41

synthetic antimalarial lumefantrine. Arteether has similar activity. Being acetals, artemether and arteether are both extensively decomposed in acidic conditions, but are stable in alkali. The ester artesunic acid (**artesunate**) is also used in injection form, but is rather unstable in alkaline solution, hydrolysing to dihydroartemisinin. The ether artelinic acid is considerably more stable. These two compounds have a rapid action and particular application in the treatment of potentially fatal cerebral malaria. Dihydroartemisinin is a more active antimalarial than artemisinin and appears to be the main metabolite of these drugs in the body; unfortunately, dihydroartemisinin displays some neurotoxicity. Nevertheless, these agents rapidly clear the blood of parasites, though they do not have a prophylactic effect. **Artemisone** (Figure 5.40), a new semi-synthetic thiomorpholine dioxide derivative of dihydroartemisinin, is currently in clinical trials. Chemically, these agents are quite unlike any other class of current antimalarial agent, are well tolerated with no major side-effects, and so far no drug resistance is evident. As new analogues are introduced, they may well provide an important group of drugs in the fight against this life-threatening disease.

Currently, there is no shortage of *Artemisia* for drug use. However, as artemisinin derivatives become more widely used, the longer term supply of artemisinin for drug manufacture may need addressing. A significant development then is the use of genetic engineering to produce artemisinic acid in culture based on the biosynthetic pathway deduced for artemisinin (Figure 5.39). Genes from *A. annua* encoding the enzymes amorphadiene synthase and the cytochrome P-450-dependent monooxygenase that performs the three-step oxidation of amorphadiene to artemisinic acid were expressed in yeast *Saccharomyces cerevisiae* (Figure 5.41). To provide sufficient FPP precursor in the yeast host, several genes linked to the MVA → FPP transformation were upregulated; overexpression of the gene for HMGCoA reductase (see page 190) had the most significant effect. To avoid FPP being channelled off into steroid biosynthesis, the gene encoding for squalene synthase (see page 235) was downregulated. The net effect was high production of artemisinic acid (up to 100 mg l^{-1}) in the culture medium. This breakthrough is being commercialized to provide artemisinic acid for semi-synthesis of antimalarial drugs.

The relationship between a peroxide linkage and antimalarial activity is strengthened by the isolation of other sesquiterpene peroxides which have similar levels of activity as artemisinin. Thus, roots of the vine yingzhao (*Artabotrys uncinatus*; Annonaceae) which is also used as a traditional remedy for malaria, contain the bisabolyl derivatives yingzhaosu A and yingzhaosu C (Figure 5.40), the latter structure containing an aromatic ring of isoprenoid origin (compare the monoterpenes thymol and carvacrol, page 204). **Arterolane** (Figure 5.40) is a totally synthetic 1,2,4-trioxolane developed from the artemisinin template that shows potent antimalarial activity and is currently in clinical trials.

Despite intensive studies, there is as yet no generally accepted mechanism of action for artemisinin and derivatives. The malarial parasite utilizes the host's haemoglobin as a food source. However, haem, which is a soluble iron–porphyrin material released from haemoglobin as a result of proteolytic digestion, is toxic to *Plasmodium*, so it is normally converted into the insoluble non-toxic form haemozoin (malarial pigment) by enzymic polymerization. Agents like chloroquine (see page 382) may interfere with this polymerization process, or simply prevent the haemozoin from crystallizing. Of the various mechanisms proposed for the mode of action of artemisinin, initial reductive cleavage of the peroxide function by the Fe^{II} centre of haem to generate an oxygen radical species is widely accepted. This goes on to produce a carbon-centred radical that can interact with and effectively alkylate the porphyrin ring of haem. Alternatively, specific proteins or other biomolecules may suffer alkylation, resulting in death of the malarial parasite. It has also been suggested that artemisinin may act on a single enzymic target, a Ca^{2+}-ATPase; evidence for this comes from the observation that parasites with a mutant enzyme were insensitive to artemisinin.

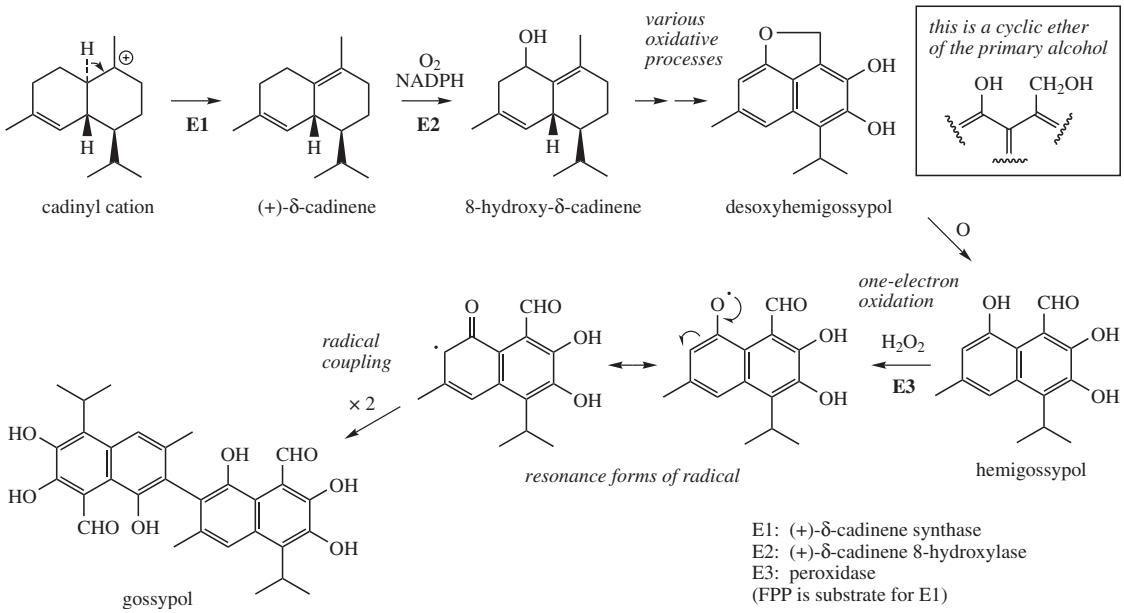


Figure 5.42

Gossypol (Figure 5.42) is an interesting and unusual example of a dimeric sesquiterpene in which loss of hydrogen has led to an aromatic system (compare the phenolic monoterpenes thymol and carvacrol, page 204). The cadinyl carbocation via **δ-cadinene** is involved in generating the basic aromatic sesquiterpene unit **hemigossypol**, and then dimerization is simply an example of phenolic oxidative coupling *ortho* to the phenol groups (Figure 5.42). The coupling is catalysed by an H_2O_2-dependent peroxidase enzyme. Gossypol is found in immature flower buds and seeds of the cotton plant (*Gossypium* species; Malvaceae), though it was originally isolated in small amounts from cottonseed oil. Its toxicity renders cottonseed oil unsafe for human consumption, but it has been used in China as a male infertility agent [Box 5.7].

Box 5.7

Gossypol

Gossypol occurs in the seeds of cotton (*Gossypium* species, e.g. *G. hirsutum, G. herbaceum, G. arboreum, G. barbadense*; Malvaceae) in amounts of 0.1–0.6%. Cotton is a major plant crop, and a considerable amount of cottonseed is produced as a by-product; this is mainly used as a feedstuff for cattle. Cottonseed, though protein-rich, has limited food use because of the cardiotoxic and hepatotoxic effects of gossypol in humans and animals other than ruminants. The contraceptive effects of gossypol were discovered when subnormal fertility in some Chinese rural communities was traced back to the presence of gossypol in dietary cottonseed oil. Gossypol acts as a male contraceptive, altering sperm maturation, spermatozoid motility, and inactivation of sperm enzymes necessary for fertilization. Extensive clinical trials in China have shown the antifertility effect is reversible after stopping the treatment provided that consumption has not been too prolonged. Cases of irreversible infertility have resulted from longer periods of drug use.

The gossypol molecule is chiral due to restricted rotation about the aryl–aryl linkage, and can thus exist as two atropisomers which do not easily racemize (Figure 5.43). Only the (−)-isomer is pharmacologically active as a contraceptive, whereas most of the toxic symptoms appear to be associated with the (+)-isomer. Most species of *Gossypium* (except *G. barbadense*)

Box 5.7 (continued)

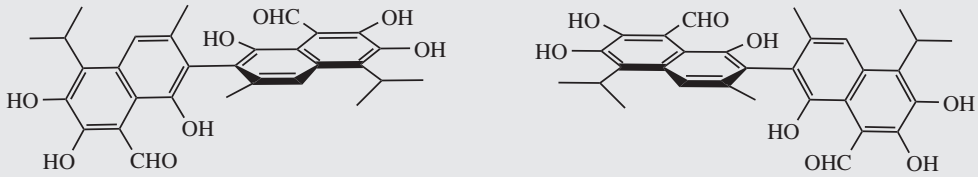


Figure 5.43

produce gossypol where the (+)-isomer predominates over the (−)-isomer, with amounts varying according to species and cultivar. The relative proportions of the two isomers appear to be controlled by regioselective coupling of two hemigossypol molecules by the peroxidase enzyme (Figure 5.42). Racemic (±)-gossypol (but neither of the enantiomers) complexes with acetic acid, so that suitable treatment of cottonseed extracts can separate the racemate from the excess of (+)-isomer. The racemic form can then be resolved. Other plants in the Gossypieae tribe of the Malvaceae also produce gossypol, with the barks of *Thespia populnea* (3.3%) and *Montezuma speciosissima* (6.1%) being particularly rich sources. Unfortunately, gossypol from these sources is almost entirely the (+)-form that lacks contraceptive activity. An exception is *Thespia danis*, where aerial parts contain predominantly (72%) (−)-gossypol; yields are about 0.2%. It is possible to reduce gossypol levels in cottonseed by interfering with the enzyme δ-cadinene synthase in the plant; this has been achieved by expression of the antisense gene, or by gene silencing through RNA interference: supplying a short strand of RNA to target that portion of mRNA responsible for a particular enzyme. Such techniques may lead to production of strains yielding non-toxic cottonseed oil.

The formation of sesquiterpenes by a carbocation mechanism also means that there is considerable scope for rearrangements of the Wagner–Meerwein type. So far, only occasional hydride migrations have been invoked in rationalizing the examples considered. Obviously, fundamental skeletal rearrangements will broaden the range of natural sesquiterpenes even further. That such processes do occur has been proven beyond doubt by appropriate labelling experiments, and a single example will be used as illustration. The **trichothecenes** are a group of fungal toxins found typically in infected grain foodstuffs [Box 5.8]. Their name comes from the fungal genus *Trichothecium*, but most of the known structures are derived from cultures of *Fusarium* species. A typical trichothecene contaminant is **3-acetyldeoxynivalenol**, which is produced from the less-substituted trichothecene **isotrichodermol** by a sequence of oxygenation, esterification, and de-esterification reactions (Figure 5.44). The trichothecenes have their origins in **nerolidyl diphosphate**, and ring closure of the bisabolyl cation derived from it generates a new carbocation with a five-membered ring (Figure 5.44). At this stage, a series of one hydride and two methyl migrations occur to give a cation, which loses a proton to produce the specific trichothecene precursor **trichodiene**. These migrations are fully backed up by experimental data and, although not immediately predictable, can be rationalized satisfactorily by consideration of the cation suitably bound to the enzyme surface, as shown in Figure 5.44. The sequence is initiated by a 1,4-hydride shift, which is spatially allowed by the relative proximity of the centres. Two 1,2-methyl shifts then follow, and it is important to note that each migrating group attacks the opposite side of the centre from which the previous group is departing, i.e. inverting the configuration at these centres. Accordingly, a concerted sequence of migrations is feasible; a more vivid example of this type of concerted process is seen in the formation of triterpenoids and steroids (see page 236). Loss of a proton and generation of a double bond then terminates the process, giving trichodiene. Oxygenation of trichodiene gives, in several steps, **isotrichotriol**. Two of the hydroxylations are at activated allylic positions; hydroxylation on the five-membered ring, therefore, will occur before the epoxidation. Ether formation, involving perhaps protonation, loss of water, and generation of an allylic cation completes the pathway to the basic trichothecene structure as in **isotrichodermol**.

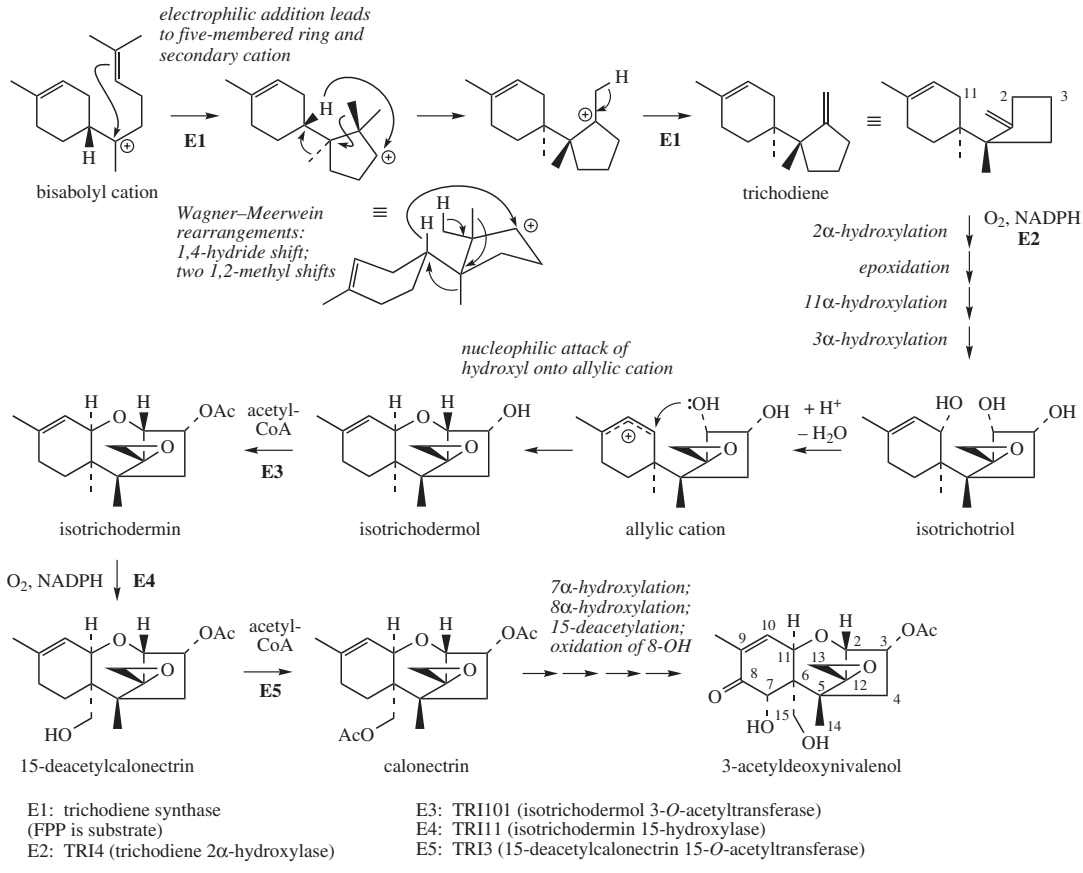


Figure 5.44

Box 5.8

Trichothecenes

The trichothecenes are a group of sesquiterpene toxins produced by several fungi of the genera *Fusarium, Myrothecium, Trichothecium*, and *Trichoderma*, which are parasitic on cereals such as maize, wheat, rye, barley, and rice. About 180 different structures have been identified, with some of these being isolated from plants of the genus *Baccharis* (Compositae/Asteraceae), where a symbiotic plant–fungus relationship may account for their production. Examples of trichothecene structures most commonly encountered as food contaminants include deoxynivalenol (DON; vomitoxin), diacetoxyscirpenol, T-2 toxin, and verrucarin A (Figure 5.45). The double bond and the epoxide group in the basic trichothecene skeleton are essential for toxicity, and the number of oxygen substituents and ester functions also contribute. Macrocyclic ester functions as seen in verrucarin A tend to produce the most toxic examples. Although these compounds are more toxic when injected, oral toxicity is relatively high, and lethal amounts can easily be consumed because of the nature of the host plants. They are sufficiently toxic to warrant routine analysis of foodstuffs such as wheat and flour, and also flour-derived products, e.g. bread, since they survive processing and the high temperatures used in baking. DON levels above 1 ppm are considered hazardous for human consumption. It is relevant to note that, when mammals ingest these compounds, a degree of de-epoxidation can occur, ascribed to gut microflora, thus providing a level of detoxification by removing a structural feature necessary for toxicity.

As their main mechanism of action, these compounds inhibit protein biosynthesis by binding to the ribosome and inhibiting peptidyl transferase activity (see page 422). They also inhibit DNA biosynthesis. A major human condition known to be caused by

Box 5.8 (continued)

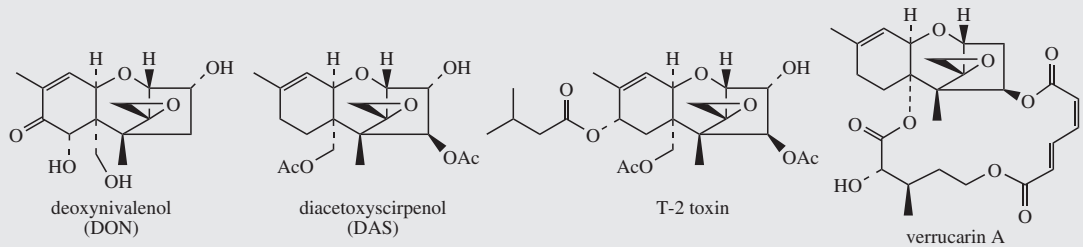


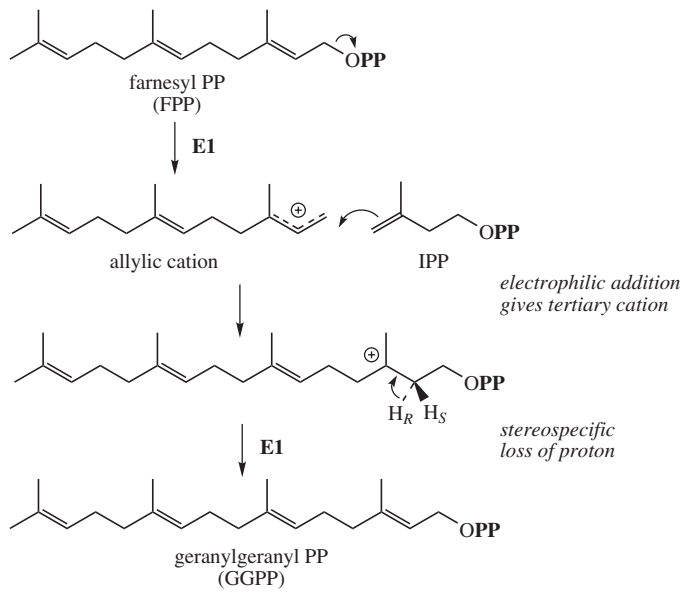
Figure 5.45

trichothecenes is alimentary toxic aleukia (ATA), characterized by destruction of the skin, haemorrhaging, inflammation, sepsis, a decrease in red and white blood corpuscles, bone marrow atrophy, and a high mortality rate. A severe outbreak of ATA was recorded in the former Soviet Union shortly after the Second World War when food shortages led to the consumption of grain that had overwintered in the field. This had become badly contaminated with *Fusarium sporotrichioides* and, hence, T-2 toxin. It is estimated that tens of thousands died as a result. Many trichothecene derivatives have been tested as potential anticancer agents, but they have proved too toxic for clinical use.

Finally, it is worth noting how many of the sesquiterpene derivatives described above are found in plants belonging to the daisy family, the Compositae/Asteraceae. Whilst sesquiterpenes are by no means restricted to this family, the Compositae/Asteraceae undoubtedly provides a very rich source.

DITERPENES (C_{20})

The diterpenes arise from **geranylgeranyl diphosphate** (**GGPP**), which is formed by addition of a further IPP molecule to FPP in the same manner as described for the lower terpenoids (Figure 5.46). One of the



E1: geranylgeranyl diphosphate synthase

Figure 5.46

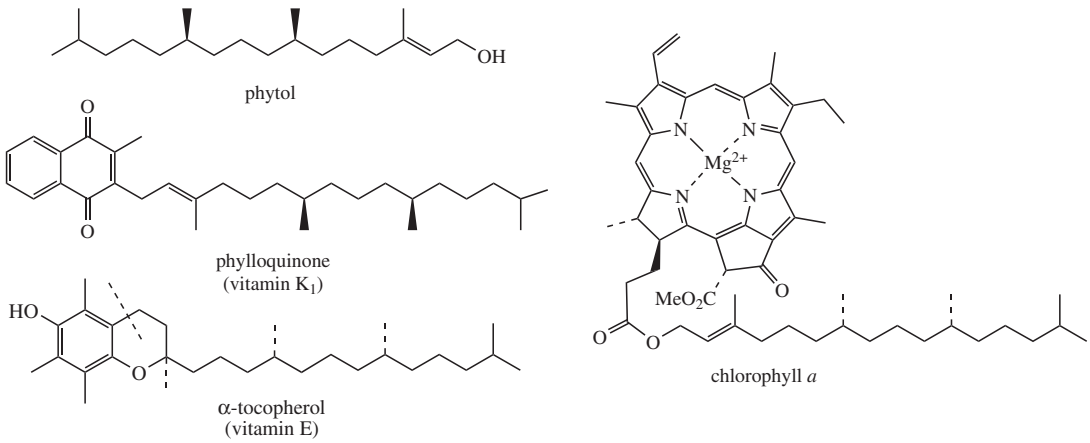


Figure 5.47

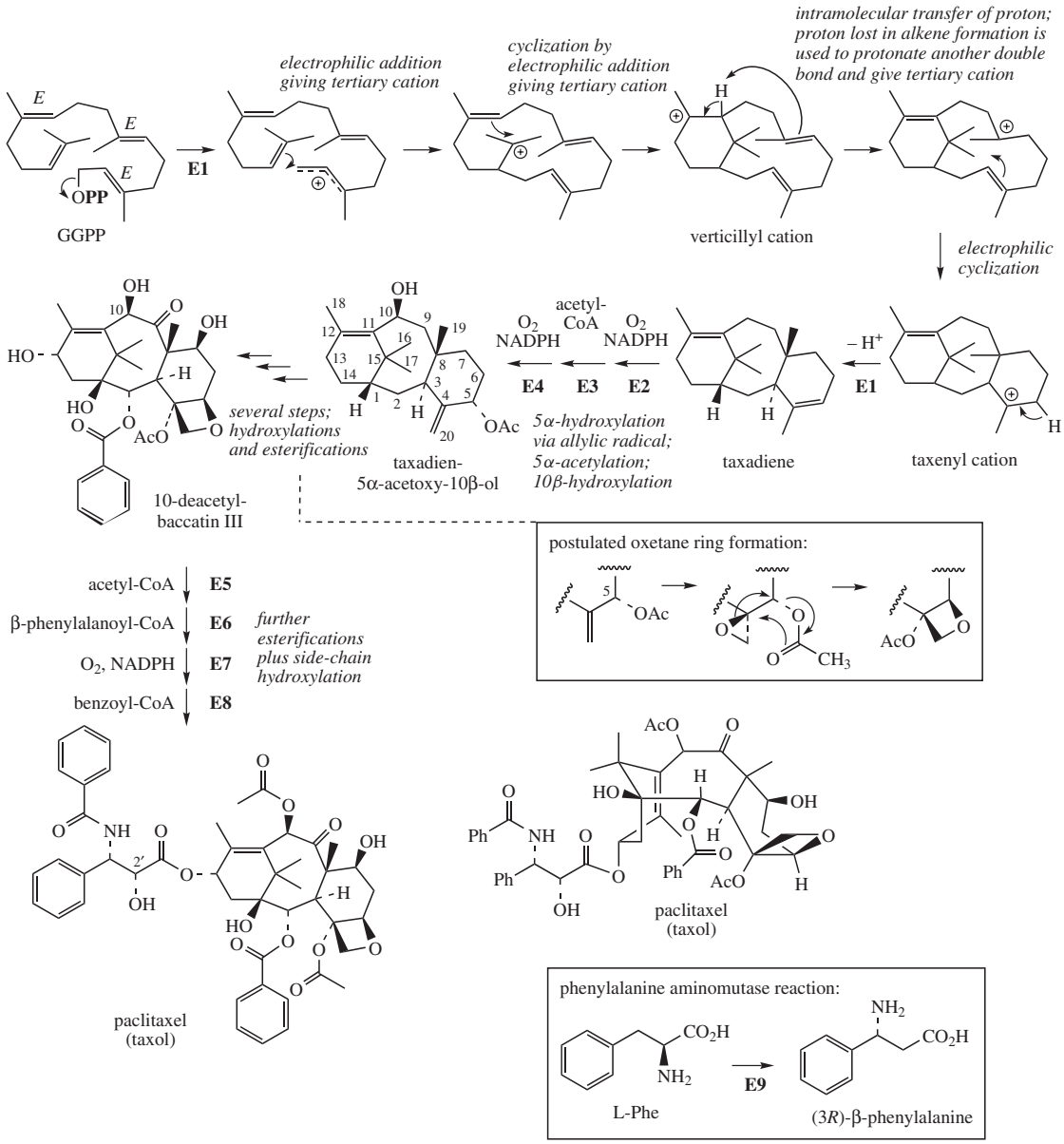
simplest and most important of the diterpenes is **phytol** (Figure 5.47), a reduced form of geranylgeraniol, which forms the lipophilic side-chain of the chlorophylls, e.g. **chlorophyll *a*** (Figure 5.47). Related haem molecules, the porphyrin components of haemoglobin, lack such lipophilic side-chains. Available evidence suggests that GGPP is involved in forming the ester linkage, and the three reduction steps necessary to form the phytol ester occur after attachment to the chlorophyll molecule. A phytyl substituent is also found in **vitamin K_1** (**phylloquinone**; Figure 5.47), a naphthoquinone derivative found in plants, though other members of the vitamin K group (**menaquinones**) from bacteria have unsaturated terpenoid side-chains of variable length. The phytyl group of phylloquinone is introduced by alkylation of dihydroxynaphthoic acid with phytyl diphosphate and a similar phytylation of homogentisic acid features in the formation of the E group vitamins (**tocopherols**). These compounds are discussed further under shikimate derivatives (see page 178).

Cyclization reactions of GGPP mediated by carbocation formation, plus the potential for Wagner–Meerwein rearrangements, will allow many structural variants of diterpenoids to be produced. The toxic principle 'taxine' from English yew (*Taxus baccata*; Taxaceae) has been shown to be a mixture of at least 11 compounds based on the **taxadiene** skeleton, which can readily be rationalized as in Figure 5.48, employing the same mechanistic principles as seen with mono- and sesqui-terpenes. A novel feature discovered is the enzyme's ability to utilize the proton released during alkene formation to protonate another double bond; this effectively relocates the cationic centre. Although these compounds are sometimes classified as diterpenoid alkaloids, the nitrogen atom is not incorporated into the diterpene skeleton, as exemplified by **taxol** (**paclitaxel**; Figure 5.48) from Pacific yew (*Taxus brevifolia*) [Box 5.9]. The side-chains in taxol containing aromatic rings are derived from shikimate via phenylalanine. The nitrogen atom derives from the amino acid β-phenylalanine, a rearranged version of the α-amino acid L-phenylalanine. Virtually all of the enzymes and genes for the whole sequence from GGPP to taxol are now characterized. Some enzymes show relatively broad substrate specificity. Taxol is an important anticancer agent, with a broad spectrum of activity against some cancers that do not respond to other agents.

Box 5.9

Taxus brevifolia and Taxol (Paclitaxel)

A note on nomenclature: the name taxol was given to a diterpene ester with anticancer properties when it was first isolated in 1971 from *Taxus brevifolia*. However, the name Taxol had already been registered as a trademark; this was subsequently used when the natural product was exploited commercially as a drug. Accordingly, the generic name paclitaxel has been assigned to the



E1: taxadiene synthase
E2: taxoid 5α-hydroxylase
E3: taxadienol *O*-acetyltransferase
E4: taxoid 10β-hydroxylase
E5: 10β-*O*-acetyltransferase
E6: baccatin III 13-*O*-phenylpropanoyltransferase
E7: β-phenylalanoylbaccatin III 2′-hydroxylase
E8: taxane *N*-benzoyltransferase
E9: phenylalanine aminomutase

Figure 5.48

Box 5.9 (continued)

compound, and the literature now contains an unhappy mixture of the two names, though the original name taxol is most often employed.

The anticancer drug taxol (Figure 5.48) is extracted from the bark of the Pacific yew, *T. brevifolia* (Taxaceae), a slow-growing shrub/tree found in the forests of northwest Canada (British Columbia) and the USA (Washington, Oregon, Montana, Idaho, and north California). Although the plant is not rare, it does not form thick populations, and it needs to be mature (about 100 years old) to be large enough for exploitation of its bark. The bark from about three mature trees is required to provide 1 g of taxol, and a course of treatment may need 2 g of the drug. Current demand for taxol is in the region of 250 kg per annum. Harvesting has been strictly regulated, but it was immediately realized that this source would not provide a satisfactory long-term supply of the drug. Taxol is now obtained by alternative means.

All parts of *T. brevifolia* contain a wide range of diterpenoid derivatives termed taxanes, which are structurally related to the toxic constituents found in other *Taxus* species, e.g. the common English yew, *T. baccata*. Nearly 400 taxanes have been characterized from various *Taxus* species, and taxol is a member of a small group of compounds possessing a four-membered oxetane ring and a complex ester side-chain in their structures; both of these features are essential for antitumour activity. Taxol is found predominantly in the bark of *T. brevifolia*, but in relatively low amounts (about 0.01–0.02%). Up to 0.033% of taxol has been recorded in some samples of leaves and twigs, but generally the taxol content is much lower than in the bark. The content of some other taxane derivatives in the bark is considerably higher, e.g. up to 0.2% baccatin III (Figure 5.49). Other taxane derivatives characterized include 10-deacetyltaxol, 10-deacetylbaccatin III, cephalomannine, and 10-deacetylcephalomannine (Figure 5.49).

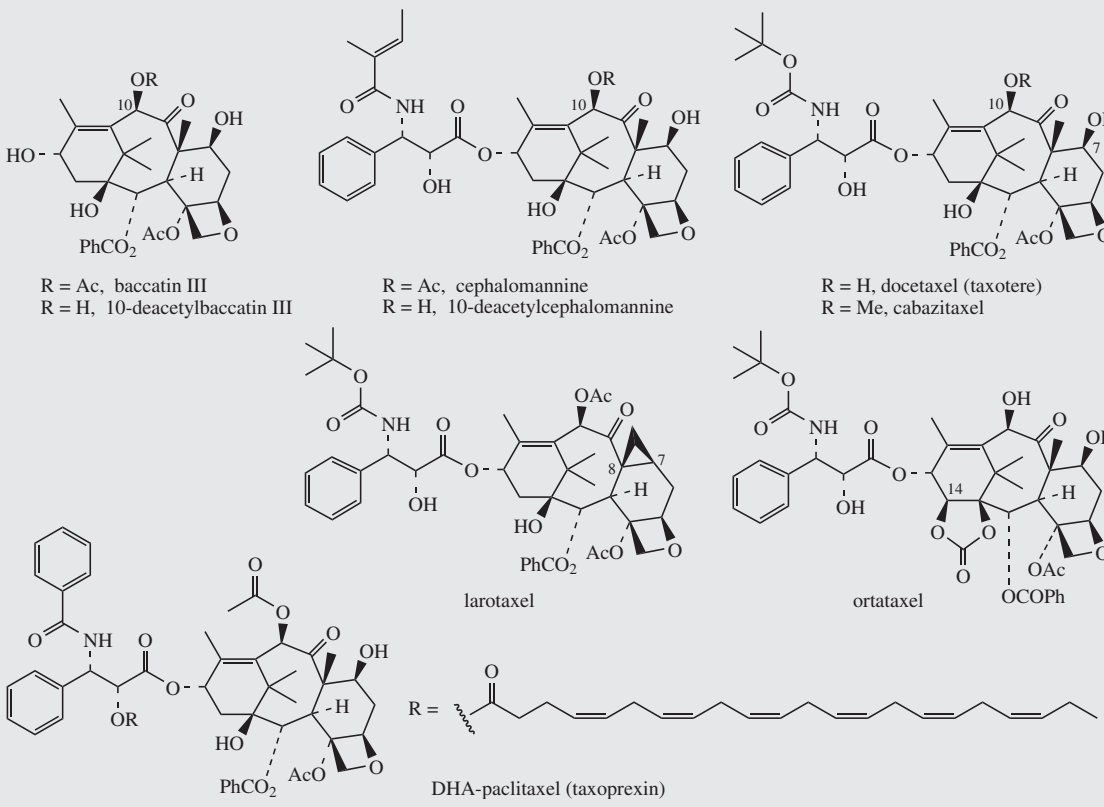


Figure 5.49

Box 5.9 (continued)

A satisfactory solution currently exploited for the supply of taxol and derivatives for drug use is to produce these compounds by semi-synthesis from more accessible structurally related materials. Both baccatin III and 10-deacetylbaccatin III (Figure 5.49) have been efficiently transformed into taxol, the latter being the preferred substrate. 10-Deacetylbaccatin III is readily extracted from the leaves and twigs of English yew *T. baccata*, and although the content is variable, it is generally present at much higher levels (up to 0.2%) than taxol can be found in *T. brevifolia. T. baccata* is widely planted as an ornamental tree in Europe and the USA and is much faster growing than the Pacific yew. Yew trees, typically 8–10 years old, can be harvested by pruning off top growth on a regular basis. In Europe, there are also arrangements whereby unwanted clippings from ornamental yew hedges may be channelled into drug production.

Semi-synthesis provided reliable supplies to establish taxol as an important anticancer drug. However, the chemical conversion of 10-deacetylbaccatin III is still complex enough to consider alternative approaches. Taxol-producing cell cultures of Chinese yew *T. chinensis* have been established and are currently grown in fermentors for commercial drug production. Taxol is isolated and purified from the fermentation broth; taxol yields in the region of 20 mg l^{-1} can be obtained from cultures of various *Taxus* species.

Initial optimism for obtaining taxol by microbial culture was tempered by the very low levels produced. Fungi such as *Taxomyces adreanae* isolated from the inner bark of *T. brevifolia*, and *Pestalotiopsis microspora* from the inner bark of the Himalayan yew (*T. wallachiana*), appear to have inherited the ability to synthesize taxol by gene transfer from the host tree. At best, taxol levels measured were only about 70 μg l^{-1} of culture medium, and thus commercially insignificant. However, the use of microorganisms and enzymes to specifically hydrolyse ester groups from mixtures of structurally related plant-derived taxanes in crude extracts and thus improve the yields of 10-deacetylbaccatin III has been reported.

Paclitaxel (**Taxol®**) is used clinically in the treatment of ovarian and breast cancers, non-small-cell lung cancer, small-cell lung cancer, and cancers of the head and neck. A disadvantage associated with taxol is its very low water solubility. **Docetaxel** (**Taxotere®**; Figure 5.49) is a side-chain analogue of taxol; it was one of the intermediates produced during a semi-synthesis of taxol from 10-deacetylbaccatin III. It was found to be more active than taxol, and, crucially, was more water-soluble. It is used against ovarian and breast cancers and non-small-cell lung cancer.

Clinical development of taxol did not proceed until it was realized that it possessed a mode of action different from the drugs available at the time. Taxol acts as an antimitotic by binding to microtubules, promoting their assembly from tubulin, and stabilizing them against depolymerization during cell division. The resultant abnormal tubulin–microtubule equilibrium disrupts the normal mitotic spindle apparatus and blocks cell proliferation. Vincristine and vinblastine (see page 375), and podophyllotoxin (see page 155), also interfere with the tubulin system, but bind to the protein tubulin in the mitotic spindle, preventing polymerization and assembly into microtubules. Taxol thus shares a similar target, but has a different mechanism of action to the other antimitotics. In recent years, other natural products have been discovered that have the same or a similar mode of action; for example, the epothilones (see page 85) and discodermolide (see page 90). Taxol has also been shown to bind to a second target, a protein which normally blocks the process of apoptosis (cell death). Inhibition of this protein allows apoptosis to proceed.

Several taxol analogues are in clinical trials. Considerable effort is being directed to improve solubility and to counter a tendency to induce multiple drug resistance. Four compounds in advanced clinical trials are **DHA–paclitaxel** (**Taxoprexin®**), **cabazitaxel, larotaxel**, and **ortataxel** (Figure 5.49). The first is an ester of taxol at the side-chain hydroxyl with the polyunsaturated fatty acid DHA (see page 49); polyunsaturated fatty acids are able to deliver cytotoxic drugs selectively to cancer cells. The other three structures are based on taxotere. Cabazitaxel is the 7,10-dimethyl ether and larotaxel is a 7,8-cyclopropane analogue. Ortataxel is an orally active cyclic carbonate derivative produced from the natural 14β-hydroxy-10-deacetylbaccatin III.

The latex of some plants in the genus *Euphorbia* (Euphorbiaceae) can cause poisoning in humans and animals, skin dermatitis, cell proliferation, and tumour promotion (co-carcinogen activity). Many species of *Euphorbia* are regarded as potentially toxic, and the latex can produce severe irritant effects, especially on mucous membranes and the eye. Most of the biological effects are due to diterpene esters, e.g. esters of **phorbol** (Figure 5.50), which activate protein kinase C, an important and widely distributed enzyme responsible for phosphorylating many biochemical entities. The permanent activation of protein kinase C is thought to lead to the uncontrolled cancerous growth. The most commonly encountered ester of phorbol is 12-*O*-myristoylphorbol 13-acetate (Figure 5.50), one of the most potent tumour promoters known. The origins of phorbol are not known, but may be rationalized as in Figure 5.50. Cyclization of GGPP generates a cation containing a 14-membered ring system. Loss of a proton via cyclopropane ring formation leads to **casbene**, an antifungal metabolite produced by the castor oil plant, *Ricinus communis* (Euphorbiaceae). Casbene, via the ring closures shown in Figure 5.50, is then likely to be the precursor of the phorbol ring system.

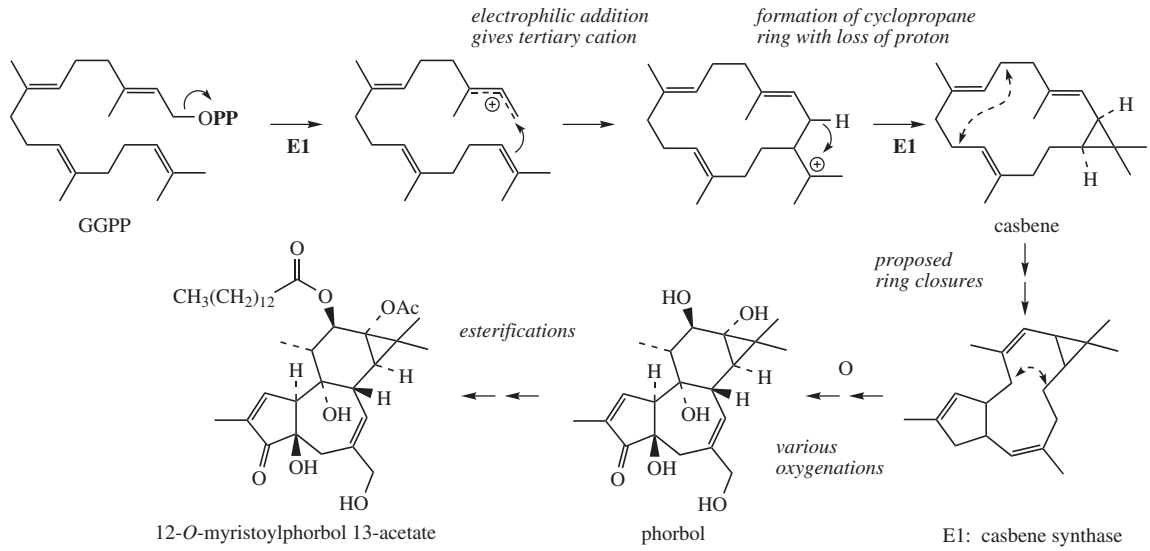
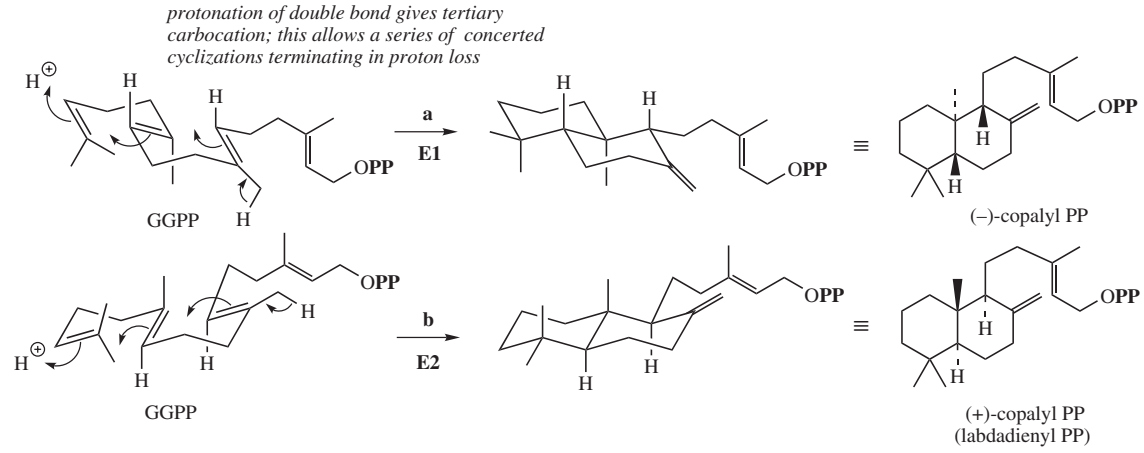


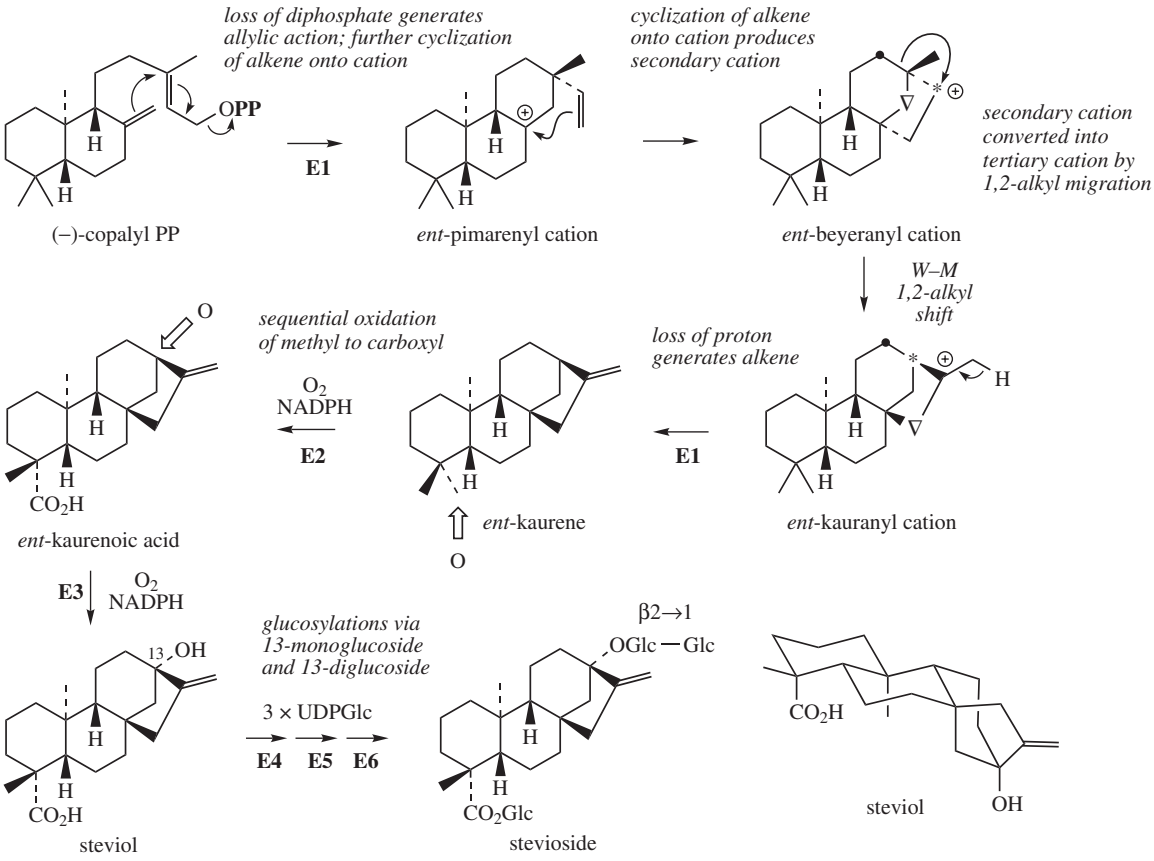
Figure 5.50

In contrast to the cyclization mechanisms shown in Figures 5.48 and 5.50, where loss of diphosphate generates the initial carbocation, many of the natural diterpenes have arisen by a different mechanism. Carbocation formation is initiated by protonation of the double bond at the head of the chain, leading to a first cyclization sequence. Loss of the diphosphate later on also produces a carbocation and facilitates further cyclization. The early part of the sequence resembles that involved in hopanoid biosynthesis (see page 242), and to some extent triterpenoid and steroid biosynthesis (see page 238), though in the latter cases it is the opening of the epoxide ring in the precursor squalene oxide rather than protonation of an alkene that is responsible for generation of the cationic intermediates. Protonation of GGPP can initiate a concerted cyclization sequence, terminated by loss of a proton from a methyl, yielding **(–)-copalyl PP** (Figure 5.51, a). The stereochemistry in this product is controlled by the folding



E1: (–)-copalyl diphosphate synthase (*ent*-kaurene synthase A)
E2: (+)-copalyl diphosphate synthase (part of bifunctional abietadiene synthase)

Figure 5.51



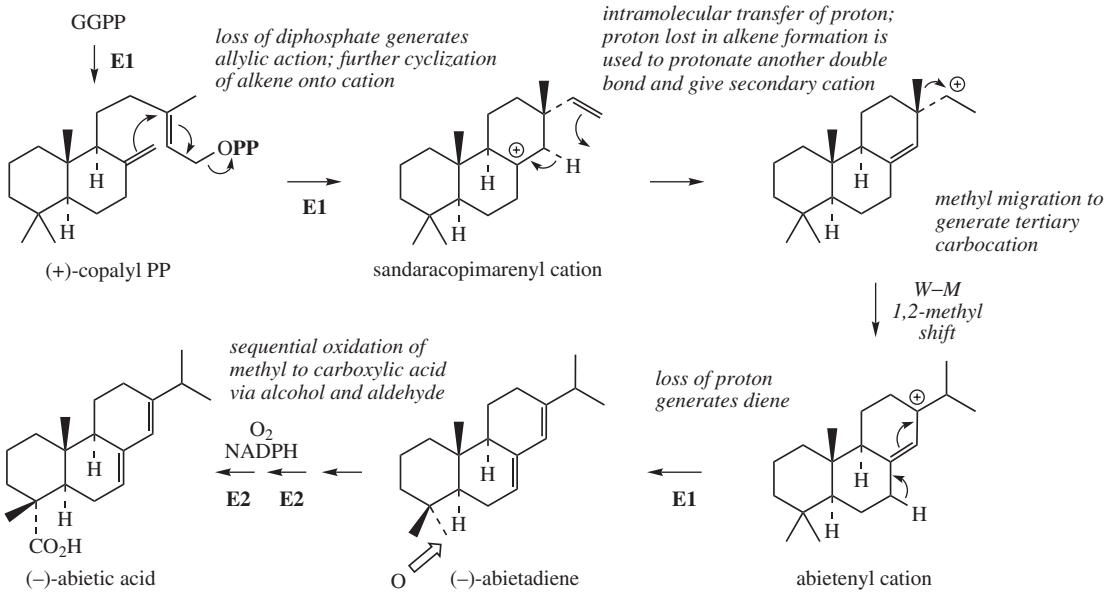
E1: kaurene synthase (*ent*-kaurene synthase B)
E2: *ent*-kaurene oxidase
E3: *ent*-kaurenoic acid 13-hydroxylase
E4, E5, E6: glucosyltransferases

Figure 5.52

of the substrate on the enzyme surface, though an alternative folding can lead to **(+)-copalyl PP** (labdadienyl PP), the enantiomeric product with opposite configurations at the newly generated chiral centres (Figure 5.51, b).

From (−)-copalyl PP, a sequence of cyclizations and a rearrangement, all catalysed by a single enzyme kaurene synthase, leads to ***ent*-kaurene** (Figure 5.52). As shown, this involves loss of the diphosphate leaving group enabling carbocation-mediated formation of the third ring system, and subsequent production of the fourth ring. Then follows a Wagner–Meerwein migration, effectively contracting the original six-membered ring to a five-membered one, whilst expanding the five-membered ring to give a six-membered ring. The driving force is transformation of a secondary carbocation to give a tertiary one, but this also results in the methyl group no longer being at a bridgehead, and what appears at first glance to be merely a confusing change in stereochemistry. Loss of a proton from this methyl generates the exocyclic double bond of ***ent*-kaurene** and provides an exit from the carbocationic system. The prefix *ent* is used to indicate enantiomeric; the most common stereochemistry is that found in (+)-copalyl PP (Figure 5.51) and derivatives, so the kaurene series is termed 'enantiomeric'.

ent-Kaurene is the precursor of **stevioside** (Figure 5.52) in the plant *Stevia rebaudiana* (Compositae/Asteraceae) by relatively simple oxidation, hydroxylation, and glucosylation reactions. Both glucosyl ester and glucoside linkages are present in stevioside, and these help to confer an intensely sweet taste to this and related compounds. Stevioside is present in the plant leaf in quite large amounts (3–10%), is some 200–300 times as sweet as sucrose, and is being used in many countries as a commercial non-calorific sweetening agent.



E1: abietadiene synthase (bifunctional enzyme)
E2: abietadienol/al oxidase

Figure 5.53

The alternative stereochemistry typified by (+)-copalyl PP can be seen in the structure of **abietic acid** (Figure 5.53), the major component of the rosin fraction of turpentine from pines and other conifers (Table 5.1). Initially, the tricyclic system is built up as in the pathway to *ent*-kaurene (Figure 5.52), via the same mechanism, but generating the enantiomeric series of compounds. An internal proton transfer (compare taxadiene page 225) relocates the carbocationic centre to the side-chain, which then undergoes a methyl migration (Figure 5.53). Finally, loss of a proton leads to the diene **abietadiene**. The whole sequence from GGPP to abietadiene is catalysed by a single bifunctional enzyme; the first activity is a (+)-copalyl PP synthase, and the product is then transferred to the second active site. In the first step, carbocation formation is proton initiated, whilst the second is ionization initiated. **Abietic acid** results from sequential oxidation of the 4α-methyl. Wounding of pine trees leads to an accumulation at the wound site of both monoterpenes and diterpenes, and fractionation by distillation gives turpentine oil and rosin. The volatile monoterpenes seem to act as a solvent to allow deposition of the rosin layer to seal the wound. The diterpenes in rosin have both antifungal and insecticidal properties.

Extensive modification of the copalyl diterpene skeleton is responsible for generation of the **ginkgolides**, highly oxidized diterpene trilactones which are the active principles of *Ginkgo biloba* (Ginkgoaceae) [Box 5.10]. Apart from the C_{20} skeleton, hardly any typical terpenoid features are recognizable. However, a speculative scheme involving several rearrangements, ring cleavage, and formation of lactone rings can broadly explain its origin and provide a link between these extremely complex ginkgolide structures and the more familiar diterpenes (Figure 5.54). Detailed evidence is lacking. What is known is that levopimaradiene and abietatriene are precursors; levopimaradiene synthase is also a bifunctional enzyme (compare abietadiene synthase) that converts GGPP into levopimaradiene via (+)-copalyl PP. The unusual *tert*-butyl substituent arises as a consequence of the A ring cleavage. **Bilobalide** (Figure 5.54) contains a related C_{15}-skeleton, and is most likely a partially degraded ginkgolide. *Ginkgo* is the world's oldest tree species, and its leaves are a highly popular and currently fashionable health supplement, taken in the anticipation that it can delay some of the degeneration of the faculties normally experienced in old age.

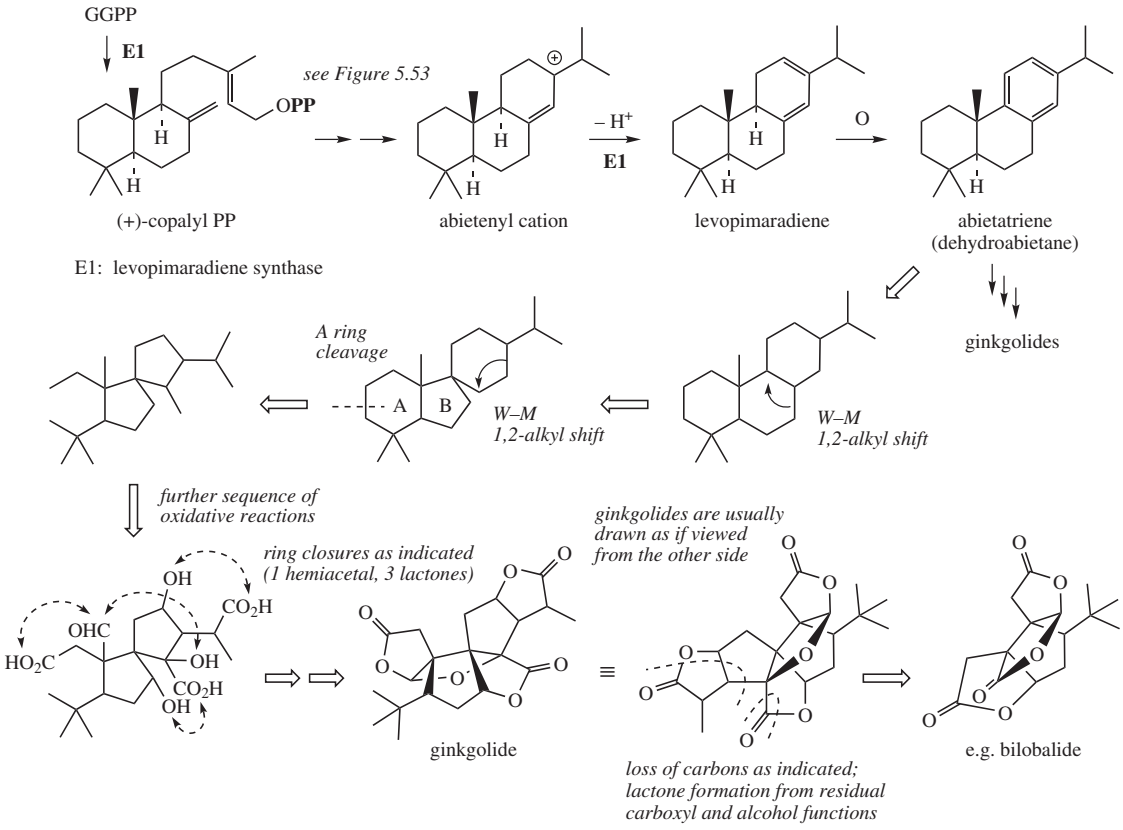


Figure 5.54

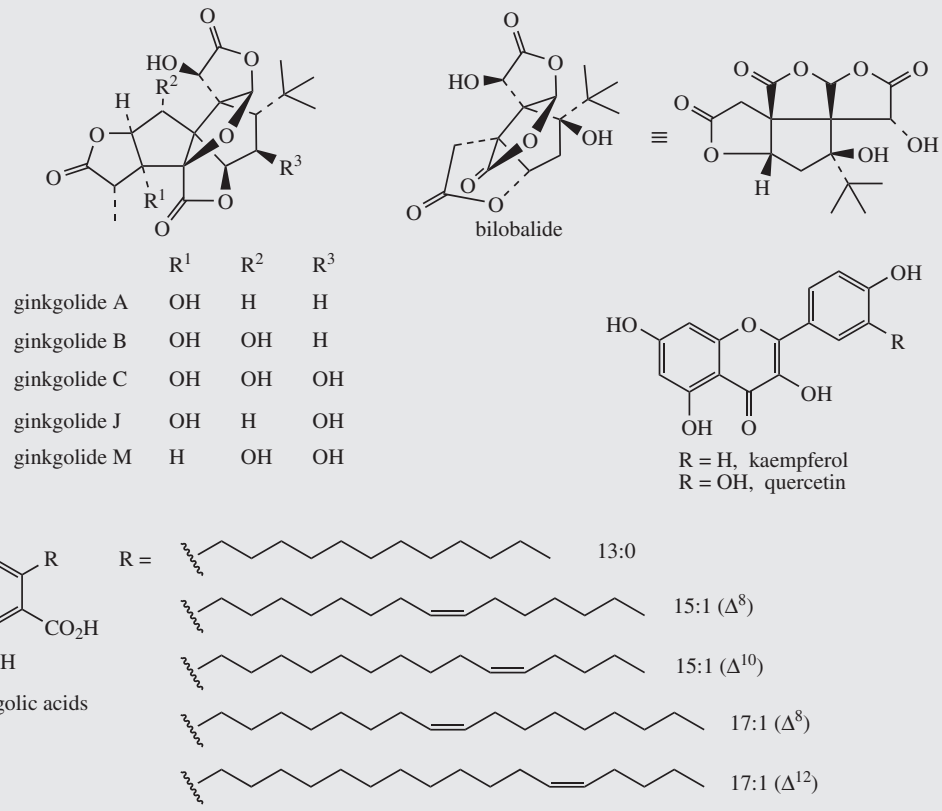
Box 5.10

Ginkgo biloba

Ginkgo biloba is a primitive member of the gymnosperms and the only survivor of the Ginkgoaceae, all other species being found only as fossils. It is a small tree native to China, but widely planted as an ornamental, and cultivated for drug use in Korea, France, and the United States. Standardized extracts of the leaves are marketed against cerebral vascular disease and senile dementia. Extracts have been shown to improve peripheral and cerebrovascular circulation. The decline in cognitive function and memory processes in old age can be due to disturbances in brain blood circulation, and thus **ginkgo** may exert beneficial effects by improving this circulation, and assist with other symptoms such as vertigo, tinnitus, and hearing loss. Virtually all clinical studies report positive results regarding cerebral insufficiency.

The active constituents have been characterized as mixtures of terpenoids and flavonoids. The dried leaves contain 0.1–0.25% terpene lactones, comprising predominantly the five ginkgolides (A, B, C, J, and M) and bilobalide (Figure 5.55). Bilobalide accounts for about 30–40% of the mixture, whilst ginkgolide A is the predominant ginkgolide (about 30%). The ginkgolides are diterpenoid in nature, whilst bilobalide is described as sesquiterpenoid. However, bilobalide bears such a structural similarity to the ginkgolides, that it is most probably a degraded ginkgolide. The ginkgolides have been shown to have potent and selective antagonistic activity towards platelet-activating factor (PAF, see page 44), which is implicated in many physiological processes. More recently, ginkgolides have been shown to extert antagonistic effects on glycine receptors; together with GABA receptors, glycine receptors are the main inhibitory receptors in the central nervous system. Bilobalide lacks PAF antagonistic activity but shows neuroprotective effects. The flavonoid content of the dried leaves is 0.5–1.0%, and consists of a mixture of mono-, di-, and tri-glycosides of the flavonols kaempferol and quercetin (Figure 5.55; see also page 171) and some biflavonoids. These probably

Box 5.10 (continued)



	R^1	R^2	R^3
ginkgolide A	OH	H	H
ginkgolide B	OH	OH	H
ginkgolide C	OH	OH	OH
ginkgolide J	OH	H	OH
ginkgolide M	H	OH	OH

Figure 5.55

also contribute to the activity of ginkgo, and may act as radical scavengers. Ginkgo leaf is also a rich source of shikimic and quinic acids (see page 138).

Extracts of ginkgo for drug use are usually standardized to contain flavonoid glycosides and terpene lactones in a ratio of 24% to 6%, or 27% to 7%. Standardized extracts should also contain less than 5 ppm of alkyl phenols, since these are known to be allergenic, capable of inducing contact dermatitis, as well as being potentially cytotoxic and mutagenic. The ginkgolic acids (also known as anacardic acids, see page 118) form the main group of alkyl phenols found in ginkgo leaf (up to 1.7%). These phenolic acids have long, mainly unsaturated, alkyl chains and are related to the fatty acids (Figure 5.55). Ginkgo may be combined with ginseng (see page 245) in the treatment of geriatric disorders. Ginkgo and the ginkgolides are undergoing extensive investigation in conditions where there are high PAF levels, e.g. shock, burns, ulceration, and inflammatory skin disease.

In **forskolin** (colforsin; Figure 5.56), the third ring is heterocyclic rather than carbocyclic. The basic skeleton of forskolin can be viewed as the result of quenching of the cation by water as opposed to proton loss, followed by S_N2'-type nucleophilic substitution onto the allylic diphosphate (or nucleophilic substitution onto the allylic cation generated by loss of diphosphate) (Figure 5.56). A series of oxidative modifications will then lead to forskolin. This compound has been isolated from roots of *Coleus forskohlii* (*Plectranthus barbatus*) (Labiatae/Lamiaceae), a plant used in Indian traditional medicine, and has been found to lower blood pressure and have cardioprotective properties. Forskolin has become a valuable pharmacological tool as a potent stimulator of adenylate cyclase activity, and has shown promising potential for the treatment of glaucoma,

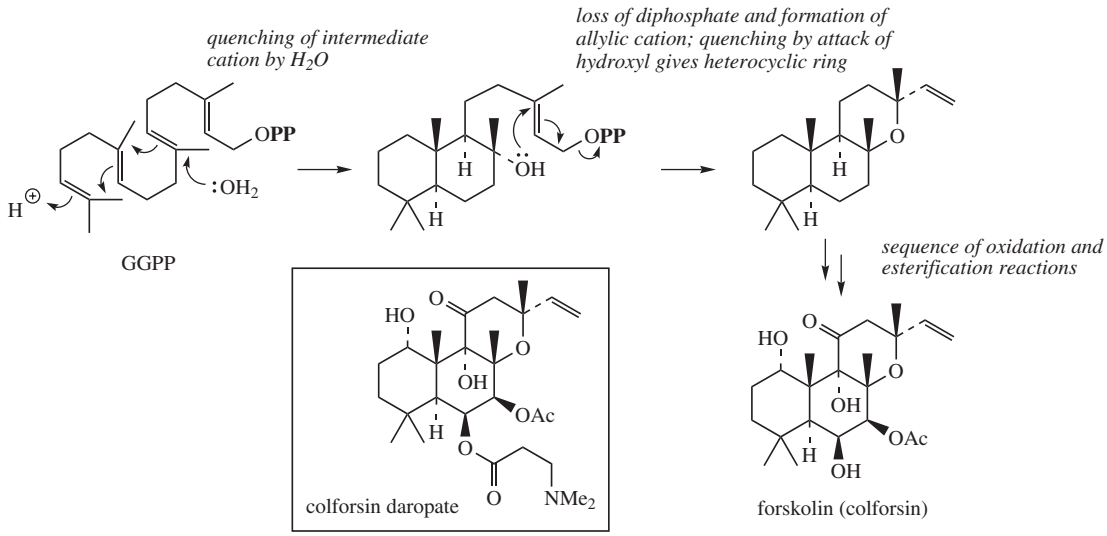


Figure 5.56

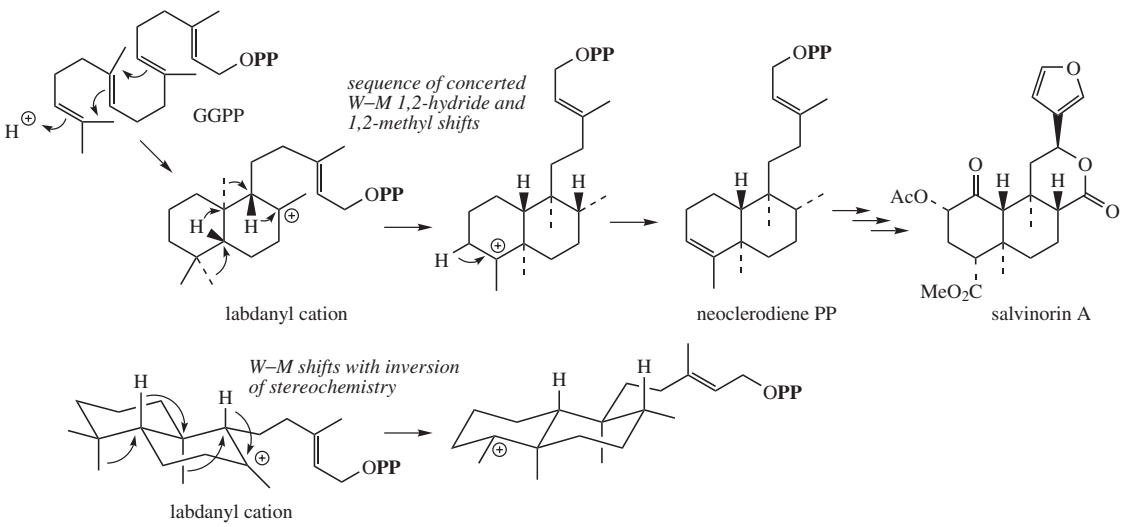


Figure 5.57

congestive heart failure, hypertension, and bronchial asthma. The water-soluble derivative colforsin daropate has been introduced for drug use.

In Figure 5.56, the GGPP-derived cation is quenched by the addition of water; in other cases, it may precipitate a series of concerted Wagner–Meerwein hydride and methyl migrations, resulting in a rearranged skeleton with different stereochemistry (Figure 5.57). Similar processes are encountered in triterpenoid and steroid biosynthesis (see page 238). The first migration is of hydride, generating a new tertiary carbocation, with successive methyl, hydride, and methyl migrations; the sequence is terminated by proton loss, giving **neoclerodiene PP**. The migrating groups are each positioned *anti* to one another, one group entering whilst the other leaves from the opposite side of the stereocentre. This inverts the configuration at each appropriate centre. Undoubtedly arising from this system is the

neoclerodane diterpene **salvinorin A**, the active hallucinogen from *Salvia divinorum* (Labiatae/Lamiaceae), a traditional medicine of some Mexican Indians. Salvinorin A is currently attracting a lot of interest, since it is a selective agonist at the kappa-opioid receptor (see page 352). It is the first non-nitrogenous agonist known, and is a valuable lead for the development of other selective ligands that may have therapeutic potential in a range of conditions, including pain, nausea, and depression.

SESTERTERPENES (C_{25})

Although many examples of this group of natural terpenoids are now known, they are found principally in fungi and marine organisms, and span relatively few structural types. The origins of **ophiobolene** and **ophiobolin A** (Figure 5.58) from cyclization of **geranylfarnesyl PP** (**GFPP**) in the plant pathogen *Helminthosporium maydis* is shown in Figure 5.58, and provides no novel features except for an experimentally demonstrated 1,5-hydride shift. GFPP arises by a continuation of the chain extension process, adding a further IPP unit to GGPP. Ophiobolin A shows a broad spectrum of biological activity against bacteria, fungi, and nematodes. The most common type of marine sesterterpenoid is exemplified by **sclarin**, and this structure can be envisaged as the result of a concerted cyclization sequence (Figure 5.59) analogous to that seen with GGPP in the diterpenoids, and with squalene in the hopanoids (see below).

TRITERPENES (C_{30})

Triterpenes are not formed by an extension of the now familiar process of adding IPP to the growing chain. Instead, two molecules of FPP are joined tail-to-tail to yield

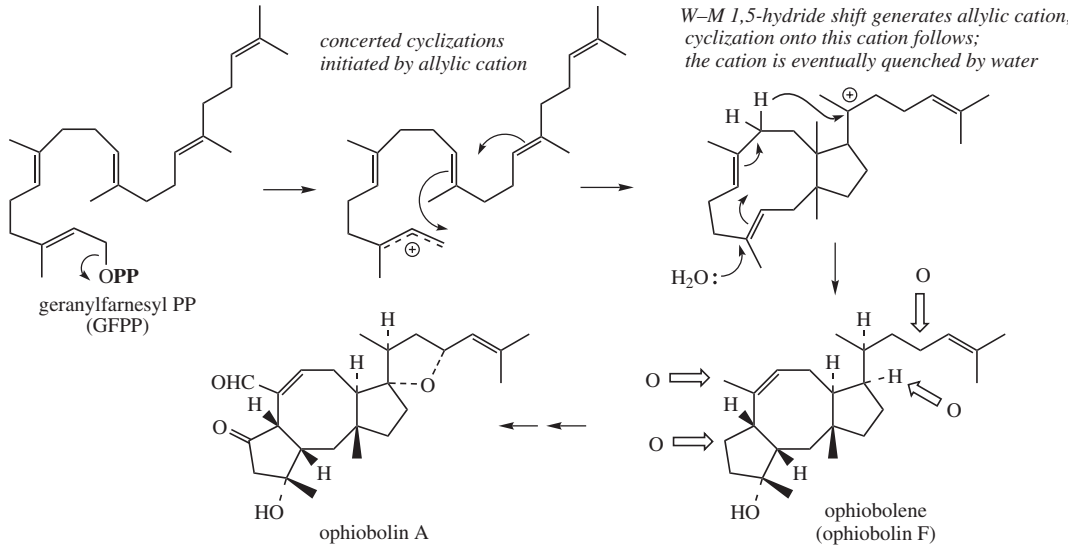


Figure 5.58

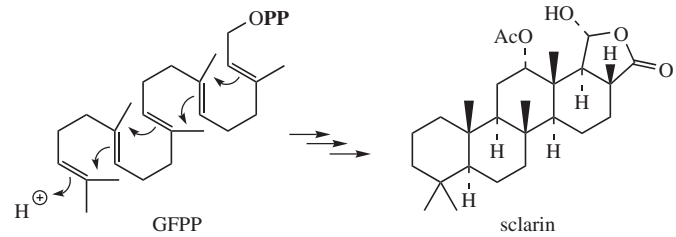


Figure 5.59

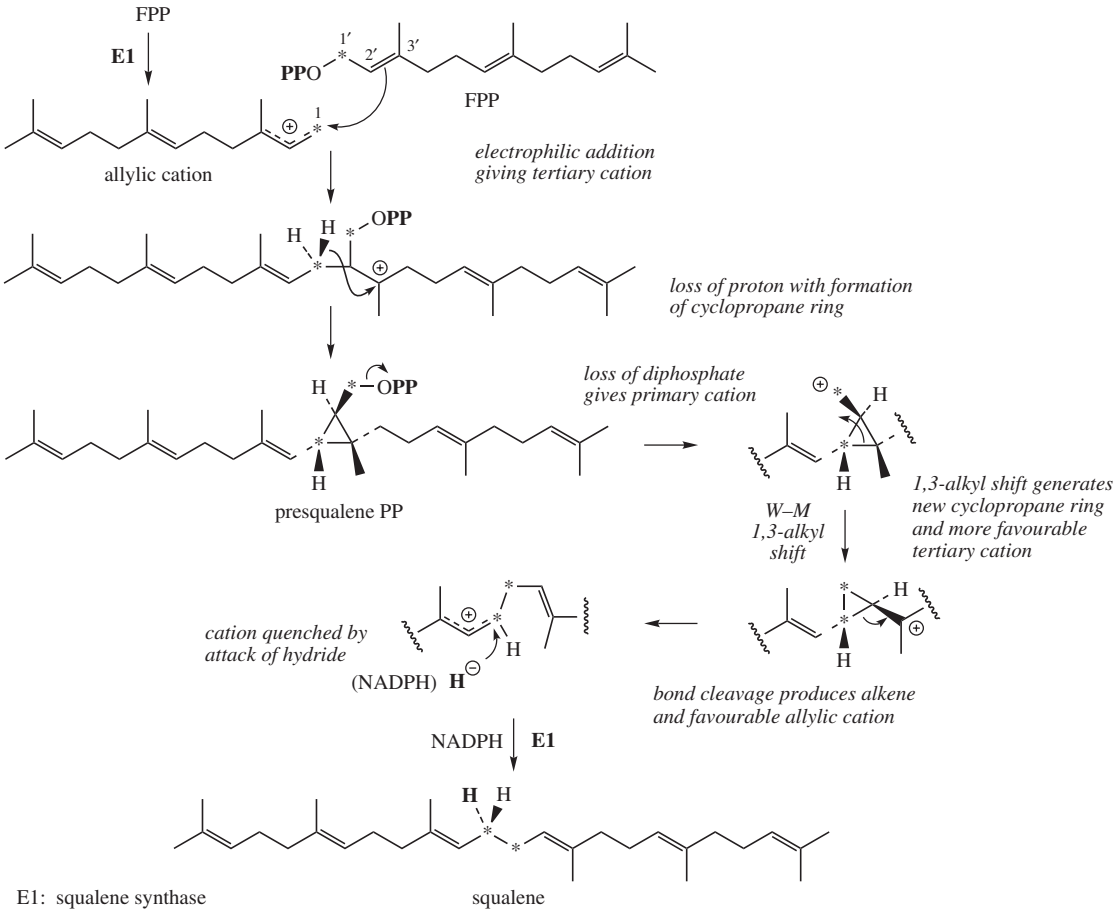


Figure 5.60

squalene (Figure 5.60); in general, FPP is formed from MVA (see page 192). Squalene is a hydrocarbon originally isolated from the liver oil of shark (*Squalus* sp.), but was subsequently found in rat liver and yeast, and these systems were used to study its biosynthetic role as a precursor of triterpenes and steroids. Several seed oils are now recognized as quite rich sources of squalene, e.g. *Amaranthus cruentus* (Amaranthaceae). During the coupling process, which on paper merely requires removal of the two diphosphate groups, a proton from a C-1 position of one molecule of FPP is lost and a proton from NADPH is inserted. Difficulties with formulating a plausible mechanism for this unlikely reaction were resolved when **presqualene diphosphate**, an intermediate in the process, was isolated from rat liver. Its characterization as a cyclopropane derivative immediately ruled out all the hypotheses current at the time.

The formation of presqualene PP in Figure 5.60, is initiated by attack of the 2,3-double bond of FPP onto the farnesyl cation, which is mechanistically equivalent to normal chain extension using IPP. The resultant tertiary cation is discharged by loss of a proton and formation of a cyclopropane ring, giving presqualene PP. An exactly analogous sequence was used for the origins of irregular monoterpenes (see page 205). Obviously, to then form squalene, C-1s of the two FPP units must eventually be coupled, whilst presqualene PP formation has actually joined C-1 of one molecule to C-2 of the other. To account for the subsequent change in bonding of the two FPP units, a further cyclopropane cationic intermediate is proposed. Loss of diphosphate from presqualene PP would give an unfavourable primary cation, which via Wagner–Meerwein rearrangement can generate a tertiary carbocation and achieve the required C-1–C-1′ bond.

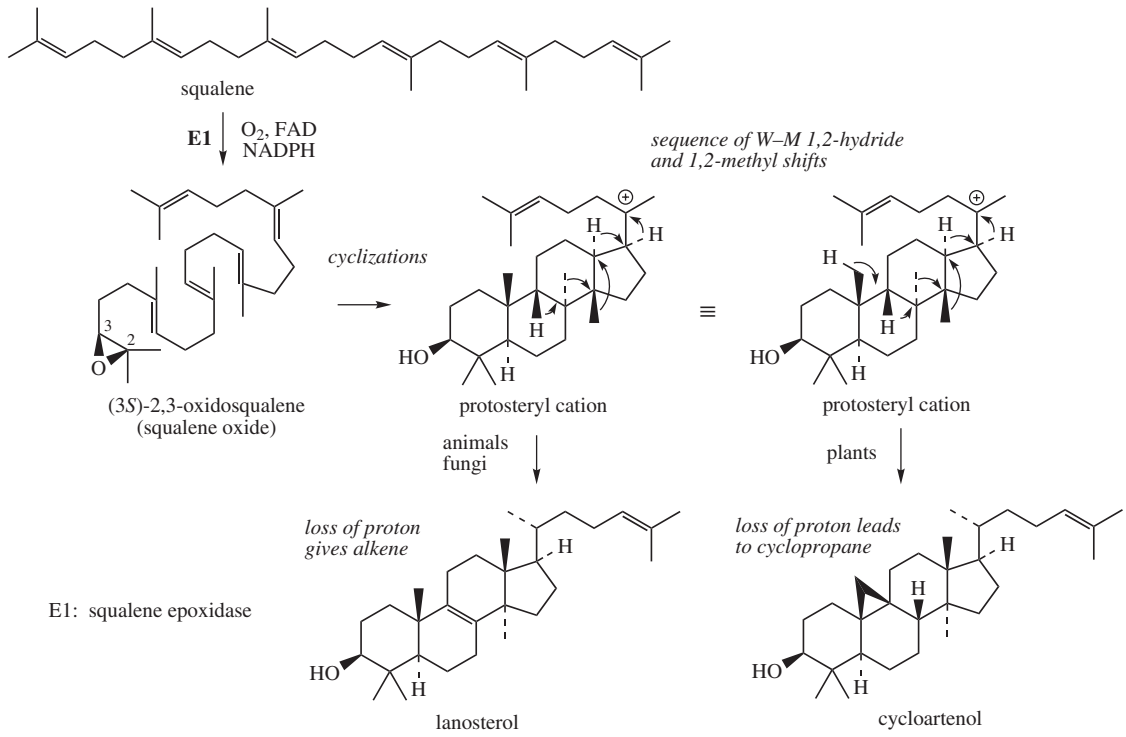


Figure 5.61

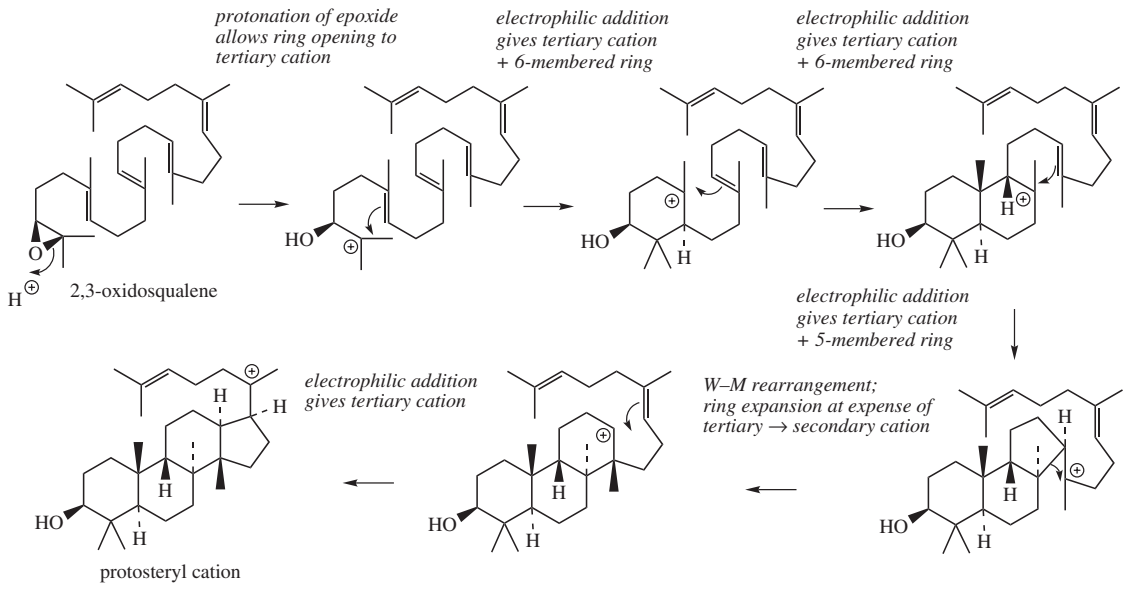


Figure 5.62

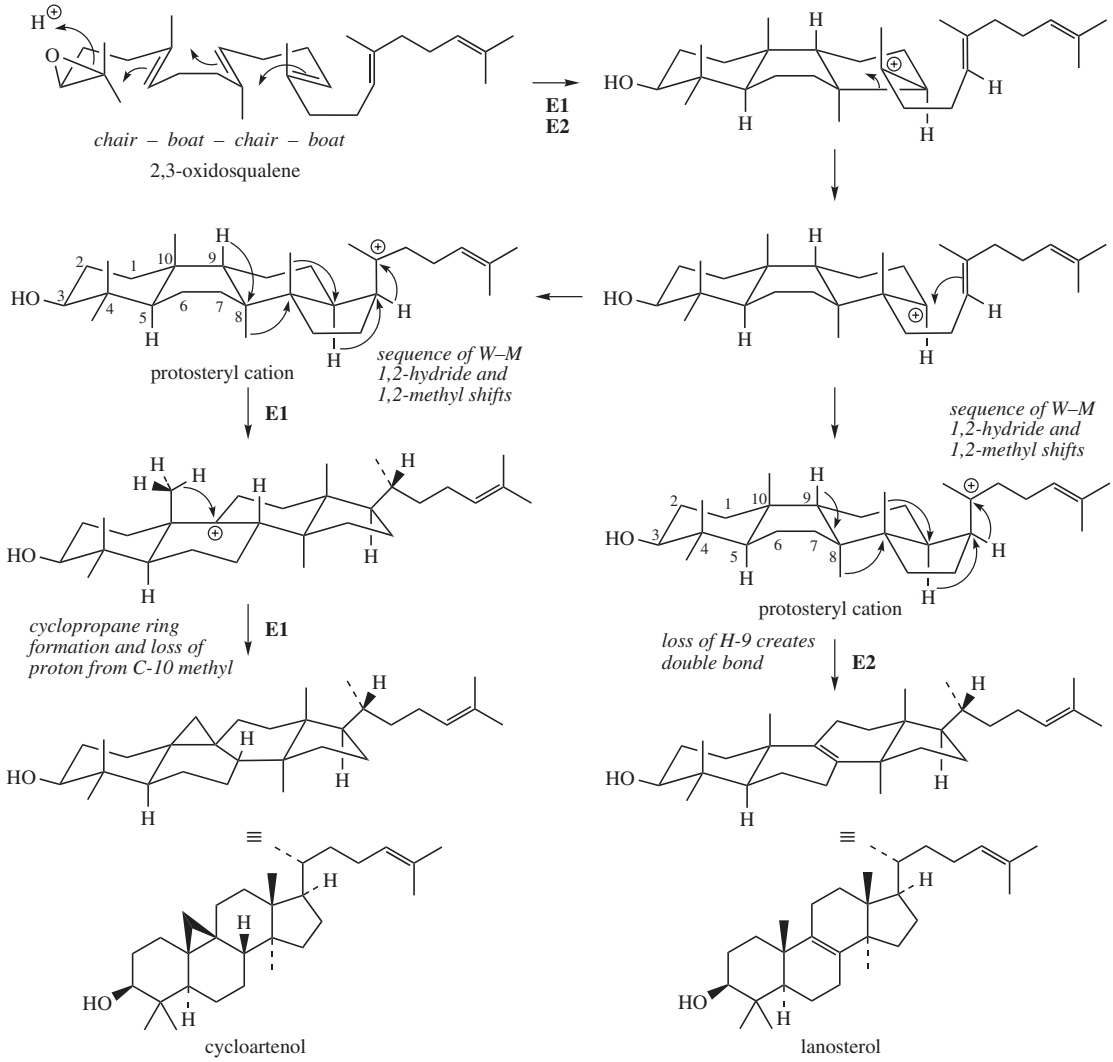
Breaking the original but now redundant C-1–C-2′ bond can give an allylic cation, and the generation of **squalene** is completed by supply of hydride from NADPH. Squalene synthase catalyses all steps in the sequence, the formation of presqualene PP from FPP, its subsequent rearrangement, and NADPH-dependent reduction.

Cyclization of squalene is via the intermediate **2,3-oxidosqualene** (squalene-2,3-oxide; Figure 5.61), produced in a reaction catalysed by squalene epoxidase, a flavoprotein requiring O_2 and NADPH cofactors. If oxidosqualene is suitably positioned and folded on the enzyme surface, then the polycyclic triterpene structures formed can be rationalized in terms of a series of cyclizations, followed by a sequence of concerted Wagner–Meerwein migrations of hydride and methyl groups (Figure 5.61). The cyclizations are carbocation-mediated and proceed in a stepwise sequence (Figure 5.62). Thus, protonation of the epoxide group will allow opening of this ring and generation of the preferred tertiary carbocation, suitably placed to allow electrophilic addition to a double bond, formation of a six-membered ring, and production of a new tertiary carbocation. This process continues twice more, generating the preferred tertiary carbocation (Markovnikov addition) after each ring formation, though the third ring formed is consequently a five-membered one. This is expanded to a six-membered ring via a Wagner–Meerwein 1,2-alkyl shift, resulting in some relief of ring strain, though sacrificing a tertiary carbocation for a secondary one. A further electrophilic addition generates the tertiary protosteryl cation (Figure 5.62). The stereochemistries in this cation are controlled by the type of folding achieved on the enzyme surface, and this probably also limits the extent of the cyclization process. Thus, if the folded oxidosqualene approximates to a *chair–boat–chair–boat* conformation (Figure 5.63), the transient **protosteryl cation** will be produced with these conformational characteristics. This cation then undergoes a series of Wagner–Meerwein 1,2-shifts, first migrating a hydride and generating a new cation, migrating the next hydride, then a methyl and so on until a proton is lost forming a double bond and thus creating **lanosterol** (Figure 5.63). The stereochemistry of the protosteryl cation in Figure 5.63 shows how favourable this sequence will be, and emphasizes that, in the ring system, the migrating groups are positioned *anti* to each other, one group entering whilst the other leaves from the opposite side of the stereocentre (compare diterpenes, page 233). This, of course, inverts configurations at each appropriate centre. No *anti* group is available to migrate to C-9 (steroid numbering), and the reaction terminates by loss of the proton H-9. Lanosterol is a typical animal triterpenoid, and the precursor for cholesterol and other sterols in animals (see page 248) and fungi (see page 254). In plants, its intermediate role is taken by **cycloartenol** (Figure 5.63), which contains a cyclopropane ring, generated by inclusion of carbon from the methyl at C-10. For cycloartenol, H-9 is not lost, but migrates to C-8, and the carbocation so formed is quenched by cyclopropane formation and loss of one of the methyl protons. For many plant steroids, this cyclopropane ring has then to be reopened (see page 251). Most natural triterpenoids and steroids contain a 3-hydroxyl group, the original epoxide oxygen from oxidosqualene.

An additional feature of the protosteryl cation is that the C-10 methyl and H-5 also share an *anti*-axial relationship, and are also susceptible to Wagner–Meerwein rearrangements, so that the C-9 cation formed in the cycloartenol sequence may then initiate further migrations. This can be terminated by formation of a 5,6-double bond (Figure 5.64), as in the pathway to the **cucurbitacins**, a group of highly oxygenated triterpenes encountered in the Cucurbitaceae, the cucumber/melon/marrow family. These compounds are characteristically bitter tasting, purgative, and extremely cytotoxic.

Should **oxidosqualene** be folded in a roughly *chair–chair–chair–boat* conformation by binding to another type of cyclase enzyme (Figure 5.65), then an identical carbocation mechanism ensues. However, the transient **dammarenyl cation** formed has different stereochemical features to the protosteryl cation. Whilst a series of Wagner–Meerwein migrations can occur, there is relatively little to be gained on purely chemical grounds, since these would invert stereochemistry and destroy what is already a very favourable conformation. Instead, the dammarenyl cation typically undergoes further carbocation-promoted cyclizations, without any major changes to the ring system already formed. Occasionally, the migrations do occur: **euphol** (Figure 5.65) from *Euphorbia* species (Euphorbiaceae) is a stereoisomer of lanosterol.

If the Wagner–Meerwein rearrangements do not take place, then the dammarenyl cation could be quenched with water (Figure 5.66), giving the epimeric **dammarenediols**, as found in Dammar resin from *Balanocarpus heimii* (Dipterocarpaceae) and ginseng (*Panax ginseng*; Araliaceae) [Box 5.11]. Alternatively, the alkyl migration shown gives the baccharenyl cation, relieving some ring strain by creating a six-membered ring, despite sacrificing a tertiary carbocation for a secondary one. A pentacyclic ring system can now be formed by cyclization onto the double bond, giving a new five-membered ring and the tertiary lupenyl cation. Although this appears to



E1: oxidosqualene:cycloartenol cyclase (cycloartenol synthase)
E2: oxidosqualene:lanosterol cyclase (lanosterol synthase)

Figure 5.63

contradict the reasoning used above for the dammarenyl → baccharenyl transformation, the contribution of the enzyme involved must also be considered in each case. A five-membered ring is not highly strained, as evidenced by all the natural examples encountered. Loss of a proton from the lupanyl cation gives **lupeol**, found in lupin (*Lupinus luteus*; Leguminosae/Fabaceae). Ring expansion in the lupanyl cation by bond migration gives the oleanyl system, and labelling studies have demonstrated this ion is discharged by hydride migrations and loss of a proton, giving the widely distributed **β-amyrin**. Formation of the isomeric **α-amyrin** involves first the migration of a methyl in the oleanyl cation, then discharge of the new taraxasteryl cation by three hydride migrations and loss of a proton. Loss of a proton from the non-migrated methyl in the taraxasteryl cation is an alternative way of achieving a neutral molecule, and yields **taraxasterol** found in dandelion (*Taraxacum officinale*; Compositae/Asteraceae). Comparison with α-amyrin shows the subtly different stereochemistry

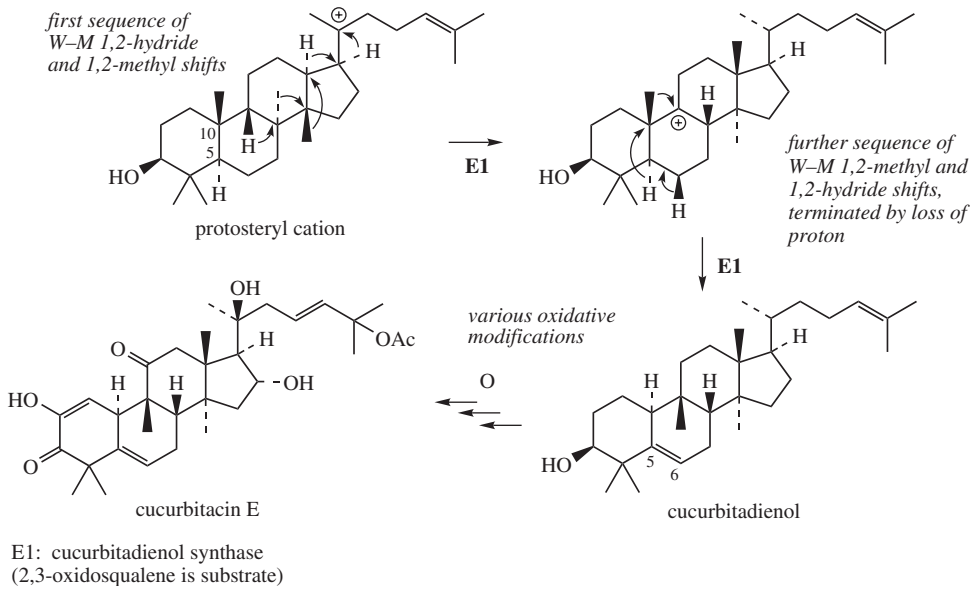


Figure 5.64

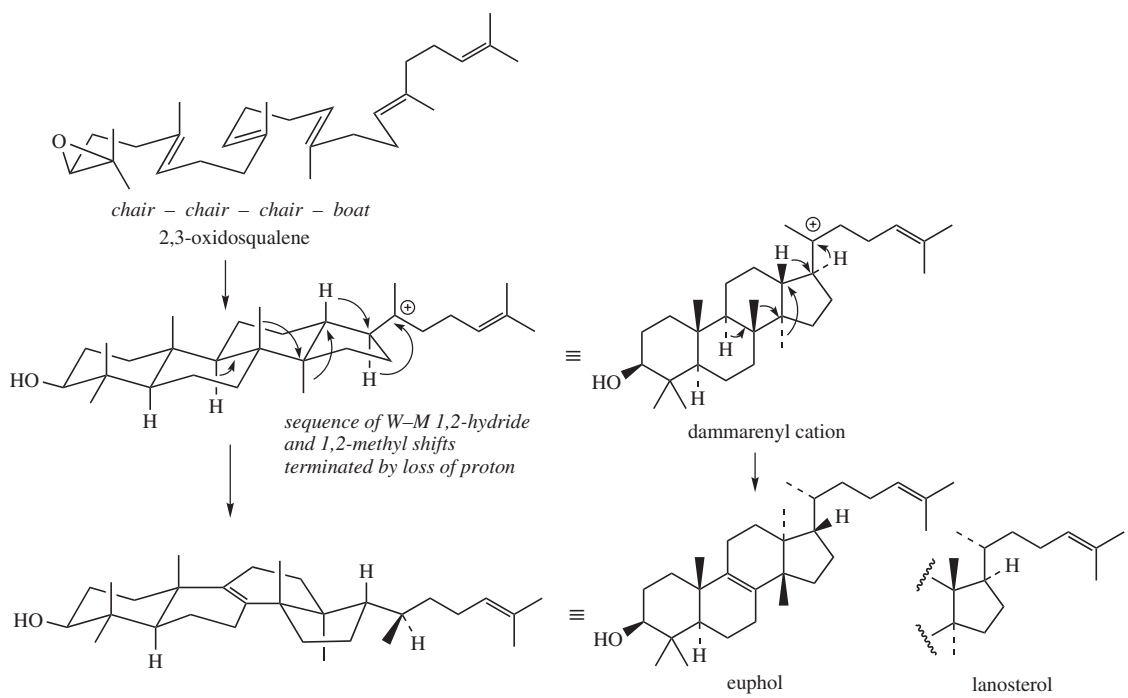


Figure 5.65

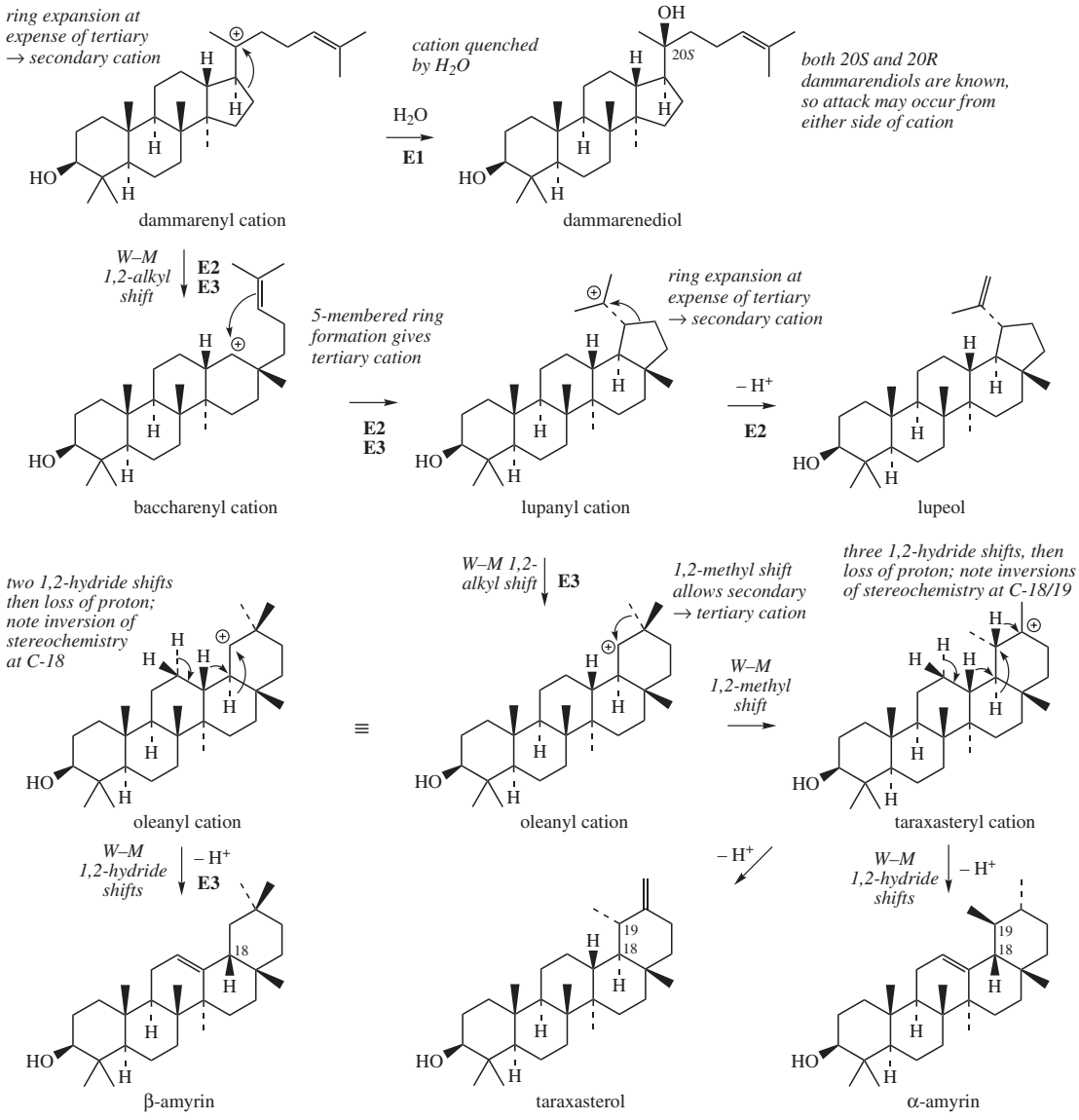


Figure 5.66

present, because the inversions of configuration caused by hydride migrations have not occurred. Where evidence is available, these extensive series of cyclizations and Wagner–Meerwein rearrangements appear to be catalysed by a single enzyme which converts oxidosqualene into the final product, e.g. lanosterol, cycloartenol, lupeol, or β-amyrin. Enzymes can be monofunctional, synthesizing a single product, or in some cases the triterpenoid cyclase is multifunctional, catalysing formation of a range of structurally related products, e.g. lupeol, α-amyrin, taraxasterol, and others. Remarkably, changing a single amino acid residue in lupeol synthase from olive (*Olea europaea*; Oleaceae) was sufficient to alter the protein's function; it became instead a producer of β-amyrin.

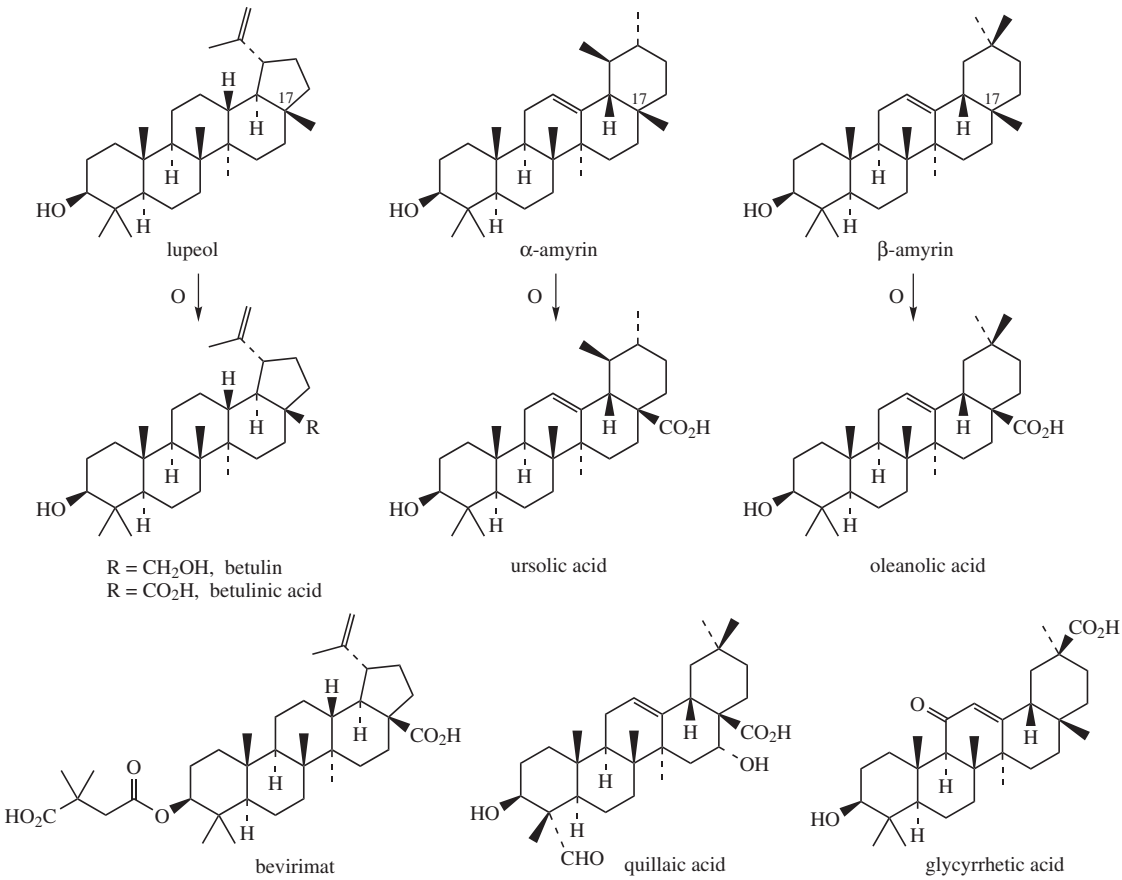


Figure 5.67

Further modification of these hydrocarbons to **triterpene acids** is frequently encountered (Figure 5.67). Olives (*Olea europaea*; Oleaceae) contain large quantities of **oleanolic acid**, in which the C-17 methyl of β-amyrin has been oxidized to a carboxylic acid. The leaves of olive contain a mixture of oleanolic acid and **betulinic acid**, the latter an oxidized version of lupeol. The corresponding derivative of α-amyrin is **ursolic acid**, found in bearberry (*Arctostaphylos uva-ursi*; Ericaceae). These are presumably produced by cytochrome P-450-dependent oxidations, but details are not known. Oxidative transformations at other methyl groups and/or ring carbon atoms are required to produce triterpenoids such as **glycyrrhetic acid** and **quillaic acid** (see triterpenoid saponins below). The alcohol **betulin** is a major component in the bark of white birch (*Betula alba*; Betulaceae), where it comprises up to 24% of the outer bark layer; large amounts of the rarer betulinic acid may be obtained by selective chemical oxidation of betulin. Derivatives of betulinic acid and betulin are currently attracting considerable interest in HIV therapeutics by inhibiting viral growth in a manner different from that of current HIV drugs. The most promising compound, currently in clinical trials, is the semi-synthetic 3′,3′-dimethylsuccinyl ester **bevirimat** (Figure 5.67).

Bacterial membranes frequently contain **hopanoids** (Figure 5.68), triterpenoid compounds that appear to take the place of the sterols that are typically found in the membranes of higher organisms, helping to maintain structural integrity and to control permeability. Hopanoids are also the characteristic triterpenes in ferns. Hopanoids arise from squalene by a similar carbocation cyclization mechanism, but do not involve the initial epoxidation to oxidosqualene. Instead, the carbocation is produced by protonation (compare the cyclization of GGPP to copalyl PP, page 228), and the resultant compounds tend to lack the characteristic 3-hydroxyl group, e.g. **hopene** from

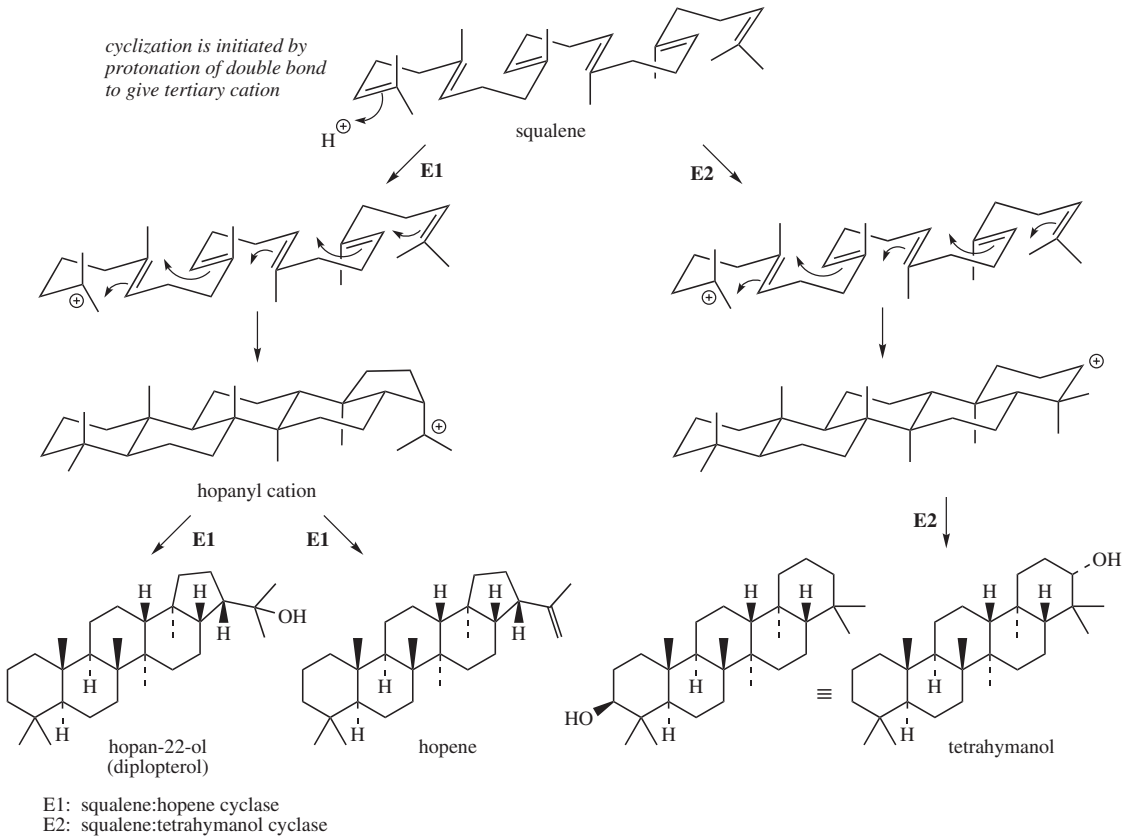
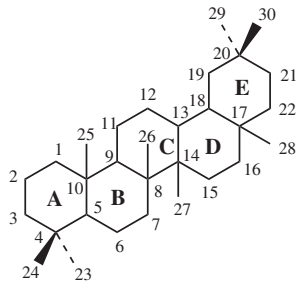


Figure 5.68

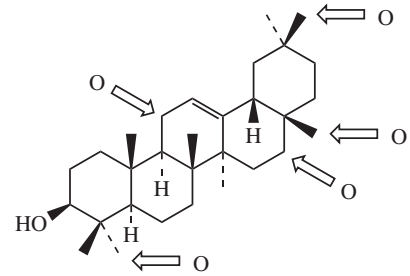
Alicyclobacillus acidocaldarius (Figure 5.68). On the other hand, **tetrahymanol** from the protozoan *Tetrahymena pyriformis*, because of its symmetry, might appear to have a 3-hydroxyl group, but this is derived from water and not molecular oxygen, as would be the case if oxidosqualene were involved. As in formation of the protosteryl cation (page 237), Markovnikov additions followed by Wagner–Meerwein ring expansions (rather than anti-Markovnikov additions) may occur during the cyclization mechanisms shown in Figure 5.68.

Triterpenoid Saponins

The pentacyclic triterpenoid skeletons exemplified by lupeol, α-amyrin, and β-amyrin (Figure 5.67) are frequently encountered in the form of triterpenoid saponin structures. Saponins are glycosides which, even at low concentrations, produce a frothing in aqueous solution, because they have surfactant and soap-like properties. The name comes from the Latin *sapo*, meaning soap, and plant materials containing saponins were originally used for cleansing clothes, e.g. soapwort (*Saponaria officinalis*; Caryophyllaceae) and quillaia or soapbark (*Quillaja saponaria*; Rosaceae). These materials also cause haemolysis, lysing red blood cells by increasing the permeability of the plasma membrane, and thus they are highly toxic when injected into the bloodstream. Some saponin-containing plant extracts have been used as arrow poisons. However, saponins are relatively harmless when taken orally, and some of our valuable food materials, e.g. beans, lentils, soybeans, spinach, and oats, contain significant amounts. Sarsaparilla (see page 264) is rich in steroidal saponins, but is widely used in the manufacture of non-alcoholic drinks. Toxicity is



pentacyclic triterpenoid skeleton



potential sites for oxidation (β-amyrin type)

Figure 5.69

minimized during ingestion by low absorption, and by hydrolysis. Acid-catalysed hydrolysis of saponins liberates sugar(s) and an aglycone (sapogenin) which can be either triterpenoid or steroidal (see page 259) in nature. Some plants may contain exceptionally high amounts of saponins, e.g. about 10% in quillaia bark.

Triterpenoid saponins are rare in monocotyledons, but abundant in many dicotyledonous families. Several medicinally useful examples are based on the β-amyrin subgroup (Figure 5.69), and many of these possess carboxylic acid groups derived by oxidation of methyl groups, those at positions 4 (C-23), 17 (C-28) and 20 (C-30) on the aglycone ring system being subject to such oxidation. Less oxidized formyl (–CHO) or hydroxymethyl ($-CH_2OH$) groups may be encountered, and positions 11 and 16 may also be oxygenated. Sugar residues are usually attached to the 3-hydroxyl, with one to six monosaccharide units, the most common being glucose, galactose, rhamnose, and arabinose, with uronic acid units (glucuronic acid and galacturonic acid) also featuring (see page 488). Thus, quillaia bark contains a saponin mixture with **quillaic acid** (Figure 5.67) as the principal aglycone [Box 5.11]. The medicinally valuable root of liquorice (*Glycyrrhiza glabra*; Leguminosae/Fabaceae) contains **glycyrrhizin**, a mixture of potassium and calcium salts of **glycyrrhizic acid** (Figure 5.70), which is composed of the aglycone **glycyrrhetic acid** and two glucuronic acid units [Box 5.11]. Ginseng is a herbal drug derived from the roots of *Panax ginseng* (Araliaceae) that is widely held to counter stress and improve general well-being. A group of saponins based on the dammarane skeleton and termed **ginsenosides** (see Figure 5.72) are most likely the biologically active components [Box 5.11].

Box 5.11

Liquorice

Liquorice (licorice; glycyrrhiza) is the dried unpeeled rhizome and root of the perennial herb *Glycyrrhiza glabra* (Leguminosae/Fabaceae). A number of different varieties are cultivated commercially, including *G. glabra* var. *typica* (Spanish liquorice) in Spain, Italy, and France, and *G. glabra* var. *glandulifera* (Russian liquorice) in Russia. Russian liquorice is usually peeled before drying. *G. uralensis* (Manchurian liquorice) from China is also commercially important. Much of the liquorice is imported in the form of an extract, prepared by extraction with water, then evaporation to give a dark black solid. Most of the liquorice produced is used in confectionery and for flavouring, including tobacco, beers, and stouts. Its pleasant sweet taste and foaming properties are due to saponins. Liquorice root contains about 20% of water-soluble extractives, and much of this (typically 3–5% of the root, but up to 12% in some varieties) is comprised of glycyrrhizin, a mixture of the potassium and calcium salts of glycyrrhizic (=glycyrrhizinic) acid (Figure 5.70). Glycyrrhizic acid is a diglucuronide of the aglycone glycyrrhetic (=glycyrrhetinic) acid. The bright yellow colour of liquorice root is provided by flavonoids (1–1.5%), including liquiritigenin and isoliquiritigenin and their corresponding glucosides (see page 169). Considerable amounts (5–15%) of sugars (glucose and sucrose) are also present.

Glycyrrhizin is reported to be 50–150 times as sweet as sucrose, and liquorice has thus long been used in pharmacy to mask the taste of bitter drugs. Its surfactant properties have also been exploited in various formulations, as have its demulcent

Box 5.11 (continued)

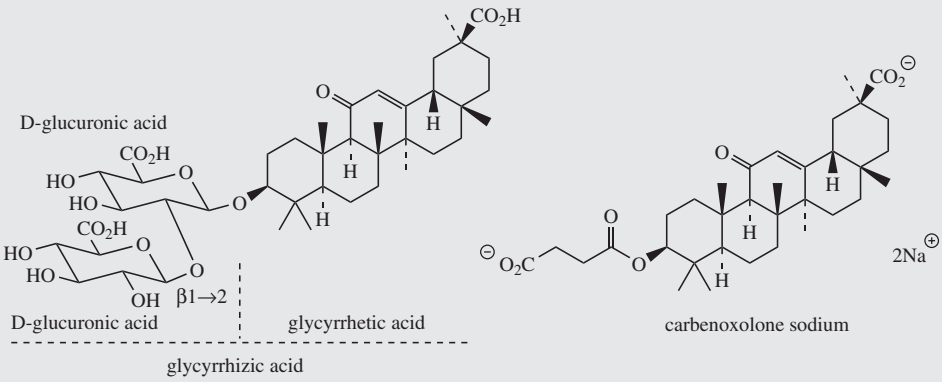


Figure 5.70

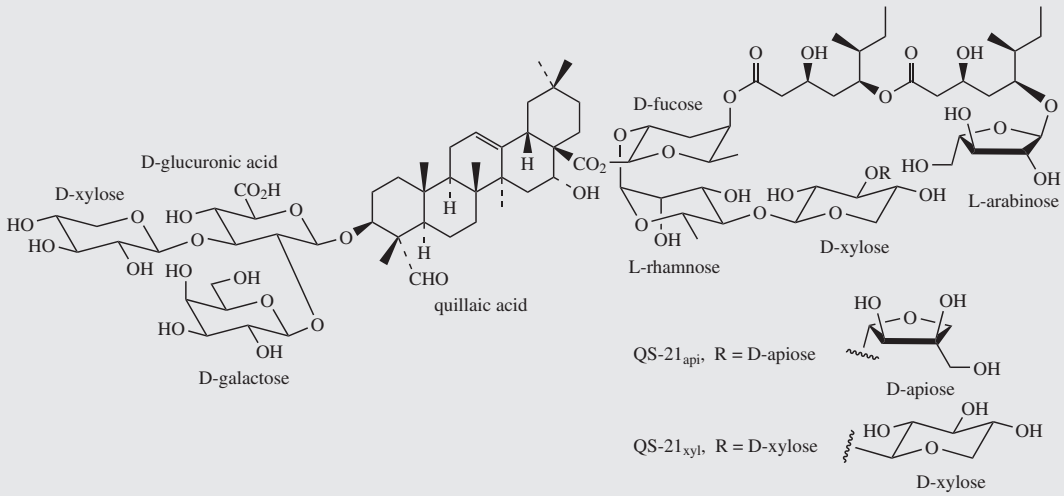


Figure 5.71

and mild expectorant properties. More recently, some corticosteroid-like activity has been recognized, with liquorice extracts displaying mild anti-inflammatory and mineralocorticoid activities. These have been exploited in the treatment of rheumatoid arthritis, Addison's disease (chronic adrenocortical insufficiency), and various inflammatory conditions. Glycyrrhetic acid has been implicated in these activities, and has been found to inhibit enzymes that catalyse the conversion of prostaglandins and glucocorticoids into inactive metabolites. This results in increased levels of prostaglandins, e.g. PGE_2 and $PGF_{2\alpha}$ (see page 60), and of hydrocortisone (see page 278). A semi-synthetic derivative of glycyrrhetic acid, the hemisuccinate **carbenoxolone sodium** (Figure 5.70), has been widely prescribed for the treatment of gastric ulcers, and also duodenal ulcers. Because of side-effects, typically loss of potassium and increase in sodium levels, it has been superseded by newer drugs.

Quillaia

Quillaia bark or soapbark is derived from the tree *Quillaja saponaria* (Rosaceae) and other *Quillaja* species found in Chile, Peru, and Bolivia. The bark contains up to 10% saponins, a mixture known as 'commercial saponin' which is used as a foaming agent

Box 5.11 (continued)

in beverages and emulsifier in foods. Quillaia's surfactant properties are occasionally exploited in pharmaceutical preparations, where in the form of quillaia tincture it is used as an emulsifying agent, particularly for fats, tars, and volatile oils. The bark contains a mixture of saponins which on hydrolysis liberates quillaic acid (Figure 5.67) as the aglycone, together with sugars, uronic acids, and acids from ester functions.

Saponins from quillaia are also showing great promise as immunoadjuvants, substances added to vaccines and other immunotherapies designed to enhance the body's immune response to the antigen. One such agent, QS-21A (Figure 5.71), is a mixture of the two saponins QS-21_{api} and QS-21_{xyl}, each incorporating a quillaic acid triterpenoid core, flanked on either side by complex oligosaccharides, one of which includes a chain comprised of two fatty acyl components. These compounds were isolated from the 21st chromatographic fraction of the *Q. saponaria* extract.

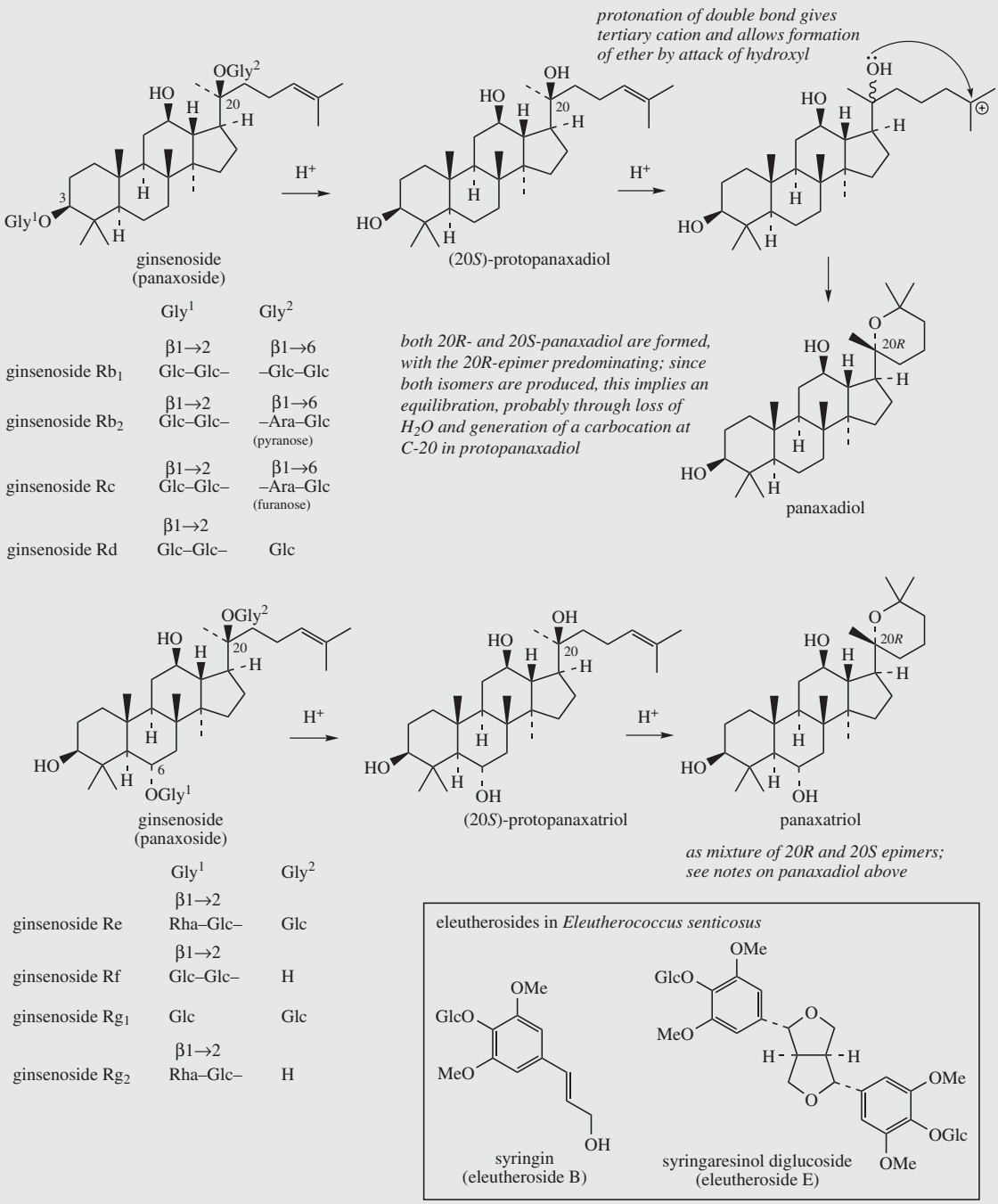
Ginseng

The roots of the herbaceous plants *Panax ginseng* (Araliaceae) from China, Korea and Russia, and related *Panax* species, e.g. *P. quinquefolium* (American ginseng) from the USA and Canada and *P. notoginseng* (Sanchi-ginseng) from China, have been widely used in China and Russia for the treatment of a number of diseases, including anaemia, diabetes, gastritis, insomnia, sexual impotence, and as a general restorative, promoting health and longevity. Interest in the drug has increased considerably in recent years, and ginseng is widely available as a health food in the form of powders, extracts, and teas. The dried and usually peeled root provides white ginseng, whereas red ginseng is obtained by steaming the root, this process generating a reddish-brown caramel-like colour, and reputedly enhancing biological activity. **Ginseng** is classified as an 'adaptogen', an agent that helps the body to adapt to stress, improving stamina and concentration, and providing a normalizing and restorative effect. It is also widely promoted as an aphrodisiac. The Korean root is highly prized and the most expensive. Long-term use of ginseng can lead to symptoms similar to those of corticosteroid poisoning, including hypertension, nervousness, and sleeplessness in some subjects, yet hypotension and tranquillizing effects in others.

The benefits of ginseng treatment are by no means confirmed at the pharmacological level, though ginseng does possess antioxidant activity, can affect both central nervous system and neuroendocrine functions, can alter carbohydrate and lipid metabolism, and modulates immune function. Many of the secondary metabolites present in the root have now been identified. It contains a large number of triterpenoid saponins based on the dammarane subgroup, saponins that have been termed ginsenosides by Japanese investigators, or panaxosides by Russian researchers. These are derivatives of two main aglycones, protopanaxadiol and protopanaxatriol (Figure 5.72), though the aglycones liberated on acid hydrolysis are panaxadiol and panaxatriol respectively. Acid-catalysed cyclization in the side-chain produces an ether ring (Figure 5.72). Sugars are present in the saponins on the 3- and 20-hydroxyl groups in the diol series, and the 6- and 20-hydroxyl groups in the triol series. Over 30 ginsenosides have been characterized from the different varieties of ginseng, with ginsenoside Rb_1 (Figure 5.72) of the diol series typically being the most abundant constituent. Ginsenoside Rg_1 (Figure 5.72) is usually the major component representative of the triol series. Some other variants are shown in Figure 5.72. Particularly in white ginseng, many of the ginsenosides are also present as esters with malonic acid. Steaming to prepare red ginseng causes partial hydrolysis of esters and glycosides; there is undoubtedly greater antioxidant activity in the steamed product, though this is derived mainly from various phenolic constituents. Ginsenosides Rb_1 and Rg_1 appear to be the main representatives in *P. ginseng*, ginsenosides Rb_1, Rg_1, and Rd in *P. notoginseng*, and ginsenosides Rb_1, Re, and malonylated Rb_1 in *P. quinquefolium*. The pentacyclic triterpenoid sapogenin oleanolic acid (Figure 5.67) is also produced by hydrolysis of the total saponins of *P. ginseng*, and is present in some saponin structures (chikusetsusaponins). The saponin contents of *P. notoginseng* (about 12%) and *P. quinquefolium* (about 6%) are generally higher than that of *P. ginseng* (1.5–2%). Pharmacological studies on individual ginsenosides are being facilitated by development of enzymic procedures for interconverting the structures, e.g. by selective hydrolysis of sugar groups, or by selective glycosylations. The aglycone mixture of protopanaxadiol and protopanaxatriol (**pandimex**®) is being tested clinically as an anticancer agent.

The root of *Eleutherococcus senticosus* (*Acanthopanax senticosus*) (Araliaceae) is used as an inexpensive substitute for ginseng, and is known as **Russian** or **Siberian ginseng**. This material is held to have adaptogenic properties similar to *P. ginseng*, and a number of eleutherosides have been isolated. However, the term eleutheroside has been applied to compounds of different chemical classes, and the main active anti-stress constituents appear to be lignan glycosides, e.g. eleutheroside E (≡syringaresinol diglucoside; Figure 5.72) (compare page 152) and phenylpropane glycosides, e.g. eleutheroside B (≡syringin). The leaves of Russian ginseng contain a number of saponins based on oleanolic acid (Figure 5.67), but these are quite different to the ginsenosides/panaxosides found in *Panax*. Whilst there is sufficient evidence to support the beneficial adaptogen properties for *E. senticosus*, detailed pharmacological confirmation is not available.

Box 5.11 (continued)



	Gly[1]	Gly[2]
	β1→2	β1→6
ginsenoside Rb_1	Glc–Glc–	–Glc–Glc
	β1→2	β1→6
ginsenoside Rb_2	Glc–Glc–	–Ara–Glc (pyranose)
	β1→2	β1→6
ginsenoside Rc	Glc–Glc–	–Ara–Glc (furanose)
	β1→2	
ginsenoside Rd	Glc–Glc–	Glc

	Gly[1]	Gly[2]
	β1→2	
ginsenoside Re	Rha–Glc–	Glc
	β1→2	
ginsenoside Rf	Glc–Glc–	H
ginsenoside Rg_1	Glc	Glc
	β1→2	
ginsenoside Rg_2	Rha–Glc–	H

Figure 5.72

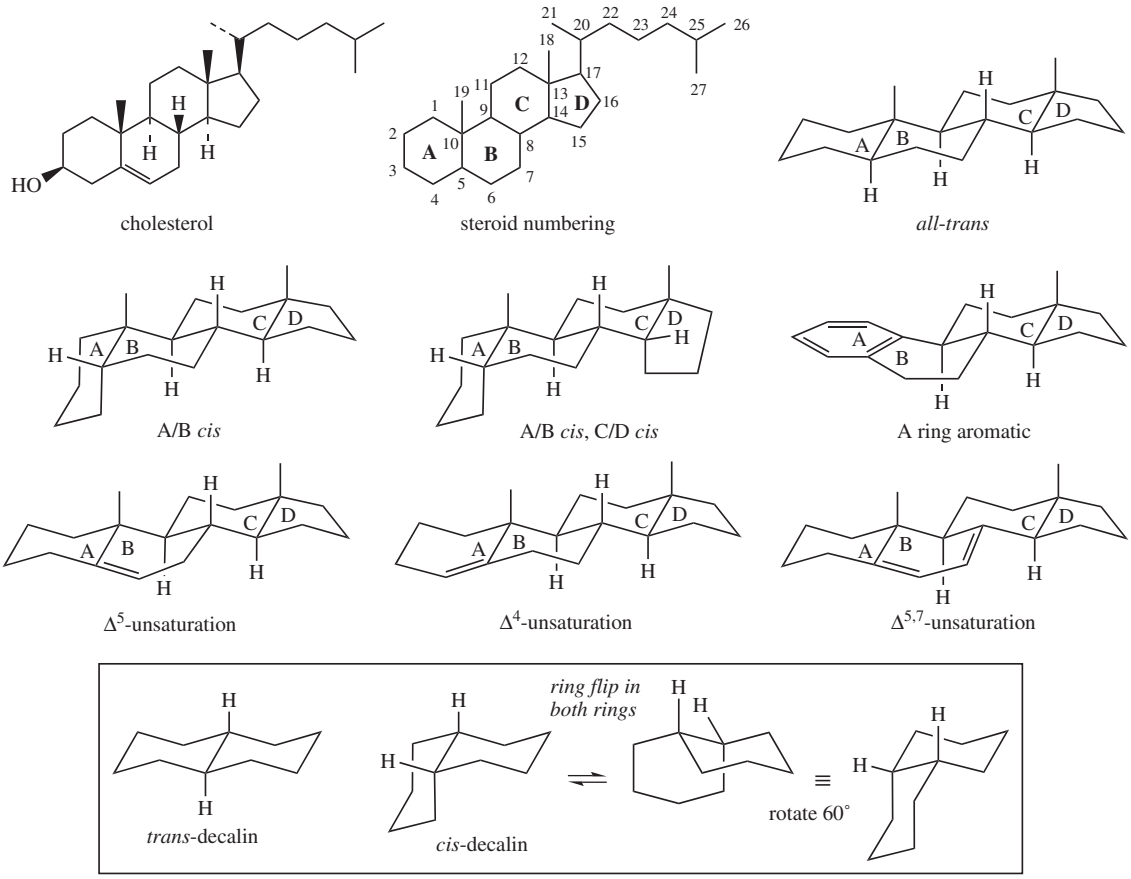


Figure 5.73

STEROIDS

The steroids are modified triterpenoids containing the tetracyclic ring system of lanosterol (Figure 5.61), but lacking the three methyl groups at C-4 and C-14. **Cholesterol** (Figure 5.73) typifies the fundamental structure, but further modifications, especially to the side-chain, help to create a wide range of biologically important natural products, e.g. sterols, steroidal saponins, cardioactive glycosides, bile acids, corticosteroids, and mammalian sex hormones. Because of the profound biological activities encountered, many natural steroids and a considerable number of synthetic and semi-synthetic steroidal compounds are routinely employed in medicine. The markedly different biological activities observed emanating from compounds containing a common structural skeleton are in part ascribed to the functional groups attached to the steroid nucleus, and in part to the overall shape conferred on this nucleus by the stereochemistry of ring fusions.

Stereochemistry and Nomenclature

Ring systems containing six-membered or five-membered rings can be *trans*-fused as exemplified by *trans*-decalin or *cis*-fused as in *cis*-decalin (Figure 5.73). The *trans*-fusion produces a flattish molecule when two chair conformations are present. The only conformational mobility allowed is to less favourable boat forms. Bridgehead hydrogen atoms (or other substituents) are axial to both of the rings. In contrast, the *cis*-fused decalin is basically a bent molecule; it is found to be flexible, in that alternative conformers are possible, with both rings still being in chair form. Bridgehead substituents are axial to one ring, whilst being equatorial to the other, in each conformer. However, this conformational flexibility will be lost if either ring is then fused to a third ring, so is not encountered in steroid structures

In natural steroids, there are examples of the A/B ring fusion being *trans* or *cis*, or having unsaturation, either

Δ^4 or Δ^5. In some compounds, notably the oestrogens, ring A can even be aromatic; clearly, there can then be no bridgehead substituent at C-10 and, therefore, the normal C-10 methyl (C-19) must be lost. All natural steroids have a *trans* B/C fusion, though *cis* forms can be made synthetically. The C/D fusion is also usually *trans*, though there are notable exceptions, such as the cardioactive glycosides. Of course, such comments apply equally to some of the triterpenoid structures already considered. However, it is in the steroid field where the relationship between stereochemistry and biological activity is most marked. The overall shapes of some typical steroid skeletons are shown in Figure 5.73.

Systematic **steroid nomenclature** is based on a series of parent hydrocarbons, including **gonane**, **estrane**, **androstane**, **pregnane**, **cholane**, **cholestane**, **ergostane**, **campestane**, **stigmastane**, and **poriferastane** (Figure 5.74). The triterpenoid hydrocarbons **lanostane** and **cycloartane** are similarly used in systematic nomenclature and are also included in Figure 5.74. It is usual to add only unsaturation (ene/yne) and the highest priority functional group as suffixes to the root name; other groups are added as prefixes. Stereochemistry of substituents is represented by α (on the lower face of the molecule when it is drawn according to customary conventions as in Figure 5.74), or β (on the upper face). Ring fusions may be designated by using α or β for the appropriate bridgehead hydrogen, particularly those at positions 5 and 14, which will define the A/B and C/D fusions respectively, e.g. 5β-cholestane has the A/B rings *cis*-fused. Since the parent hydrocarbon assumes that ring fusions are *trans*, the stereochemistry for ring fusions is usually only specified where it is *cis*. Cholesterol is thus cholest-5-en-3β-ol. The term *nor* is affixed to indicate loss of a carbon atom, e.g. 19-norsteroids (see page 288) lack C-19 (the methyl at C-10).

Cholesterol

In animals, the triterpenoid alcohol **lanosterol** (C_{30}) is converted into **cholesterol** (C_{27}) (Figure 5.75), a process that, as well as the loss of three methyl groups, requires reduction of the side-chain double bond, and generation of a $\Delta^{5,6}$ double bond in place of the $\Delta^{8,9}$ double bond. The sequence of these steps is, to some extent, variable and

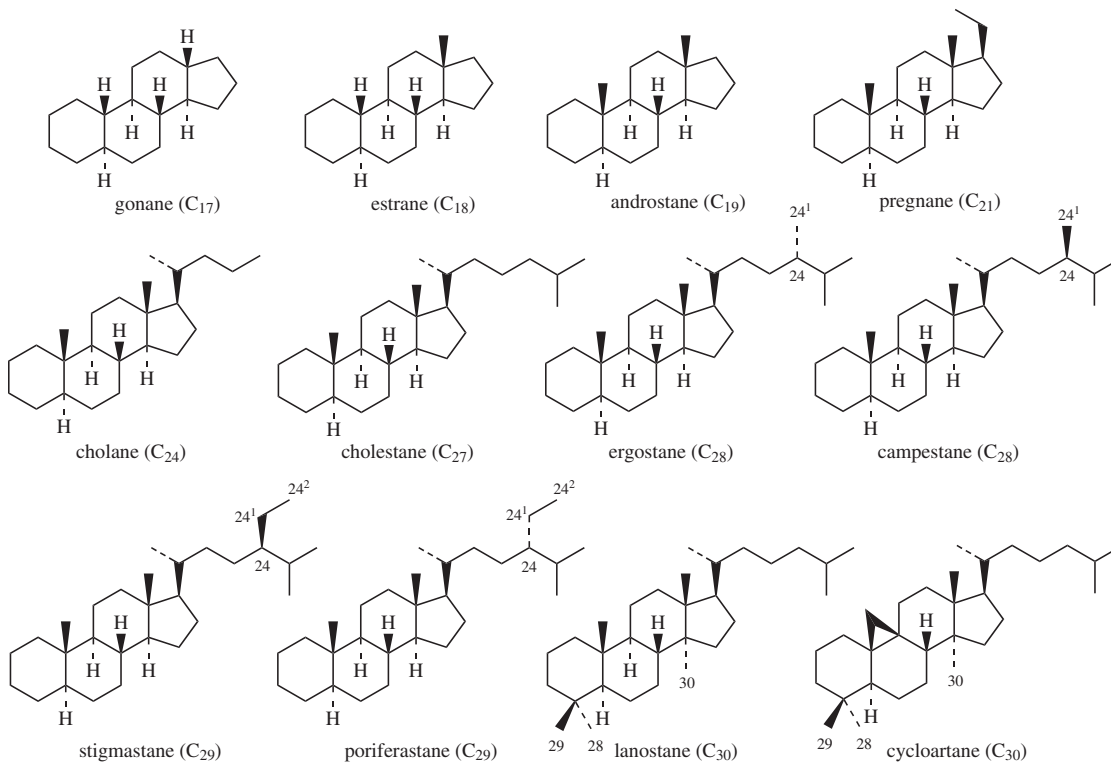


Figure 5.74

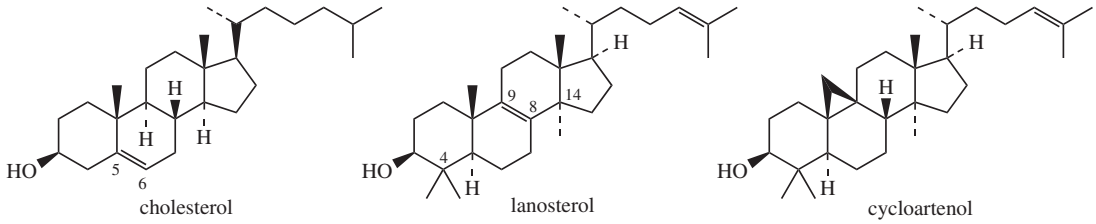


Figure 5.75

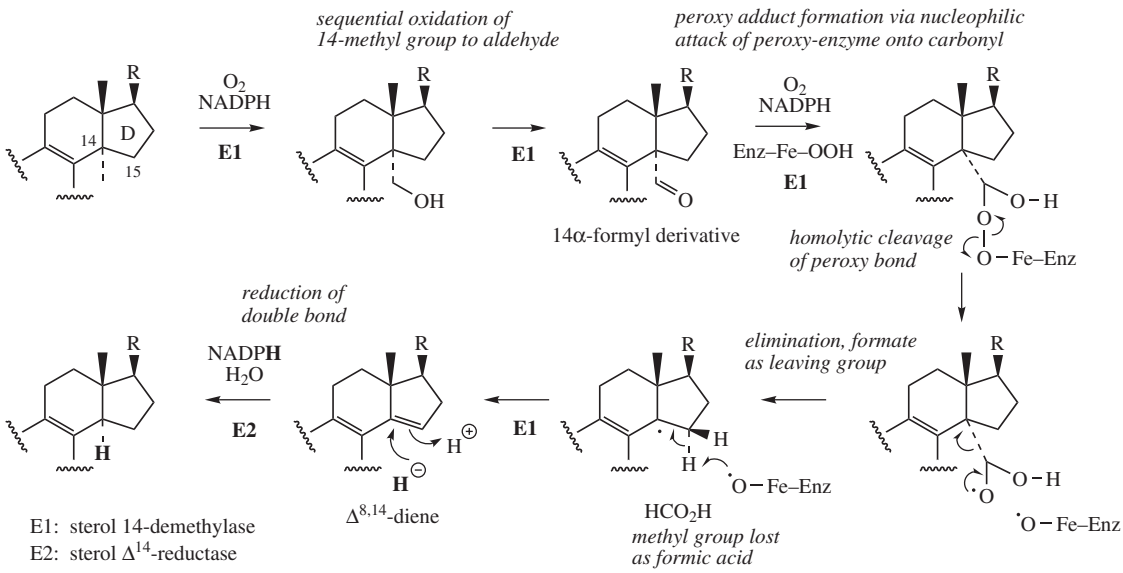


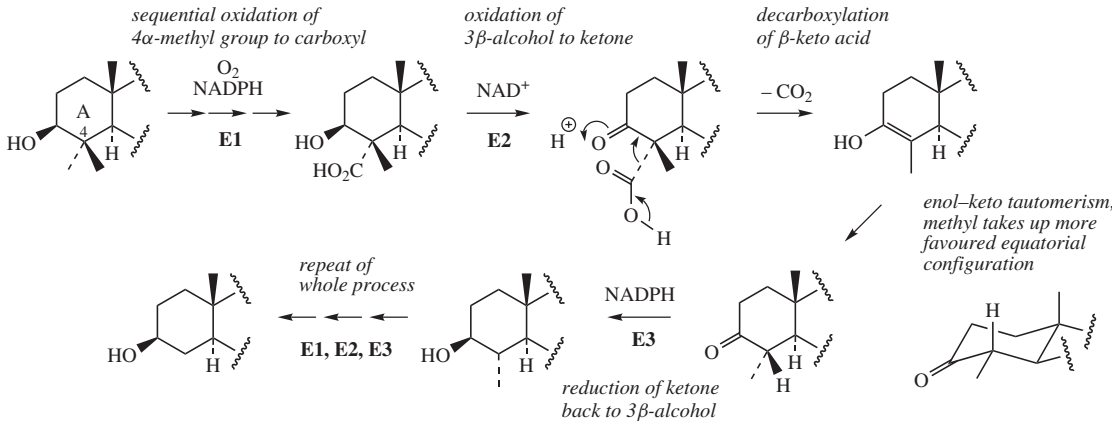
Figure 5.76

dependent on the organism involved. Accordingly, these individual transformations are considered rather than the overall pathway.

The methyl at C-14 is usually the one lost first, and this is removed as formic acid. The reaction is catalysed by a cytochrome P-450 monooxygenase which achieves two oxidation reactions to give the 14α-formyl derivative (Figure 5.76), and loss of this formyl group giving the $\Delta^{8,14}$ diene, most probably via homolytic cleavage of the peroxy adduct as indicated (compare similar peroxy adducts and mechanisms involved in side-chain cleavage from ring D, page 292, and in A ring aromatization, page 292). The 14-demethyl sterol is then obtained by an NADPH-dependent reduction step, the 15-proton being derived from water.

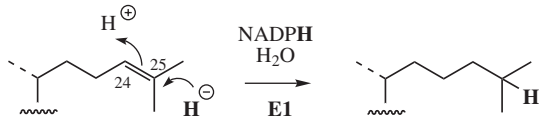
Loss of the C-4 methyl groups occurs sequentially, usually after removal of the 14α-methyl. Both carbon atoms are oxidized to carboxyl groups, then cleaved off via a decarboxylation mechanism (Figure 5.77). This is facilitated by oxidizing the 3-hydroxyl to a ketone, thus generating intermediate β-keto acids. In this sequence, the enolate is restored to a ketone in which the remaining C-4 methyl takes up the more favourable equatorial (4α) orientation. In animals and yeasts, the two methyl groups are removed by the same enzyme systems; plants are found to employ a different set of enzymes for the removal of the α- and β-methyl groups.

The side-chain Δ^{24} double bond is reduced by an NADPH-dependent reductase, hydride from the coenzyme being added at C-25, with H-24 being derived from water (Figure 5.78). The Δ^{8} double bond is effectively migrated to Δ^{5} via Δ^{7} and the $\Delta^{5,7}$ diene (Figure 5.79). This sequence involves an allylic isomerization, a dehydrogenation, and a reduction. Newly introduced protons at C-9 and C-8 originate from water, and that at C-7 from NADPH.



E1+E2+E3: sterol 4α-methyl oxidase (sterol C-4 demethylase)
E1: methylsterol monooxygenase
E2: sterol-4α-carboxylate 3-dehydrogenase (decarboxylating)
E3: 3-ketosteroid reductase

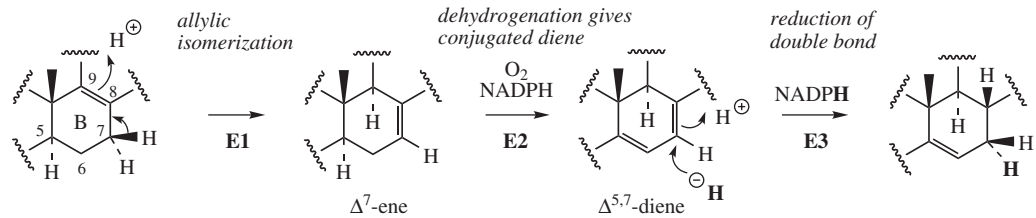
Figure 5.77



E1: sterol Δ^{24}-reductase

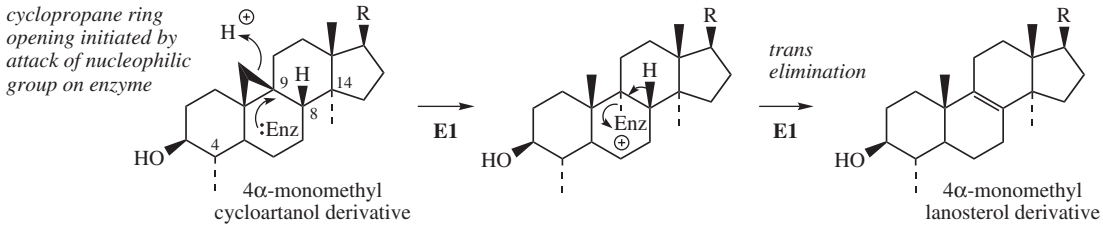
Figure 5.78

The role of lanosterol in non-photosynthetic organisms (animals, fungi) is taken in photosynthetic organisms (plants, algae) by the cyclopropane triterpenoid **cycloartenol** (Figure 5.75). This cyclopropane feature is found in a number of plant sterols, though the majority of plant steroids contain the typical C-10 methyl group. This means that, in addition to the lanosterol → cholesterol modifications outlined above, a further mechanism to reopen the cyclopropane ring is necessary. This is shown in Figure 5.80. Acid-catalysed ring opening is the most likely mechanism, and this can be achieved non-enzymically, but under severe conditions. It is suggested, therefore, that a nucleophilic group from the enzyme attacks C-9, opening the cyclopropane ring and incorporating a proton from water. A *trans* elimination then generates the Δ^8 double bond. The stereochemistry at C-8 (Hβ) is unfavourable for a concerted mechanism involving loss of H-8 with cyclopropane ring opening. The cyclopropane ring-opening process seems to occur only with 4α-monomethyl sterols. In plants, removal of the first 4-methyl group (4α; note the remaining 4β-methyl group then takes up the α-orientation as in Figure 5.77) is also known to precede loss of the 14α-methyl. Accordingly, the general substrate shown in Figure 5.80 has both 4α- and 14α-methyl groups;



E1: sterol Δ^8-Δ^7-isomerase (cholestenol Δ-isomerase)
E2: Δ^7-sterol Δ^5-dehydrogenase (lathosterol oxidase)
E3: sterol Δ^7-reductase (7-dehydrocholesterol reductase)

Figure 5.79



E1: cycloeucalenol cycloisomerase (cycloeucalenol-obtusifoliol isomerase)

Figure 5.80

side-chain alkylation (see page 252) also commonly precedes these other modifications. The specificity of the cyclopropane ring-opening enzyme means cycloartenol is not converted into lanosterol, and lanosterol is thus absent from virtually all plant tissues. Cholesterol [Box 5.12] is almost always present in plants, though in only trace amounts, and is formed via cycloartenol.

Phytosterols

The major sterol found in mammals is the C_{27} compound cholesterol, which acts as a precursor for other steroid structures such as sex hormones and corticosteroids. The main sterols in plants, fungi, and algae are characterized by extra one-carbon or two-carbon substituents on

Box 5.12

Cholesterol

Cholesterol (Figure 5.75) is the principal animal sterol, and since it is a constituent of cell membranes has been found in all animal tissues. It maintains membrane fluidity, microdomain structure, and permeability. It functions as a precursor for steroid hormones (pages 279, 291), bile acids (page 276), vitamin D (page 257), and lipoproteins, but is also correlated with cardiovascular disease, atherosclerosis, hypercholesterolaemia, and gallstone disease. Human gallstones are almost entirely composed of cholesterol precipitated from the bile.

Although the processes involved are quite complex, there appears to be a clear correlation between human blood cholesterol levels and heart disease. Atherosclerosis is a hardening of the arteries caused by deposition of cholesterol, cholesterol esters, and other lipids in the artery wall, causing a narrowing of the artery and, thus, an increased risk of forming blood clots (thrombosis). Normally, most of the cholesterol serves a structural element in cell walls, whilst the remainder is transported via the blood and is used for synthesis of steroid hormones, vitamin D, or bile acids. Transport of cholesterol is facilitated by formation of lipoprotein carriers, comprising protein and phospholipid shells surrounding a core of cholesterol, in both free and esterified forms. Low density lipoproteins (LDLs) carry larger amounts of cholesterol and deposit it around the body. High-density lipoproteins (HDLs) pick up free cholesterol and return it to the liver to be degraded. Risk of atherosclerosis increases with increasing levels of LDL cholesterol, and is reduced with increasing levels of HDL cholesterol. Blood LDL cholesterol levels are thus a good statistical indicator of the potential risk of a heart attack. Current recommendations are that total cholesterol levels in blood should be lower than 4 mmol/l, of which LDL cholesterol concentration should be less than 2 mmol/l. The risks can be lessened by avoiding foods rich in cholesterol, e.g. eggs, reducing the intake of foods containing high amounts of saturated fatty acids such as animal fats, and replacing these with vegetable oils and fish that are rich in polyunsaturated fatty acids (see page 49). Blood LDL cholesterol levels may also be reduced by incorporating into the diet plant sterol esters or plant stanol esters, which reduce the absorption of cholesterol (see page 255). In humans, dietary cholesterol is actually a smaller contributor to LDL cholesterol levels than is dietary saturated fat. Cholesterol biosynthesis may also be inhibited by drug therapy using specific inhibitors of HMG-CoA reductase in the mevalonate pathway, e.g. lovastatin and related compounds (see page 98).

Cholesterol is one of the primary sources for the semi-synthesis of medicinal steroids. It is currently available in quantity via the brains and spinal cords of cattle as a by-product of meat production. Large quantities are also extractable from lanolin, the fatty material coating sheep's wool. This is a complex mixture of esters of long-chain fatty acids (including straight-chain, branched-chain, and hydroxy acids) with long-chain aliphatic alcohols and sterols. Cholesterol is a major sterol component. Saponification of crude lanolin gives an alcohol fraction (lanolin alcohols or wool alcohols) containing about 34% cholesterol and 38% lanosterol/dihydrolanosterol. Wool alcohols are also used as an ointment base.

the side-chain, attached at C-24. These substituent carbon atoms are numbered 24^1 and 24^2 (Figure 5.74); some older publications may use 28 and 29. The widespread plant sterols **campesterol** and **sitosterol** (Figure 5.81) are respectively 24-methyl and 24-ethyl analogues of cholesterol. **Stigmasterol** contains additional unsaturation in the side-chain, a *trans*-Δ^{22} double bond, a feature seen in many plant sterols, but never in mammalian ones. The introduction of methyl and ethyl groups at C-24 generates a new chiral centre, and the 24-alkyl groups in campesterol, sitosterol, and stigmasterol are designated α. The predominant sterol found in fungi is **ergosterol** (Figure 5.81),

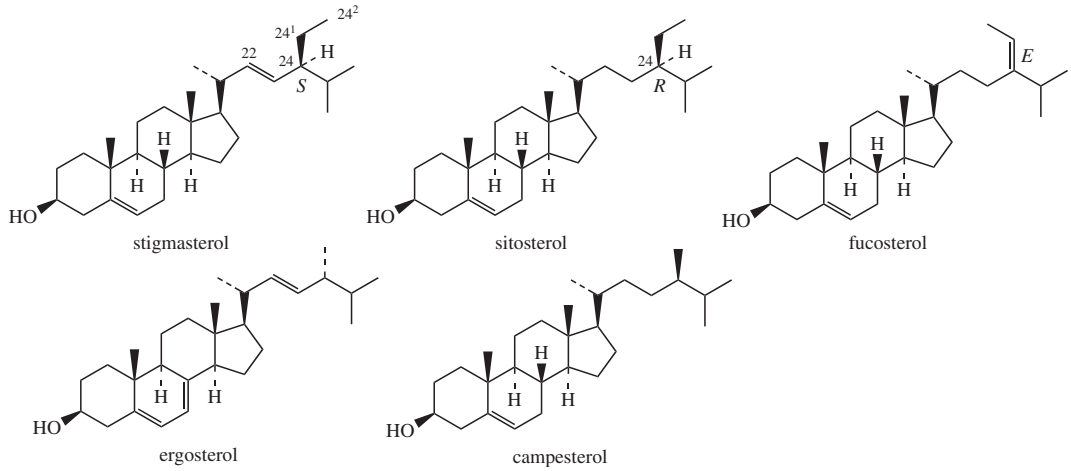


Figure 5.81

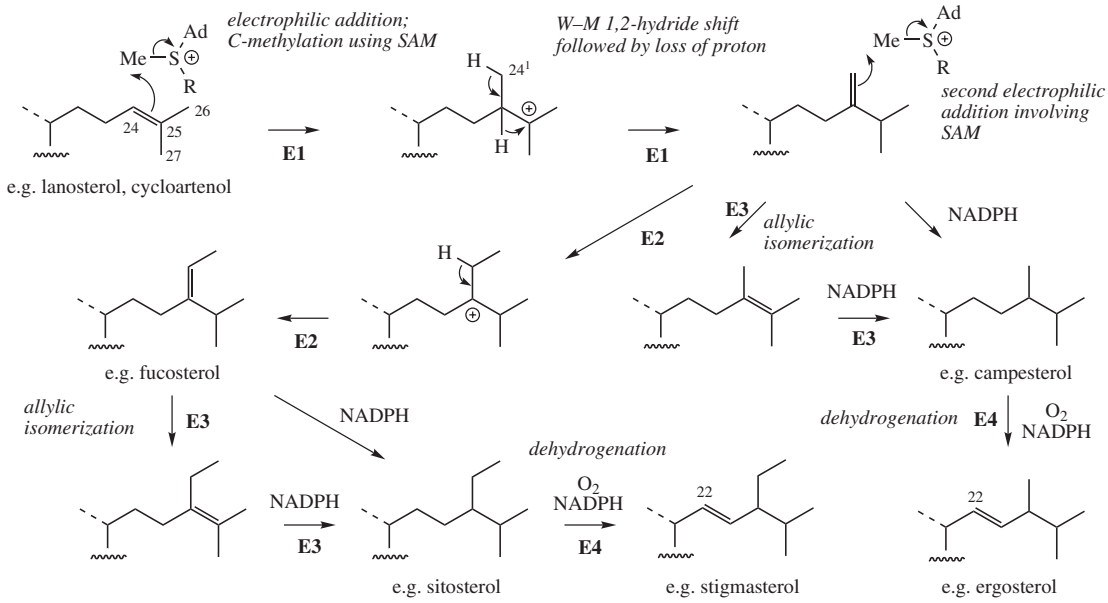
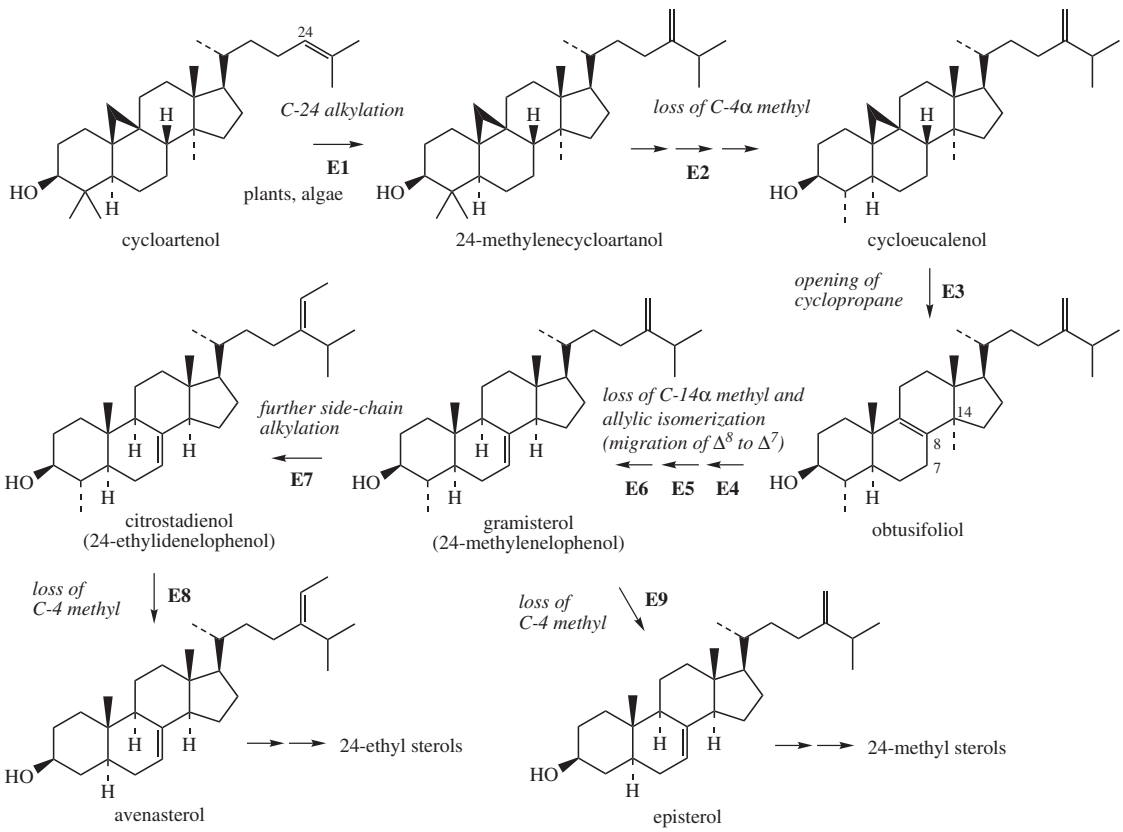


Figure 5.82

which has a β-oriented 24-methyl, as well as a *trans*-Δ^{22} double bond and additional Δ^7 unsaturation. The descriptors α and β unfortunately, and also confusingly, do not relate to similar terms used with the steroid ring system, but are derived from consideration of Fischer projections for the side-chain, substituents to the left being designated α and those to the right as β. Systematic *RS* nomenclature is thus preferred, but note that this defines sitosterol as 24*R*, whilst stigmasterol, because of its extra double bond, is 24*S*. The majority of plant sterols have a 24α-methyl or 24α-ethyl substituent, whilst algal sterols tend to have 24β-ethyl groups, and fungi 24β-methyl groups. The most abundant sterol in brown algae (*Fucus* spp.; Fucaceae) is **fucosterol** (Figure 5.81), which demonstrates a further variant, a 24-ethylidene substituent. Such groups can have *E*-configurations, as in fucosterol, or the alternative *Z*-configuration. Sterols are found predominantly in free alcohol form, but also as esters with long-chain fatty acids (e.g. palmitic, oleic, linoleic, and α-linolenic acids), as glycosides, and as glycosides acylated with fatty acids. These sterols, termed phytosterols, are structural components of membranes in plants, algae, and fungi and affect the permeability of these membranes. They also appear to play a role in cell proliferation.

The source of the extra methyl or ethyl side-chain carbons in both cases is SAM, and to achieve alkylation, the side-chain must have a Δ^{24} double bond, i.e. the side-chains as seen in lanosterol and cycloartenol. The precise mechanisms involved have been found to vary according to organism, but some of the demonstrated sequences are given in Figure 5.82. Methylation of the Δ^{24} double bond at C-24 via SAM yields a carbocation which undergoes a hydride shift and loss of a proton from C-24[1] to generate the 24-methylene side-chain. This can be



E1: sterol 24-*C*-methyltransferase
E2, E8, E9: sterol 4α-methyl oxidase (sterol C-4 demethylase)
E3: cycloeucalenol cycloisomerase (cycloeucalenol-obtusifoliol isomerase)
E4: sterol C-14 demethylase
E5: sterol Δ^{14}-reductase
E6: sterol Δ^8-Δ^7-isomerase (cholestenol Δ-isomerase)
E7: 24-methylenesterol *C*-methyltransferase

Figure 5.83

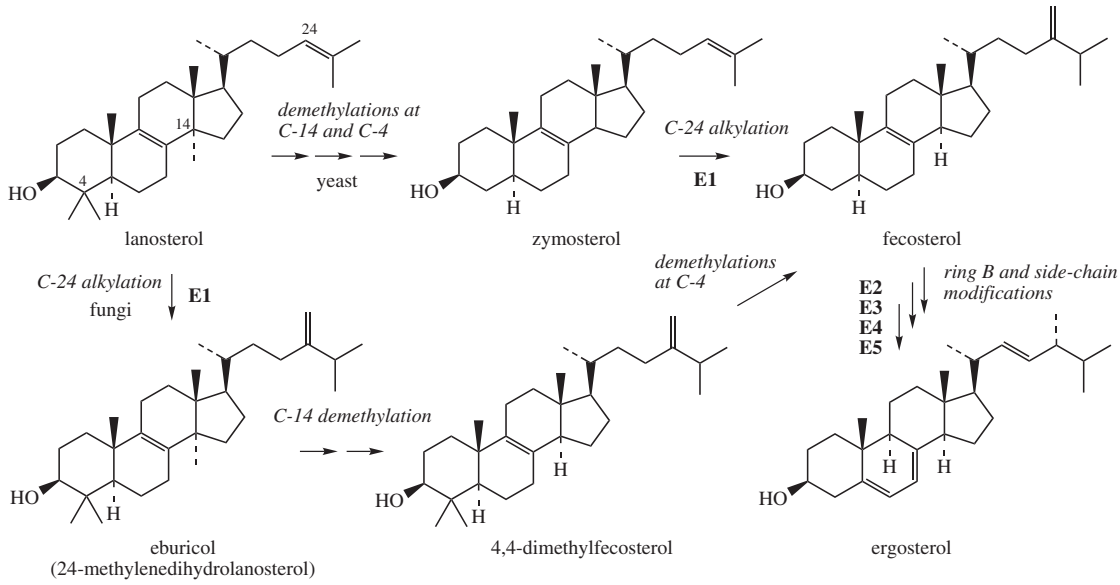


Figure 5.84

reduced to a 24-methyl either directly or after allylic isomerization. Alternatively, the 24-methylene derivative acts as substrate for a second methylation step with SAM, producing a carbocation. Discharge of this cation by proton loss produces a 24-ethylidene side-chain, and reduction or isomerization/reduction gives a 24-ethyl group. The *trans*-Δ^{22} double bond is introduced only after alkylation at C-24 is completed. No stereochemistry is intended in Figure 5.82. It is apparent that stereochemistries in the 24-methyl, 24-ethyl, and 24-ethylidene derivatives could be controlled by the reduction processes or by proton loss as appropriate. It is more plausible for different stereochemistries in the 24-methyl and 24-ethyl side-chains to arise from reduction of different double bonds, rather than reduction of the same double bond in two different ways. In practice, other mechanisms involving a 25(26)-double bond are also found to operate.

The substrates for alkylation are found to be cycloartenol in plants and algae, and lanosterol in fungi. The second methylation step in plants and algae usually involves **gramisterol** (24-methylenelophenol; Figure 5.83). This indicates that the processes of side-chain alkylation and the steroid skeleton modifications, i.e. loss of methyl groups, opening of the cyclopropane ring, and migration of double bond, tend to run concurrently rather than sequentially. Accordingly, the range of plant and algal sterol derivatives includes products containing side-chain alkylation, that retain one or more skeletal methyl groups, and perhaps possess a cyclopropane ring, as well as those more abundant examples such as sitosterol and stigmasterol based on a cholesterol-type skeleton.

Most fungal sterols originate from lanosterol, so less variety is encountered. The pathway from lanosterol to **ergosterol** proceeds via **zymosterol** in yeast, but via **eburicol** (24-methylenedihydrolanosterol) in filamentous fungi; both pathways lead through **fecosterol** (Figure 5.84). Thus, in yeast, there is initially the same sequence of demethylations as in mammals, but then zymosterol undergoes side-chain alkylation. In fungi, side-chain alkylation occurs before the demethylations. The pathways converge at fecosterol, and ergosterol is produced by further ring B and side-chain modifications. Some useful antifungal agents, e.g. ketoconazole and miconazole, are specific inhibitors of the 14α-demethylases in fungi, but have minimal effect on cholesterol biosynthesis in humans; there is a degree of selectivity towards fungal over mammalian enzymes. Inability to synthesize the essential sterol components of their membranes proves fatal for fungi. Similarly, 14-demethylation in plants proceeds via **obtusifoliol** (Figure 5.83) and plants are unaffected by azole derivatives developed as agricultural fungicides.

The antifungal effect of polyene antibiotics, such as amphotericin and nystatin, depends on their ability to bind strongly to ergosterol in fungal membranes and not to cholesterol in mammalian cells (see page 81).

Sitosterol and **stigmasterol** (Figure 5.81) are produced commercially from soya beans (*Glycine max*; Leguminosae/Fabaceae) as raw materials for the semi-synthesis of medicinal steroids (see pages 282, 294) [Box 5.13]. For many years, only stigmasterol was utilized, since the Δ^{22} double bond allowed chemical degradation of the side-chain to be effected with ease. The utilization of sitosterol was not realistic until microbiological processes for removal of the saturated side-chain became available.

Box 5.13

Soya Bean Sterols

Soya beans or soybeans (*Glycine max*; Leguminosae/Fabaceae) are grown extensively in the United States, China, Japan, and Malaysia as a food plant. They are used as a vegetable, and provide a high protein flour, an important edible oil (Table 3.1), and an acceptable non-dairy soybean milk. The flour is increasingly used as a meat substitute. Soy sauce is obtained from fermented soybeans and is an indispensible ingredient in Chinese cookery. The seeds also contain substantial amounts (about 0.2%) of sterols. These include stigmasterol (about 20%), sitosterol (about 50%), and campesterol (about 20%) (Figure 5.81), the first two of which are used for the semi-synthesis of medicinal steroids. In the seed, about 40% of the sterol content is in the free form, the remainder being combined in the form of glycosides or as esters with fatty acids. The oil is usually solvent extracted from the dried flaked seed using hexane. The sterols can be isolated from the oil after basic hydrolysis as a by-product of soap manufacture, and form part of the unsaponifiable matter.

The efficacy of dietary plant sterols in reducing cholesterol levels in laboratory animals has been known for many years. This has more recently led to the introduction of **plant sterol esters** as food additives, particularly in margarines and dairy products, as an aid to reducing blood levels of low density lipoprotein (LDL) cholesterol, known to be a contributory factor in atherosclerosis and the incidence of heart attacks (see page 251). Plant sterol esters are usually obtained by esterifying sitosterol from soya beans with fatty acids to produce a fat-soluble product. Regular consumption of this material (recommended 1.3 g per day) is shown to reduce blood LDL cholesterol levels by 10–15%. The plant sterols are more hydrophobic than cholesterol and have a higher affinity for micelles involved in fat digestion, effectively decreasing intestinal cholesterol absorption. The plant sterols themselves are poorly absorbed from the gastrointestinal tract. Of course, the average diet will naturally include small amounts of plant sterol esters.

Related materials used in a similar way are **plant stanol esters**. Stanols are obtained by hydrogenation of plant sterols, and will consist mainly of sitostanol (from sitosterol and stigmasterol) and campestanol (from campesterol) (Figure 5.85); these are then esterified with fatty acids. Regular consumption of plant stanol esters (recommended 3.4 g per day) is shown to reduce blood LDL cholesterol levels by an average of 14%. Much of the material used in preparation of plant stanol esters originates from tall oil, a by-product of the wood pulping industry. This contains campesterol, sitosterol, and also sitostanol. The stanols are usually transesterified with rapeseed oil, which is rich in unsaturated fatty acids (see page 47). Stanols tend to be more effective in reducing cholesterol levels than sterols. It has proved possible to engineer food plants such as rapeseed and soya to produce the reduced sterols, i.e. stanols, so that the chemical reduction steps may be rendered unnecessary. This has been achieved by incorporating a *Streptomyces*-derived 3-hydroxysteroid oxidase gene into the plant host. This enzyme is an FAD-dependent

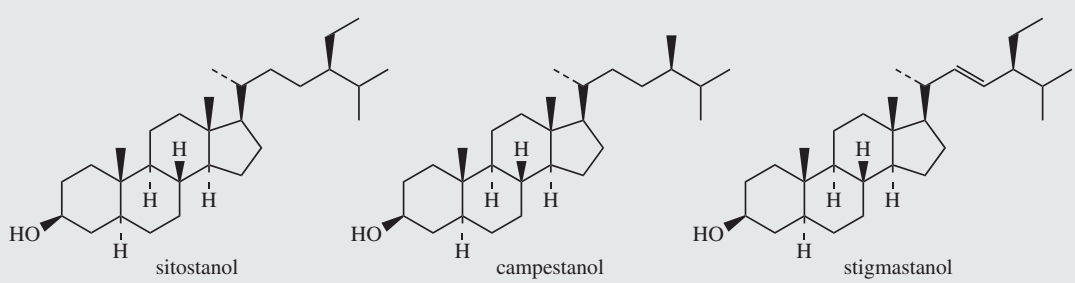


Figure 5.85

Box 5.13 (continued)

bifunctional enzyme that catalyses oxidation and isomerization in ring A of cholesterol and other steroids (compare page 264); additional reductive steps are presumably achieved by endogenous plant enzymes (Figure 5.86). Genetically modified soya seeds contained sitostanol, stigmastanol, and campestanol (Figure 5.85), with up to 80% of each normal sterol content being replaced by the reduced compound. Side-chain double bonds were not reduced.

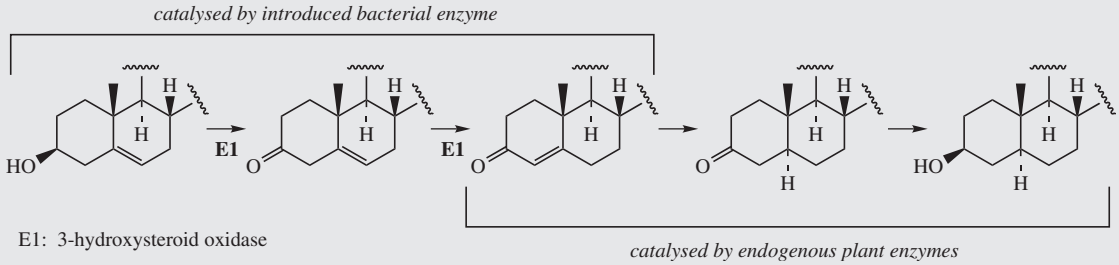


Figure 5.86

Fusidic acid (Figure 5.87), an antibacterial agent from *Fusidium coccineum* [Box 5.14], has no additional side-chain alkylation, but has lost one C-4 methyl and undergone hydroxylation and oxidation of a side-chain methyl. The stereochemistry in fusidic acid is not typical of most steroids, and ring B adopts a boat conformation; the *chair–boat–chair–chair* molecular shape is comparable to the protosteryl cation (Figure 5.63, page 238). Its relationship to the protosteryl cation is shown in Figure 5.87, and its formation probably involves initial proton loss, which thus prevents the normal carbocation-mediated migrations.

Vitamin D

Vitamin D_3 (**colecalciferol, cholecalciferol**) [Box 5.15] is a sterol metabolite formed photochemically in animals from **7-dehydrocholesterol** by the sun's irradiation of the skin (Figure 5.88). 7-Dehydrocholesterol is the immediate $\Delta^{5,7}$ diene precursor of cholesterol (see

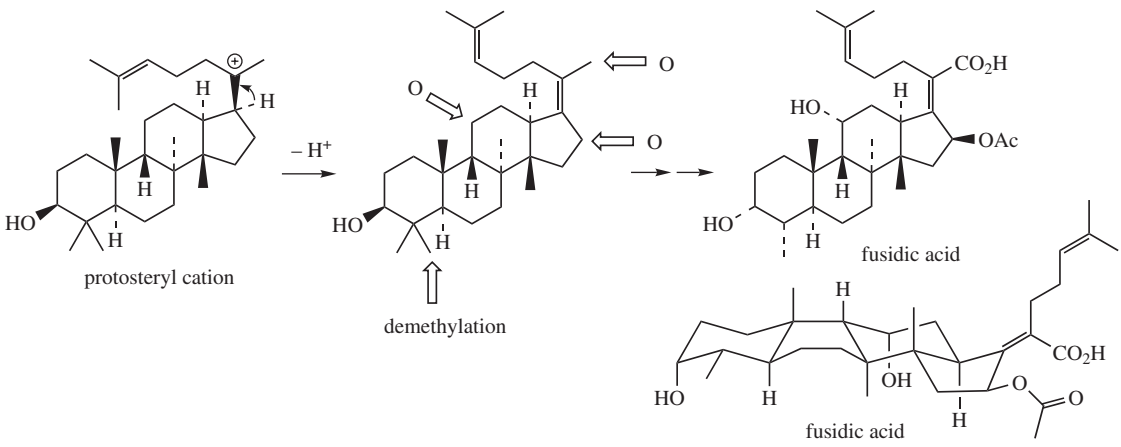


Figure 5.87

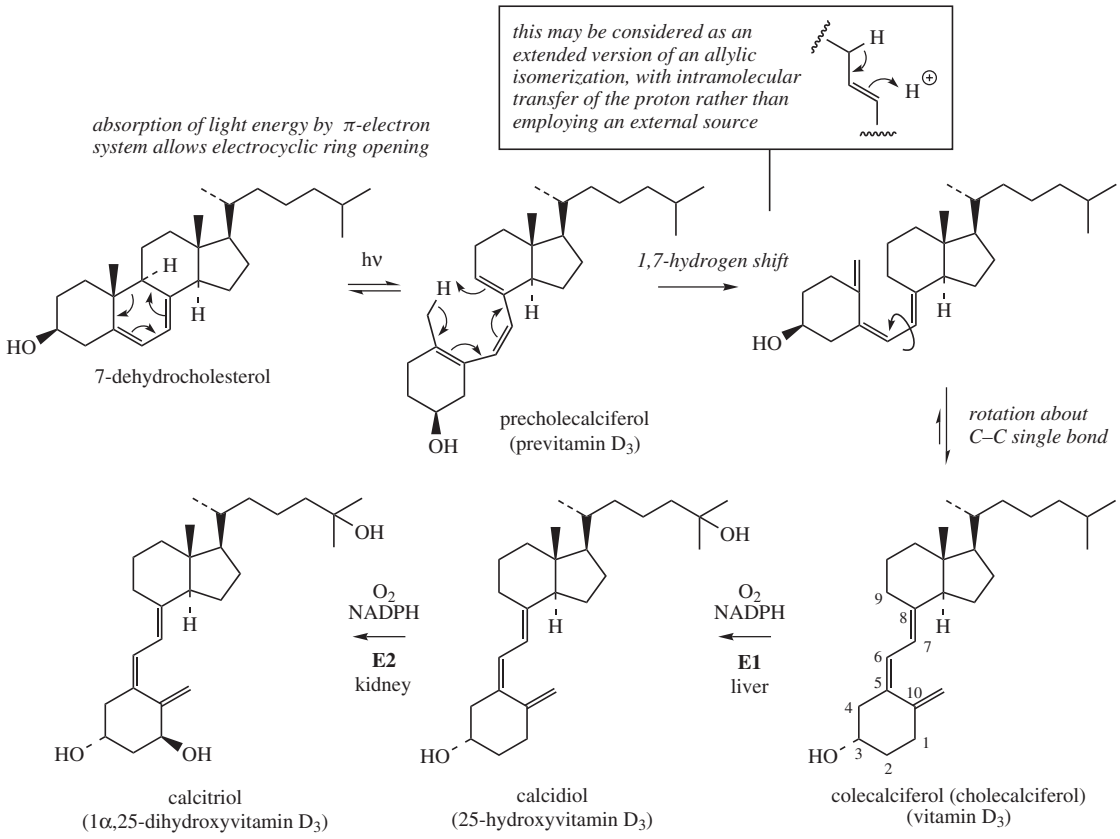

this may be considered as an extended version of an allylic isomerization, with intramolecular transfer of the proton rather than employing an external source
absorption of light energy by π-electron system allows electrocyclic ring opening
hν
1,7-hydrogen shift
7-dehydrocholesterol
precholecalciferol (previtamin D3)
rotation about C–C single bond
O2 NADPH
E2 kidney
O2 NADPH
E1 liver
calcitriol (1α,25-dihydroxyvitamin D3)
calcidiol (25-hydroxyvitamin D3)
colecalciferol (cholecalciferol) (vitamin D3)
E1: cholecalciferol 25-monooxygenase
E2: calcidiol 1-monooxygenase
ergosterol
pre-ergocalciferol (previtamin D2)
ergocalciferol (vitamin D2)

Figure 5.88

Box 5.14

Fusidic Acid

Fusidic acid (Figure 5.87) is a steroidal antibiotic produced by cultures of the fungus *Acremonium fusidioides* (formerly *Fusidium coccineum*). It has also been isolated from several *Cephalosporium* species. Fusidic acid and its salts are narrow-spectrum antibiotics active against Gram-positive bacteria, especially *Staphylococcus*. It is primarily used, as its salt **sodium fusidate**, in infections caused by penicillin-resistant *Staphylococcus* species, especially in osteomyelitis, since fusidic acid concentrates in bone. It is usually administered in combination with another antibiotic to minimize development of resistance. Fusidic acid reversibly inhibits bacterial protein biosynthesis at the translocation step by binding to the larger subunit of the ribosome (see page 422).

page 250), and a photochemical reaction allows ring opening to precholecalciferol. A thermal 1,7-hydrogen shift follows to give colecalciferol (vitamin D_3). Vitamin D_3 is manufactured photosynthetically by the same sequence. **Vitamin D_2** (**ergocalciferol**) is formed from ergosterol in exactly the same way, and is found naturally in plants and yeasts. Large amounts are produced semi-synthetically by the sequence shown in Figure 5.88, using ergosterol from yeast (*Saccharomyces cerevisiae*). Vitamin D_3 is not itself the active form of the vitamin; in the body it is hydroxylated first to **calcidiol** and then to **calcitriol** (Figure 5.88). Colecalciferol and calcitriol have also been found in several plant species, especially members of the Solanaceae family.

Systematic nomenclature of vitamin D derivatives utilizes the obvious relationship to steroids, and the term *seco* (ring-opened) is incorporated into the root name (compare secologanin as a ring-opened analogue of loganin, page 207). The numbering system for steroids is also retained, and vitamin D_3 becomes a derivative of 9,10-secocholestane, namely (5*Z*,7*E*)-9,10-secocholesta-5,7,10(19)-trien-3β-ol, '9,10' indicating the site of ring cleavage. Note that it is necessary to indicate the configuration of two of the double bonds. The β-configuration for the 3-hydroxyl and the α-configuration for the 1-hydroxy in calcitriol are potentially confusing until one appreciates that the bottom ring has been turned over in the cholesterol–vitamin D relationship.

Box 5.15

Vitamin D

Vitamin D_3 (**colecalciferol, cholecalciferol**; Figure 5.88) is the main form of the fat-soluble vitamin D found in animals, though **vitamin D_2** (**ergocalciferol**; Figure 5.88) is a constituent of plants and yeasts. Vitamin D_3 is obtained in the diet from liver and dairy products such as butter, cream, and milk, whilst large amounts can be found in oily fish and fish liver oils, e.g. cod liver oil and halibut liver oil (Table 3.1, page 46). A major source for normal requirements is the casual exposure of the skin to sunlight, when the sterol 7-dehydrocholesterol is converted into colecalciferol by UV irradiation. With a proper diet, and sufficient exposure to sunshine, vitamin D deficency should not occur. Vitamin D deficiency leads to rickets, an inability to calcify the collagen matrix of growing bone, and is characterized by a lack of rigidity in the bones, particularly in children. In adults, osteoporosis may occur. In most countries, foods such as milk and cereals are usually fortified with vitamin D_3, obtained commercially by UV irradiation of 7-dehydrocholesterol which is produced in quantity by semi-synthesis from cholesterol. Vitamin D_2 has a similar activity in humans and is manufactured by UV irradiation of yeast, thereby transforming the ergosterol content. Other compounds with vitamin D activity have also been produced: vitamin D_4 from 22,23-dihydroergosterol, vitamin D_5 from 7-dehydrositosterol, vitamin D_6 from 7-dehydrostigmasterol, and vitamin D_7 from 7-dehydrocampesterol. Vitamin D_1 was an early preparation later shown to be a mixture of vitamin D_2 and a photochemical by-product lumisterol. Lumisterol (9β,10α-ergosterol) is formed by recyclization of pre-ergocalciferol (Figure 5.88), giving the ergosterol stereoisomer. Vitamin D is unstable to heat, light, and air.

Vitamin D_3 is not itself the active form of the vitamin, and in the body it is hydroxylated firstly to 25-hydroxyvitamin D_3 (calcidiol; Figure 5.88) by an enzyme in the liver, and then to 1α,25-dihydroxyvitamin D_3 (calcitriol) by a kidney enzyme; both enzymes are cytochrome P-450 systems. Calcitriol is then transported to the bones, intestine, and other organs. It stimulates calcium absorption from the intestine, reabsorption from the kidney, and mobilization from bone. **Calcitriol** and other analogues, e.g. **alfacalcidol, dihydrotachysterol**, and **paricalcitol** (Figure 5.89), are available for use where chronic vitamin D deficiency is due to liver or kidney malfunction. The long-term use of calcitriol and alfacalcidol (1α-hydroxyvitamin D_3) in the treatment of osteoporosis may lead to toxic effects arising from elevated serum calcium levels.

Box 5.15 (continued)

Vitamin D is also known to have other physiological functions, including a role in immune suppression, hormone secretion, and the differentiation of both normal and malignant cells. Vitamin D derivatives, including **calcipotriol, tacalcitol**, and **maxacalcitol** (Figure 5.89), are widely used in the topical treatment of psoriasis, to inhibit the cell proliferation characteristic of this condition. In maxacalcitol, the side-chain modification (22-methylene replaced by an ether link) appears to retain the non-calcaemic action whilst reducing calcaemic activity, the latter probably because of more rapid oxidative metabolism in the liver.

Vitamin D_2 is also employed as a rodenticide. High doses are toxic to rats and mice, since the vitamin causes fatal hypercalcaemia.

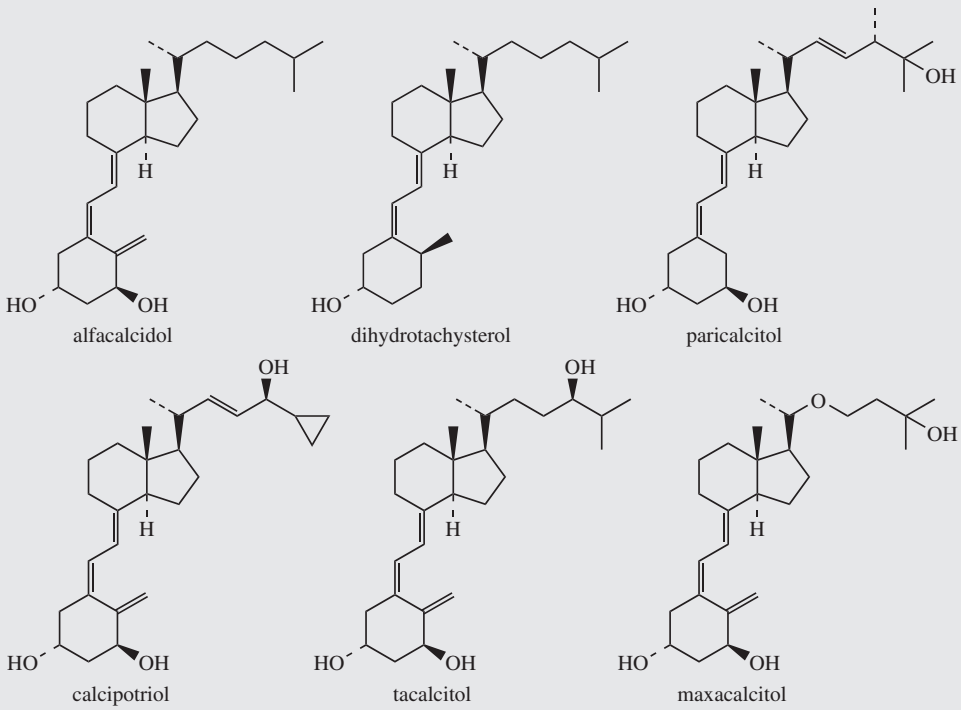
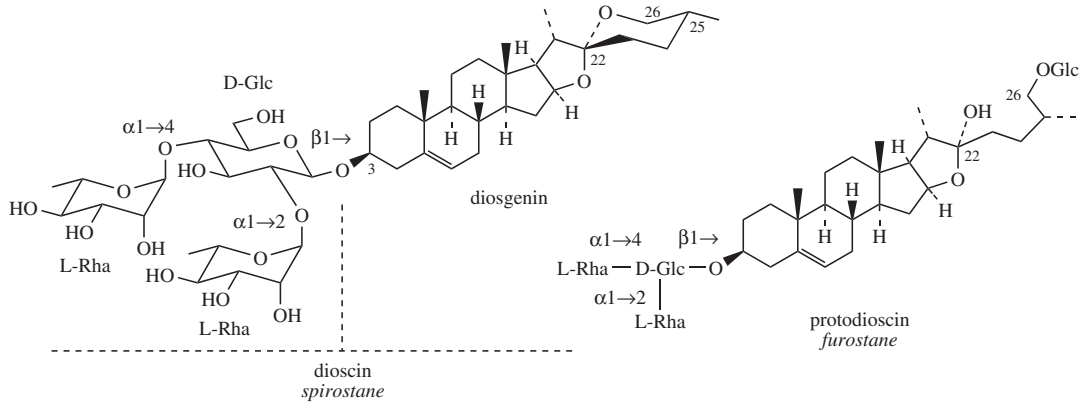


Figure 5.89

Steroidal Saponins

Steroidal saponins have similar biological properties to the triterpenoid saponins, e.g. surfactant and haemolytic activities (see page 242), but are less widely distributed in nature. They are found in many monocot families, especially the Dioscoreaceae (e.g. *Dioscorea*), the Agavaceae (e.g. *Agave, Yucca*) and the Liliaceae (e.g. *Smilax, Trillium*). Their sapogenins are C_{27} sterols in which the side-chain of cholesterol has undergone modification to produce either a spiroketal (spirostane saponins), e.g. **dioscin**, or a hemiketal (furostane saponins), e.g. **protodioscin** (Figure 5.90). As described below, the furostanes feature in the biosynthetic sequence to the spirostanes. All the steroidal saponins have the same configuration at the centre C-22, but stereoisomers at C-25 exist, e.g. **yamogenin** (Figure 5.91), and often mixtures of the C-25 stereoisomers co-occur in the same plant. Sugars are found at position 3, with a second glycoside function at C-26 in the furostanes. The sugar moiety at position 3 typically contains fewer monosaccharide units than are found with triterpenoid saponins; one to three monosaccharide units are most common. In general, the more sugar residues there are attached, the greater is the haemolytic activity. The sugar at C-26 in furostanes is usually glucose.



Characteristic features of steroidal saponins:

spirostanes:
- spiroketal at C-22
- common configuration at C-22
- sugar residues on 3β-hydroxyl

furostanes:
- hemiketal at C-22
- common configuration at C-22
- sugar residues on 3β-hydroxyl and 26-hydroxyl

Figure 5.90

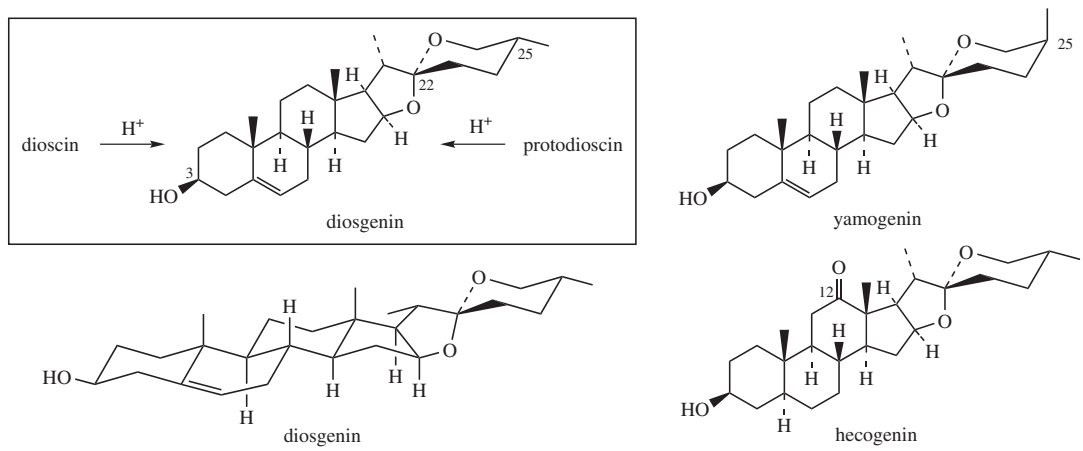


Figure 5.91

Acid hydrolysis of either dioscin or protodioscin liberates the aglycone **diosgenin** (Figure 5.91); the hydrolytic conversion of protodioscin into diosgenin is analogous to the biosynthetic sequence. The three-dimensional shape of diosgenin is indicated in Figure 5.91.

The spiroketal function is derived from the cholesterol side-chain by a series of oxygenation reactions, hydroxylating one of the terminal methyl groups and at C-16, and then producing a ketone function at C-22 (Figure 5.92). This proposed intermediate is transformed into the hemiketal and then the spiroketal. The chirality at C-22 is fixed by the stereospecificity in the formation of the ketal, whilst the different possible stereochemistries at C-25 are dictated by whether C-26 or C-27 is hydroxylated in the earlier step. Enzymic glycosylation at the 3-hydroxyl of spirostane sapogenins has been reported, but knowledge of other steps at the enzymic level is lacking. Furostane derivatives, e.g. **protodioscin** (Figure 5.90), can co-occur with spirostanes, and undoubtedly represent glycosylation of the intermediate hemiketal at the 26-hydroxyl. These compounds are readily hydrolysed and then

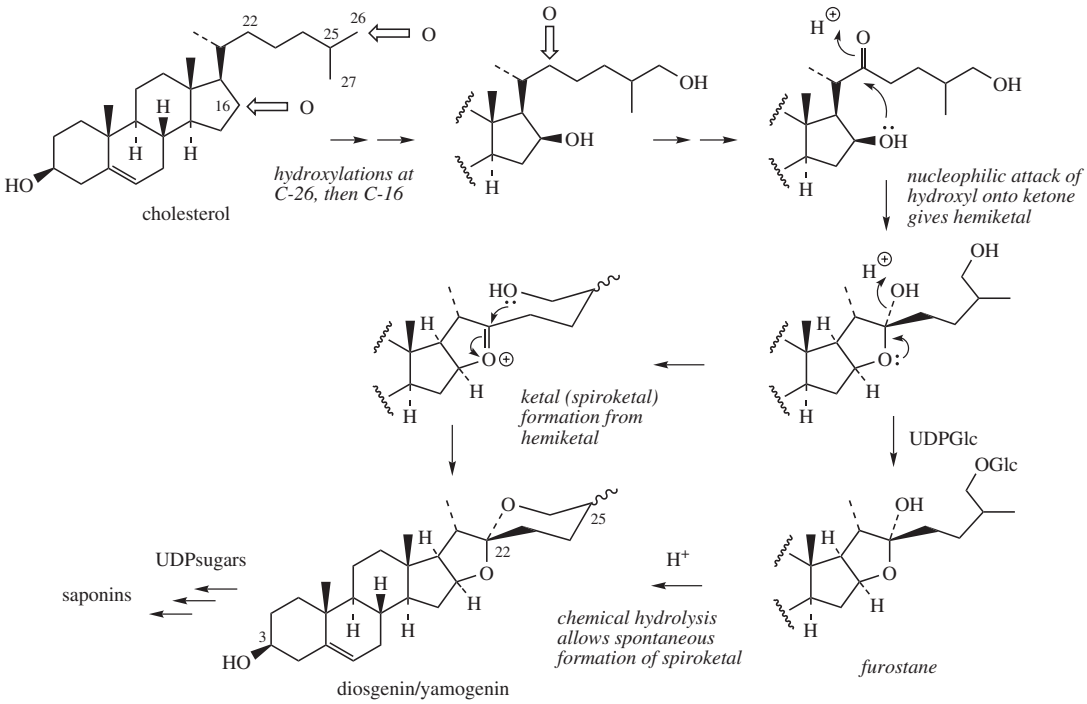


Figure 5.92

spontaneously cyclize to the spiroketal. Allowing homogenized fresh plant tissues to stand and autolyse through the action of endogenous glycosidase enzymes not only achieves cyclization of such open-chain saponins, but also can hydrolyse off the sugar units at C-3, thus yielding the aglycone or sapogenin. This is a standard approach employed in commercial production of steroidal sapogenins, which are important starting materials for the semi-synthesis of steroidal drugs [Box 5.16]. **Diosgenin** is the principal example and is obtained from Mexican yams (*Dioscorea* spp.; Dioscoreaceae). Fenugreek (*Trigonella foenum-graecum*; Leguminosae/Fabaceae) is another potentially useful commercial source of diosgenin. Sisal (*Agave sisalana*; Agavaceae) is also used commercially, yielding **hecogenin** (Figure 5.91), a 12-keto derivative with *trans*-fused A/B rings, the result of reduction of the Δ^5 double bond.

Undoubtedly related to the furostans and spirostans is the sterol glycoside **OSW-1** (Figure 5.93) isolated from the bulbs of *Ornithogalum saundersiae* (Liliaceae). This material has shown quite remarkable cytotoxicity, inhibiting the growth of various tumour cells, and being 10–100 times more potent than some clinical anticancer agents. It appears to damage the mitichondrial membrane in

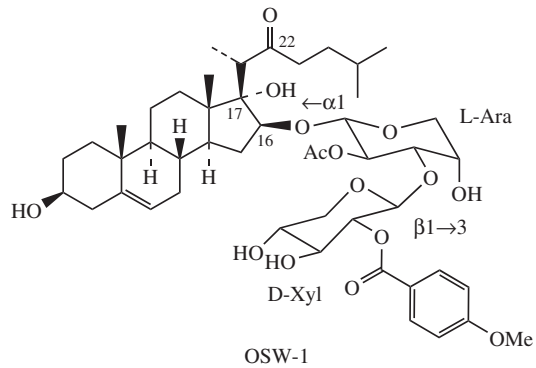


Figure 5.93

cancer cells, so its mechanism of action is also different to current anticancer agents. In addition, it shows little toxicity towards normal cells. The aglycone part of OSW-1 is structurally similar to early intermediates in the spirostanol pathway, in that it possesses the 22-keto and 16-hydroxyl functions, with an additional hydroxyl at C-17. 17α-Hydroxylation is characteristic of corticosteroids (see page 278) and the oestrogen/androgen biosynthetic pathway (see page 291).

Box 5.16

Dioscorea

About 600 species of *Dioscorea* (Dioscoreaceae) are known, and a number of these are cultivated for their large starchy tubers, commonly called yams, which are an important food crop in many parts of the world. Important edible species are *Dioscorea alata* and *D. esculenta* (Southeast Asia), *D. rotundata* and *D. cayenensis* (West Africa) and *D. trifida* (America). A number of species accumulate quite high levels of saponins in their tubers, which make them bitter and inedible, but these provide suitable sources of steroidal material for drug manufacture.

Dioscorea spp. are herbaceous, climbing, vine-like plants, the tuber being totally buried, or sometimes protruding from the ground. Tubers weigh anything up to 5 kg, but in some species the tubers have been recorded to reach weights as high as 40–50 kg. Drug material is obtained from both wild and cultivated plants, with plants collected from the wild having been exploited considerably more than cultivated ones. Commercial cultivation is less economic, requiring a 4–5-year growing period and some form of support for the climbing stems. Much of the world's production has come from Mexico, where tubers from *D. composita* (barbasco), *D. mexicana*, and *D. floribunda*, mainly harvested from wild plants, are utilized. The saponin content of the tubers varies, usually increasing as tubers become older. Typically, tubers of *D. composita* may contain 4–6% total saponins, and *D. floribunda* 6–8%. Other important sources of *Dioscorea* used commercially now include India (*D. deltoidea*), South Africa (*D. sylvatica*) and China (*D. collettii, D. pathaica*, and *D. nipponica*).

Sapogenins are isolated by chopping the tubers, allowing them to ferment for several days, and then completing the hydrolysis of saponins by heating with aqueous acid. The sapogenins can then be solvent extracted. The principal sapogenin in the species given above is diosgenin (Figure 5.91), with small quantities of the 25β-epimer yamogenin (Figure 5.91). Demand for diosgenin for pharmaceuticals is huge, equivalent to 10 000 t of *Dioscorea* tuber per annum, and it is estimated that about 60% of all steroidal drugs are derived from diosgenin.

Powdered *Dioscorea* (wild yam) root or extract is also marketed to treat the symptoms of menopause as an alternative to hormone replacement therapy (HRT; see page 294). Although there is a belief that this increases levels of progesterone, which is then used as a biosynthetic precursor of other hormones, there is little definitive evidence that diosgenin is metabolized in the human body to progesterone, and any beneficial effects may arise from diosgenin itself.

Fenugreek

The seeds of fenugreek (*Trigonella foenum-graecum*; Leguminosae/Fabaceae) are an important spice material, and are ingredients in curries and other dishes. The plant is an annual and is grown widely, especially in India, both as a spice and as a forage crop. Seeds can yield, after hydrolysis, 1–2% of sapogenins, principally diosgenin (Figure 5.91) and yamogenin (Figure 5.91). Although yields are considerably lower than from *Dioscorea*, the ease of cultivation of fenugreek and its rapid growth make the plant a potentially viable crop for steroid production in temperate countries. Field trials of selected high-yielding strains have been conducted.

Sisal

Sisal (*Agave sisalana*; Agavaceae) has long been cultivated for fibre production, being the source of sisal hemp, used for making ropes, sacking, and matting. The plant is a large, rosette-forming succulent with long, tough, spine-tipped leaves containing the very strong fibres. The main area of sisal cultivation is East Africa (Tanzania, Kenya), with smaller plantations in other parts of the world. The sapogenin hecogenin (Figure 5.91) was initially produced from the leaf waste (0.6–1.3% hecogenin) after the fibres had been stripped out. The leaf waste was concentrated, allowed to ferment for several days, and then treated with steam under pressure to complete hydrolysis of the saponins. Filtration then produced a material containing about 12% hecogenin, plus other sapogenins. This was refined further in the pharmaceutical industry. Other sapogenins present include tigogenin and neotigogenin (Figure 5.94).

As the demand for natural fibres declined due to the availability of synthetics, so did the supply of sisal waste and, thus, hecogenin. In due course, hecogenin became a more valuable commodity than sisal, and efforts were directed specifically towards hecogenin production. This has resulted in the cultivation of *Agave* hybrids with much improved hecogenin content. The highest levels (2.5%) of hecogenin recorded have been found in a Mexican species *Agave sobria* var *roseana*.

The fermented sap of several species of Mexican *Agave*, especially *Agave tequilana*, provides the alcoholic beverage pulque. Distillation of the fermented sap produces tequila.

Box 5.16 (continued)

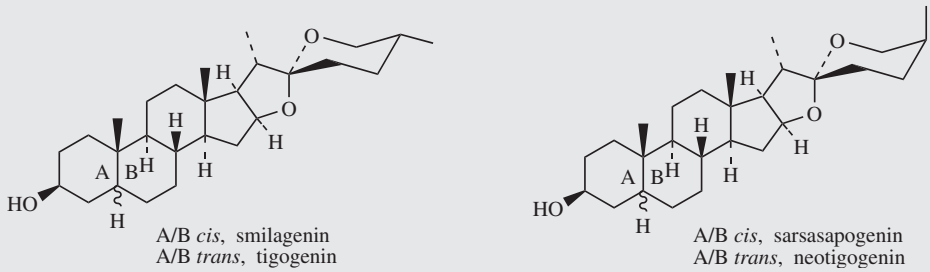


Figure 5.94

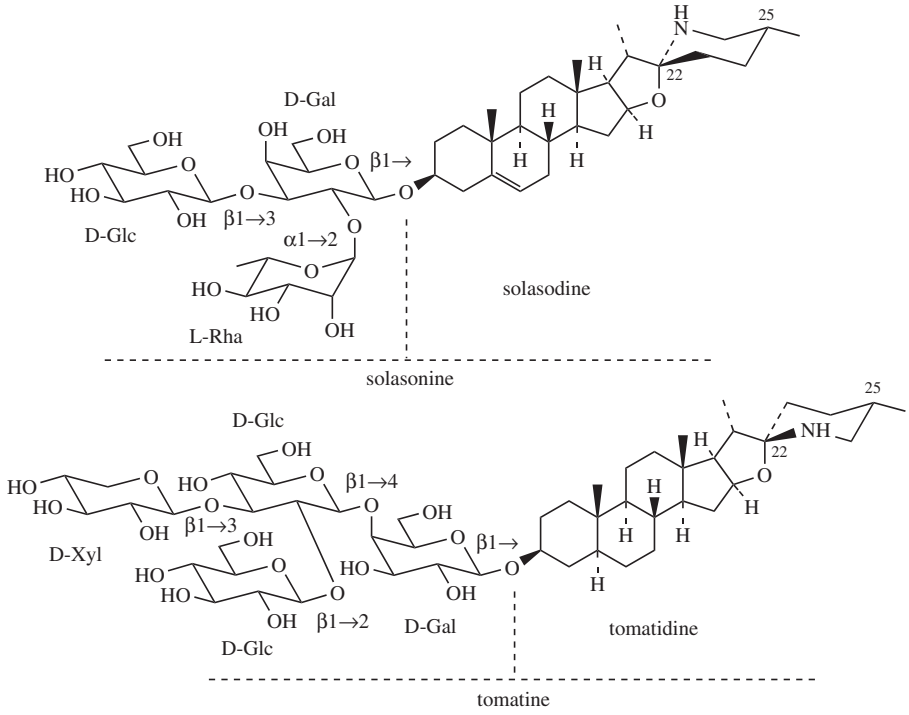


Figure 5.95

Some steroidal alkaloids are nitrogen analogues of steroidal saponins and display similar properties, such as surface activity and haemolytic activity, but these compounds *are* toxic when ingested. These types of compound, e.g. **solasonine** (Figure 5.95; aglycone **solasodine**), are found in many plants of the genus *Solanum* (Solanaceae), and such plants must thus be regarded as potentially toxic. In contrast to the oxygen analogues, all compounds have the same stereochemistry at C-25 (methyl always equatorial), whilst isomers at C-22 do exist, e.g. **tomatine** (Figure 5.95; aglycone **tomatidine**) from tomato (*Lycopersicon esculente*; Solanaceae). The nitrogen atom is introduced by a transamination reaction, typically employing an amino acid as donor (see page

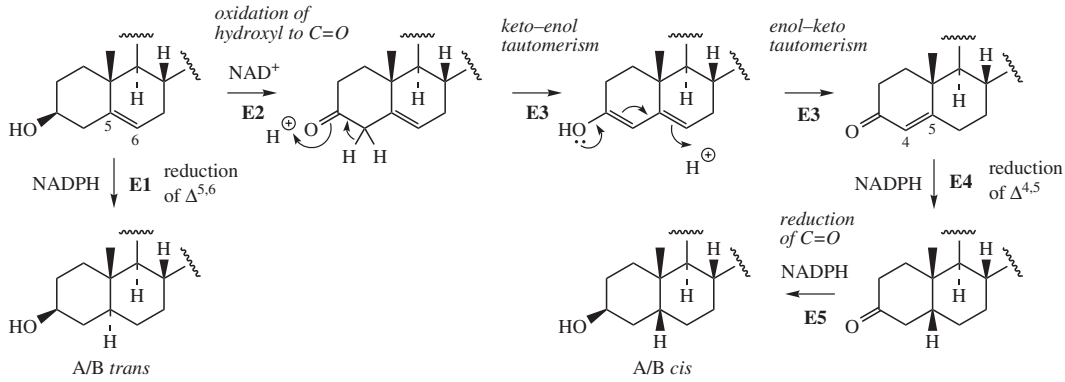


Figure 5.96

409). Since the production of medicinal steroids from steroidal saponins requires preliminary degradation to remove the ring systems containing the original cholesterol side-chain, it is immaterial whether these rings contain oxygen or nitrogen. Thus, plants rich in solasodine or tomatidine can also be employed for commercial steroid production (see page 410).

Smilagenin and **sarsasapogenin** (Figure 5.94) found in sarsaparilla (*Smilax* spp.; Liliaceae/Smilacaceae) are reduced forms of diosgenin and yamogenin respectively. These contain *cis*-fused A/B rings, whilst the corresponding *trans*-fused systems are present in **tigogenin** and **neotigogenin** (Figure 5.94) found in *Digitalis purpurea* along with cardioactive glycosides (see page 269). All four stereoisomers are derived from cholesterol, and the stereochemistry of the A/B ring fusion appears to be controlled by the nature of the substrate being reduced. Enzymic reduction of the isolated Δ^5 double bond yields the *trans*-fused system, whereas reduction of a Δ^4 double bond (1,4-addition to a conjugated ketone) gives the alternative *cis*-fused system (Figure 5.96). Accordingly, to obtain the A/B *cis* fusion, the Δ^5 unsaturation of cholesterol is changed to Δ^4 by oxidation of the 3-hydroxyl and allylic isomerization to the conjugated 4-ene-3-one system, and this is followed by reduction of both functional groups (Figure 5.96) (compare biosynthesis of progesterone, page 267). The sarsaparilla saponins are not present in sufficient quantities to be commercially important for steroid production, but quite large amounts of sarsasapogenin can be extracted from the seeds of *Yucca brevifolia* (Agavaceae) [Box 5.17].

Box 5.17

Sarsaparilla

Sarsaparilla consists of the dried roots of various *Smilax* species (Liliaceae/Smilacaceae), including *S. aristolochiaefolia, S. regelii*, and *S. febrifuga*, known respectively as Mexican, Honduran, and Ecuadorian sarsaparilla. The plants are woody climbers indigenous to Central America. Sarsaparilla has a history of use in the treatment of syphilis, rheumatism, and skin diseases, but is now mainly employed as a flavouring in the manufacture of non-alcoholic drinks. It has some potential as a raw material for the semi-synthesis of medicinal steroids, being a source of sarsasapogenin and smilagenin (Figure 5.94). The roots contain 1.8–2.4% steroidal saponins, including parillin (Figure 5.97).

Yucca

Yucca brevifolia (Agavaceae) has been explored as a potential source of sarsasapogenin for steroid production, especially at times when market prices of diosgenin from *Dioscorea* became prohibitively expensive. The plant grows extensively in the

Box 5.17 (continued)

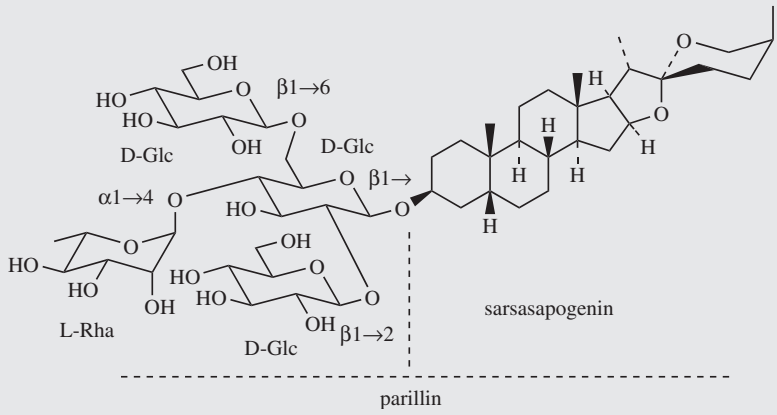


Figure 5.97

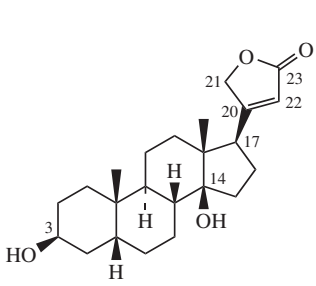
Mojave Desert in California, and high levels of sarsasapogenin (8–13%) are present in the seeds. This means the plants can be harvested regularly without damage. The subsequent stabilization of *Dioscorea* prices in the 1970s stopped any further commercial exploitation.

Cardioactive Glycosides

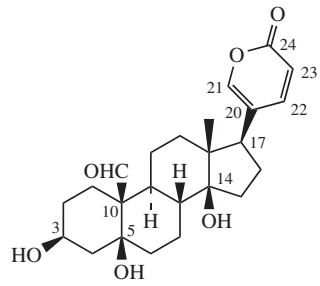
Many of the plants known to contain cardiac or cardiotonic glycosides have long been used as arrow poisons (e.g. *Strophanthus*) or as heart drugs (e.g. *Digitalis*). They are used medicinally to strengthen a weakened heart and allow it to function more efficiently, though the dosage must be controlled very carefully, since the therapeutic dose is so close to the toxic dose. The cardioactive effects of *Digitalis* were discovered as a result of its application in the treatment of dropsy, an accumulation of water in the body tissues. *Digitalis* alleviated dropsy indirectly by its effect on the heart, improving the blood supply to the kidneys and so removing excess fluid.

The therapeutic action of cardioactive glycosides depends on the structure of the aglycone, and on the type and number of sugar units attached. Two types of aglycone are recognized, **cardenolides** (e.g. **digitoxigenin** from *D. purpurea*), which are C_{23} compounds, and **bufadienolides** (e.g. **hellebrigenin** from *Helleborus niger*), which are C_{24} structures (Figure 5.98). Stereochemistry is very important for activity, and these compounds have *cis* fusions for both the A/B and C/D ring junctions, 3β- and 14β-hydroxyl groups with the glycoside function at C-3, and an α,β-unsaturated lactone grouping at C-17β. This lactone ring is five-membered in the cardenolides and six-membered in the bufadienolides. The hellebrigenin structure shows two other modifications not found in the basic steroid skeleton, namely a hydroxyl at the bridgehead carbon C-5 and a formyl group at C-10, being an oxidized form of the normal methyl. The three-dimensional shape of digitoxigenin is shown in Figure 5.98. These basic structures arise biosynthetically by metabolism of cholesterol, in which the side-chain is cleaved to a two-carbon acetyl group, followed by incorporation of either two carbonatoms for cardenolides or three carbonatoms for bufadienolides (Figure 5.99).

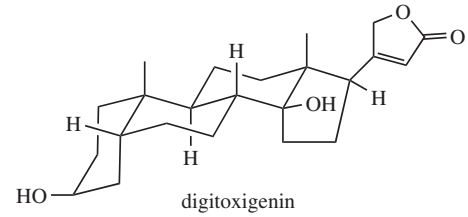
Shortening of the cholesterol side-chain is accomplished by stepwise hydroxylation at C-22 and then C-20, then cleavage of the C-20/22 bond to give **pregnenolone**, a sequence catalysed by the 'side-chain cleaving enzyme'. Pregnenolone is then oxidized in ring A to give **progesterone** (Figure 5.99). This can be reduced to give the *cis*-fused A/B system as in 5β-pregnan-3,20-dione (compare Figure 5.96) which is the substrate for 14β-hydroxylation, i.e. inverting the stereochemistry at this centre. Inversion is atypical for hydroxylation by monooxygenases, which are found to hydroxylate with retention of configuration. Whatever the mechanism of this hydroxylation, no Δ^8 or Δ^{15} double-bond intermediates are involved. Hydroxylation in the side-chain at C-21 follows. The lactone ring is



digitoxigenin
cardenolide



hellebrigenin
bufadienolide



digitoxigenin

Characteristic features of cardiac glycosides:

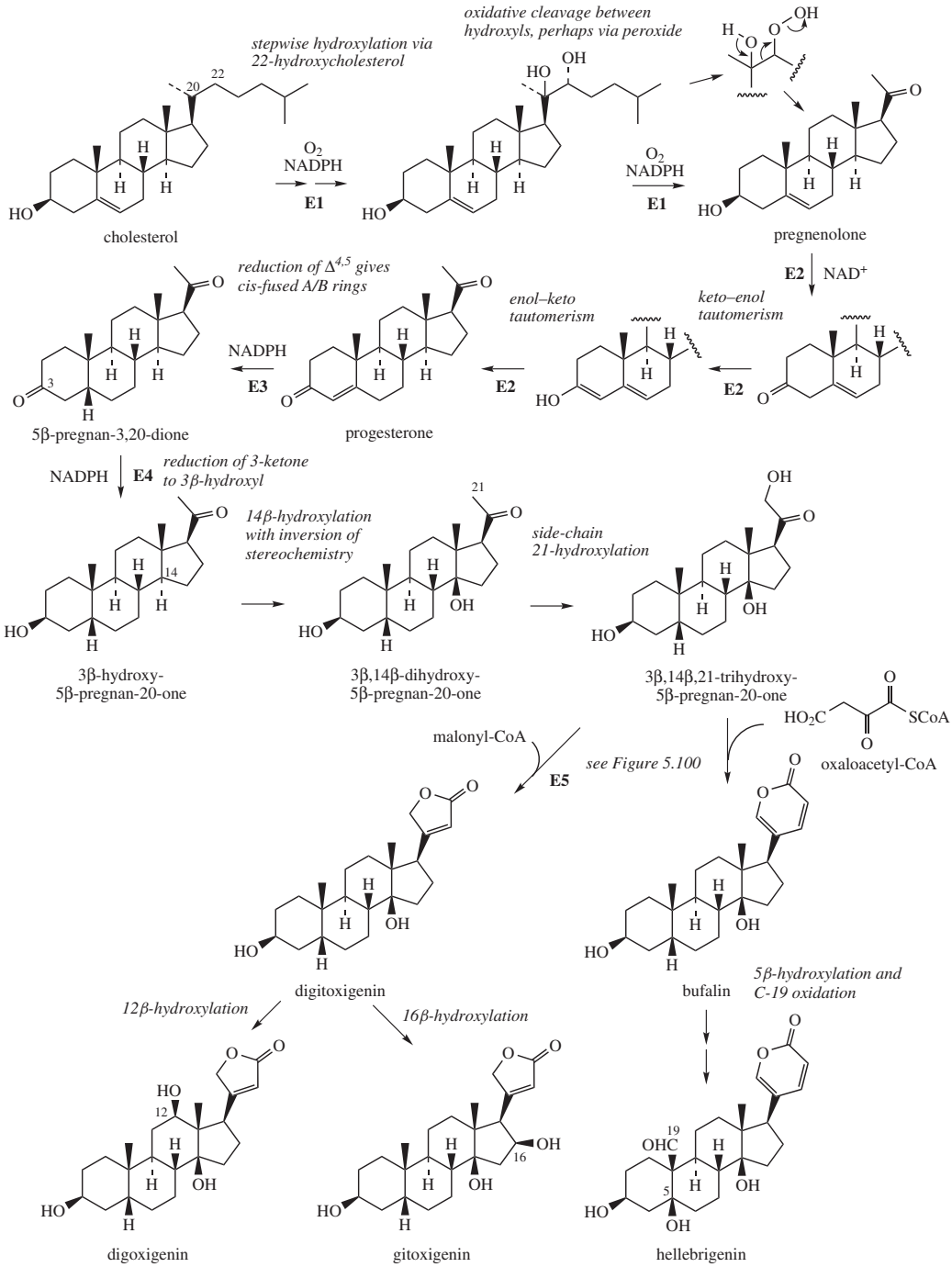
- *cis*-fused A/B and C/D rings
- 14β-hydroxyl
- unsaturated lactone at C-17β
- sugar residues on 3β-hydroxyl

Figure 5.98

created at this stage. An intermediate malonate ester is involved, and ring formation probably occurs via the aldol addition process shown in Figure 5.100 to give the cardenolide **digitoxigenin**, the carboxyl carbon of the malonate ester being lost by decarboxylation during the process (compare the role of malonate in the acetate pathway). Digitoxigenin is the precursor of **digoxigenin** and **gitoxigenin** by specific hydroxylations (Figure 5.99). Oxaloacetate is the source of the extra carbon atoms in bufadienolides; a similar esterification/aldol reaction sequence may be proposed (Figure 5.100). This would produce **bufalin** (Figure 5.99), a bufadienolide structure found in the skin of toad (*Bufo* spp.), from which this class of compound was originally isolated and has subsequently taken the general name. Note that, in the subsequent formation of **hellebrigenin** (Figure 5.99), hydroxylation at C-5 occurs with the expected retention of stereochemistry, not with inversion as seen at C-14.

The fundamental pharmacological activity of the cardioactive glycosides resides in the aglycone portion, but is considerably modified by the nature of the sugar at C-3. This increases water solubility and binding to heart muscle. The sugar unit may have one to four monosaccharides; many (e.g. D-**digitoxose** and D-**digitalose**; Figure 5.101) are unique to this group of compounds. About 20 different sugars have been characterized, and with the exception of D-glucose, they are 6-deoxy- (e.g. L-rhamnose, D-digitalose) or 2,6-dideoxy- (e.g. D-digitoxose, D-cymarose) hexoses, some of which are also 3-methyl ethers (e.g. D-digitalose and D-**cymarose**; Figure 5.101). In plants, cardiac glycosides are confined to the angiosperms, but are found in both monocotyledons and dicotyledons. The cardenolides are more common, and the plant families the Apocynaceae (e.g. *Strophanthus*), Liliaceae (e.g. *Convallaria*), and Scrophulariaceae (e.g. *Digitalis*) yield medicinal agents [Box 5.18]. The rarer bufadienolides are found in some members of the Liliaceae (e.g. *Urginea*) [Box 5.18] and Ranunculaceae (e.g. *Helleborus*), as well as in toads. Monarch butterflies and their larvae are known to accumulate in their bodies a range of cardenolides which they ingest from their food plant, the common milkweed (*Asclepias syriaca*; Asclepiadaceae). This makes them unpalatable to predators such as birds. Endogenous *Digitalis*-like compounds have also been detected, albeit in very small quantities, in mammalian tissues. **Ouabain** (see below, Figure 5.106), first isolated from the bark of the African ouabio tree (*Acokanthera ouabio*; Apocynaceae), is now known to occur naturally in blood plasma, adrenal glands, and the hypothalamus of mammals; it is a major component in the seeds of *Strophanthus gratus* [Box 5.18]. **19-Norbufalin** (Figure 5.101) is found in the lens of human eyes, at higher levels if these are cataract-afflicted, and it is believed to regulate ATPase activity under some physiological and pathological conditions.

Spirostane saponins are known to co-occur with cardenolide glycosides in *D. purpurea* (see page 264), indicating the plant is able to synthesize structures in which A/B and C/D ring fusions may be *trans–trans* or



E1: cholesterol monooxygenase (side-chain cleaving) (side-chain cleaving enzyme)
E2: Δ^5-3β-hydroxysteroid dehydrogenase/Δ^5-Δ^4-ketosteroid isomerase
E3: progesterone 5β-reductase
E4: 3β-hydroxysteroid dehydrogenase
E5: 21-hydroxypregnane 21-*O*-malonyltransferase

Figure 5.99

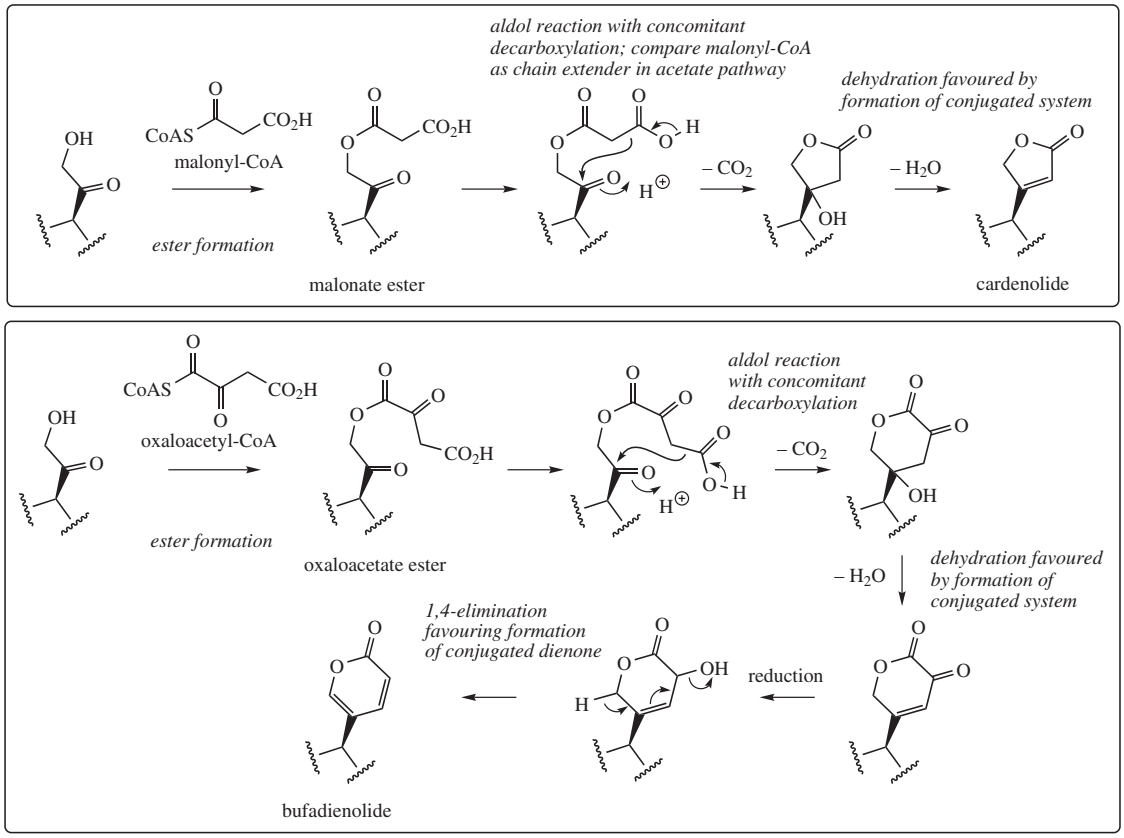


Figure 5.100

cis–cis. Other variants are possible, in that compounds such as **digipurpurogenin II** (Figure 5.102) have also been found in *D. purpurea*. This compound has the C/D ring fusion *cis*, with a 14β-hydroxyl group, characteristic of the cardiac glycosides, yet the A/B rings are typical of more common sterols. The glycoside **P57A3** of this compound found in the South African succulents *Hoodia pilifera* and *H. gordonii* is attracting considerable interest as an appetite suppressant. Indigenous people have used the plants as a substitute for food and water, and appetite-suppressant properties were subsequently confirmed, leading to the isolation of P57A3 and its glycoside (a cymaroside). A *Hoodia* herbal appetite suppressant has been patented. Though these compounds are pregnane derivatives, they bear a strong relationship to cardenolides via the 12β- and 14β-oxygenation, and the chain of 2,6-dideoxy sugar units.

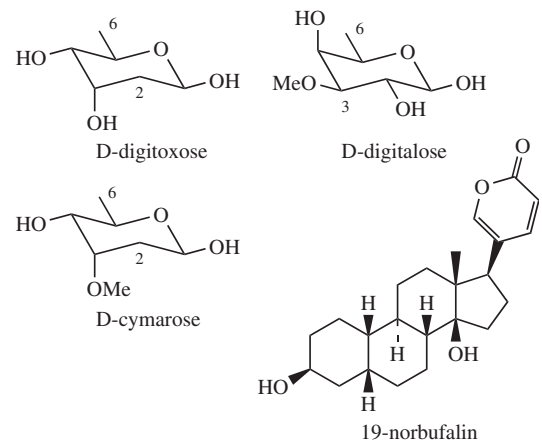


Figure 5.101

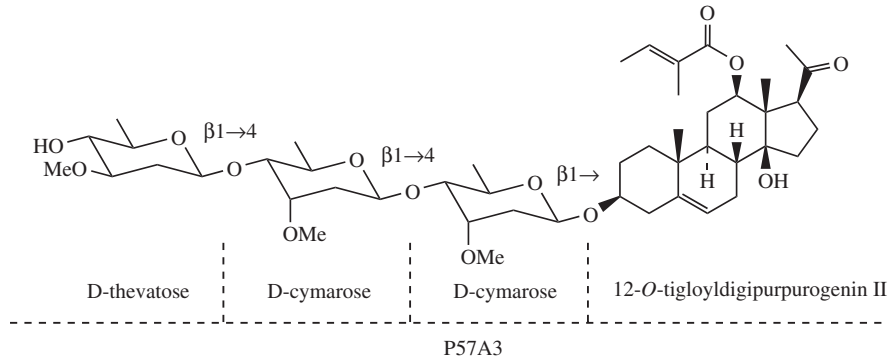


Figure 5.102

Box 5.18

Digitalis purpurea

Digitalis leaf consists of the dried leaf of the red foxglove *Digitalis purpurea* (Scrophulariaceae). The plant is a biennial herb, common in Europe and North America, which forms a low rosette of leaves in the first year and its characteristic spike of purple (occasionally white) bell-shaped flowers in the second year. It is potentially very toxic, but the leaf is unlikely to be ingested by humans. *D. purpurea* is cultivated for drug production, principally in Europe, the first year leaves being harvested then rapidly dried at 60°C as soon as possible after collection. This procedure is necessary to inactivate hydrolytic enzymes which would hydrolyse glycoside linkages in the cardioactive glycosides, giving rise to less active derivatives. Even so, some partial hydrolysis does occur. Excess heat may also cause dehydration in the aglycone to biologically inactive Δ^{14}-anhydro compounds.

Because of the pronounced cardiac effects of digitalis, the variability in the cardiac glycoside content, and also differences in the range of structures present due to the effects of enzymic hydrolysis, the crude leaf drug is usually assayed biologically rather than chemically. Prepared digitalis is a biologically standardized preparation of powdered leaf, its activity being assessed on cardiac muscle of guinea pig or pigeon and compared against a standard preparation. It may be diluted to the required activity by mixing in powdered digitalis of lower potency, or inactive materials such as lucerne (*Medicago sativa*) or grass. The crude drug is hardly ever used now, having been replaced by the pure isolated glycosides.

The cardioactive glycoside content of *D. purpurea* leaf is 0.15–0.4%, consisting of about 30 different structures. The major components are based on the aglycones digitoxigenin, gitoxigenin, and gitaloxigenin (Figure 5.103), the latter being a formate ester. The glycosides comprise two series of compounds, those with a tetrasaccharide *glucose*–(*digitoxose*)$_3$– unit and those with a trisaccharide (*digitoxose*)$_3$– unit. The latter group (the secondary glycosides) is produced by partial hydrolysis from the former group (the primary glycosides) during drying by the enzymic action of a β-glucosidase which removes the terminal glucose. Thus, the principal glycosides in the fresh leaves, namely purpureaglycoside A and purpureaglycoside B (Figure 5.103), are partially converted into digitoxin and gitoxin respectively (Figure 5.103), which normally predominate in the dried leaf. These transformations are indicated schematically in Figure 5.104. In the fresh leaf, purpureaglycoside A can constitute about 50% of the glycoside mixture, whilst in the dried leaf the amounts could be negligible if the plant material is old or poorly stored. The gitaloxigenin-based glycosides are relatively unstable, and the formyl group on the aglycone is readily lost by hydrolysis. Other minor glycosides are present, but neither the fresh nor dried leaf contains any significant quantities of the free aglycones.

Glycosides of the gitoxigenin series are less active than the corresponding members of the digitoxigenin-derived series. **Digitoxin** is the only compound routinely used as a drug, and it is employed in congestive heart failure and treatment of cardiac arrhymias, particularly atrial fibrillation.

Digitalis lanata

Digitalis lanata (Scrophulariaceae), the Grecian foxglove, is a perennial or biennial herb from southern and central Europe. It differs in appearance from the red foxglove by its long narrow smoother leaves and its smaller flowers of a yellow–brown colour. It is cultivated in Europe, the United States, and South America and is harvested and dried in a similar manner to *D. purpurea*. It has not featured as a crude drug, but is used exclusively for the isolation of individual cardiac glycosides, principally digoxin and lanatoside C (Figure 5.105).

Box 5.18 (continued)

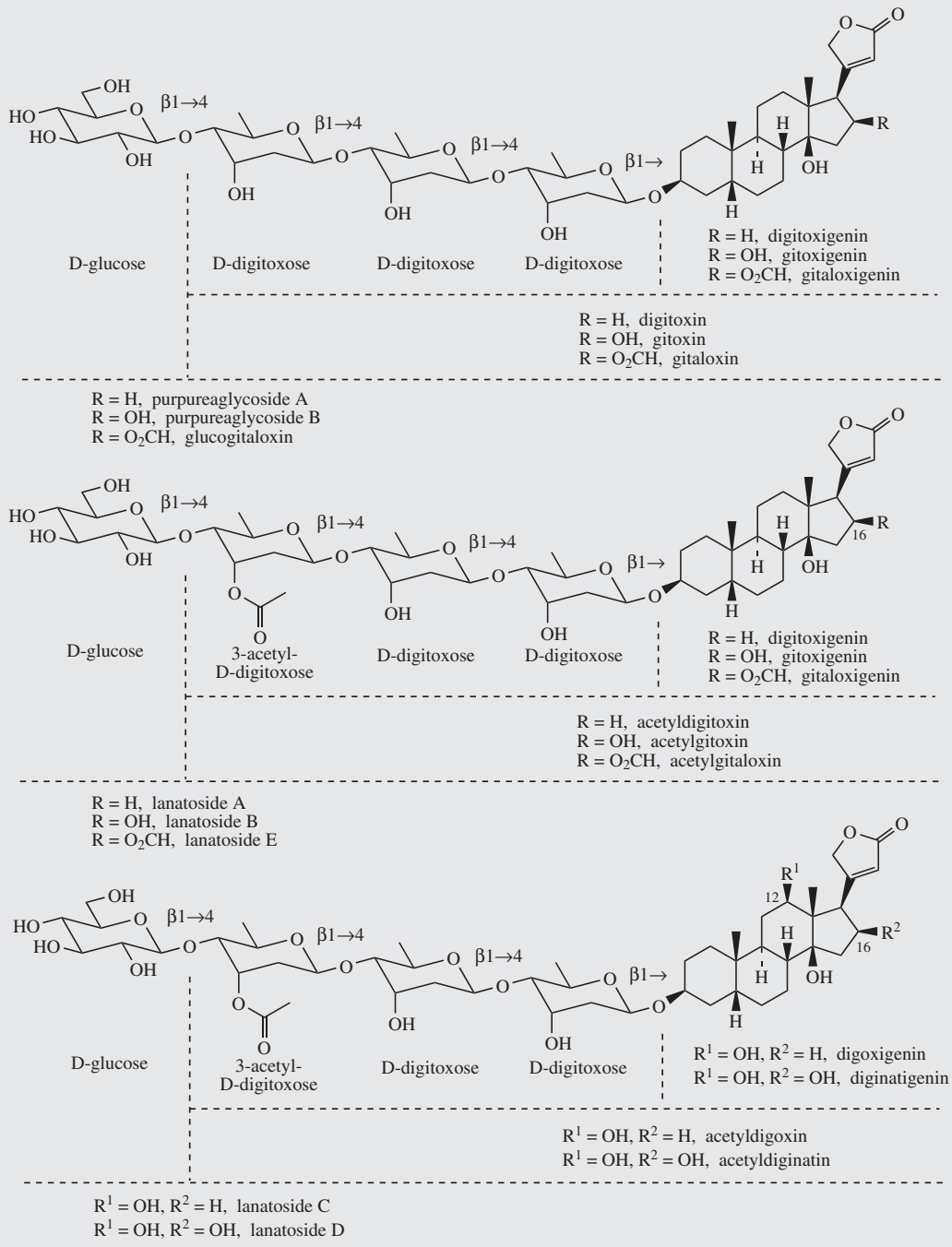


Figure 5.103

Box 5.18 (continued)

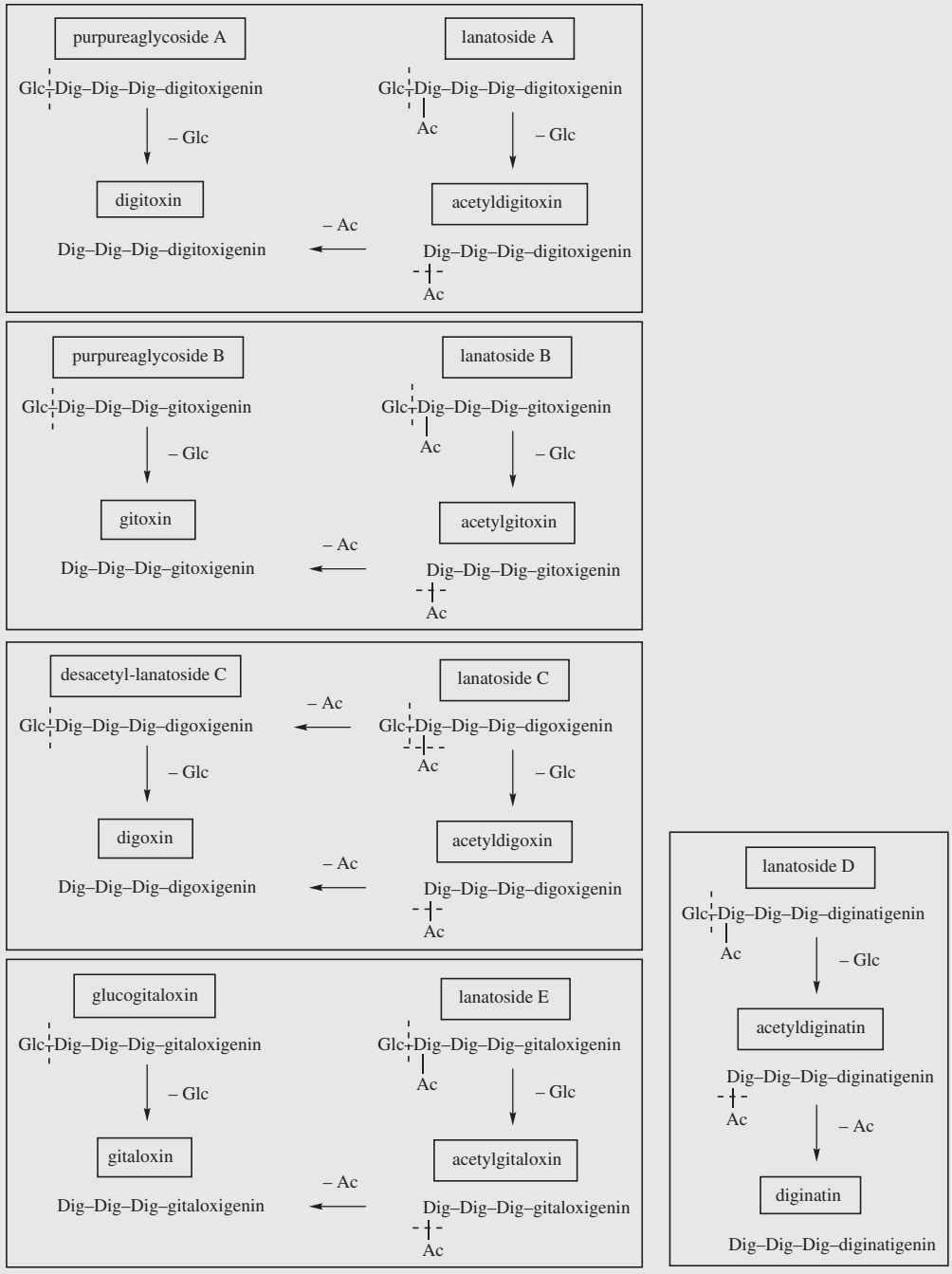
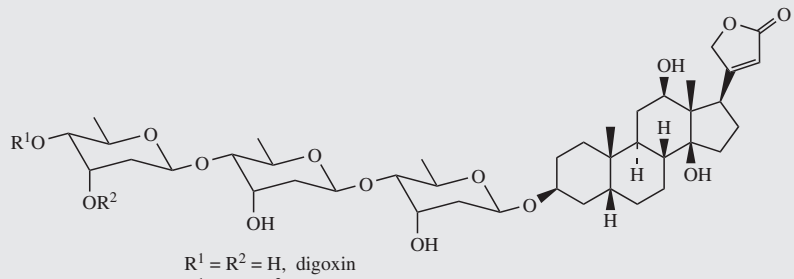


Figure 5.104

Box 5.18 (continued)



$R^1 = R^2 = H$, digoxin
$R^1 = Me$, $R^2 = H$, medigoxin (metildigoxin)
$R^1 = Glc$, $R^2 = Ac$, lanatoside C
$R^1 = Glc$, $R^2 = H$, deslanoside (desacetyl-lanatoside C)

Figure 5.105

The total cardenolide content of up to 1% is two to three times that found in *D. purpurea*. The main constituents resemble those of *D. purpurea*, but contain an acetyl ester function on the third digitoxose, that furthest from the aglycone. This acetyl group makes the compounds easier to isolate from the plant material and they crystallize more readily. Drying of the leaf is similarly accompanied by some partial hydrolysis of the original fresh leaf constituents through enzymic action, and both the terminal glucose and the acetyl group may be hydrolysed off, extending the range of compounds isolated. The *D. lanata* cardiac glycosides are based on five aglycones: digitoxigenin, gitoxigenin and gitaloxigenin, as found in *D. purpurea*, plus digoxigenin and diginatigenin (Figure 5.103), which do not occur in *D. purpurea*. The primary glycosides containing the acetylated tetrasaccharide unit *glucose–acetyldigitoxose–*(*digitoxose*)$_2$– are called lanatosides. Lanatosides A and C (Figure 5.103) constitute the major components in the fresh leaf (about 50–70%) and are based on the aglycones digitoxigenin and digoxigenin respectively. Lanatosides B, D, and E (Figure 5.103) are minor components derived from gitoxigenin, diginatigenin, and gitaloxigenin respectively. Enzymic hydrolysis of the lanatosides generally involves loss of the terminal glucose prior to removal of the acetyl function, so that compounds like acetyldigitoxin and acetyldigoxin, as well as digitoxin and digoxin, are present in the dried leaf as decomposition products from lanatosides A and C respectively. These transformations are also indicated in simplified form in Figure 5.104.

Digoxin (Figure 5.105) has a rapid action and is more quickly eliminated from the body than digitoxin; therefore, it is the most widely used of the cardioactive glycosides. It is more hydrophilic than digitoxin, binds less strongly to plasma proteins, and is mainly eliminated by the kidneys, whereas digitoxin is metabolized more slowly by the liver. Digoxin is used in congestive heart failure and atrial fibrillation. **Lanatoside C** and **deslanoside** (**desacetyl-lanatoside C**) (Figure 5.105) have also been employed, though not to the same extent. They have very rapid action and are suited for treatment of cardiac emergencies by injection. The semi-synthetic derivative **medigoxin** or **metildigoxin** (methyl replacing the glucose in lanatoside C; Figure 5.105) has also been available, being more active through better bioavailability.

The cardioactive glycosides increase the force of contractions in the heart, thus increasing cardiac output and allowing more rest between contractions. The primary effect on the heart appears to be inhibition of the ion transport activity of the enzyme Na^+/K^+-ATPase in the cell membranes of heart muscle, specifically inhibiting the Na^+ pump, thereby raising the intracellular Na^+ concentration. The resultant decrease in the Na^+ gradient across the cell membrane reduces the energy available for transport of Ca^{2+} out of the cell, leads to an increase in intracellular Ca^{2+} concentration, and provides the positive ionotropic effect and increased force of contractions. The improved blood circulation also tends to improve kidney function, leading to diuresis and loss of oedema fluid often associated with heart disease. However, the diuretic effect, historically important in the treatment of dropsy, is more safely controlled by other diuretic drugs.

To treat congestive heart failure, an initial loading dose of the cardioactive glycoside is followed by regular maintenance doses, the amounts administered depending on drug bioavailability and subsequent metabolism or excretion. Because of the extreme toxicity associated with these compounds (the therapeutic level is 50–60% of the toxic dose; a typical daily dose is only about 1 mg), dosage must be controlled very carefully. Bioavailability has sometimes proved erratic and can vary between different manufacturers' formulations, so patients should not be provided with different preparations during their treatment. Individual patients also excrete the glycosides or metabolize them by hydrolysis to the aglycone at different rates, and ideally these processes should be monitored. Levels of the drug in blood plasma can be measured quite rapidly by radioimmunoassay using a specific antibody. A preparation of **digoxin-specific antibody fragments** derived from sheep is available both for assay

Box 5.18 (continued)

and also as a means of reversing life-threatening digoxin overdose. It has also successfully reversed digitoxin overdose, thus demonstrating a somewhat broader specificity. The value of digoxin treatment for heart failure where the heartbeat remains regular has recently been called into question. It still remains a recognized treatment for atrial fibrillation.

Many other species of *Digitalis*, e.g. *D. dubia, D. ferruginea, D. grandiflora, D. lutea, D. mertonensis, D. nervosa, D. subalpina*, and *D. thaspi* contain cardioactive glycosides in their leaves, and some have been evaluated and cultivated for drug use. Gene engineering may be able to increase the amount of cardioactive glycosides obtainable. Thus, transgenic *D. minor* plants expressing an additional gene for the mevalonate pathway enzyme HMG-CoA reductase from *Arabidopsis thaliana* have shown increased sterol and cardenolide production.

Strophanthus

Strophanthus comprises the dried ripe seeds of *Strophanthus kombé* or *S. gratus* (Apocynaceae), which are tall vines from equatorial Africa. *S. kombé* has a history of use by African tribes as an arrow poison, and the seeds contain 5–10% cardenolides, a mixture known as K-strophanthin. This has little drug use today, though was formerly used medicinally as a cardiac stimulant. The main glycoside (about 80%) is K-strophanthoside (Figure 5.106) with smaller amounts of K-strophanthin-β and cymarin, related to K-strophanthoside as shown. These are derivatives of the aglycone strophanthidin. *S. gratus* contains 4–8% of **ouabain** (G-strophanthin; Figure 5.106), the rhamnoside of ouabigenin. Ouabigenin is rather unusual in having additional hydroxylation

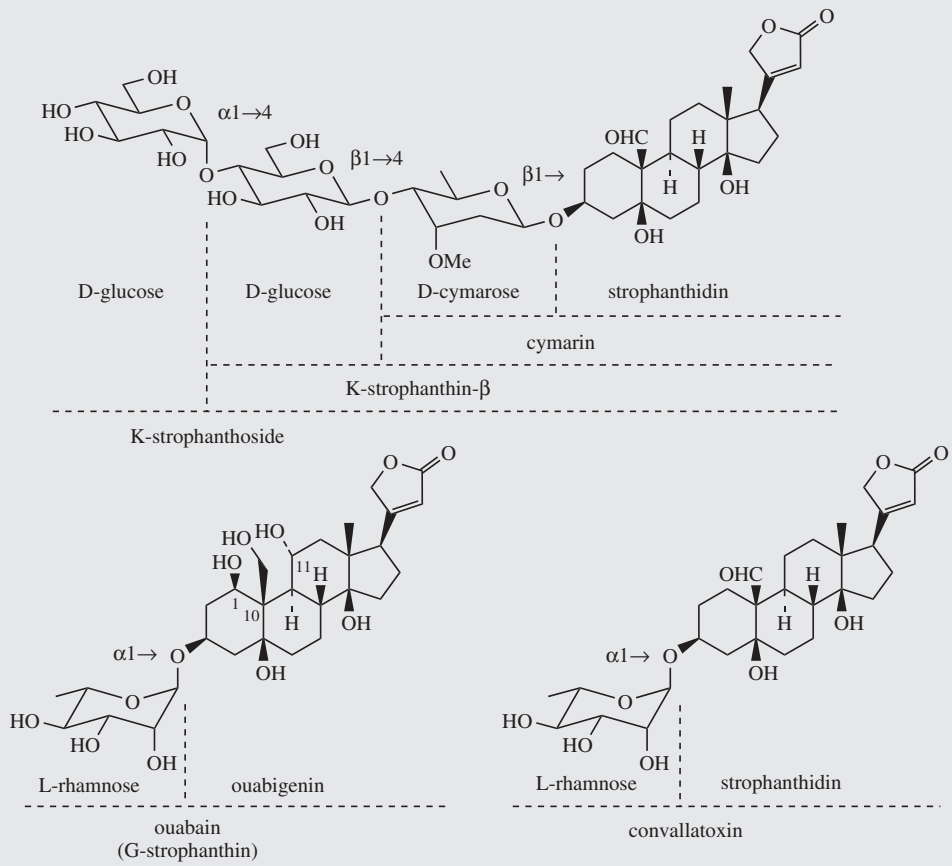


Figure 5.106

Box 5.18 (continued)

at 1β and 11α, as well as a hydroxymethyl at C-10. Ouabain is a stable, crystalline material which is often employed as the biological standard in assays for cardiac activity. It is a potent cardiac glycoside, acts quickly, but wears off rapidly. It is very polar, with rapid renal elimination and must be injected because it is so poorly absorbed orally. It has been used for emergency treatment in cases of acute heart failure.

Convallaria

The dried roots and tops of lily of the valley *Convallaria majalis* (Liliaceae/Convallariaceae) contain cardioactive glycosides (0.2–0.3%) and in the past have been used in some European countries rather than digitalis. The effects are similar, but the drug is less cumulative. This plant is widely cultivated as an ornamental, particularly for its intensely perfumed small white flowers, and must be considered potentially toxic. The major glycoside (40–50%) is convallatoxin (Figure 5.106), the rhamnoside of strophanthidin.

Squill

Squill (white squill) consists of the dried sliced bulbs of the white variety of *Urginea maritima* (formerly *Scilla maritima*; also known as *Drimia maritima*) (Liliaceae/Hyacinthaceae) which grows on seashores around the Mediterranean. The plant contains bufadienolides (up to 4%), principally scillaren A and proscillaridin A (Figure 5.107). The aglycone of scillaren A is scillarenin, which is unusual in containing a Δ^4 double bond and, thus, lacks the *cis* A/B ring fusion found in the majority of cardiac glycosides. Squill is not usually used for its cardiac properties, as the glycosides have a short duration of action. Instead, squill is employed for its expectorant action in preparations such as Gee's linctus. Large doses cause vomiting and a digitalis-like action on the heart.

Red squill is a variety of *U. maritima* which contains an anthocyanin pigment (see page 170) and bufadienolides which are different from those of the white squill. The main glycosides are glucoscilliroside and scilliroside (Figure 5.107), glucosides of scillirosidin. This chemical variety should not be present in medicinal squill, and has mainly been employed as a rodenticide. Rodents lack a vomiting reflex and are poisoned by the cardiac effects, whilst vomiting will occur in other animals and humans due to the emetic properties of the drug. The use of red squill as a rodenticide is now considered inhumane.

Toxic Plants: Cardioactive Glycosides

Many plants containing cardioactive glycosides are widely grown as ornamentals and must be considered toxic and treated with due care and respect. These include *Digitalis* species, *Convallaria majalis, Helleborus* species, and oleander (*Nerium oleander*; Apocynaceae).

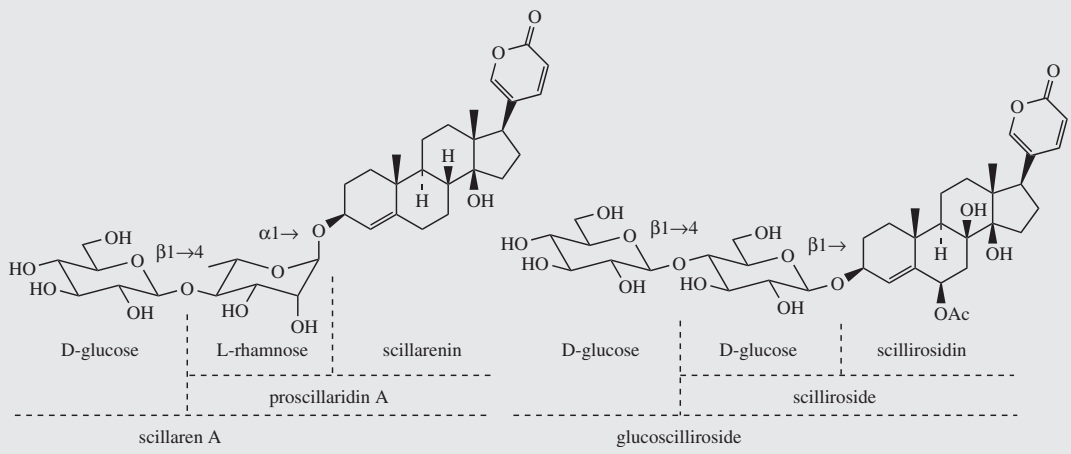
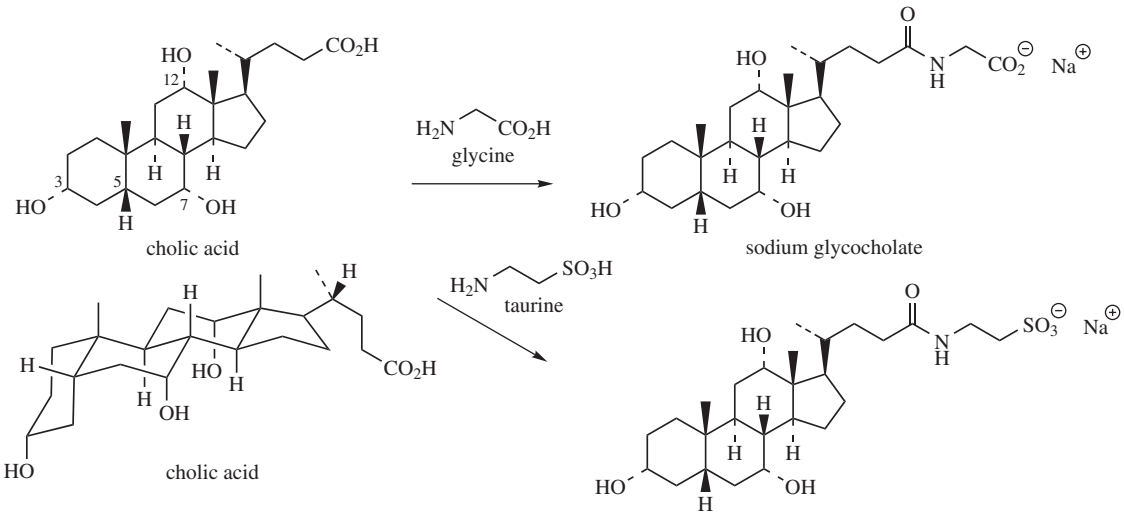


Figure 5.107



Characteristic features of bile acids:

- C_{24} cholane skeleton
- *cis*-fusion of A/B rings
- C_5-carboxylic acid side-chain
- 3α- and 7α-hydroxyls

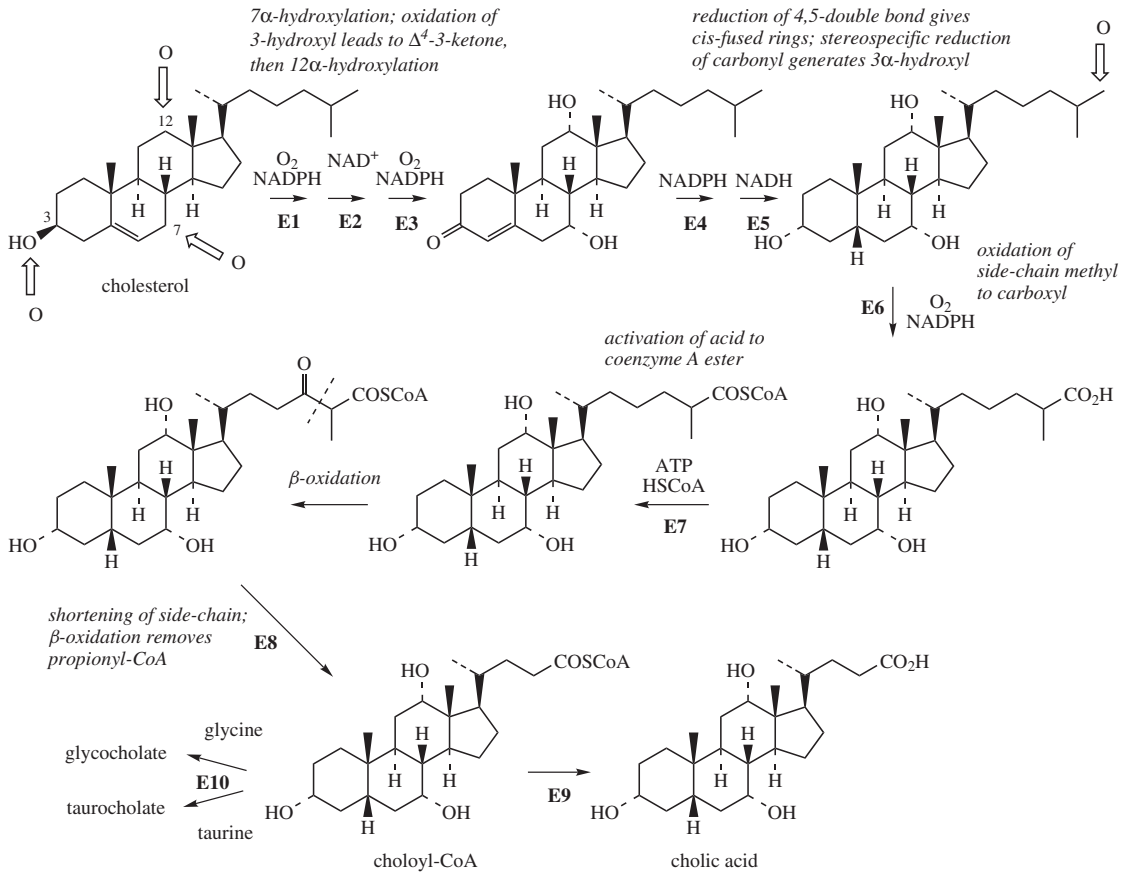
Figure 5.108

Bile Acids

The **bile acids** are C_{24} steroidal acids, e.g. **cholic acid** (Figure 5.108), which occur in salt form in bile, secreted into the gut to emulsify fats and encourage digestion [Box 5.19]. The carboxyl group is typically bound via an amide linkage to glycine (about 75%) or taurine (about 25%), e.g. cholic acid is found as **sodium glycocholate** and **sodium taurocholate**. Conjugation with glycine or taurine increases the water solubility of bile salts under physiological conditions. Taurine (2-aminoethanesulphonic acid) was first isolated from ox bile, but is now known to be widely distributed in animal tissues. Metabolism to bile acids is the principal way in which mammals degrade cholesterol absorbed from the diet. Cholesterol is extremely hydrophobic; its removal is dependent upon increasing hydrophilicity, achieved by the introduction of several polar groups into the molecule. The *cis* fusion of rings A and B confers a curvature to the steroidal skeleton, and the polar hydroxyl groups are all positioned on the lower α face, contrasting with the non-polar upper β face. Because of this ambiphilicity, they can form micelles and act as detergents.

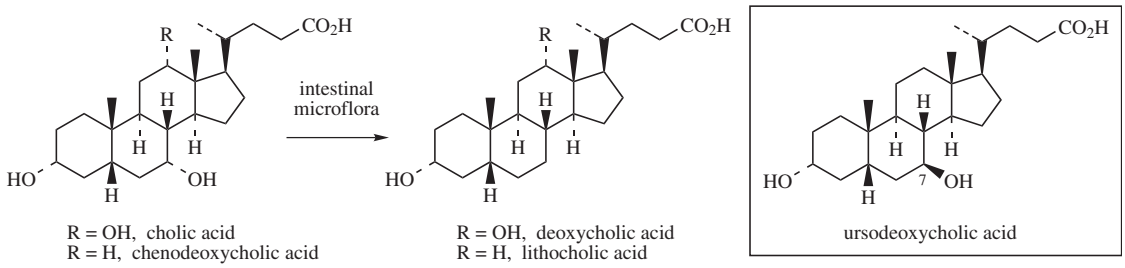
The bile acids are formed in the liver from cholesterol by a sequence which also removes three carbon atoms from the side-chain (Figure 5.109). This is achieved by initial oxidation of one of the side-chain methyl groups to an acid, followed by a β-oxidation sequence as seen with fatty acids (see Figure 2.11), removing the three-carbon unit as propionyl-CoA. Other essential features of the molecule are introduced earlier. The A/B ring system is *cis*-fused, and this is achieved by reduction of a Δ^4 rather than a Δ^5 double bond (see page 264). Migration of the double bond is accomplished via the 3-ketone; when this is reduced back to a hydroxyl, the configuration at C-3 is changed to 3α. The pathway to **chenodeoxycholic acid** (Figure 5.110) is essentially the same, though diverges early on by omitting the 12α-hydroxylation step. Alternative pathways ('acidic' pathways) to bile acids are known, though the one described (the 'classical' or 'neutral' pathway) predominates. These other pathways are characterized by initial 27-hydroxylation in the side-chain, and side-chain modifications tend to precede modifications in the sterol nucleus.

Both cholic acid and chenodeoxycholic acid are formed in the liver, stored in the gall bladder, and released into the intestine; they are termed primary bile acids. However, the 7α-hydroxyl functions of these compounds can be removed by intestinal microflora, so that mammalian bile also contains **deoxycholic acid** and **lithocholic acid** (Figure 5.110), which are termed secondary bile acids. The bile salts are then usually reabsorbed and stored in



E1: cholesterol 7α-hydroxylase (cholesterol 7α-monooxygenase)
E2: cholest-5-ene-3β,7α-diol 3β-dehydrogenase
E3: 7α-hydroxycholest-4-ene-3-one 12α-hydroxylase
E4: Δ^4-3-ketosteroid 5β-reductase
E5: 3α-hydroxysteroid dehydrogenase
E6: cholestanetriol 26-monooxygenase
E7: cholestanate-CoA ligase
E8: propanoyl-CoA *C*-acyltransferase
E9: cholate-CoA ligase
E10: bile acid-CoA: amino acid *N*-acyltransferase

Figure 5.109



R = OH, cholic acid
R = H, chenodeoxycholic acid

R = OH, deoxycholic acid
R = H, lithocholic acid

Figure 5.110

the gall bladder, although they are also excreted as the body's main means of eliminating excess cholesterol. Inability to remove cholesterol by bile acid synthesis and excretion may contribute to atherosclerosis and gallstone disease; gallstones often contain more than 70% of cholesterol (see page 251). 7α-Hydroxylation of cholesterol is the critical and rate-limiting step. This is catalysed by the cytochrome P-450-dependent monooxygenase cholesterol 7α-hydroxylase, and any deficiency of this enzyme leads to high levels of total serum and LDL cholesterol (see page 251).

Box 5.19

Bile Acids

Bile acids are obtained by purification from fresh ox bile taken from beef carcasses as a by-product of the meat trade. Bile acids are still important as starting materials for the semi-synthesis of other medicinal steroids, being a cheap and readily accessible raw material. The 7-epimer of chenodeoxycholic acid, **ursodeoxycholic acid** (ursodiol; Figure 5.110), is a minor secondary bile acid in humans, but a major component in bear bile. It is produced semi-synthetically from cholic acid or chenodeoxycholic acid and used medicinally to dissolve cholesterol gallstones as an alternative to surgery. By suppressing synthesis of both cholesterol and cholic acid, they contribute to removal of biliary cholesterol and, consequently, a gradual dissolution of gallstones which can have formed due to supersaturation. Partial or complete dissolution requires treatment over a period of many months, and is not effective for radio-opaque gallstones, which contain appreciable levels of calcium salts. Anion-exchange resins such as colestyramine (cholestyramine) and colestipol are used as cholesterol-lowering drugs to bind bile acids and prevent their reabsorption. This promotes hepatic conversion of cholesterol into bile acids, thus increasing breakdown of LDL cholesterol, and is of value in treating high-risk coronary patients.

Adrenocortical Hormones/Corticosteroids

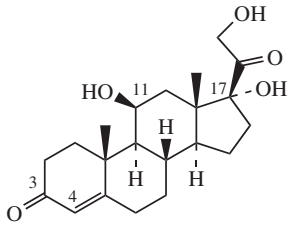
A large number of steroid hormones have been isolated and characterized from the adrenal glands. Since they are produced by the adrenal cortex, the outer part of the adrenal glands near the kidneys, they are termed **adrenocortical hormones** or **corticosteroids** [Box 5.20]. They contain a pregnane C_{21} skeleton and fall into two main activity groups, the **glucocorticoids** and the **mineralocorticoids**, although it is difficult to separate entirely the two types of activity in one molecule. Glucocorticoids are concerned with the synthesis of carbohydrate from protein and the deposition of glycogen in the liver. They also play an important role in inflammatory processes. Mineralocorticoids are concerned with the control of electrolyte balance, active compounds promoting the retention of Na^+ and Cl^-, and the excretion of K^+.

Examples of natural glucocorticoids include **hydrocortisone** (**cortisol**) and **corticosterone**, whilst **aldosterone** and **deoxycorticosterone** (**cortexone**) typify mineralocorticoids (Figure 5.111). Deoxycorticosterone has also been found in plants. Some common features of these molecules are the β-$CO.CH_2OH$ side-chain at C-17 and, frequently, an α-hydroxy also at this position. Ring A usually contains a Δ^4-3-keto functionality. The 11β-hydroxy is essential for glucocorticoid activity. In aldosterone, the principal mineralocorticoid hormone, the methyl group (C-18) has been oxidized to an aldehyde, and this is able to react with the 11β-hydroxyl, so that aldosterone exists predominantly in the hemiacetal form. This essentially eliminates the glucocorticoid activity.

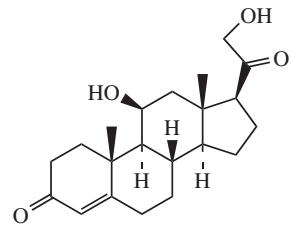
The corticosteroids are produced from cholesterol via **pregnenolone** and **progesterone**. This involves side-chain cleavage as seen in the biosynthesis of cardioactive glycosides (see page 267), and the same sequence of reactions is operative. From progesterone, the formation of **deoxycorticosterone, corticosterone**, and **hydrocortisone** (**cortisol**) (Figure 5.112) requires only a series of hydroxylation steps, catalysed by cytochrome P-450-dependent hydroxylases with NADPH and O_2 cofactors. Thus, positions 17, 21, and 11 may be hydroxylated, and the exact order can in fact vary from that shown in Figure 5.112, according to species. It can be seen that production of hydrocortisone from cholesterol actually utilizes cytochrome P-450-dependent enzymes in four of the five steps. The further oxidation of C-18 to an aldehyde via the alcohol allows formation of **aldosterone** from corticosterone, again involving a P-450 system.

Semi-Synthesis of Corticosteroids

The medicinal use of corticosteroids was stimulated by reports of the dramatic effects of **cortisone** (Figure 5.113) on patients suffering from rheumatoid arthritis in the late 1940s and early 1950s. The cortisone employed



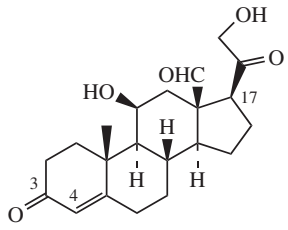
hydrocortisone
(cortisol)



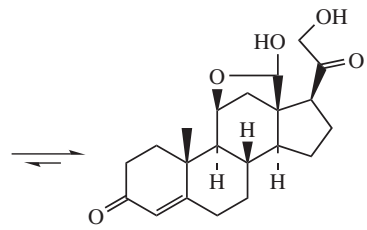
corticosterone

Characteristic features of glucocorticoids:

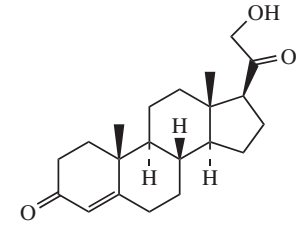
- C_{21} pregnane skeleton
- 17β-CO.CH_2OH side-chain
- 11β-hydroxyl
- Δ^4-3-keto (usually)
- 17α-hydroxyl (usually)



aldosterone



aldosterone
(hemiacetal form)



deoxycorticosterone
(cortexone)

Characteristic features of mineralocorticoids:

- C_{21} pregnane skeleton
- 17β-CO.CH_2OH side-chain
- Δ^4-3-keto (usually)

Figure 5.111

was isolated from the adrenal glands of cattle, and later was produced semi-synthetically by a laborious process from **deoxycholic acid** (see page 276) isolated from ox bile, a sequence necessitating over 30 chemical steps. In due course, it was shown that cortisone itself was not the active agent; it was reduced in the liver to **hydrocortisone** as the active agent (Figure 5.113). Two dehydrogenase enzymes regulate hydrocortisone levels, both of them acting in a unidirectional sense. One of these activates cortisone by reduction, and though prevalent in the liver is found in a wide range of tissues. The other enzyme is found predominantly in mineralocorticoid target tissues, e.g. kidney and colon, where it protects receptors against excess hydrocortisone by converting it to inactive cortisone. Increased demand for cortisone and hydrocortisone (cortisol) led to exploitation of alternative raw materials, particularly plant sterols and saponins. A major difficulty in any semi-synthetic conversion was the need to provide the 11β-hydroxyl group, which was essential for glucocorticoid activity.

Sarmentogenin (Figure 5.114) had been identified as a natural 11-hydroxy cardenolide in *Strophanthus sarmentosus*, but it was soon appreciated that the amounts present in the seeds, and the limited quantity of plant material available, would not allow commercial exploitation of this compound. As an alternative to using a natural 11-oxygenated substrate, compounds containing a 12-oxygen substituent might be used instead, in that this group provides activation and allows chemical modification at the adjacent site. Indeed, this was a feature of the semi-synthesis of cortisone from deoxycholic acid, which contains a 12α-hydroxyl. However, it was the 12-keto steroidal sapogenin **hecogenin** (Figure 5.114) from sisal (*Agave sisalana*; Agavaceae) (see page 262) that made possible the economic production of cortisone on a commercial scale. This material is still used

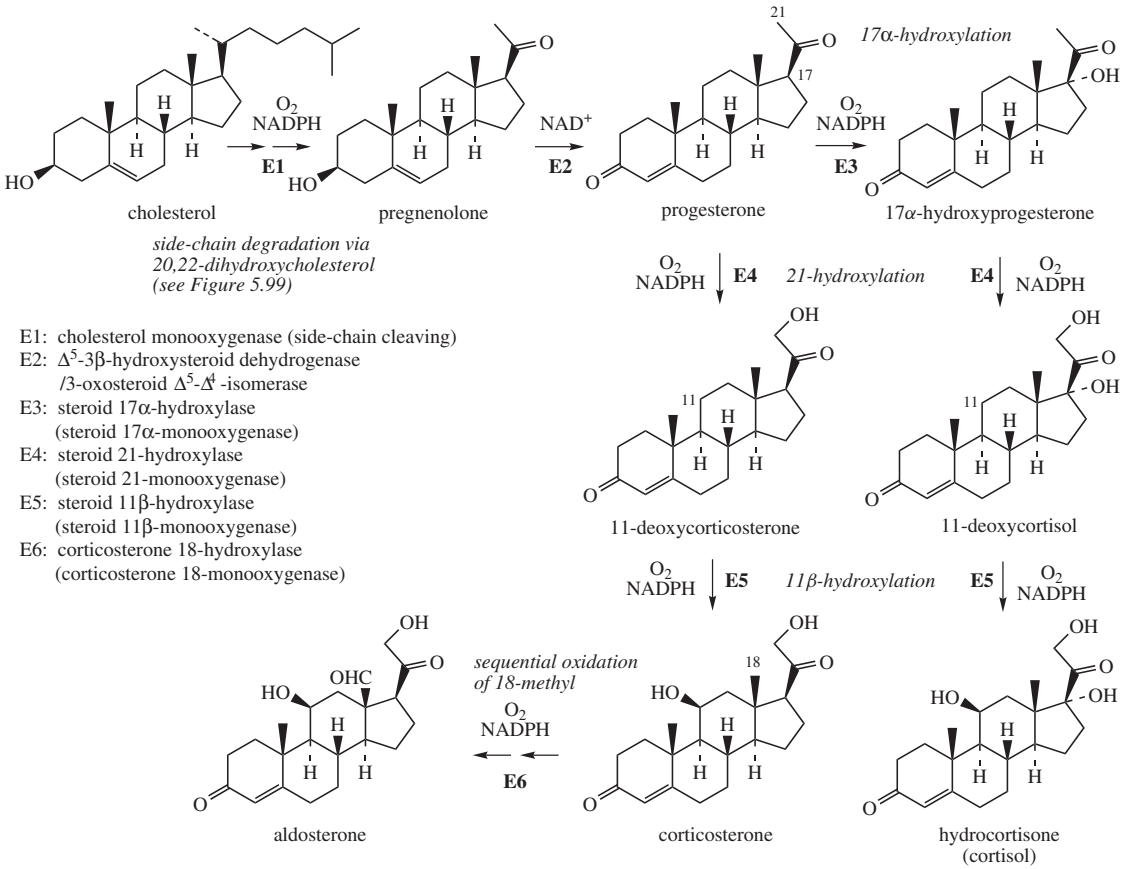


Figure 5.112

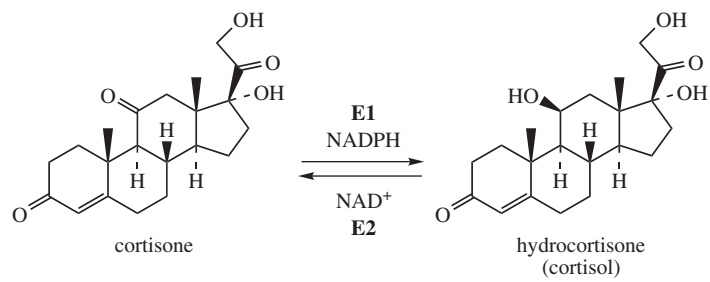


Figure 5.113

in the semi-synthesis of steroidal drugs, and the critical modifications in ring C are shown in Figure 5.115. Bromination α to the 12-keto function generates the 11α-bromo derivative, which on treatment with base gives the 12-hydroxy-11-ketone by a base-catalysed keto–enol tautomerism mechanism. The 12-hydroxyl is then removed by hydride displacement of the acetate using calcium in liquid ammonia. The 11-keto sapogenin derived

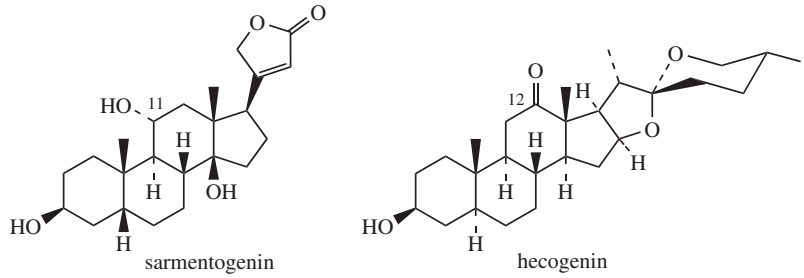


Figure 5.114

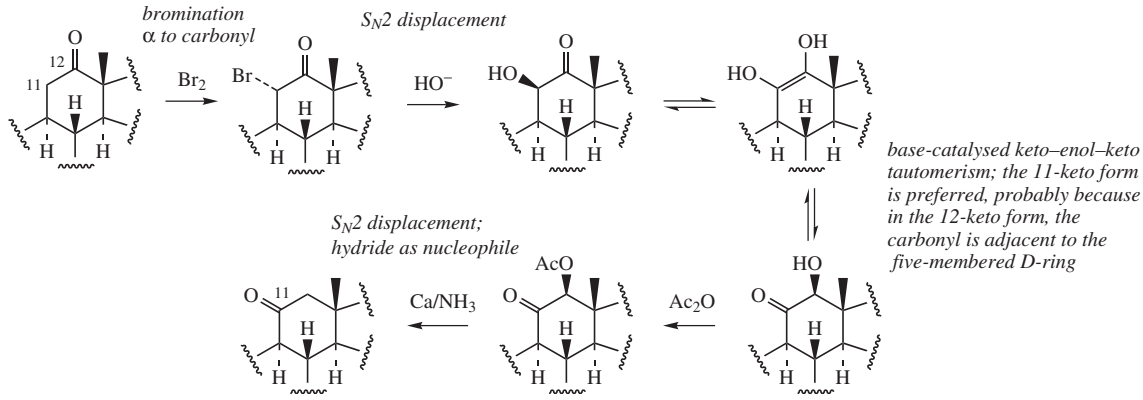


Figure 5.115

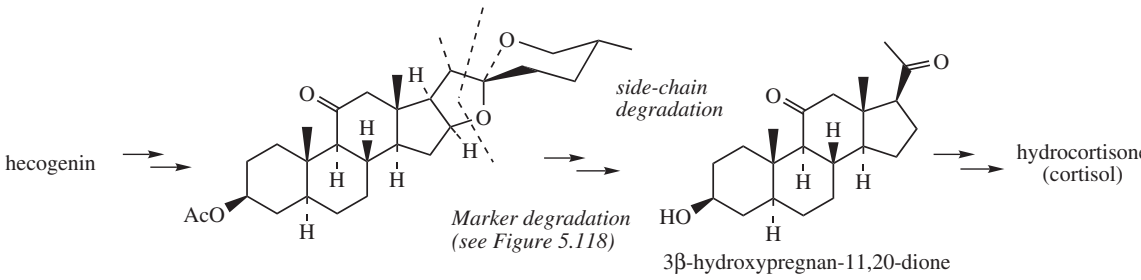


Figure 5.116

by this sequence is subjected to the side-chain degradation used with other sapogenins, e.g. diosgenin (see later, Figure 5.118), giving the 11-ketopregnane (Figure 5.116). This compound can then be used for conversion into cortisone, hydrocortisone, and other steroid drugs.

Of much greater importance was the discovery in the mid 1950s that hydroxylation at C-11 could be achieved via a microbial fermentation. **Progesterone** was transformed by *Rhizopus arrhizus* into **11α-hydroxyprogesterone** (Figure 5.117) in yields of up to 85%. More recently, *R. nigricans* has been employed to give even higher yields. 11α-Hydroxyprogesterone is then converted into hydrocortisone by chemical means, the 11β configuration being introduced via oxidation to the 11-ketone followed by a stereospecific reduction step.

Progesterone could be obtained in good yields (about 50%) from **diosgenin** extracted from Mexican yams (*Dioscorea* species; Dioscoreaceae; see page

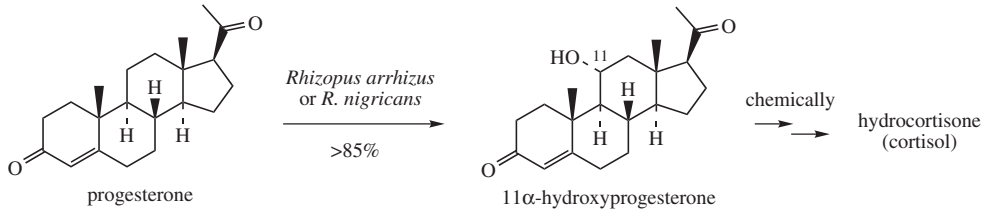


Figure 5.117

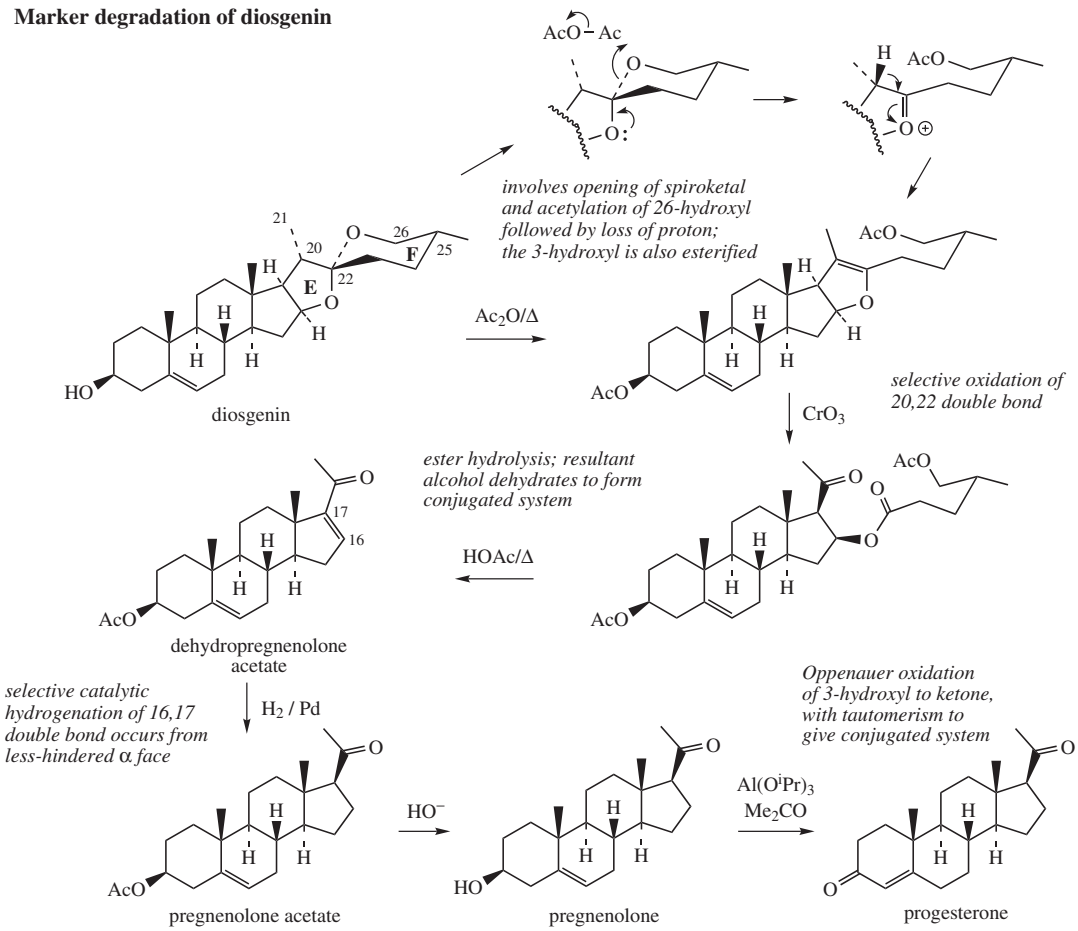


Figure 5.118

262) or **stigmasterol** from soya beans (*Glycine max*; Leguminosae/Fabaceae; see page 255). Steroidal sapogenins such as diosgenin may be degraded by the **Marker degradation** (Figure 5.118), which removes the spiroketal portion, leaving carbon atoms C-20 and C-21 still attached to contribute to the pregnane system. Initial treatment with acetic anhydride produces the 3,26-diacetate, by opening the ketal, dehydrating in ring E and acetylating the remaining hydroxyl groups. The double bond in ring E is then selectively oxidized to give a product, which now contains the unwanted side-chain carbon atoms as an ester

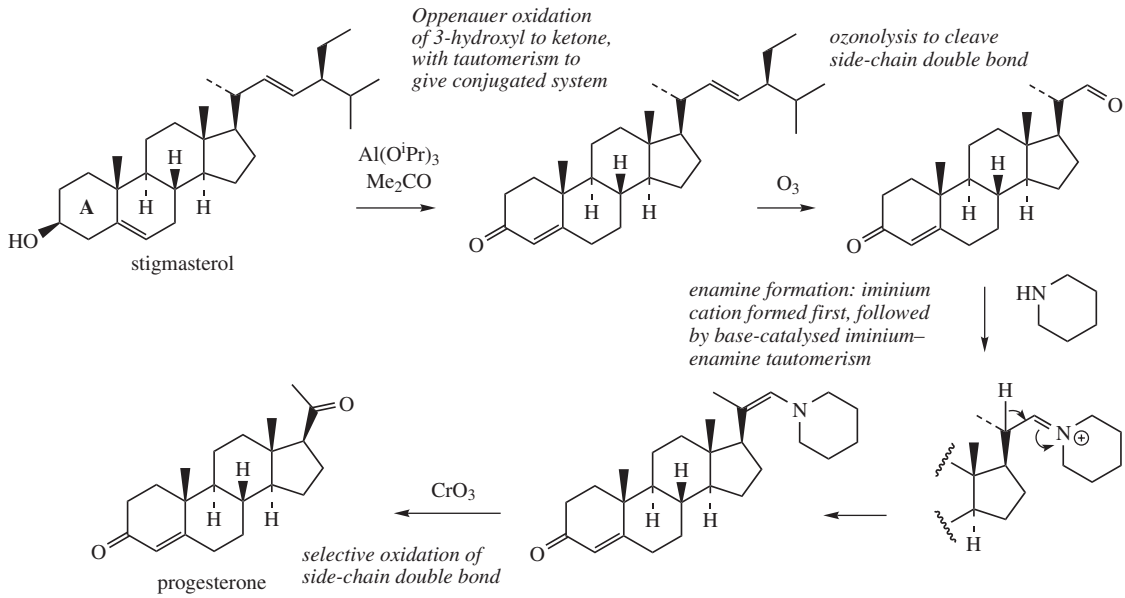


Figure 5.119

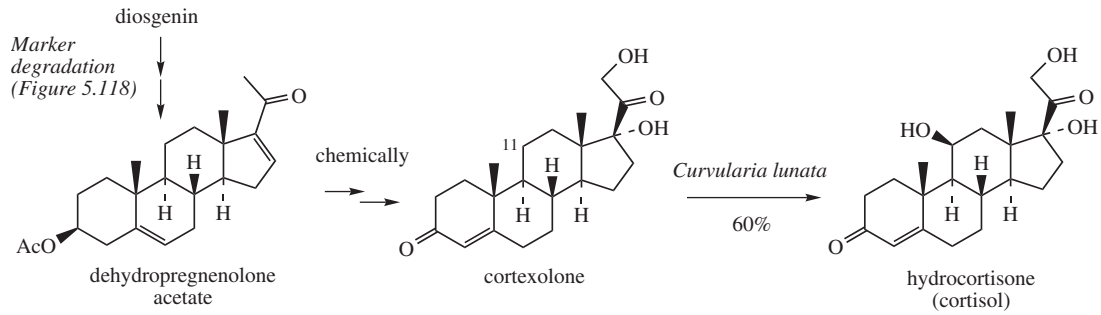
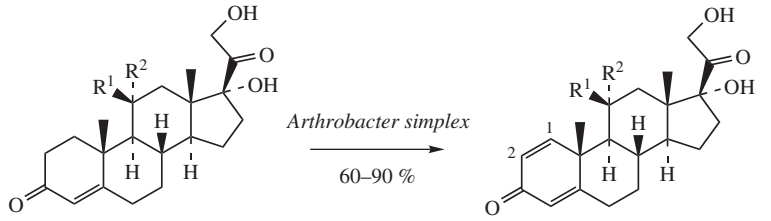


Figure 5.120

function, easily removed by hydrolysis. Under the conditions used, the product is the α,β-unsaturated ketone dehydropregnenolone acetate. Hydrogenation of the double bond is achieved in a regioselective and stereoselective manner, addition of hydrogen being from the less-hindered α-face to give pregnenolone acetate. **Progesterone** is obtained by hydrolysis of the ester function and Oppenauer oxidation to give the preferred α,β-unsaturated ketone (compare page 264). It is immediately obvious from Figure 5.119 that, since the objective is to remove the unwanted ring F part of the sapogenin, features like the stereochemistry at C-25 are irrelevant, and the same general degradation procedure can be used for other sapogenins. It is equally applicable to the nitrogen-containing analogues of sapogenins, e.g. **solasodine** (Figure 5.95). In such compounds, the stereochemistry at C-22 is also quite immaterial.

Degradation of the sterol **stigmasterol** to progesterone is achieved by the sequence shown in Figure 5.119. The double bond in the side-chain allows cleavage by ozonolysis, and the resultant aldehyde is chain-shortened via formation of an enamine with piperidine. This can be selectively oxidized to progesterone. In this sequence, the ring A transformations are more conveniently carried out as the first reaction. A similar route can be used for the fungal sterol **ergosterol**, though an additional step is required for reduction of the Δ^7 double bond.



R^1R^2 = O, cortisone
R^1 = OH, R^2 = H, hydrocortisone (cortisol)

R^1R^2 = O, prednisone
R^1 = OH, R^2 = H, prednisolone

Figure 5.121

An alternative sequence from diosgenin to hydrocortisone has been devised, making use of another microbiological hydroxylation, this time a direct 11β-hydroxylation of the steroid ring system (Figure 5.120). The fungus *Curvularia lunata* is able to 11β-hydroxylate **cortexolone** to **hydrocortisone** in yields of about 60%. Although a natural corticosteroid, cortexolone may be obtained in large amounts by chemical transformation from 16-dehydropregnenolone acetate, an intermediate in the Marker degradation of diosgenin (Figure 5.118).

Some steroid drugs are produced by total synthesis, but, in general, the ability of microorganisms to biotransform steroid substrates has proved invaluable in exploiting inexpensive natural steroids as sources of drug materials. It is now possible via microbial fermentation to hydroxylate the steroid nucleus at virtually any position and with defined stereochemistry. These processes are, in general, more expensive than chemical transformations, and are only used commercially when some significant advantage is achieved, e.g. replacement of several chemical steps. For example, the therapeutic properties of cortisone and hydrocortisone can be further improved by the microbial introduction of a 1,2-double bond, giving **prednisone** and **prednisolone** respectively (Figure 5.121). These agents surpass the parent hormones in antirheumatic and antiallergic activity with fewer side-effects. As with cortisone, prednisone is converted in the body by reduction into the active agent, in this case prednisolone.

Box 5.20

Corticosteroid Drugs

Glucocorticoids are primarily used for their antirheumatic and anti-inflammatory activities. They give valuable relief to sufferers of rheumatoid arthritis and osteoarthritis, and find considerable use for the treatment of inflammatory conditions by suppressing the characteristic development of swelling, redness, heat, and tenderness. One of the key ways in which they exert their action is by interfering with prostaglandin biosynthesis, via production of a peptide that inhibits the phospholipase enzyme responsible for release of arachidonic acid from phospholipids (see page 62). However, these agents merely suppress symptoms; they do not provide a cure for the disease. Long-term usage may result in serious side-effects, including adrenal suppression, osteoporosis, ulcers, fluid retention, and increased susceptibility to infections. Because of these problems, steroid drugs are rarely the first choice for inflammatory treatment, and other therapies are usually tried first. Nevertheless, corticosteroids are widely used for inflammatory conditions affecting the ears, eyes, and skin, and in the treatment of burns. Some have valuable antiallergic properties, helping in reducing the effects of hay fever and asthma. In some disease states, e.g. Addison's disease, the adrenal cortex is no longer able to produce these hormones, and replacement therapy becomes necessary. The most common genetic deficiency is lack of the 21-hydroxylase enzyme in the biosynthetic pathway, necessary for both hydrocortisone and aldosterone biosynthesis (Figure 5.112). This can then lead to increased synthesis of androgens (see Figure 5.141).

Mineralocorticoids are primarily of value in maintaining electrolyte balance where there is adrenal insufficiency.

Natural corticosteroid drugs **cortisone** (as **cortisone acetate**) and **hydrocortisone** (**cortisol**) (Figure 5.113) are valuable in replacement therapies, and hydrocortisone is one of the most widely used agents for topical application in the treatment of inflammatory skin conditions. The early use of the natural corticosteroids for anti-inflammatory activity tended to show up some serious side-effects on water, mineral, carbohydrate, protein, and fat metabolism. In particular, the mineralocorticoid activity is usually considered an undesirable effect. In an effort to optimize anti-inflammatory activity, many thousands of chemical

Box 5.20 (continued)

modifications to the basic structure were tried. Three structural changes proved particularly valuable. Introduction of a Δ^1 double bond modifies the shape of ring A and was found to increase glucocorticoid over mineralocorticoid activity, e.g. **prednisone** and **prednisolone** (Figure 5.121), which are about four times more potent than cortisone/hydrocortisone as anti-inflammatory agents. A 9α-fluoro substituent increased all activities, whereas 16α- or 16β-methyl groups reduced the mineralocorticoid activity without affecting the glucocorticoid activity.

The discovery that 9α-fluoro analogues had increased activity arose indirectly from attempts to epimerize 11α-hydroxy compounds into the active 11β-hydroxy derivatives (Figure 5.122). Thus, when an 11α-tosylate ester was treated with acetate, a base-catalysed elimination was observed rather than the hoped-for substitution, which is hindered by the axial methyl groups. This *syn* elimination suggests that an E1 mechanism is involved. The same $\Delta^{9(11)}$-ene can also be obtained by dehydration of the 11β-alcohol using thionyl chloride. The alkene was then treated with aqueous bromine to generate an 11-hydroxy compound. Addition of HOBr to the 9(11)-double bond proceeds via electrophilic attack from the less-hindered α-face, giving the cyclic bromonium ion, and then ring opening involves β-attack of hydroxide at C-11. Attack at C-9 is sterically hindered by the methyl at C-10. The 9α-bromocortisol (as its 21-acetate) produced in this way was less active as an anti-inflammatory agent than cortisol 21-acetate by a factor of three, and 9α-iodocortisol acetate was also less active by a factor of 10. For fluoro compounds, fluorine has to be introduced indirectly; this was achieved from the β-epoxide formed by base treatment of the 9α-bromo-11β-hydroxy analogue (Figure 5.122). The derivative 9α-fluorocortisol 21-acetate (*flu*orohy*drocortisone* acetate; **fludrocortisone acetate**; Figure 5.123) was found to be about 11 times more active than cortisol acetate. However, its mineralocorticoid activity was also increased some 300-fold, so that its potential anti-inflammatory activity has no clinical relevance, and this drug is only employed for its mineralocorticoid activity; in practice, it is the only mineralocorticoid agent routinely used. The introduction of a 9α-fluoro substituent into prednisolone causes powerful Na^+ retention. These effects can be reduced (though usually not eliminated entirely) by introducing a substituent at C-16, either a 16α-hydroxy or a 16α/16β-methyl.

The 16α-hydroxyl can be introduced microbiologically, e.g. as in the conversion of 9α-fluoroprednisolone into **triamcinolone** (Figure 5.124). The ketal formed from triamcinolone and acetone, **triamcinolone acetonide** (Figure 5.124), provides the most satisfactory means of administering this anti-inflammatory. **Methylprednisolone** (Figure 5.123) is a 6α-methyl derivative of prednisolone showing a modest increase in activity over the parent compound. A 6-methyl group can be supplied via reaction

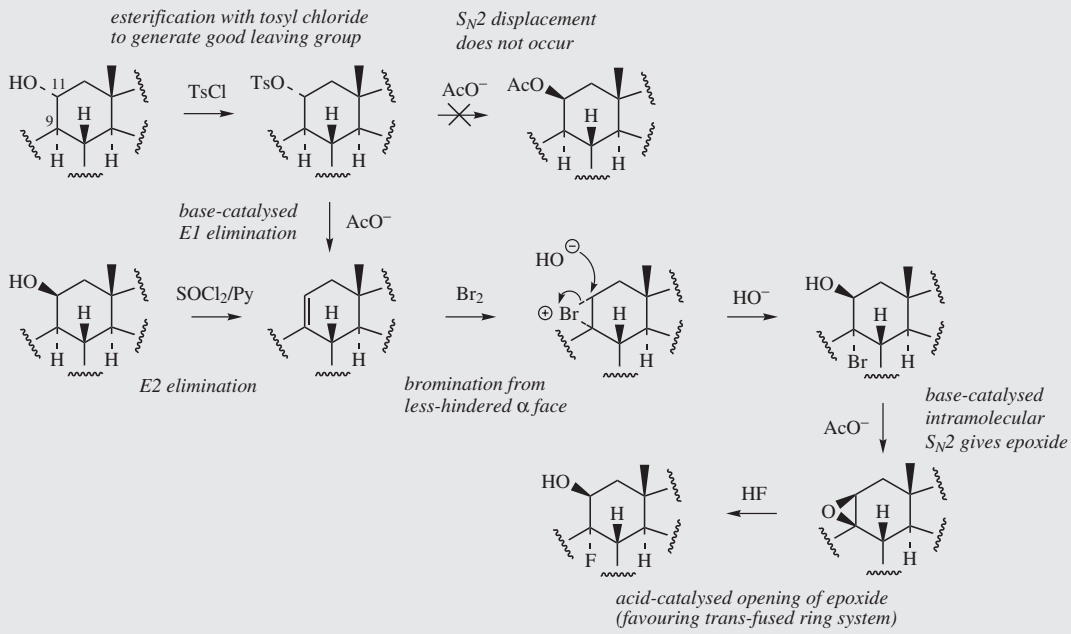


Figure 5.122

Box 5.20 (continued)

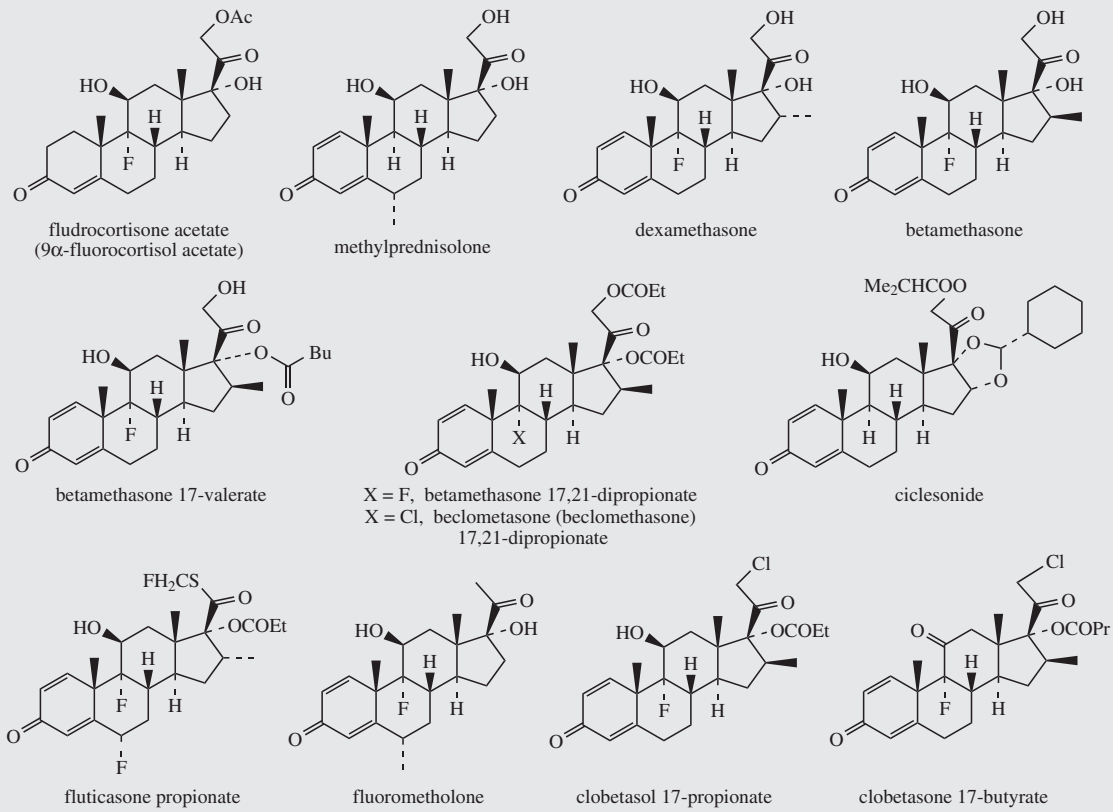


Figure 5.123

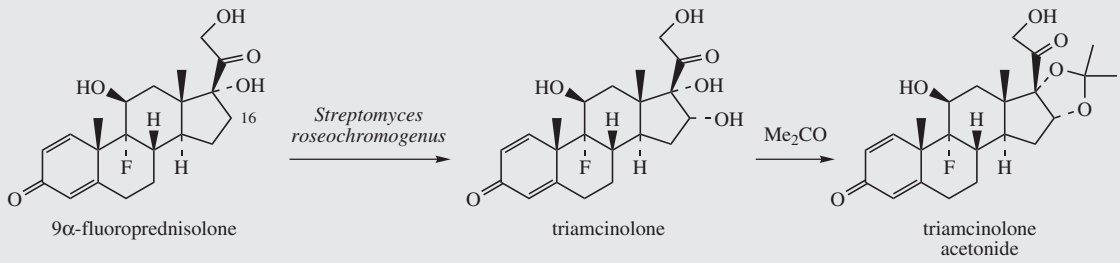


Figure 5.124

of the Grignard reagent MeMgBr with a suitable 5,6-epoxide derivative. **Dexamethasone** and **betamethasone** (Figure 5.123) exemplify respectively 16α- and 16β-methyl derivatives in drugs with little, if any, mineralocorticoid activity. The 16-methyl group is easily introduced by a Grignard reaction with an appropriate α,β-unsaturated Δ^{16}-20-ketone. Betamethasone, for topical application, is typically formulated as a C-17 ester with valeric acid (**betamethasone 17-valerate**), or as the 17,21-diester with

Box 5.20 (continued)

propionic acid (**betamethasone 17,21-dipropionate**; Figure 5.123). The 9α-chloro compound **beclometasone 17,21-dipropionate** (**beclomethasone 17,21-dipropionate**) is an important agent with low systemic distribution used as an inhalant for the control of asthma. **Ciclesonide** (Figure 5.123) is a promising asthma drug that is essentially devoid of oral activity, but is activated by endogenous esterases through hydrolysis of the 21-ester function. **Fluticasone propionate** (Figure 5.123) is also used in asthma treatment, and is representative of compounds where the 17-side-chain has been modified to a carbothiate (sulfur ester).

Although the anti-inflammatory activity of hydrocortisone is lost if the 21-hydroxyl group is not present, considerable activity is restored when a 9α-fluoro substituent is introduced. **Fluorometholone** (Figure 5.123) is a corticosteroid which exploits this relationship and is of value in eye conditions. Other agents are derived by replacing the 21-hydroxyl with a halogen, e.g. **clobetasol 17-propionate** and **clobetasone 17-butyrate** (Figure 5.123) which are effective topical drugs for severe skin disorders.

Many other corticosteroids are currently available for drug use. The structures of some of these are given in Figure 5.125, grouped according to the most characteristic structural features, namely 16-methyl, 16-hydroxy, and 21-chloro derivatives. In **rixemolone** (Figure 5.126), a recently introduced anti-inflammatory for ophthalmic use, neither a 21-hydroxy nor a 9α-fluoro substituent is present, but instead there are methyl substituents at positions 21, 17α, and 16α. Rimexolone has significant advantages in eye conditions over drugs such as dexamethasone, in that it does not significantly raise intraocular pressure. The recently introduced **deflazacort** (Figure 5.126) is a drug with high glucocorticoid activity, but does not conveniently fit into any of the general groups in that it contains an oxazole ring spanning C-16 and C-17.

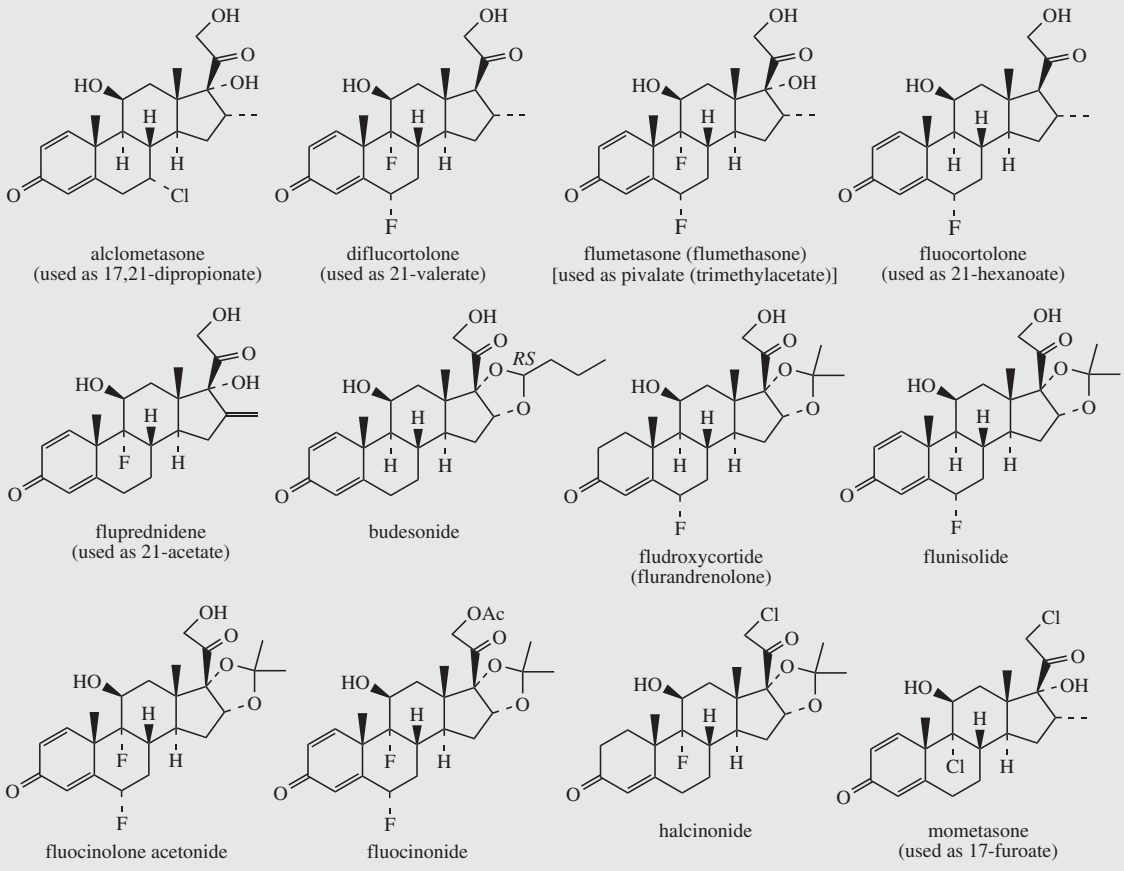


Figure 5.125

Box 5.20 (continued)

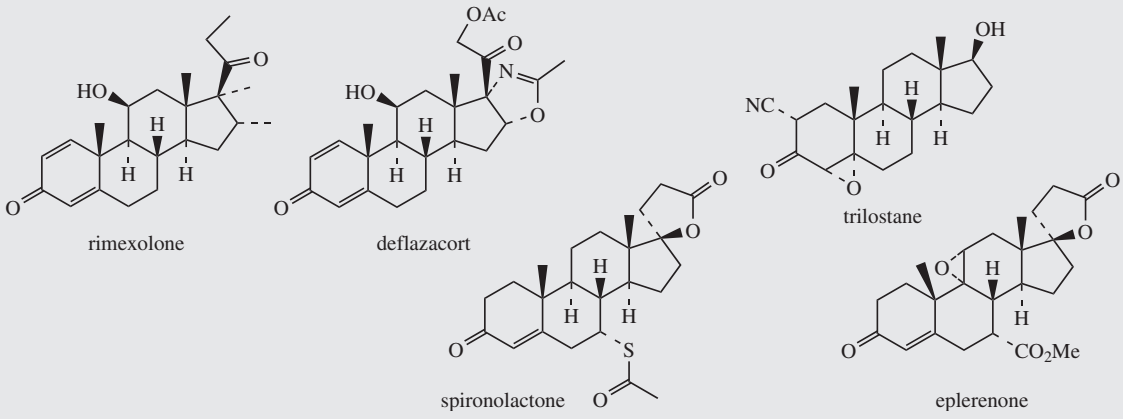
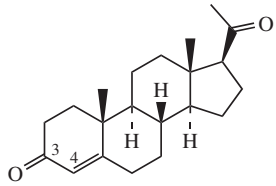


Figure 5.126

Trilostane (Figure 5.126) is an adrenocortical suppressant which inhibits synthesis of glucocorticoids and mineralocorticoids and has value in treating Cushing's syndrome, a condition characterized by a moon-shaped face and caused by excessive glucocorticoids. This drug is an inhibitor of the dehydrogenase–isomerase which transforms pregnenolone into progesterone (Figure 5.112).

Spironolactone (Figure 5.126) is an antagonist of the endogenous mineralocorticoid aldosterone and inhibits the sodium-retaining action of aldosterone whilst also decreasing the potassium-secreting effect. Classified as a potassium-sparing diuretic, it is employed in combination with other diuretic drugs to prevent excessive potassium loss. Low doses of spironolactone are beneficial in severe heart failure. Progesterone (page 288) is also an aldosterone antagonist; the spironolactone structure differs from progesterone in its 7α-thioester substituent, and replacement of the 17β side-chain with a 17α-spirolactone. **Eplerenone** (Figure 5.126), a 9,10-epoxy spironolactone analogue, is a newer aldosterone antagonist that has fewer side-effects than spironolactone due to more selective binding to the mineralocorticoid receptor. It is used for the treatment of hypertension and heart failure.



Characteristic features of progestogens:

- C_{21} pregnane skeleton
- Δ^4-3-keto

Figure 5.127

Progestogens

Progestogens (progestins; gestogens) are female sex hormones, concerned with preparing the uterus for pregnancy, and then maintaining the necessary conditions [Box 5.21]. There is only one naturally occurring progestational steroid and that is **progesterone** (Figure 5.127), which is secreted by the corpus luteum following release of an ovum. Progesterone is also an intermediate in the biosynthesis of the corticosteroids, e.g. hydrocortisone and aldosterone (see page 279), and its derivation from cholesterol via pregnenolone has also been seen in the formation of cardioactive glycosides (see page 267).

Box 5.21

Progestogen Drugs

Quantities of **progesterone** (Figure 5.127) for drug use are readily available by semi-synthesis using the Marker degradation (see page 281). However, progesterone is poorly absorbed, is rapidly metabolized in the liver, and is not suitable for oral use. Many semi-synthetic analogues have been produced, and it was thus appreciated that the α,β-unsaturated ketone system in ring A was essential for activity. The side-chain function at C-17 could be modified, and ethisterone (17α-ethynyltestosterone; Figure 5.128), originally developed as a potential androgen, was found to be active orally as a progestational agent. This structure incorporates an ethynyl side-chain at C-17, a feature of several semi-synthetic steroidal hormones used as drugs. This group, referred to as 'ethinyl' in drug molecules, is introduced by nucleophilic attack of acetylide anion onto a C-17 carbonyl (Figure 5.128), attack coming from the α-face, the methyl C-18 hindering approach from the β-face. The substrate androstenolone is readily obtained from the Marker degradation intermediate dehydropregnenolone acetate (Figure 5.118). The oxime (Figure 5.128) is treated with a sulfonyl chloride in pyridine and undergoes a Beckmann rearrangement in which C-17 migrates to the nitrogen, giving the amide. This amide is also an enamine and can be hydrolysed to the 17-ketone. Acetylation or other esterification of the 17-hydroxyl in progestogens increases lipid solubility and extends the duration of action by inhibiting metabolic degradation.

Though considerably better than progesterone, the oral activity of ethisterone is still relatively low, and better agents were required. An important modification from ethisterone was the 19-nor analogue, **norethisterone** (US: norethindrone) and its ester **norethisterone acetate** (US: norethindrone acetate) (Figure 5.129). Attention was directed to the 19-norsteroids by the observation that 19-nor-14β,17α-progesterone (Figure 5.129), obtained by degradation of the cardioactive glycoside strophanthidin (see page 273), displayed eight times higher progestational activity than progesterone, despite lacking the methyl C-19, and having the unnatural configurations at the two centres C-14 (C/D rings *cis*-fused) and C-17. Norethisterone can be synthesized from the oestrogen estrone (see page 291), which already lacks the C-9 methyl, or from androstenolone (Figure 5.128) by a sequence which allows oxidation of C-19 to a carboxyl, which is readily lost by decarboxylation when adjacent to the α,β-unsaturated ketone system.

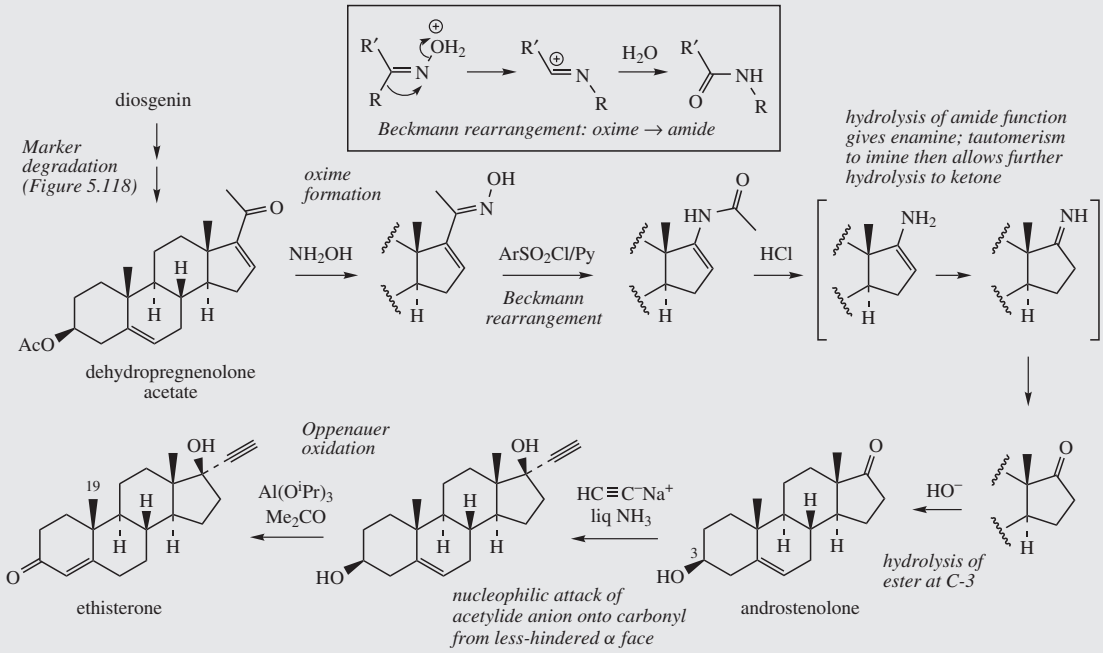


Figure 5.128

Box 5.21 (continued)

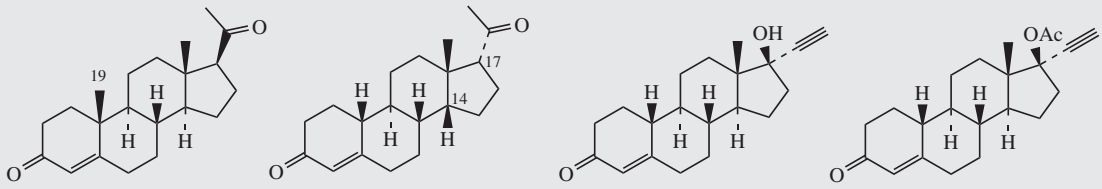


Figure 5.129

Although ethisterone and norethisterone are structurally C_{21} pregnane derivatives, they may also be regarded as 17-ethynyl derivatives of testosterone (see page 296), the male sex hormone, and 19-nortestosterone respectively. It is thus convenient to classify the progestogens as either progesterone derivatives or 17-ethynyl-testosterone derivatives, with the latter group then being subdivided into estranes or 13-ethylgonanes. Structures of some currently available progestogen drugs are shown in Figure 5.130. Semi-synthetic progesterone structures, still containing the 17-acetyl side-chain, tend to be derivatives of 17α-hydroxyprogesterone, a biosynthetic intermediate on the way to hydrocortisone (Figure 5.112), which also has progesterone-like activity. Relatively few examples of this class remain in use. **Medroxyprogesterone acetate** contains an additional 6α-methyl, introduced to block potential deactivation by metabolic hydroxylation, and is 100–300 times as potent as ethisterone on oral administration. **Megestrol acetate** contains a 6-methyl group and an additional Δ^6 double bond, whilst **dydrogesterone** unusually has different stereochemistries at C-9 and C-10. Norethisterone and norethisterone acetate (Figure 5.129), along with **etynodiol (ethynodiol) diacetate** (Figure 5.130), are 17-ethynylestranes. **Norgestrel** (Figure 5.130) is representative of 17-ethynyl progestogens with an ethyl group replacing the 13-methyl. Although these can be obtained by semi-synthesis from natural 13-methyl compounds, norgestrel is produced by total synthesis as the racemic compound. Since only the laevorotatory enantiomer which has the natural configuration is biologically active, this enantiomer, **levonorgestrel**, is now replacing the racemic form for drug use. In **desogestrel**, further features are the modification of an 11-oxo function to an 11-methylene, and removal of the 3-ketone. **Norgestimate** and **norelgestromin** are characterized by conversion of the 3-keto into an oxime group.

During pregnancy, the corpus luteum continues to secrete progesterone for the first 3 months, after which the placenta becomes the supplier of both progesterone and oestrogen. Progesterone prevents further ovulation and relaxes the uterus to prevent the fertilized egg being dislodged. In the absence of pregnancy, a decline in progesterone levels results in shedding of the uterine endometrium and menstruation. Progestogens are useful in many menstrual disorders and as **oral contraceptives**, either alone at low dosage (progestogen-only contraceptives, e.g. norethisterone, levonorgestrel) or in combination with oestrogens (combined oral contraceptives, e.g. ethinylestradiol + norethisterone, ethinylestradiol + levonorgestrel). The combined oestrogen–progestogen preparation inhibits ovulation, but normal menstruation occurs when the drug is withdrawn for several days each month. The low-dosage progestogen-only pill appears to interfere with the endometrial lining to inhibit fertilized egg implantation and thickens cervical mucus, making a barrier to sperm movement. The progestogen-only formulation is less likely to cause thrombosis, a serious side-effect sometimes experienced from the use of oral contraceptives. There appears to be a slightly higher risk of thrombosis in patients using the so-called 'third-generation' oral contraceptive pills containing the newer progestogens **desogestrel** and **gestodene**. Current oral contraceptives have a much lower hormone content than the early formulations of the 1960s and 1970s, typically about 10% of the progestogen and 50% of the oestrogen content. Deep muscular injections of medroxyprogesterone or norethisterone esters and implants of **etonogestrel** can be administered to provide long-acting contraception. A high dose of **levonorgestrel** is the drug of choice for emergency contraception after unprotected intercourse, i.e. the 'morning-after' pill. Hormone replacement therapy (HRT) in non-hysterectomized women also involves progestogen–oestrogen combinations (see page 294), whilst progestogens such as norethisterone, megestrol acetate, and medroxyprogesterone acetate also find limited application in the treatment of breast cancers.

The structure of **drospirenone** (Figure 5.130) is quite unlike other progestogens, but is based on that of the aldosterone antagonist spironolactone (Figure 5.126, page 287). This new agent is a progestogen with antimineralocorticoid activity. The oestrogen content in a combined contraceptive pill tends to increases aldosterone levels, causing side-effects such as sodium and water retention, leading to swelling and increased blood pressure; these side-effects are thus reduced. Drospirenone also displays antiandrogenic activity, and this may be of value in patients who suffer effects such as androgen-related skin disorders.

Box 5.21 (continued)

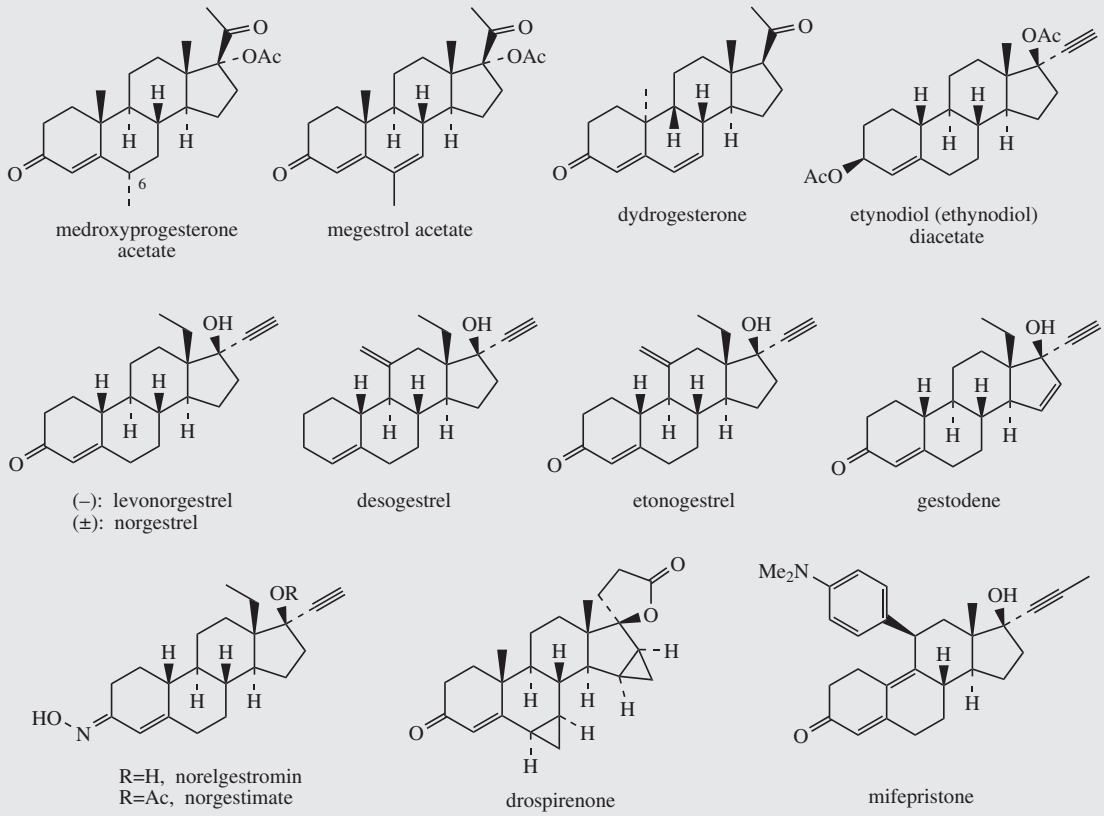


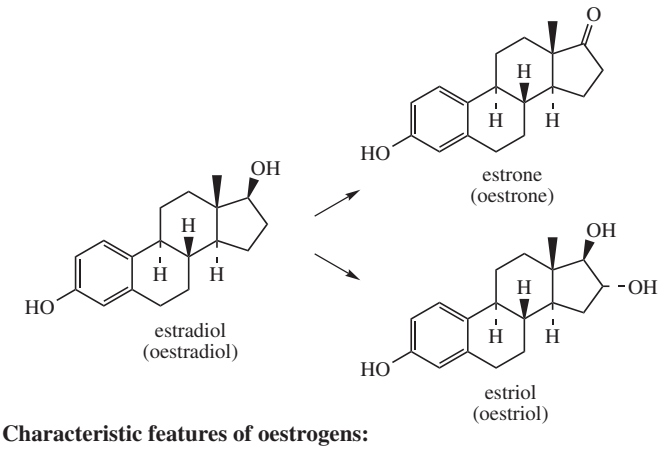
Figure 5.130

Mifepristone (Figure 5.130) is a progestogen antagonist used orally as an abortifacient to terminate pregnancy. This drug has a higher affinity for the progesterone receptor than does the natural hormone and prevents normal responses. This leads to loss of integrity of the uterine endometrial lining and detachment of the implanted fertilized egg.

Oestrogens

The **oestrogens** (US spelling: **estrogens**) are female sex hormones produced in the ovaries, and also in the placenta during pregnancy. They are responsible for the female sex characteristics and, together with progesterone, control the menstrual cycle [Box 5.22]. Oestrogens were first isolated from the urine of pregnant women, in which levels increase some 50-fold during the pregnancy. In horses, levels rise by as much as 500 times during pregnancy. Oestrogens occur both in free form and as glucuronides and sulfates at position 3; they are not restricted to females, since small amounts are produced in the male testis. The principal and most potent example is **estradiol** (also **oestradiol**, but US spelling has been generally adopted), though only low levels are found in urine, and larger amounts of the less-active metabolites **estrone** (**oestrone**) and the 16α-hydroxylated derivative **estriol** (**oestriol**) are present (Figure 5.131). Estrone has also been found in significant quantities in some plant seeds, e.g. pomegranate and date palm. These compounds have an aromatic A ring, a consequence of which is that C-19, the methyl on C-10, is absent. There is no carbon side-chain at C-17, and the basic C_{18} skeleton is termed estrane.

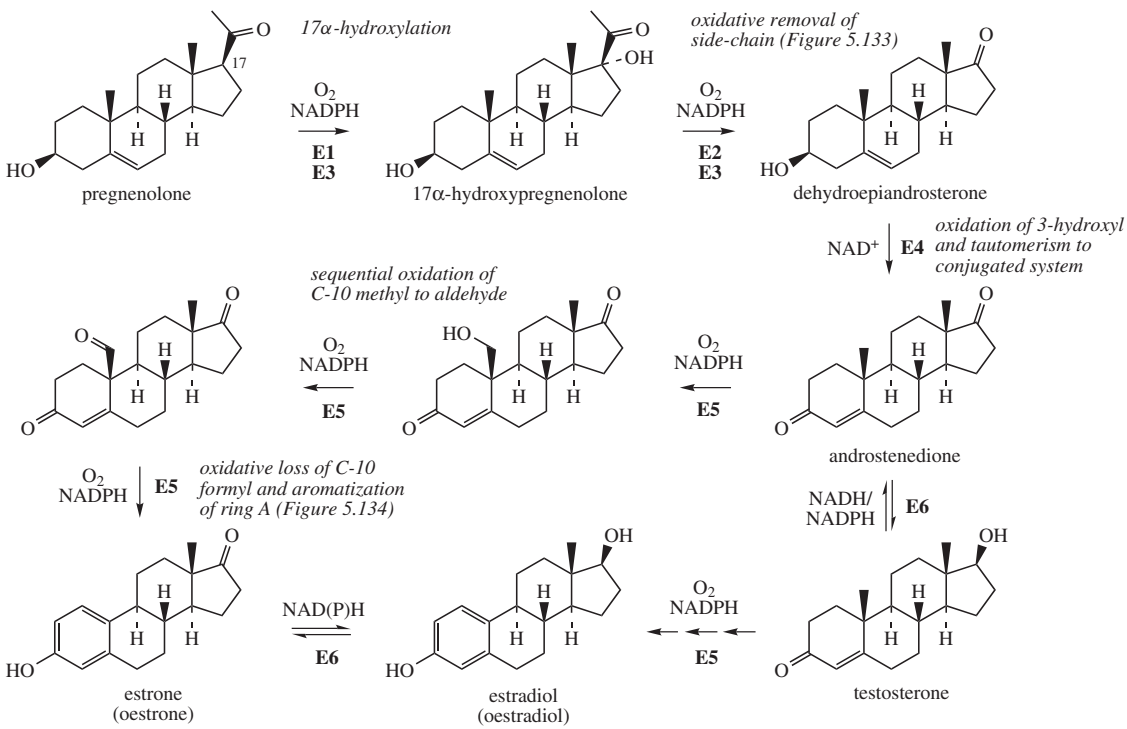
The biosynthetic pathway to estradiol and estrone (Figure 5.132) proceeds from cholesterol via



Characteristic features of oestrogens:

- C_{18} estrane skeleton
- aromatic A ring (consequently no methyl at C-10)
- no side-chain

Figure 5.131



E1: steroid 17α-hydroxylase
E2: steroid C-17/C-20 lyase
E3: steroid 17α-hydroxylase-17,20-lyase (CPY17) (bifunctional)
E4: Δ^5-3β-hydroxysteroid dehydrogenase /3-oxosteroid Δ^5-Δ^4-isomerase
E5: cytochrome P-450 aromatase (P450arom; CPY19)
E6: 17β-hydroxysteroid dehydrogenase

Figure 5.132

pregnenolone and bears a resemblance to the hydrocortisone pathway (Figure 5.112) in the early 17-hydroxylation step. Indeed, in humans, the same cytochrome P-450-dependent enzyme catalyses 17-hydroxylation of progesterone (leading to corticosteroids) and of pregnenolone (leading to oestrogens and androgens). Further, in the presence of cytochrome b_5, it catalyses the next step in oestrogen biosynthesis, the C-17–C-20 cleavage. Overall, this enzyme plays a significant role in controlling the direction of steroid synthesis. For hydrocortisone biosynthesis, 17α-hydroxyprogesterone is transformed by 21-hydroxylation, whereas 17α-hydroxypregnenolone is oxidized in the α-hydroxyketone function in oestrogen biosynthesis, cleaving off the two-carbon side-chain as acetic acid. The product is the 17-ketone **dehydroepiandrosterone**, which is the most abundant steroid in the blood of young adult humans, with levels peaking at about 20 years of age, then declining as the person ages. Apart from its role as a precursor of hormones, it presumably has other physiological functions, though these still remain to be clarified (see page 297). A mechanism for the side-chain cleavage reaction, initiated by attack of an enzyme-linked peroxide, is shown in Figure 5.133 and is analogous to that proposed for loss of the 14-methyl group during cholesterol biosynthesis (see page 249). Oxidation and tautomerism in rings A/B then give **androstenedione** (Figure 5.133).

Both androstenedione and its reduction product the male sex hormone **testosterone** are substrates for aromatization in ring A, with loss of C-19, leading to **estrone** and **estradiol** respectively (Figure 5.132). This sequence is also catalysed by a single cytochrome P-450-dependent enzyme, called **aromatase**, and the reaction proceeds via sequential oxidation of the methyl, with its final elimination as formic acid (Figure 5.134). The mechanism

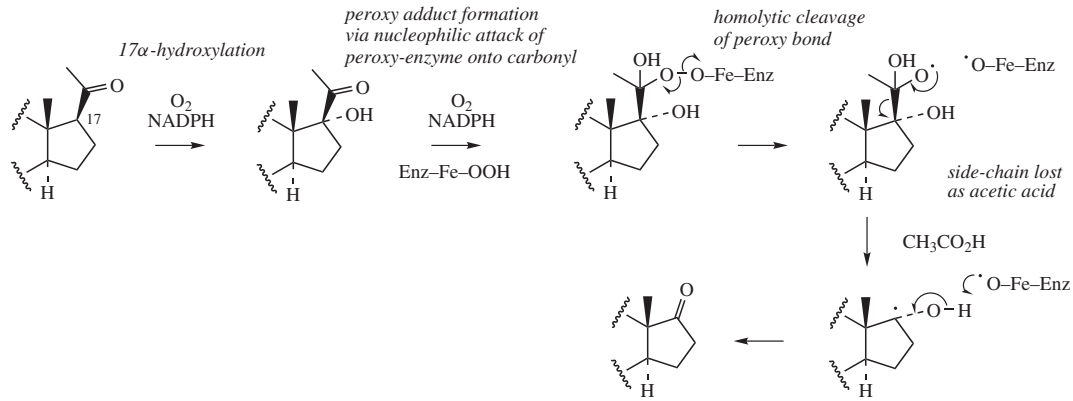


Figure 5.133

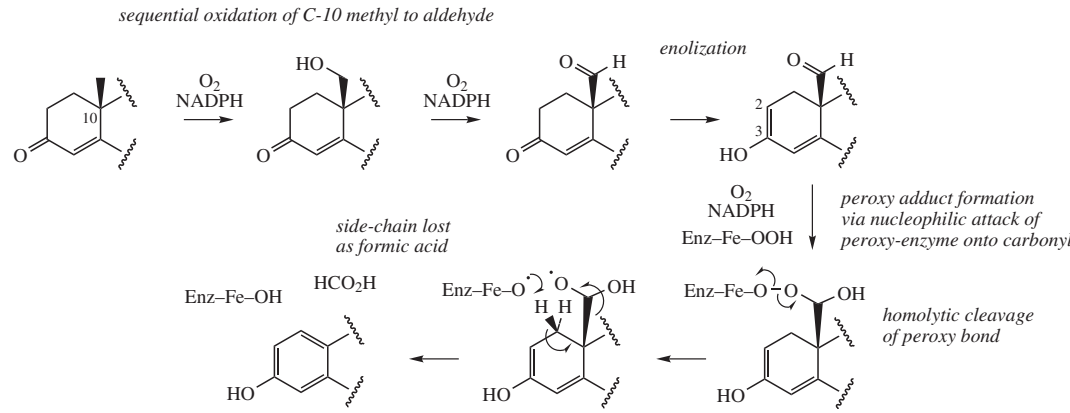


Figure 5.134

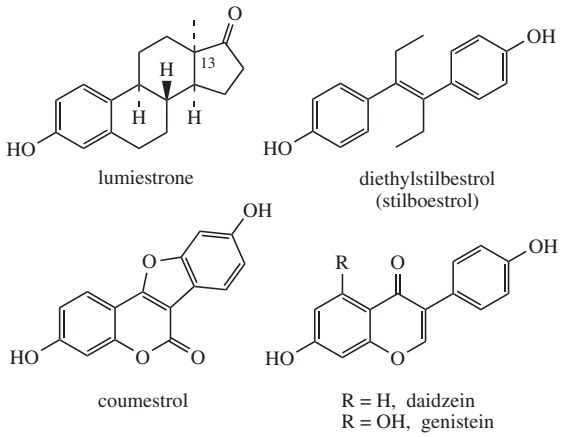


Figure 5.135

suggested is analogous to that of the side-chain cleavage reaction. The 2,3-enolization is also enzyme-catalysed and is a prerequisite for aromatization. As with other steroid hormones, the exact order of some of the steps, including formation of the Δ^4-3-keto function, 17-hydroxylation, reduction of the 17-keto, and aromatization in ring A, can vary according to the organism or the site of synthesis in the body. Since many breast tumours require oestrogens for growth, the design of **aromatase inhibitors** has become an important target for anticancer drug research [Box 5.22].

The aromatic ring makes the oestrogen molecule almost planar (see page 247) and is essential for activity. Changes which remove the aromaticity, e.g. partial reduction, or alter stereochemistry, give analogues with reduced or no activity. Thus, exposure of estrone to UV light leads to inversion of configuration at C-13 adjacent to the carbonyl function and, consequently, to formation of a *cis*-fused C/D ring system. The product, **lumiestrone** (Figure 5.135), is no longer biologically active. Some planar non-steroidal structures can also demonstrate oestrogenic activity as a result of a similar shape and relative spacing of oxygen functions. Thus, the synthetic **diethylstilbestrol** (**stilboestrol**; Figure 5.136) has been widely used as an oestrogen drug, and **coumestrol, daidzein**, and **genistein** (Figure 5.136) are naturally occurring isoflavonoids with oestrogenic properties from plants such as alfalfa, clovers, and soya beans and are termed phyto-oestrogens (see page 177). Dietary natural isoflavonoids are believed to give some protection against breast cancers and are also recommended to alleviate the symptoms of menopause.

Box 5.22

Oestrogen Drugs

Oestrogens suppress ovulation, and with progestogens they form the basis of combined oral contraceptives (see page 289) and hormone replacement therapy (HRT). They are also used to supplement natural oestrogen levels where these are insufficient, as in some menstrual disorders, and to suppress androgen formation and, thus, tumour growth of cancers dependent on androgens, e.g. prostate cancers. Oestrogens appear to offer a number of beneficial effects to women, including protection against osteoporosis, heart attacks, and possibly Alzheimer's disease. However, some cancers, e.g. breast and uterine cancers, are dependent on a supply of oestrogen for growth, especially during the early stages, so high oestrogen levels are detrimental.

Steroidal oestrogens for drug use were originally obtained by processing pregnancy urines, but the dramatic increase in demand resulting from the introduction of oral contraceptives required development of semi-synthetic procedures. Androstenolone formed via the Marker degradation of diosgenin plus side-chain removal (Figure 5.128) may be transformed to a dione by catalytic reduction of the Δ^5 double bond and oxidation of the 3-hydroxyl (Figure 5.136). This then allows production of androstadienedione by dibromination and base-catalysed elimination of HBr. Alternatively, it is now possible to achieve the synthesis of androstadienedione in a single step by a microbiological fermentation of either sitosterol obtained from soya beans (see page 255), or of cholesterol obtained in large quantities from the woolfat of sheep or from the spinal cord of cattle (see page 251). These materials lack unsaturation in the side-chain and were not amenable to simple chemical oxidation processes, e.g. as with stigmasterol (see page 282). Their exploitation required the development of suitable biotransformations, and this objective has now achieved through fermentation with *Mycobacterium phlei* (Figure 5.136). The aromatization step to estrone can be carried out in low yields by vapour-phase free-radical-initiated thermolysis, or more recently with considerably better yields using a dissolving-metal reductive thermolysis. In both processes, the methyl at C-10 is lost. This sequence gives estrone, from which **estradiol** (**oestradiol**) may be obtained by reduction of the 17-carbonyl. However, by far the most commonly used medicinal oestrogen is **ethinylestradiol** (**ethinyloestradiol**; Figure 5.136), which is 12 times as effective as estradiol when administered orally. This analogue can be synthesized from estrone by treatment with sodium acetylide in liquid ammonia, which attacks from the less-hindered α-face (compare page 282). The ethynyl substituent prevents oxidation at C-17; metabolism of estradiol leads to

Box 5.22 (continued)

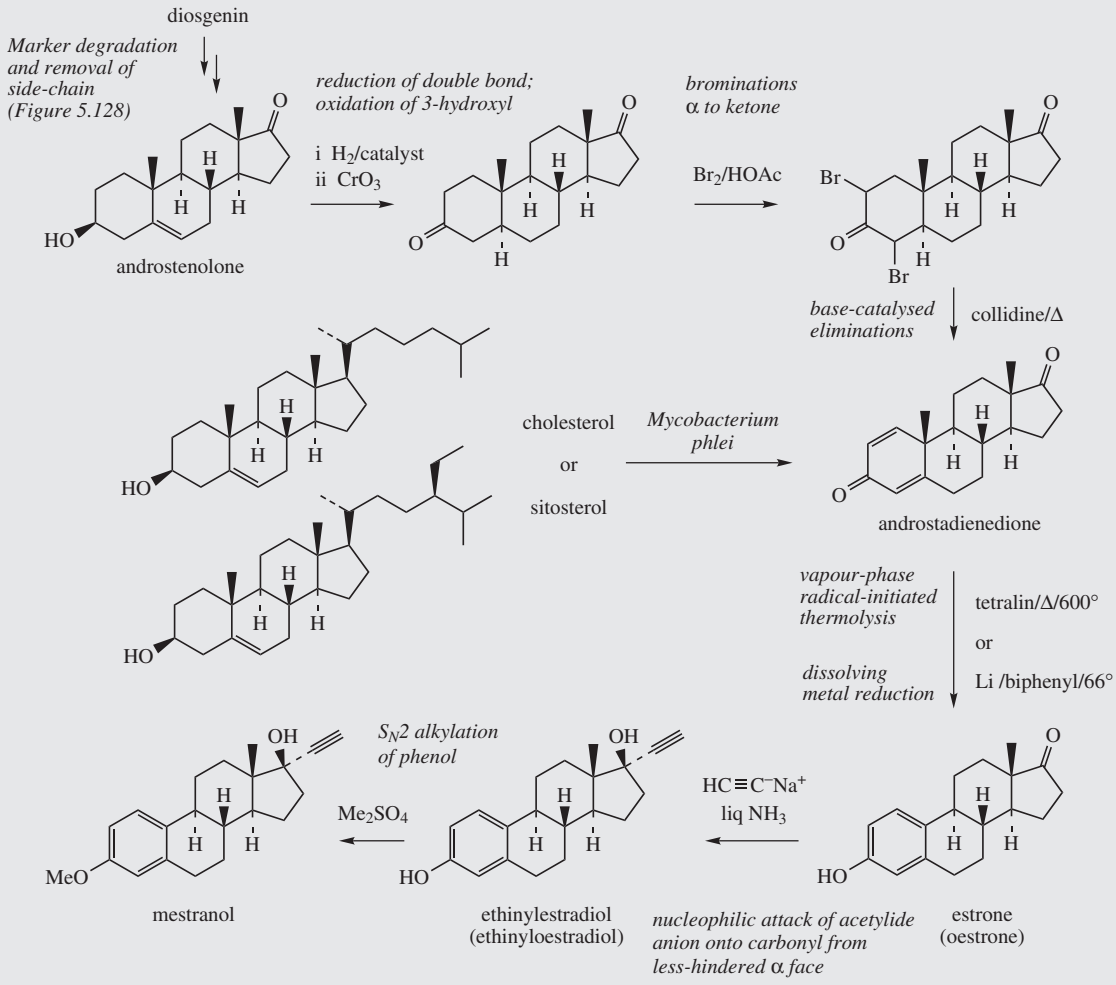


Figure 5.136

the less-active estrone. To retain oestrogenic activity, structural modifications appear effectively limited to the addition of the 17α-ethynyl group and to substitution on the 3-hydroxyl. The phenol group allows synthesis of other derivatives, for example, the 3-methyl ether **mestranol** (Figure 5.136); this acts as a pro-drug, being oxidized in the liver to ethinylestradiol. The ester **estradiol valerate** (**oestradiol valerate**) facilitates prolonged action through slower absorption and metabolism.

The lower activity metabolites **estriol** (**oestriol**; about 2% activity of estradiol) and **estrone** (**oestrone**; about 33% activity) (Figure 5.131) are sometimes used in **hormone replacement therapy (HRT)**. Oestrogen and progesterone levels decline naturally at menopause when the menstrual cycle ceases. The sudden reduction in oestrogen levels can lead to a number of unpleasant symptoms, including tiredness, hot flushes, vaginal dryness, and mood changes. HRT reduces these symptoms and delays other

Box 5.22 (continued)

long-term consequences of reduced oestrogen levels, including osteoporosis and atherosclerosis. HRT currently provides the best therapy for preventing osteoporosis, a common disease in post-menopausal women. Osteoporosis is characterized by a generalized loss of bone mass, leading to increased risk of fracture, and a sharp reduction in endogenous oestrogen levels is recognized as a critical factor. However, it has been discovered that HRT also increases the risk of developing some types of cancer, especially breast, ovarian, and endometrial cancers. Oestrogen and progestogen combinations are used in HRT unless the woman has had a hysterectomy, in which case oestrogen alone is prescribed. Natural oestrogen structures are preferred to the synthetic structures, such as ethinylestradiol or mestranol. Before the availability of plant-derived semi-synthetic oestrogens, extraction of urine from pregnant women and pregnant horses allowed production of oestrogen mixtures for drug use. **Conjugated equine oestrogens** are still widely prescribed for HRT and are obtained by extraction from the urine of pregnant mares and subsequent purification, predominantly in Canada and the USA. Animal welfare groups voice concern over the conditions the animals endure and urge women to reject these drug preparations in favour of plant-derived alternatives. Conjugated equine oestrogens provide a profile of natural oestrogens based principally on estrone and equilin (Figure 5.137). They consist of a mixture of oestrogens in the form of sodium salts of their sulfate esters, comprising mainly estrone (50–60%) and equilin (20–30%), with smaller amounts of 17α- and 17β-dihydroequilin, and 17α-estradiol. The semi-synthetic **estropipate** (Figure 5.137) is also a conjugated oestrogen, the piperazine salt of estrone sulfate.

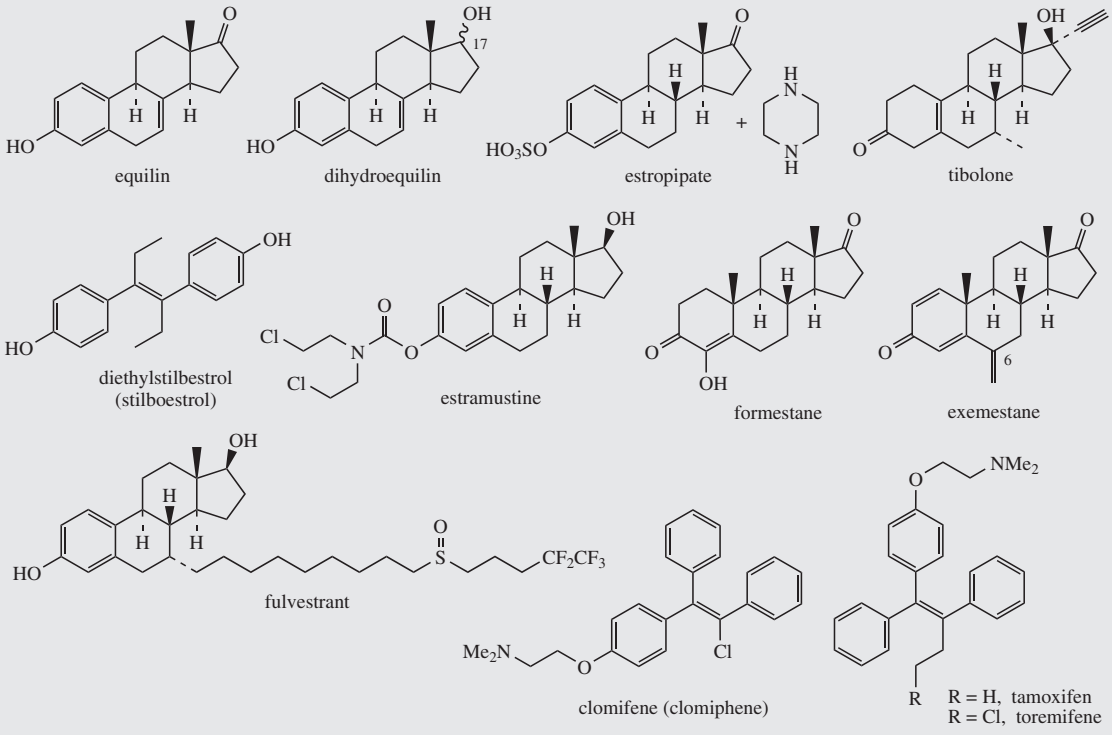


Figure 5.137

Box 5.22 (continued)

Phyto-oestrogens are predominantly isoflavonoid derivatives found in food plants and are used as dietary supplements to provide similar benefits to HRT, especially in countering some of the side-effects of the menopause in women. These compounds are discussed under isoflavonoids (see page 177). **Dioscorea** (**wild yam**) root or extract (see page 262) is also marketed to treat the symptoms of menopause as an alternative to HRT. Although there is a belief that this increases levels of progesterone, which is then used as a biosynthetic precursor of other hormones, there is little definitive evidence that diosgenin is metabolized in the human body to progesterone.

The structure of **tibolone** (Figure 5.137) appears to resemble that of a progestogen more than it does an oestrogen. Although it does not contain an aromatic A ring, the 5(10)-double bond ensures a degree of planarity. This agent combines both oestrogenic and progestogenic activity, and also has weak androgenic activity; it has been introduced for short-term treatment of symptoms of oestrogen deficiency.

Diethylstilbestrol (**stilboestrol**; Figure 5.137) is the principal non-steroidal oestrogen drug, but now finds only occasional use, since there are safer and better agents available. It may sometimes be used in treating breast cancer in postmenopausal women.

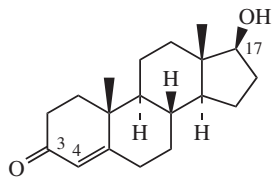
In **estramustine** (Figure 5.137), estradiol is combined with a cytotoxic alkylating agent of the nitrogen mustard class via a carbamate linkage. This drug has a dual function, a hormonal effect by suppressing androgen (testosterone) formation and an antimitotic effect from the mustine residue. It is used for treating prostate cancers.

Aromatase Inhibitors

Formestane (Figure 5.137), the 4-hydroxy derivative of androstenedione, was the first steroid aromatase inhibitor to be used clinically. However, oral availability was poor and it has been superseded by **exemestane**; the prominent structural modification is the exocyclic alkene at position 6. As analogues of androstenedione, these agents bind to and inhibit aromatase, reducing synthesis of oestrogens. They are of value in treating advanced breast cancer in post-menopausal patients.

Oestrogen Receptor Antagonists (Anti-oestrogens)

Breast cancer is dependent on a supply of oestrogen, and a major success in treating this disease has been the introduction of **tamoxifen** (Figure 5.137). This drug contains the stilbene skeleton seen in diethylstilbestrol, but acts as an oestrogen-receptor antagonist rather than as an agonist in breast tissue, and thus deprives the cells of oestrogen. However, it is an agonist in bone and uterine tissue. The chlorinated analogue **toremifene** is also available, but is used primarily in post-menopausal women. **Fulvestrant** (Figure 5.137) is a newer agent in this group; it is estradiol-based, and can be effective against tamoxifen-resistant breast cancer. Oestrogen antagonists can also be used as fertility drugs, occupying oestrogen receptors and interfering with feedback mechanisms, thus inducing ova release. **Clomifene** (clomiphene; Figure 5.137), and to a lesser extent tamoxifen, are used in this way, but can lead to multiple pregnancies.



testosterone

Characteristic features of androgens:

- C_{19} androstane skeleton
- no side-chain
- Δ^4-3-keto
- 17β-hydroxyl

Figure 5.138

Androgens

The primary male sex hormone, or **androgen**, is **testosterone** (Figure 5.138). This is secreted by the testes and is responsible for development and maintenance of the male sex characteristics. Androgens also have a secondary physiological effect, an anabolic activity which stimulates growth of bone and muscle and promotes storage of protein [Box 5.23]. The biosynthetic pathway to testosterone is included in Figure 5.132, where it can feature as an intermediate in the pathway to oestrogens. Low levels of testosterone are also synthesized in females in the ovary. Testosterone lacks any side-chain and has a 17β-hydroxyl as in estradiol, but still contains the methyl C-19 and the Δ^4-3-one system in ring A. This C_{19} skeleton is designated androstane.

Box 5.23

Androgen Drugs

Testosterone can be produced from androstenolone (Figure 5.128) by chemical routes, requiring reduction of the 17-carbonyl and oxidation of the 3-hydroxyl, with the use of appropriate protecting groups. A simple high-yielding process (Figure 5.139) exploits microbiological conversion with yeast, in which fermentation first under aerobic conditions oxidizes the 3-hydroxyl and then in the absence of air reduces the 17-keto group.

Testosterone is not active orally, since it is easily metabolized in the liver; it has to be implanted or injected in the form of esters. Transdermal administration from impregnated patches has also proved successful and is now the method of choice for treating male sexual impotence caused by low levels of sex hormones (hypogonadism). Testosterone may also be prescribed for menopausal women as an adjunct to HRT (see page 294) to improve sex drive, and occasionally in the treatment of oestrogen-dependent breast cancer. The ester testosterone undecanoate is orally active, as is **mesterolone** (Figure 5.140), which features introduction of a 1α-methyl group and reduction of the Δ^4 double bond.

The ratio of androgenic to anabolic activity can vary in different molecules. Considerable effort has been put into producing steroids with low androgenic activity but high anabolic activity to use for various metabolic and endocrine disorders. However, it is difficult to remove the androgenic activity completely from anabolic steroids. **Nandrolone** (19-nortestosterone; Figure 5.140) is probably the only example currently in clinical use. Abuse of these materials by athletes wishing to promote muscle development and strength is considerably more frequent. Androgenic activity can affect the sexual characteristics of women, making them more masculine, whilst prolonged use of these drugs can lower fertility in either sex and endanger long-term health by increasing the risks of heart and liver disease or cancer.

The progestogen **cyproterone acetate** (Figure 5.140) is a competitive androgen antagonist or anti-androgen that reduces male libido and fertility, and finds use in the treatment of severe hypersexuality and sexual deviation in the male, as well as in prostate cancer. **Finasteride** and the structurally similar **dutasteride** (Figure 5.140) are also anti-androgens that have value in prostate conditions. These are 4-aza-steroids and specific inhibitors of the 5α-reductase involved in testosterone metabolism. This enzyme reduces the 4,5-double bond and converts testosterone into dihydrotestosterone, which is actually a more potent androgen. High levels of dihydrotestosterone are implicated in prostate cancer and benign prostatic hyperplasia; inhibition of 5α-reductase helps to reduce prostate tissue growth. Finasteride has also been noted to prevent hair loss in men, and is marketed to treat male-pattern baldness. Continuous use for 3–6 months is necessary; unfortunately, the effects are reversed 6–12 months after the treatment is discontinued. **Abiraterone** (Figure 5.140) is an inhibitor of the steroid 17α-hydroxylase-17,20-lyase (CYP17) and thus blocks formation of testosterone (and also oestrogens). This agent is proving very successful in clinical trials for prostate cancer treatment. Structurally, this is a side-chain modification of pregnenolone, the normal substrate for the enzyme.

Dehydroepiandrosterone (DHEA; Figure 5.132) is a precursor of androgens and oestrogens; it is the most abundant steroid in the blood of young adult humans, levels peaking at about 20 years of age and then declining as the person ages. Whilst this hormone has a number of demonstrated biological activities, its precise physiological functions remain to be clarified. This material has become popular in the hope that it will maintain youthful vigour and health, and counter the normal symptoms of ageing. These claims are as yet unsubstantiated, but taking large amounts of this androgen and oestrogen precursor can lead to side-effects associated with high levels of these hormones, e.g. increased risk of prostate cancer in men or of breast cancer in women, who may also develop acne and facial hair. DHEA is not a precursor of glucocorticoids, mineralocorticoids, or of progestogens.

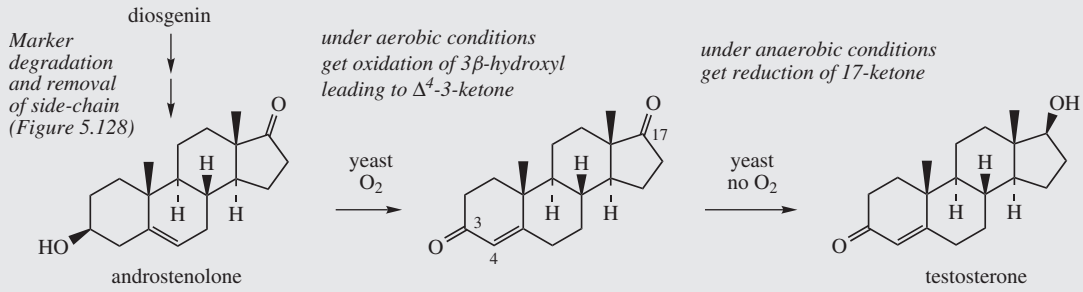


Figure 5.139

Box 5.23 (continued)

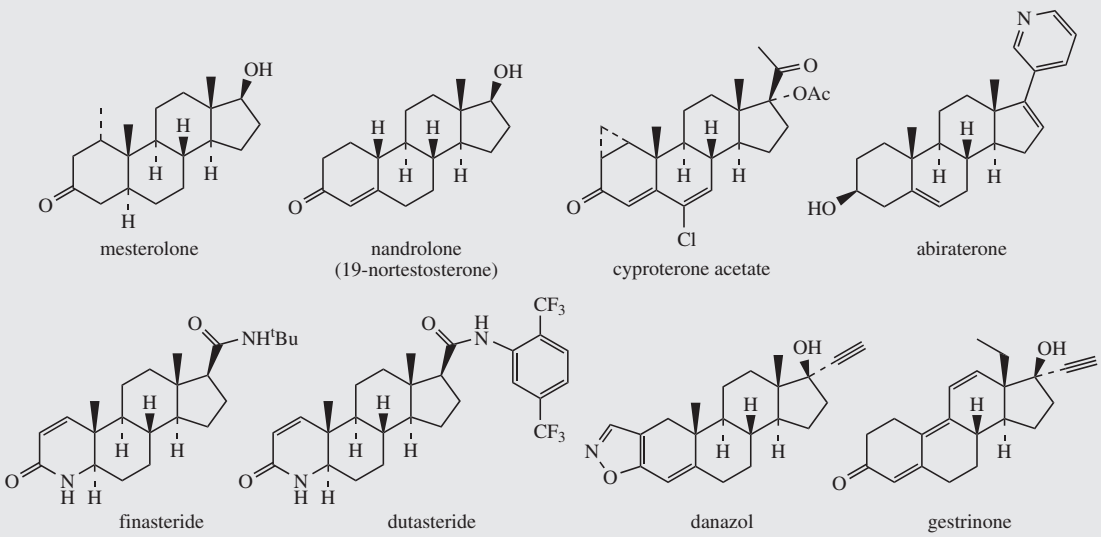


Figure 5.140

Danazol and **gestrinone** (Figure 5.140) are inhibitors of pituitary gonadotrophin release, combining weak androgenic activity with anti-oestrogenic and antiprogestogenic activity. These highly modified structures bear one or more of the features we have already noted in discussions of androgens, progestogens, and oestrogens, possibly accounting for their complex activity. These compounds are used particularly to treat endometriosis, where endometrial tissue grows outside the uterus.

It is particularly worthy of note that the routes to corticosteroids, progestogens, oestrogens, and androgens involve common precursors or partial pathways, and some enzymes display multiple functions or perhaps broad specificity. The main relationships are summarized in Figure 5.141. Consequently, these processes need to be under very tight control for a person's normal physiological functions and characteristics to be maintained. Human 17β-hydroxysteroid dehydrogenases, which control androgen and oestrogen potency, provide a suitable example. These enzymes interconvert weak (17-ketosteroid) and potent (17β-hydroxysteroid) hormones, and three isoforms have been characterized. Two of these isoforms function with NADPH; one predominantly converts estrone into estradiol, the other androstenedione into testosterone. The third isoform uses NAD^+ as cofactor and mainly functions in the oxidative direction, converting both estradiol and testosterone into the less-active ketones. According to their activities and tissue localization, these enzyme isoforms help to achieve the required functional balance of the various hormones.

TETRATERPENES (C_{40})

The tetraterpenes are represented by only one group of compounds, the **carotenoids**, though several hundred natural structural variants are known. These compounds play a role in photosynthesis, but they are also found in non-photosynthetic plant tissues, in fungi, and in bacteria. Formation of the tetraterpene skeleton, e.g. **phytoene**, involves tail-to-tail coupling of two molecules of **geranylgeranyl diphosphate (GGPP)** in a sequence essentially analogous to that seen for squalene and triterpenes (Figure 5.142). A cyclopropyl compound, **prephytoene diphosphate** (compare presqualene diphosphate, page 235), is an intermediate in the sequence, and the main difference between the tetraterpene and triterpene pathways is how the resultant allylic cation is discharged. For squalene formation, the allylic cation accepts a hydride ion from NADPH, but for phytoene biosynthesis, a proton is lost, generating a double bond in the centre of the molecule, and thus a short conjugated chain is developed. In plants and fungi, this new double bond has the *Z* (*cis*) configuration, whilst in bacteria it is *E* (*trans*). This

Steroid hormone biosynthetic interrelationships

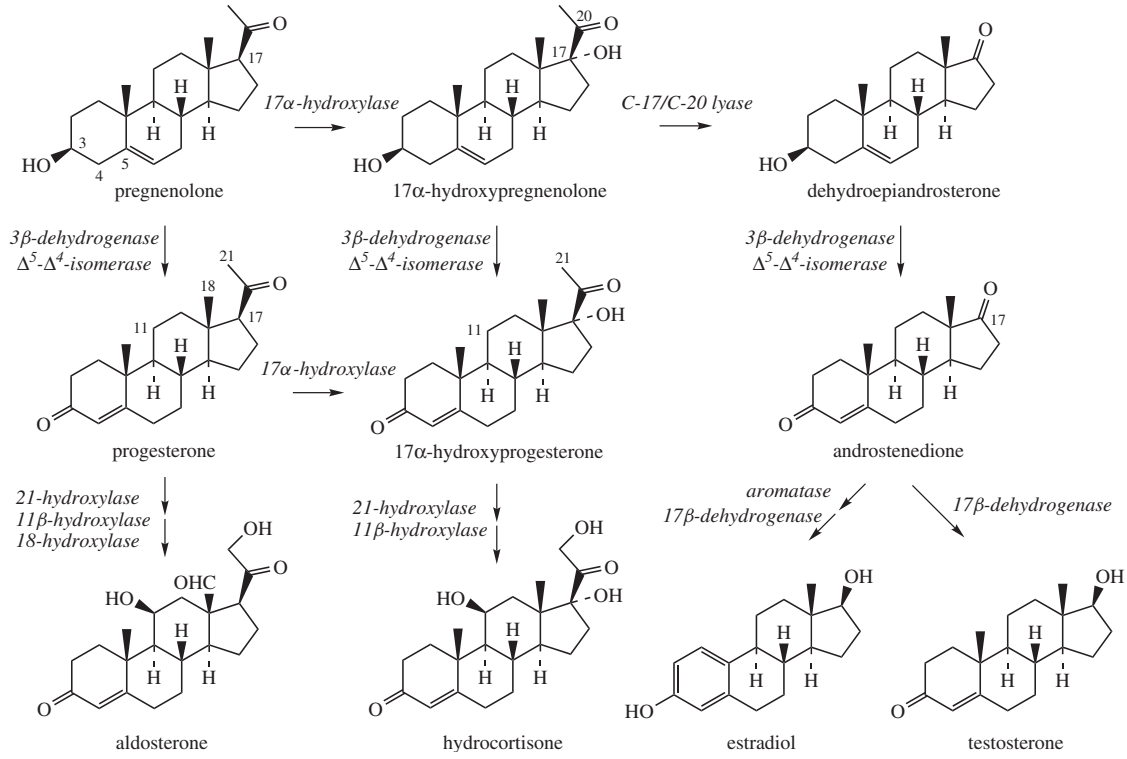


Figure 5.141

triene system prevents the types of cyclization seen with squalene. Conjugation is extended then by a sequence of desaturation reactions, removing pairs of hydrogen atoms alternately from each side of the triene system, giving eventually **lycopene** (Figure 5.142), which, in common with the majority of carotenoids, has the *all-trans* configuration. In bacteria and fungi, a single phytoene desaturase enzyme converts phytoene into lycopene. In plants, two desaturase enzymes and an isomerase are involved, the isomerase being responsible for changing the configurations of those double bonds introduced initially in *cis* form. The central double bond appears to be isomerized as part of the first dehydrogenation step.

The extended π-electron system confers colour to the carotenoids, and accordingly they contribute yellow, orange, and red pigmentations to plant tissues; phytoene is colourless. Lycopene is the characteristic carotenoid pigment in ripe tomato fruit (*Lycopersicon esculente*; Solanaceae). The orange colour of carrots (*Daucus carota*; Umbelliferae/Apiaceae) is caused by **β-carotene** (Figure 5.143), though this compound is widespread in higher plants. β-Carotene and other natural carotenoids (Figure 5.143) are widely employed as colouring agents for foods, drinks, confectionery, and drugs. β-Carotene displays additional cyclization of the chain ends, which can be rationalized by the carbocation mechanism shown in Figure 5.144. The new proton introduced is derived from water; depending on which proton is then lost from the cyclized cation, three different cyclic alkene systems can arise at the end of the chain, described as β-, γ-, or ε-ring systems. The γ-ring system is the least common. **α-Carotene** (Figure 5.143) has a β-ring at one end of the chain and an ε-type at the other, and is representative of carotenoids lacking symmetry. **γ-Carotene** (a precursor of β-carotene) and **δ-carotene** (a precursor of α-carotene) illustrate carotenoids where only one end of the chain has become cyclized. Oxygenated carotenoids (termed xanthophylls) are also widely distributed, and the biosynthetic origins of the oxygenated rings found in some of these, such as **zeaxanthin, lutein**, and **violaxanthin** (Figure 5.143), all common green-leaf carotenoids, are

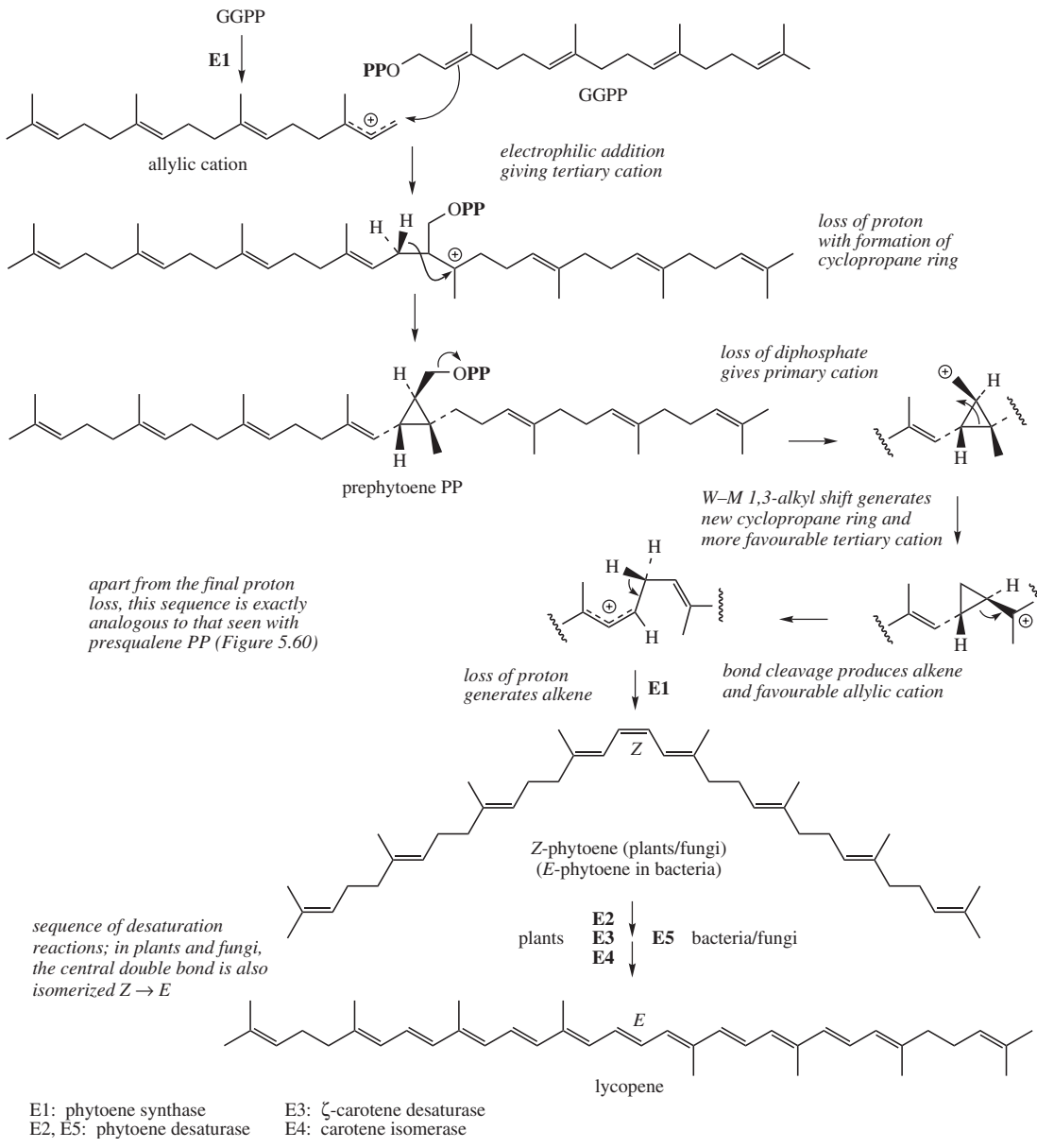


Figure 5.142

shown in Figure 5.144. The epoxide grouping in violaxanthin allows further chemical modifications, such as ring contraction to a cyclopentane, exemplified by **capsanthin** (Figure 5.143), the brilliant red pigment of sweet peppers (*Capsicum annuum*; Solanaceae), or formation of an allene as in **fucoxanthin**, an abundant carotenoid in brown algae (*Fucus* species; Fucaceae). **Astaxanthin** (Figure 5.143) is commonly found in marine animals and is responsible for the pink/red coloration of crustaceans, shellfish, and fish such as salmon. These animals are unable to synthesize carotenoids, and astaxanthin is produced by modification of plant carotenoids, e.g. β-carotene, obtained in the diet.

Carotenoids function along with chlorophylls in photosynthesis as accessory light-harvesting pigments, effectively extending the range of light absorbed by the

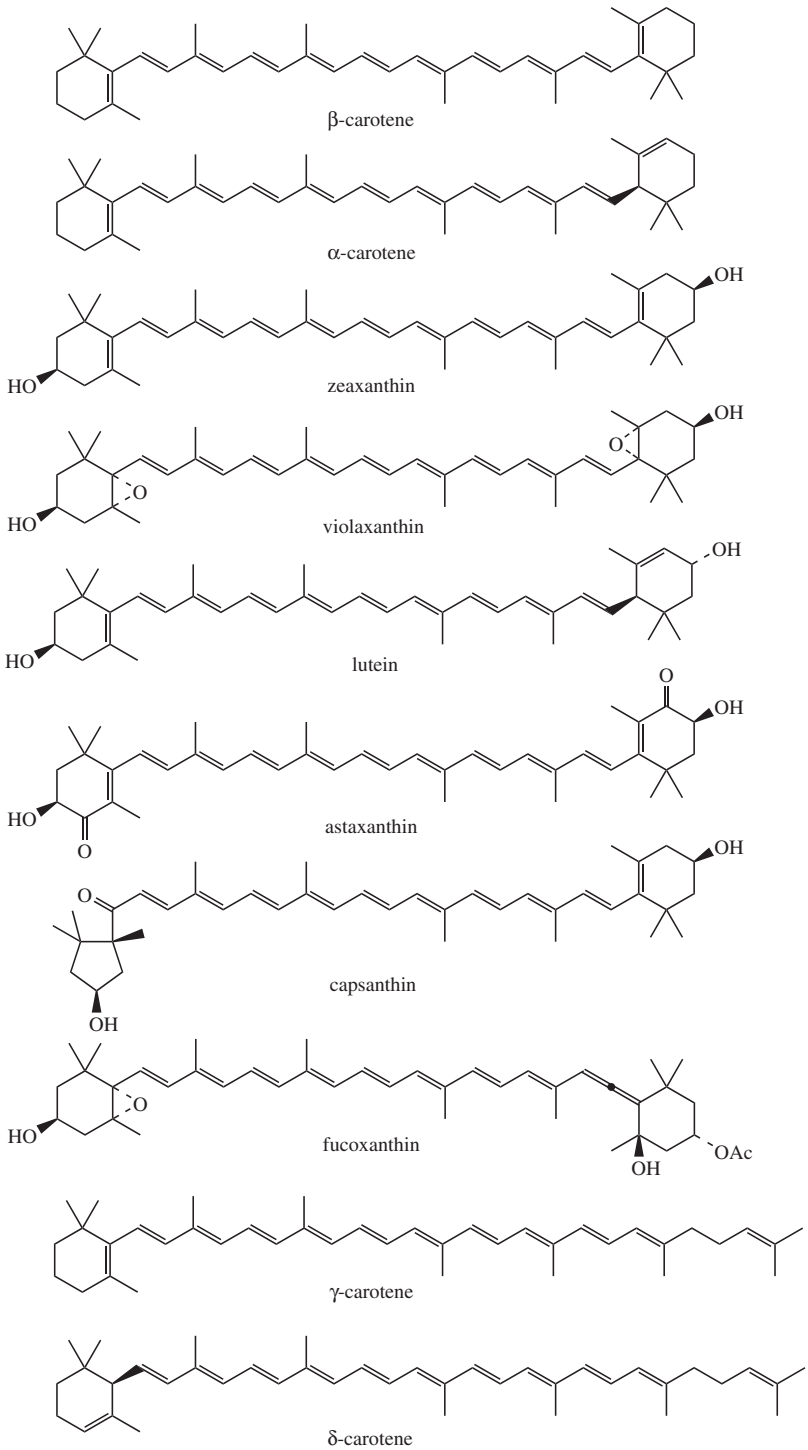

β-carotene
α-carotene
OH
HO
zeaxanthin
O
violaxanthin
lutein
astaxanthin
capsanthin
fucoxanthin
OAc
γ-carotene
δ-carotene

Figure 5.143

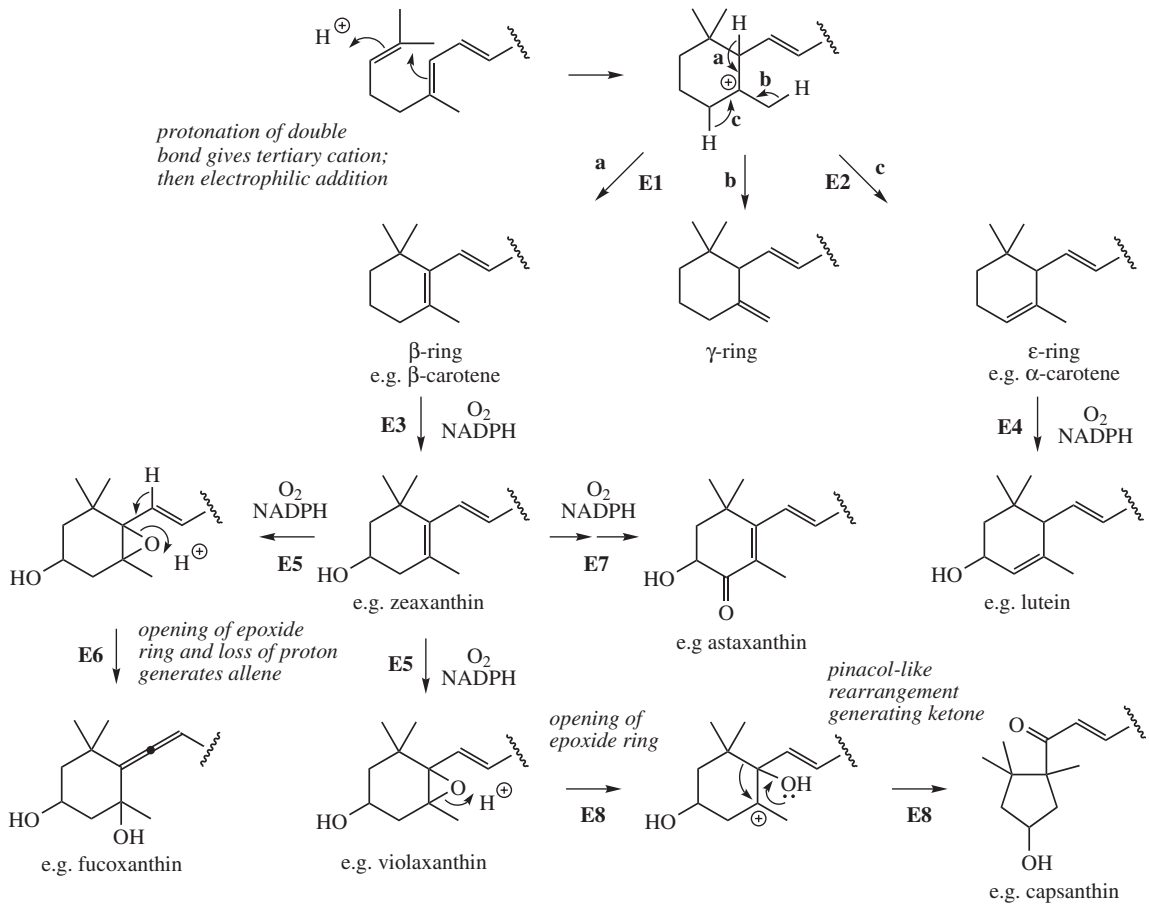


Figure 5.144

photosynthetic apparatus. They also serve as important protectants for plants and algae against photo-oxidative damage, quenching toxic oxygen species. Some herbicides (bleaching herbicides) act by inhibiting carotenoid biosynthesis, and the unprotected plant is subsequently killed by photo-oxidation. Recent research also suggests that carotenoids are important antioxidant molecules in humans, quenching singlet oxygen and scavenging peroxyl radicals, thus minimizing cell damage and affording protection against some forms of cancer. The most significant dietary carotenoid in this respect is **lycopene**, with tomatoes and processed tomato products featuring as the predominant source. The extended conjugated system allows free-radical addition reactions and hydrogen abstraction from positions allylic to this conjugation. Tomato fruits accumulate considerable amounts of lycopene, and thus provide an excellent vehicle for genetic engineering in the carotenoid field. It is possible to diminish carotenoid levels dramatically, or divert the lycopene into other carotenoids, e.g. β-carotene, by expressing new genes controlling the cyclization reactions.

The A group of vitamins are important metabolites of carotenoids [Box 5.24]. **Vitamin A_1** (**retinol**; Figure 5.145) effectively has a diterpene structure, but it is derived in mammals by oxidative metabolism of a tetraterpenoid, mainly β-carotene, taken in the diet.

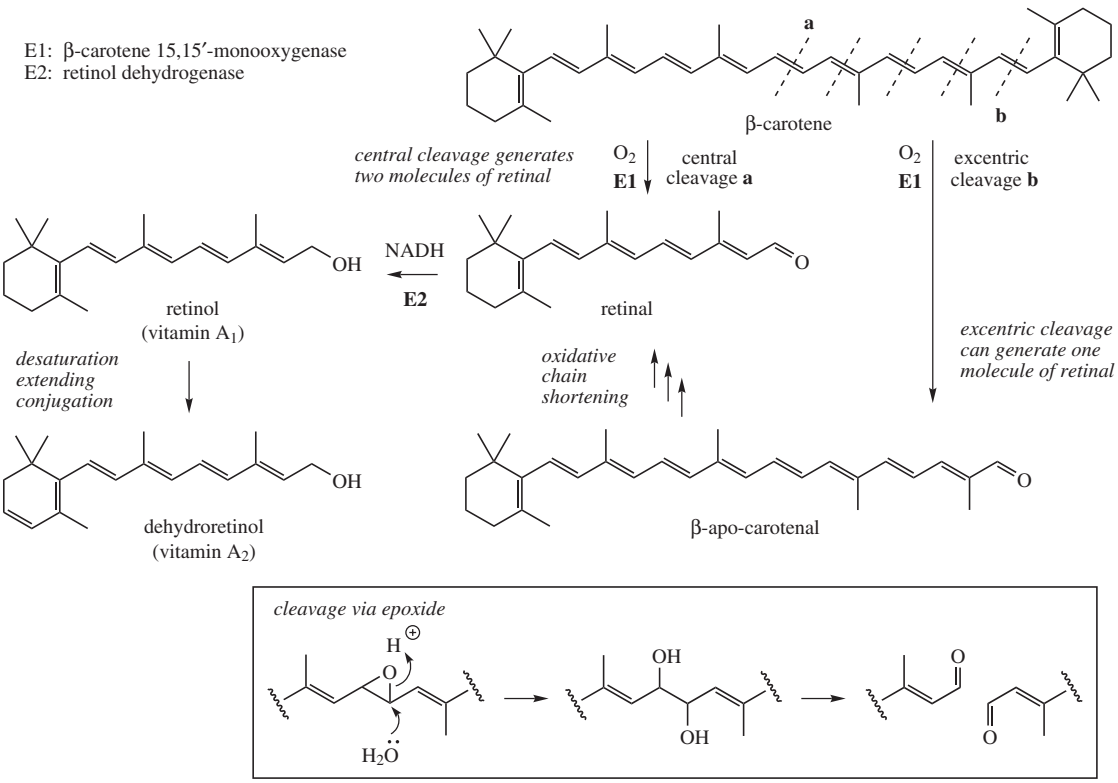
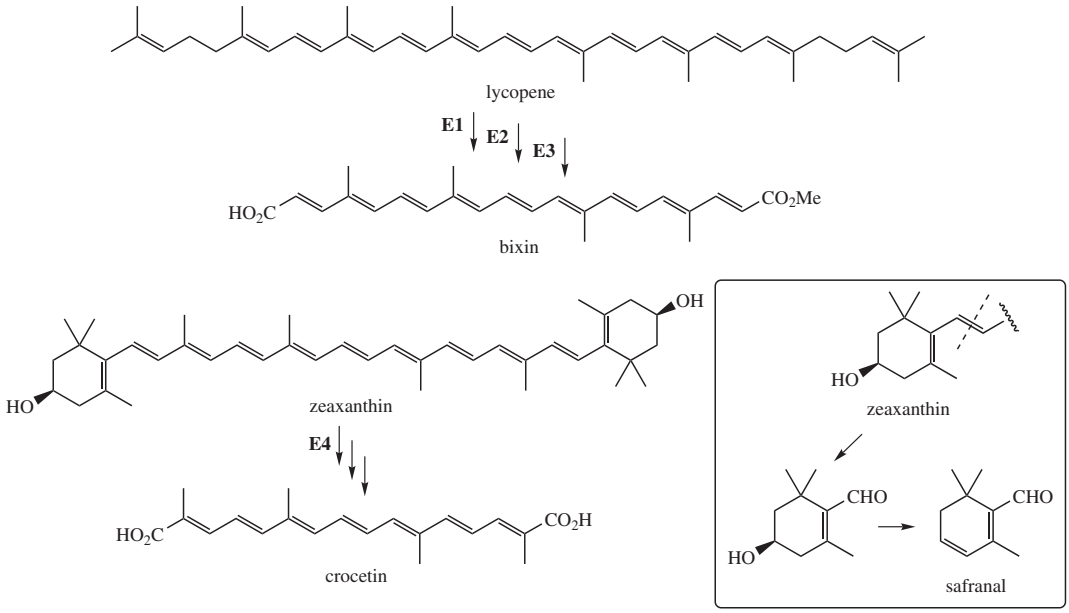


Figure 5.145

Cleavage occurs in the mucosal cells of the intestine and is catalysed by an O_2-dependent monooxygenase, most probably via an intermediate epoxide. This can theoretically yield two molecules of the intermediate aldehyde **retinal**, which is subsequently reduced to the alcohol retinol (Figure 5.145). Although β-carotene cleaved at the central double bond is capable of giving rise to two molecules of retinol, there is evidence that cleavage can also occur at other double bonds, so-called excentric cleavage (Figure 5.145). Further chain shortening from the larger cleavage product, an apocarotenal, then produces retinal, but only one molecule can be produced per molecule of β-carotene. **Vitamin A_2** (**dehydroretinol**; Figure 5.145) is an analogue of retinol containing a cyclohexadiene ring system; the corresponding aldehyde and retinal are also included in the A group of vitamins. Retinol and its derivatives are found only in animal products, and these provide some of our dietary needs. Cod liver oil and halibut liver oil are rich sources that are used as dietary supplements. However, carotenoid sources are equally important. These need to have at least one non-hydroxylated ring system of the β-type, e.g. β-carotene, α-carotene, and γ-carotene; about 50 natural carotenoids fall in this group.

Cleavage of carotenoid precursors also explains the formation of **bixin** and **crocetin** (Figure 5.146); indeed, these compounds are classified as apocarotenoids. Large amounts (up to 10%) of the red pigment bixin are found in the seed coats of annatto (*Bixa orellana*; Bixaceae), and bixin is widely used as a natural food colorant, especially for cheese and other dairy products. Bixin originates from cleavage of the non-cyclic carotenoid lycopene. Crocetin, in the form of esters with gentiobiose [D-Glc(β1→6)D-Glc], is the major pigment in stigmas of *Crocus sativus* (Iridaceae) which comprise the extremely expensive spice saffron. Crocetin is known to be produced from the cyclic carotenoid zeaxanthin; the monoterpenoid **safranal** (Figure 5.146) also contributes to the aroma and taste of saffron, and is formed from the remaining portion of the carotenoid structure.



E1: lycopene cleavage (di)oxygenase
E2: bixin aldehyde dehydrogenase
E3: norbixin carboxyl methyltransferase
E4: zeaxanthin 7,8(7′,8′)-cleavage (di)oxygenase

Figure 5.146

Box 5.24

Vitamin A

Vitamin A_1 (**retinol**) and **vitamin A_2** (dehydroretinol) (Figure 5.145) are fat-soluble vitamins found only in animal products, particularly eggs, dairy products, and animal livers and kidneys. Fish liver oils, e.g. cod liver oil, halibut liver oil (see Table 3.1) are particularly rich sources. They exist as the free alcohols, or as esters with acetic and palmitic acid. Vitamin A_2 has about 40% of the activity of vitamin A_1. Carotenoid precursors (provitamins) are widely distributed in plants, and after ingestion, these are subsequently transformed into vitamin A in the liver. Green vegetables and plant sources rich in carotenoids, such as carrots, help to provide adequate levels. A deficiency of vitamin A leads to vision defects, including impairment at low light levels (night blindness) and a drying and degenerative disease of the cornea. It is also necessary for normal growth of young animals. Retinoids (vitamin A and analogues) are now known to act as signalling molecules which regulate diverse aspects of cell differentiation, embryonic development, growth, and vision.

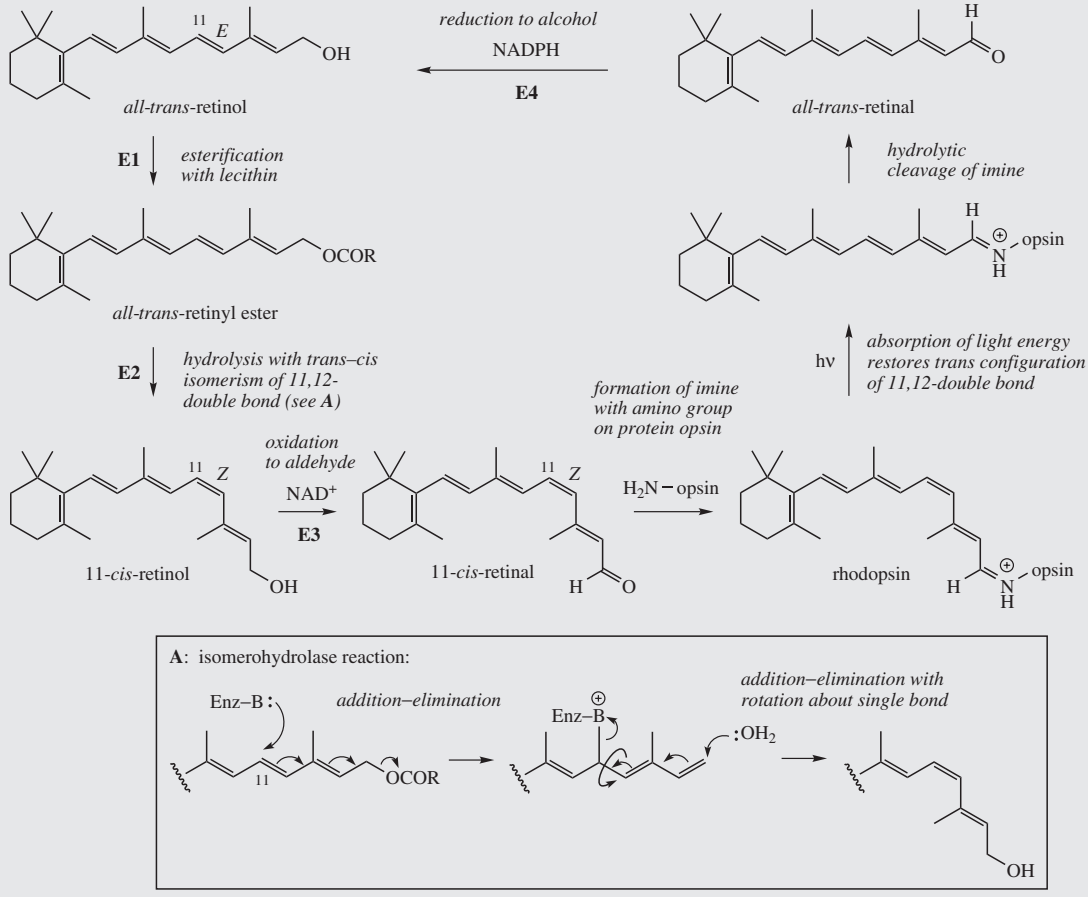
For the processes of vision, *all-trans*-retinol needs to be converted into 11-*cis*-retinol and then 11-*cis*-retinal. The isomerization is accomplished via esterification, which creates a better leaving group, and a likely S_N2' addition–elimination mechanism that allows rotation about a single bond is shown in Figure 5.147. 11-*cis*-Retinal is then bound to the protein opsin in the retina via an imine linkage to give the red visual pigment rhodopsin; its sensitivity to light involves isomerization of the *cis*-retinal portion back to the *all-trans* form, thus translating the light energy into a molecular change which triggers a nerve impulse to the brain. The absorption of light energy promotes an electron from a π- to a π*-orbital, thus temporarily destroying the double-bond character and allowing rotation. A similar *cis–trans* isomerization affecting cinnamic acids was discussed under coumarins (see page 161). *all-trans*-Retinal is then subsequently released from the protein by hydrolysis, and the process can continue.

Vitamin A is relatively unstable, and sensitive to oxidation and light. Antioxidant stabilizers such as vitamin E and vitamin C are sometimes added. It is more stable in oils such as the fish liver oils, which are thus good vehicles for administering the vitamin. Synthetic material is also used. Excessive intake of vitamin A can lead to toxic effects, including pathological changes in the skin, hair loss, blurred vision, and headaches.

all-trans-Retinoic acid, the biologically most active metabolite of vitamin A, has been found to play a major role in the regulation of gene expression, in cellular differentiation, and in the proliferation of epithelial cells. Synthetic retinoic acid (**tretinoin**) and **isotretinoin** (13-*cis*-retinoic acid) (Figure 5.148) are used as topical or oral treatments for acne vulgaris, reducing levels of dehydroretinol and modifying skin keratinization. Dehydroretinol levels in the skin become markedly elevated in

Box 5.24 (continued)

conditions such as eczema and psoriasis. **Acitretin** (Figure 5.148) is an aromatic analogue which can give relief in severe cases of psoriasis. All these materials can produce toxic side-effects, including increased sensitivity to UV light. Tretinoin has also proven useful in cancer chemotherapy, particularly in acute promyelocytic leukaemia.



E1: lecithin retinol acyl transferase
E2: isomerohydrolase
E3: 11-*cis*-retinol dehydrogenase
E4: *all-trans*-retinol dehydrogenase

Figure 5.147

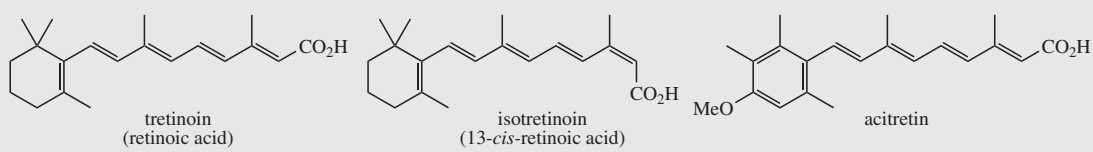


Figure 5.148

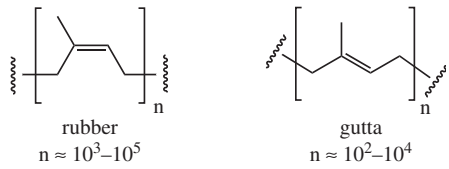


Figure 5.149

HIGHER TERPENOIDS

Terpenoid fragments containing several isoprene units are found as alkyl substituents in shikimate-derived quinones (see page 178). Thus, ubiquinones typically have C_{40}–C_{50} side-chains, plastoquinones usually C_{45}, and menaquinones up to C_{65}. The alkylating agents are polyprenyl diphosphates, formed simply by increasing the chain length with further addition of IPP residues. Even longer polyisoprene chains are encountered in some natural polymers, especially rubber and gutta percha. **Rubber** (Figure 5.149), from the rubber tree *Hevea brasiliensis* (Euphorbiaceae), is unusual in possessing an extended array of *cis* (*Z*) double bonds rather than the normal *trans* configuration. **Gutta percha**, from *Palaquium gutta* (Sapotaceae), on the other hand, has *trans* (*E*) double bonds. The *cis* double bonds in rubber are known to arise by loss of the *pro-S* proton (H_S) from C-2 of IPP (contrast loss of H_R which gives a *trans* double bond) (Figure 5.150). However, a small (up to C_{20}) *trans*-allylic diphosphate initiator is actually used for the beginning of the chain before the extended *cis* chain is elaborated.

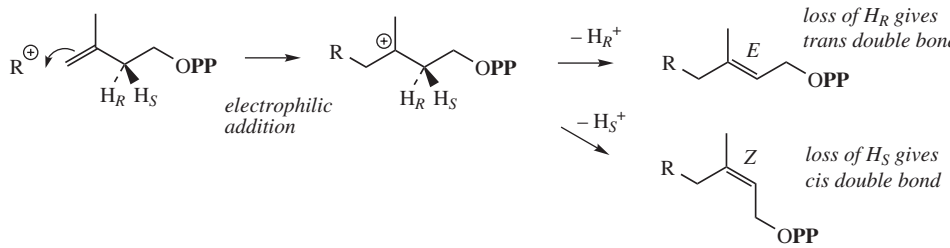


Figure 5.150

FURTHER READING

Biosynthesis, General

Ajikumar PK, Tyo K, Carlsen S, Mucha O, Phon TH and Stephanopoulos G (2008) Terpenoids: opportunities for biosynthesis of natural product drugs using engineered microorganisms. *Molecular Pharmaceutics* **5**, 167–190.

Bouvier F, Rahier A and Camara B (2005) Biogenesis, molecular regulation and function of plant isoprenoids. *Prog Lipid Res* **44**, 357–429.

Christianson DW (2006) Structural biology and chemistry of the terpenoid cyclases. *Chem Rev* **106**, 3412–3442.

Dairi T (2005) Studies on biosynthetic genes and enzymes of isoprenoids produced by actinomycetes. *J Antibiot* **58**, 227–243.

Dewick PM (2002) The biosynthesis of C_5–C_{25} terpenoid compounds. *Nat Prod Rep* **19**, 181–222.

Dubey VS, Bhalla R and Luthra R (2003) An overview of the non-mevalonate pathway for terpenoid biosynthesis in plants. *J Biosci* **28**, 637–646.

Eisenreich W, Bacher A, Arigoni D and Rohdich F (2004) Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cell Mol Life Sci* **61**, 1401–1426.

Hunter WN (2007) The non-mevalonate pathway of isoprenoid precursor biosynthesis. *J Biol Chem* **282**, 21573–21577.

Kuzuyama T and Seto H (2003) Diversity of the biosynthesis of the isoprene units. *Nat Prod Rep* **20**, 171–183.

Rodríguez-Concepción M and Boronat A (2002) Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol* **130**, 1079–1089.

Rohdich F, Bacher A and Eisenreich W (2004) Perspectives in anti-infective drug design. The late steps in the biosynthesis of the universal terpenoid precursors, isopentenyl diphosphate and dimethylallyl diphosphate. *Bioorg Chem* **32**, 292–308.

Schuhr CA, Radykewicz T, Sagner S, Latzel C, Zenk MH, Arigoni D, Bacher A, Rohdich F and Eisenreich W (2003) Quantitative assessment of crosstalk between the two isoprenoid biosynthesis pathways in plants by NMR spectroscopy. *Phytochem Rev* **2**, 3–16.

Segura MJR, Jackson BE and Matsuda SPT (2003) Mutagenesis approaches to deduce structure–function relationships in terpene synthases. *Nat Prod Rep* **20**, 304–317.

Isoprene

Sharkey TD and Yeh S (2001) Isoprene emission from plants. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 407–36.

Volatile Oils

Bentley R (2006) The nose as a stereochemist. Enantiomers and odor. *Chem Rev* **106**, 4099–4112.

Dudareva N, Pichersky E and Gershenzon J (2004) Biochemistry of plant volatiles. *Plant Physiol* **135**, 1893–1902.

Edris AE (2007) Pharmaceutical and therapeutic potentials of essential oils and their individual volatile constituents: a review. *Phytother Res* **21**, 308–323.

Kalemba D and Kunicka A (2003) Antibacterial and antifungal properties of essential oils. *Curr Med Chem* **10**, 813–829.

Pichersky E and Gershenzon J (2002) The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr Opin Plant Biol* **5**, 237–243.

Pyrethrins

Khambay BPS (2002) Pyrethroid insecticides. *Pesticide Outlook* 49–54.

Iridoids

Birkett MA and Pickett JA (2003) Aphid sex pheromones: from discovery to commercial production. *Phytochemistry* **62**, 651–656.

Dinda B, Debnath S and Harigaya Y (2007) Naturally occurring iridoids. A review, part 1. *Chem Pharm Bull* **55**, 159–222; Naturally occurring secoiridoids and bioactivity of naturally occurring iridoids and secoiridoids. A review, part 2. *Chem Pharm Bull* **55**, 689–728.

Grant L, McBean DE, Fyfe L and Warnock AM (2007) A review of the biological and potential therapeutic actions of *Harpagophytum procumbens*. *Phytother Res* **21**, 199–209.

Jensen SR, Franzyka H and Wallander E (2002) Chemotaxonomy of the Oleaceae: iridoids as taxonomic markers. *Phytochemistry* **60**, 213–231.

Oudin A, Courtois M, Rideau M and Clastre M (2007) The iridoid pathway in *Catharanthus roseus* alkaloid biosynthesis. *Phytochem Rev* **6**, 259–276.

Sesquiterpenoids

Staneva JD, Todorova MN and Evstatieva LN (2008) Sesquiterpene lactones as chemotaxonomic markers in genus *Anthemis*. *Phytochemistry* **69**, 607–618.

Artemisinin

Covello PS, Teoh KH, Polichuk DR, Reed DW and Nowak G (2007) Functional genomics and the biosynthesis of artemisinin. *Phytochemistry* **68**, 1864–1871.

Efferth T (2007) Antiplasmodial and antitumor activity of artemisinin – from bench to bedside. *Planta Med* **73**, 299–309.

Gelb MH (2007) Drug discovery for malaria: a very challenging and timely endeavor. *Curr Opin Chem Biol* **11**, 440–445.

Jansen FH and Soomro SA (2007) Chemical instability determines the biological action of the artemisinins. *Curr Med Chem* **14**, 3243–3259.

Jefford CW (2001) Why artemisinin and certain synthetic peroxides are potent antimalarials. Implications for the mode of action. *Curr Med Chem* **8**, 1803–1826.

Li Y and Wu Y-L (2003) An over four millennium story behind qinghaosu (artemisinin) – a fantastic antimalarial drug from a traditional Chinese herb. *Curr Med Chem* **10**, 2197–2230.

O'Neill PM and Posner GH (2004) A medicinal chemistry perspective on artemisinin and related endoperoxides. *J Med Chem* **47**, 2945–2964.

Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R and Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943.

Robert A, Dechy-Cabaret O, Cazelles J and Meunier B (2002) From mechanistic studies on artemisinin derivatives to new modular antimalarial drugs. *Acc Chem Res* **35**, 167–174.

Utzinger J, Shuhuac X, Keiser J, Minggan C, Jiang Z and Tanner M (2001) Current progress in the development and use of artemether for chemoprophylaxis of major human schistosome parasites. *Curr Med Chem* **8**, 1841–1859.

Wright CW (2005) Plant derived antimalarial agents: new leads and challenges. *Phytochem Rev* **4**, 55–61.

Wu Y (2002) How might qinghaosu (artemisinin) and related compounds kill the intraerythrocytic malaria parasite? A chemist's view. *Acc Chem Res* **35**, 255–259.

Gossypol

Kovacic P (2003) Mechanism of drug and toxic actions of gossypol: focus on reactive oxygen species and electron transfer. *Curr Med Chem* **10**, 2711–2718.

Trichothecenes

Grove JF (2007) The trichothecenes and their biosynthesis. *Prog Chem Org Nat Prod* **88**, 63–130.

Diterpenoids

Chinou I (2005) Labdanes of natural origin – biological activities (1981–2004). *Curr Med Chem* **12**, 1295–1317.

Keeling CI and Bohlmann J (2006) Diterpene resin acids in conifers. *Phytochemistry* **67**, 2415–2423.

Peters RJ (2006) Uncovering the complex metabolic network underlying diterpenoid phytoalexin biosynthesis in rice and other cereal crop plants. *Phytochemistry* **67**, 2307–2317.

Trapp S and Croteau R (2001) Defensive resin biosynthesis in conifers. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 689–724.

Taxol

Altmann K-H and Gertsch J (2007) Anticancer drugs from nature – natural products as a unique source of new microtubule-stabilizing agents. *Nat Prod Rep* **24**, 327–357.

Croteau R, Ketchum REB, Long RM, Kaspera R and Wildung MR (2006) Taxol biosynthesis and molecular genetics. *Phytochem Rev* **5**, 75–97.

Fang W-S and Liang X-T (2005) Recent progress in structure activity relationship and mechanistic studies of taxol analogues. *Mini Rev Med Chem* **5**, 1–12.

Ganem B and Franke RR (2007) Paclitaxel from primary taxanes: a perspective on creative invention in organozirconium chemistry. *J Org Chem* **72**, 3981–3987.

Ganesh T (2007) Improved biochemical strategies for targeted delivery of taxoids. *Bioorg Med Chem* **15**, 3597–3623.

Kingston DGI (2001) Taxol, a molecule for all seasons. *Chem Commun* 867–880.

Kingston DGI (2007) The shape of things to come: structural and synthetic studies of taxol and related compounds. *Phytochemistry* **68**, 1844–1854.

Kingston DGI, Jagtap PG, Yuan H and Samala L (2002) The chemistry of taxol and related taxoids. *Prog Chem Org Nat Prod* **84**, 53–225.

Ojima I (2008) Guided molecular missiles for tumor-targeting chemotherapy – case studies using the second-generation taxoids as warheads. *Acc Chem Res* **41**, 108–119.

Skwarczynski M, Hayashi Y and Kiso Y (2006) Paclitaxel prodrugs: toward smarter delivery of anticancer agents. *J Med Chem* **49**, 7253–7269.

Walker K and Croteau R (2001) Taxol biosynthetic genes. *Phytochemistry* **58**, 1–7.

Zhao CF, Yu LJ, Li LQ and Xiang F (2007) Simultaneous identification and determination of major taxoids from extracts of *Taxus chinensis* cell cultures. *Z Naturforsch Teil C* **62**, 1–10.

Ginkgo

Nakanishi K (2005) Terpene trilactones from *Gingko biloba*: from ancient times to the 21st century. *Bioorg Med Chem* **13**, 4987–5000.

Strømgaard K and Nakanishi K (2004) Chemistry and biology of terpene trilactones from *Ginkgo biloba*. *Angew Chem Int Ed* **43**, 1640–1658.

Stevioside

Brandle JE and Telmer PG (2007) Steviol glycoside biosynthesis. *Phytochemistry* **68**, 1855–1863.

Geuns JMC (2003) Molecules of interest. Stevioside. *Phytochemistry* **64**, 913–921.

Salvinorin

Grundmann O, Phipps SM, Zadezensky I and Butterweck V (2007) *Salvia divinorum* and salvinorin A: and update on pharmacology and analytical methodology. *Planta Med* **73**, 1039–1046.

Imanshahidi M and Hosseinzadeh H (2006) The pharmacological effects of *Salvia* species on the central nervous system. *Phytother Res* **20**, 427–437.

Prisinzano TE and Rothman RB (2008) Salvinorin A analogs as probes in opioid pharmacology. *Chem Rev* **108**, 1732–1743.

Triterpenoids

Abe I (2007) Enzymatic synthesis of cyclic triterpenes. *Nat Prod Rep* **24**, 1311–1331.

Charlton-Menys V and Durrington PN (2007) Squalene synthase inhibitors: clinical pharmacology and cholesterol-lowering potential. *Drugs* **67**, 11–16.

Chen JC, Chiu MH, Nie RL, Cordell GA and Qiu SX (2005) Cucurbitacins and cucurbitane glycosides: structures and biological activities. *Nat Prod Rep* **22**, 386–399; errata 794–795.

Dzubak P, Hajduch M, Vydra D, Hustova A, Kvasnica M, Biedermann D, Markova L, Urban M and Sarek J (2006) Pharmacological activities of natural triterpenoids and their therapeutic implications. *Nat Prod Rep* **23**, 394–411.

Hoshino T and Sato T (2002) Squalene–hopene cyclase: catalytic mechanism and substrate recognition. *Chem Commun* 291–301.

Krasutsky PA (2006) Birch bark research and development. *Nat Prod Rep* **23**, 919–942.

Phillips DR, Rasbery JM, Bartel B and Matsuda SPT (2006) Biosynthetic diversity in plant triterpene cyclization. *Curr Opin Plant Biol* **9**, 305–314.

Wendt KU (2005) Enzyme mechanisms for triterpene cyclization: new pieces of the puzzle. *Angew Chem Int Ed* **44**, 3966–3971.

Wendt KU, Schultz GE, Corey EJ and Liu DR (2000) Enzyme mechanisms for polycyclic triterpene formation. *Angew Chem Int Ed* **39**, 2812–2833.

Yogeeswari P and Sriram D (2005) Betulinic acid and its derivatives: a review on their biological properties. *Curr Med Chem* **12**, 657–666.

Triterpenoid Saponins

Kalinowska M, Zimowski J, P1czkowski C and Wojciechowski ZA (2005) The formation of sugar chains in triterpenoid saponins and glycoalkaloids. *Phytochem Rev* **4**, 237–257.

Osbourn AE (2003) Saponins in cereals. *Phytochemistry* **62**, 1–4.

Sparg SG, Light ME and van Staden J (2004) Biological activities and distribution of plant saponins. *J Ethnopharmacol* **94**, 219–243.

Townsend B, Jenner H and Osbourn A (2006) Saponin glycosylation in cereals. *Phytochem Rev* **5**, 109–114.

Xu R, Fazio GC and Matsuda SPT (2004) On the origins of triterpenoid skeletal diversity *Phytochemistry* **65**, 261–291.

Liquorice

Asl MN and Hosseinzadeh H (2008) Review of pharmacological effects of *Glycyrrhiza* sp. and its bioactive compounds. *Phytother Res* **22**, 709–724.

Baltina LA (2003) Chemical modification of glycyrrhizic acid as a route to new bioactive compounds for medicine. *Curr Med Chem* **10**, 155–171.

Fiore C, Eisenhut M, Ragazzi E, Zanchin G and Armanini D (2005) A history of the therapeutic use of liquorice in Europe. *J Ethnopharmacol* **99**, 317–324.

Ginseng

Davydov M and Krikorian AD (2000) *Eleutherococcus senticosus* (Araliaceae) as an adaptogen: a closer look. *J Ethnopharmacol* **72**, 345–393.

Leung KW, Yung KKL, Mak NK, Yue PYK, Luo H-B, Cheng Y-K, Fan TPD, Yeung HW, Ng TB and Wong RNS (2007) Angiomodulatory and neurological effects of ginsenosides. *Curr Med Chem* **14**, 1371–1380.

Panossian A and Wagner H (2005) Stimulating effect of adaptogens: an overview with particular reference to their efficacy following single dose administration. *Phytother Res* **19**, 819–838.

Park JD, Rhee DK and Lee YH (2005) Biological activities and chemistry of saponins from *Panax ginseng*. *Phytochem Rev* **4**, 159–175.

Cholesterol

Jain KS, Kathiravan MK, Somani RS and Shishoo CJ (2007) The biology and chemistry of hyperlipidemia. *Bioorg Med Chem* **15**, 4674–4699.

Libby P, Aikawa M and Schönbeck U (2000) Cholesterol and atherosclerosis. *Biochim Biophys Acta* **1529**, 299–309.

Risley JM (2002) Cholesterol biosynthesis: lanosterol to cholesterol. *J Chem Educ* **79**, 377–384.

Vance DE and Van den Bosch H (2000) Cholesterol in the year 2000. *Biochim Biophys Acta* **1529**, 1–8.

Wang M and Briggs MR (2004) HDL: The metabolism, function, and therapeutic importance. *Chem Rev* **104**, 119–137.

Waterham HR (2006) Defects of cholesterol biosynthesis. *FEBS Lett* **580**, 5442–5449.

Sterols

Benveniste P (2004) Biosynthesis and accumulation of sterols. *Annu Rev Plant Biol* **55**, 429–457.

Moreaua RA, Whitaker BD and Hicks KB (2002) Phytosterols, phytostanols, and their conjugates in foods: structural diversity, quantitative analysis, and health-promoting uses. *Prog Lipid Res* **41**, 457–500.

Nes WD (2000) Sterol methyl transferase: enzymology and inhibition. *Biochim Biophys Acta* **1529**, 63–88.

Nes WD (2003) Enzyme mechanisms for sterol *C*-methylations. *Phytochemistry* **64**, 75–95.

Schaller H (2003) The role of sterols in plant growth and development. *Prog Lipid Res* **42**, 163–175.

Vitamin D

Boland R, Skliar M, Curino A and Milanesi L (2003) Vitamin D compounds in plants. *Plant Sci* **164**, 357–369.

Chen TC, Chimeh F, Lu Z, Mathieu J, Person KS, Zhang A, Kohn N, Martinello S, Berkowitz R and Holick MF (2007) Factors that influence the cutaneous synthesis and dietary sources of vitamin D. *Arch Biochem Biophys* **460**, 213–217.

Nagpal S, Lu J and Boehm MF (2001) Vitamin D analogs: mechanism of action and therapeutic applications. *Curr Med Chem* **8**, 1661–1679.

Cardioactive Glycosides

Gobbini M and Cerri A (2005) *Digitalis*-like compounds: the discovery of the *o*-aminoalkyloxime group as a very powerful substitute for the unsaturated γ-butyrolactone moiety. *Curr Med Chem* **12**, 2343–2355.

Paula S, Tabet MR and Ball WJ (2005) Interactions between cardiac glycosides and sodium/potassium-ATPase: three-dimensional structure–activity relationship models for ligand binding to the E_2–P_i form of the enzyme versus activity inhibition. *Biochemistry* **44**, 498–510.

Schoner W (2002) Endogenous cardiac glycosides, a new class of steroid hormones. *Eur J Biochem* **269**, 2440–2448.

Bile Acids

Nonappaa and Maitra U (2008) Unlocking the potential of bile acids in synthesis, supramolecular/materials chemistry and nanoscience. *Org Biomol Chem* **6**, 657–669.

Corticosteroids

Gupta R, Jindal DP and Kumar G (2004) Corticosteroids: the mainstay in asthma therapy. *Bioorg Med Chem* **12**, 6331–6342.

Whitehouse MW (2005) Drugs to treat inflammation: a historical introduction. *Curr Med Chem* **12**, 2031–2942.

Steroid Hormones

Ahmed S, Owen CP, James K, Sampson L and Patel CK (2002) Review of estrone sulfatase and its inhibitors – an important new target against hormone dependent breast cancer. *Curr Med Chem* **9**, 263–273.

Brožic P, Lanišnik Rižner T and Gobec S (2008) Inhibitors of 17β-hydroxysteroid dehydrogenase type 1. *Curr Med Chem* **15**, 137–150.

Brunoa RD and Njar VCO (2007) Targeting cytochrome P450 enzymes: a new approach in anti-cancer drug development. *Bioorg Med Chem* **15**, 5047–5060.

Gao W, Bohl CE and Dalton JT (2005) Chemistry and structural biology of androgen receptor. *Chem Rev* **105**, 3352–3370.

Jordan VC (2003) Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 1. Receptor interactions. *J Med Chem* **46**, 883–908.

Moreira VM, Salvador JAR, Vasaitis TS and Njar VCO (2008) CYP17 inhibitors for prostate cancer treatment – an update. *Curr Med Chem* **15**, 868–899.

Poirier D (2003) Inhibitors of 17β-hydroxysteroid dehydrogenases. *Curr Med Chem* **10**, 453–477.

Saiah E (2008) The role of 11β-hydroxysteroid dehydrogenase in metabolic disease and therapeutic potential of 11β HSD1 inhibitors. *Curr Med Chem* **15**, 642–649.

Stanczyk FZ (2003) All progestins are not created equal. *Steroids* **68**, 879–890.

Yen SSC (2001) Dehydroepiandrosterone sulfate and longevity: new clues for an old friend. *Proc Natl Acad Sci USA* **98**, 8167–8169.

Carotenoids

Auldridge ME, McCarty DR and Klee HJ (2006) Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Curr Opin Plant Biol* **9**, 315–321.

DellaPenna D and Pogson BJ (2006) Vitamin synthesis in plants: tocopherols and carotenoids. *Annu Rev Plant Biol* **57**, 711–738.

Enfissi EMA, Fraser PD and Bramley PM (2006) Genetic engineering of carotenoid formation in tomato. *Phytochem Rev* **5**, 59–65.

Fraser PD and Bramley PM (2004) The biosynthesis and nutritional uses of carotenoids. *Prog Lipid Res* **43**, 228–265.

Grotewold E (2006) The genetics and biochemistry of floral pigments. *Annu Rev Plant Biol* **57**, 761–780.

Herbers K (2003) Vitamin production in transgenic plants. *Plant Physiol* **160**, 821–829.

Leuenberger MG, Engeloch-Jarret C and Woggon W-D (2001) The reaction mechanism of the enzyme-catalyzed central cleavage of β-carotene to retinal. *Angew Chem Int Ed* **40**, 2614–2617.

Römer S and Fraser PD (2005) Recent advances in carotenoid biosynthesis, regulation and manipulation. *Planta* **221**, 305–308.

Vitamin A

Kagechika H and Shudo K (2005) Synthetic retinoids: recent developments concerning structure and clinical utility. *J Med Chem* **48**, 5875–5883.

Lidén M and Eriksson U (2006) Understanding retinol metabolism: structure and function of retinol dehydrogenases. *J Biol Chem* **281**, 13001–13004.

Njar VCO, Gediya L, Purushottamachar P, Chopra P, Vasaitis TS, Khandelwal A, Mehta J, Huynh C, Belosay A and Patel J (2006) Retinoic acid metabolism blocking agents (RAMBAs) for treatment of cancer and dermatological diseases. *Bioorg Med Chem* **14**, 4323–4340.

Higher Terpenoids

Swiezewska E and Danikiewicz W (2005) Polyisoprenoids: structure, biosynthesis and function. *Prog Lipid Res* **44**, 235–258.