3

THE ACETATE PATHWAY: FATTY ACIDS AND POLYKETIDES

Polyketides constitute a large class of natural products grouped together on purely biosynthetic grounds. Their diverse structures can be explained as being derived from poly-β-keto chains, formed by coupling of acetic acid (C_2) units via condensation reactions, i.e.

$$n CH_3CO_2H \rightarrow —[CH_2CO]_n—$$

Included in such compounds are the fatty acids, polyacetylenes, prostaglandins, macrolide antibiotics and many aromatic compounds, e.g. anthraquinones and tetracyclines.

The formation of the poly-β-keto chain could be envisaged as a series of Claisen reactions, the reverse of which are involved in the β-oxidation sequence for the metabolism of fatty acids (see page 18). Thus, two molecules of acetyl-CoA could participate in a Claisen condensation giving acetoacetyl-CoA, and this reaction could be repeated to generate a poly-β-keto ester of appropriate chain length (Figure 3.1). However, a study of the enzymes involved in fatty acid biosynthesis showed this simple rationalization could not be correct and that a more complex series of reactions was operating. It is now known that fatty acid biosynthesis involves initial carboxylation of acetyl-CoA to malonyl-CoA, a reaction involving ATP, CO_2 (as bicarbonate, HCO_3^-), and the coenzyme biotin as the carrier of CO_2 (see page 17).

The conversion of acetyl-CoA into malonyl-CoA increases the acidity of the α-hydrogen atoms, thus providing a better nucleophile for the Claisen condensation. In the biosynthetic sequence, no acylated malonic acid derivatives are produced, and no label from [^{14}C]bicarbonate is incorporated, so the carboxyl group introduced into malonyl-CoA is simultaneously lost by a decarboxylation reaction during the Claisen condensation (Figure 3.1). Accordingly, the carboxylation step helps to activate the α-carbon and facilitate Claisen condensation, and the carboxyl is immediately removed on completion of this task. An alternative rationalization is that decarboxylation of the malonyl ester is used to generate the acetyl enolate anion without any requirement for a strong base.

The pathways to fatty acids, macrolides, and aromatic polyketides branch early. The chain extension process of Figure 3.1 continues for aromatics, generating a highly reactive poly-β-keto chain that is stabilized by association with groups on the enzyme surface until chain assembly is complete and cyclization reactions occur. However, for fatty acids, the carbonyl groups are reduced before attachment of the next malonate group. Partial reduction processes, leading to a mixture of methylenes, hydroxyls, and carbonyls, are characteristic of macrolides (see page 68).

FATTY ACID SYNTHASE: SATURATED FATTY ACIDS

The processes of fatty acid biosynthesis are well studied and are known to be catalysed by the enzyme **fatty acid synthase** (FAS). FASs from various organisms show significant structural differences. In animals, FAS is a large multifunctional protein with seven discrete functional domains, providing all of the catalytic activities required. All domains are on a single polypeptide, encoded by a single gene, though the enzyme exists as a homodimer and requires both units for activity. Fungal FAS is also a multifunctional enzyme, but the seven component activities are distributed over two non-identical polypeptides α and β, and the enzyme is an $\alpha_6\beta_6$ dodecamer. The

Medicinal Natural Products: A Biosynthetic Approach, 3rd Edition. Paul M. Dewick
 ISBN: 978-0-470-74168-9

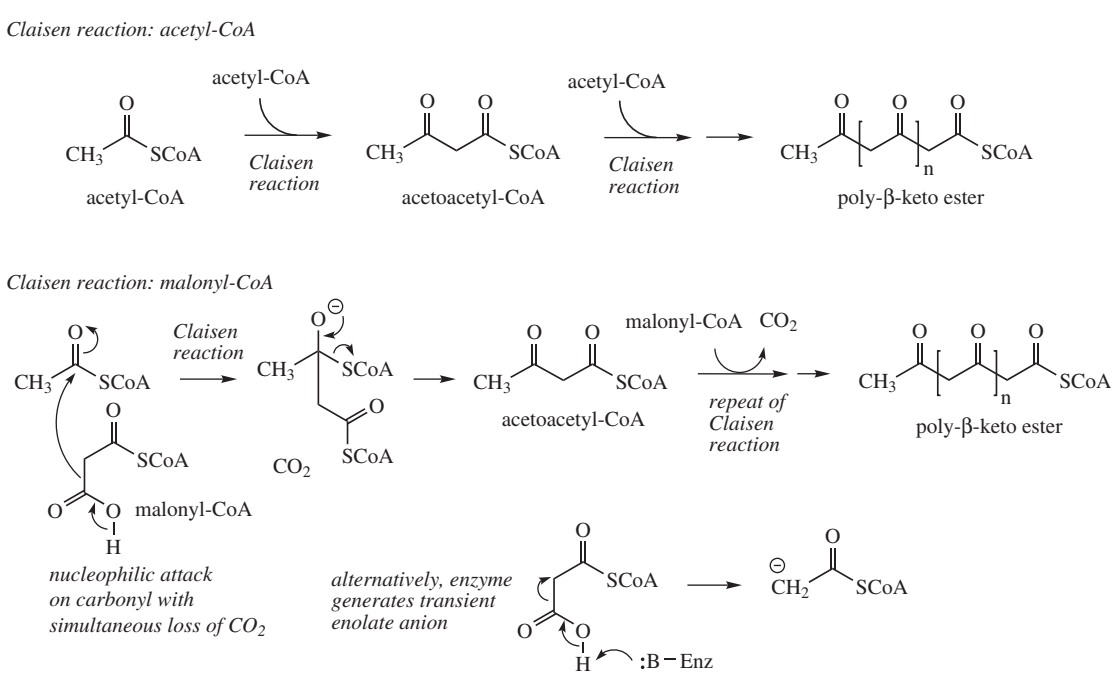


Figure 3.1

multifunctional protein systems in animals and fungi are termed **type I FASs**. Bacteria and plants possess a **type II FAS** that is an assembly of separable enzymes, encoded by seven different genes. Nevertheless, all of the different FAS systems perform effectively the same task employing the same mechanisms.

Acetyl-CoA and malonyl-CoA themselves are not involved in the condensation step: they are converted into enzyme-bound thioesters, the malonyl ester by means of an acyl carrier protein (ACP) (Figure 3.2). The Claisen reaction follows giving acetoacetyl-ACP (β-ketoacyl-ACP; R = H), which is reduced stereospecifically to the corresponding β-hydroxy ester, consuming NADPH in the reaction. Then follows elimination of water giving the *E* (*trans*) α,β-unsaturated ester. Reduction of the double bond again utilizes NADPH and generates a saturated acyl-ACP (fatty acyl-ACP; R = H) which is two carbon atoms longer than the starting material. This can feed back into the system, condensing again with malonyl-ACP and going through successive reduction, dehydration and reduction steps, gradually increasing the chain length by two carbon atoms for each cycle until the required chain length is obtained. At that point, the fatty acyl chain can be released as a fatty acyl-CoA or as the free acid. Fungi typically produce the CoA esters by a reversal of the loading reaction, whilst mammals release the free acid by the action of a thioesterase (TE). The chain length actually elaborated is probably controlled by the specificity of the TE enzyme.

The FAS protein is known to contain an ACP binding site, and also an active-site cysteine residue in the ketosynthase (KS) domain. Acetyl and malonyl groups are successively transferred from coenzyme A esters and attached to the thiol groups of Cys and ACP respectively by the same malonyl/acetyl transferase (MAT; Figure 3.3). The Claisen condensation occurs (KS), and the processes of reduction (ketoreductase, KR), dehydration (dehydratase, DH), and reduction (enoyl reductase, ER) then occur whilst the growing chain is attached to ACP. The ACP carries a phosphopantetheine group exactly analogous to that in coenzyme A, and this provides a long flexible arm, enabling the growing fatty acid chain to reach the active site of each enzyme in the complex, allowing the different chemical reactions to be performed without releasing intermediates from the enzyme (compare polyketide synthesis on page 70 and peptide synthesis on page 438). The sequence of enzyme activities along the protein chain of the enzyme complex does not correspond with the order in which they are employed (see Figure 3.4). Then the chain is transferred to the thiol of Cys, and the process can continue. Making the process even more efficient, the dimeric animal FAS contains

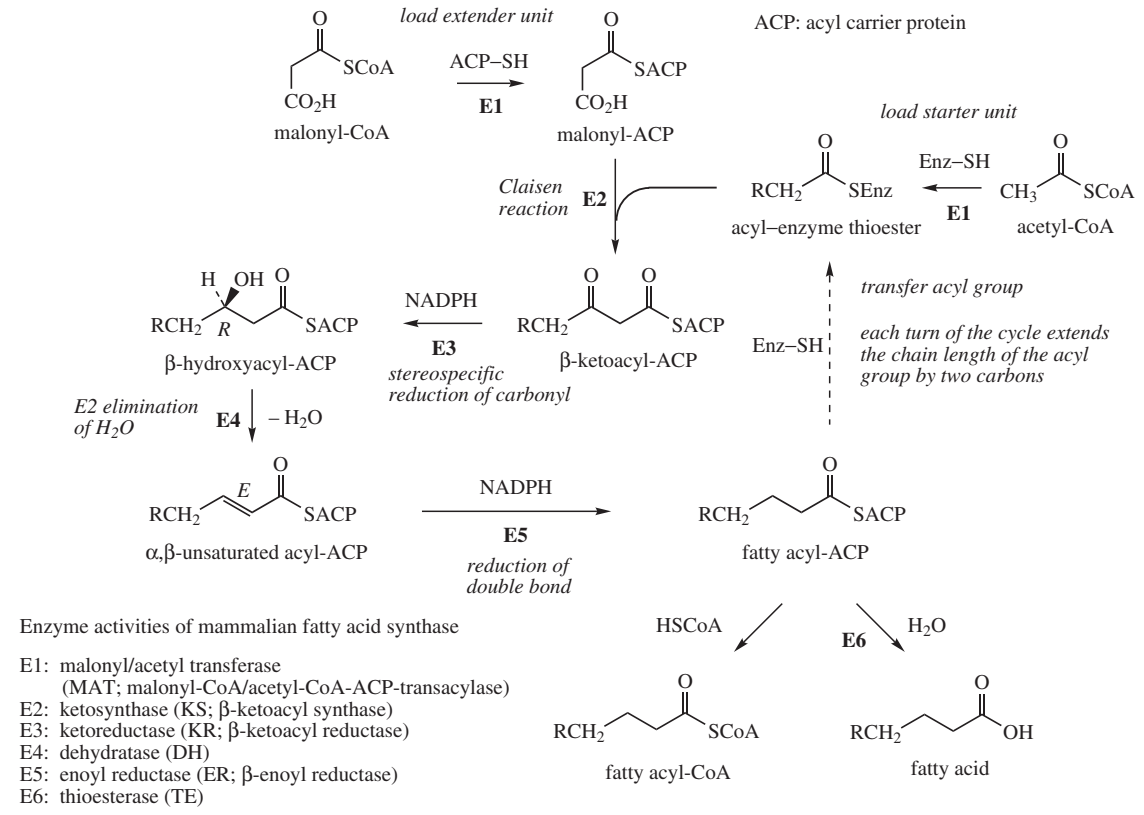


Figure 3.2

two catalytic centres and is able to generate two growing chains at the same time. The monomeric subunits are also arranged so that, although each monomer can catalyse the full set of reactions, there also exists the potential for utilizing domains on the other subunit for the initial reactions of substrate loading (MAT) and Claisen condensation (KS).

Although fungal FAS is also a type I multifunctional protein, there are some differences compared with the mammalian enzyme. Thus, the functional domains are arranged on two separate polypeptides (Figure 3.4), and the complex is a dodacamer rather than a dimer. Acetyl-CoA and malonyl-CoA are loaded onto the system by separate enzyme activities, namely acetyl transferase (AT) and malonyl/palmitoyl transferase (MPT), the latter also being involved in releasing the product palmitoyl-CoA from the ACP. A further enzyme activity is also included, phosphopantetheinyl transferase (PPT), which is required for activating the ACP by inserting the phosphopantetheine arm. The dissociated FAS II system found in most bacteria and plants consists of a series of discrete proteins, each of which catalyses an individual reaction of the fatty acid biosynthetic pathway. In some cases, two or more enzymes are able to perform the same chemical reaction, but have differing substrate specificities. For example, three distinct KS activities with characteristic chain-length specificities have been identified in plants. One condenses acetyl-CoA with malonyl-ACP, one is responsible for further elongation to C_{16}, and a third one elongates C_{16} to C_{18}. This allows both C_{16} and C_{18} metabolites to be tapped off for different functions. Unusually, the tuberculosis-causing bacillus *Mycobacterium tuberculosis* is known to possess both FAS I and FAS II systems. It uses FAS I to produce acyl chains in the range C_{14}–C_{16}, and FAS II to elongate these structures to mycolic acids, very long-chain fatty acids with 24–56 carbon atoms that are components of the mycobacterial cell wall. Because of the fundamental differences in mammalian FAS I and bacterial FAS II, there is considerable potential for exploiting these differences and developing selective inhibitors of fatty acid synthesis as antibacterial agents. Indeed, this approach has been

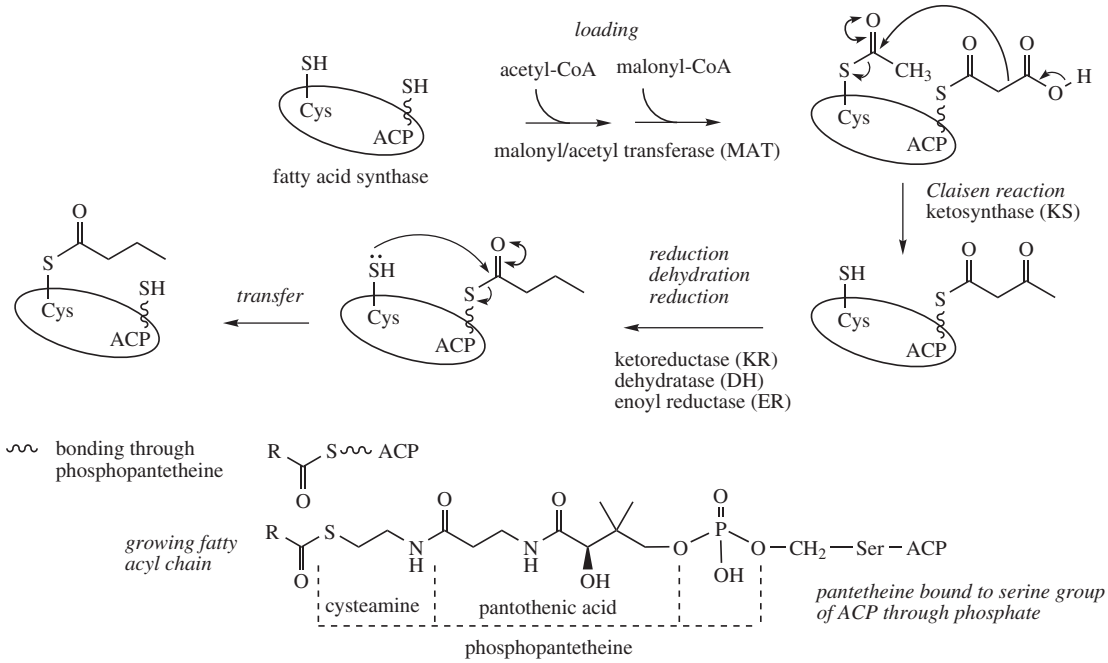


Figure 3.3

domain organization: mammalian FAS

— KS — MAT — DH — ER — KR — ACP — TE —

domain organization: fungal FAS

α — ACP — KR — KS — PPT —

β — AT — ER — DH — MPT —

ACP: acyl carrier protein
AT: acetyl transferase
DH: dehydratase
ER: enoyl reductase
KR: ketoreductase
KS: ketosynthase
MAT: malonyl/acetyl transferase
MPT: malonyl/palmitoyl transferase
PPT: phosphopantetheinyl transferase
TE: thioesterase

Figure 3.4

successful with agents such as isoniazid (antituberculosis) and triclosan (general biocide), both inhibitors of bacterial ER. Considerable effort is being invested in targeting FAS II in *Plasmodium falciparum*, the causative agent of malaria.

The combination of one acetate starter unit with seven malonate extender units would give the C_{16} fatty acid, palmitic acid, and with eight malonates the C_{18} fatty acid, stearic acid (Figure 3.5). Note that the two carbon atoms at the head of the chain (methyl end) are provided by acetate, not malonate, whilst the remainder are derived from malonate. However, malonate itself is produced by carboxylation of acetate. This means that all carbon atoms in the fatty acid originate from acetate, but malonate will only provide the C_2 chain extension units and not the C_2 starter group. The linear combination of acetate C_2 units as in Figure 3.2 explains why the common fatty acids are straight-chained and possess an even number of carbon atoms. Natural fatty acids may contain from 4 to 30, or even more, carbon atoms, the most abundant being those with 16 or 18 carbon atoms. Some naturally occurring fatty acids are shown in Figure 3.6. The rarer fatty acids containing an odd number of carbon atoms typically originate from incorporation of a different starter unit, e.g. propionic acid, or can arise by loss of one carbon from an even-numbered acid. Other structures (see page 53) can arise by utilizing different starter units and/or different extender units.

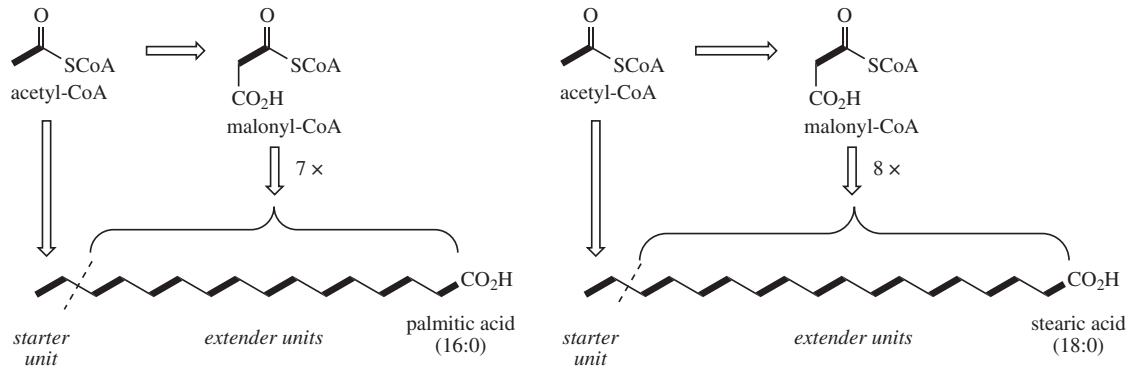
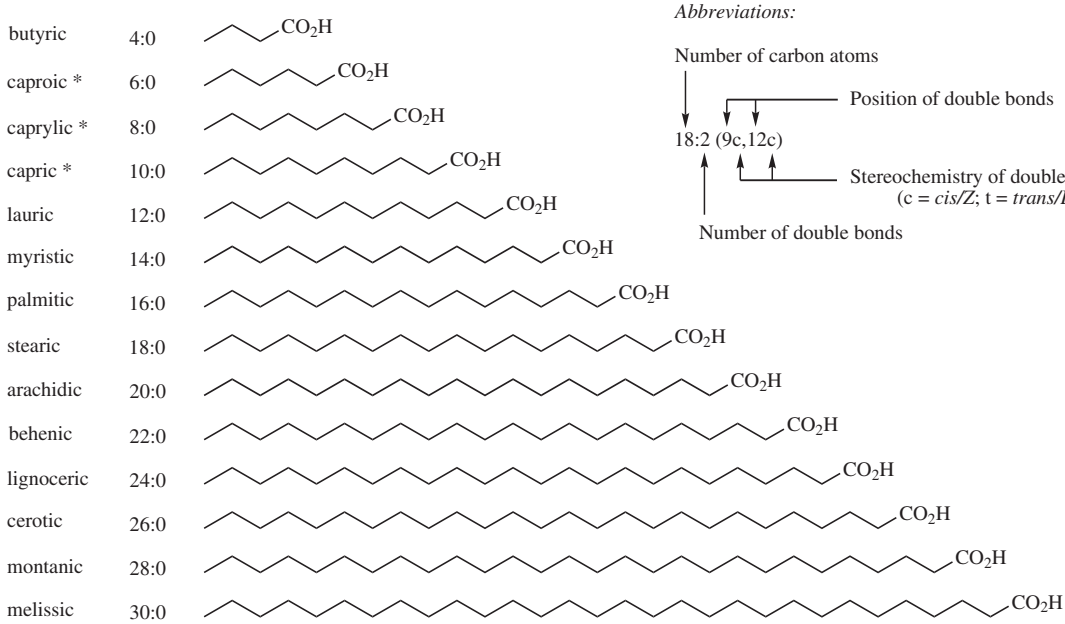


Figure 3.5

Fatty acids are mainly found in ester combination with glycerol in the form of triglycerides (Figure 3.7). These materials are called **fats** or **oils**, depending on whether they are solid or liquid at room temperature. If all three esterifying acids are the same, then the triglyceride is termed simple, whereas a mixed triglyceride is produced if two or more of the fatty acids are different. Most natural fats and oils are composed largely of mixed triglycerides. In this case, isomers can exist, including potential optical isomers, since the central carbon will become chiral if the primary alcohols of glycerol are esterified with different fatty acids. In practice, only one of each pair

Common naturally occurring fatty acids

Saturated



* To avoid confusion, systematic nomenclature (hexanoic, octanoic, decanoic) is recommended

Figure 3.6 (*continued overleaf*)

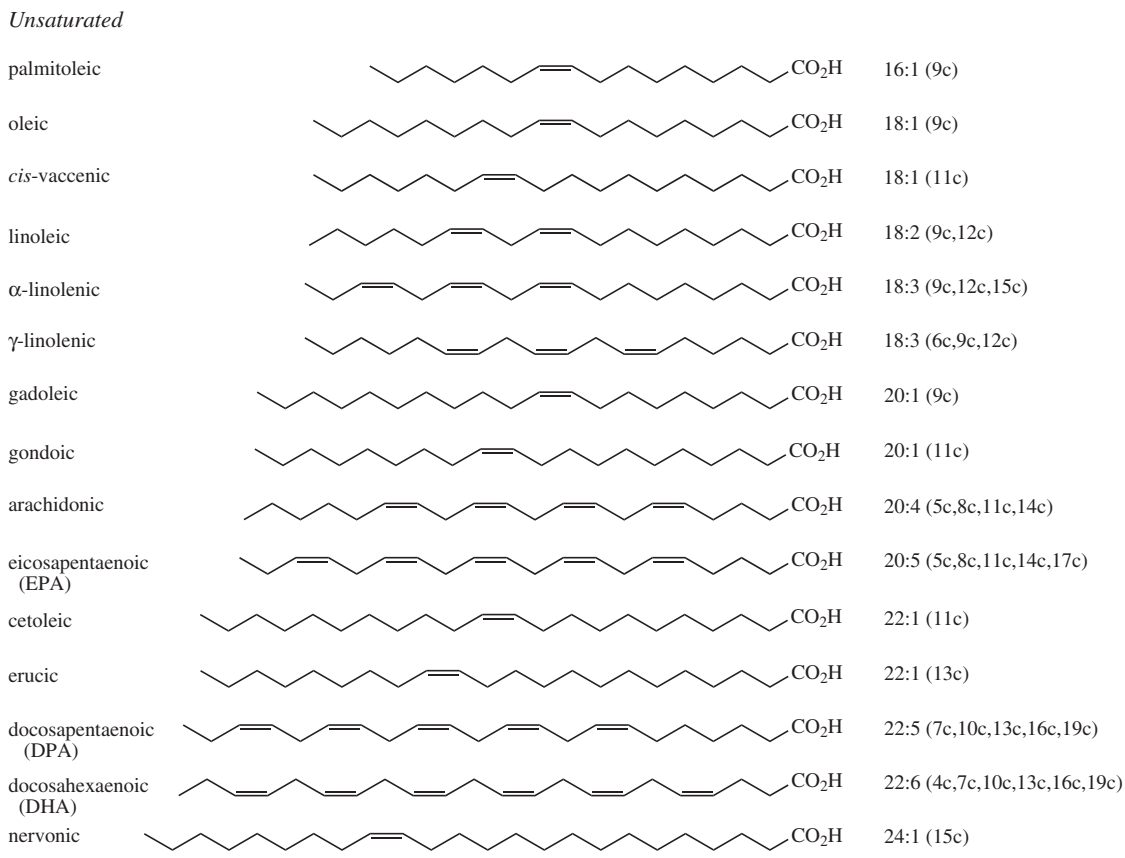


Figure 3.6 (*continued*)

of enantiomers is formed in nature. Triglycerides are produced predominantly from glycerol 3-phosphate by esterification with fatty acyl-CoA residues, the phosphate being removed prior to the last esterification (Figure 3.7). The diacyl ester of glycerol 3-phosphate is also known as a **phosphatidic acid**, and is the basis of phospholipid structures. In these structures, the phosphate is also esterified with an alcohol, which is usually choline, ethanolamine, serine, or *myo*-inositol, e.g. **phosphatidyl choline** (Figure 3.8). **Phospholipids** are important structural components of cell membranes, and because of the polar and non-polar regions in their structure, they have detergent-like properties. They are also able to form liposomes, which have considerable potential as drug delivery systems. A particularly important natural phospholipid is **platelet-activating factor** (**PAF**; Figure 3.8), which resembles a phosphatidylcholine, though this compound possesses an ether linkage to a long-chain fatty alcohol, usually hexadecanol, rather than an ester linkage. The central hydroxyl of glycerol is esterified, but to acetic acid rather than to a long-chain fatty acid. PAF functions at nanomolar concentrations, activates blood platelets, and contributes to diverse biological effects, including thrombosis, inflammatory reactions, allergies, and tissue rejection. Long-chain alcohols are reduction products from fatty acyl-CoA esters; they also feature in natural **waxes**. Waxes are complex mixtures of esters of long-chain fatty acids, usually C_{20}–C_{24}, with long-chain monohydric alcohols or sterols. The alcohol is esterified using a fatty acyl-CoA.

UNSATURATED FATTY ACIDS

Animal fats contain a high proportion of glycerides of saturated fatty acids and tend to be solids, whilst those from plants and fish contain predominantly unsaturated fatty acid esters and tend to be liquids. Some of the common naturally occurring unsaturated fatty acids are also included in Figure 3.6. A convenient shorthand

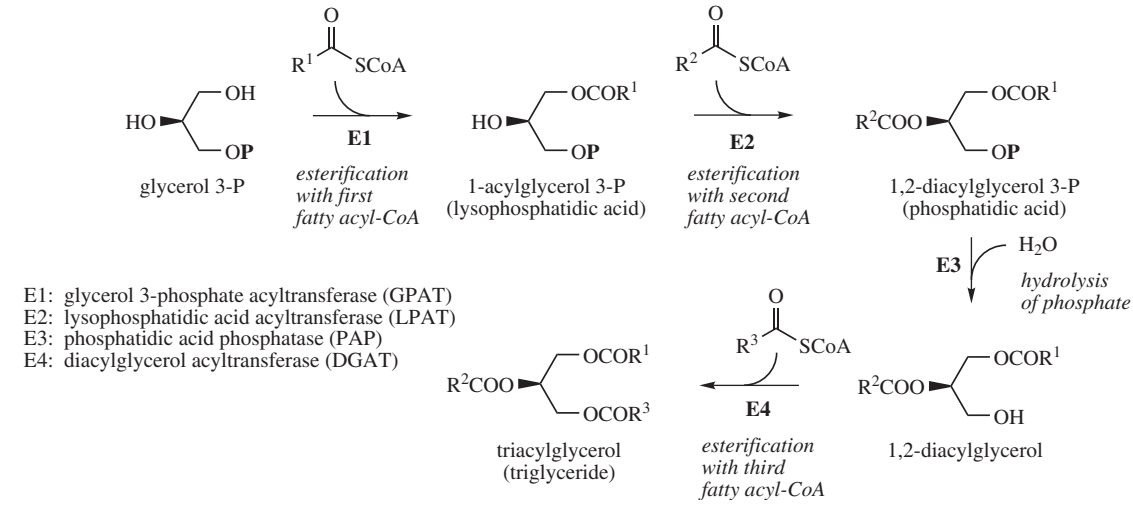


Figure 3.7

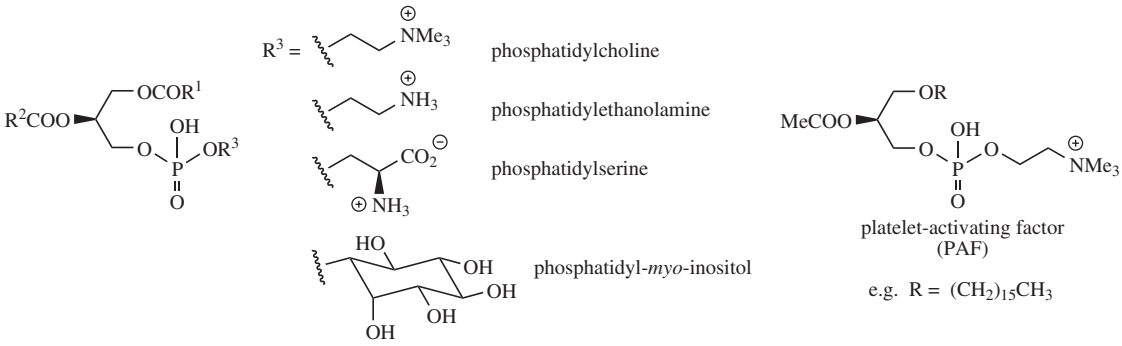


Figure 3.8

representation for fatty acids indicating chain length with number, position and stereochemistry of double bonds is also presented in Figure 3.6. A less systematic numbering starting from the methyl (the ω end) may also be encountered. Major groups of fatty acids are designated ω–3 (omega-3), ω–6 (omega-6), ω–9 (omega-9), etc. (or sometimes n–3, n–6, n–9), if there is a double bond that number of carbon atoms from the methyl terminus. This has some value in relating structures when an unsaturated fatty acid is biosynthetically elongated from the carboxyl end as during prostaglandin biosynthesis (see page 50). Double bonds at position 9 are common, but unsaturation can occur at other positions in the chain. Polyunsaturated fatty acids tend to have their double bonds in a non-conjugated array as a repeating unit —(CH═CHCH$_2$)$_n$—. In virtually all cases, the stereochemistry of the double bond is *Z* (*cis*), thus introducing a 'bend' into the alkyl chain. This interferes with the close association and aggregation of molecules that is possible in saturated structures and helps to maintain the fluidity in oils and cellular membranes.

Fats and oils represent long-term stores of energy for most organisms, being subjected to oxidative metabolism as required. Major oils which are produced commercially for use as foods, toiletries, medicinals, or pharmaceutical formulation aids are listed in Table 3.1. Typical fatty acid analyses are shown, though it must be appreciated that these figures can vary quite widely. For instance, plant oils show significant variation according to the climatic conditions under which the plant was grown. In colder climates, a higher proportion of polyunsaturated fatty acids is produced, so that the plant can maintain the fluidity

Table 3.1 Fixed oils and fats.[a,b,c]

Oil	Source	Part used	Oil content (%)	Typical fatty acid composition (%)	Uses, notes
Almond	*Prunus amygdalus* var. *dulcis*, or var. *amara* (Rosaceae)	seed	40–55	oleic (62–86), linoleic (7–30), palmitic (4–9), stearic (1–2)	emollient base, toiletries, carrier oil (aromatherapy)
Arachis (groundnut, peanut)	*Arachis hypogaea* (Leguminosae/Fabaceae)	seed	45–55	oleic (35–72), linoleic (13–43), palmitic (7–16), stearic (1–7), behenic (1–5), arachidic (1–3)	food oil, emollient base
Borage (starflower)	*Borago officinalis* (Boraginaceae)	seed	28–35	linoleic (38), γ-linolenic (23–26), oleic (16), palmitic (11)	dietary supplement for γ-linolenic (gamolenic) acid content in treatment of premenstrual tension, breast pain, eczema
Butterfat	cow *Bos taurus* (Bovidae)	milk	2–5	palmitic (29), oleic (28), stearic (13), myristic (12), butyric (4), lauric (3), caproic (2), capric (2), palmitoleic (2)	food
Castor	*Ricinus communis* (Euphorbiaceae)	seed	35–55	ricinoleic (80–90), oleic (4–9), linoleic (2–7), palmitic (2–3), stearic (2–3)	emollient base, purgative, soap manufacture; castor seeds contain the highly toxic, but heat-labile protein ricin (see page 435)
Coconut	*Cocos nucifera* (Palmae/Arecaceae)	seed kernel	65–68	lauric (43–53), myristic (15–21), palmitic (7–11), caprylic (5–10), capric (5–10), oleic (6–8), stearic (2–4)	soaps, shampoos; fractionated coconut oil containing only short- to medium-length fatty acids (mainly caprylic and capric) is a dietary supplement
Cod-liver	cod *Gadus morrhua* (Gadidae)	fresh liver	50	oleic (24), DHA (14), palmitic (11), EPA (6), palmitoleic (7), stearic (4), myristic (3)	dietary supplement due to presence of EPA and DHA, plus vitamins A (see page 304) and D (see page 258); halibut-liver oil from halibut *Hippoglossus vulgaris* (Pleurnectideae) has similar properties and is used in the same way
Cottonseed	*Gossypium hirsutum* (Malvaceae)	seed	15–36	linoleic (33–58), palmitic (17–29), oleic (13–44), stearic (1–4)	solvent for injections, soaps; cotton seeds also contain 1.1–1.3% gossypol (see page 220) and small amounts of cyclopropenoid fatty acids, e.g. sterculic and malvalic acids (see page 55)
Evening primrose	*Oenothera biennis* (Onagraceae)	seed	24	linoleic (65–80), γ-linolenic (7–14), oleic (9), palmitic (7)	dietary supplement for γ-linolenic (gamolenic) acid content in treatment of premenstrual tension, breast pain, eczema

Fish	various, including herring *Clupea harengus* (Clupeidae) (Europe) and menhaden *Brevoortia* spp. (Clupeidae) (N America)	whole fish – by-product of fishing and fish-meal industry	up to 20% (herring)	herring: cetoleic (7–30), oleic (9–25), gondoic (7–24), palmitic (10–19), EPA (4–15), palmitoleic (6–12), myristic (5–8), DHA (2–8), stearic (1–2)	dietary supplement for EPA and DHA (see page 49); refined and concentrated oils are produced containing higher amounts of EPA (15–30%) and DHA (10–20%)
Honesty	*Lunaria annua* (Cruciferae/ Brassicaceae)	seed	30–40	erucic (43), nervonic (25), oleic (24)	nervonic acid is being investigated for the treatment of multiple sclerosis; the disease is characterized by a deficiency in nervonic acid
Lard	pig *Sus scrofa* (Suidae)	abdominal fat		oleic (45), palmitic (25), stearic (12), linoleic (10), palmitoleic (3)	food
Linseed (flaxseed)	*Linum usitatissimum* (Linaceae)	seed	35–44	α-linolenic (30–60), oleic (39), linoleic (15), palmitic (7), stearic (4)	liniments, dietary supplement for α-linolenic acid content; formerly the basis of paints, reacting with oxygen, polymerizing, and drying to a hard film
Maize (corn)	*Zea mays* (Graminae/Poaceae)	embryo	33–39	linoleic (34–62), oleic (19–50), palmitic (8–19), stearic (0–4)	food oil, dietary supplement, solvent for injections
Olive	*Olea europaea* (Oleaceae)	fruits	15–40	oleic (56–85), palmitic (8–20), linoleic (4–20), stearic (1–4)	food oil, emollient base
Palm kernel	*Elaeis guineensis* (Palmae/Arecaceae)	kernel	45–50	lauric (40–52), myristic (14–18), oleic (9–16), palmitoleic (6–10), caprylic (3–6), capric (3–5), stearic (1–4), linoleic (1–3)	soaps; fractionated palm oil is a solid obtained by fractionation and hydrogenation and is used as a suppository base
Rapeseed	*Brassica napus* (Cruciferae/ Brassicaceae)	seed	40–50	erucic (30–60), oleic (9–25), linoleic (11–25), gadoleic (5–15), α-linolenic (5–12), palmitic (0–5)	food oil, using varieties producing lower levels of erucic acid where the main components are now oleic (48–60%), linoleic (18–30%), α-linolenic (6–14%), and palmitic (3–6%) acids; erucic acid is known to accumulate and cause lesions in heart muscle; erucic acid is used as a plasticizer in PVC clingfilm
Sesame	*Sesamum indicum* (Pedaliaceae)	seed	44–54	oleic (35–50), linoleic (35–50), palmitic (7–12), stearic (4–6)	food oil, soaps, solvent for injections, carrier oil (aromatherapy)

(*continued overleaf*)

Table 3.1 (*continued*)

Oil	Source	Part used	Oil content (%)	Typical fatty acid composition (%)	Uses, notes
Soya (soybean)	*Glycine max* (Leguminosae/ Fabaceae)	seed	18–20	linoleic (44–62), oleic (19–30), palmitic (7–14), α-linolenic (4–11), stearic (1–5)	food oil, dietary supplement, carrier oil (aromatherapy); soya oil contains substantial amounts of the sterols sitosterol and stigmasterol (see page 255)
Suet (mutton tallow)	sheep *Ovis aries* (Bovidae)	abdominal fat		stearic (32), oleic (31), palmitic (27), myristic (6)	foods
Suet (beef tallow)	cow *Bos taurus* (Bovidae)	abdominal fat		oleic (48), palmitic (27), palmitoleic (11), stearic (7), myristic (3)	foods
Sunflower	*Helianthus annuus* (Compositae/ Asteraceae)	seed	22–36	linoleic (50–70), oleic (20–40), palmitic (3–10), stearic (1–10)	food oil, carrier oil (aromatherapy)
Theobroma	*Theobroma cacao* (Sterculiaceae)	kernel	35–50	oleic (35), stearic (35), palmitic (26), linoleic (3)	suppository base, chocolate manufacture; theobroma oil (cocoa butter) is a solid

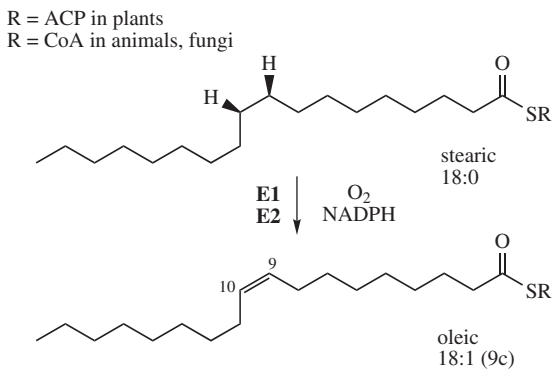
[a] The oil yields and fatty acid compositions given in the above table are typical values, and can vary widely. The quality of an oil is determined principally by its fatty acid analysis. Structures of the fatty acids are shown in Figure 3.6 (see page 43, 44)

[b] The term fat or oil has no precise significance, merely describing whether the material is a solid (fat) or liquid (oil) at room temperature. Most commercial oils are obtained from plant sources, particularly seeds and fruits, and the oil is extracted by cold or hot expression, or less commonly by solvent extraction with hexane. The crude oil is then refined by filtration, steaming, neutralization to remove free acids, washing and bleaching as appropriate. Many food oils are then partially hydrogenated to produce semi-solid fats. Animal fats and fish oils are usually extracted by steaming, the higher temperature deactivating enzymes which would otherwise begin to hydrolyse the glycerides.

[c] Oils and fats feature as important food components and cooking oils, some 80% of commercial production being used as human food, whilst animal feeds account for another 6%. Most of the remaining production is used as the basis of soaps, detergents, and pharmaceutical creams and ointments. A number of oils are used as diluents (carrier or base oils) for the volatile oils employed in aromatherapy.

of its storage fats and membranes. The melting points of these materials depend on the relative proportions of the various fatty acids, reflecting primarily the chain length and the amount of unsaturation in the chain. Saturation and increasing chain length in the fatty acids gives a more solid fat at room temperature. Thus, butterfat and cocoa butter (theobroma oil) contain a relatively high proportion of saturated fatty acids and are solids. Palm kernel and coconut oils are both semi-solids having a high concentration of the saturated C_{12} acid **lauric acid**. A characteristic feature of olive oil is its very high **oleic acid** (18:1) content, whilst rapeseed oil possesses high concentrations of long-chain C_{20} and C_{22} fatty acids, e.g. **erucic acid** (22:1). Typical fatty acids in fish oils have high unsaturation and also long chain lengths, e.g. **eicosapentaenoic acid (EPA)** (20:5) and **docosahexaenoic acid (DHA)** (22:6) in cod liver oil.

Unsaturated fatty acids can arise by more than one biosynthetic route, but in most organisms the common mechanism is by desaturation of the corresponding alkanoic acid, with further desaturation in subsequent steps. Most eukaryotic organisms possess a Δ^9-desaturase enzyme that introduces a *cis* double bond into a saturated fatty acid, requiring O_2 and NADPH or NADH cofactors. The mechanism of desaturation does not involve any intermediates hydroxylated at C-9 or C-10, and the requirement for O_2 is as an acceptor at the end of an electron transport chain. A stearoyl (C_{18}) thioester is the usual substrate giving an oleoyl derivative (Figure 3.9), coenzyme A esters being utilized by animal and fungal enzymes, and ACP esters by plant systems. In some systems, desaturation may take place on fatty acids bound as lipids or phospholipids.

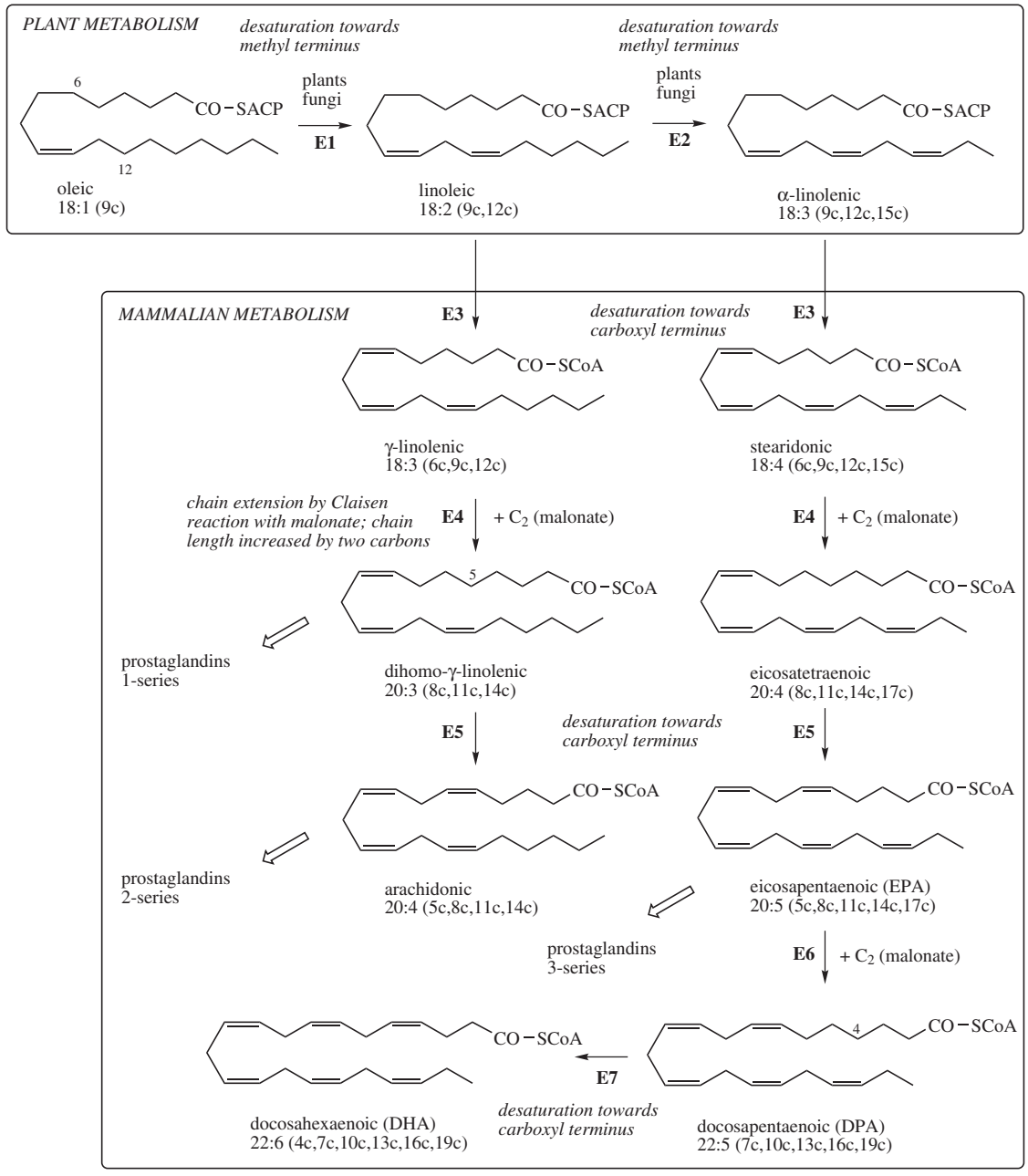


E1: stearoyl-ACP Δ^9-desaturase
E2: stearoyl-CoA Δ^9-desaturase

Figure 3.9

The position of further desaturation then depends very much on the organism. Non-mammalian enzymes tend to introduce additional double bonds between the existing double bond and the methyl terminus, e.g. oleic acid → linoleic acid → α-linolenic acid (Figure 3.10). Animals introduce new double bonds towards the carboxyl group. They also lack the Δ^{12} and Δ^{15} desaturase enzymes, so, although **linoleic acid** and **α-linolenic acid** are necessary for the synthesis of polyunsaturated acids that lead to prostaglandins (see page 59) and leukotrienes (see page 64), animals must obtain these materials in the diet, mainly from plants. Δ^6-Desaturation (towards the carboxyl) of linoleic acid leads to γ-linolenic acid, whilst analogous desaturation of α-linolenic acid gives stearidonic acid. Prostaglandins are derived from C_{20} polyunsaturated fatty acid precursors, so addition of two extra carbon atoms is required. This is achieved by an elongase enzyme acting upon γ-linolenic acid, giving **dihomo-γ-linolenic acid** ($\Delta^{8,11,14}$-eicosatrienoic acid). The additional two carbon atoms derive from malonate, by the usual chain extension mechanism. Dihomo-γ-linolenic acid is the precursor of prostaglandins in the 'one' series; the 'two' series derives from **arachidonic acid** ($\Delta^{5,8,11,14}$-eicosatetraenoic acid), formed by further Δ^5-desaturation, again towards the carboxyl terminus. The 'three' series of prostaglandins are derivatives of **$\Delta^{5,8,11,14,17}$-eicosapentaenoic acid** (**EPA**). This is an analogue of arachidonic acid formed from α-linolenic acid via stearidonic acid with a similar chain extension process and Δ^5-desaturation.

Further chain extension of EPA gives **docosapentaenoic acid** (**DPA**), and then Δ^4-desaturation gives **DHA**. DHA is vital for proper visual and neurological development in infants, and deficiency has been associated with cognitive decline and the onset of Alzheimer's disease in adults. Thus, a range of metabolites necessary for good health, including prostaglandins, leukotrienes, and these long-chain polyunsaturated fatty acids, are produced from the plant fatty acids linoleic acid and α-linolenic acid, which have to be obtained in the diet. Accordingly, these plant fatty acids are referred to as '**essential fatty acids**' (**EFAs**). Further, since beneficial fatty acids, including α-linolenic acid, EPA, DPA, and DHA, have a double bond three carbon atoms from the methyl end of the chain, they are grouped together under the term ω–3 fatty acids (**omega-3 fatty acids**). Marine fish represent a major source of the nutritionally relevant longer chain fatty acids, especially EPA and DHA, which are not found in seed oils of higher plants. Regular consumption of oily fish, e.g. herring, tuna, mackerel, or use of fish oil supplements, is claimed to reduce the risk of heart attacks and atherosclerosis. Fish obtain most of



E1: Δ^{12}-desaturase E3: Δ^{6}-desaturase E5: Δ^{5}-desaturase E7: Δ^{4}-desaturase
E2: Δ^{15}-desaturase E4: C_{18} elongase E6: C_{20} elongase

Note: the names given are for the appropriate fatty acid; the structures shown are the thioesters involved in the conversions

Figure 3.10

these fatty acids by consumption of marine microalgae, which are considered to be the primary producers. It has also been found that these microalgae utilize another approach to synthesize unsaturated fatty acids, involving a multifunctional complex analogous to polyketide synthases (PKSs; see page 67) and a sequence that does not require desaturases and elongases. Some of these microalgae, e.g. *Crypthecodinium cohnii* and *Schizochytrium* spp., are currently exploited for the commercial production of DHA-enriched oils, which are also more palatable than fish-derived products. Limited fish stocks and expensive microbial culture processes encourage researchers to consider alternative plant-based production of these beneficial long-chain polyunsaturated fatty acids. There has already been considerable success in producing transgenic oilseed crops that synthesize fatty acids such as arachidonic acid, EPA, and DHA. Arachidonic acid itself has not been found in higher plants, but does occur in some algae, mosses, and ferns. All of the omega-3 fatty acids are beneficial to health, but the plant and fish oils provide different compounds, and the distinction between the two groups should be recognized.

Although most plant-derived oils contain high amounts of unsaturated fatty acid glycerides, including those of linoleic and α-linolenic acids, the conversion of linoleic acid into **γ-linolenic acid** can be blocked or inhibited in certain conditions in humans. This restricts synthesis of prostaglandins. In such cases, the use of food supplements, e.g. **evening primrose oil** from *Oenothera biennis* (Onagraceae) or **borage oil** from *Borago officinalis* (Boraginaceae), which are rich in γ-linolenic esters (see Table 3.1), can be valuable and help in the disorder. These plants are somewhat unusual in their ability to desaturate linoleic esters towards the carboxyl terminus via a Δ^6-desaturase, rather than towards the methyl terminus as is more common in plants. Expression of Δ^6-desaturase genes, either from plants or suitable fungi, can be used

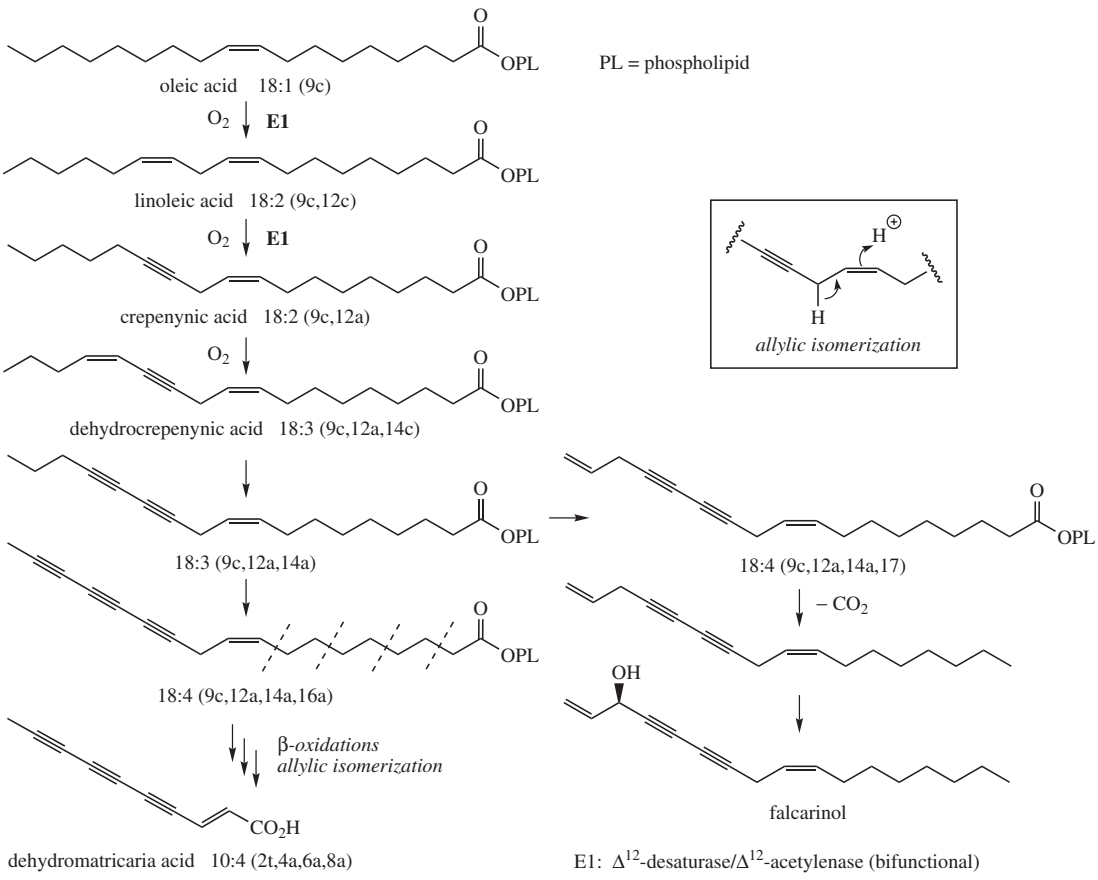


Figure 3.11

to increase the production of γ-linolenic and stearidonic acids in hosts such as soybeans.

In the vast majority of unsaturated fatty acids, the double bonds have the *Z*/*cis* configuration. Fatty acids with *trans* double bonds do occur naturally at relatively low levels in meat and dairy products as a by-product of fermentation in ruminant animals. However, *trans*-unsaturation may be introduced during the partial hydrogenation of polyunsaturated fats that is commonly practised during food processing to produce semi-solid fats from oils. There now seems to be good correlation between the consumption of *trans*-fats and the occurrence of coronary heart disease and atherosclerosis, though precise mechanisms are yet to be determined. Nevertheless, levels of *trans*-fats in foodstuffs are now closely monitored.

Many unsaturated compounds found in nature contain one or more acetylenic bonds, and these are predominantly produced by further desaturation of olefinic systems in fatty acid-derived molecules. They are surprisingly widespread in nature and are found in many organisms, but they are especially common in plants of the Compositae/Asteraceae, the Umbelliferae/Apiaceae, and fungi of the group Basidiomycetes. These compounds tend to be highly unstable and some are even explosive if sufficient amounts are accumulated. Since only very small amounts are present in plants, this does not present any widespread hazard. Whilst fatty acids containing several double bonds usually have these in a non-conjugated array, molecules containing triple bonds tend to possess conjugated unsaturation. This gives the compounds intense and highly characteristic UV spectra, which aids their detection and isolation.

The processes of desaturation are exemplified in Figure 3.11, in which oleic acid (probably bound as a phospholipid) features as a precursor of **crepenynic acid** and **dehydrocrepenynic acid**. A bifunctional desaturase/acetylenase system catalysing these two steps has been characterized in the plant *Crepis alpina* (Compositae/Asteraceae). The acetylenic bond is now indicated by 'a' in the semi-systematic shorthand nomenclature. Chain shortening by β-oxidation (see page 18) is often a feature of these pathways, and formation of the C_{10} acetylenic acid **dehydromatricaria acid** proceeds through C_{18} intermediates, losing eight carbon atoms, presumably via four β-oxidations. In the latter part of the pathway, the *Z* double bond from oleic acid moves into conjugation with the polyacetylene chain via an allylic isomerization, giving the more favoured *E*-configuration. Loss of the carboxyl function is also encountered, and **falcarinol** is an example of such structures, most likely derived via dehydrocrepenynic acid. Some noteworthy acetylenic structures (though they are no longer acids and components of fats) are given in Figure 3.12. **Cicutoxin** from the water hemlock (*Cicuta virosa*; Umbelliferae/Apiaceae) and **oenanthotoxin** from the hemlock water dropwort (*Oenanthe crocata*; Umbelliferae/Apiaceae) are extremely toxic to mammals, causing persistent vomiting and convulsions, leading to respiratory paralysis. Ingestion of the roots of these plants may frequently lead to fatal poisoning. **Falcarinol** is a constituent of *Falcaria vulgaris* (Umbelliferae/Apiaceae), *Oenanthe crocata, Hedera helix* (Araliaceae), and several other plants, and is known to cause contact dermatitis in certain individuals when the plants are handled. Falcarinol (sometimes called panaxynol) and the structurally related **panaxytriol** are also characteristic polyacetylene components of ginseng (*Panax ginseng*; Araliaceae; see page 245). **Wyerone** from the broad bean (*Vicia faba*; Leguminosae/Fabaceae) has antifungal properties, and its structure exemplifies how the original straight chain may be cross-linked to produce a ring system. The furan ring is believed to originate from a conjugated diyne.

Primary amides of unsaturated fatty acids have been characterized in humans and other mammals, and although their biological role is not fully understood, they may represent a group of important signalling molecules.

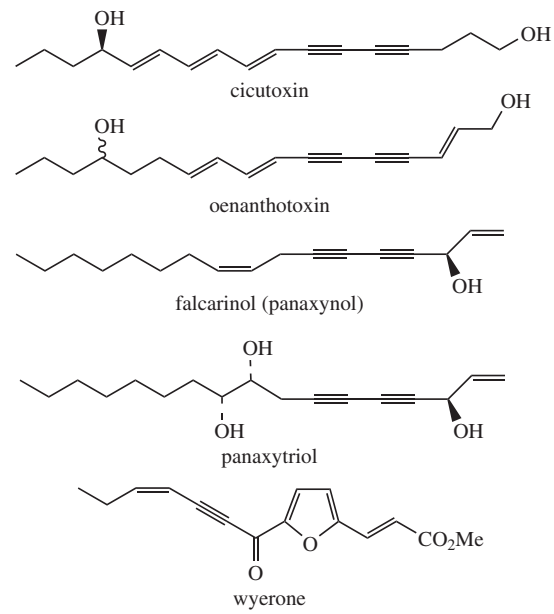


Figure 3.12

Oleamide, the simple amide of oleic acid, has been shown to be a sleep-inducing lipid, and the amide of erucic acid, **erucamide**, stimulates the growth of blood vessels. The ethanolamide of arachidonic acid, **anandamide**, appears to be the natural ligand for receptors to which cannabinoids bind; this and related structures will be considered later (see page 122). The herbal preparation echinacea [Box 3.1] is derived from the roots of *Echinacea purpurea* (Compositae/Asteraceae) and is used for its immunostimulant properties, particularly as a prophylactic and treatment for the common cold. At least some of its activity arises from a series of alkamides (also termed alkylamides), amides of polyunsaturated acids with isobutylamine. These acids are predominantly C_{11} and C_{12} diene-diynes (Figure 3.13).

Box 3.1

Echinacea

Echinacea consists of the dried roots of *Echinacea purpurea, Echinacea angustifolia*, or *Echinacea pallida* (Compositae/Asteraceae), herbaceous perennial plants indigenous to North America, and widely cultivated for their large daisy-like flowers, which are usually purple or pink. Herbal preparations containing the dried root, or extracts derived from it, are hugely popular, being promoted as immunostimulants, particularly as prophylactics and treatments for bacterial and viral infections, e.g. the common cold. Tests have validated stimulation of the immune response, though the origins of this activity cannot be ascribed to any specific substance. Activity has variously been assigned to lipophilic alkamides, polar caffeic acid derivatives, high molecular weight polysaccharide material, or to a combination of these. Compounds in each group have been demonstrated to possess some pertinent activity, e.g. immunostimulatory, anti-inflammatory, antibacterial, or antiviral effects.

The alkamides comprise a complex mixture of unsaturated fatty acids as amides with 2-methylpropanamine (isobutylamine) or 2-methylbutanamine, amines which are probably decarboxylation products from valine and isoleucine respectively. The acid portions are predominantly C_{11} and C_{12} diene-diynes or tetraenes (Figure 3.13). These compounds are found throughout the plant, though relative proportions of individual components vary considerably. The root of *Echinacea purpurea* contains at least 12 alkamides (about 0.6%), of which C_{12} diene-diynes predominate; levels of these compounds fall significantly during drying and storage. Caffeic acid derivatives present include caffeic acid (see page 149), chlorogenic acid (5-*O*-caffeoylquinic acid, see page 150), caftaric acid (2-*O*-caffeoyltartaric acid), and cichoric acid (2,3-di-*O*-caffeoyltartaric acid) (Figure 3.13). Cichoric acid is a major component (0.6–2.1%) in *Echinacea purpurea*, but only minor in the other species.

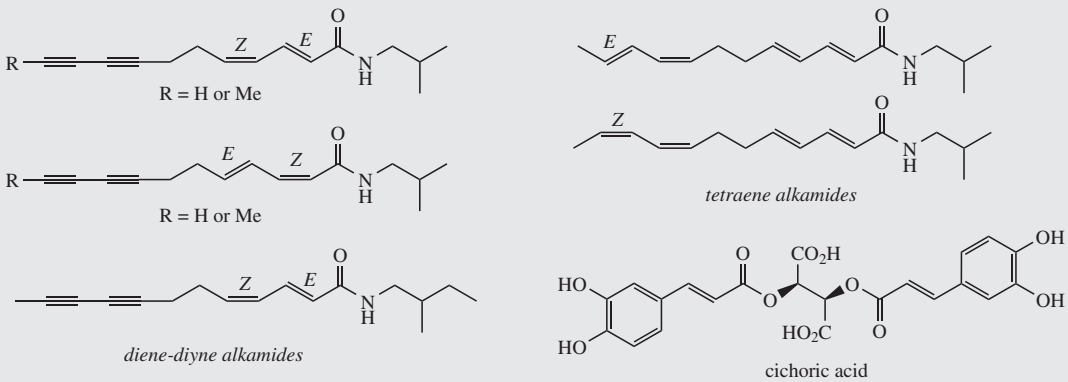


Figure 3.13

UNCOMMON FATTY ACIDS

Most fatty acids are undoubtedly primary metabolites. However, a considerable amount of structural diversity is encountered in this group of compounds, with some structures having a rather limited natural distribution, so it is more appropriate to think of these as secondary metabolites. The distinction is actually unnecessary, since it is impossible to consider one group without the other; this is a typical 'grey' area. Some structures arise by further modification of the basic straight-chain systems

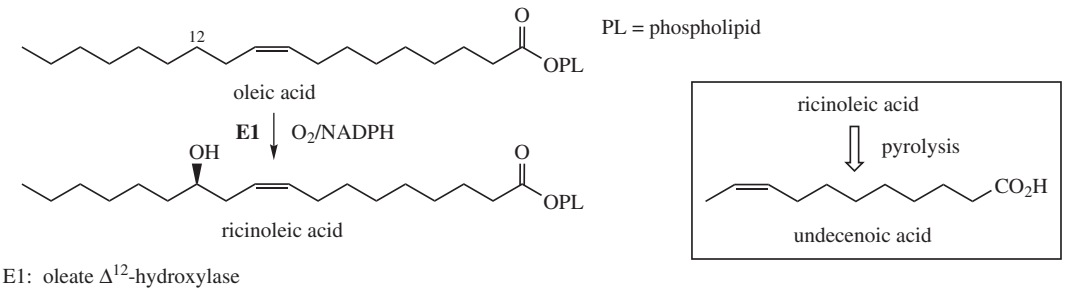


Figure 3.14

already discussed, whilst others require a fundamental change in the nature of the starter and extender units employed in the biosynthesis.

Ricinoleic acid (Figure 3.14) is the 12-hydroxy derivative of oleic acid and is the major fatty acid found in castor oil, expressed from seeds of the castor oil plant (*Ricinus communis*; Euphorbiaceae). It is formed by direct hydroxylation of oleic acid, esterified as part of a phospholipid, by the action of an O_2- and NADPH-dependent mixed-function oxidase. This is not of the cytochrome P-450 type, but structurally and mechanistically resembles the fatty acid desaturase enzymes. Castor oil has a long history of use as a domestic purgative, but it is now mainly employed as a cream base. **Undecenoic acid** (Δ^9-undecenoic acid) can be obtained from ricinoleic acid by thermal degradation, and as the zinc salt or in ester form, it is used in a number of fungistatic preparations. Fatty acid hydroxylases of the cytochrome P-450 type have also been identified; the position of hydroxylation can vary according to enzyme, though hydroxylation of the terminal methyl (ω-hydroxylation) is quite common.

Epoxy fatty acids like **vernolic acid** (Figure 3.15) have been found in substantial quantities in the seed oil of some plant species, including *Vernonia galamensis* and *Stokesia laevis* (both Compositae/Asteraceae). Vernolic acid is formed by direct epoxidation of linoleic acid (as a phospholipid ester), but, as with hydroxylation, different types of enzyme have been identified. These plants also exploit desaturase-like systems, whereas the same transformation in *Euphorbia lagascae* (Euphorbiaceae) is catalysed by a cytochrome P-450-dependent enzyme.

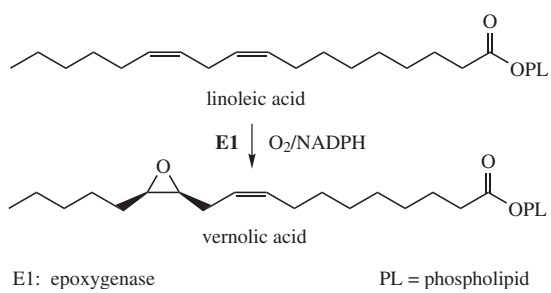


Figure 3.15

Whilst straight-chain fatty acids are the most common, branched-chain acids have been found to occur in mammalian systems, e.g. in wool fat and butter fat. They are also characteristic fatty acid constituents of the lipid part of cell walls in some pathogenic bacteria. Several mechanisms appear to operate in their formation. Methyl side-chains can be introduced by a *C*-alkylation mechanism using SAM. **Tuberculostearic acid** (Figure 3.16) is a *C*-methyl derivative of stearic acid found in the cell wall of *Mycobacterium tuberculosis*, the bacterium causing tuberculosis, and it provides a diagnostic marker for the disease. It is derived from oleic acid by alkylation on C-10, initiated by the double-bond electrons. Methyl transfer is followed by a Wagner–Meerwein 1,2-hydride shift in the carbocation, then proton loss to give the 10-methylene derivative. Finally, the double bond is reduced in an NADPH-dependent reaction to give tuberculostearic acid. This sequence of transformations is also seen during *C*-methylation of sterol side-chains (see page 252). Alternatively, loss of a proton from the first-formed carbocation intermediate via cyclopropane ring formation leads to **dihydrosterculic acid**. This is known to be dehydrogenated to **sterculic acid**, an unusual fatty acid containing a highly strained cyclopropene ring. Sterculic acid is present in the seed oil from *Sterculia foetida* (Sterculiaceae), and with similar cyclopropene acids, e.g. malvalic acid, is present in edible cottonseed oil from *Gossypium* species (Malvaceae). **Malvalic acid** is produced from sterculic acid by chain shortening from the carboxyl end (Figure 3.16). Sterculic acid is an inhibitor of the Δ^9-desaturase which converts stearic acid into oleic

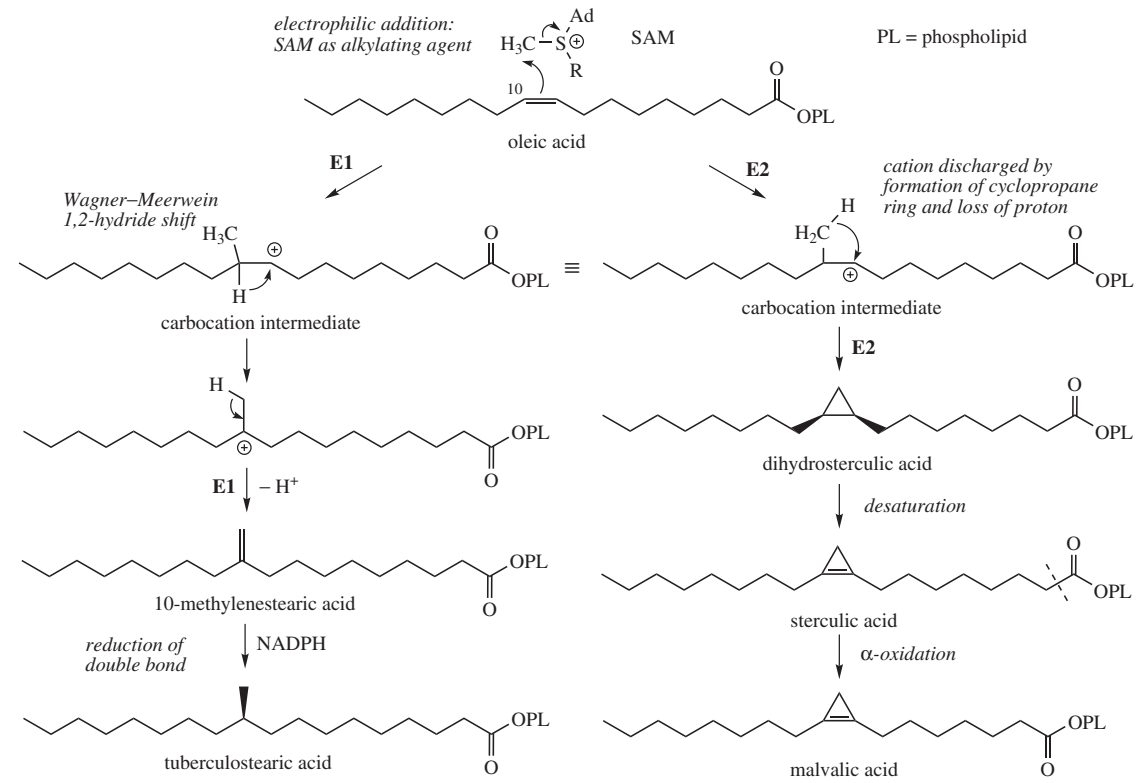


Figure 3.16

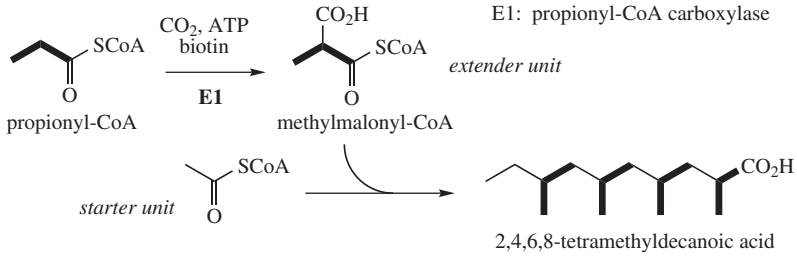


Figure 3.17

acid and is potentially harmful to humans in that it can alter membrane permeability and inhibit reproduction. Seed oils containing cyclopropene fatty acids destined for human consumption require suitable treatment to remove these undesirables.

Methyl side-chains can also be introduced by using methylmalonyl-CoA instead of malonyl-CoA as the chain extending agent (Figure 3.17). Methylmalonyl-CoA arises by biotin-dependent carboxylation of propionyl-CoA in exactly the same way as malonyl-CoA was formed (see page 17). **2,4,6,8-Tetramethyldecanoic acid** found in the preen gland wax of the goose (*Anser anser*) is produced from an acetyl-CoA starter and four methylmalonyl-CoA chain extender units. This introduces the concept of

changing the starter and/or extender units whilst retaining the same FAS type of mechanism. It is a concept that gets developed much further with macrolide antibiotic structures (see page 68). Various bacteria contain a range of **iso-fatty acids** that help to control membrane fluidity. These iso-fatty acids are produced by malonate chain extension of a number of different starter units derived from branched-chain amino acids, modified into CoA-esters. Thus, leucine, valine, and isoleucine may generate isovaleryl-CoA, isobutyryl-CoA, and 2-methylbutyryl-CoA respectively as starter groups (Figure 3.18). Several other examples where these amino acids provide building blocks will be met in due course.

Chaulmoogric and **hydnocarpic acids** (Figure 3.19) are uncommon cyclopentenyl fatty acids found in chaulmoogra oil, an oil expressed from seeds of *Hydnocarpus wightiana* (Flacourtiaceae). These acids are known to arise by malonate chain extension of the coenzyme A ester of 2-cyclopentenyl carboxylic acid as an alternative starter unit to acetate. Chaulmoogra oil provided for many years the only treatment for the relief of leprosy, these two acids being strongly bactericidal towards the leprosy infective agent *Mycobacterium leprae*. Purified salts and esters of hydnocarpic and chaulmoogric acids were subsequently employed, until they were then themselves replaced by more effective synthetic agents.

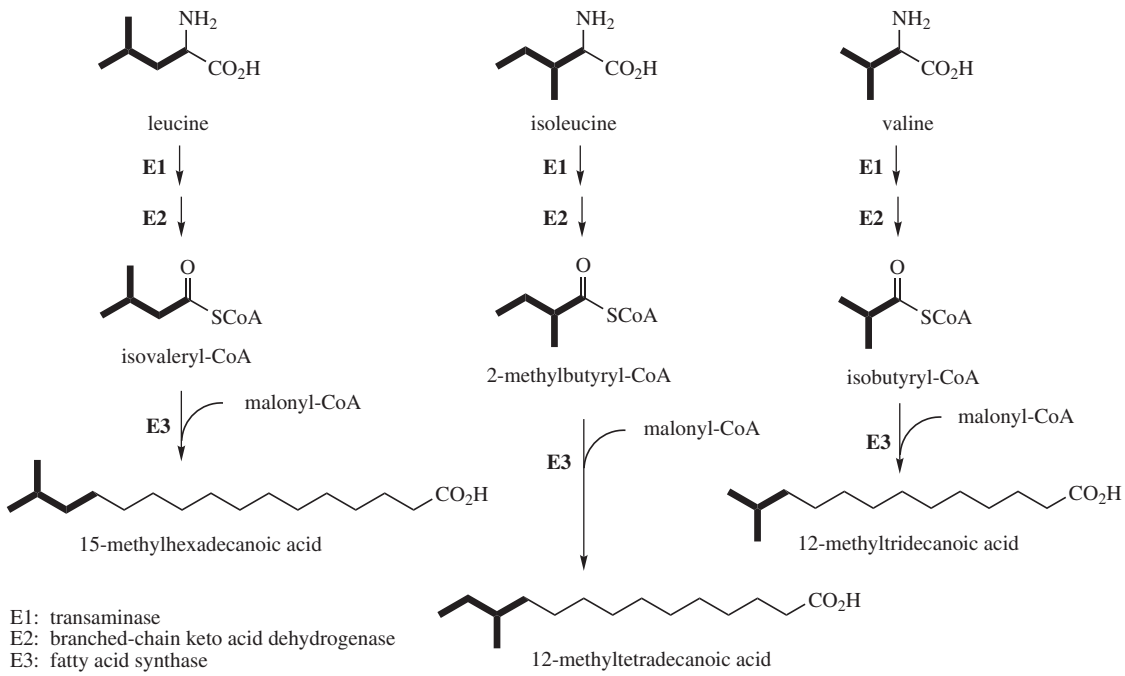


Figure 3.18

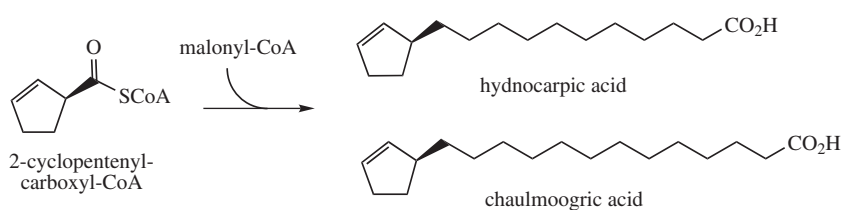


Figure 3.19

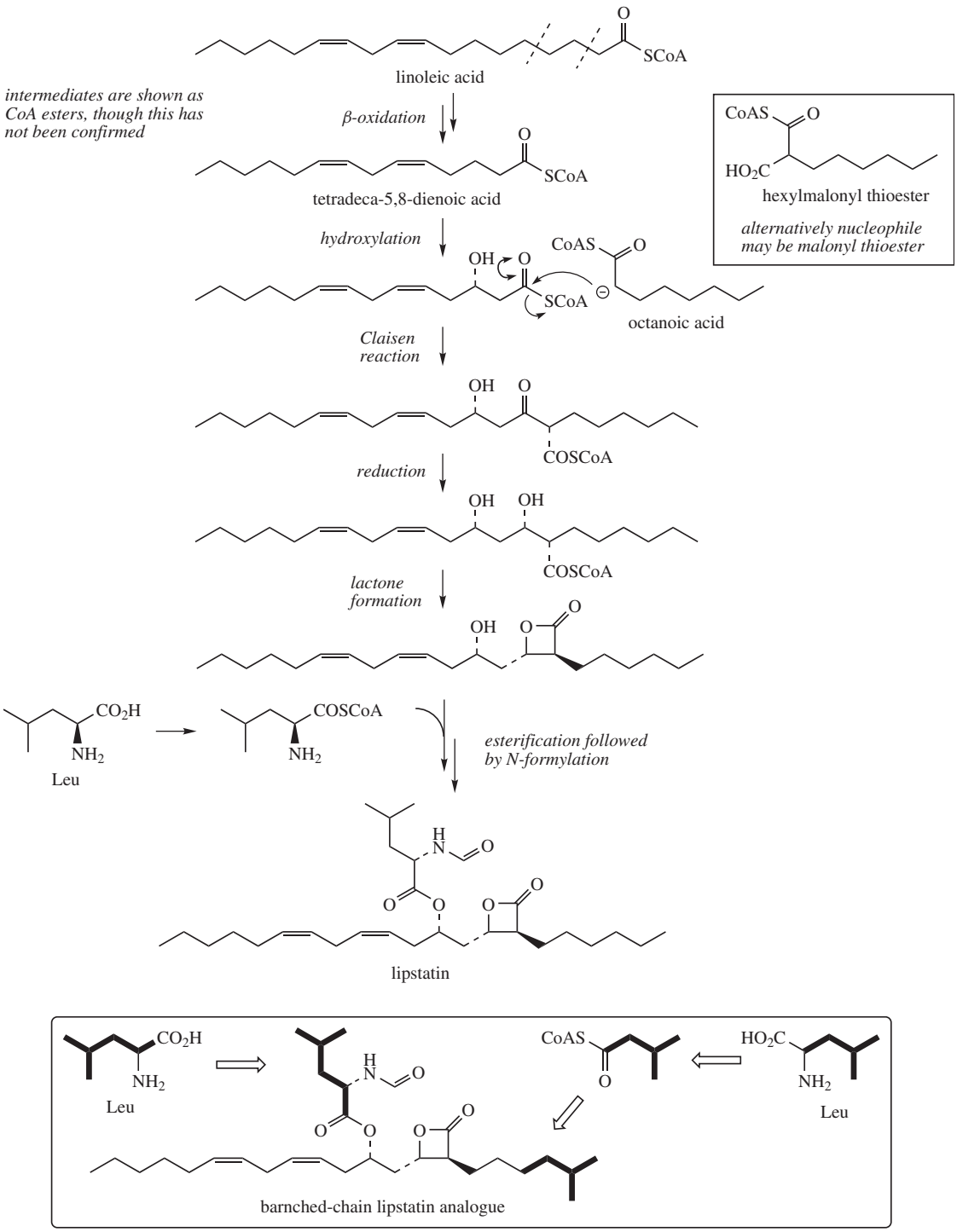

linoleic acid
SCoA
intermediates are shown as CoA esters, though this has not been confirmed
β-oxidation
tetradeca-5,8-dienoic acid
hydroxylation
CoAS
OH
SCoA
octanoic acid
Claisen reaction
COSCoA
reduction
lactone formation
CO2H
NH2
Leu
esterification followed by N-formylation
lipstatin
HO2C
hexylmalonyl thioester
alternatively nucleophile may be malonyl thioester
barnched-chain lipstatin analogue

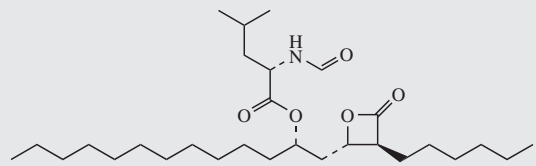
Figure 3.20

Some branched-chain systems can be rationalized as a combination of two separate fatty acid chains coupled in a Claisen reaction. This has been shown for the β-lactone derivative **lipstatin** found in *Streptomyces toxytricini*, a compound that generated considerable interest because of its ability to inhibit pancreatic lipase; it has since been developed into an anti-obesity drug [Box 3.2]. Lipstatin is formed from the two fatty acids tetradeca-5,8-dienoic acid and octanoic acid; the first of these is believed to originate from linoleic acid derived from sunflower oil in the culture medium (Figure 3.20). There is evidence that the nucleophilic species is more likely to be the activated hexylmalonyl thioester rather than the octanoyl thioester. Claisen coupling is followed by simple reduction and lactonization (esterification). The remaining portion of the molecule is leucine derived, and esterification appears to precede introduction of the *N*-formyl group; the *N*-formyl group is presumably an oxidized *N*-methyl. A branched-chain analogue of lipstatin is a minor metabolite in the cultures (Figure 3.20). This incorporates two molecules of leucine, one of which acts as a starter unit for the methyloctanoate fatty acid portion, most likely via isovaleryl-CoA as seen above.

Box 3.2

Lipstatin

Lipstatin was isolated from the mycelium of *Streptomyces toxytricini* cultures and shown to possess marked inhibitory activity towards pancreatic lipase, the key enzyme for intestinal fat digestion. The β-lactone function in the lipophilic molecule irreversibly inactivates lipase by covalent reaction with a serine residue at the catalytic site; the reaction closely parallels that of serine residues with β-lactam antibiotics (see page 464). Tetrahydrolipstatin (**orlistat**) (Figure 3.21), obtained by catalytic hydrogenation of lipstatin, was selected for further development; it is more stable and crystallizes readily, although it is somewhat less active. Orlistat is now manufactured by total synthesis.



orlistat (tetrahydrolipstatin)

Figure 3.21

Orlistat reduces the absorption of dietary fat and is used in conjunction with a low-fat calorie-reduced diet to reduce body mass in obese patients. Only trace amounts of the drug are absorbed systemically, and the primary effect is local lipase inhibition within the digestive tract; dietary fat is thus excreted undigested. Absorption of fat-soluble vitamins, especially vitamin D, is also inhibited, and vitamin supplements are usually co-administered.

PROSTAGLANDINS

The prostaglandins are a group of modified C_{20} fatty acids first isolated from human semen and initially assumed to be secreted by the prostate gland [Box 3.3]. They are now known to occur widely in animal tissues, but only in tiny amounts, and they have been found to exert a wide variety of pharmacological effects on humans and animals. They are active at very low, hormone-like concentrations and can regulate blood pressure, contractions of smooth muscle, gastric secretion, and platelet aggregation. Their potential for drug use is extremely high, but it has proved difficult to separate the various biological activities into individual agents.

The basic prostaglandin skeleton is that of a cyclized C_{20} fatty acid containing a cyclopentane ring, a C_7 side-chain with the carboxyl function, and a C_8 side-chain with the methyl terminus. Prostaglandins

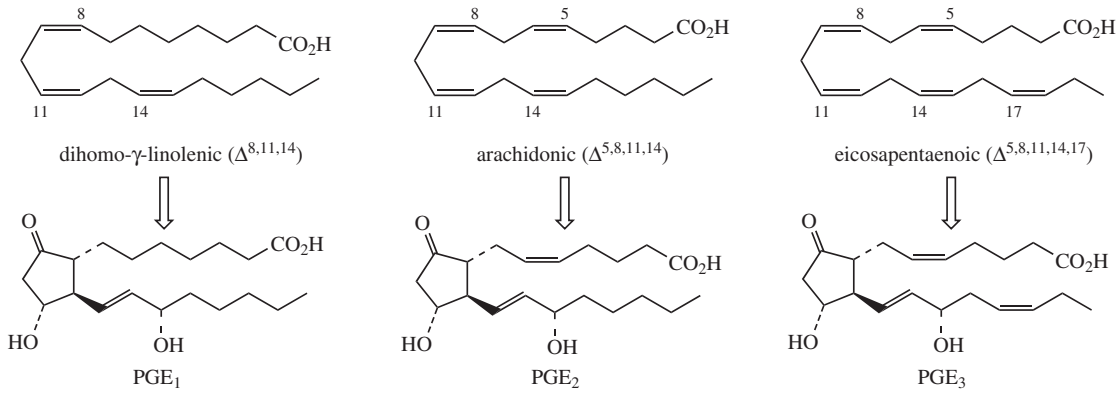


Figure 3.22

are biosynthesized from three EFAs, namely **$\Delta^{8,11,14}$-eicosatrienoic acid** (**dihomo-γ-linolenic acid**), **$\Delta^{5,8,11,14}$-eicosatetraenoic acid** (**arachidonic acid**), and **$\Delta^{5,8,11,14,17}$-eicosapentaenoic acid**, which yield prostaglandins of the 1-, 2-, and 3-series respectively (Figure 3.22; see Figure 3.24 below for principles of nomenclature). The three precursors lead to products of similar structure, but with varying levels of unsaturation in the two side-chains. Some of the structures elaborated from arachidonic acid are shown in Figure 3.23.

In the first reaction, arachidonic acid is converted into **prostaglandin H_2** (**PGH_2**) by way of **prostaglandin G_2** (**PGG_2**) by prostaglandin H synthase. This is a bifunctional enzyme comprised of an oxygenase (**cyclooxygenase, COX**) and a peroxidase, both requiring haem as cofactor. COX incorporates two molecules of oxygen, liberating a compound with both cyclic and acyclic peroxide functions, and also creating five chiral centres. In arachidonic acid, the methylene group flanked by two double bonds is susceptible to radical oxidation, initiated by a tyrosyl radical in the enzyme. This leads to incorporation of oxygen and formation of a cyclic peroxide function. Radical addition to an alkene function allows ring formation; then, a further incorporation of oxygen, in what is effectively a repeat of the earlier sequence, leads to an acyclic peroxide and produces PGG_2. The hydrogen atom terminating the synthesis is obtained from the same tyrosine function in the enzyme, thus propagating the radical reaction. The acyclic peroxide group in PGG_2 is then cleaved by the peroxidase component of prostaglandin H synthase to yield **PGH_2**. This unstable peroxide occupies a central role and can be modified in several different ways. Reductive cleavage gives **prostaglandin $F_{2\alpha}$** (**$PGF_{2\alpha}$**); other modifications can be rationally accommodated by initial cleavage of the cyclic peroxide to the diradical, though alternative ionic mechanisms may also be proposed. Quenching of the radicals by capture and loss of hydrogen atoms would provide either **prostaglandin E_2** (**PGE_2**) or **prostaglandin D_2** (**PGD_2**). Reduction of PGE_2 is known to provide a minor pathway to $PGF_{2\alpha}$. The bicyclic system in **prostaglandin I_2** (**PGI_2**; **prostacyclin**) is envisaged as arising by involvement of a side-chain double bond, then loss of a hydrogen atom, all catalysed by prostacyclin synthase, a cytochrome P-450-dependent enzyme. Prostaglandin structures representative of the 1-series, e.g. **PGE_1**, or of the 3-series, e.g. **PGE_3**, can be formed in a similar way from the appropriate fatty acid precursor (Figure 3.22).

The basic skeleton of the prostaglandins is termed **prostanoic acid**, and derivatives of this system are collectively known as prostanoids. The term **eicosanoids** is also used to encompass prostaglandins, thromboxanes, and leukotrienes, which are all derived from C_{20} fatty acids (eicosanoic acids). Semi-systematic nomenclature of prostaglandins is based on the substitution pattern in the five-membered ring, denoted by a letter suffix (Figure 3.24), and the number of double bonds in the side-chains is given by a numerical subscript. Greek letters α and β are used to indicate the configuration at C-9, α indicating that the substituent is below the plane (as found in natural prostaglandins) and β indicating that the substituent is above the plane (as in some synthetic analogues). 'Prostaglandin' is usually abbreviated to PG. Prostaglandins A, B, and C are inactive degradation products from natural prostaglandins.

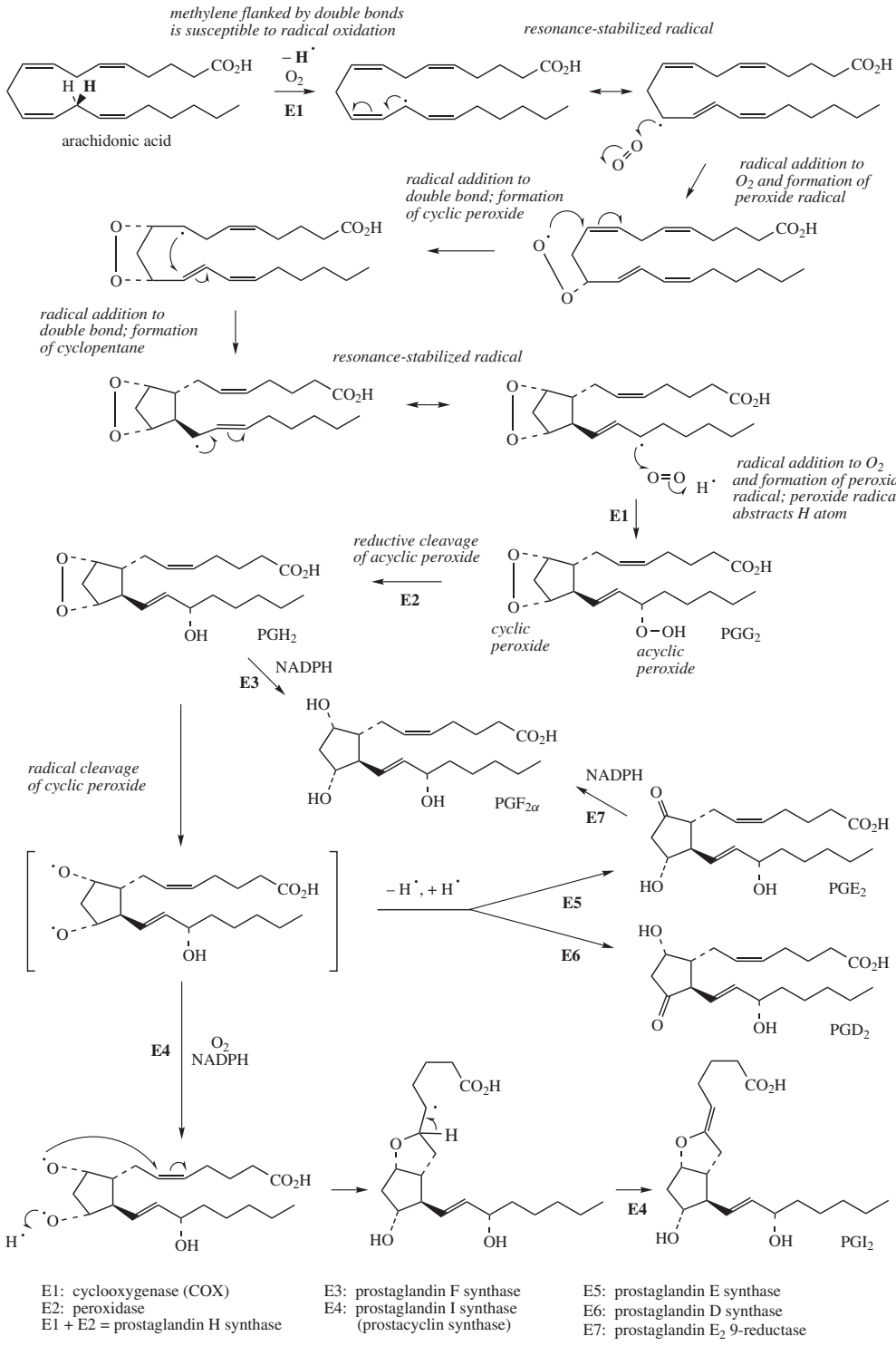

methylene flanked by double bonds is susceptible to radical oxidation
resonance-stabilized radical
arachidonic acid
E1
radical addition to double bond; formation of cyclic peroxide
radical addition to O2 and formation of peroxide radical
radical addition to double bond; formation of cyclopentane
resonance-stabilized radical
radical addition to O2 and formation of peroxide radical; peroxide radical abstracts H atom
reductive cleavage of acyclic peroxide
E2
PGH2
cyclic peroxide
acyclic peroxide
PGG2
NADPH
E3
PGF2α
radical cleavage of cyclic peroxide
E7
PGE2
E5
E6
PGD2
E4
PGI2
E1: cyclooxygenase (COX)
E2: peroxidase
E1 + E2 = prostaglandin H synthase
E3: prostaglandin F synthase
E4: prostaglandin I synthase (prostacyclin synthase)
E5: prostaglandin E synthase
E6: prostaglandin D synthase
E7: prostaglandin E2 9-reductase

Figure 3.23

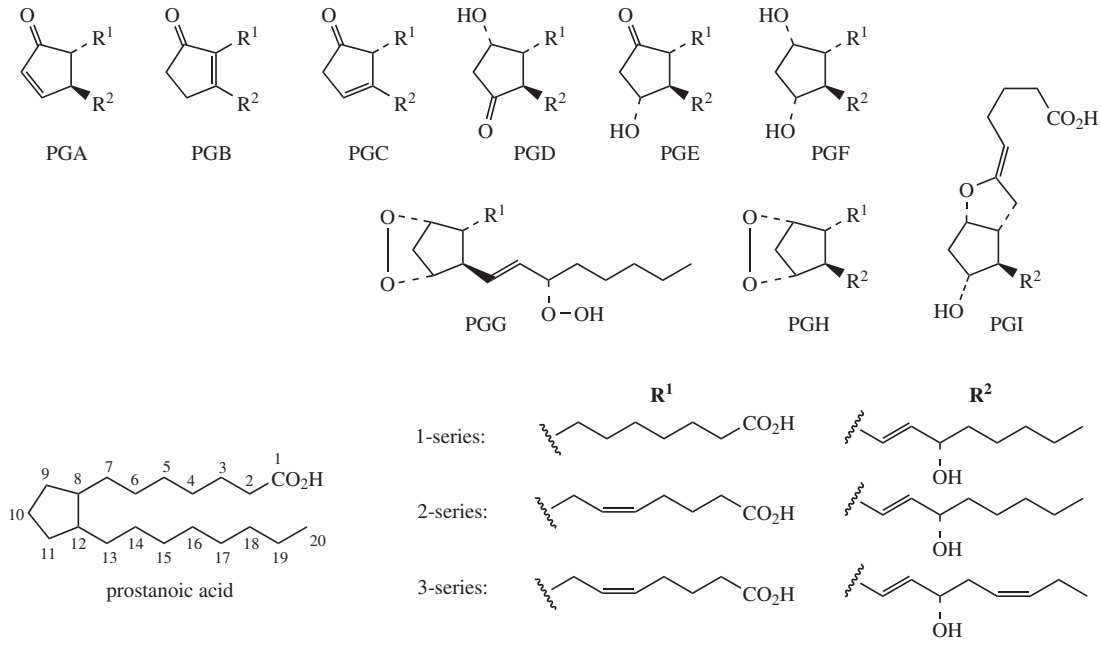


Figure 3.24

Box 3.3

Prostaglandins

Prostaglandins occur in nearly all mammalian tissues, but only at very low concentrations. PGE_1 and $PGF_{1\alpha}$ were initially isolated from sheep seminal plasma, but these compounds and PGD_2, PGE_2, and $PGF_{2\alpha}$ are widely distributed. Animal sources cannot supply sufficient amounts for drug usage. The soft coral *Plexaura homomalla* (sea whip) from the Caribbean has been identified as having very high (2–3%) levels of prostaglandin esters, predominantly the methyl ester of PGA_2 (1–2%) with related structures. Prostaglandins of the A-, E-, and F-types are widely distributed in soft corals, especially *Plexaura*, though many corals produce prostaglandin structures not found in animals. For several years, the sea whip coral was used as a source of research material, but this was neither a satisfactory nor renewable natural source. Considerable effort was exerted on the total synthesis of prostaglandins and their interconversions, and the high level of success achieved opened up the availability of compounds for pharmacological testing and subsequent drug use. Synthetic analogues have also been developed to modify or optimize biological activity. The studies have demonstrated that biological activity is effectively confined to the natural enantiomers; the unnatural enantiomer of PGE_1 had only 0.1% of the activity of the natural isomer.

The prostaglandins display a wide range of pharmacological activities, including contraction and relaxation of smooth muscle of the uterus, the cardiovascular system, the intestinal tract, and of bronchial tissue. They also inhibit gastric acid secretion, control blood pressure and suppress blood platelet aggregation, as well as acting as mediators of inflammation, fever, and allergy. Some of these effects are consistent with the prostaglandins acting as second messengers, modulating transmission of hormone stimulation and, thus, metabolic response. They are synthesized in response to various stimuli in a variety of cells, released immediately after synthesis, and act in the vicinity of their synthesis to maintain local homeostasis. Some prostaglandins in the A and J series have demonstrated potent antitumour properties.

Since the prostaglandins control many important physiological processes in animal tissues, their drug potential is high, but the chances of precipitating unwanted side-effects are also high, and this has so far limited their therapeutic use. There is, however, much additional scope for controlling the production of natural prostaglandins in body tissues by means of specific inhibitors. Indeed, it has been found that some established non-steroidal anti-inflammatory drugs (NSAIDs), e.g. aspirin, indometacin, and

Box 3.3 (continued)

ibuprofen, inhibit early steps in the prostaglandin biosynthetic pathway that transform the unsaturated fatty acids into cyclic peroxides. The local production of prostaglandins such as PGE_2 can sensitize pain nerve endings and increase blood flow, promoting pain, swelling, and redness. On the other hand, PGI_2 and PGE_2 are protective to the stomach, and inhibiting their formation explains the gastrointestinal toxicity associated with prolonged and high-dose use of NSAIDs. Aspirin is now known to inactivate irreversibly the COX activity (arachidonic acid $\rightarrow$ PGG_2), though not the peroxidase activity ($PGG_2 \rightarrow PGH_2$), by selective acetylation of a serine residue of the enzyme; ibuprofen and indometacin compete with arachidonic acid at the active site and are reversible inhibitors of COX. A more recent discovery is that two forms of the COX enzyme exist, designated COX-1 and COX-2. COX-1 is expressed constitutively in most tissues and cells and is thought to control synthesis of those prostaglandins important for normal cellular functions, such as gastrointestinal integrity and vascular homeostasis. In simplistic terms, COX-1 is considered a 'housekeeping' enzyme. COX-2 is not normally present, but is inducible in certain cells in response to inflammatory stimuli, resulting in enhanced prostaglandin release in the central nervous system and inflammatory cells with the characteristic inflammatory response. Its inhibition appears to produce the analgesic, antipyretic, and anti-inflammatory effects of NSAIDs. However, popular NSAIDs do not discriminate between the two COX enzymes, and so this leads to both therapeutic effects via inhibition of COX-2 and adverse effects such as gastrointestinal problems, ulcers, and bleeding via inhibition of COX-1. Because of differences in the nature of the active sites of the two enzymes, it has now been possible to develop agents that can inhibit COX-2 rather than COX-1 as potential new anti-inflammatory drugs. The first of these 'coxib' drugs, celecoxib and rofecoxib, were introduced for relief of pain and inflammation in osteoarthritis and rheumatoid arthritis, though rofecoxib was subsequently withdrawn because of significant cardiovascular side-effects. Second-generation agents etoricoxib and lumiracoxib have since been introduced. The widely used analgesic paracetamol (US: acetaminophen) is usually classified as an NSAID, though it has little anti-inflammatory activity. Its ability to inhibit pain and fever is suggested to emanate from inhibition of a COX-1 variant, termed COX-3. The anti-inflammatory activity of corticosteroids correlates with inhibition of phospholipase enzymes and preventing the release of arachidonic acid from storage phospholipids, but expression of COX-2 is also inhibited by glucocorticoids.

The role of of essential fatty acids (see page 49) such as linoleic and γ-linolenic acids, obtained from plant ingredients in the diet, can now readily be appreciated. Without a source of arachidonic acid, or compounds which can be converted into arachidonic acid, synthesis of prostaglandins would be compromised, and this would seriously affect many normal metabolic processes. A steady supply of prostaglandin precursors is required; prostaglandins are continuously being synthesized and then degraded. Prostaglandins are rapidly degraded by processes which include oxidation of the 15-hydroxyl to a ketone, reduction of the 13,14-double bond, and oxidative degradation of both side-chains.

A major area of application of prostaglandins as drugs is in obstetrics, where they are used to induce abortions during the early to middle stages of pregnancy, or to induce labour at term. Dinoprost ($PGF_{2\alpha}$) is not usually prescribed, since it is rapidly metabolized in body tissues (half-life less than 10 min), and the modified version **carboprost** (15-methyl $PGF_{2\alpha}$; Figure 3.25) has been developed to reduce deactivation by blocking oxidation at position 15. Carboprost is produced by oxidizing the 15-hydroxyl in a suitably protected $PGF_{2\alpha}$, then alkylating the 15-carbonyl with a Grignard reagent. Carboprost is effective at much reduced dosage than dinoprost and is of value in augmenting labour at term, especially in cases where ergometrine (see page 394) or oxytocin (see page 430) are ineffective. The natural prostaglandin structure **dinoprostone** (PGE_2) is also used primarily to induce labour; the 16-dimethyl prostaglandin **gemeprost** is currently the preferred prostaglandin for abortions. These agents are usually administered vaginally.

Alprostadil (PGE_1) differs from PGE_2 by having unsaturation in only one side-chain. Though having effects on uterine muscle, it also has vasodilator properties, and these are exploited for maintaining new-born infants with congenital heart defects, facilitating blood oxygenation prior to corrective surgery. The very rapid metabolism of PGE_1 means this drug must be delivered by continuous intravenous infusion. Alprostadil is also of value in male impotence, self-injectable preparations being used to achieve erection of the penis. An interesting modification to the structure of PGE_1 is found in the analogue **misoprostol**. This compound has had the oxygenation removed from position 15, transferred to position 16, plus alkylation at position 16 to reduce metabolism (compare 15-methyl $PGF_{2\alpha}$ above). These modifications result in an orally active drug which inhibits gastric secretion effectively and can be used to promote healing of gastric and duodenal ulcers. In combination with non-specific NSAIDs, it can significantly lower the incidence of gastrointestinal side-effects such as ulceration and bleeding. It also induces labour, and is occasionally used for this purpose.

Epoprostenol (PGI_2, prostacyclin) reduces blood pressure and also inhibits platelet aggregation by reducing calcium concentrations. It is employed to inhibit blood clotting during renal dialysis, but its very low half-life (about 3 min) again necessitates continuous intravenous administration. The tetrahydrofuran ring is part of an enol ether and is readily opened by hydration, leading to 6-ketoprostaglandin $F_{1\alpha}$ (Figure 3.26). **Iloprost** (Figure 3.25) is a stable carbocyclic analogue of potential use in the treatment of thrombotic diseases.

Box 3.3 (continued)

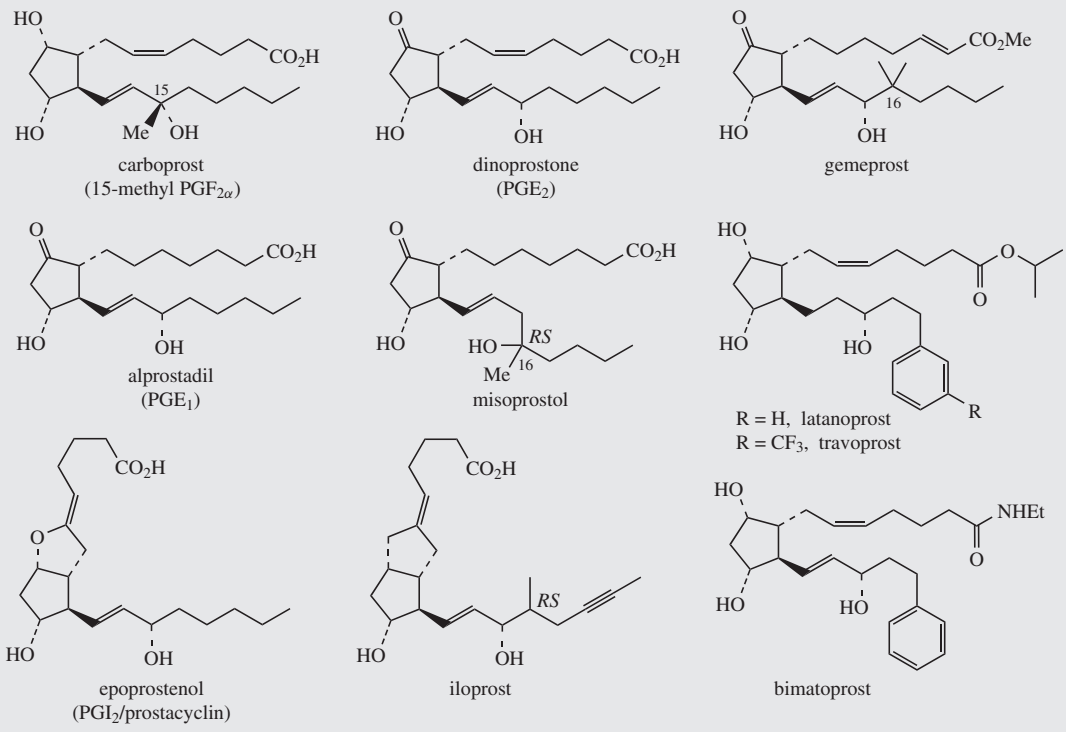


Figure 3.25

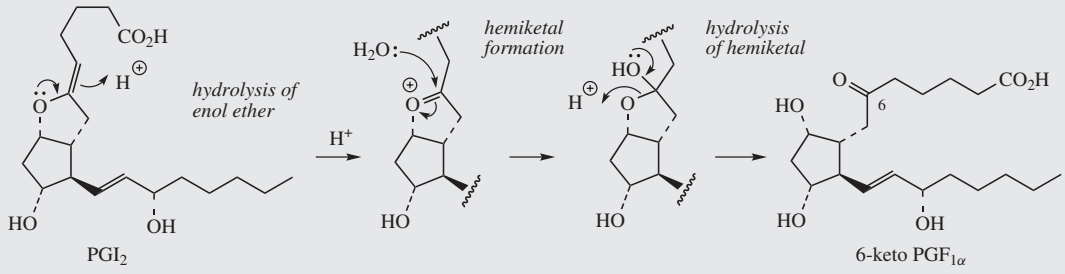


Figure 3.26

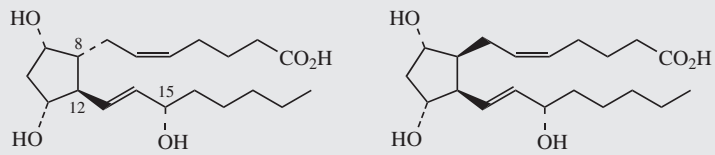
Latanoprost, travoprost, and **bimatoprost** (Figure 3.25) are recently introduced prostaglandin analogues which increase the outflow of aqueous humour from the eye. They are thus used to reduce intraocular pressure in the treatment of the eye disease glaucoma. Bimatoprost is an amide derivative related to prostaglandin amides (prostamides) obtained by the action of COX-2 on anandamide, the natural ligand for cannabinoid receptors (see page 122). The pharmacological properties of prostamides cannot readily be explained simply by interaction with prostaglandin receptors; bimatoprost also inhibits prostaglandin F synthase. An unusual side-effect of bimatoprost is also being exploited cosmetically: it makes eyelashes grow longer, thicker, and darker.

Box 3.3 (continued)

Isoprostanes

Isoprostanes represent a new class of prostaglandin-like compounds produced *in vivo* in humans and animals by non-enzymic radical-mediated oxidation of membrane-bound polyunsaturated fatty acids independent of the COX enzyme. An isomer of $PGF_{2\alpha}$ in which the two alkyl substituents on the five-membered ring were arranged *cis* rather than *trans* was detected in human urine and was the first of these compounds to be characterized. This compound was initially termed 8-*iso*-$PGF_{2\alpha}$, or 8-*epi*-$PGF_{2\alpha}$; however, as many more variants in the isoprostane series were discovered, it is now termed 15-F_{2t}-IsoP (Figure 3.27). The first number refers to the position of the hydroxyl, F_2 relates the compound to the prostaglandin class, and then t (*trans*) or c (*cis*) defines the relationship of the side-chain to the ring hydroxyls. The isoprostanes can be viewed as arising by a radical mechanism which resembles the enzyme-mediated formation of prostaglandins shown in Figure 3.23. Varying side-chain substituents arise by utilizing different double bonds from the several available in the cyclization mechanism and incorporating an oxygen atom from molecular oxygen at different positions. Many variants are formed because chemical processes rather than enzyme-controlled processes are employed. Isoprostanes derived similarly from docosahexaeneoic acid (DHA), a major lipid component of brain tissue, have also been isolated, and these are termed neuroprostanes. Structurally related compounds are also found in plants; these are derived from α-linolenic acid and are termed phytoprostanes (see page 204).

Interest in these isoprostanoid derivatives stems partly from the finding that certain compounds possess biological activity, probably via interaction with receptors for prostaglandins. For example, 15-F_{2t}-IsoP is a potent vasoconstrictor and also aggregates platelets. There is also evidence that prostaglandins, including PGE_2, PGD_2, and $PGF_{2\alpha}$, can also be formed via epimerization reactions on isoprostanes, i.e. prostaglandin synthesis by a chemical pathway independent of COX. Perhaps the most promising application relates to their origin via radical peroxidation of unsaturated fatty acids. Radicals are implicated in inflammatory and degenerative diseases such as atherosclerosis, cancer, and Alzheimer's disease. Isoprostane analysis of urine or serum provides non-invasive monitoring of oxidative damage as an insight into these disease states.



$PGF_{2\alpha}$ 15-F_{2t}-IsoP

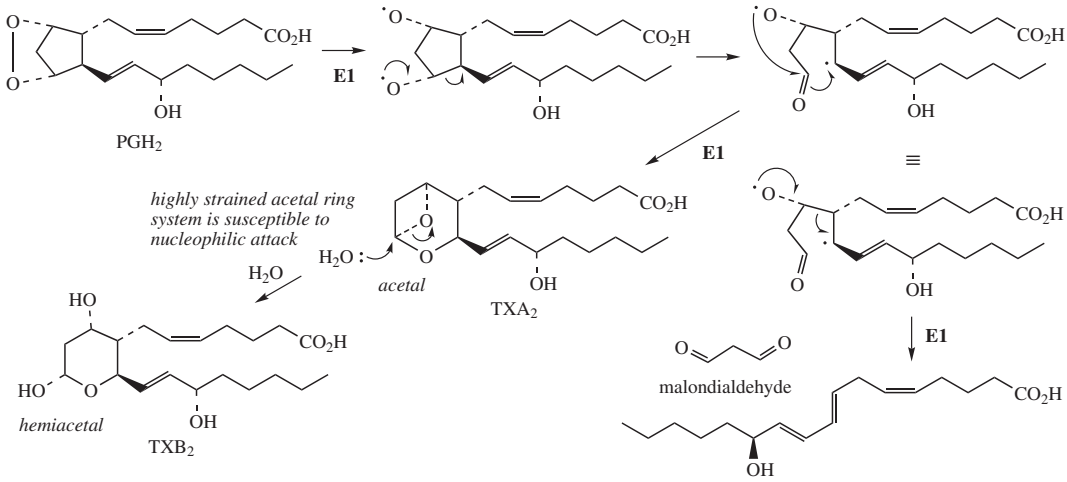
Figure 3.27

THROMBOXANES

An intriguing side-branch from the prostaglandin pathway leads to thromboxanes (Figure 3.28) [Box 3.4]. The peroxide and cyclopentane ring functions of PGH_2 are cleaved and restructured to form **thromboxane A_2** (**TXA_2**), which contains a highly strained four-membered oxetane ring. A hypothetical radical mechanism is shown; the enzyme also produces equimolar amounts of hydroxyheptadecatrienoic acid and malondialdehyde by a fragmentation reaction. TXA_2 is highly unstable and reacts readily with nucleophiles. In an aqueous environment, it reacts to yield the biologically inactive hemiacetal **thromboxane B_2** (**TXB_2**).

LEUKOTRIENES

Yet another variant for the metabolism of arachidonic acid is the formation of leukotrienes, a series of fatty acid derivatives with a conjugated triene functionality, and first isolated from leukocytes [Box 3.5]. In a representative pathway (others have been characterized) (Figure 3.29), **arachidonic acid** is converted into a hydroperoxide, the point of oxygenation being C-5, rather than C-11 as in the prostaglandin pathway (Figure 3.23). This compound loses water via formation of an epoxide ring, giving **leukotriene A_4** (**LTA_4**). This unstable allylic epoxide may hydrolyse by conjugate addition to give **leukotriene B_4** (**LTB_4**), or alternatively the epoxide may be



E1: thromboxane A_2 synthase

Figure 3.28

Box 3.4

Thromboxanes

The thromboxanes were isolated from blood platelets, and whilst TXA_2 showed high biological activity, TXB_2 was effectively inactive. TXA_2 is a potent stimulator of vasoconstriction and platelet aggregation. Aggregation of blood platelets to form a clot or thrombus is caused by increasing cytoplasmic calcium concentrations and thus deforming the platelets, which then fuse together. TXA_2 has the opposite effect to PGI_2, and presumably the development of thrombosis reflects an imbalance in the two activities. Both compounds are produced from the same precursor, PGH_2, which is converted in the blood platelets into TXA_2, and in the blood vessel wall into PGI_2. Thromboxanes A_3 and B_3 have also been isolated from blood platelets. These are derived from $\Delta^{5,8,11,14,17}$-eicosapentaenoic acid and relate structurally to prostaglandins in the 3-series. TXA_3 is not strongly aggregatory towards blood platelets. The highly unstable nature of the biologically active thromboxanes has made their synthesis difficult, and drug use of natural structures will probably be impracticable. It is likely that most efforts will be directed towards thromboxane antagonists to help reduce blood platelet aggregation in thrombosis patients. The value of aspirin in preventing cardiovascular disease is now known to be related to inhibition of thromboxane A_2 biosynthesis in platelets.

Box 3.5

Leukotrienes

The leukotrienes are involved in allergic responses and inflammatory processes. An antigen–antibody reaction can result in the release of compounds such as histamine (see page 398) or materials termed slow reacting substance of anaphylaxis (SRSA). These substances are then mediators of hypersensitive reactions such as hay fever and asthma. Structural studies have identified SRSA as a mixture of LTC_4, LTD_4, and LTE_4. These cysteine-containing leukotrienes are powerful bronchoconstrictors and vasoconstrictors, and induce mucus secretion, the typical symptoms of asthma. LTE_4 is some 10–100-fold less active than LTD_4, so that degradation of the peptide side-chain represents a means of eliminating leukotriene function. LTB_4 appears to facilitate migration of leukocytes in inflammation, and is implicated in the pathology of psoriasis, inflammatory bowel disease, and arthritis. The biological effects of leukotrienes are being actively researched to define the cellular processes involved. This may lead to the development of agents to control allergic and inflammatory reactions. Drugs inhibiting the formation of LTC_4 and LTB_4 are in clinical trials, whilst montelukast and zafirlukast have been introduced as orally active leukotriene (LTD_4) receptor antagonists for the prophylaxis of asthma.

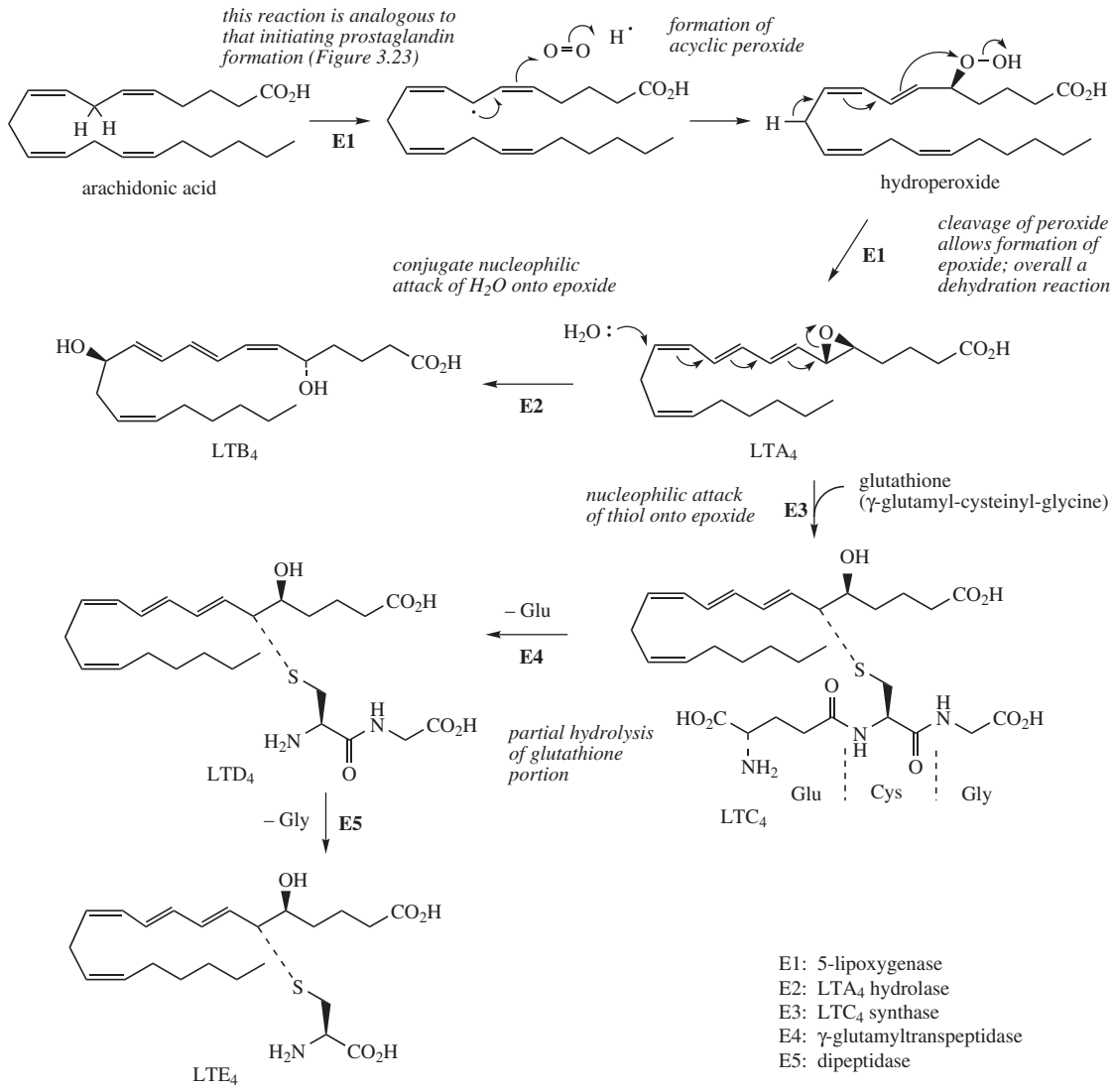


Figure 3.29

attacked directly by a nucleophile, in this case the sulfur atom of the tripeptide glutathione (γ-glutamyl-cysteinyl-glycine; Figure 3.29). The adduct produced in the latter reaction is termed **leukotriene C_4** (**LTC_4**). Partial hydrolysis in the tripeptide fragment then leads to **leukotriene D_4** (**LTD_4**) and **leukotriene E_4** (**LTE_4**). Analogues, e.g. LTA_3 and LTA_5, are also known, and these are derived from **$\Delta^{5,8,11}$-eicosatrienoic acid** and **$\Delta^{5,8,11,14,17}$-eicosapentaenoic acid** respectively. The subscript numeral indicates the total number of double bonds in the leukotriene chain.

POLYKETIDE SYNTHASES: GENERALITIES

During fatty acid biosynthesis, the growing chain is constructed by a Claisen reaction, catalysed by the ketosynthase activity in the FAS complex. This reaction initially produces a β-ketoester, and the ketone group is reduced after each condensation step and before the next round of chain extension (Figure 3.2). The overall reduction is achieved by a three-stage process: reduction to an alcohol (KR), dehydration to the conjugated ester (DH), then reduction of the double bond (ER). These reactions

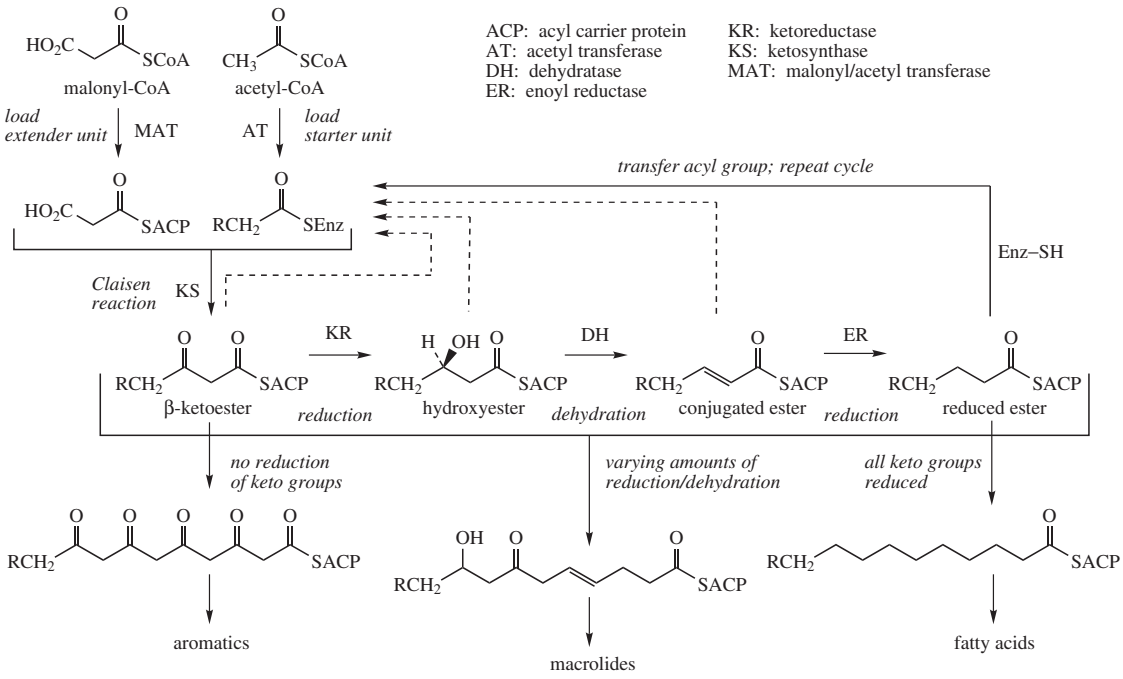


Figure 3.30

are shown again in Figure 3.30. This figure also shows how the same general processes may be modified to produce other polyketide systems, namely aromatics and macrolides; these processes are catalysed by **polyketide synthases (PKSs)**. For aromatics, the reduction sequence is omitted completely (or almost completely; see page 101) so that the β-ketoester re-enters the extension cycle, giving a poly-β-ketoester with alternate keto and methylene functions. This highly reactive poly-β-keto backbone undergoes further enzyme-catalysed intramolecular cyclization reactions which are responsible for generating a range of aromatic structures (see page 99). Macrolides are characterized by having carbon chains that have undergone partial, complete, or no reduction, according to the position in the chain. The chain may contain a mixture of hydroxyl groups, carbonyl groups, double bonds, and methylene groups. This means that the enzyme activities KR, DH, and ER are not all active during a particular extension cycle, and that a partially modified system, a β-ketoester, hydroxy ester, conjugated ester, or reduced ester, feeds back into the extension cycle. The order in which the modifications occur (or do not occur) is closely controlled by the enzyme.

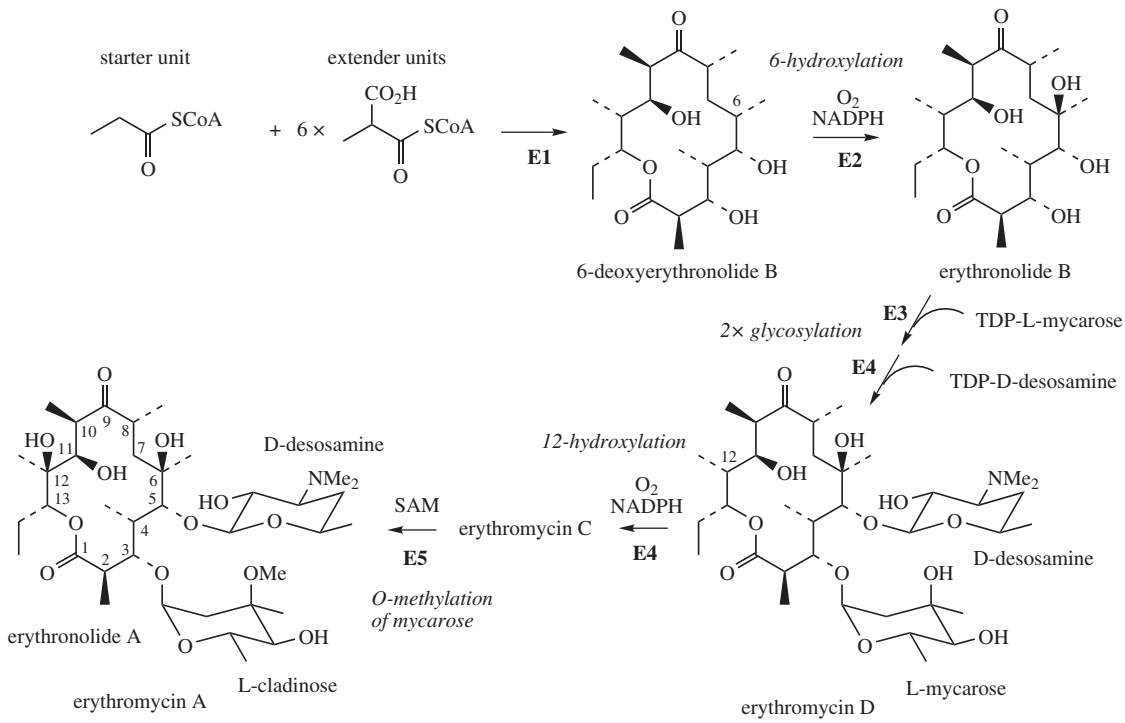
A detailed study of the genes, protein amino acid sequences, and mechanistic similarities in various PKS enzymes has led to three general types being distinguished. As with the type I and type II FASs (see page 40), **type I PKSs** are very large multifunctional proteins with individual functional domains, whilst **type II PKSs** are composed of a complex of individual monofunctional proteins. Type I systems can also be subdivided into 'iterative' (i.e. repeating) and 'non-iterative' categories. Iterative systems (like the FASs) use their functional domains repeatedly to produce a particular polyketide. Non-iterative systems possess a distinct active site for every enzyme-catalysed step. Type II systems are of the iterative type. Both types of system use ACP to activate acyl-CoA substrates and to channel the growing polyketide intermediates. Type I enzymes are responsible for macrolide biosynthesis, where the variability at each step in the biosynthetic pathway gives rise to much more structural diversity than encountered with fatty acids. The usual starter unit employed is either acetyl-CoA or propionyl-CoA, whilst malonyl-CoA or methylmalonyl-CoA is the main extender unit. Aromatic compounds are usually products from type II enzymes; though variation in starter units may be encountered, the chain extender unit is always malonyl-CoA. PKSs responsible for chain extension of cinnamoyl-CoA starter units leading to flavonoids and stilbenes (see page 116)

do not fall into either of the above categories. These enzymes differ from the other types in that they are homodimeric proteins, they utilize coenzyme A esters rather than ACPs, and they employ a single active site to perform a series of decarboxylation, condensation, cyclization, and aromatization reactions. They have now been termed **type III PKSs** or chalcone synthase-like PKSs, though many examples outside of the flavonoid group are now recognized. Type I PKSs are found in bacteria and fungi, type II PKSs are restricted to bacteria, whilst type III PKSs are found in plants, bacteria, and fungi.

Despite the differences in molecular architecture, the chemical aspects of chain construction are effectively the same in all PKSs. In all cases, these natural products are produced by Claisen reactions involving thioesters, and a basic chain that is composed of a linear sequence of C_2 units is constructed. Even when the starter unit is modified, or methylmalonyl-CoA is used as an extender unit instead of malonyl-CoA, the fundamental chain is effectively the same and is produced in the same mechanistic manner. Therefore, notwithstanding the various building blocks employed, these compounds are all still considered as derived via the acetate pathway. It is also more fitting to discuss the compounds in groups according to structural aspects rather than the protein type involved in their construction.

POLYKETIDE SYNTHASES: MACROLIDES

The **macrolides** are a large family of compounds, many with antibiotic activity, characterized by a macrocyclic lactone (sometimes lactam) ring. Rings are commonly 12-, 14-, or 16-membered, though polyene macrolides (see page 81) are larger in the range 26–38 atoms. The largest natural macrolide structure discovered has a 66-membered ring. The macrolide antibiotics provide us with excellent examples of natural products conforming to the acetate pathway, but composed principally of propionate units, or mixtures of propionate and acetate units. There is now extensive genetic evidence from a variety of systems to show that macrolide assembly is



E1: EryA (6-deoxyerythronolide B synthase; DEBS)
E2: EryF (6-deoxyerythronolide B hydroxylase)
E3: EryBV (mycarosyl transferase)
E4: EryCIII (desosaminyl transferase)
E5: EryK (erythromycin 12-hydroxylase)
E6: EryG (methyltransferase)

Figure 3.31

most often accomplished by non-iterative type I PKSs. Conceptually, polyketide synthesis by this group of enzymes is probably the easiest to envisage and understand, and provides the best grounding for appreciation of other systems. The PKS provides a biological production line of multifunctional proteins organized as discrete modules. Each module contains the enzyme activities for an extension cycle with its subsequent modifications. The developing polyketide chain attached to an ACP is thus modified according to the appropriate enzyme activities present, and is then passed on to another ACP prior to the next condensation and modification. The linear sequence of modules in the enzyme corresponds to the generated sequence of extender units in the polyketide product.

Erythromycin A (Figure 3.31) from *Saccharopolyspora erythraea* is a valuable antibacterial drug [Box 3.6] and contains a 14-membered macrocycle composed entirely of propionate units, both as starter and extension units, the latter via methylmalonyl-CoA. In common with many antibacterial macrolides, sugar units, including amino sugars, are attached through glycoside linkages. These unusual 6-deoxy sugars are frequently restricted to this group of natural products. In erythromycin A, the sugars are L-cladinose and D-desosamine. Chain extension and appropriate reduction processes lead to an enzyme-bound polyketide in which one carbonyl group has suffered total reduction, four have been reduced to alcohols, whilst one carbonyl is not reduced, and remains throughout the sequence. These processes ultimately lead to release of the modified polyketide as the macrolide ester **6-deoxyerythronolide B**, a demonstrated intermediate in the pathway to erythromycins (Figure 3.31). The stereochemistry in the chain is controlled by the condensation and reduction steps during chain extension, but a reassuring feature is that there appears to be a considerable degree of stereochemical uniformity throughout the known macrolide antibiotics. In the later stages of the biosynthesis of erythromycin, hydroxylations at carbon atoms 6 and 12, and addition of sugar units, are achieved. Methylation of a sugar hydroxyl is the last step in the sequence.

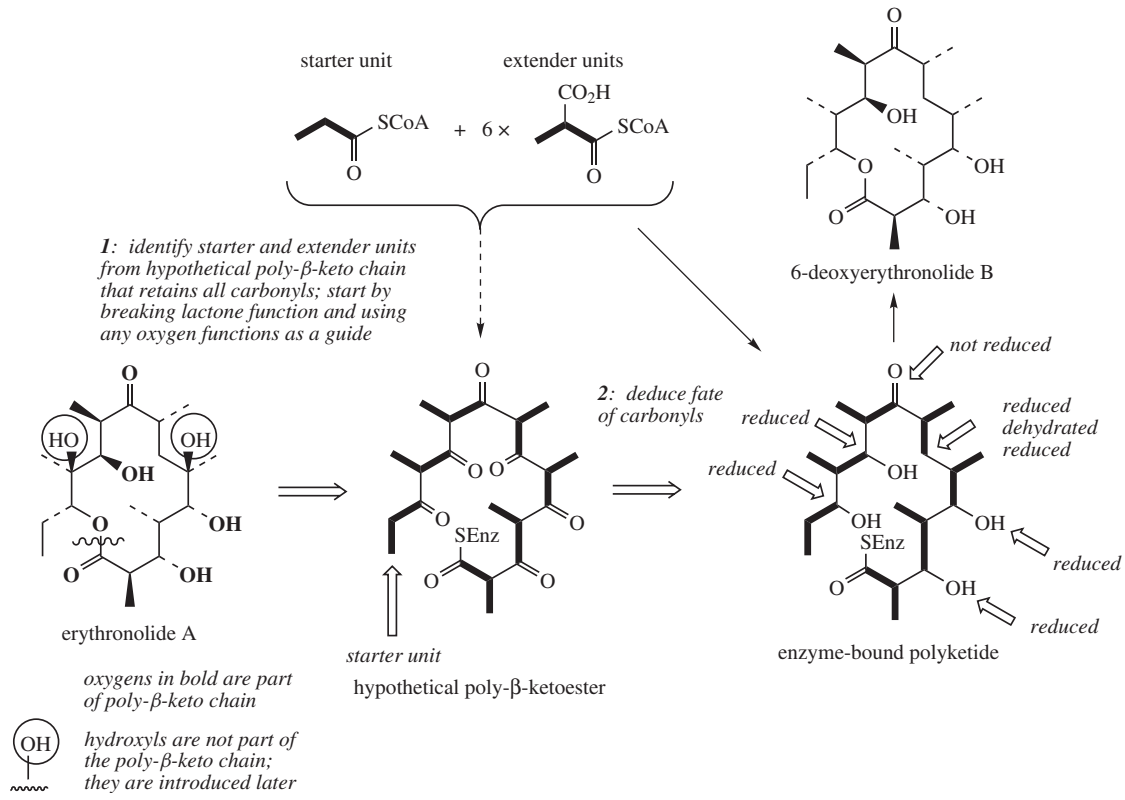


Figure 3.32

These latter modifications illustrate that features of the biosynthesis can be considered according to the timing of the steps in the sequence. The PKS controls a number of structural modifications on the basic polyketide, especially the reductive processes, whereas others are post-PKS changes. Here, hydroxylations, glycosylations, and methylation are post-PKS changes. It is usually possible to demonstrate these later conversions by appropriate interconversions and the isolation of enzymes catalysing the individual steps.

Since many macrolide structures are quite complex, and a variety of starter and extender units may be employed, it is appropriate here to demonstrate how simple analysis of the structure can help rationalize the biosynthetic origins of the basic macrolide. It also provides pointers to which modifications are likely to occur later in the pathway. Thus **erythronolide A**, the aglycone of erythromycin A, should be related to the hypothetical non-cyclized poly-β-ketoester, as shown in Figure 3.32. The positions of oxygen functions in erythronolide A are markers for the poly-β-keto system; the hydroxyls at C-6 and C-12 are not accommodated by alternate oxygenation, and can be deduced to be introduced later. Since a combination of acetyl starter and malonyl extender would give a straight-chain poly-β-ketoester, it is possible to identify the actual starter and extender

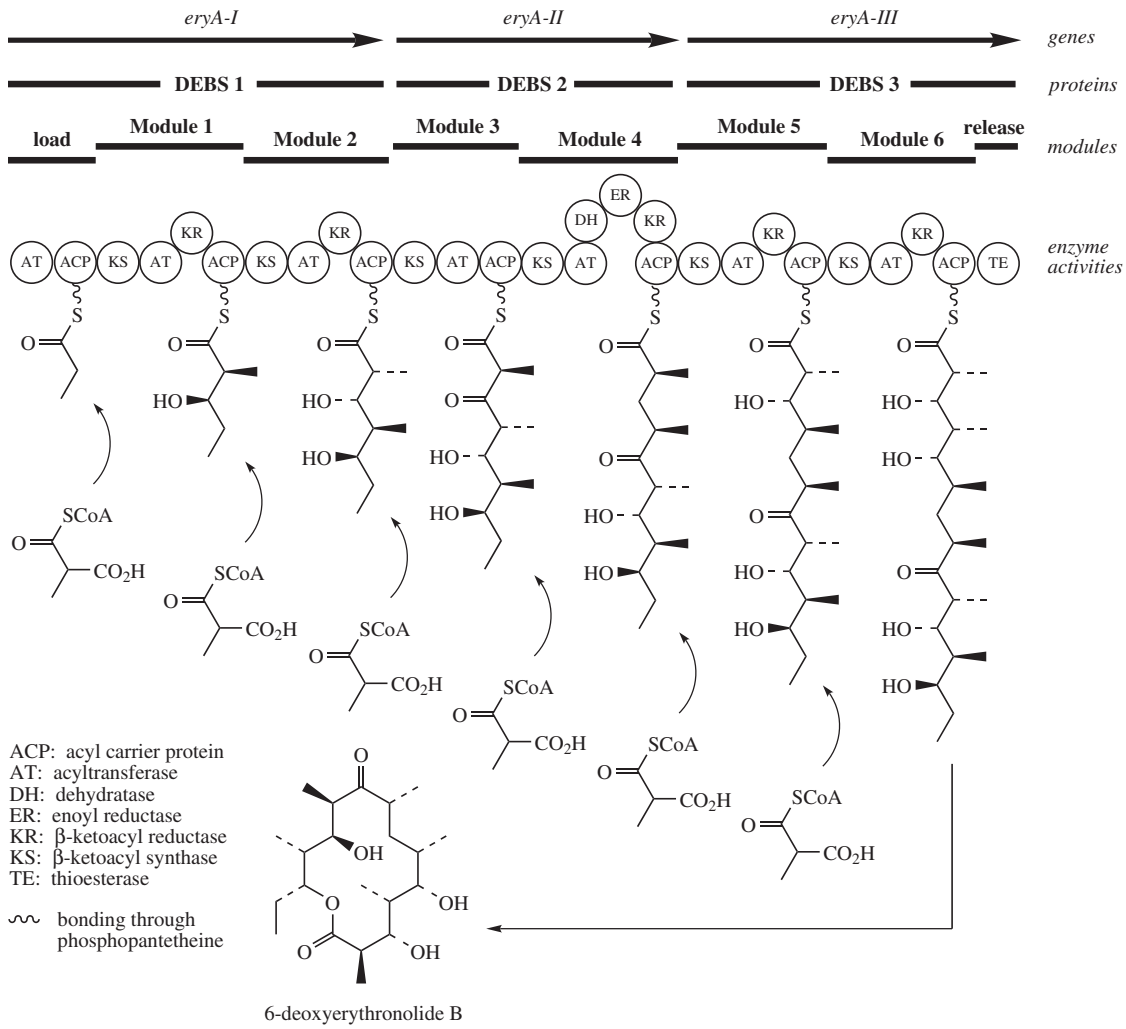


Figure 3.33

units used from the extra substituents at the start and between the carbonyls. In this case, a combination of propionyl-CoA as starter and six methylmalonyl-CoA extenders can be deduced. Comparison with erythronolide A then allows the fate of the carbonyl groups during polyketide assembly to be appraised. Cyclization of this polyketide upon release from the enzyme should generate 6-deoxyerythronolide B.

6-Deoxyerythronolide B synthase (DEBS) is a modular type I PKS involved in erythromycin biosynthesis; its organization and function are illustrated in Figure 3.33. The enzyme contains three subunits (DEBS-1, DEBS-2, and DEBS-3), each encoded by a gene (*eryA*-I, *eryA*-II, and *eryA*-III). It has a linear sequence of six modules, each of which contains the activities needed for one cycle of chain extension, and a loading domain associated with the first module. A minimal module contains a β-ketoacyl synthase (KS), an acyltransferase (AT), and an ACP that together would catalyse a two-carbon chain extension. The specificity of the AT for either malonyl-CoA or an alkyl-malonyl-CoA determines which two-carbon chain extender is used; in DEBS, the extender units are all methylmalonyl-CoA. The starter unit used is similarly determined by the specificity of the AT in the loading domain; for DEBS, this is propionyl-CoA. After each condensation reaction, the oxidation state of the β-carbon is determined by the presence of a β-ketoacyl reductase (KR), a KR plus a dehydratase (DH), or a KR plus DH plus an enoyl reductase (ER) in the appropriate module. Note that these modifications occur on the β-carbonyl, which is actually provided by the preceding module, i.e. module 1 modifies the oxidation state of the starter unit, etc. Thus, in DEBS, module 3 lacks any β-carbon-modifying domains, modules 1, 2, 5, and 6 contain KR domains and are responsible for production of hydroxy substituents, whereas module 4

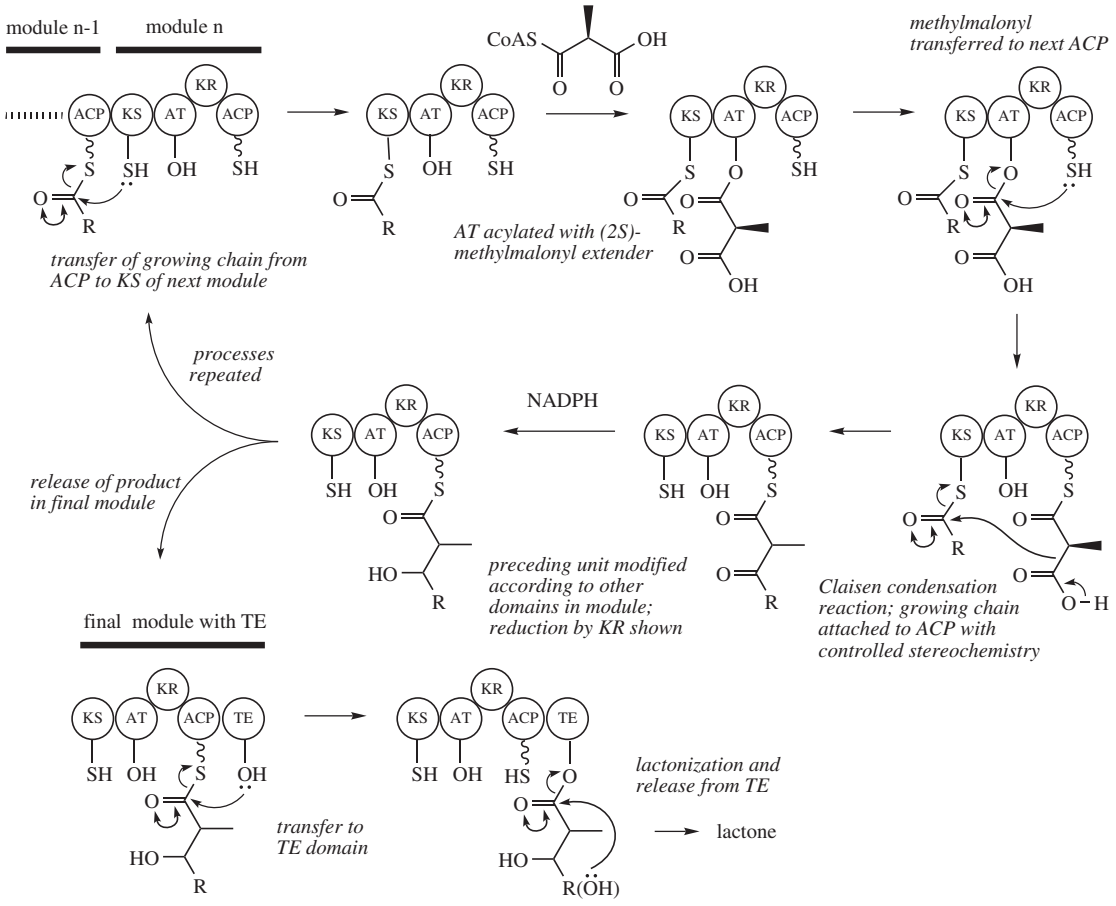


Figure 3.34

contains the complete KR, DH, and ER set, and results in complete reduction to a methylene. The chain is finally terminated by a thioesterase (TE) activity which releases the polyketide from the enzyme and allows cyclization.

The reactions involved in a typical module are outlined in Figure 3.34. The KS is responsible for the Claisen reaction and the growing chain is transferred to it from the ACP domain of the previous module. Initially, a starter unit, e.g. propionyl-CoA, would be bonded to the AT of the loading domain, then transferred to the loading domain's ACP – these steps are not shown in Figure 3.34, though analogous ones form part of the sequence given. The extender unit, here methylmalonyl-CoA, is attached to the AT domain, then transferred to the next downstream ACP. At this stage, the Claisen reaction is catalysed and the growing chain is thus attached to the ACP. The resultant stereochemistry of the methyl groups is also controlled by the enzyme; only (2*S*)-methylmalonyl-CoA is utilized, but epimerization occurs in some cycles, changing the stereochemistry. The reduction processes, if any, are now catalysed according to the presence of KR, DH, and ER domains in this module. In the example shown, reduction of the ketone to a hydroxyl is catalysed by the KR domain. The ACP is linked to the chain through a phosphopantetheine group, allowing the various domains to be accessed, as seen in FASs. Transfer of the modified chain to KS in the next module allows the sequence to be repeated. Alternatively, at the end of the sequence, transfer to the TE allows lactonization (via one of the alcohol functions) and release from the enzyme.

Overall, the AT specificity and the catalytic domains on each module determine the structure and stereochemistry of each two-carbon extension unit, the order of the modules specifies the sequence of the units, and the number of modules determines the size of the polyketide chain. The vast structural diversity of natural polyketides arises from combinatorial possibilities of arranging modules containing the various catalytic domains, the sequence and number of modules, the stereochemistry of associated side-chains, and the post-PKS enzymes which

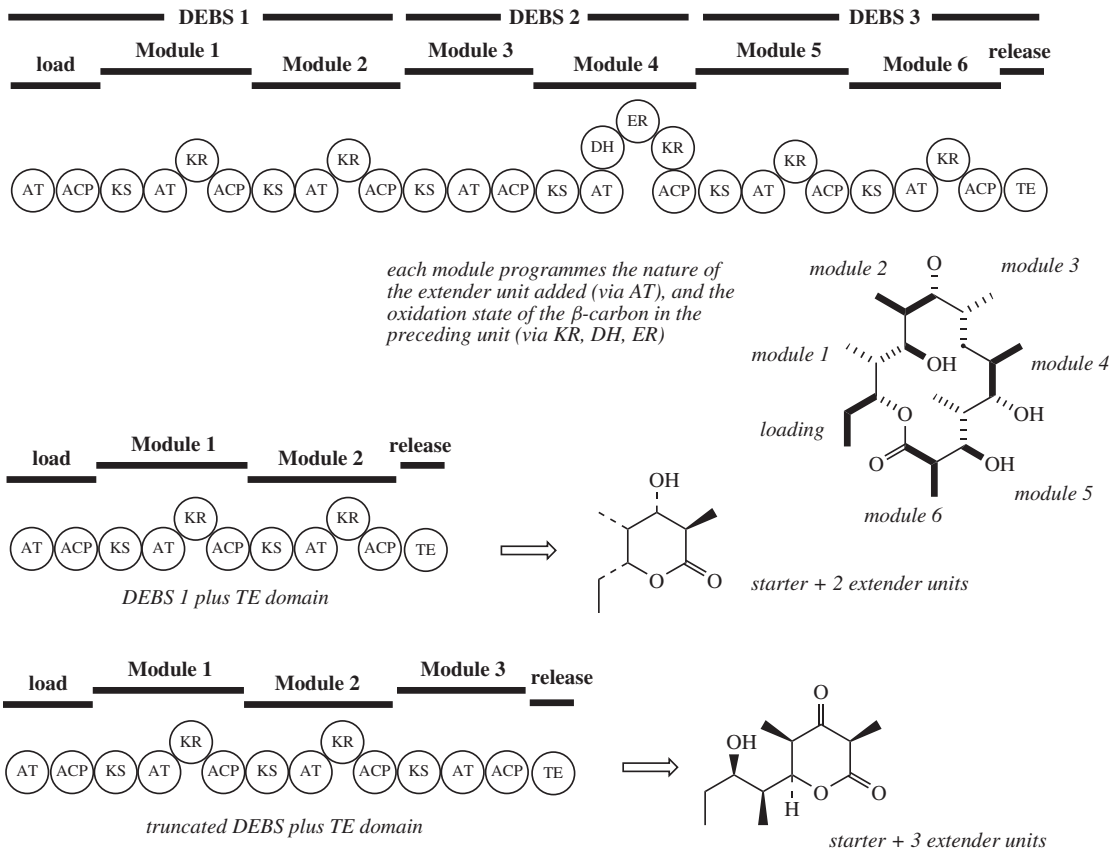


Figure 3.35

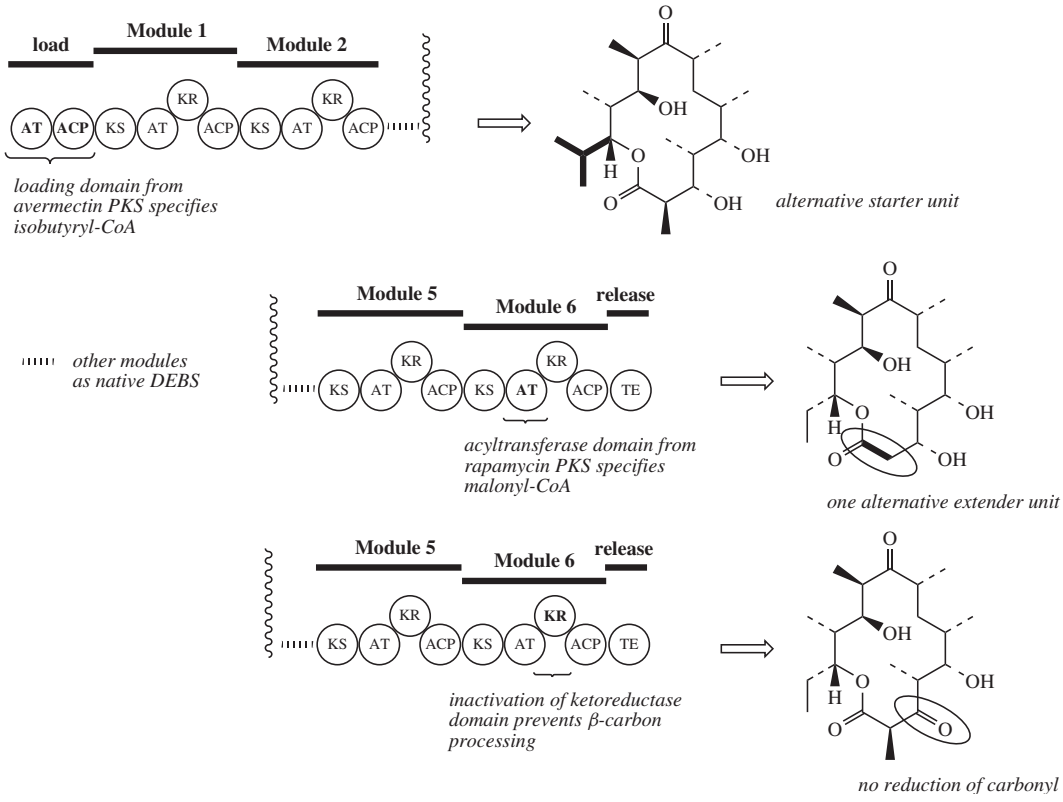


Figure 3.35 (*continued*)

subsequently modify the first-formed product by oxidation, glycosylation, etc., e.g. 6-deoxyerythronolide B → erythromycin A. Genetic engineering now offers vast opportunities for rational modification of the resultant polyketide structure, especially in PKS enzymes of the modular type I, which prove particularly adaptable. A few representative examples of more than a hundred successful experiments leading to engineered polyketides using the DEBS protein are shown in Figure 3.35.

Reducing the size of the gene sequence so that it encodes fewer modules results in the formation of smaller polyketides, characterized by the corresponding loss of extender units; in these examples, the gene encoding the chain-terminating TE also has to be attached to complete the biosynthetic sequence. Such products are naturally considerably smaller than erythromycin, but they have value as synthons containing defined functions and stereochemistry and can be used in the total synthesis of other natural product derivatives (see page 91). Replacing the loading domain of DEBS with that from another PKS, e.g. that producing avermectin (see page 80), alters the specificity of the enzyme for the starter unit. The loading module of the avermectin-producing PKS actually has a much broader specificity than that for DEBS; Figure 3.35 shows the utilization of isobutyryl-CoA, which features in the natural biosynthesis of avermectin B_{1b}. Other examples include the replacement of an AT domain (in DEBS specifying a methylmalonyl extender) with a malonyl-specific AT domain from the rapamycin-producing PKS (see page 84), and inactivation of a KR domain by modifying a catalytic amino acid, thus stopping any β-carbon processing for that module with consequent retention of a carbonyl group. Not all experiments in gene modification are successful, and even when they are, yields can be disappointingly lower than in the natural system. There is always a fundamental requirement that enzymes catalysing steps after the point of modification need to have sufficiently broad substrate specificities to accept and process the abnormal compounds being synthesized; this becomes more unlikely where two or more genetic changes have been made. Nevertheless,

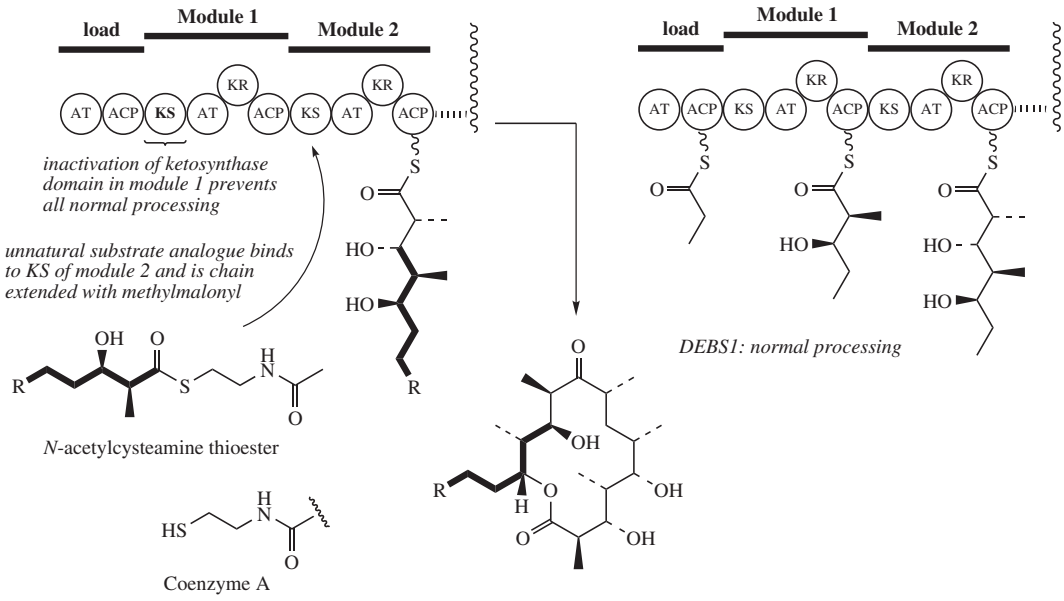


Figure 3.36

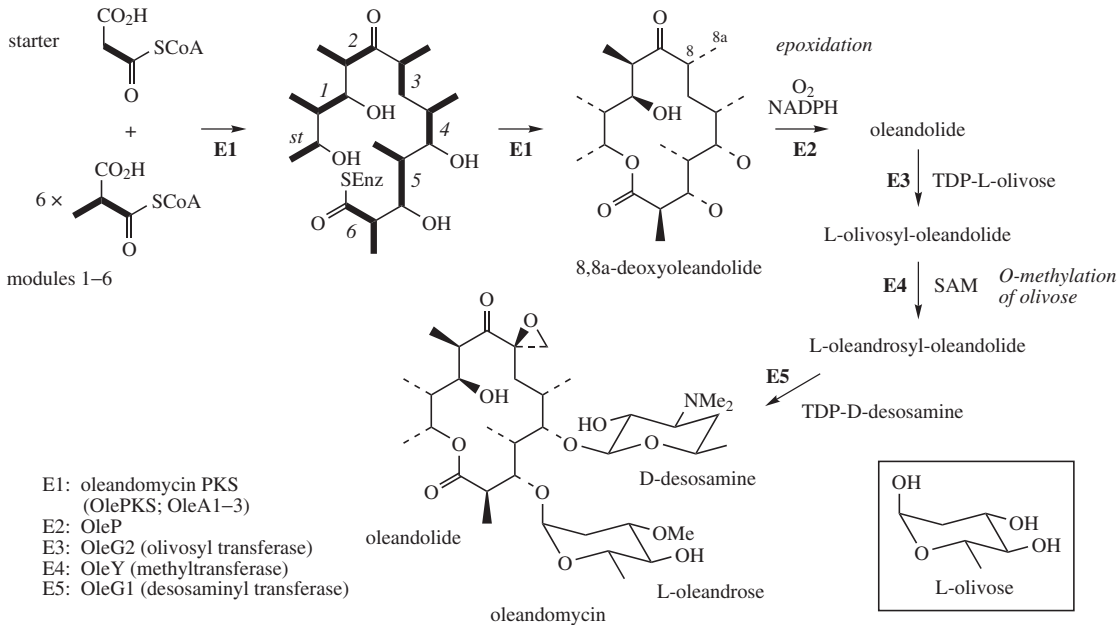


Figure 3.37

some multiple modifications have been successful, and it has also been possible to exploit changes in a combinatorial fashion using different expression vectors for the individual subunits, thus creating a library of polyketides, which may then be screened for potential biological activity.

Another approach is to disable a particular enzymic function, again by altering a specific amino acid, and to supply an alternative synthetic substrate analogue. Provided this substrate is not too dissimilar to the natural substrate, it may then be processed by the later enzymes. This is exemplified in Figure 3.36, where a number of artifical substrates may be utilized to produce a range of erythromycin derivatives inaccessible by semi-synthesis. The KS domain in module 1 is disabled by replacement of the active site cysteine, and the alternative substrate is supplied in the form of an *N*-acetylcysteamine thioester. The *N*-acetylcysteamine group acts as a simple analogue of coenzyme A, which is acceptable to and processed by the PKS.

A combination of propionate and acetate units is used to produce the 14-membered macrocyclic ring of **oleandomycin** (Figure 3.37) from *Streptomyces antibioticus* [Box 3.6], but otherwise, many of the structural features and the stereochemistry of oleandomycin resemble those of erythromycin A. Structural analysis as described above for erthyromycin would predict that the starter unit should

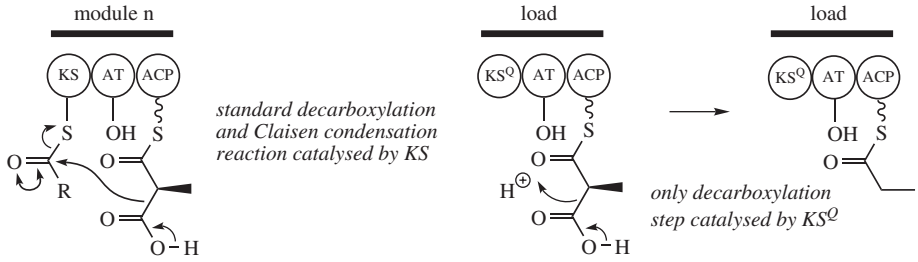


Figure 3.38

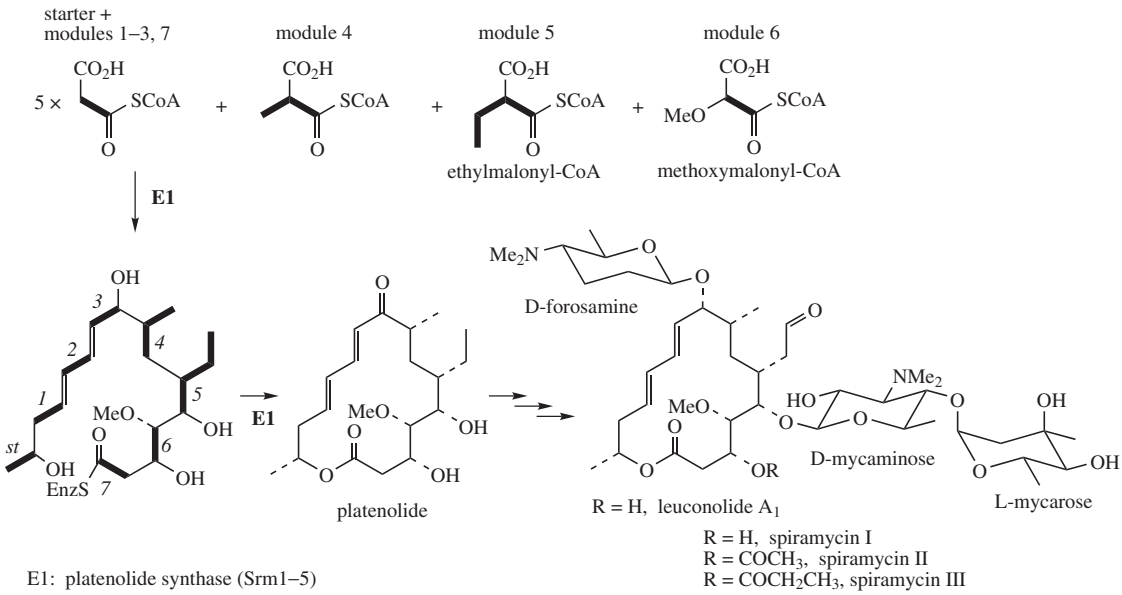
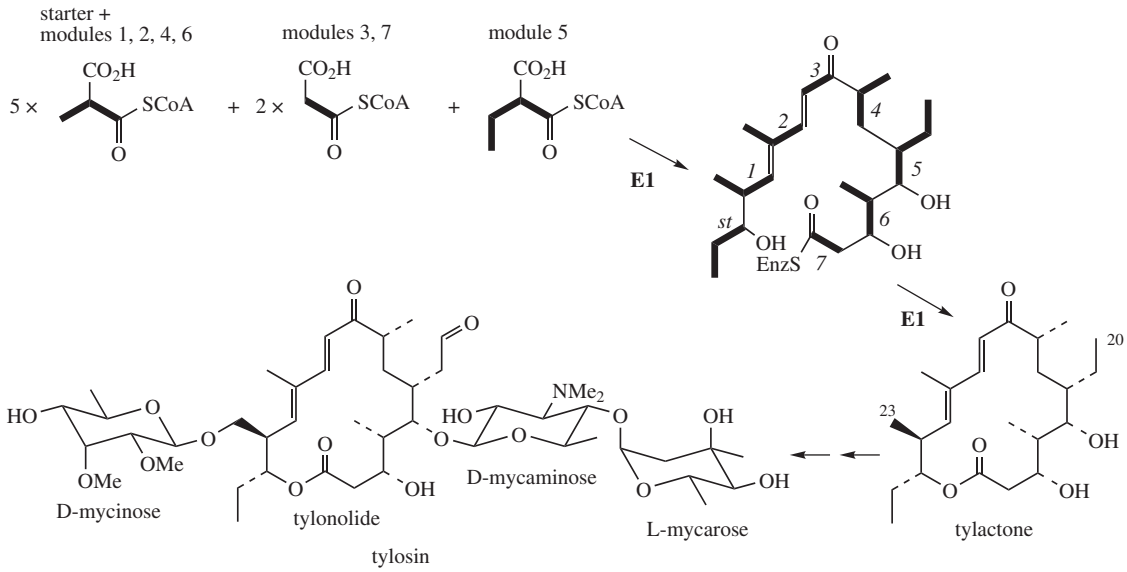


Figure 3.39

be acetyl-CoA, whilst six propionates, via methylmalonyl-CoA, supply the extension units (Figure 3.37). The oleandomycin PKS (OlePKS) is organized essentially the same as DEBS, though with one significant difference. It has been discovered that the loading domain in the OlePKS contains an AT specific for malonyl-CoA (not acetyl-CoA) and contains an additional KS domain (KS^Q). In KS^Q, an active site Cys residue is replaced by Gln and cannot make the essential thioester linkage; Q is the single-letter abbreviation for glutamine (see page 423). Consequently, the KS^Q domain cannot participate in any Claisen reaction, and its role is to decarboxylate the ACP-bound substrate to acetyl-ACP for use as the starter (Figure 3.38); many other examples with this type of mechanism are known. Hence, oleandomycin is derived from malonyl-CoA and six methylmalonyl-CoAs. The 8-methyl group is subsequently modified by a P-450 monooxygenase to give an epoxide function. The sugar units in oleandomycin are D-desosamine (as in erythromycin) and L-oleandrose; as with erythromycin, *O*-methylation of one sugar unit occurs after attachment to the macrolide.

Spiramycin I (Figure 3.39) [Box 3.6] from *Streptomyces ambofaciens* has a 16-membered lactone ring and is built up from a combination of six acetate units (one as starter, but all derived from malonyl-CoA), one propionate extender (from methylmalonyl-CoA), together with two further variants. First, a butyrate chain extender is incorporated via ethylmalonyl-CoA and yields an extension unit having an ethyl side-chain. Second, a methoxymalonyl-CoA extender incorporates an extension unit with a methoxy side-chain; this function might plausibly be expected to arise from hydroxylation followed by methylation. The relationships are outlined in Figure 3.39; the first-formed product is platenolide. The platenolide PKS is composed of seven modules plus a loading domain, organized into five subunits: loading domain plus modules 1 and 2 on the first subunit, module 3 on the second, modules 4 and 5 on subunit 3, and modules 6 and 7 on subunits 4 and 5. Platenolide can be modified to a number of other macrolides. For the spiramycins, the ethyl side-chain is subsequently oxidized to generate an aldehyde. Spiramycin I also contains a conjugated diene, the result of carbonyl reductions being followed by dehydration during chain assembly. **Tylosin** (Figure 3.40) [Box 3.6] from *Streptomyces fradiae* has many structural resemblances to the spiramycins, but can be analysed as a propionate starter (actually derived from methylmalonyl-CoA) with chain extension from two malonyl-CoA units, four methylmalonyl-CoA units, and one ethylmalonyl-CoA unit. Oxidation at C-20 in the ethyl side-chain to give an aldehyde is known to occur after introduction of the first sugar, D-mycaminose, and this is followed by hydroxylation of the methyl C-23; both enzymes involved are cytochrome P-450 systems.



E1: tylactone PKS (Tyl1–5)

Figure 3.40

Box 3.6

Macrolide Antibiotics

The macrolide antibiotics are macrocyclic lactones with a ring size typically 12–16 atoms, and with extensive branching through methyl substituents. Two or more sugar units are attached through glycoside linkages; these sugars tend to be unusual 6-deoxy structures often restricted to this class of compounds. Examples include L-cladinose, L-mycarose, D-mycinose, and L-oleandrose. At least one sugar is an amino sugar, e.g. D-desosamine, D-forosamine, and D-mycaminose. These antibiotics have a narrow spectrum of antibacterial activity, principally against Gram-positive microorganisms. Their antibacterial spectrum resembles, but is not identical to, that of the penicillins, so they provide a valuable alternative for patients allergic to the penicillins. Erythromycin is the principal macrolide antibacterial currently used in medicine.

Erythromycins

The erythromycins (Figure 3.41) are macrolide antibiotics produced by cultures of *Saccharopolyspora erythraea* (formerly *Streptomyces erythreus*). The commercial product **erythromycin** is a mixture containing principally erythromycin A plus small amounts of erythromycins B and C (Figure 3.41). Erythromycin activity is predominantly against Gram-positive bacteria, and the antibiotic is prescribed for penicillin-allergic patients. It is also used against penicillin-resistant *Staphylococcus* strains, in the treatment of respiratory tract infections, and systemically for skin conditions such as acne. It is the antibiotic of choice for infections of *Legionella pneumophila*, the cause of legionnaire's disease. Erythromycin exerts its antibacterial action by inhibiting protein biosynthesis in sensitive organisms. It binds reversibly to the larger 50S subunit of bacterial ribosomes and blocks the translocation step in which the growing peptidyl-tRNA moves from the aminoacyl acceptor site to the peptidyl donor site on the ribosome (see page 422). The antibiotic is a relatively safe drug with few serious side-effects. Nausea and vomiting may occur, and if high doses are prescribed, a temporary loss of hearing might be experienced. Hepatotoxicity may also occur at high dosage. Unfortunately, the drug has a particularly vile taste.

Erythromycin is unstable under acidic conditions, undergoing degradation to inactive compounds by a process initiated by the 6-hydroxyl attacking the 9-carbonyl to form a hemiketal, and by a similar reaction involving the 12-hydroxyl. Dehydration to an enol ether may follow, though the principal product is the spiroketal anhydroerythromycin A (Figure 3.42). The keto form is the only material with significant antibacterial activity, and the enol ether shown is responsible for gastrointestinal side-effects. Thus, to protect oral preparations of erythromycin against gastric acid, they are formulated as enteric coated tablets, or as insoluble esters (e.g. ethyl succinate esters) which are then hydrolysed in the intestine. Esterification produces a taste-free product, and typically involves the hydroxyl of the amino sugar desosamine. To reduce acid instability, semi-synthetic analogues of erythromycin have also been developed. **Clarithromycin** (Figure 3.41) is a 6-*O*-methyl derivative of erythromycin A; this modification blocks 6,9-hemiketal formation as in Figure 3.42. **Azithromycin** (Figure 3.41) is a ring-expanded aza-macrolide in which the carbonyl function has been removed. In both analogues, the changes enhance activity compared with that of erythromycin, and cause fewer gastrointestinal side-effects. Clarithromycin is used as part of the drug regimen for eradication of *Helicobacter pylori* in the treatment of gastric and duodenal ulcers.

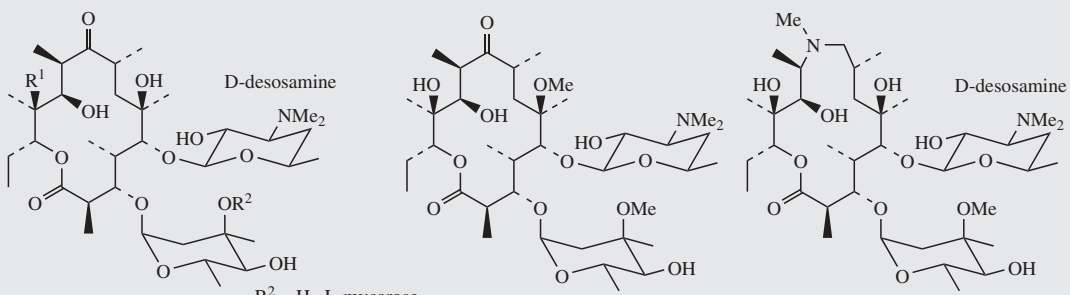


Figure 3.41

Box 3.6 (continued)

Bacterial resistance to erythromycin has become significant and has limited its therapeutic use against many strains of *Staphylococcus*. Several mechanisms of resistance have been implicated, one of which is a change in permeability of the bacterial cell wall. Differences in permeability also appear to explain the relative insensitivity of Gram-negative bacteria to erythromycin when compared with Gram-positive bacteria. Resistant bacteria may also modify the chemical nature of the binding site on the ribosome, thus preventing antibiotic binding, and some organisms are now known to metabolize the macrolide ring to yield inactive products. New derivatives based upon 3-keto-6-*O*-methylerythromycin A, termed ketolides, are providing important advances in this respect. These agents retain potent antibacterial activity towards erythromycin-susceptible organisms, but improved activity towards MLS_B (macrolide–lincosamide–streptogramin B)-resistant bacteria. In general, the ketolide structures differ from erythromycin by removal of the cladinose sugar and replacement with a 3-keto group, the presence of a 6-methoxy group, and attaching a cyclic carbamate at C-11/C-12. The first of these drugs in use is **telithromycin** (Figure 3.43).

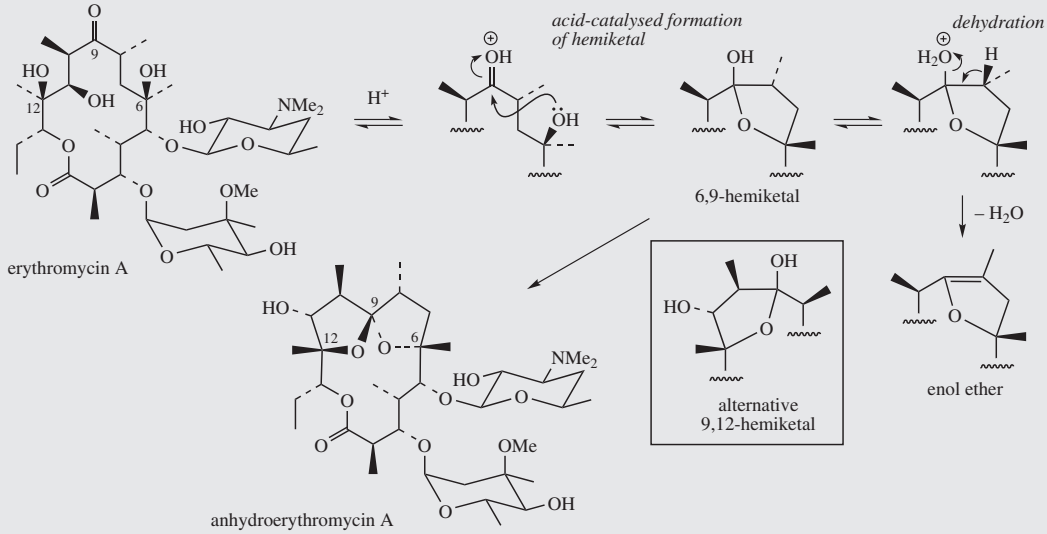


Figure 3.42

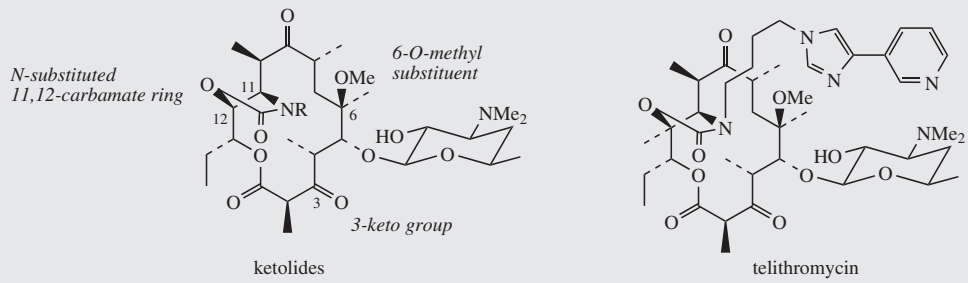


Figure 3.43

Oleandomycin and Spiramycin

Oleandomycin (Figure 3.37) is produced by fermentation cultures of *Streptomyces antibioticus* and has been used medicinally as its triacetyl ester **troleandomycin** against Gram-positive bacterial infections; it is rarely employed now. The **spiramycins**

Box 3.6 (continued)

(Figure 3.39) are macrolides produced by cultures of *Streptomyces ambofaciens*. The commercial antibiotic is a mixture containing principally spiramycin I, together with smaller amounts (10–15% each) of the acetyl ester spiramycin II and the propionyl ester spiramycin III. This antibiotic has recently been introduced into medicine for the treatment of toxoplasmosis, infections caused by the protozoan *Toxoplasma gondii*.

Tylosin

Tylosin (Figure 3.40) is an important veterinary antibiotic. It is produced by *Streptomyces fradiae*, and is used to control chronic respiratory diseases caused by *Mycoplasma galliseptum* in poultry and to treat Gram-positive infections in pigs.

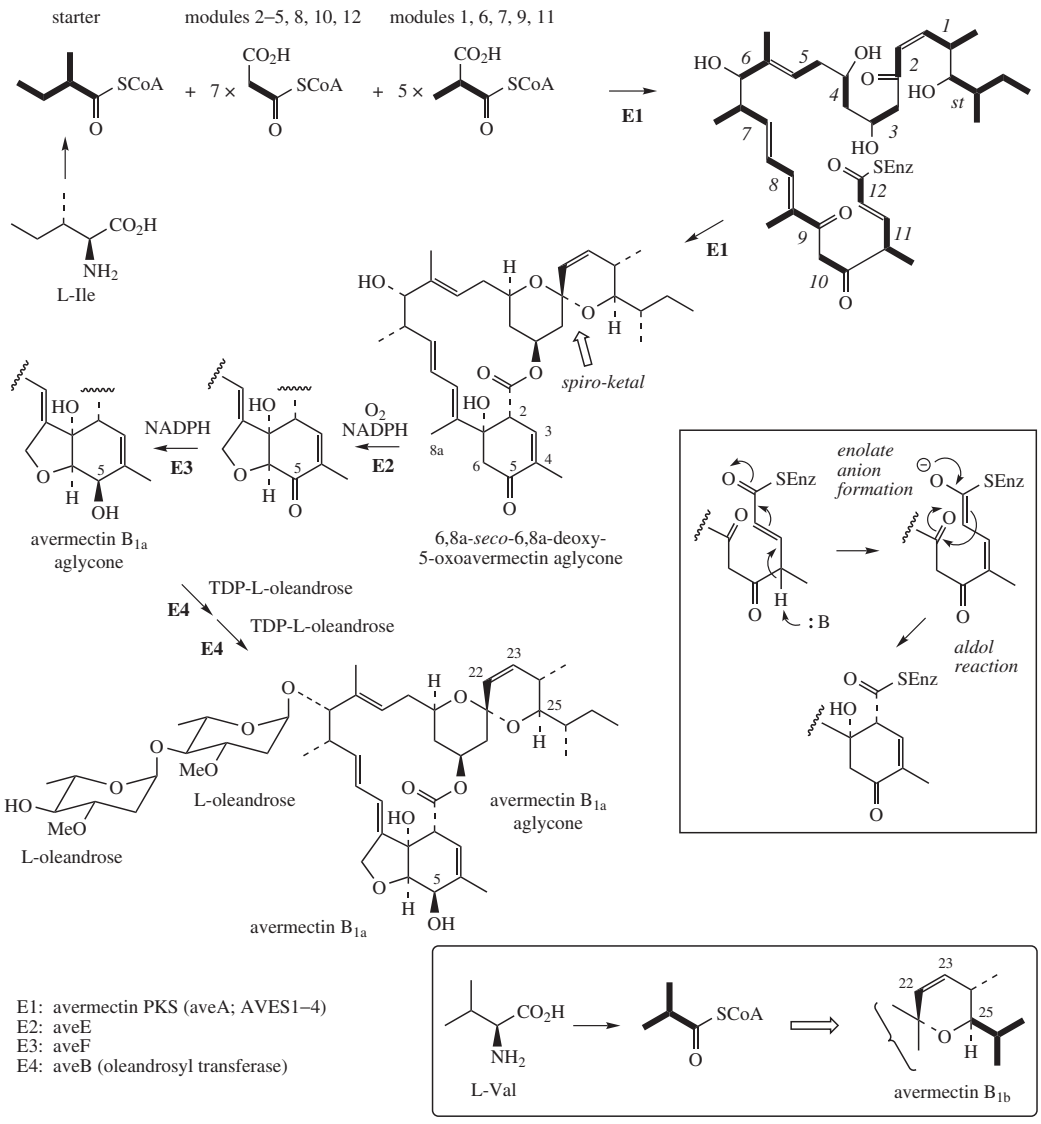


Figure 3.44

The **avermectins** (Figure 3.44) have no antibacterial activity, but possess anthelmintic, insecticidal, and acaricidal properties, and these are exploited in human and veterinary medicine [Box 3.7]. The avermectins are also 16-membered macrolides, but their structures are made up from a much longer polyketide chain (starter plus 12 extenders) which is also used to form oxygen heterocycles fused to the macrolide. **Avermectin B_{1a}** exemplifies a typical structure, and the basic carbon skeleton required to produce this can be postulated as in Figure 3.44. The starter unit in this case would be 2-methylbutyryl-CoA, which is derived from the amino acid L-isoleucine (compare iso-fatty acids, page 56). Both malonyl-CoA and methylmalonyl-CoA are then utilized as extender units. The simple structural analysis approach does not pick up two variants: extender 10 contains an unmodified keto function, and there is a double bond between extenders 11 and 12, not present in the final structure (C-2,3 in avermectin). Further, the product of the avermectin PKS is not a simple macrolide, but a structure in which additional cyclizations have also been catalysed. The spiroketal system is easily accounted for as a combination of ketone and two alcohol functions. The cyclohexenone ring arises by an aldol reaction as shown, and is dependent upon the 2,3-double bond and 5-keto function just mentioned. This reaction may precede lactone ring formation, based on the increased acidity of thioesters relative to oxygen lactones. The avermectin PKS is composed of four polypeptide subunits: load plus modules 1 and 2; modules 3–6; modules 7–9; modules 10–12. Intriguingly, module 7 contains a DH domain and module 10 a KR domain, but both of these are non-functional, so do not contribute to the oxygenation state of the extender unit added; in the other systems discussed so far, variation in extender modification is simply a result of the appropriate domain being absent. Post-PKS modifications include oxidative cyclization to produce the tetrahydrofuran ring, which appears to be catalysed by a single cytochrome P-450-dependent enzyme. Glycosylation is accomplished, unusually, by the repeated action of a single glycosyltransferase.

Box 3.7

Avermectins

The avermectins are a group of macrolides with strong anthelmintic, insecticidal, and acaricidal properties, but with low toxicity to animals and humans. They are produced by cultures of *Streptomyces avermectilis*. Some eight closely related structures have been identified, with avermectins B_{1a} (Figure 3.44) and B_{2a} (Figure 3.45) being the most active antiparasitic compounds. **Abamectin** (a mixture of about 85% avermectin B_{1a} and about 15% avermectin B_{1b}) is used on agricultural crops to control mites and insects. **Ivermectin** (Figure 3.45) is a semi-synthetic 22,23-dihydro derivative of avermectin B_{1a} (the commercial product also contains up to 20% dihydroavermectin B_{1b}) and was first used in veterinary practice against insects, ticks, mites, and roundworms. Although it is a broad spectrum nematocide against roundworms, it is inactive against tapeworms and flatworms, or against bacteria and fungi. It is an extremely potent agent, and is effective at very low dosages. It is now the drug of choice for use against filarial and several other worm parasites in humans, e.g. river blindness caused by the nematode *Onchocerca volvulus*. The avermectins and ivermectin target the glutamate-gated chloride channels unique to nematodes, insects, ticks, and arachnids, resulting in neuro-muscular paralysis and death.

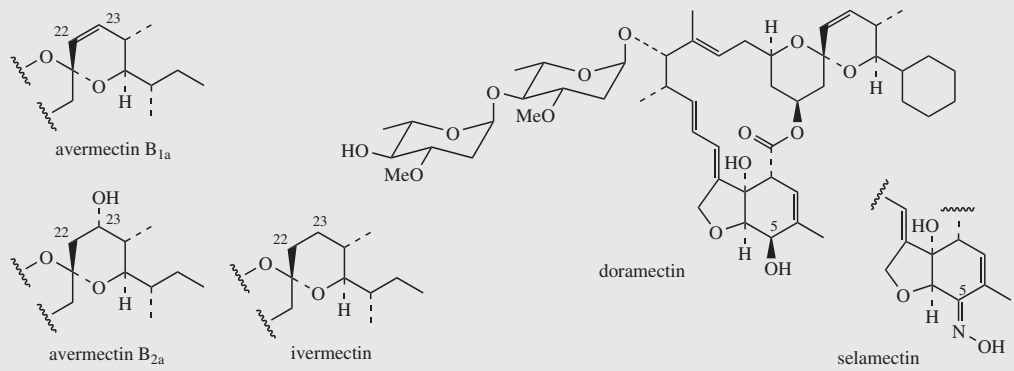


Figure 3.45

Box 3.7 (continued)

Doramectin is a newer analogue of avermectin B_{1a} (Figure 3.45) with extended biological activity, used to protect cattle against internal and external parasites. It is produced in a mutant of *Streptomyces avermectilis* that is unable to synthesize the branched-chain starter acids required for avermectin biosynthesis. The organism is able to utilize exogenously supplied acids as starter groups and incorporate them into modified avermectins; doramectin is produced by the use of cyclohexanecarboxylic acid in the culture medium (compare penicillins, page 461). It has also proved possible to express in this mutant biosynthetic genes for the production of cyclohexanecarbonyl-CoA, a natural compound involved in the biosynthesis of some other antibiotics. This engineered mutant synthesizes doramectin without the need for cyclohexanecarboxylic acid supplementation. **Selamectin** is the 5-oxime of doramectin, used to control fleas and internal parasites in dogs and cats.

Avermectins are usually isolated as a mixture in which the main *a* component has a 2-butyl group (derived from isoleucine) at C-25, whilst the minor *b* component has a 2-propyl group instead, e.g. **avermectin B_{1b}**. In this case, the starter group is 2-methylpropionyl-CoA, derived from the amino acid L-valine. The A-series of avermectins are the 5-methoxy analogues of the B-series.

Even larger macrolides are encountered in the **polyene macrolides**, most of which have antifungal properties, but not antibacterial activity [Box 3.8]. The macrolide ring size ranges from 26 to 38 atoms, and this also accommodates a conjugated polyene of up to seven *E* double bonds. Relatively few methyl groups are attached to the ring, and thus malonyl-CoA is utilized more frequently than methylmalonyl-CoA as chain extender. Typical examples are **amphotericin B** (Figure 3.46) from *Streptomyces nodosus* and **nystatin A_1** from *Streptomyces noursei*. These have very similar structures and are derived from the same basic precursors (Figure 3.46). The ring size becomes contracted due to cross-linking by formation of a hemiketal; this modification occurs before or during release from the PKS. The two compounds have slightly different hydroxylation patterns, part of which is introduced by a post-PKS hydroxylation. The two areas of conjugation in nystatin A_1 are extended into a heptaene system in amphotericin B. Both compounds are carboxylic acids, a result of oxidation of a propionate-derived methyl group, and glycosylated with the amino sugar D-mycosamine. The PKSs involved in amphotericin and nystatin biosynthesis are also closely related. They are composed of six polypeptide subunits: load; modules 1 and 2; modules 3–8; modules 9–14; modules 15–17; module 18. Both proteins also contain some non-functional domains.

Box 3.8

Polyene Antifungals

The polyene antifungals are a group of macrocyclic lactones with very large 26–38-membered rings. They are characterized by the presence of a series of conjugated *E* double bonds and are classified according to the longest conjugated chain present. Medicinally important ones include the heptaene amphotericin B and the tetraene nystatin. There are relatively few methyl branches in the macrocyclic chain. The polyenes have no antibacterial activity but are useful antifungal agents. Their activity is a result of binding to sterols in the eukaryotic cell membrane; this action explains the lack of antibacterial activity, because bacterial cells do not contain sterol components (see page 241). Fungal cells are also attacked rather than mammalian cells, since the antibiotics bind about 10-fold more strongly to ergosterol, the major fungal sterol (see page 252), than to cholesterol, the main animal sterol component (see page 251). This binding modifies the cell wall permeability and leads to formation of transmembrane pores that allow K^+ ions, sugars, and proteins to be lost from the microorganism. Though binding to cholesterol is less than to ergosterol, it is responsible for the observed toxic side-effects of these agents on humans. The polyenes are relatively unstable, undergoing light-catalysed decomposition, and are effectively insoluble in water. This insolubility actually protects the antibiotic from gastric decomposition, allowing oral treatment of infections in the intestinal tract.

Amphotericin is an antifungal polyene produced by cultures of *Streptomyces nodosus* and contains principally the heptaene amphotericin B (Figure 3.46) together with structurally related compounds, e.g. the tetraene amphotericin A (about 10%) which is the 28,29-dihydro analogue of amphotericin B. Amphotericin A is much less active than amphotericin B. Amphotericin is active against most fungi and yeasts, but it is not absorbed from the gut, so oral administration is restricted to the treatment of intestinal candidiasis. It is administered intravenously for treating potentially life-threatening systemic fungal infections. However, it then becomes highly protein bound, resulting in poor penetration and slow elimination from the body. After parenteral administration, toxic side-effects, including nephrotoxicity, are relatively common. Close supervision and monitoring of the patient is thus

Box 3.8 (continued)

necessary, especially since the treatment may need to be prolonged. Lipid formulations of amphotericin are much less toxic and have proven a significant advance. *Candida* infections in the mouth or on the skin may be treated with appropriate formulations. More recently, amphotericin has been shown to be among the few agents that can slow the course of prion disease in animal models.

Nystatin is a mixture of tetraene antifungals produced by cultures of *Streptomyces noursei*. The principal component is nystatin A_1 (Figure 3.46), but the commercial material also contains nystatin A_2 and A_3; these have additional glycoside residues. Nystatin is too toxic for intravenous use, but has value for oral treatment of intestinal candidiasis, as lozenges for oral infections, and as creams for topical control of *Candida* species.

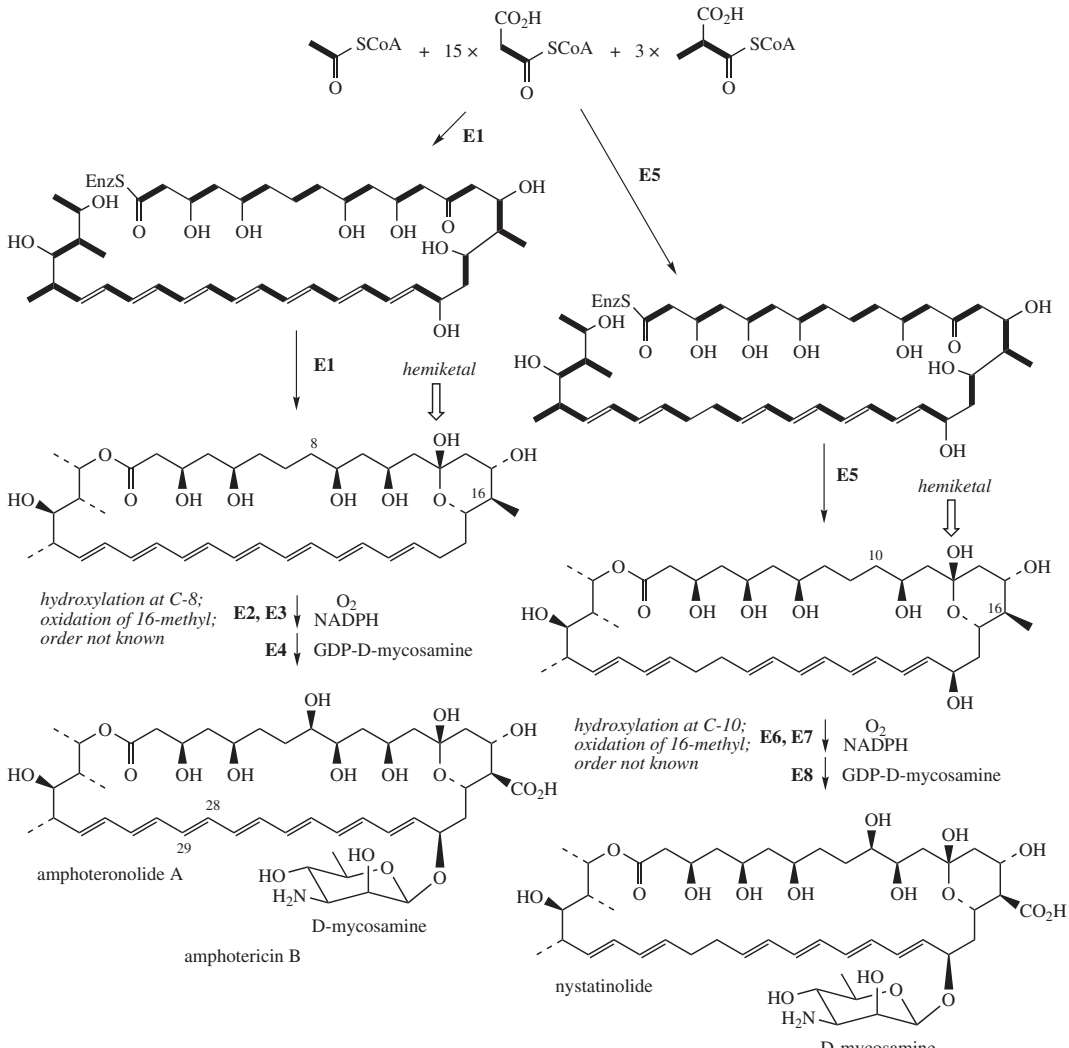


Figure 3.46

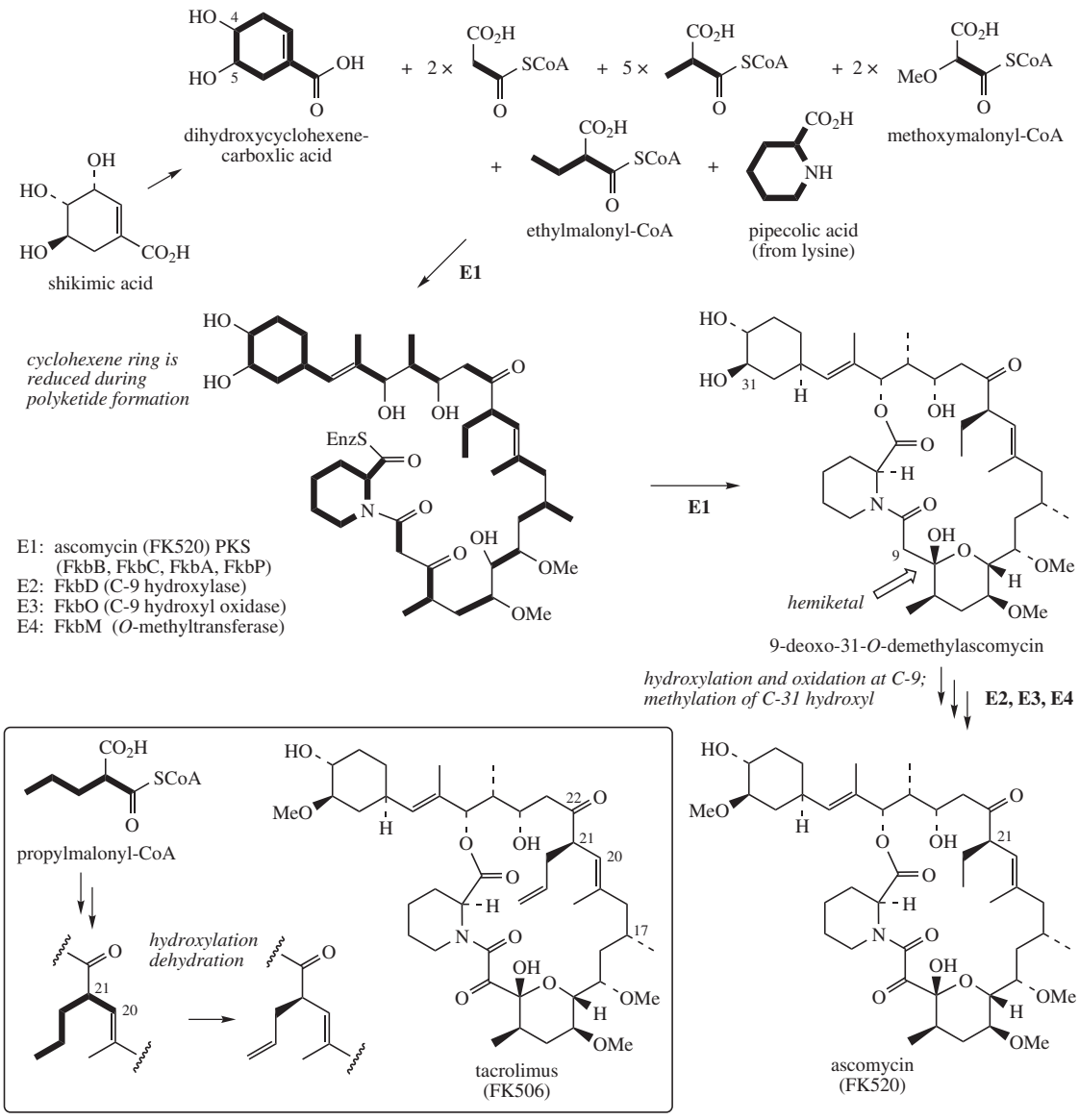
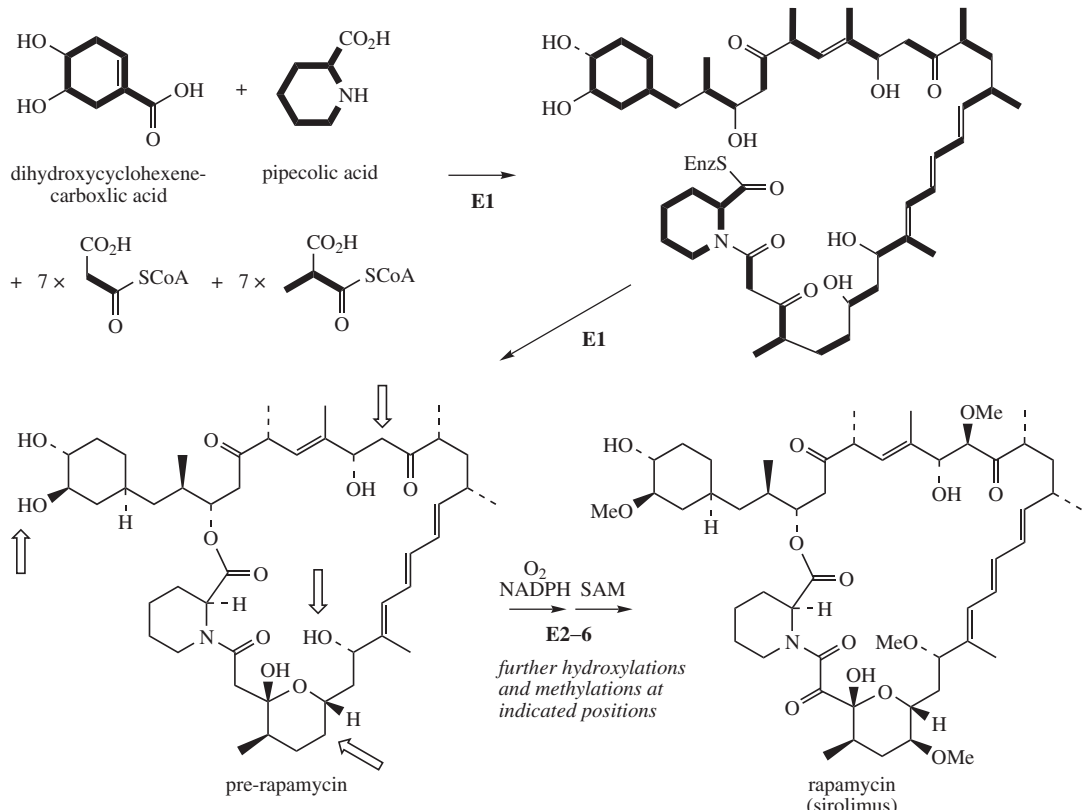


Figure 3.47

Two unusual and clinically significant macrolides are **ascomycin** (**FK520**), from *Streptomyces hygroscopicus* var. *ascomyceticus*, and **tacrolimus** (**FK506**), isolated from *Streptomyces tsukubaensis* (Figure 3.47). These compounds contain a 23-membered macrolactone that also incorporates an *N*-heterocyclic ring and are identical apart from the substituent at C-21. The piperidine ring and adjacent carbonyl are incorporated as pipecolic acid (see page 330) via an amide linkage onto the end of the growing chain. Chain assembly is catalysed by an enzyme combination that is partly a PKS and partly a non-ribosomal peptide synthetase (NRPS). Though PKSs and NRPSs utilize different substrates (carboxylic thioesters and amino acid thioesters respectively), they exhibit remarkable similarities in the modular arrangement of catalytic domains and product assembly mechanism (see page 438). PKS–NRPS hybrid systems appear to have evolved, and these produce metabolites

incorporating amino acid units as well as carboxylic acids. In the cases of ascomycin and tacrolimus, the predominantly PKS system also contains an NRPS-like protein, responsible for incorporating the pipecolic acid moiety. The main chains are derived principally from acetate and propionate, the fragments of which can readily be identified. Two methoxymalonyl-CoA extenders are utilized for ascomycin production and most likely for tacrolimus as well. Methoxymalonyl-CoA is also an extender in spiramycin biosynthesis (see page 75), and its production from a C_3 glycolytic pathway intermediate such as glyceric acid is also part of the PKS system. Ethylmalonyl-CoA (for ascomycin) or propylmalonyl-CoA (for tacrolimus) provides the remaining extender. In the case of tacrolimus, post-PKS hydroxylation and dehydration accounts for the allyl side-chain at C-21.

The starter unit is 4,5-dihydroxycyclohex-1-enecarboxylic acid, a reduction product from shikimate; it is incorporated as the free acid, perhaps by an AMP-activated form analogous to that used in the NRPS mechanism (see page 440). A further surprise is that the cyclohexene ring is reduced during transfer from the load domain to module 1. Methylation of the 5-hydroxyl is a post-PKS modification. Although **rapamycin** (**sirolimus**) (Figure 3.48) contains a very large 31-membered macrocycle, several portions of the structure are identical to those in ascomycin and tacrolimus. Dihydroxycyclohexenecarboxylic acid and pipecolic acid are again utilized in its formation, whilst the rest of the skeleton is supplied by simple acetate and propionate residues. Methoxymalonyl-CoA is not used as an extender, and post-PKS modifications include the more common hydroxylation/methylation sequences. Ascomycin, tacrolimus, and derivatives are effective immunosuppressants, and they are proving to be valuable drugs in organ transplant surgery and in the treatment of skin problems [Box 3.9].



E1: rapamycin PKS (rapA–C, rapP)
E2, E3: rapJ, rapN (hydroxylases)
E4–6: rapI, rapM, rapQ (methyltransferases)

Figure 3.48

Box 3.9

Tacrolimus and Sirolimus

Tacrolimus (FK-506) (Figure 3.47) is a macrolide immunosuppressant isolated from cultures of *Streptomyces tsukubaensis*. It is used in liver and kidney transplant surgery. Despite the significant structural differences between tacrolimus and the cyclic peptide cyclosporin A (ciclosporin; see page 454), these two agents have a similar mode of action. They both inhibit T-cell activation in the immunosuppressive mechanism by binding first to a receptor protein giving a complex which then inhibits a phosphatase enzyme called calcineurin. The resultant aberrant phosphorylation reactions prevent appropriate gene transcription and subsequent T-cell activation. Structural similarities between the region C-17 to C-22 in tacrolimus and fragments of the cyclosporin A peptide chain have been postulated to account for this binding. Tacrolimus is up to 100 times more potent than cyclosporin A; it produces similar side-effects, including neurotoxicity and nephrotoxicity. Tacrolimus is also used topically to treat moderate to severe atopic eczema; the ascomycin derivative **pimecrolimus** (Figure 3.49) is also used for mild to moderate conditions.

Sirolimus (rapamycin) (Figure 3.48) is produced by cultures of *Streptomyces hygroscopicus* and is also used as an immunosuppressant drug. Although tacrolimus and sirolimus possess a common structural unit, and both inhibit T-cell activation, they appear to achieve this by somewhat different mechanisms. The first-formed sirolimus–receptor protein binds not to calcineurin, but to a different protein. It exerts its action later in the cell cycle by blocking growth-factor-driven cell proliferation. **Everolimus** (Figure 3.49) is a semi-synthetic derivative of sirolimus with better oral bioavailability now available for kidney and heart transplants. **Zotarolimus** (Figure 3.49) inhibits cell proliferation, preventing scar tissue formation following cardiovascular surgery; this agent is delivered via a coronary stent. Some semi-synthetic sirolimus derivatives have also been investigated as anticancer drugs. These also display modifications to the C-40 hydroxyl group in the side-chain cyclohexane ring. **Temsirolimus** is now approved for treatment of renal cell carcinoma, and **deforolimus** is in advanced clinical trials against a variety of cancers (Figure 3.49).

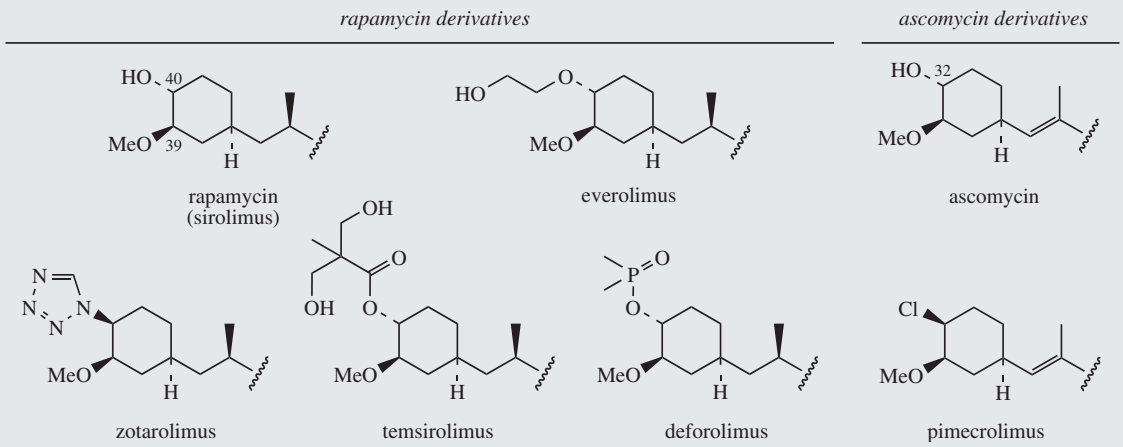


Figure 3.49

Attracting considerable interest at the present time are the **epothilones** (Figure 3.50), a group of macrolides produced by cultures of the myxobacterium *Sorangium cellulosum*. They are formed by a multienzyme system composed of nine PKS modules and one NRPS module distributed over six protein subunits. These compounds appear to employ an unusual starter unit containing a thiazole ring; this is constructed from the amino acid cysteine and an acetate unit by EpoB, the NRPS component of the enzyme (see also similar ring systems in bacitracin, page 443, and bleomycin, page 449). The loading module EpoA contains a decarboxylating KS domain, and the acetyl starter is thus derived from malonyl-CoA

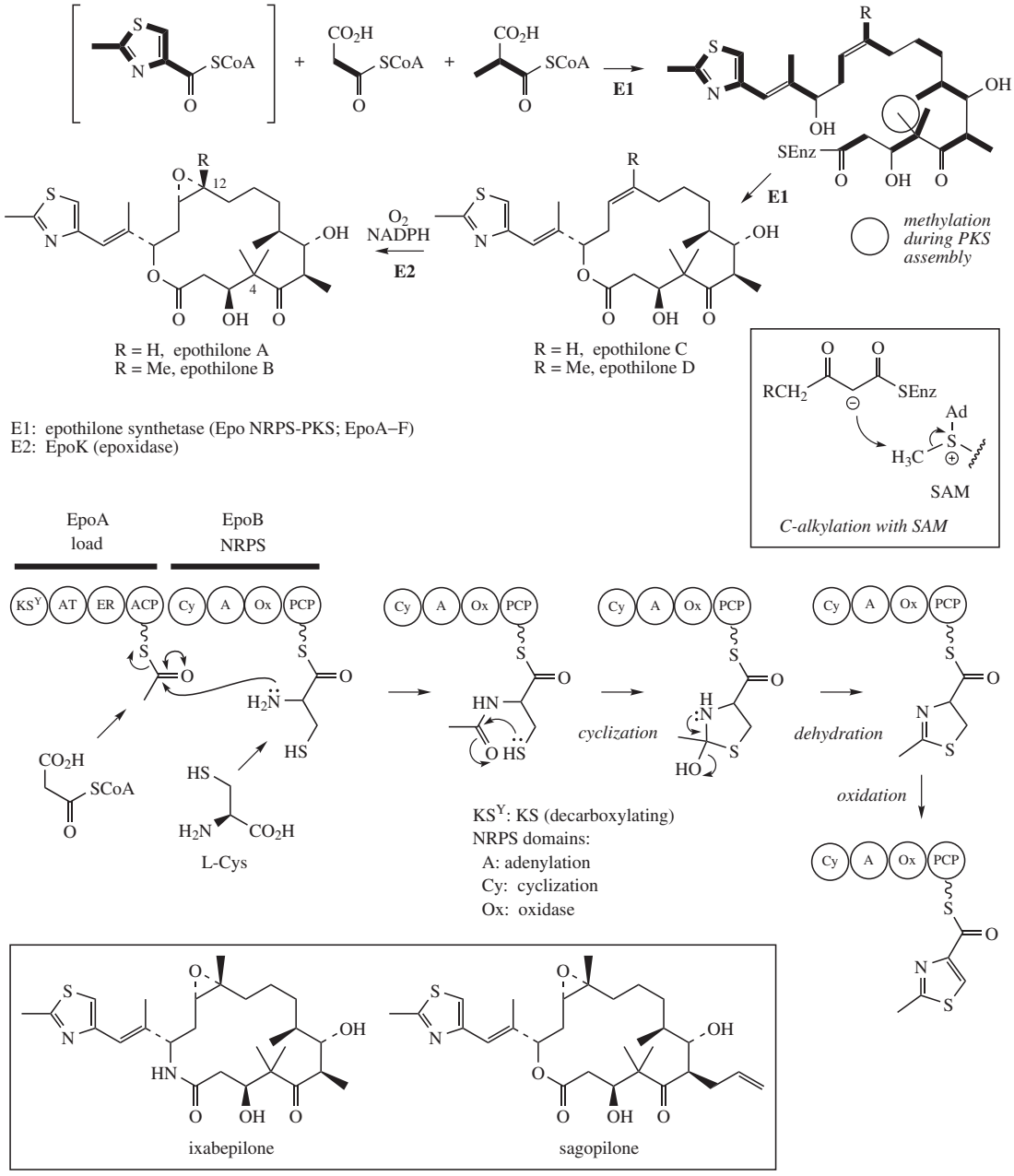


Figure 3.50

(compare oleandomycin, page 73). The macrolide ring also contains an extra methyl group at C-4, the result of methylation during polyketide chain assembly by a methyltransferase domain in module 8, part of the EpoE protein; this undoubtedly occurs whilst a β-diketone function is present. The other interesting feature is that this bacterium produces epothilone A and epothilone B in a ratio of about 2:1. These compounds differ in the nature of the substituent at C-12, which is hydrogen in epothilone A but a methyl group in epothilone B.

Genetic evidence shows that the PKS enzyme can accept either malonyl-CoA or methylmalonyl-CoA extender units for this position. Thus, epothilone B is constructed from three malonate and five methylmalonate extender units, whilst epothiolone A requires four units of each type. The first-formed products are epothilones C and D; these alkenes are converted into epoxides by a post-PKS cytochrome P-450-dependent epoxidase activity. The epothilones display marked antitumour properties with a mode of action paralleling that of the highly successful anticancer drug taxol (see page 224). However, the epothilones have a much higher potency (2000–5000 times) and are active against cell lines which are resistant to taxol and other drugs. There appears to be considerable potential for developing the epothilones or analogues into valuable anticancer drugs, and several derivatives

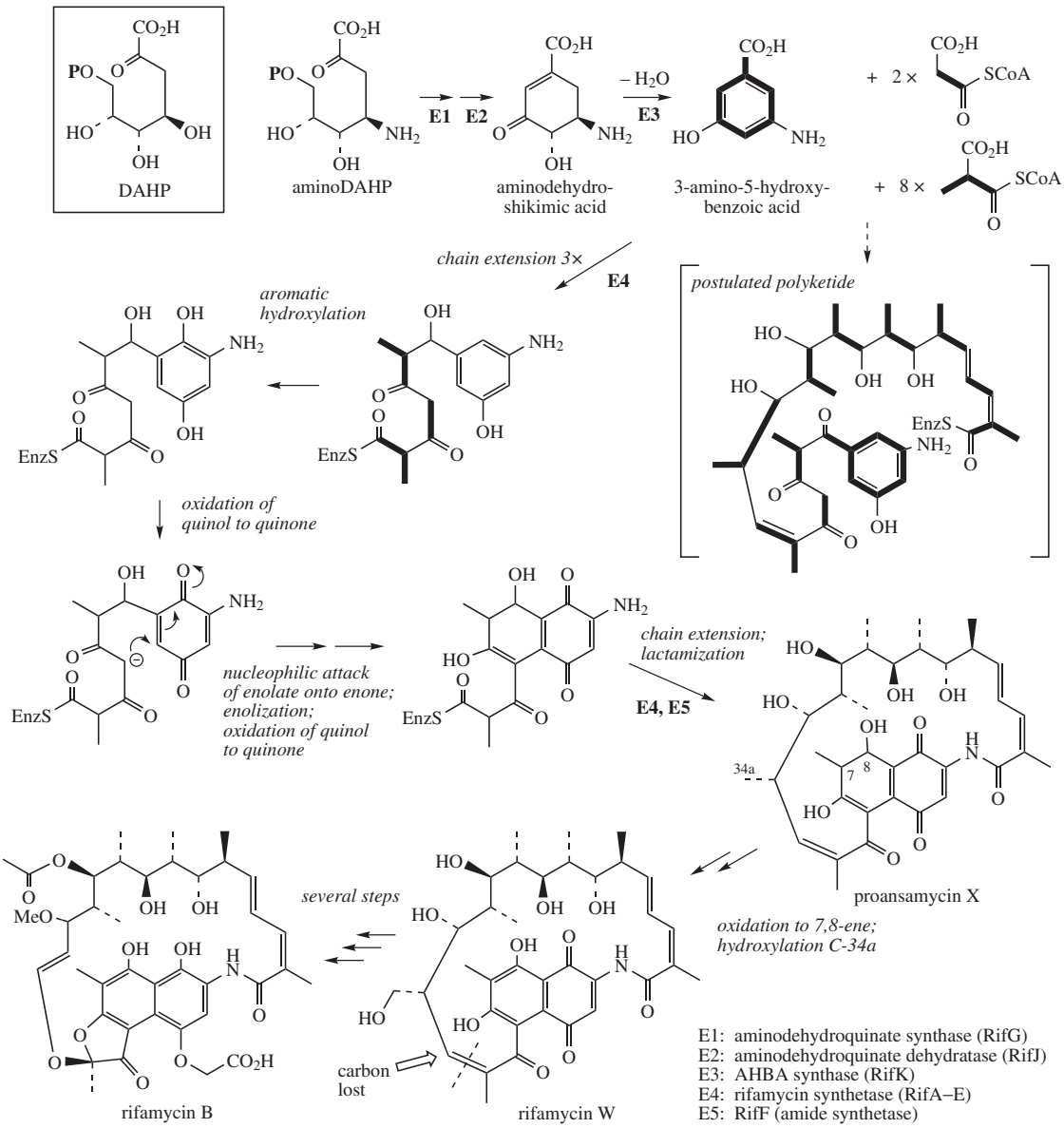


Figure 3.51

are undergoing clinical trials. These include epothilone B (**patupilone**) and its allyl analogue **sagopilone**; the lactam **ixabepilone** has recently been approved for drug use in the treatment of breast cancer (Figure 3.50).

A further group of macrolides are termed **ansa macrolides**; these have non-adjacent positions on an aromatic ring bridged by a long aliphatic chain (Latin: *ansa* = handle). The aromatic portion may be a substituted naphthalene or naphthaquinone, or alternatively a substituted benzene ring. The macrocycle in the ansamycins is closed by an amide rather than an ester linkage, i.e. ansamycins are lactams. The only ansamycins currently used therapeutically are **rifamycins** [Box 3.10], semi-synthetic naphthalene-based macrocycles produced from **rifamycin B** (Figure 3.51). Rifamycin synthetase comprises a sequence of five proteins, with an NRPS-like loading module and 10 PKS elongation modules. The nitrogen of the lactam ring derives from the starter unit, **3-amino-5-hydroxybenzoic acid** (Figure 3.51), which is a simple phenolic acid derivative produced by an unusual variant of the shikimate pathway (see Chapter 4). AminoDAHP is formed via aminosugar metabolism, and then the standard shikimate pathway continues with amino analogues. 3-Amino-5-hydroxybenzoic acid is produced from aminodehydroshikimic acid by dehydration. In the biosynthesis of **rifamycin B** (Figure 3.51) in *Amycolatopsis mediterranei*, this starter unit plus two malonyl-CoA and eight methylmalonyl-CoA extenders are employed to fabricate **proansamycin X** as the first product released from the enzyme. Simple structural analysis might predict an enzyme-bound intermediate as shown in Figure 3.51; it is now known that the naphthoquinone ring system is constructed during chain assembly, and a proposed mechansim is outlined. **Rifamycin W** and then the antibiotic rifamycin B are the result of several further modifications, though much of the detail requires further investigation. Post-PKS modifications include cleavage of the double bond, loss of one carbon, and formation of the ketal.

Box 3.10

Rifamycins

The rifamycins are ansamycin antibiotics produced by cultures of *Amycolatopsis mediterranei* (formerly *Nocardia mediterranei* or *Streptomyces mediterranei*). The crude antibiotic mixture was found to contain five closely related substances rifamycins A–E, but if the organism was cultured in the presence of sodium diethyl barbiturate (barbitone or barbital), the product was almost entirely rifamycin B (Figure 3.52). Rifamycin B has essentially no antibacterial activity, but may be converted chemically,

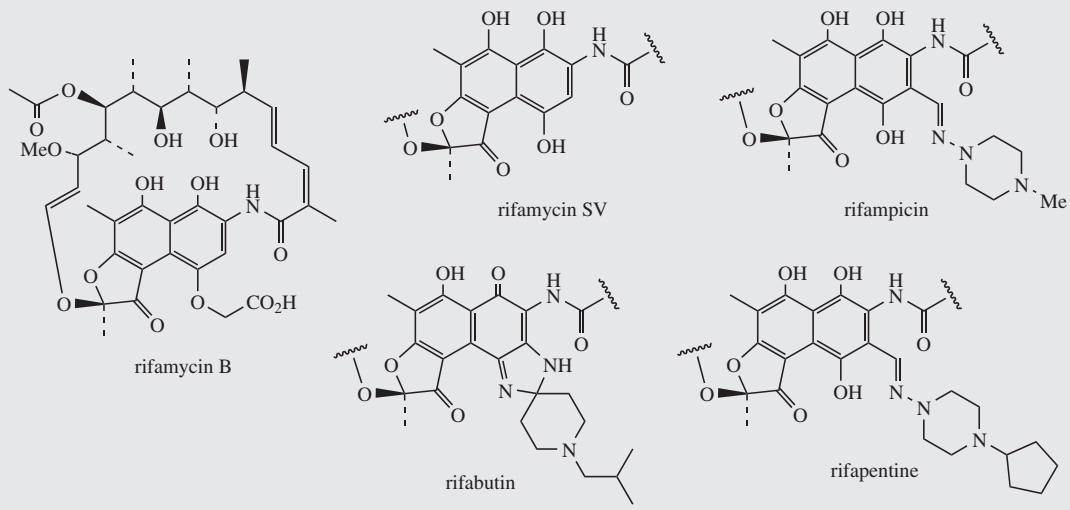


Figure 3.52

Box 3.10 (continued)

enzymically, or by biotransformation into rifamycin SV (Figure 3.52), a highly active antibacterial agent, and the first rifamycin to be used clinically. Further chemical modifications of rifamycin SV have produced better clinically useful drugs. Rifamycin SV is actually the immediate biosynthetic precursor of rifamycin B, and this conversion can be genetically blocked, resulting in the accumulation of rifamycin SV. Strain optimization has not been fruitful, however, and most commercial fermentations still rely on producing rifamycin B.

The most widely used agent is the semi-synthetic derivative **rifampicin** (Figure 3.52). Rifampicin has a wide antibacterial spectrum, with high activity towards Gram-positive bacteria and a lower activity towards Gram-negative organisms. Its most valuable activity is towards *Mycobacterium tuberculosis*, and rifampicin is a key agent in the treatment of tuberculosis, usually in combination with at least one other drug to reduce the chances for development of resistant bacterial strains. It is also useful in control of meningococcal meningitis and leprosy. Rifampicin's antibacterial activity arises from inhibition of RNA synthesis by binding to DNA-dependent RNA polymerase. RNA polymerase from mammalian cells does not contain the peptide sequence to which rifampicin binds, so RNA synthesis is not affected. In contrast to the natural rifamycins, which tend to have poor absorption properties, rifampicin is absorbed satisfactorily after oral administration and is also relatively free of toxic side-effects. The most serious side-effect is disturbance of liver function. A trivial, but to the patient potentially worrying, side-effect is discoloration of body fluids, including urine, saliva, sweat, and tears, to a red–orange colour, a consequence of the naphthalene/naphthoquinone chromophore in the rifamycins. **Rifabutin** (Figure 3.52) is a newly introduced derivative, which also has good activity against the *Mycobacterium avium* complex frequently encountered in patients with AIDS. **Rifapentine** (Figure 3.52) has a longer duration of action than the other anti-tubercular rifamycins.

Zearalenone (Figure 3.53), a toxin produced by the fungus *Gibberella zeae* and several species of *Fusarium*, has a relatively simple structure which is derived entirely from acetate/malonate units. It can be envisaged as a cyclization product from a polyketide where consecutive retention of several carbonyl functions will then allow formation of an aromatic ring by aldol condensation and enolizations near the carboxyl terminus – the processes of aromatic ring formation will be considered in more detail shortly (see page 99). Other parts of the chain have suffered considerable reductive modification (Figure 3.53). This simple structural analysis is essentially correct, though a novel feature here is that two proteins differing in their reducing nature are actually involved. Initial chain extension is carried out by a 'reducing' non-iterative type I PKS that constructs a C_{12} intermediate, which is then transferred to a second protein that lacks any KR, ER, and DH domains and carries out 'non-reducing' chain extension. The second protein possesses a starter unit-ACP transacylase (SAT) domain that loads the C_{12} chain for three further extensions.

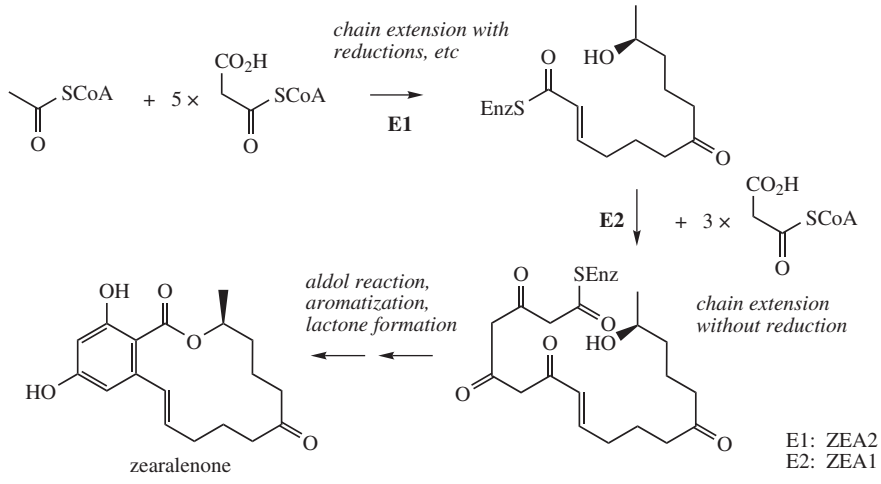


Figure 3.53

From a structural point of view, zearalenone is a remarkable example of an acetate-derived metabolite containing all types of oxidation level seen during the fatty acid extension cycle, i.e. carbonyl, secondary alcohol (eventually forming part of the lactone), alkene, and methylene, as well as having a portion which has cyclized to an aromatic ring because no reduction processes occur in that fragment of the chain.

POLYKETIDE SYNTHASES: LINEAR POLYKETIDES AND POLYETHERS

The macrolide systems described above are produced by formation of an intramolecular ester or amide linkage, utilizing appropriate functionalities in the growing polyketide chain. Macrolide formation does not always occur, and similar acetate/propionate precursors might also be expected to yield molecules which are essentially linear in nature. It is now possible to appreciate how some organisms, particularly marine microalgae, are able to synthesize polyunsaturated fatty acids such as EPA (Figure 3.54) directly by the use of PKS enzymes, instead of desaturating and elongating FAS-derived fatty acids (see page 49). The sequence proposed in the EPA-producing marine bacterium *Shewanella pneumatophori* allows chain extension with malonyl-CoA, introducing unsaturation via the KR/DH components. However, this would normally lead to a conjugated system of *trans* double bonds, as seen in the polyene antifungals (see page 82); the polyunsaturated fatty acids usually have a non-conjugated array of *cis* double bonds. EPA formation in *Shewanella* requires two additional enzyme activities, a 2,3-isomerase and a 2,2-isomerase, probably acting in concert with the DH (Figure 3.54). The first of these is an allylic isomerization (see page 51), whilst the second is a *trans–cis* isomerization, usually considered to require energy absorption (compare coumarins, page 161, and retinol, page 304).

Discodermolide (Figure 3.55) does contain a lactone moiety, though this is in a six-membered ring. Nevertheless, the bulk of the molecule can be rationalized as a linear array of malonate- and methylmalonate-derived units. In common with the fatty acids, double bonds in the chain are also *cis*. This compound is synthesized by the deep-sea marine sponge *Discodermia dissoluta*, and is of considerable interest as the most potent natural promoter of tubulin assembly yet discovered. Taxol (see page 224) and the epothilones (see page 85) have a similar mode of action. It has also been

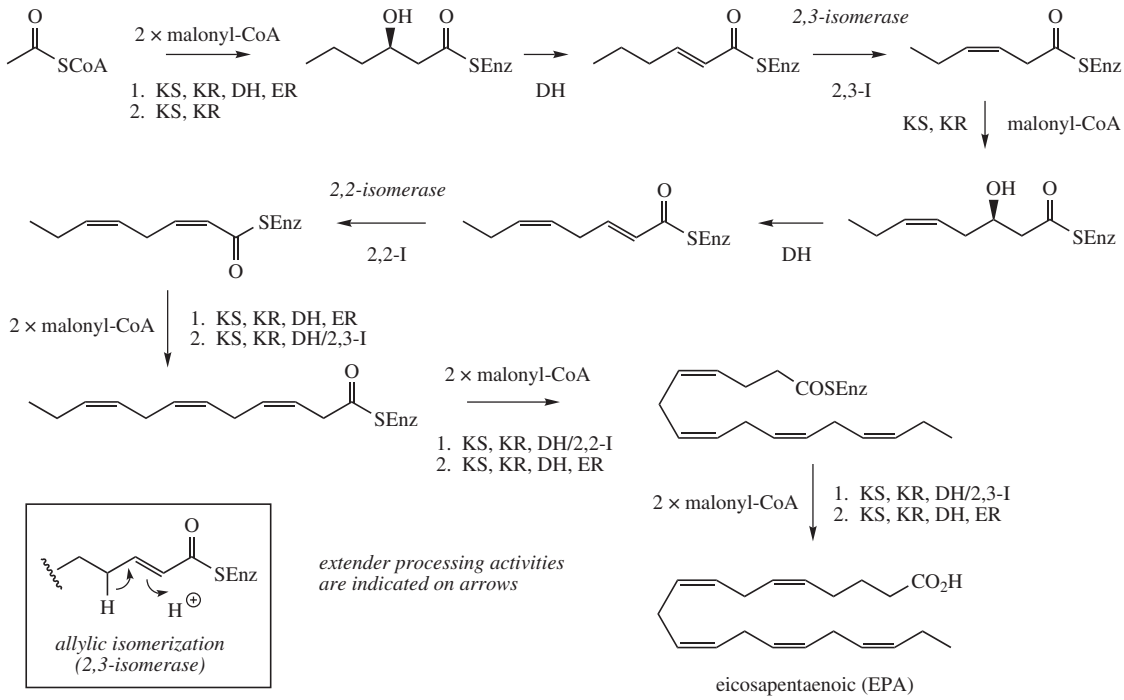


Figure 3.54

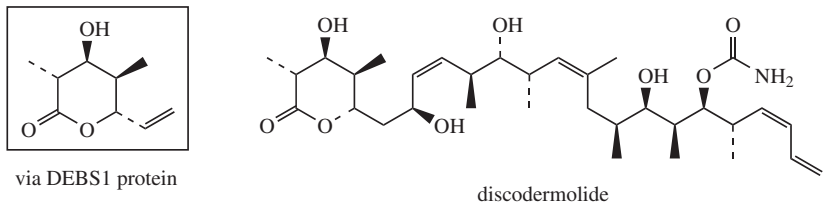


Figure 3.55

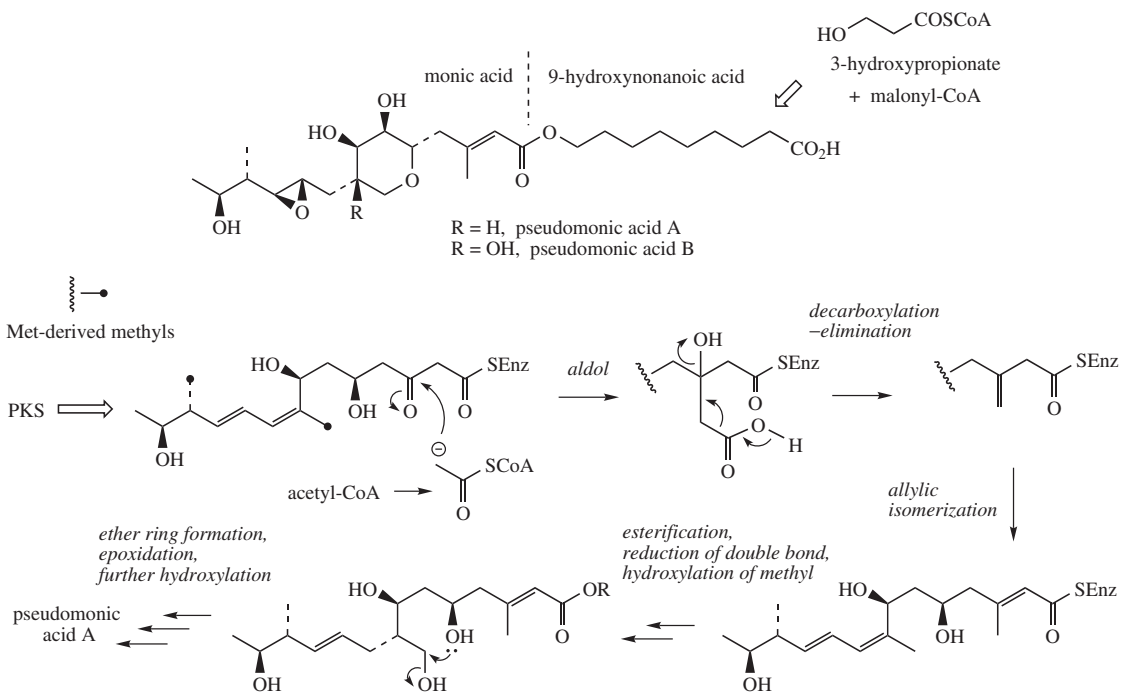


Figure 3.56

found that taxol and discodermolide display synergistic action in combination. Discodermolide itself has proved rather too toxic for drug use, and studies have been restricted by limited supplies. It has been possible to prepare key intermediates for total synthesis by using genetically modified DEBS1 proteins (see page 71) expressed in a suitable host.

Mupirocin is an antibiotic used clinically for the treatment of bacterial skin infections and for controlling *Staphylococcus aureus*, particularly methicillin-resistant *Staphylococcus aureus* (MRSA), when other antibiotics are ineffective. It has an uncommon mode of action, binding selectively to bacterial isoleucyl-tRNA synthase, which prevents incorporation of isoleucine into bacterial proteins. Mupirocin is produced by *Pseudomonas fluorescens* and consists of a mixture of pseudomonic acids, the major components being **pseudomonic acid A** (about 90%) and pseudomonic acid B (about 8%) (Figure 3.56). These are esters of monic acids with a C_9 saturated fatty acid, 9-hydroxynonanoic acid. The mupirocin gene cluster contains several domains resembling bacterial type I PKS and type I FAS systems, and other genes probably responsible for production of precursors and for later modifications. Parts of the sequence are particularly noteworthy, however. During PKS-mediated elaboration of the monic acid fragment, methyltransferase domains incorporate methyl groups, as seen with epothilone (page 86). A further

methyl group originates from the methyl of an acetate precursor. Since the protein MupH bears considerable similarity to HMG-CoA synthase, the enzyme involved in mevalonate biosynthesis (see page 190), a similar aldol addition mechanism is proposed in this reaction (see also streptogramin A formation, Figure 7.27). Decarboxylation–elimination (again compare mevalonate biosynthesis) catalysed by MupH is then followed by an allylic elimination. This 'methylation' looks much more complicated than an SAM-mediated reaction, but is a rather neat way of inserting a methyl group onto a carbonyl, which is not possible with SAM. Pseudomonic acid A formation in due course requires formation of an ether ring system and epoxidation of a double bond. Ether ring formation is thought to proceed through hydroxylation of one of the SAM-derived methyls.

Molecules such as **lasalocid A** (Figure 3.57) from *Streptomyces lasaliensis* and **monensin A** (Figure 3.58) from *Streptomyces cinnamonensis* are representatives of a large group of compounds called **polyether antibiotics**. These and related compounds are of value in veterinary medicine, being effective in preventing and controlling coccidiae, and also having the ability to improve the efficiency of food conversion in ruminants. The polyether

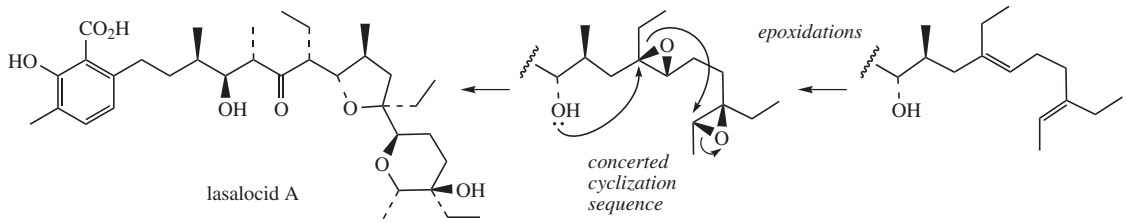


Figure 3.57

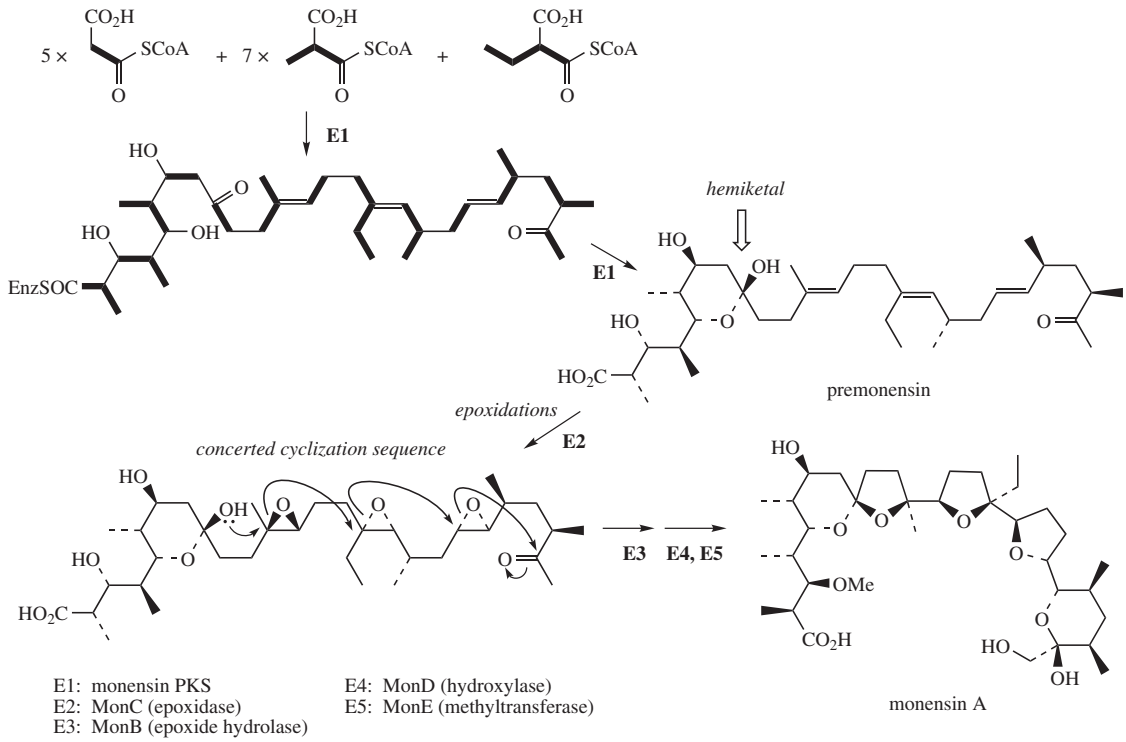


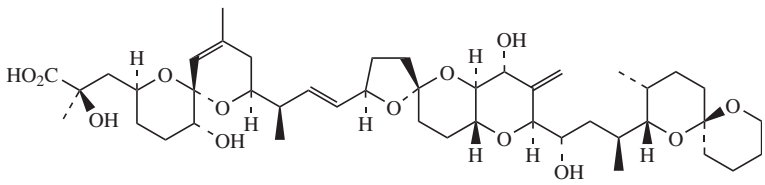
Figure 3.58

antibiotics are characterized by the presence of a number of tetrahydrofuran and/or tetrahydropyran rings along the basic chain. The polyether acts as an ionophore, increasing influx of sodium ions into the parasite, causing a resultant and fatal increase in osmotic pressure. Current thinking is that these ring systems arise via a cascade cyclization mechanism involving epoxide intermediates, as shown for lasalocid A (Figure 3.57). In the biosynthesis of **monensin A** (Figure 3.58), a modular PKS assembles a polyketide chain from malonate, methylmalonate, and ethylmalonate precursors to produce the triene shown, which is released as the hemiketal; the loading module contains a KS^Q domain and the starter is malonate (see page 73). Following triepoxide formation, a concerted stereospecific cyclization sequence initiated by the hemiketal hydroxyl could proceed as indicated. Hydroxylation of the starter methyl group and methylation of a hydroxyl are late steps.

Even more remarkable polyether structures are found in some toxins produced by marine dinoflagellates, which are in turn taken up by shellfish, thus passing on their toxicity to the shellfish. **Okadaic acid** (Figure 3.59) and related polyether structures from *Dinophysis* species are responsible for diarrhoeic shellfish poisoning in mussels, causing severe diarrhoea, nausea, and vomiting in consumers of contaminated shellfish in many parts of the world. Okadaic acid is now known to be a potent inhibitor of protein phosphatases, and causes dramatic increase in the phosphorylation of many proteins, leading to the observed cellular responses. It has become a useful pharmacological tool in the study of protein phosphorylation processes. **Brevetoxin A** and **brevetoxin B** (Figure 3.60) are examples of the toxins associated with 'red tide' blooms of dinoflagellates which affect fishing and also tourism, especially in Florida and the Gulf of Mexico. The red tide toxins are derived from *Karenia brevis* (formerly *Gymnodimium breve*) and are the causative agents of neurotoxic shellfish poisoning, leading to neurological disorders and gastrointestinal troubles. The toxins are known to bind to sodium channels, keeping them in an open state. Fatalities among marine life, e.g. fish, dolphins, and whales, and in humans are associated with these toxins synthesized by organisms at the base of the marine food chain. These compounds are postulated to be produced from a polyunsaturated fatty acid by epoxidation of double bonds, and then a concerted sequence of epoxide ring openings leading to the extended polyether structure of *trans*-fused rings (Figure 3.60). The carbon skeletons of okadaic acid and the brevetoxins do not seem to arise from acetate/propionate units in the systematic manner as seen with monensin A; biosynthetic studies have shown that fragments from the citric acid cycle and other precursors become involved. This may reflect some significant differences between terrestrial and marine microorganisms.

Ciguatoxin (Figure 3.60) is one of the more complex examples of a polyether structure found in nature, though it is certainly not the largest. This compound is found in the moray eel (*Gymnothorax javanicus*) and in a variety of coral reef fish, such as red snapper (*Lutjanus bohar*), though some 100 different species of fish may cause food poisoning (ciguatera) in tropical and subtropical regions. Ciguatoxin is remarkably toxic even at microgram levels; ciguatera is characterized by vomiting, diarrhoea, and neurological problems. Most sufferers slowly recover, though this may take a considerable time, and few cases are fatal due principally to the very low levels of toxin actually present in the fish. The dinoflagellate *Gambierdiscus toxicus* is ultimately responsible for polyether production, synthesizing a less toxic analogue which is passed through the food chain and eventually modified into the very toxic ciguatoxin by the fish. More than 20 ciguatoxin-like structures have been identified.

Halichondrins are both polyethers and macrolides, and were originally isolated from the Pacific marine sponge *Halichondria okadai*. They have received considerable attention because of their complex structures and their extraordinary antitumour activity. **Halichondrin B** (Figure 3.61) was shown to affect tubulin polymerization, though at a site distinct from the *Catharanthus* alkaloid site (see page 375). To



okadaic acid

Figure 3.59

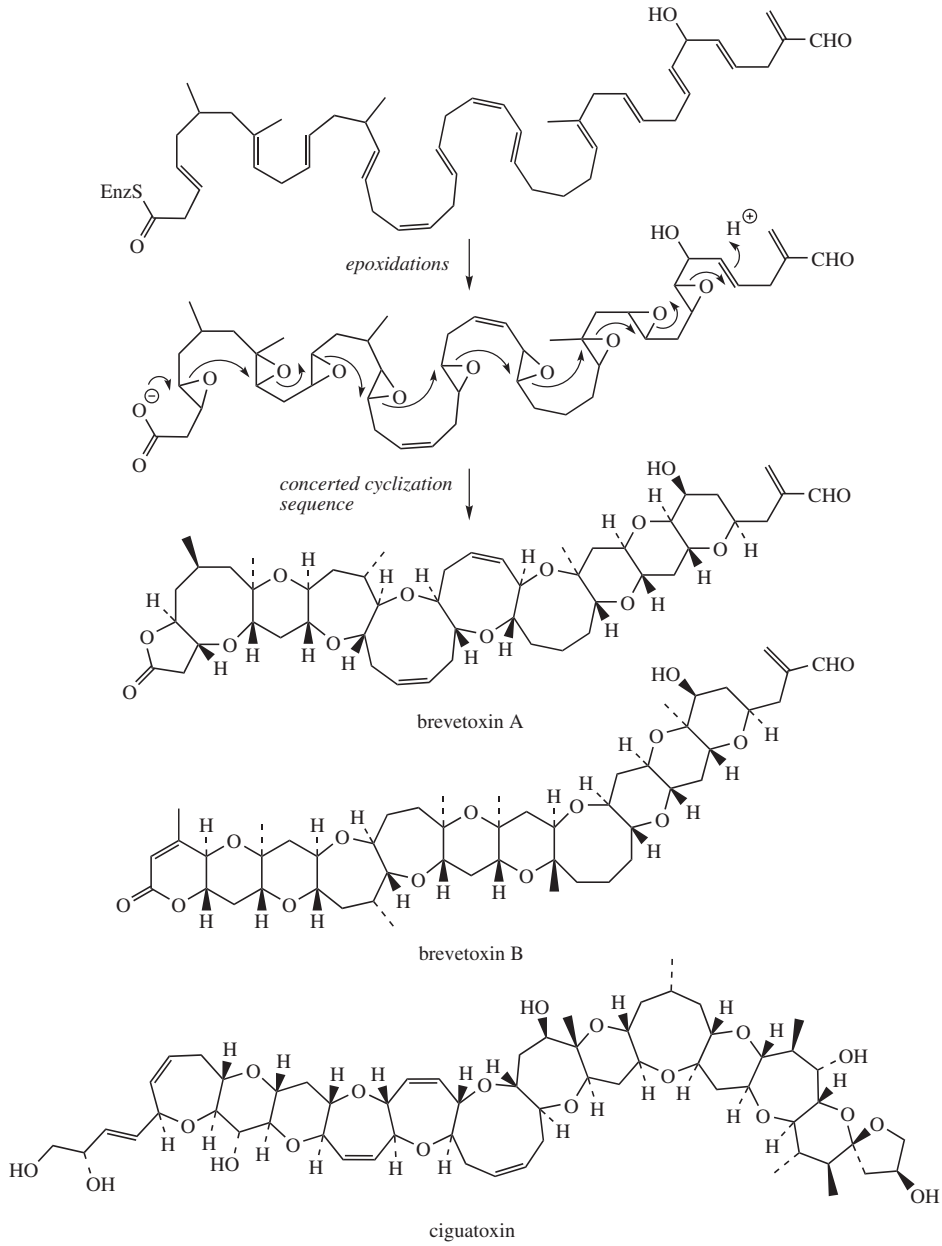


Figure 3.60

facilitate clinical studies, halichondrins were produced by total synthesis, leading to the discovery that it was only necessary to have the macrocyclic portion to maintain the biological activity; the macrocyle still contains several complex ether linkages, however. Substitution of a ketone for the lactone function increased stability; in due course, the modified derivative **eribulin** (Figure 3.61) has proved successful in the treatment of advanced breast cancer and has been introduced as a drug.

The **zaragozic acids** (**squalestatins**) are structurally rather different from the polyethers discussed so far, but

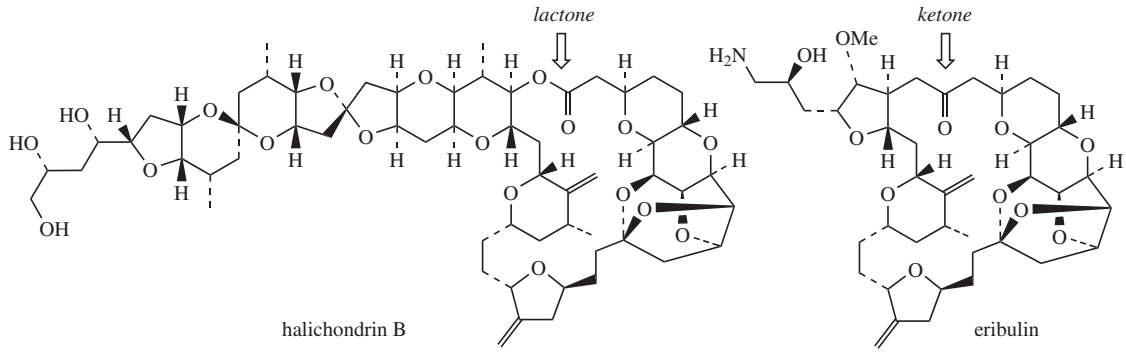


Figure 3.61

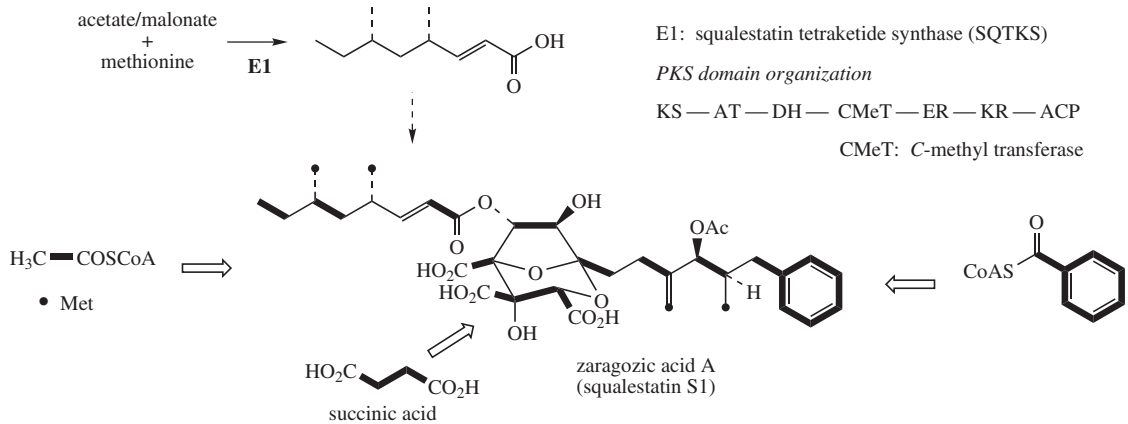


Figure 3.62

they are primarily acetate derived, and the central ring system is suggested to be formed by an epoxide-initiated process resembling the polyether derivatives just described. Thus, **zaragozic acid A** (squalestatin S1; Figure 3.62) is known to be constructed from two acetate-derived chains and a C_4 unit such as the Krebs cycle intermediate oxaloacetate (succinic acid can act as precursor). One chain has a benzoyl-CoA starter (from the shikimate pathway, see page 157), and both contain two methionine-derived side-chain substituents (Figure 3.62). An iterative (repeating) type I PKS catalysing synthesis of one of these has been characterized. In common with a number of fungal PKSs, a *C*-methyl transferase domain is included in the protein, and *C*-methylation is part of the PKS activity (see also epothilones, page 86, and mupirocin, page 91). Iterative PKSs do not have the modular system to control the order and composition of extender units added. Instead, programming of the sequence is encoded in the PKS itself, in a manner that remains to be clarified.

The heterocyclic ring system can be envisaged as arising via nucleophilic attack onto oxaloacetic acid, formation of a diepoxide, then a concerted sequence of reactions as indicated (Figure 3.63). The zaragozic acids are produced by a number of fungi, including *Sporomiella intermedia* and *Leptodontium elatius*, and are of considerable interest since they are capable of reducing blood cholesterol levels in animals by acting as potent inhibitors of the enzyme squalene synthase (see page 235). This is achieved by mimicking the steroid precursor presqualene PP and irreversibly inactivating the enzyme; both compounds have lipophilic side-chains flanking a highly polar core. They have considerable medicinal potential for reducing the incidence of coronary-related deaths (compare the statins, below).

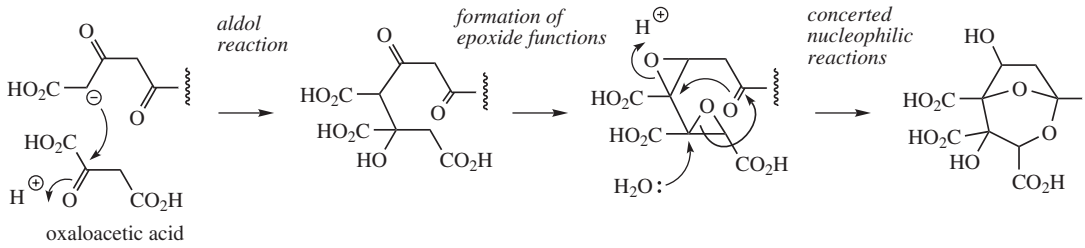


Figure 3.63

DIELS–ALDER CYCLIZATIONS

A number of cyclic structures, typically containing cyclohexane rings, are known to be formed via the acetate pathway and can be rationalized as involving an enzymic Diels–Alder reaction (Figure 3.64), though only one example will be considered here.

Thus, **lovastatin** arises from a C_{18} polyketide chain with *C*-methylation, and relatively few of the oxygen functions are retained in the final product. A secondary polyketide chain, with methylation, provides the ester side-chain (Figure 3.65). A Diels–Alder reaction accounts for formation of the decalin system as shown, and this is now known to occur at the hexaketide intermediate stage, catalysed by the PKS itself, though it is not known just how this is accomplished. The PKS (lovastatin nonaketide synthase) is an iterative type I system and carries a *C*-methyltransferase domain. Unusually, though the PKS carries an ER domain, this function is not activated unless an accessory protein LovC is also present. The product of lovastatin nonaketide synthase is the lactone dihydromonacolin L, and this is subjected to a sequence of post-PKS modifications, leading to monacolin J. The ester side-chain is produced by a non-iterative type I PKS, lovastatin diketide synthase, which also carries a *C*-methyltransferase domain. It carries no TE domain, and the fatty acyl chain appears to be delivered directly from the PKS protein in conjunction with the esterifying enzyme.

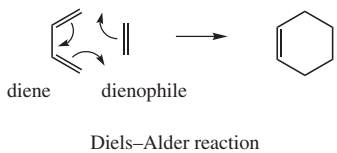


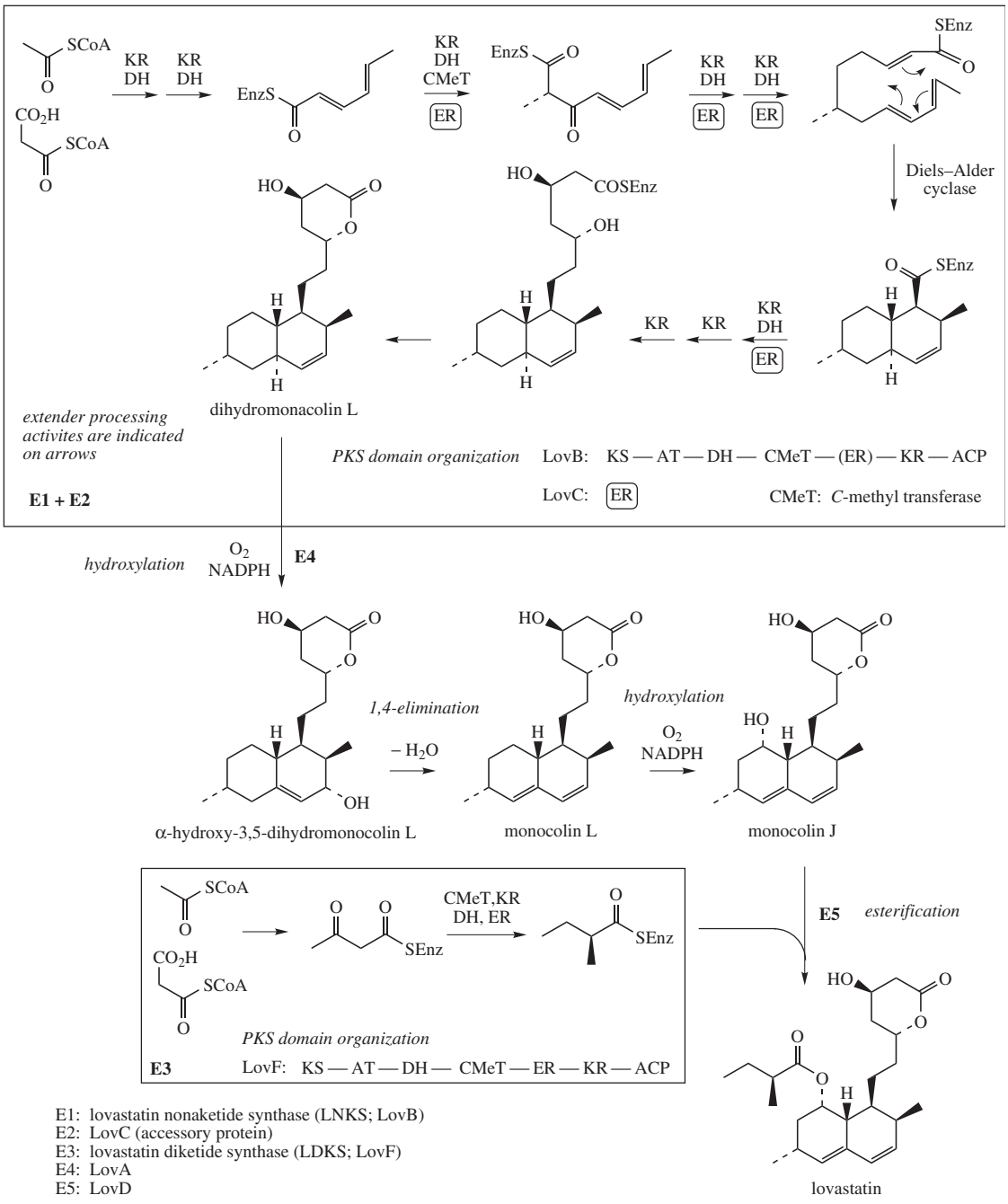
Figure 3.64

The *C*-methylation processes, quite common with fungal PKS systems, are easily interpreted as in Figure 3.66. The ketoester intermediate may be methylated with SAM via an enolate anion mechanism; the product then re-enters the standard chain extension process.

Lovastatin was isolated from cultures of *Aspergillus terreus* and was found to be a potent inhibitor of hydroxymethylglutaryl-CoA (HMG-CoA) reductase, a rate-limiting enzyme in the mevalonate pathway (see page 189). Analogues of lovastatin (statins) find drug use as HMG-CoA reductase inhibitors, thus lowering blood cholesterol levels in patients [Box 3.11]. These drugs have rapidly become the top-selling drugs worldwide.

POLYKETIDE SYNTHASES: AROMATICS

In the absence of any reduction processes, the growing poly-β-keto chain needs to be stabilized on the enzyme surface until the chain length is appropriate, at which point cyclization or other reactions can occur. A poly-β-keto ester is very reactive, and there are various possibilities for undergoing intramolecular Claisen or aldol reactions, dictated of course by the nature of the enzyme and how the substrate is folded. Methylenes flanked by two carbonyls are activated, allowing formation of carbanions/enolates and subsequent reaction with ketone or ester carbonyl groups, with a natural tendency to form strain-free six-membered rings. Aromatic compounds are typical products from type II and type III PKSs, though there are a few examples where type I PKSs produce aromatic rings. Whilst the chemical aspects of polyketide synthesis are common to all systems, a very significant difference compared with most of the type I systems described under macrolides is that the processes are iterative rather than non-iterative. In a non-iterative system, the sequence of events is readily related to the sequence of enzymic domains as each is brought into play. It is less



E1: lovastatin nonaketide synthase (LNKS; LovB)
E2: LovC (accessory protein)
E3: lovastatin diketide synthase (LDKS; LovF)
E4: LovA
E5: LovD

Figure 3.65

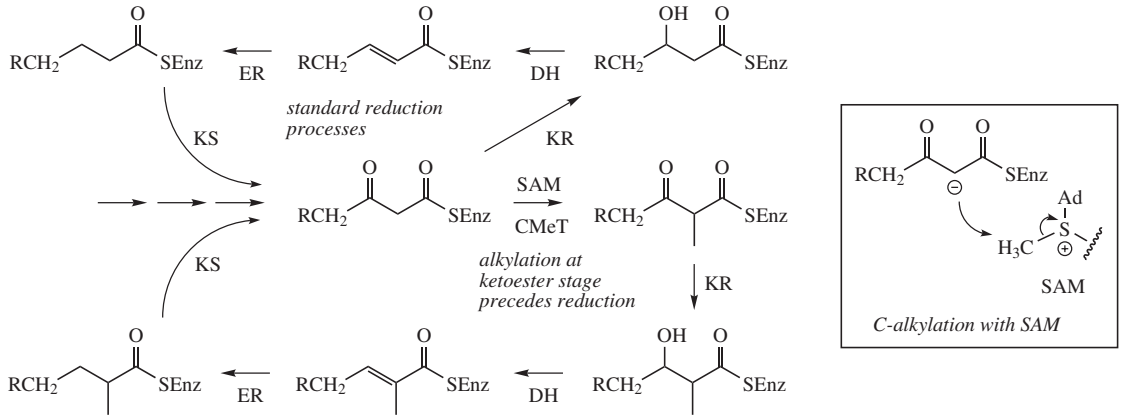


Figure 3.66

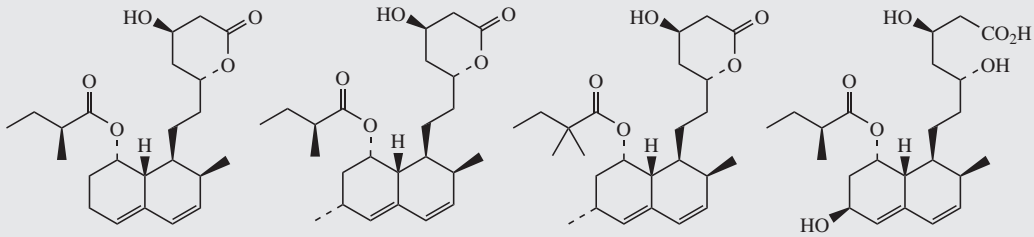
Box 3.11

Statins

Mevastatin (formerly compactin; Figure 3.67) is produced by cultures of *Penicillium citrinum* and *Penicillium brevicompactum*, and was shown to be a reversible competitive inhibitor of HMG-CoA reductase, dramatically lowering sterol biosynthesis in mammalian cell cultures and animals, and reducing total and low-density lipoprotein (LDL) cholesterol levels (see page 251). Mevastatin in its ring-opened form (Figure 3.68) mimics the half-reduced substrate mevaldate hemithioacetal during the two-stage reduction of HMG-CoA to mevalonate (see page 190), and the affinity of this agent towards HMG-CoA reductase is some 10^4-fold greater than that of the normal substrate. High blood cholesterol levels contribute to the incidence of coronary heart disease, so mevastatin and analogues are of potential value in treating high-risk coronary patients, and several agents are already in use. These are generically known as statins, and they have rapidly become the top-selling drugs worldwide.

As the liver synthesizes less cholesterol, this in turn stimulates the production of high-affinity LDL receptors on the surface of liver cells. Consequently, the liver removes more LDL from the blood, leading to the reduction of blood levels of both LDL and cholesterol. The clinical use of statins has also revealed that these drugs promote beneficial effects on cardiovascular functions that do not correlate simply with their ability to lower cholesterol levels. Statins are now routinely prescribed to slow the progression of coronary artery disease and to reduce mortality from cardiovascular disease. They have been suggested to have potentially useful anti-inflammatory and anticancer activities, and some are also being tested against Alzheimer's disease and osteoporosis.

Lovastatin (formerly called mevinolin or monacolin K; Figure 3.67) is produced by *Monascus ruber* and *Aspergillus terreus* and is slightly more active than mevastatin. It was the first statin to be marketed, but has since been superseded by more active



mevastatin | lovastatin (mevinolin; monacolin K) | simvastatin | pravastatin

Figure 3.67

Box 3.11 (continued)

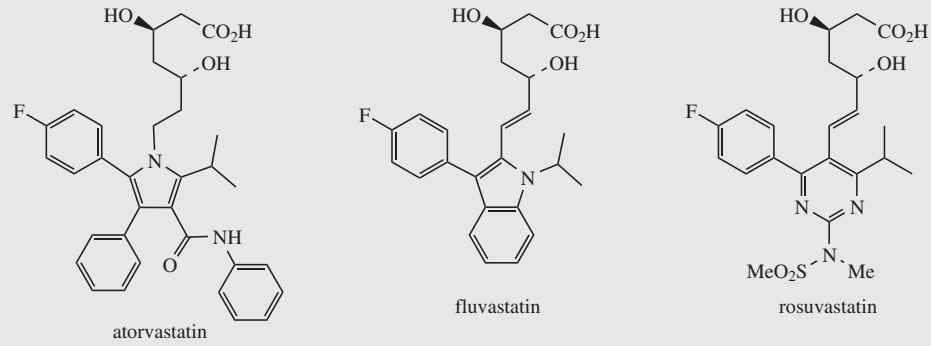


Figure 3.67 (*continued*)

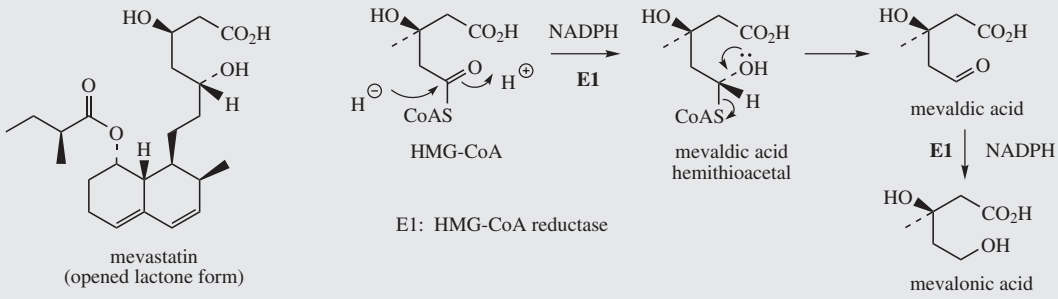


Figure 3.68

agents. **Simvastatin** is obtained fron lovastatin by ester hydrolysis and then re-esterification, and is two to three times as potent as lovastatin. **Pravastatin** is prepared from mevastatin by microbiological hydroxylation using *Streptomyces carbophilus* and is consequently more hydrophilic than the other drugs, with an activity similar to lovastatin. Lovastatin and simvastatin are both lactones, but at physiological pHs they exist in equilibrium with the open-ring hydroxy acids; only the open-ring form is biologically active. Other agents currently in use are synthetic, though they feature the same 3,5-dihydroxycarboxylic acid side-chain as in pravastatin. **Atorvastatin, fluvastatin**, and **rosuvastatin** have all been introduced recently, and others are in development.

clear how a repeating process can be closely programmed to produce the final product, and this is somehow encoded in the PKS itself. Thus, the mechanisms of Figure 3.69 will need to be repeated by reuse of the same domains. Type II systems do contain two KS domains (KS_α and KS_β), though only the KS_α domain is implicated in the Claisen condensation. KS_β lacks the active site cysteine and appears analogous to the KS^Q of some macrolide PKSs (see page 73), so it may be involved in loading and decarboxylation of malonyl-CoA. Additional domains for KR, cyclase, or aromatase activities may be present. Type III systems differ from the others in the use of coenzyme A esters rather than ACPs (Figure 3.69). The PKS is a single protein, though it exists as a homodimer in its native form. It is not appropriate here to attempt subdivisions according to the PKS type, and compounds will be considered primarily according to structural features.

Cyclizations

The poly-β-keto ester (Figure 3.70), formed from four acetate units (one acetate starter group and three malonate chain extension units), is capable of being folded in at least two ways, A and B (Figure 3.70). For A, ionization

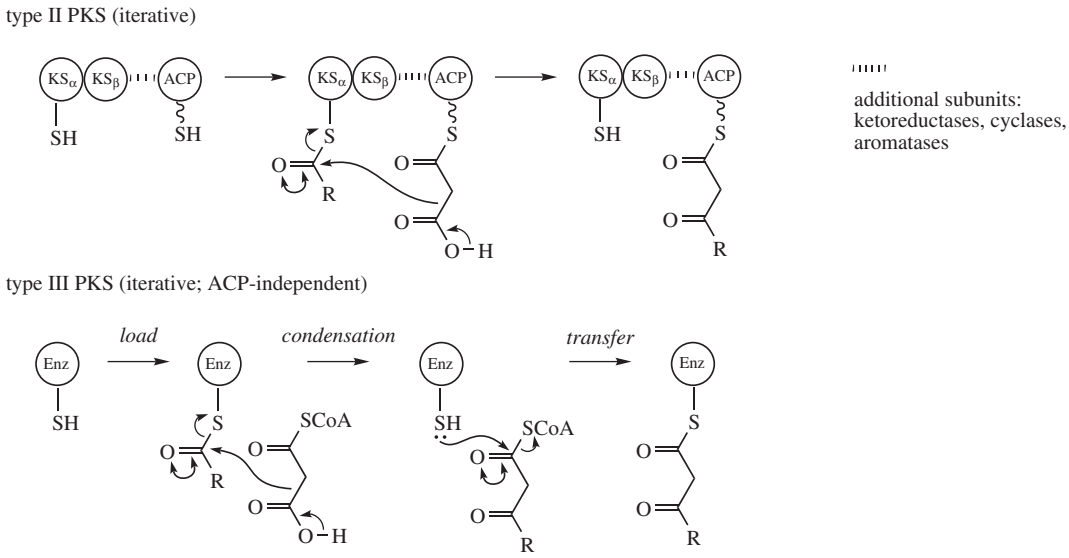


Figure 3.69

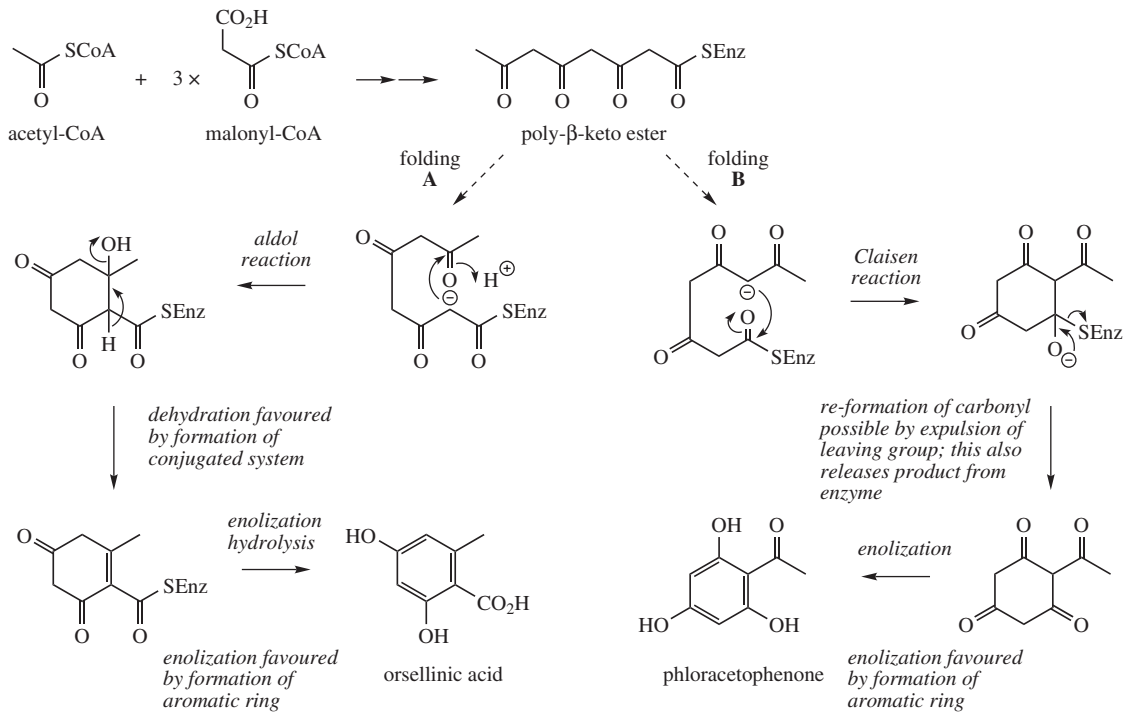


Figure 3.70

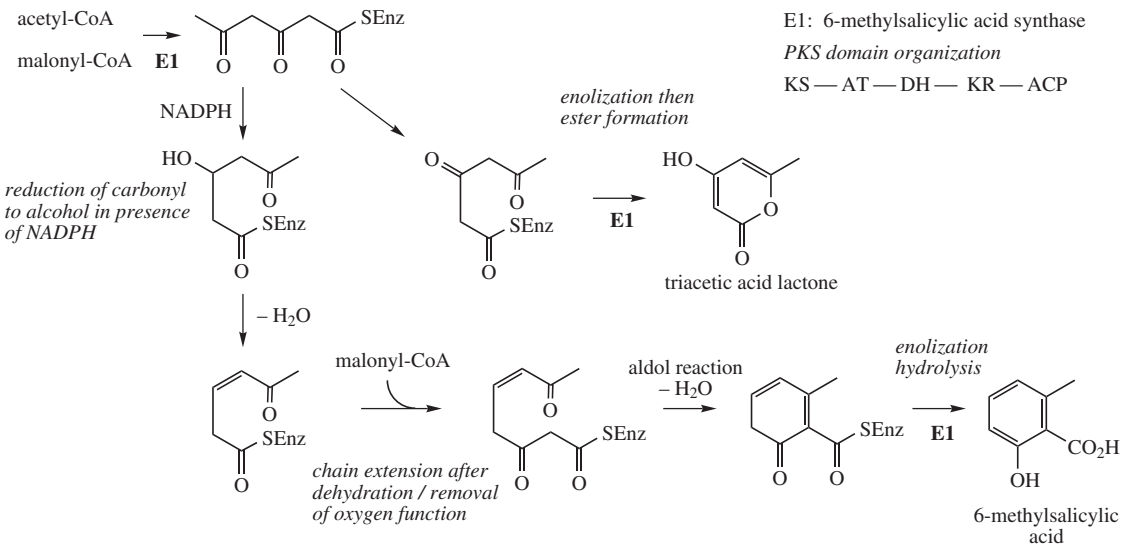


Figure 3.71

of the α-methylene allows aldol addition onto the carbonyl six carbon atoms distant along the chain, giving the tertiary alcohol. Dehydration occurs as in most chemical aldol reactions, giving the alkene, and enolization follows to attain the stability conferred by the aromatic ring. The enzyme–thioester linkage is then hydrolysed to produce **orsellinic acid**. Alternatively, folding of the polyketo ester as in B allows a Claisen reaction to occur, which, although mechanistically analogous to the aldol reaction, is terminated by expulsion of the thiol leaving group and direct release from the enzyme. Enolization of the cyclohexatrione produces **phloracetophenone**. As with other PKSs, the whole sequence of reactions is carried out by an enzyme complex which converts acetyl-CoA and malonyl-CoA into the final product without giving any detectable free intermediates, thus combining PKS and polyketide cyclase activities.

A distinctive feature of an aromatic ring system derived through the acetate pathway is that several of the carbonyl oxygens of the poly-β-keto system are retained in the final product. These end up on alternate carbon atoms around the ring system. Of course, one or more might be used in forming a carbon–carbon bond, as in orsellinic acid. Nevertheless, this oxygenation on alternate carbon atoms, a *meta* oxygenation pattern, is usually easily recognizable and points to the biosynthetic origin of the molecule. This *meta* oxygenation pattern contrasts with that seen on aromatic rings formed via the shikimate pathway (see page 140).

6-Methylsalicylic acid (Figure 3.71) is a metabolite of *Penicillium patulum*, and differs from orsellinic acid by the absence of a phenol group at position 4. It is also derived from acetyl-CoA and three molecules of malonyl-CoA, and the 'missing' oxygen function is removed during the biosynthesis. Orsellinic acid is not itself deoxygenated to 6-methylsalicylic acid. The enzyme 6-methylsalicylic acid synthase is one of the smallest PKSs known and is of the iterative type I group. It requires NADPH as cofactor and removes the oxygen function by employing its KR domain just once. Reduction of the ketone to an alcohol, followed by a dehydration step, occurs on a six-carbon intermediate after the second chain extension (Figure 3.71). Further chain extension is then followed by aldol condensation, enolization, and release from the enzyme to generate 6-methylsalicylic acid. Important evidence for reduction occurring at the C_6 stage comes from the formation of triacetic acid lactone (Figure 3.71) as a 'derailment product' if NADPH is omitted from the enzymic incubation. The PKS is iterative, the enzyme activities acting repeatedly, but there is some programming that allows KR to participate only after the second extension round. In the absence of NADPH, KR remains inactive. There is no TE domain; perhaps release from the enzyme is not simple hydrolysis.

The folding of a polyketide chain can be established by labelling studies, feeding carbon-labelled sodium acetate to the appropriate organism and establishing the position of labelling in the final product by chemical degradation and counting (for the radioactive isotope ^{14}C)

or by NMR spectrometry (for the stable isotope ^{13}C). ^{13}C NMR spectrometry is also valuable in establishing the location of intact C_2 units derived from feeding $^{13}C_2$-labelled acetate (see page 36). This is exemplified in Figure 3.72, where **alternariol**, a metabolite from the mould *Alternaria tenuis*, can be established to be derived from a single C_{14} polyketide chain, folded as shown, and then cyclized. Whilst the precise sequence of reactions involved is not known, paper chemistry allows us to formulate the essential features. Two aldol condensations followed by enolization in both rings would give a biphenyl, and lactonization would then lead to alternariol. The oxygenation pattern in alternariol shows alternate oxygen atoms on both aromatic rings, and an acetate origin is readily surmised, even though some oxygen atoms have been used in ring formation processes. The lone methyl 'start-of-chain' is also usually very obvious in acetate-derived compounds, though the carboxyl 'end-of-chain' can often react with convenient hydroxyl functions, which may have arisen through enolization, and lactone or ester functions are thus reasonably common. For example, the lichen metabolite **lecanoric acid** is a **depside** (an ester formed from two phenolic acids) and the result from combination of two orsellinic acid thioester molecules (Figure 3.73). The lactone ring in **6-hydroxymellein** (Figure 3.74) is formed by reaction of the end-of-chain thioester with a hydroxyl function in the starter group. The original keto group is reduced during chain extension, at the C_6 stage, as noted in the formation of 6-methylsalicylic acid. The same derailment metabolite, triacetic acid lactone, is produced by the enzyme in the absence of NADPH. Unusually, the carbonyl group reduced is that of the starter unit rather than of the extender. SAM-dependent methylation of one of the hydroxyls leads to **6-methoxymellein**, an antifungal stress metabolite of carrot (*Daucus carota*; Umbelliferae/Apiaceae). 6-Hydroxymellein synthase is a homodimeric protein, and may well be a type III plant PKS; type III systems utilize coenzyme A esters rather than ACPs.

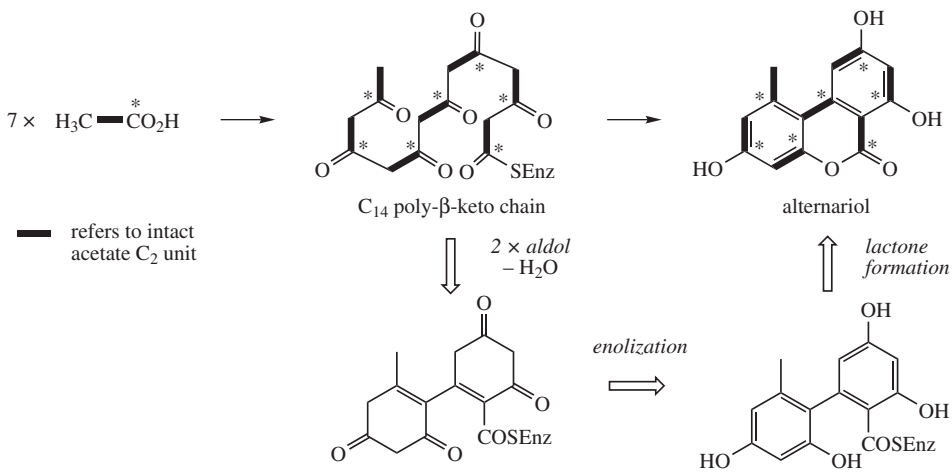


Figure 3.72

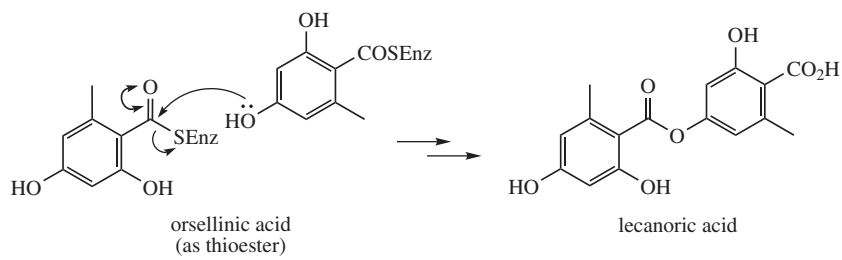


Figure 3.73

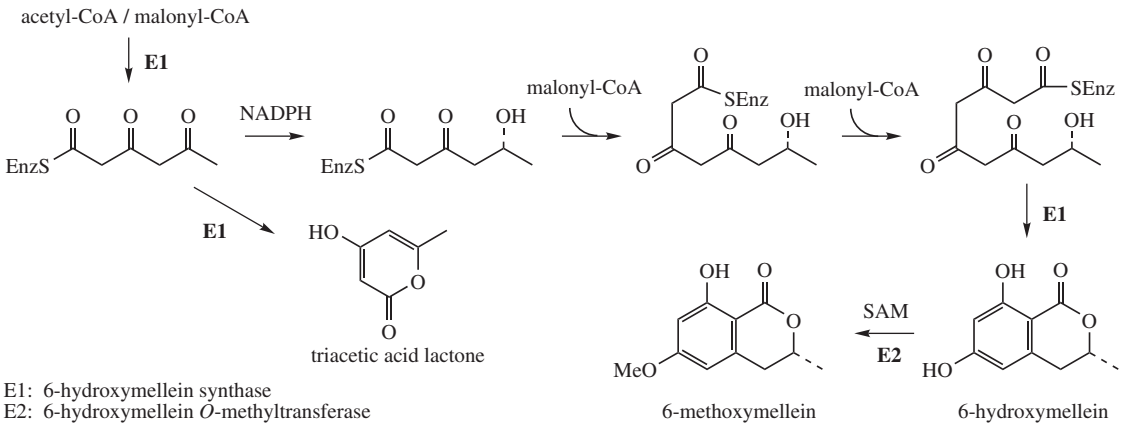


Figure 3.74

Post-Polyketide Synthase Modifications

A number of natural anthraquinone derivatives are also excellent examples of acetate-derived structures. Structural analysis of **endocrocin** (Figure 3.75) found in species of *Penicillium* and *Aspergillus* fungi suggests folding of a polyketide containing eight C_2 units to form the periphery of the carbon skeleton. Three aldol-type condensations would give a hypothetical intermediate 1, and, except for a crucial carbonyl oxygen in the centre ring, endocrocin results by enolization reactions, one of which involves the vinylogous enolization —CH_2—CH═CH—CO— → —CH═CH—CH═C(OH)—. The additional carbonyl oxygen must be introduced at some stage during the biosynthesis by an oxidative process, for which we have little information. **Emodin**, a metabolite of some *Penicillium* species, but also found in higher plants, e.g. *Rhamnus* and *Rumex* species, would appear to be formed from endocrocin by a simple decarboxylation reaction. This is facilitated by the adjacent phenol function (see page 22). *O*-Methylation of emodin would then lead to **physcion. Islandicin** is another anthraquinone pigment produced by *Penicillium islandicum*, and differs from emodin in two ways: one hydroxyl is missing and a new hydroxyl has been incorporated adjacent to the methyl. Without any evidence for the sequence of such reactions, the structure of hypotheical intermediate 2 shows the result of three aldol condensations with reduction of a carbonyl, though the latter would most likely occur during chain extension. A dehydration reaction, two oxidations, and a decarboxylation are necessary to attain the islandicin structure. In **chrysophanol, aloe-emodin** and **rhein**, the same oxygen function is lost by reduction as in islandicin, and decarboxylation also occurs. The three compounds are interrelated by a sequential oxidation of the methyl in chrysophanol to a hydroxymethyl in aloe-emodin and a carboxyl in rhein.

As noted in the case of macrolides, structural modifications undergone by the basic polyketide are also conveniently considered according to the timing of the steps in the synthetic sequence. Thus, 'missing' oxygen functions appear to be reduced out well before the folded and cyclized polyketide is detached from the enzyme, and are mediated by the KR component of the PKS enzyme complex. On the other hand, reactions like decarboxylation, *O*-methylation, and sequential oxidation of a methyl to a carboxyl are representative of post-PKS transformations occurring after the cyclization reaction. These later conversions can often be demonstrated by the isolation of separate enzymes catalysing the individual steps. Most of the secondary transformations are easily rationalized by careful consideration of the reactivity conferred on the molecule by the alternating and usually phenolic oxygenation pattern. These oxygen atoms activate adjacent sites, creating nucleophilic centres. Introduction of additional hydroxyl groups *ortho* or *para* to an existing phenol will be facilitated (see page 26), allowing the extra hydroxyl of islandicin to be inserted, for example. The *ortho*- or *para*-diphenols are themselves susceptible to further oxidation in certain circumstances, and may give rise to *ortho*- and *para*-quinones (see page 26). The quinone system in anthraquinones is built up by an oxidation of the central cyclohexadienone ring, again at a nucleophilic centre activated by the enone system. Methyl groups on an aromatic ring are also activated towards oxidation, facilitating the chrysophanol → aloe-emodin oxidation, for example. Decarboxylation, e.g. endocrocin → emodin,

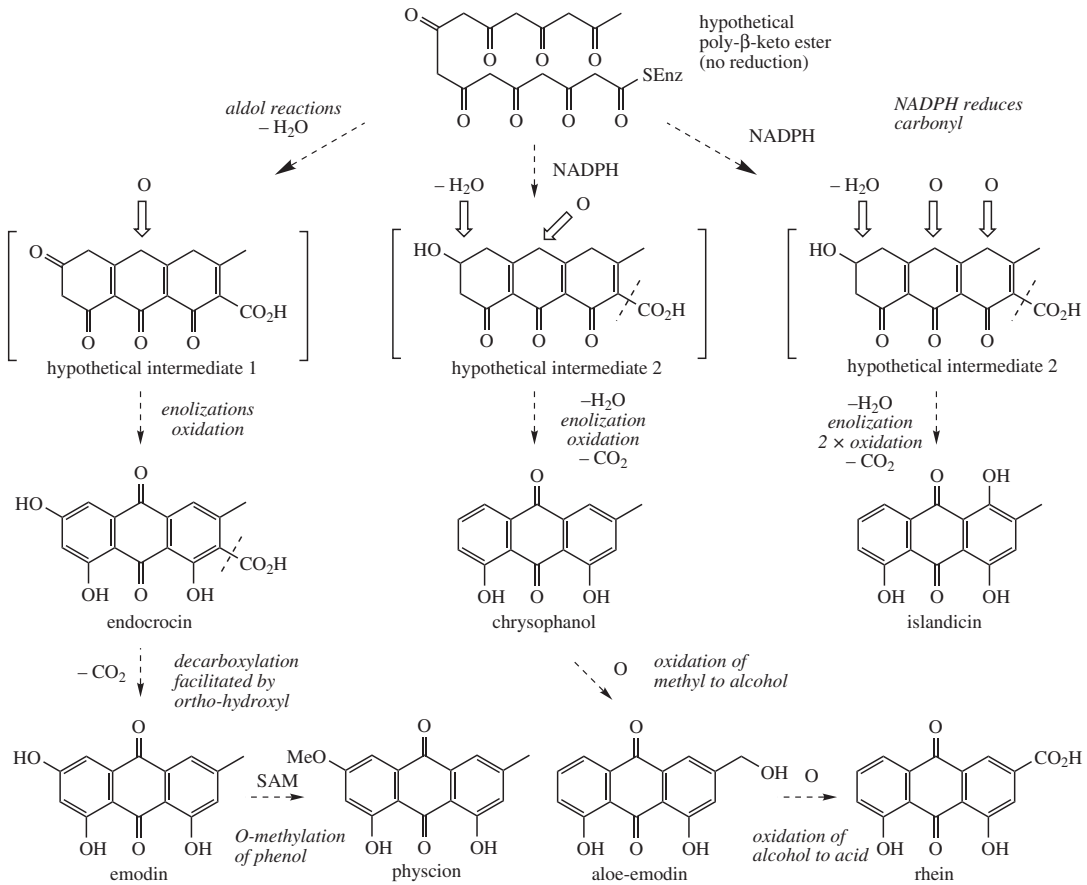


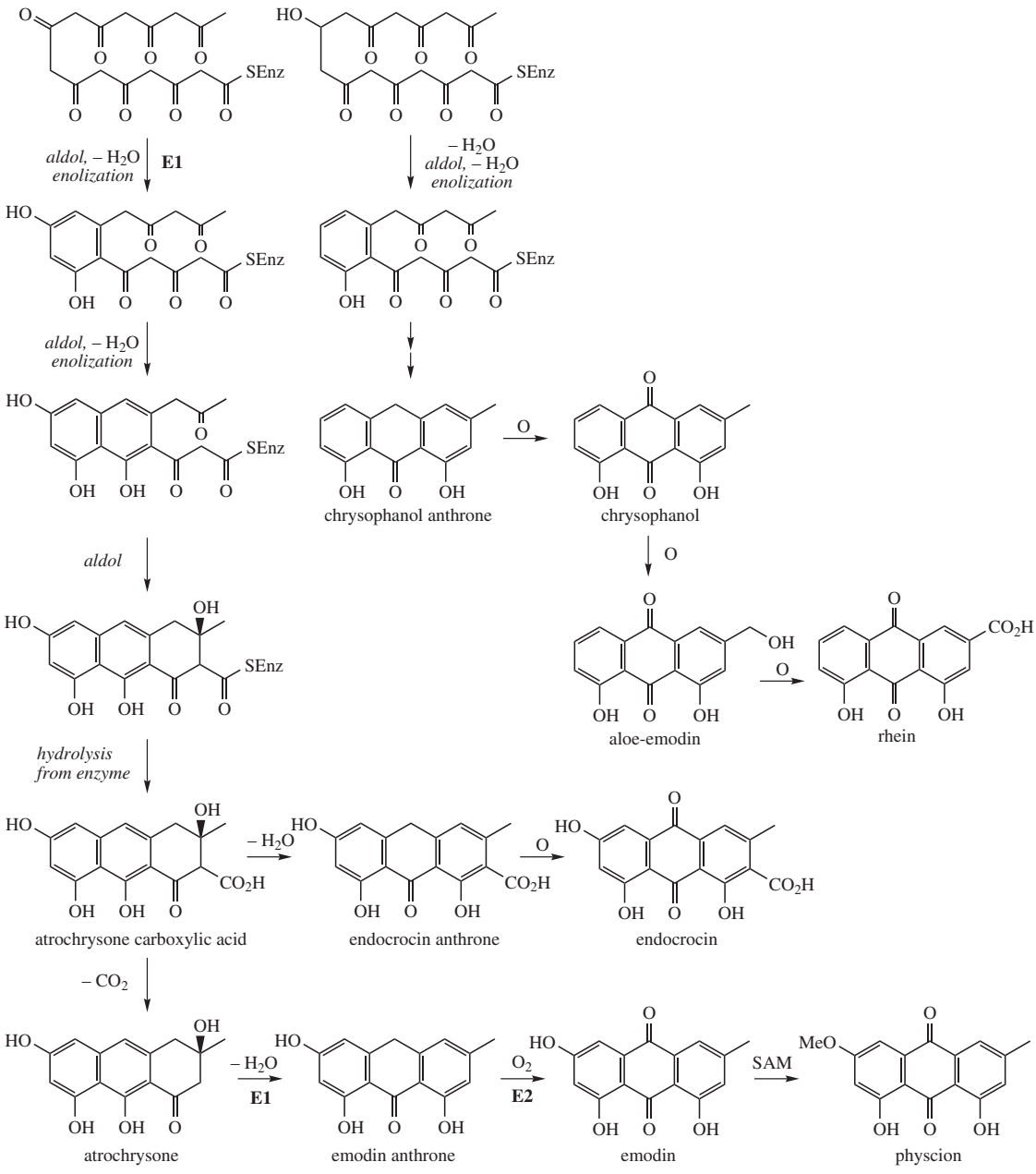
Figure 3.75

is readily achieved in the presence of an *ortho* phenol function, though a *para* phenol can also facilitate this (see page 22).

It is now appreciated that the assembly of the anthraquinone skeleton (and related polycyclic structures) is achieved in a stepwise sequence. After the polyketide chain is folded, the ring at the centre of the fold is formed first, followed in turn by the next two rings. The pathway outlined for the biosynthesis of endocrocin and emodin is shown in Figure 3.76. Mechanistically, there is little difference between this and the speculative pathway of Figure 3.75, but the sequence of reactions is altered. Decarboxylation appears to take place before aromatization of the last-formed ring system, and tetrahydroanthracene intermediates, such as atrochrysone carboxylic acid and atrochrysone, are involved. These dehydrate to the anthrones **endocrocin anthrone** and **emodin anthrone** respectively prior to introduction of the extra carbonyl oxygen as a last transformation in the production of anthraquinones. This oxygen is derived from O_2.

The interrelationships of Figure 3.76 are derived mainly from feeding experiments in fungi. A type III octaketide synthase gene from the anthraquinone-producing plant *Aloe arborescens*, when expressed in *Escherichia coli*, did not synthesize anthraquinones, but instead the two products SEK 4 and SEK 4b, not normally found in *Aloe* (Figure 3.77). These obviously result from different folding arrangements of the polyketide precursor. Thus, the expressed protein seems to lack some agent to fold the chain correctly or perhaps a cyclase element; the observed products may result from non-enzymic cyclization reactions. This protein utilizes just malonyl-CoA, so the starter unit is generated by decarboxylation (see page 73).

Emodin, physcion, chrysophanol, aloe-emodin, and rhein form the basis of a range of purgative anthraquinone



E1: emodin anthrone synthase
E2: emodin anthrone oxygenase

Figure 3.76

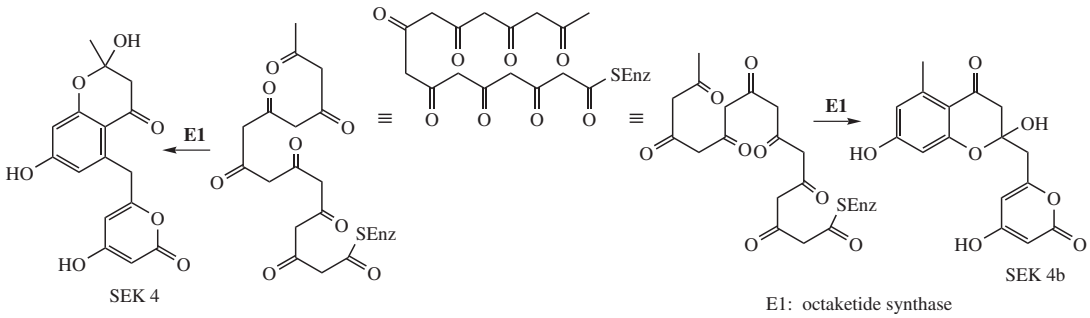


Figure 3.77

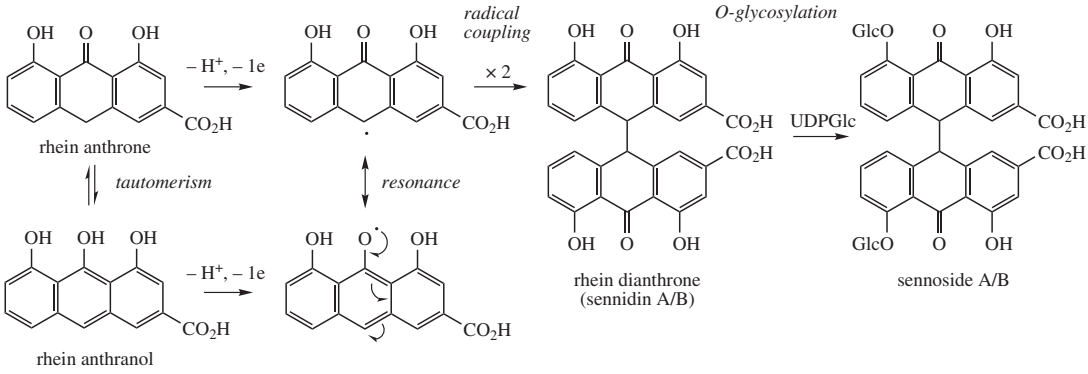


Figure 3.78

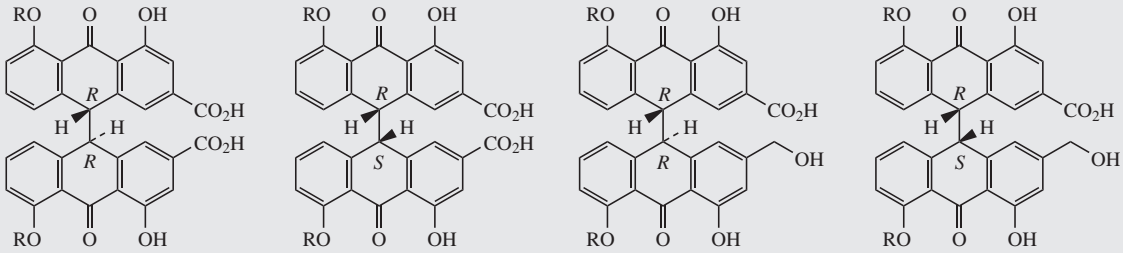
derivatives found in long-established laxatives such as senna [Box 3.12]. The free anthraquinones themselves have little therapeutic activity and need to be in the form of water-soluble glycosides to exert their action. Although simple anthraquinone *O*-glycosides are present in the drugs, the major purgative action arises from compounds such as **sennosides**, e.g. sennosides A and B (Figure 3.78), which are dianthrone *O*-glycosides. These types of derivative are likely to be produced from intermediate anthrone structures. A one-electron oxidation allows oxidative coupling (see page 28) of two anthrone systems to give a dianthrone (Figure 3.78). This can be formulated as direct oxidation at the benzylic $–CH_2–$, or via the anthranol, which is the phenolic tautomer of the anthrone. Glycosylation of the dianthrone system would then give a sennoside-like product; alternatively, glycosylation may precede the coupling.

Box 3.12

Senna

Senna leaf and fruit are obtained from *Cassia angustifolia* (Leguminosae/Fabaceae), known as Tinnevelly senna, or less commonly from *Cassia senna* (*syn Cassia acutifolia*), which is described as Alexandrian senna. The plants are low, branching shrubs, *Cassia angustifolia* being cultivated in India and Pakistan, and *Cassia senna* being produced in the Sudan, much of it from wild plants. Early harvests provide leaf material, whilst both leaf and fruit (senna pods) are obtained later on. There are no significant differences in the chemical constituents of the two sennas, or between leaf and fruit drug. However, amounts of the active constituents do vary, and this appears to be a consequence of cultivation conditions and the time of harvesting of the plant material.

Box 3.12 (continued)



R = Glc, sennoside A
R = H, sennidin A

R = Glc, sennoside B
R = H, sennidin B

R = Glc, sennoside C
R = H, sennidin C

R = Glc, sennoside D
R = H, sennidin D

Figure 3.79

The active constituents in both senna leaf and fruit are dianthrone glycosides, principally sennosides A and B (Figure 3.79). These compounds are both di-*O*-glucosides of rhein dianthrone (sennidins A and B) and liberate these aglycones on acid hydrolysis; oxidative hydrolysis (e.g. aqueous HNO_3 or H_2O_2/HCl) produces the anthraquinone rhein (Figure 3.76). Sennidins A and B are optical isomers: sennidin A is dextrorotatory (+) whilst sennidin B is the optically inactive *meso* form. Minor constituents include sennosides C and D (Figure 3.79), which are also a pair of optical isomers, di-*O*-glucosides of heterodianthrones sennidins C and D. Sennidin C is dextrorotatory, whilst sennidin D is optically inactive, approximating to a *meso* form in that the modest change in substituent does not noticeably affect the optical rotation. Oxidative hydrolysis of sennosides C and D would produce the anthraquinones rhein and aloe-emodin (Figure 3.76). Traces of other anthraquinone glycoside derivatives are also present in the plant material. Much of the sennoside content of the dried leaf appears to be formed by enzymic oxidation of anthrone glycosides during the drying process. Fresh leaves and fruits also seem to contain primary glycosides which are more potent than sennosides A and B, and which appear to be partially hydrolysed to sennosides A and B (the secondary glycosides) by enzymic activity during collection and drying. The primary glycosides contain additional glucose residues.

Senna leaf suitable for medicinal use should contain not less than 2.5% dianthrone glycosides calculated in terms of sennoside B. The sennoside content of Tinnevelly fruits is between 1.2 and 2.5%, with that of Alexandrian fruits being 2.5–4.5%. Senna preparations, in the form of powdered leaf, powdered fruit, or extracts, are typically standardized to a given sennoside content. Non-standardized preparations have unpredictable action and should be avoided. Senna is a stimulant laxative and acts on the wall of the large intestine to increase peristaltic movement. After oral administration, the sennosides are transformed by intestinal flora into rhein anthrone (Figure 3.78), which appears to be the ultimate purgative principle. The glycoside residues in the active constituents are necessary for water solubility and subsequent transportation to the site of action. Although the aglycones, including anthraquinones, do provide purgative action, these materials are conjugated and excreted in the urine after oral administration rather than being transported to the colon. Senna is a purgative drug suitable for either habitual constipation or for occasional use, and is widely prescribed.

However, further oxidative steps on a dianthrone, e.g. **emodin dianthrone** (Figure 3.80), can create a dehydrodianthrone and then allow oxidative coupling of the aromatic rings through **protohypericin** to give a naphthodianthrone, in this case **hypericin**. The reactions of Figure 3.80 can also be achieved chemically by passing air into an alkaline solution of emodin anthrone, though yields are low. **Hypericin** is found in cultures of *Dermocybe* fungi and is also a constituent of St John's Wort, *Hypericum perforatum* (Guttiferae/Hypericaceae). A protein obtained by expression of a *Hypericum perforatum* gene in *Escherichia coli* has been observed to transform the anthraquinone emodin into hypericin, presumably by a sequence similar to that shown. However, since emodin rather than emodin anthrone was the substrate, some additional step must be involved. An aldol condensation between emodin and emodin anthrone could provide the dehydrodianthrone as an alternative to the oxidative coupling. *Hypericum perforatum* is a popular herbal medicine in the treatment of depression [Box 3.13]. The naphthodianthrones have no purgative action, but hypericin can act as a photosensitizing agent in a similar manner to furocoumarins (see page 165). Thus, ingestion of hypericin results in an increased absorption of UV light and can

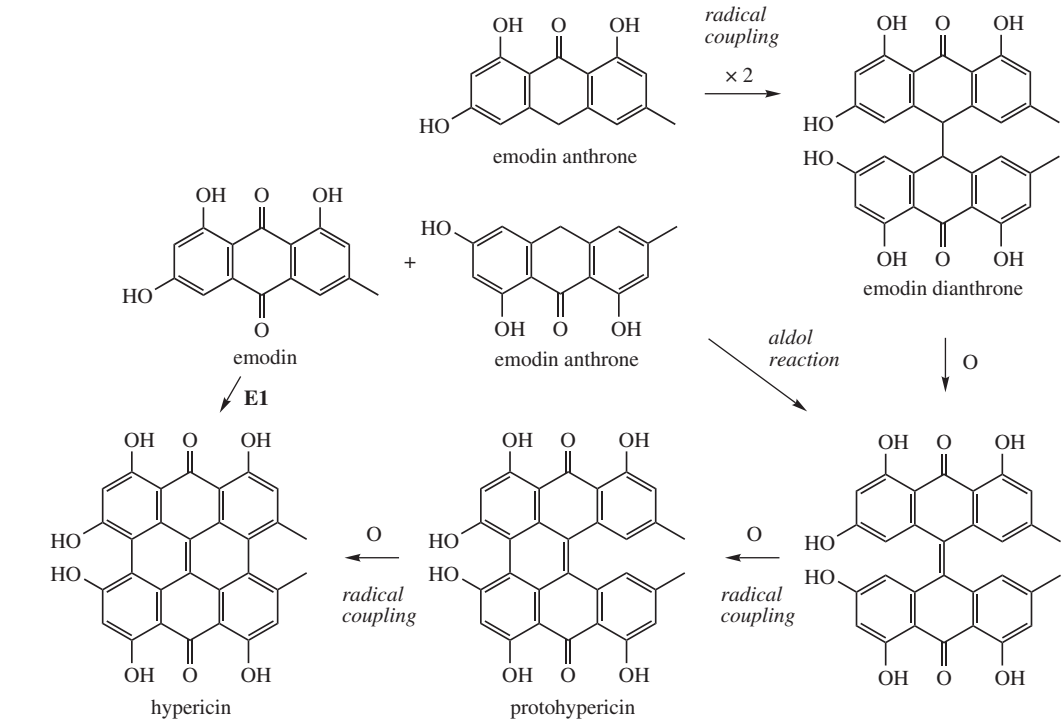


Figure 3.80

Box 3.13

Hypericum/St John's Wort

The dried flowering tops of St John's Wort (*Hypericum perforatum*; Guttiferae/Hypericaceae) have been used as a herbal remedy for many years, an extract in vegetable oil being employed for its antiseptic and wound healing properties. St John's Wort is now a major crop marketed as an antidepressant, which is claimed to be as effective in its action as the widely prescribed antidepressants of the selective serotonin re-uptake inhibitor (SSRI) class such as fluoxetine (Prozac®), and with fewer side-effects. There is considerable clinical evidence that extracts of St John's Wort are effective in treating mild to moderate depression and improving mood. However, to avoid potentially dangerous side-effects, St John's Wort should not be used at the same time as prescription antidepressants (see below). St John's Wort is a small to medium-height herbaceous perennial plant with numerous yellow flowers characteristic of this genus. It is widespread throughout Europe, where it is generally considered a weed, and has also become naturalized in North America. The tops, including flowers at varying stages of development which contain considerable amounts of the active principles, are harvested and dried in late summer.

The dried herb contains significant amounts of phenolic derivatives, including 4–5% of flavonoids (see page 170), though the antidepressant activity is considered to derive principally from naphthodianthrone structures such as hypericin (about 0.1%) and pseudohypericin (about 0.2%), and a prenylated acylphloroglucinol derivative hyperforin (Figure 3.81). The fresh plant also contains significant levels of protohypericin and protopseudohypericin (Figure 3.81), which are converted into hypericin and pseudohypericin during drying and processing, as a result of irradiation with visible light. Hyperforin is a major

Box 3.13 (continued)

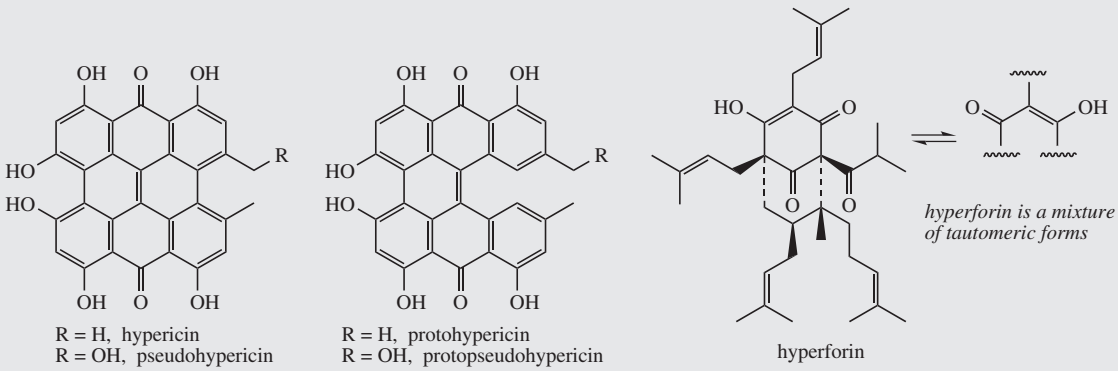


Figure 3.81

lipophilic constituent in the leaves and flowers (2–3%) and is now thought to be the major contributor to the antidepressant activity, as well as to the antibacterial properties of the oil extract. Studies show that the clinical effects of St John's Wort on depression correlate well with hyperforin content. Standardized aqueous ethanolic extracts containing 0.15% hypericin and 5% hyperforin are usually employed. The aqueous solubility of hypericin and pseudohypericin is markedly increased by the presence of flavonoid derivatives in the crude extract, particularly procyanidin B_2, a dimer of epicatechin (see page 171). Hyperforin has been demonstrated to be a powerful and non-selective inhibitor of amine reuptake, thus increasing levels of serotonin, noradrenaline (norepinephrine), and dopamine; the mechanism of its inhibition is different from conventional antidepressants. Through its enolized β-dicarbonyl system, hyperforin exists as a mixture of interconverting tautomers; this reactive system leads to hyperforin being relatively unstable in solution, though salts that are stable and storable may be produced. Smaller amounts of structurally related acylphloroglucinol derivatives are also present in the herb, e.g. adhyperforin (see page 113, Figure 3.86, which also includes the biosynthesis of hyperforin). *Hypericum perforatum* appears to be the only *Hypericum* species where hyperforin is present as a major component.

Hypericin also possesses extremely high toxicity towards certain viruses, a property that requires light and may arise via photo-excitation of the polycyclic quinone system. It is currently under investigation as an antiviral agent against HIV and hepatitis C. Antiviral activity appears to arise from an inhibition of various protein kinases, including those of the protein kinase C family. Hypericin and pseudohypericin are potent photosensitizers initiating photochemical reactions and are held responsible for hypericism, a photodermatosis seen in cattle that have consumed *Hypericum* plants present in pasture. Patients using St John's Wort as an antidepressant need to be warned to avoid overexposure to sunlight.

There is also considerable evidence that St John's Wort interacts with a number of prescription drugs, including oral contraceptives, warfarin (anticoagulant), digoxin (cardiac glycoside), theophylline (bronchodilator), indinavir (HIV protease inhibitor), and ciclosporin (immunosuppressant). In some cases, it is known to promote the cytochrome P-450-dependent metabolism of the co-administered drugs. The action of potentially life-saving drugs may thus be seriously jeopardized by co-administration of *Hypericum* extracts.

lead to dermatitis and burning. Hypericin is also being investigated for its antiviral activities, in particular for its potential activity against HIV.

A common feature of many natural products containing phenolic rings is the introduction of alkyl groups at nucleophilic sites. Obviously, the phenol groups themselves are nucleophilic, and with a suitable alkylating agent, *O*-alkyl derivatives may be formed (see page 13), e.g. the *O*-methylation of emodin to physcion (Figure 3.76). However, a phenol group also activates the ring carbon atoms at the *ortho* and *para* positions, so that these positions similarly become susceptible to alkylation, leading to *C*-alkyl derivatives. The *meta* oxygenation pattern which is a characteristic feature of acetate-derived phenolics has the effect of increasing this nucleophilicity considerably, and the process of *C*-alkylation is very

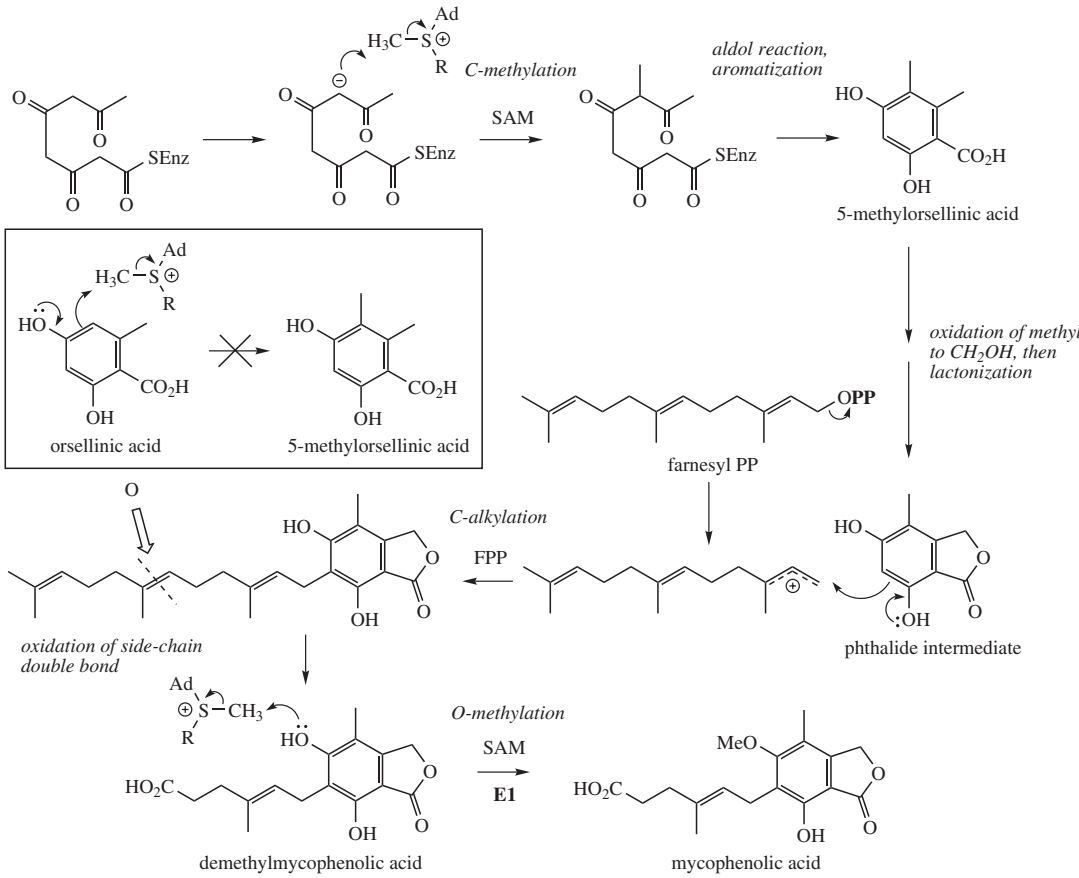


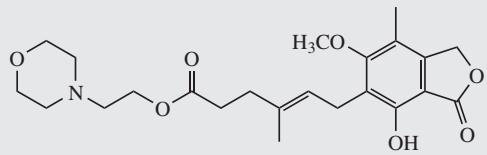
Figure 3.82

much facilitated (see page 13). Suitable natural alkylating agents are SAM and dimethylallyl diphosphate (DMAPP). Other polyprenyl diphosphate esters may also be encountered in biological alkylation reactions (e.g. see hyperforin, page 113, vitamin K, page 182). A minor inconsistency has been discovered, in that while *C*-alkylation with dimethylallyl and higher diphosphates is mediated *after* the initial polyketide cyclization product is liberated from the enzyme, there are examples where *C*-methylation occurs *before* cyclization and is catalysed by the PKS itself (compare macrolide PKSs, page 86). **5-Methylorsellinic acid** (Figure 3.82) is a simple *C*-methylated analogue of orsellinic acid found in *Aspergillus flaviceps*, and the extra methyl is derived from SAM. However, orsellinic acid (see page 101) is not a precursor of 5-methylorsellinic acid and, therefore, it is proposed that the poly-β-keto ester is methylated as part of the reactions catalysed by the PKS (Figure 3.82). Similarly, 5-methylorsellinic acid, but not orsellinic acid, is a precursor of **mycophenolic acid** in *Penicillium brevicompactum* (Figure 3.82) [Box 3.14]. Further *C*-alkylation by farnesyl diphosphate (see page 210) proceeds *after* the aromatization step, and a phthalide intermediate is the substrate involved. The phthalide is a lactone derived from 5-methylorsellinic acid by hydroxylation of its starter methyl group and reaction with the end-of-chain carboxyl. The chain length of the farnesyl alkyl group is subsequently shortened by oxidation of a double bond, giving demethylmycophenolic acid, which is then *O*-methylated, again involving SAM, to produce mycophenolic acid. Note that the *O*-methylation step occurs only after the *C*-alkylations, so that the full activating benefit of two *meta*-positioned phenols can be utilized for the *C*-alkylation.

Box 3.14

Mycophenolic acid

Mycophenolic acid (Figure 3.82) is produced by fermentation cultures of the fungus *Penicillium brevicompactum*. It has been known for many years to have antibacterial, antifungal, antiviral, and antitumour properties. It has recently been introduced into medicine as an immunosuppressant drug, to reduce the incidence of rejection of transplanted organs, particularly in kidney and heart transplants. It is formulated as the *N*-morpholinoethyl ester **mycophenolate mofetil** (Figure 3.83), which is metabolized after ingestion to mycophenolic acid; it is usually administered in combination with another immunosuppressant, ciclosporin (see page 454). The drug is a specific inhibitor of mammalian inosine monophosphate (IMP) dehydrogenase and has an antiproliferative activity on cells due to inhibition of guanine nucleotide biosynthesis. IMP dehydrogenase catalyses the NAD^+-dependent oxidation of IMP to xanthosine monophosphate (XMP), a key transformation in the synthesis of guanosine triphosphate (GTP) (see also caffeine biosynthesis, page 413). Rapidly growing cells have increased levels of the enzyme, so this forms an attractive target for anticancer, antiviral, and immunosuppressive therapy. Mycophenolate mofetil may also be used to treat severe refractory eczema.



mycophenolate mofetil

Figure 3.83

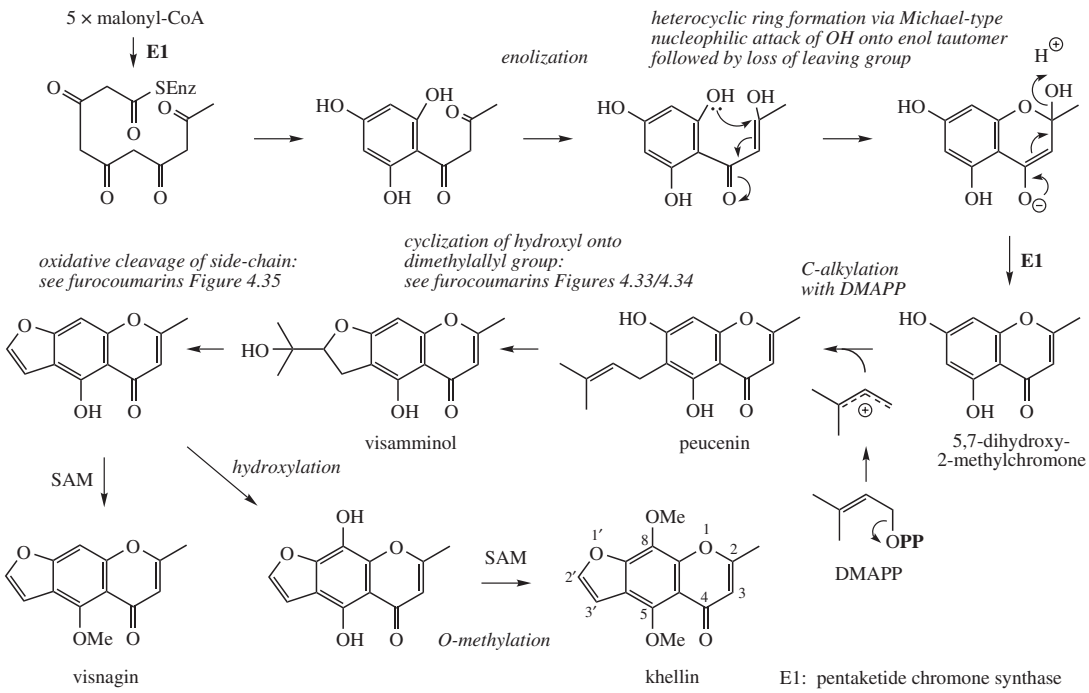


Figure 3.84

Khellin and **visnagin** (Figure 3.84) are furochromones found in the fruits of *Ammi visnaga* (Umbelliferae/Apiaceae), and the active principles of a crude plant drug which has a long history of use as an antiasthmatic agent. Figure 3.84 presents the sequence of steps utilized in the biosynthesis of these compounds, fully consistent with the biosynthetic rationale developed above. The two carbon atoms C-2′ and C-3′ forming part of the furan ring originate by metabolism of a five-carbon dimethylallyl substituent attached to C-6 (for a full discussion, see furocoumarins, page 163). The 8-methoxy group in khellin is absent in visnagin, so must be introduced late in the sequence. The key intermediate is thus **5,7-dihydroxy-2-methylchromone**. On inspection, this has the alternate acetate-derived oxygenation pattern and a methyl chain starter, so is formed from a poly-β-keto chain through Claisen condensation. The heterocyclic ring is produced by an overall dehydration reaction, formulated as a Michael-like addition reaction (see also flavonoids, page 169). Indeed, a type III PKS derived from *Aloe arborescens* has been characterized and shown to catalyse formation of the chromone from five molecules of malonyl-CoA; the starter unit is derived by decarboxylation. After formation of the furan ring via the *C*-dimethylallyl derivative peucenin and then visamminol, **visnagin** can be obtained by *O*-methylation. Alternatively, further hydroxylation *para* to the free phenol followed by two methylations yields **khellin**. The antiasthmatic properties of khellin have been exploited by developing the more polar, water-soluble derivative **cromoglicate** [Box 3.15].

Box 3.15

Khellin and Cromoglicate

The dried ripe fruits of *Ammi visnaga* (Umbelliferae/Apiaceae) have a long history of use in the Middle East as an antispasmodic and for the treatment of angina pectoris. The drug contains small amounts of coumarin derivatives, e.g. visnadin (Figure 3.85) (compare *Ammi majus*, a rich source of furocoumarins, page 165), but the major constituents (2–4%) are furochromones, including khellin, visnagin, khellol, and khellol glucoside (Figure 3.85). Both khellin and visnadin are coronary vasodilators and spasmolytic agents, with visnadin actually being the more potent agent. Khellin has been used in the treatment of angina pectoris and bronchial asthma.

The synthetic analogue **cromoglicate** (**cromoglycate**; Figure 3.85) is a most effective and widely used agent for the treatment and prophylaxis of asthma, hay fever, and allergic rhinitis. Cromoglicate contains two chromone systems with polar carboxylic acid functions, joined by a glycerol linker. The mode of action is not fully established. It was believed to prevent the release of bronchospasm mediators by stabilizing mast cell membranes, but an alternative suggestion is that it may act by inhibiting the effect of sensory nerve activation, thus interfering with bronchoconstriction. It is poorly absorbed orally and is thus admistered by inhalation or nasal spray. Eyedrops for relief of allergic conjunctivitis and nasal sprays for prophylaxis of allergic rhinitis are also available. The more potent pyridonochromone **nedocromil** (Figure 3.85) has also been introduced.

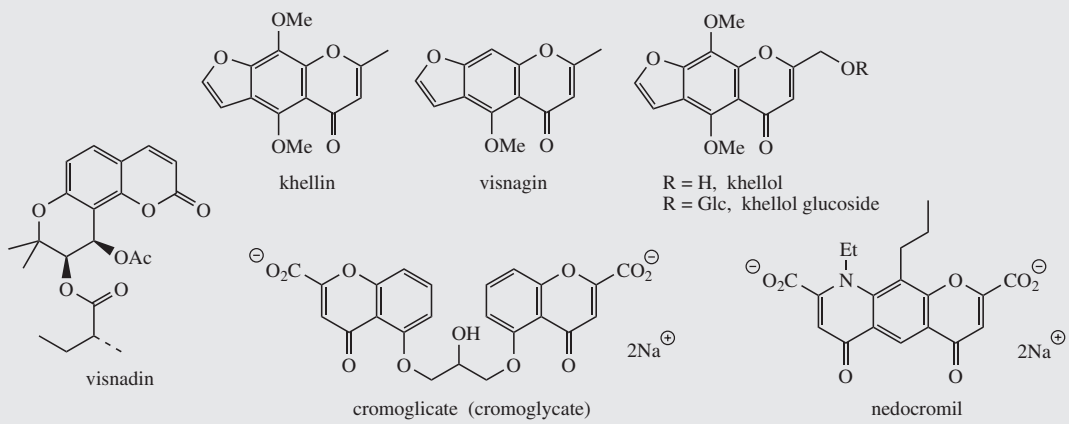


Figure 3.85

The polyketide nature of **hyperforin** (Figure 3.86) is almost entirely obscured by the added isoprenoid fragments. This compound is the predominant antidepressive agent in St John's Wort (*Hypericum perforatum*; Guttiferae/Hypericaceae) [Box 3.13], where it co-occurs with the naphthodianthrone hypericin (see page 108). The product of the type III PKS is **phlorisobutyrophenone**, an analogue of phloroacetophenone (see page 100) but using isobutyryl-CoA as starter instead of acetyl-CoA. This starter is formed from the amino acid valine, as seen in the pathways to bacterial iso-fatty acids (see page 56).

Successive *C*-alkylations with DMAPP substitute both of the remaining aromatic sites. The next alkylating agent is geranyl diphosphate, and though the mechanism is exactly the same, this alkylation must destroy the aromaticity. One more alkylation occurs; this is best formulated as a carbocation addition to the geranyl double bond, generating a tertiary carbocation, then quenching

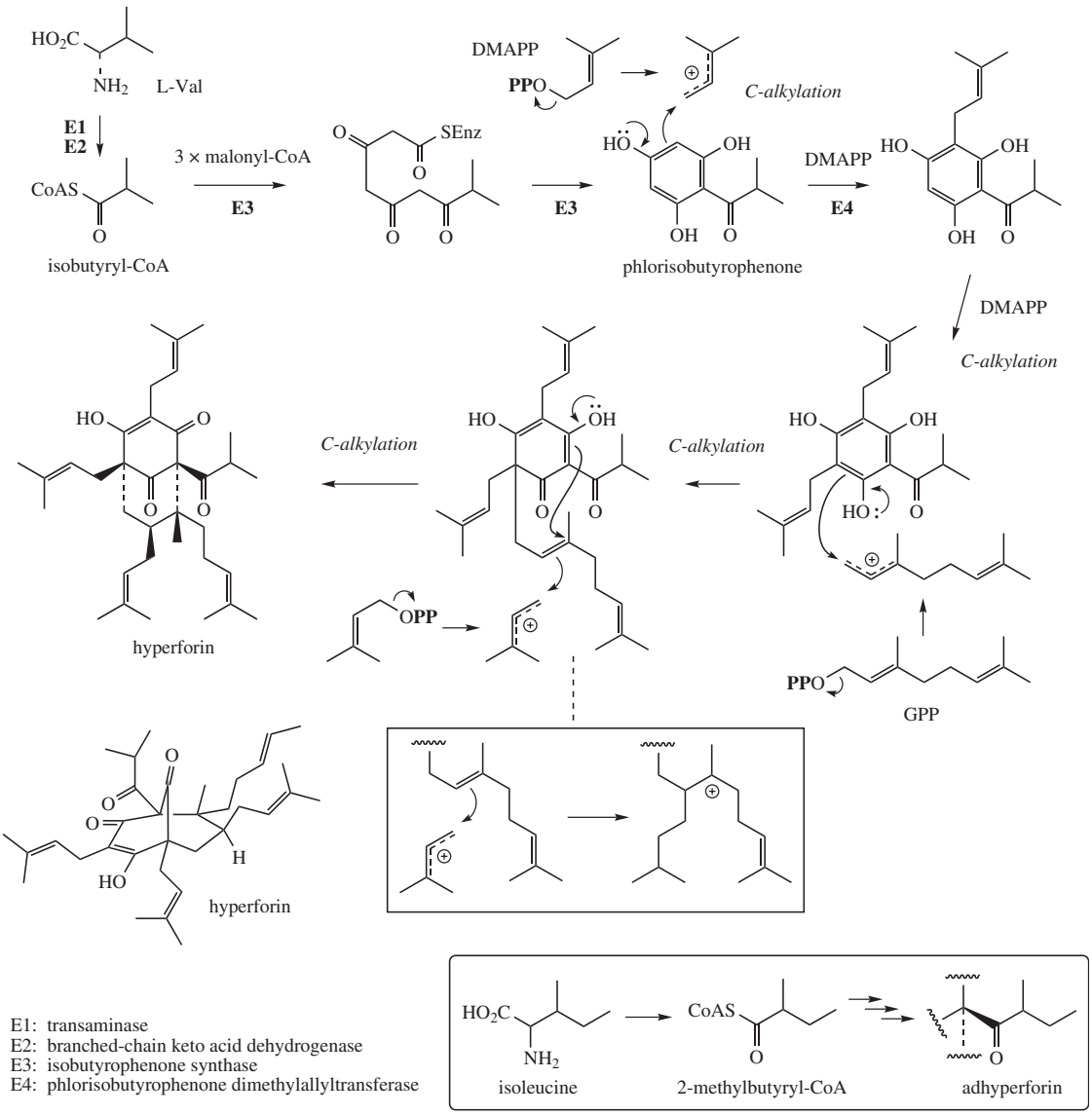


Figure 3.86

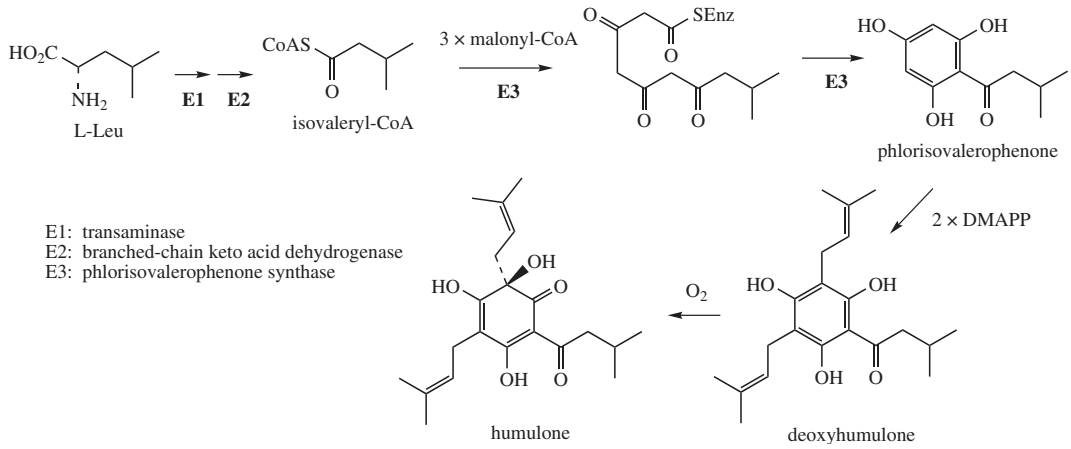


Figure 3.87

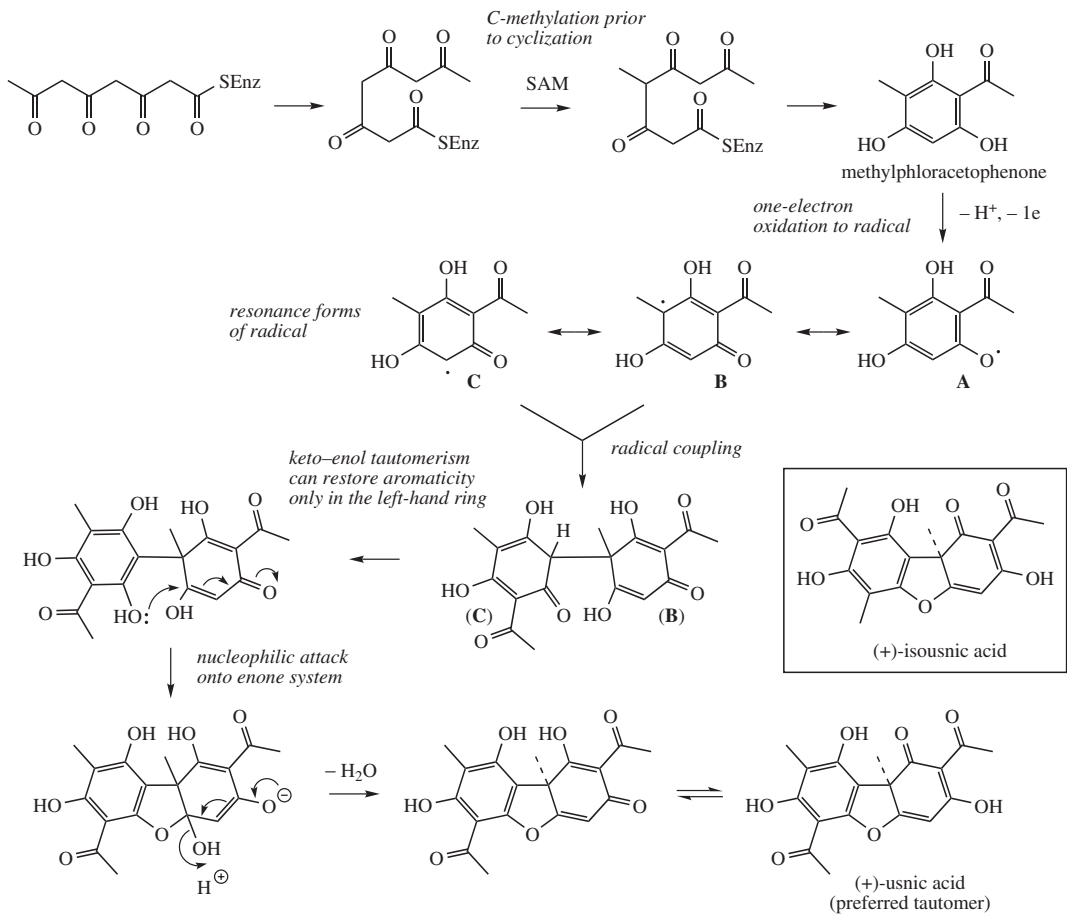


Figure 3.88

this by yet another alkylation reaction on the ring system. The structurally related **adhyperforin** (Figure 3.86) is also found in *H. perforatum*. This compound is characterized by derivation from isoleucine rather than valine, and the starter unit is thus 2-methylbutyryl-CoA (compare iso-fatty acids, page 56). Somewhat less complex structures are found in hops (*Humulus lululus*; Cannabaceae), where they are representative of a group of compounds termed bitter acids. These are responsible for the typical bitter taste of beer, and also provide foam-stabilizing and antibacterial properties. **Humulone** (Figure 3.87) is typically the major bitter acid, and is formed by oxidative transformaton of deoxyhumulone. Here, the starter unit for the polyketide is leucine-derived isovaleryl-CoA.

C-Methylation also features in the biosynthesis of **usnic acid** (Figure 3.88), an antibacterial metabolite found in many lichens, e.g. *Usnea* and *Cladonia* species, which are symbiotic combinations of alga and fungus. However, the principal structural modification encountered involves phenolic oxidative coupling (see page 28). Two molecules of **methylphloracetophenone** are incorporated, and these are known to derive from a pre-aromatization methylation reaction, not by methylation of phloracetophenone (Figure 3.88). The two molecules are joined together by an oxidative coupling mechanism which can be rationalized via the one-electron oxidation of a phenol group in methylphloracetophenone giving radical A, for which resonance forms B and C can be written. Coupling of B and C occurs. Only the left-hand ring can subsequently be restored to aromaticity by keto–enol tautomerism; this is not possible in the right-hand ring because coupling occurred on the position *para* to the original phenol and this position already contains a methyl. Instead, a heterocyclic ring is formed by attack of the phenol onto the enone system (see khellin, above). The outcome of this reaction is enzyme controlled; two equivalent phenol groups are present as potential nucleophiles, and two equivalent enone systems are also available. Potentially, four different products could be formed, but only one is typically produced. Loss of water then leads to usnic acid. Usnic acid is found in both (+)- and (−)-forms in nature, according to source. The product from using the alternative hydroxyl nucleophile in heterocyclic ring formation is **isousnic acid**, also known as a natural product in some species of lichens. Usnic acid is used as an antibacterial and preservative in toiletries.

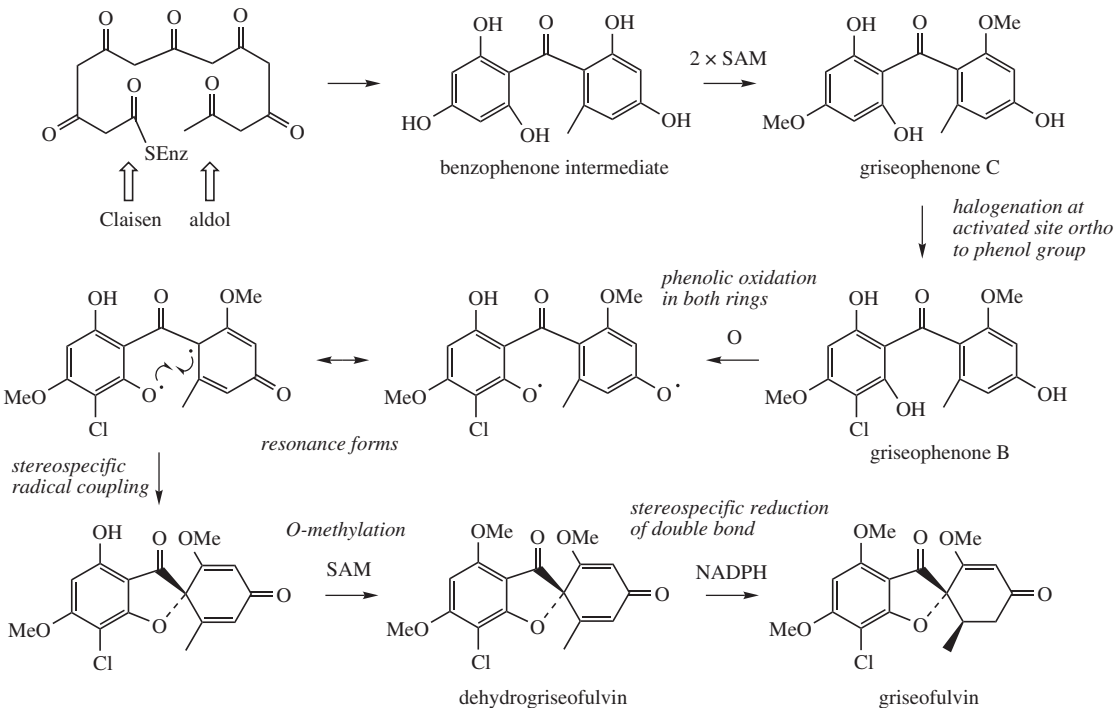


Figure 3.89

Phenolic oxidative coupling is widely encountered in natural product biosynthesis, and many other examples are described in subsequent sections. A further acetate-derived metabolite formed as a result of oxidative coupling is the antifungal agent **griseofulvin** (Figure 3.89) synthesized by cultures of *Penicillium griseofulvin* [Box 3.16]. The sequence of events leading to griseofulvin has been established in detail via feeding experiments; the pathway also includes *O*-methylation steps and the introduction of a halogen (chlorine) atom at one of the nucleophilic sites (Figure 3.89).

Initial inspection of the structure of griseofulvin shows the alternate oxygenation pattern, and also a methyl group which identifies the start of the polyketide chain. Cyclization of the C_{14} poly-β-keto chain folded as shown allows both Claisen (left-hand ring) and aldol (right-hand ring) reactions to occur, giving a benzophenone intermediate. Two selective methylations lead to griseophenone C, which is the substrate for chlorination to griseophenone B; both these compounds appear as minor metabolites in *P. griseofulvin* cultures. One-electron oxidations on a phenolic group in each ring give a diradical; radical coupling in its resonance form generates the basic grisan skeleton with its spiro centre. **Griseofulvin** is then the result of methylation of the remaining phenol group and stereospecific reduction of the double bond in dehydrogriseofulvin.

Box 3.16

Griseofulvin

Griseofulvin is an antifungal agent produced by cultures of *P. griseofulvum* and a number of other *Penicillium* species, including *P. janczewski, P. nigrum* and *P. patulum*. Griseofulvin is used to treat widespread or intractable dermatophyte infections, though has to some extent been superseded by newer antifungals. It is less effective when applied topically, but it is well absorbed from the gut and is selectively concentrated into the keratin layers of the skin. It may thus be used orally to control dermatophytes such as *Epidermophyton, Microsporium* and *Trichophyton*. Treatment for some conditions, e.g. infections in fingernails, may have to be continued for several months, but the drug is generally free of side-effects. Griseofulvin is also used orally or topically to control ringworm (*Tinea capitis*). The antifungal action appears to be through disruption of the mitotic spindle, thus inhibiting mitosis. The drug is relatively safe for use, since the effects on mitosis in sensitive fungal cells occur at concentrations that are substantially below those required to inhibit mitosis in human cells.

Starter Groups

In most of the examples of aromatics so far discussed, the basic carbon skeleton has been derived from an acetate starter group (sometimes formed via PKS decarboxylation of malonate), with malonate acting as the chain extender. The molecule has then been made more elaborate, principally via alkylation reactions and the inclusion of other carbon atoms. However, the range of natural aromatic structures which are at least partly derived from acetate is increased enormously by altering the nature of the starter group from acetate to an alternative acyl-CoA system, with malonyl-CoA again providing the chain extender. For example, the starter group for hyperforin biosynthesis is a branched-chain CoA ester derived from the amino acid leucine, and this has already been discussed (see page 113).

Chalcones and **stilbenes** are relatively simple examples of molecules in which a suitable cinnamoyl-CoA C_6C_3 precursor from the shikimate pathway (see page 149) is used as a starter group for the PKS system. If **4-hydroxycinnamoyl-CoA** (Figure 3.90) is chain extended with three malonyl-CoA units, the poly-β-keto chain can then be folded in two ways, allowing Claisen or aldol-type cyclizations to occur as appropriate. Chalcones, e.g. **naringenin-chalcone**, or stilbenes, e.g. **resveratrol**, are the end-products formed by the type III enzymes chalcone synthase and stilbene synthase respectively. Chalcone synthase and stilbene synthase enzymes share some 75–90% amino acid sequence identity, but cyclize the same substrate in different ways. The six-membered heterocyclic ring characteristic of most flavonoids, e.g. **naringenin**, is formed by nucleophilic attack of a phenol group from the acetate-derived ring onto the α,β-unsaturated ketone of the chalcone (compare khellin, page 111). Stilbenes, such as **resveratrol**, incorporate the carbonyl carbon of the cinnamoyl unit into the aromatic ring and typically lose the end-of-chain carboxyl by a decarboxylation reaction. Although some related structures, e.g. **lunularic acid** (Figure 3.90) from the liverwort *Lunularia cruciata*, still contain this carboxyl, these are rare in nature. Carboxylated stilbenes are not intermediates in the

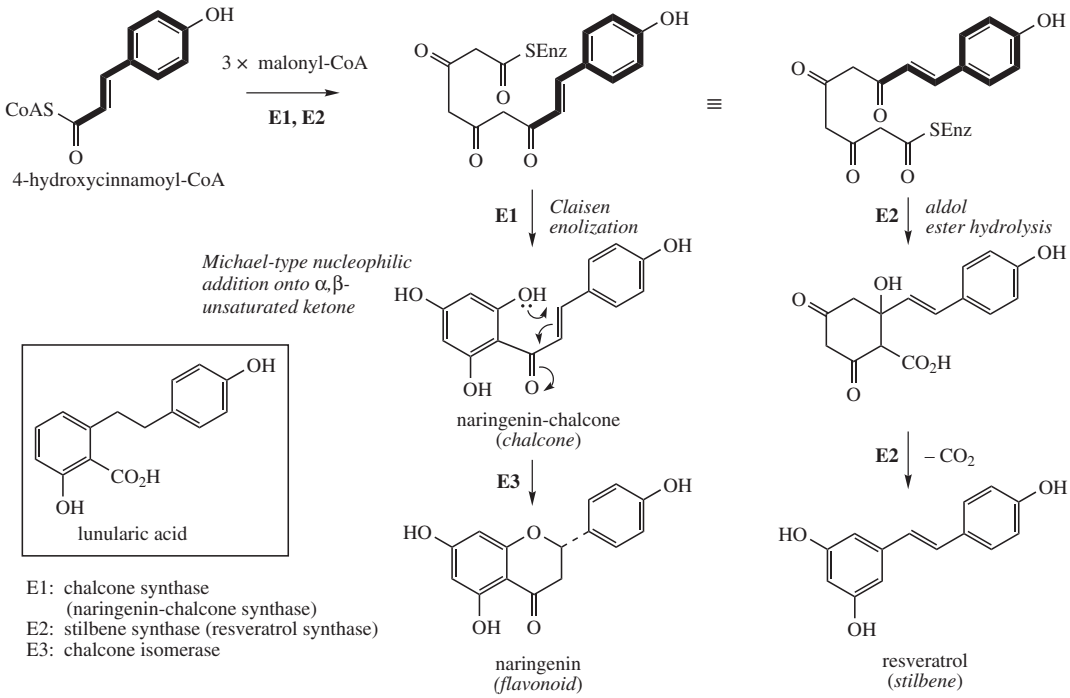


Figure 3.90

stilbene synthase reaction. Mutagenesis studies have confirmed that hydrolysis of the thioester bond and decarboxylation precede aromatization by the necessary dehydration/enolization reactions. Indeed, it appears that it is the TE activity component of the stilbene synthase enzyme that is primarily responsible for the stilbene-type cyclization (aldol reaction) rather than the chalcone-type cyclization (Claisen reaction). Through appropriate amino acid replacements it is possible to convert chalcone synthase into proteins with stilbene synthase activity. Flavonoids and stilbenes, and the further modifications that may occur, are traditionally treated as shikimate-derived compounds; they are discussed in more detail in Chapter 4 (see page 167).

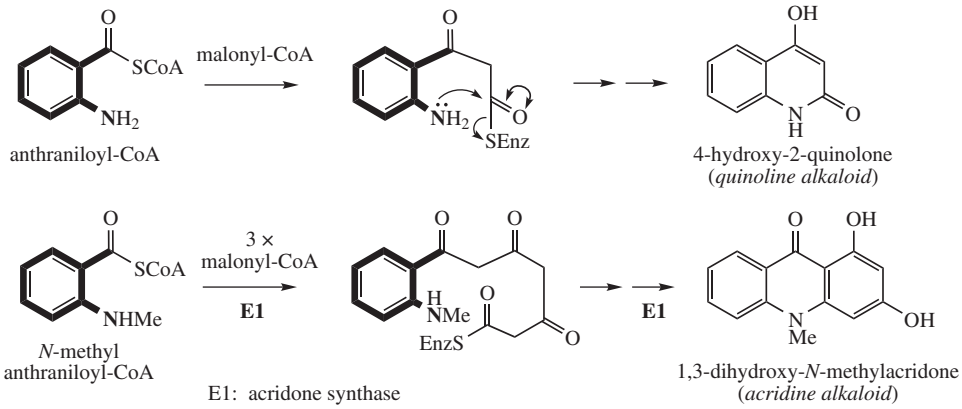


Figure 3.91

Anthranilic acid (2-aminobenzoic acid) (see page 141) is another shikimate-derived compound which, as its CoA ester anthraniloyl-CoA, can act as a starter unit for malonate chain extension. Aromatization of the acetate-derived portion then leads to **quinoline** or **acridine** alkaloids, according to the number of acetate units incorporated (Figure 3.91). Acridone synthase is a type III PKS catalysing chain extension of *N*-methylanthraniloyl-CoA to give the corresponding acridine alkaloid. It is closely related structurally to chalcone synthase; indeed, it also displays a modest chalcone synthase activity. Replacement of amino acids in three critical positions was sufficient to change its activity completely to chalcone synthase, so that it no longer accepted *N*-methylanthraniloyl-CoA as substrate. These products are usually considered as alkaloids, and are discussed in more detail in Chapter 6 (see page 395).

Fatty acyl-CoA esters are similarly capable of participating as starter groups. Fatty acid biosynthesis and aromatic polyketide biosynthesis are distinguished by the sequential reductions as the chain length increases in the former, and by the stabilization of a reactive poly-β-keto chain in the latter, with little or no reduction involved. It is thus interesting to see natural product structures

Box 3.17

Poison Ivy and Poison Oak

Poison ivy (*T. radicans* or *Rhus radicans*; Anacardiaceae) is a woody vine with three-lobed leaves that is common in North America and also found in South America. The plant may be climbing, shrubby, or may trail over the ground. It presents a considerable hazard to humans should the sap, which exudes from damaged leaves or stems, come into contact with the skin. The sap sensitizes most individuals, producing delayed contact dermatitis after a subsequent encounter. This results in watery blisters that break open, with the fluid quickly infecting other parts of the skin. The allergens may be transmitted from one person to another on the hands, on clothing, or by animals. The active principles are urushiols, a mixture of alkenyl polyphenols. In poison ivy, these are mainly catechols with C_{15} side-chains carrying varying degrees of unsaturation (Δ^8, $\Delta^{8,11}$, $\Delta^{8,11,14}$). Small amounts of C_{17} side-chain analogues are present. These catechols become oxidized to an *ortho*-quinone, which is then attacked by nucleophilic groups in proteins to yield an antigenic complex.

Poison oak (*T. toxicaria* or *Rhus toxicodendron*; Anacardiaceae) is nearly always found as a low-growing shrub and has lobed leaflets similar to those of oak. It is also common throughout North America. There appears considerable confusion over nomenclature, and *Rhus radicans* may also be termed poison oak, and *R. toxicodendron* oakleaf poison ivy. Poison oak contains similar urushiol structures in its sap as poison ivy, though heptadecylcatechols (i.e. C_{17} side-chains) predominate over pentadecylcatechols (C_{15} side-chains).

Related species of *Toxicodendron*, e.g. *T. diversilobum* (Pacific poison oak) and *T. vernix* (poison sumach, poison alder, poison dogwood), are also allergenic with similar active constituents. These plants were all formerly classified under the genus *Rhus*, but the allergen-containing species have been reclassified as *Toxicodendron*; this nomenclature is not widely used. Dilute purified extracts containing urushiols may be employed to stimulate antibody production and, thus, build up some immunity to the allergens. These plants cause major problems to human health in North America, with many thousands of reported cases each year.

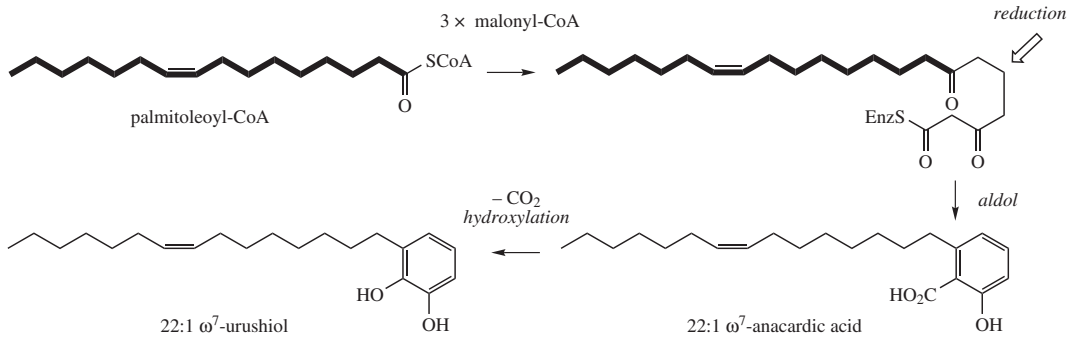


Figure 3.92

containing both types of acetate/malonate-derived chains (see also zearalenone, page 89). In plants of the Anacardiaceae, just such a pathway accounts for the formation of contact allergens called **urushiols** in poison ivy (*Toxicodendron radicans*) and poison oak (*Toxicodendron toxicaria*) [Box 3.17]. Thus, **palmitoleoyl-CoA** (Δ^9-hexadecenoyl-CoA) can act as starter group for extension by three malonyl-CoA units, with a reduction step during chain extension (Figure 3.92). Aldol cyclization then gives an **anacardic**

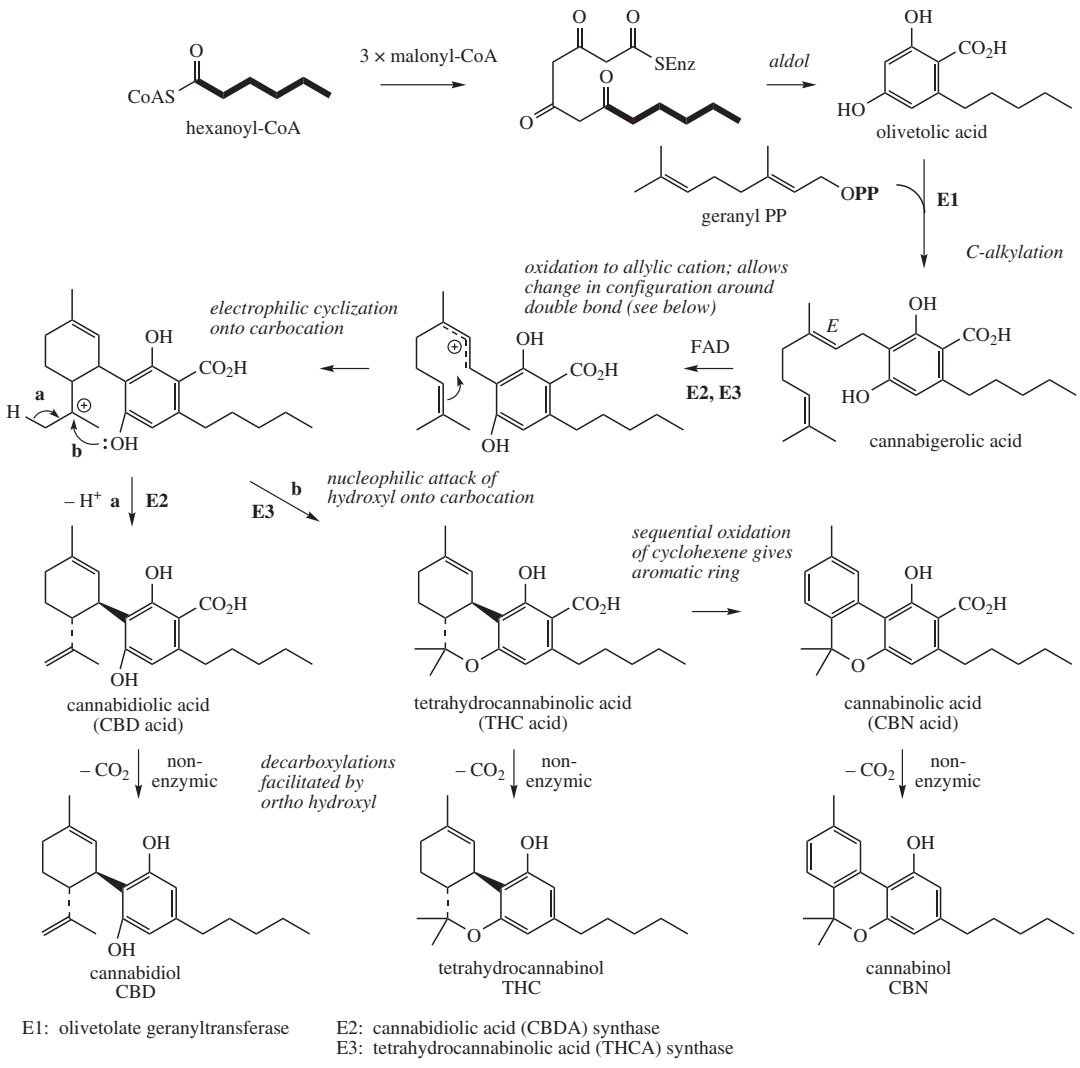


Figure 3.93

acid, which is likely to be the precursor of a **urushiol** by decarboxylation/hydroxylation. However, enzymic evidence is not yet available. It is likely that different fatty acyl-CoAs can participate in this sequence, since urushiols from poison ivy can contain up to three double bonds in the C_{15} side-chain, whilst those from poison oak also have variable unsaturation in a C_{17} side-chain (i.e. a C_{18} acyl-CoA starter). The name anacardic acid relates particularly to the presence of large quantities of these compounds in the shells of cashew nuts (*Anacardium occidentale*; Anacardiaceae); these anacardic acids have C_{15} side-chains carrying one, two, or three double bonds.

A saturated C_6 **hexanoate** starter unit is used in the formation of the **cannabinoids**, a group of terpenophenolics found in Indian hemp (*Cannabis sativa*; Cannabaceae). This plant, and preparations from it, are known under a variety of names, including hashish, marihuana, pot, bhang, charas, and dagga, and have been used for centuries for the pleasurable sensations and mild euphoria experienced after its consumption, usually by smoking. The principal psychoactive component is **tetrahydrocannabinol** (**THC**; Figure 3.93), whilst structurally similar compounds such as **cannabinol** (**CBN**) and **cannabidiol** (**CBD**), present in similar or larger amounts, are effectively inactive. In recent years, the beneficial effects of cannabis, and especially THC, in alleviating nausea and vomiting in cancer patients undergoing chemotherapy, and in the treatment of glaucoma and multiple sclerosis, has led to extensive study of cannabinoid analogues for potentially useful medicinal activity [Box 3.18].

All the cannabinoid structures contain a monoterpene C_{10} unit attached to a phenolic ring that carries a C_5 alkyl chain. The aromatic ring/C_5 chain originates from hexanoate and malonate, cyclization to a polyketide giving **olivetolic acid**, from which **cannabigerolic acid** can be obtained by *C*-alkylation with the monoterpene C_{10} unit geranyl diphosphate (Figure 3.93). Cyclization in the monoterpene unit necessitates a change in configuration of the double bond, and this involves an oxidative step in which FAD is a cofactor. It may be rationalized as involving the allylic cation, which will then allow electrophilic cyclization to proceed (compare terpenoid cyclization mechanisms, page 195). **Cannabidiolic acid** is then the result of proton loss, whilst **tetrahydrocannabinolic acid** is the product from heterocyclic ring formation via quenching with the phenol group. **CBD** and **THC** are then the respective decarboxylation products from these two compounds; decarboxylation is non-enzymic and facilitated by the adjacent phenol function (see page 22). The aromatic terpenoid-derived ring in **cannabinolic acid** and **CBN** can arise via a dehydrogenation process (compare thymol, page 204). The PKS for olivetolic acid formation has yet to be characterized; a related type III PKS system from *Cannabis* catalysed the formation of the decarboxylated derivative olivetol instead of olivetolic acid.

Box 3.18

Cannabis

Indian hemp, *Cannabis sativa* (Cannabaceae), is an annual herb indigenous to central and western Asia. It is cultivated widely in India and many tropical and temperate regions for its fibre (hemp) and seed (for seed oil). The plant is also grown for its narcotic and mild intoxicant properties, and in most countries of the world, its possession and consumption are illegal. Over many years, cannabis plants have been selected for either fibre production or drug use, the former resulting in tall plants with little pharmacological activity, whilst the latter tend to be short, bushy plants. Individual plants are almost always male or female, though the sex is not distinguishable until maturity and flowering. Seeds will produce plants of both sexes in roughly equal proportions. The active principles are secreted as a resin by glandular hairs, which are more numerous in the upper parts of female plants, and resin is produced from the time flowers first appear until the seeds reach maturity. However, all parts of the plant, both male and female, contain cannabinoids. In a typical plant, the concentration of cannabinoids increases in the order: large leaves, small leaves, flowers, bracts (which surround the ovaries), with stems containing very little. Material for drug use (ganja) is obtained by collecting the flowering tops (with little leaf) from female plants, though lower quality material (bhang) consisting of leaf from both female and male plants may be employed. By rubbing the flowering tops, the resin secreted by the glandular hairs can be released and subsequently scraped off to provide cannabis resin (charas) as an amorphous brown solid or semi-solid. A potent form of cannabis, called cannabis oil, is produced by alcoholic extraction of cannabis resin. A wide variety of names are used for cannabis products according to their nature and the geographical area. In addition to the Indian words above, the names hashish (Arabia), marihuana (Europe, USA), kief and dagga (Africa) are frequently used. The term ‘assassin’ is a corruption of ‘hashishin’, a group of 13th century murderous Persians who were said to have been rewarded for their activities with hashish. Names such as grass, dope, pot, hash, weed, and wacky backy are more likely to be in current usage. Cannabis for the illicit market is cultivated worldwide. Most of the resin originates from North Africa, Afghanistan, and Pakistan, whilst

Box 3.18 (continued)

cannabis herb comes mainly from North America and Africa. Canada, the United States, and the UK grow large quantities for legitimate use.

The quantity of resin produced by the flowering tops of high-quality Indian cannabis is about 15–20%. The amount produced by various plants is dependent on several features, however, and this will markedly alter biological properties. Thus, in general, plants grown in a tropical climate produce more resin than those grown in a temperate climate. The tall fibre-producing plants are typically low resin producers, even in tropical zones. However, the most important factor is the genetic strain of the plant, as the resin produced may contain high levels of psychoactive compounds or mainly inactive constituents. The quality of any cannabis drug is potentially highly variable.

The major constituents in cannabis are termed cannabinoids, a group of more than 60 structurally related terpenophenolics. The principal psychoactive agent is tetrahydrocannabinol (THC, Figure 3.93). This is variously referred to as Δ^1-THC or Δ^9-THC according to whether the numbering is based on the terpene portion or as a systematic dibenzopyran (Figure 3.94). Both systems are currently in use, though the systematic one is preferred. Also found, often in rather similar amounts, are cannabinol (CBN) and cannabidiol (CBD) (Figure 3.93), which have negligible psychoactive properties. These compounds predominate in the inactive resins. Many other cannabinoid structures have been characterized, including cannabigerol and cannabichromene (Figure 3.94). A range of cannabinoid acids, e.g. cannabidiolic acid, tetrahydrocannabinolic acid, and tetrahydrocannabinolic acid-B (Figure 3.94) are also present, as are some analogues of the other compounds mentioned, where a propyl side-chain replaces the pentyl group, e.g. tetrahydrocannabivarin (Figure 3.94). The latter compounds presumably arise from the use of butyrate rather than hexanoate as starter unit in the biosynthetic sequence.

The THC content of high-quality cannabis might be in the range 0.5–1% for large leaves, 1–3% for small leaves, 3–7% for flowering tops, 5–10% for bracts, 14–25% for resin, and up to 60% in cannabis oil. Higher amounts of THC are produced in selected strains known as skunk cannabis, so named because of its powerful smell; flowering tops from skunk varieties might contain 10–15% THC. The THC content in cannabis products tends to deteriorate on storage, an effect accelerated by heat and light. Cannabis leaf and resin stored under ordinary conditions rapidly lose their activity and can be essentially inactive after about 2 years. A major change which occurs is oxidation in the cyclohexene ring resulting in conversion of THC into CBN. THC is more potent when smoked than when taken orally, its volatility allowing rapid absorption and immediate effects, so smoking has become the normal means of using cannabis. Any cannabinoid acids will almost certainly be decarboxylated upon heating, and thus the smoking process will also effectively increase somewhat the levels of active cannabinoids available, e.g. THC acid → THC (Figure 3.93). The smoking of cannabis produces a mild euphoria similar to alcohol intoxication, inducing relaxation, contentment, and a sense of well-being, with some changes in perception of sound and colour. However, this is accompanied by a reduced ability to concentrate and do complicated tasks, and a loss of short-term memory. Users claim cannabis is much preferable to alcohol or tobacco, insisting it does not cause dependence, withdrawal symptoms, or lead to the use of other drugs,

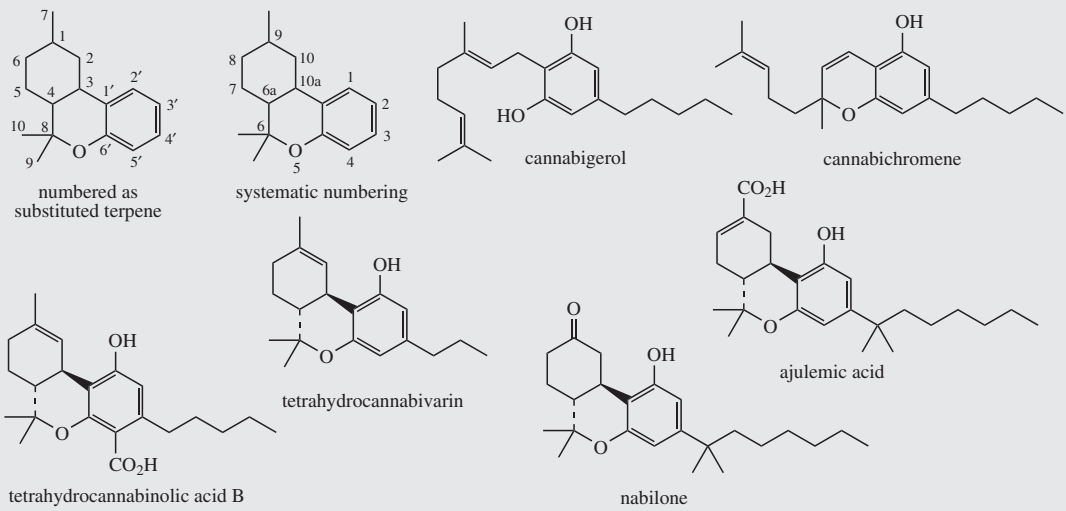


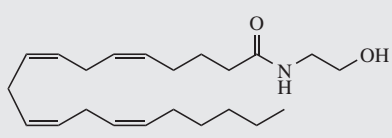
Figure 3.94

Box 3.18 (continued)

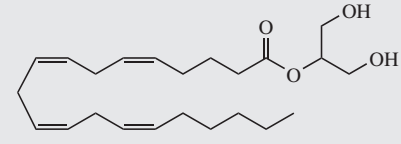
and they campaign vociferously for its legalization. However, psychological dependence does occur, and cannabis can lead to hallucinations, depression, anxiety, and panic, with the additonal risk of bronchitis and lung cancer when the product is smoked.

Cannabis has been used medicinally, especially as a mild analgesic and tranquillizer, but more effective and reliable agents replaced it, and even controlled prescribing was discontinued. In recent times, cannabis has been shown to have valuable anti-emetic properties which help to reduce the side-effects of nausea and vomiting caused by cancer chemotherapeutic agents. This activity stems from THC, and has resulted in some use of **dronabinol** (synthetic THC) and the prescribing of cannabis for a small number of patients. The synthetic THC analogue **nabilone** (Figure 3.94) has been developed as an anti-emetic drug for reducing cytotoxic-induced vomiting. Some of the psychoactive properties of THC, e.g. euphoria, mild hallucinations, and visual disturbances, may be experienced as side-effects of nabilone treatment. Cannabis has also been shown to possess properties which may be of value in other medical conditions. There is now ample evidence that cannabis can give relief to patients suffering from chronic pain, multiple sclerosis, glaucoma, asthma, migraine, epilepsy, and other conditions. Many sufferers who cannot seem to benefit from any of the current range of drugs are obtaining relief from their symptoms by using cannabis, but are breaking the law to obtain this medication. Current thinking is that cannabis offers a number of beneficial pharmacological responses and that there should be legal prescribing of cannabinoids or derivatives. Clinical trials have already confirmed the value of cannabis and/or THC taken orally for the relief of chronic pain and the painful spasms characteristic of multiple sclerosis, and in reducing intraocular pressure in glaucoma sufferers. An oral spray preparation of natural THC + CBD (**Sativex®**) for multiple sclerosis patients has been approved in some countries. In general, cannabis is only able to alleviate the symptoms of these diseases; it does not provide a cure. The non-psychoactive CBD has been shown to have anti-inflammatory properties potentially useful in arthritis treatment. The synthetic cannabinoid **ajulemic acid** (Figure 3.94) is in clinical trials for treatment of arthritic pain and inflammation; it does not produce any psychotropic actions at therapeutic dosage.

Two main cannabinoid receptors have been identified: CB_1, predominantly in the central nervous system, and CB_2, expressed mainly in the immune system. This was followed by discovery in animal brain tissue of a natural ligand for CB_1, namely anandamide (Figure 3.95), which is arachidonoylethanolamide; ananda is the Sanskrit word for bliss. Anandamide mimics several of the pharmacological properties of THC. The natural ligand of CB_2 is 2-arachidonoylglycerol (Figure 3.95); this also interacts with CB_1, and since levels of 2-arachidonoylglycerol in the brain are some 800 times higher than those of anandamide, it is now thought to be the physiological ligand for both receptors, rather than anandamide. These two compounds are the main ligands, but other related natural compounds from mammalian brain also function in the same way. These include polyunsaturated fatty acid ethanolamides, namely dihomo-γ-linolenoyl- (20:3) and adrenoyl- (22:4) ethanolamides, *O*-arachidonoylethanolamine (virodhamine) and 2-arachidonylglyceryl ether (noladin ether). The identification of these endogenous materials may open up other ways of exploiting some of the desirable pharmacological features of cannabis.



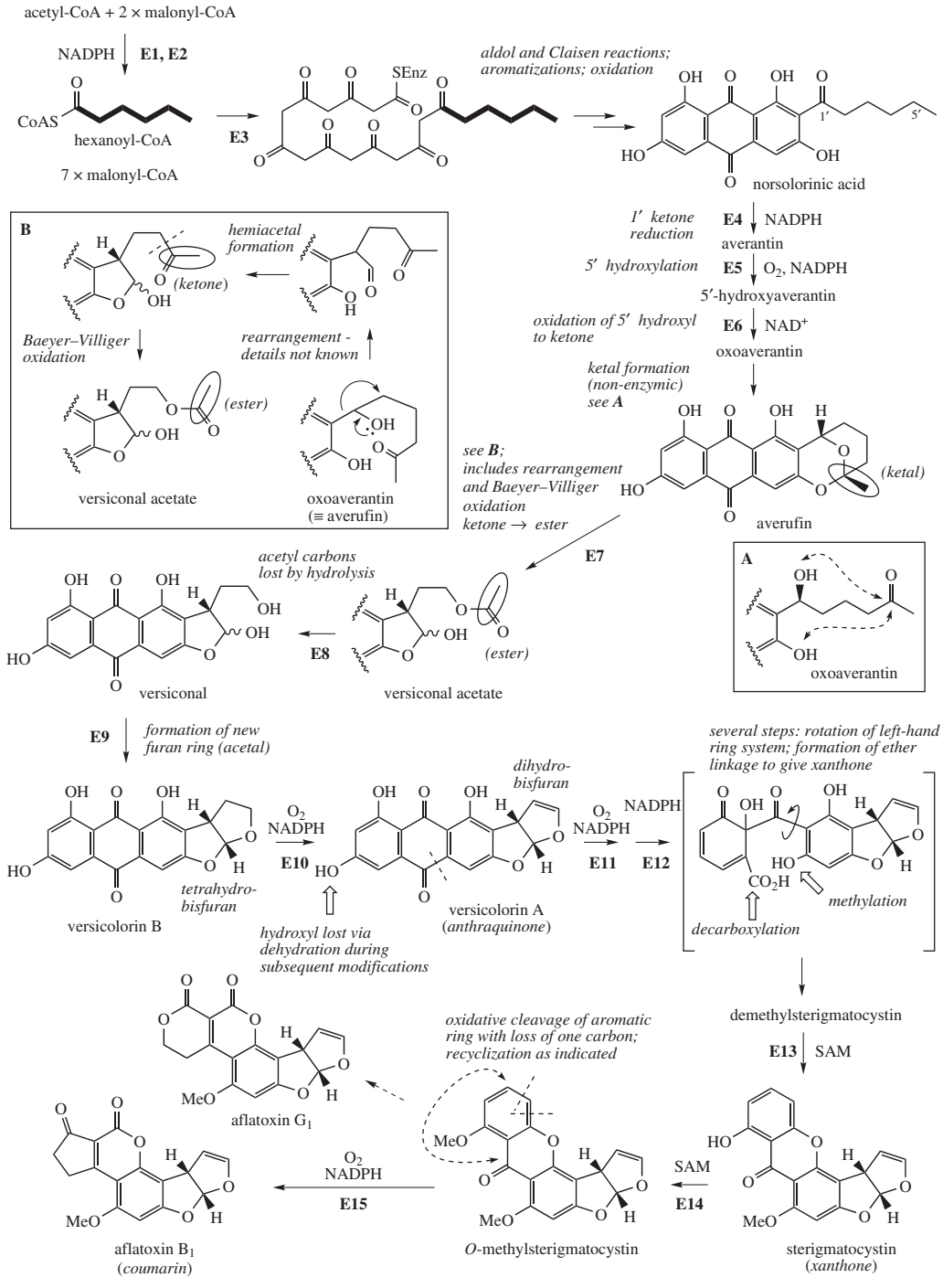
anandamide
(arachidonoylethanolamide)



2-arachidonoylglycerol

Figure 3.95

Hexanoate is also the starter unit used in the formation of the **aflatoxins**, a group of highly toxic metabolites produced by *Aspergillus flavus* and *Aspergillus parasiticus*, and probably responsible for the high incidence of liver cancer in some parts of Africa. These compounds were first detected following the deaths of young turkeys fed on mould-contaminated peanuts (*Arachis hypogaea*; Leguminosae/Fabaceae). Peanuts still remain one of the crops most likely to present a potential risk to human health because of contamination with aflatoxins [Box 3.19]. These and other food materials must be routinely screened to ensure levels of aflatoxins do not exceed certain set limits. The aflatoxin structures contain a bisfuran unit fused to an



E1, E2: Fas1, Fas2 (FAS)
E3: PksA (PKS)
E4: Nor-1 (reductase)
E5: AvnA (hydroxylase)
E6: AdhA (alcohol dehydrogenase)
E7: AvfA (oxidase)
E8: EstA (esterase)
E9: VBS (VERB synthase)
E10: VerB (desaturase)
E11: AflN (oxidase)
E12: AflM (reductase)
E13: OmtB (methyltransferase)
E14: OmtA (methyltransferase)
E15: OrdA (oxidoreductase)

Figure 3.96

aromatic ring, e.g. **aflatoxin B_1** and **aflatoxin G_1**, and their remarkably complex biosynthetic origin begins with a poly-β-keto chain derived from a hexanoyl-CoA starter and seven malonyl-CoA extender units (Figure 3.96). This gives an anthraquinone **norsolorinic acid** by the now-familiar condensation reactions, but the folding of the chain is rather different from that seen with simpler anthraquinones (see page 105). The enzyme norsolorinic acid synthase utilizes just acetyl-CoA and malonyl-CoA, and is a complex of an iterative type I PKS, together with a yeast-like type I FAS that produces the hexanoate starter (compare zearalenone, page 89). The FAS is comprised of two subunit proteins and transfers the starter to the PKS without any involvement of free hexanoyl-CoA; naturally, NADPH is a necessary cofactor for the FAS component. The six-carbon side-chain of norsolorinic acid is then oxygenated and cyclized to give the ketal **averufin**.

The remaining pathway rapidly becomes very complicated indeed and is noteworthy for several major skeletal changes and the involvement of Baeyer–Villiger oxidations (see page 27), catalysed by cytochrome P-450-dependent monooxygenases. **Versiconal acetate** is another known intermediate; its formation involves a rearrangement, currently far from clear, and a Baeyer–Villiger oxidation. The latter oxidation achieves principally the transfer of a two-carbon fragment (the terminal ethyl of hexanoate) to become an ester function. These two carbon atoms can then be lost by hydrolysis, leading to formation of **versicolorin B**, now containing the tetrahydrobisfuran moiety (an acetal), which is oxidized in **versicolorin A** to a dihydrobisfuran system. Only minimal details of the following steps are shown in Figure 3.96. They result in significant structural modifications, from the anthraquinone skeleton of versicolorin A to the xanthone skeleton of sterigmatocystin, whilst the aflatoxin end-products are coumarins. **Sterigmatocystin** is derived from versicolorin A by oxidative cleavage of the anthraquinone system catalysed by a cytochrome P-450-dependent enzyme and involving a Baeyer–Villiger cleavage, then recyclization through phenol groups to give a xanthone skeleton. Rotation in an intermediate leads to the angular product as opposed to a linear product. One phenol group is methylated and another phenol group is lost. This is not a reductive step, but a dehydration, achievable because the initial oxidation involves the aromatic ring and temporarily destroys aromaticity. **Aflatoxin B_1** formation requires oxidative cleavage of an aromatic ring in sterigmatocystin, loss of one carbon, and recyclization exploiting the carbonyl functionality; interestingly, an *O*-methylation step is required first, though the methoxy group so formed is subsequently lost. **Aflatoxin G_1** was originally thought to be derived by a Baeyer–Villiger reaction on aflatoxin B_1, but it is now known to be produced via a branch pathway from *O*-methylsterigmatocystin; the fine detail has yet to be established.

Box 3.19

Aflatoxins

Aflatoxins are potent mycotoxins produced by the fungi *A. flavus* and *A. parasiticus*. Four main naturally occurring aflatoxins, aflatoxins B_1, B_2, G_1, and G_2 (Figure 3.97), are recognized, but these can be metabolized by microorganisms and animals to other aflatoxin structures, which are also toxic. Aflatoxin B_1 is the most commonly encountered member of the group, and is also the most acutely toxic and carcinogenic example. Aflatoxin B_2 is a dihydro derivative of aflatoxin B_1, whilst aflatoxins G_1 and G_2 are an analogous pair with a six-membered lactone rather than a five-membered cyclopentenone ring. These toxins are most commonly associated with peanuts (groundnuts), maize, rice, pistachio nuts, and Brazil nuts, though other crops can be affected, and although found worldwide, they are particularly prevalent in tropical and subtropical regions. Aflatoxin M_1 (Figure 3.97)

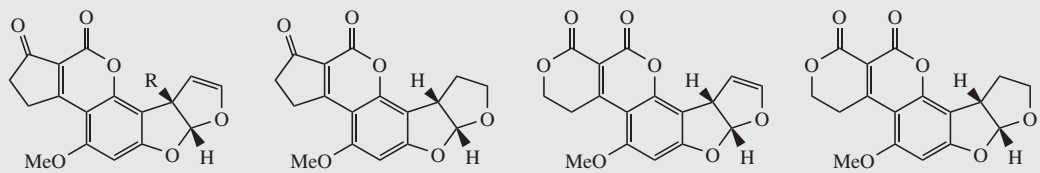


Figure 3.97

Box 3.19 (continued)

is a hydroxy derivative of aflatoxin B_1 and equally toxic. It may occur in cow's milk as a result of mammalian metabolism of aflatoxin B_1 originally contaminating the animal's food. Because these compounds fluoresce strongly under UV light, they are relatively easily detected and monitored.

The aflatoxins primarily affect the liver, causing enlargement, fat deposition, and necrosis, at the same time causing cells of the bile duct to proliferate, with death resulting from irreversible loss of liver function. In the case of aflatoxin B_1, this appears to be initiated by cytochrome P-450-dependent metabolism in the body to the epoxide (Figure 3.98). The epoxide intercalates with DNA, and in so doing becomes orientated towards nucleophilic attack from guanine residues. This leads to inhibition of DNA replication and of RNA synthesis, and initiates mutagenic activity. Fortunately, endogenous glutathione is normally available to minimize damage by reacting with dangerous electrophiles. A glutathione–toxin adduct is formed in the same way (Figure 3.98), and the polar functionalities make this adduct water soluble and excretable. Aflatoxins are also known to cause hepatic carcinomas, this varying with the species of animal. The above normal incidence of liver cancer in Africa and Asia has been suggested to be linked to the increased amounts of aflatoxins found in foodstuffs, and a tolerance level of 30 ppb has been recommended. Acute hepatitis may result from food containing aflatoxin B_1 at levels of the order of 0.1 ppm; levels of more than 1 ppm are frequently encountered.

The biosynthesis of aflatoxins proceeds through intermediates sterigmatocystin and versicolorin (see Figure 3.96). Toxins related to these structures, but differing in aromatic substituents, are also produced by various fungi. The sterigmatocystins are synthesized by species of *Aspergillus* and *Bipolaris*, and contain a reduced bifuran fused to a xanthone, whilst the versicolorins from *Aspergillus versicolor* contain the same type of reduced bisfuran system but fused to an anthraquinone. Like the aflatoxins, the sterigmatocystins are acutely toxic and carcinogenic. The versicolorins are less toxic, though still carcinogenic.

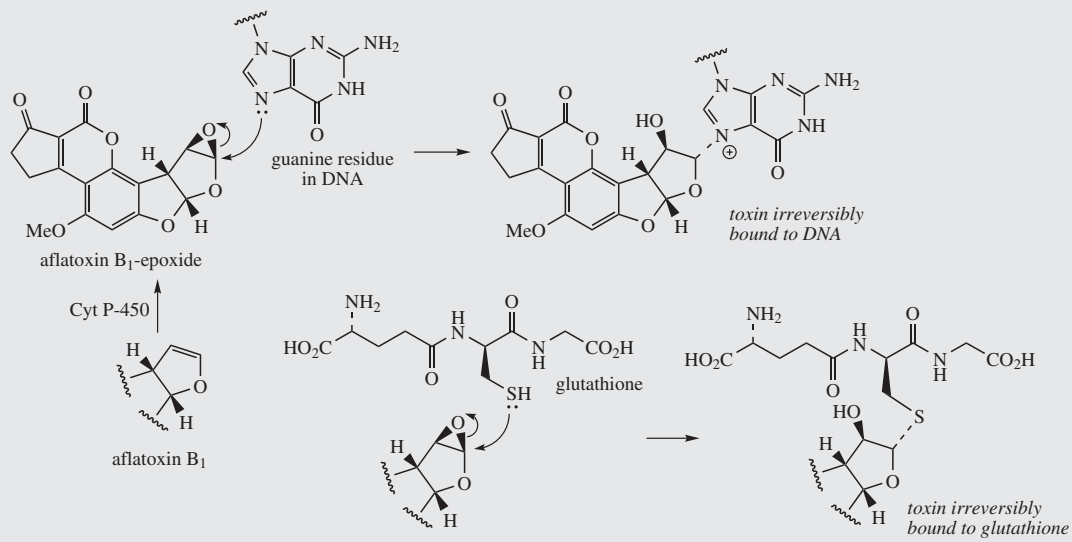
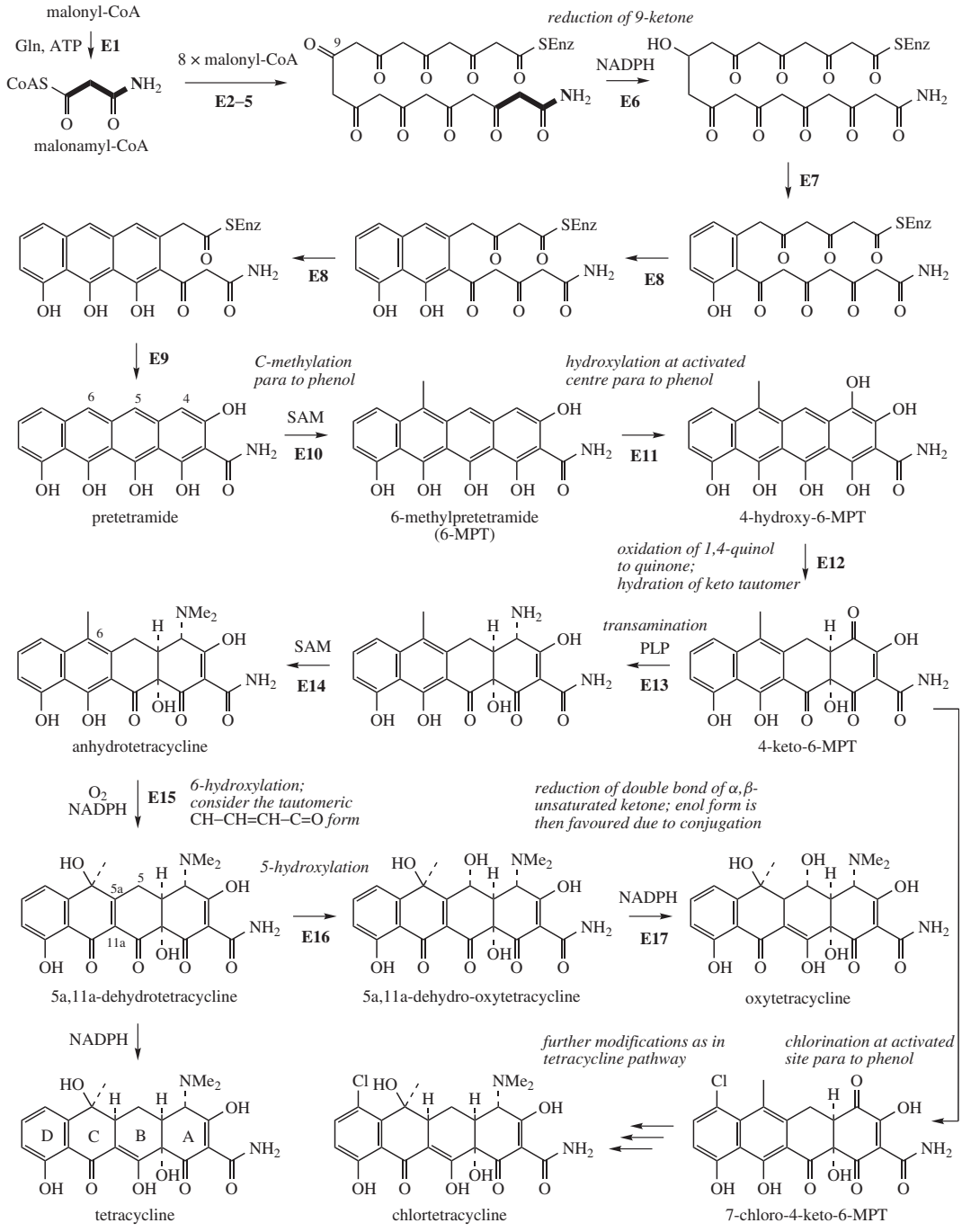


Figure 3.98

The **tetracyclines** are a group of broad-spectrum antibiotics produced by species of *Streptomyces*, and several natural and semi-synthetic members are used clinically [Box 3.20]. They contain a linear tetracyclic skeleton of polyketide origin in which the starter group is **malonamyl-CoA** (Figure 3.99), i.e. the coenzyme A ester of malonate semi-amide. Thus, all carbon atoms of the tetracycline skeleton are malonate derived, and the PKS also provides the amidotransferase activity that introduces the amino function into the starter group. The main features of the pathway (Figure 3.99) were deduced from extensive studies of mutant strains of *Streptomyces aureofaciens* with genetic blocks causing accumulation of mutant metabolites or the production of abnormal



E1: OxyD (amidotransferase)
E2: OxyA (ketosynthase)
E3: OxyB (chain length factor)
E4: OxyC (ACP)
E5: OxyP (acyltransferase)
E6: OxyJ (9-ketoreductase)
E7: OxyK (aromatase)
E8: OxyN (cyclase)
E9: OxyI (cyclase)
E10: OxyF (methyltransferase)
E11: OxyL (oxygenase)
E12: OxyG (oxygenase)
E13: OxyQ, OxyR (aminotransferase)
E14: OxyT (*N*-methyltransferase)
E15: OxyS (oxygenase)
E16: OxyE (oxygenase)
E17: TchA (reductase)

Figure 3.99

tetracylines. This organism typically produces **chlortetracycline**, whilst the parent compound **tetracycline** is in fact an aberrant product synthesized in mutants blocked in the chlorination step. The use of mutants with genetic blocks also enabled the shikimate pathway (Chapter 4) to be delineated. In that case, since a primary metabolic pathway was affected, mutants tended to accumulate intermediates and could not grow unless later components of the pathway were supplied. With the tetracyclines, a secondary metabolic pathway is involved, and the relatively broad specificity of some of the enzymes concerned allows many of the later steps to proceed even if one step, e.g. the chlorination, is not achievable. This has also proved valuable for production of some of the clinical tetracycline antibiotics. More recently, the genetic details of the pathway have been clarified, with most attention being directed towards study of **oxytetracycline** production in *Streptomyces rimosus*. Nevertheless, the bulk of the pathway is effectively the same in both organisms. The PKS is an iterative type II system, i.e. the relevant enzyme activities are provided by individual separable proteins.

The appropriate poly-β-keto ester is constructed almost as expected (Figure 3.99). However, the 9-keto reductive step is catalysed after chain assembly is complete, rather than during chain extension. This is controlled by a specific 9-ketoreductase component; reduction at the position nine carbon atoms from the carboxy terminus is a remarkably consistent feature of a number of other type II PKS systems. The tetracene ring system is built up gradually, starting with the ring at the centre of the chain fold as with anthraquinones, and the first intermediate released from the enzyme is **pretetramide** (Figure 3.99). This is the substrate for introduction of a methyl group at C-6, giving 6-methylpretetramide (full tetracene numbering is shown in Figure 3.100). Hydroxylation in ring A followed by oxidation gives a quinone; to accommodate the quinone, this will necessitate tautomerism in ring B to the keto form. The quinone is then substrate for hydration at the A/B ring fusion. An amine group is subsequently introduced stereospecifically into ring A by a transamination reaction, followed by di-*N*-methylation using SAM to give **anhydrotetracycline**. In the last steps, C-6 and C-4 are hydroxylated, and NADPH reduction of the C-5a/11a double bond generates **oxytetracycline**. Omission of the C-4 hydroxylation would provide **tetracycline**. The pathway to **chlortetracycline** involves an addition chlorination step; this occurs on 4-keto-6-methylpretetramide at the nucleophilic site *para* to the phenol in ring D. The subsequent steps are then identical.

Box 3.20

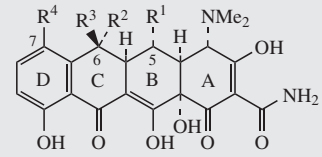
Tetracyclines

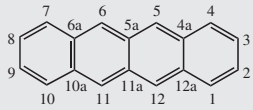
The tetracyclines (Figure 3.100) are a group of broad-spectrum, orally active antibiotics produced by cultures of *Streptomyces* species. **Chlortetracycline** isolated from *Streptomyces aureofaciens* was the first of the group to be discovered, closely followed by **oxytetracycline** from cultures of *S. rimosus*. **Tetracycline** was found as a minor antibiotic in *S. aureofaciens*, but may be produced in quantity by utilizing a mutant strain blocked in the chlorination step (Figure 3.99). Similarly, the early C-6 methylation step can also be blocked, and such mutants accumulate 6-demethyltetracyclines, e.g. demeclocycline (demethylchlorotetracycline). These reactions can also be inhibited in the normal strain of *S. aureofaciens* by supplying cultures with either aminopterin (which inhibits C-6 methylation) or mercaptothiazole (which inhibits C-7 chlorination). Oxytetracycline from *S. rimosus* lacks the chlorine substituent, but has an additional 5α-hydroxyl group, introduced late in the pathway. Only minor alterations can be made to the basic tetracycline structure to modify the antibiotic activity, and these are at positions 5, 6, and 7. Other functionalities in the molecule are all essential to retain activity. Semi-synthetic tetracyclines used clinically include **methacycline**, obtained by a dehydration reaction from oxytetracycline, and **doxycycline**, via reduction of the 6-methylene in methacycline. **Minocycline** contains a 7-dimethylamino group and is produced by a sequence involving aromatic nitration. **Lymecycline** is an example of an antibiotic developed by chemical modification of the primary amide function at C-2.

Tetracyclines have both amino and phenolic functions, and are thus amphoteric compounds; they are more stable in acid than under alkaline conditions. They are thus suitable for oral administration and are absorbed satisfactorily. However, because of the sequence of phenol and carbonyl substituents in the structures, they act as chelators and complex with metal ions, especially calcium, aluminium, iron, and magnesium. Accordingly, they should not be administered with foods such as milk and dairy products (which have a high calcium content), aluminium- and magnesium-based antacid preparations, iron supplements, etc., otherwise erratic and unsatisfactory absorption will occur. A useful feature of doxycycline and minocycline is that their absorptions are much less affected by metal ions. Chelation of tetracyclines with calcium also precludes their use in children

Box 3.20 (continued)

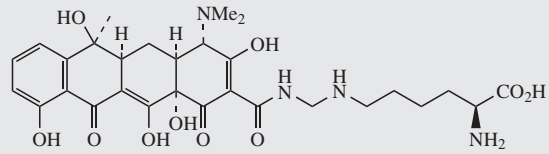
Tetracyclines



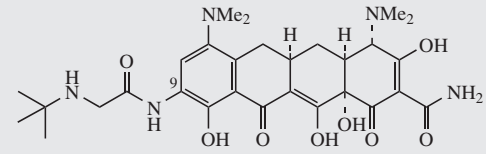


tetracene numbering

	R^1	R^2	R^3	R^4	
	5	6α	6β	7	
tetracycline	H	Me	OH	H	natural
chlortetracycline	H	Me	OH	Cl	natural
oxytetracycline	OH	Me	OH	H	natural
demeclocycline	H	H	OH	H	natural
methacycline	OH	$=CH_2$		H	semi-synthetic
doxycycline	OH	Me	H	H	semi-synthetic
minocycline	H	H	H	NMe_2	semi-synthetic



lymecycline



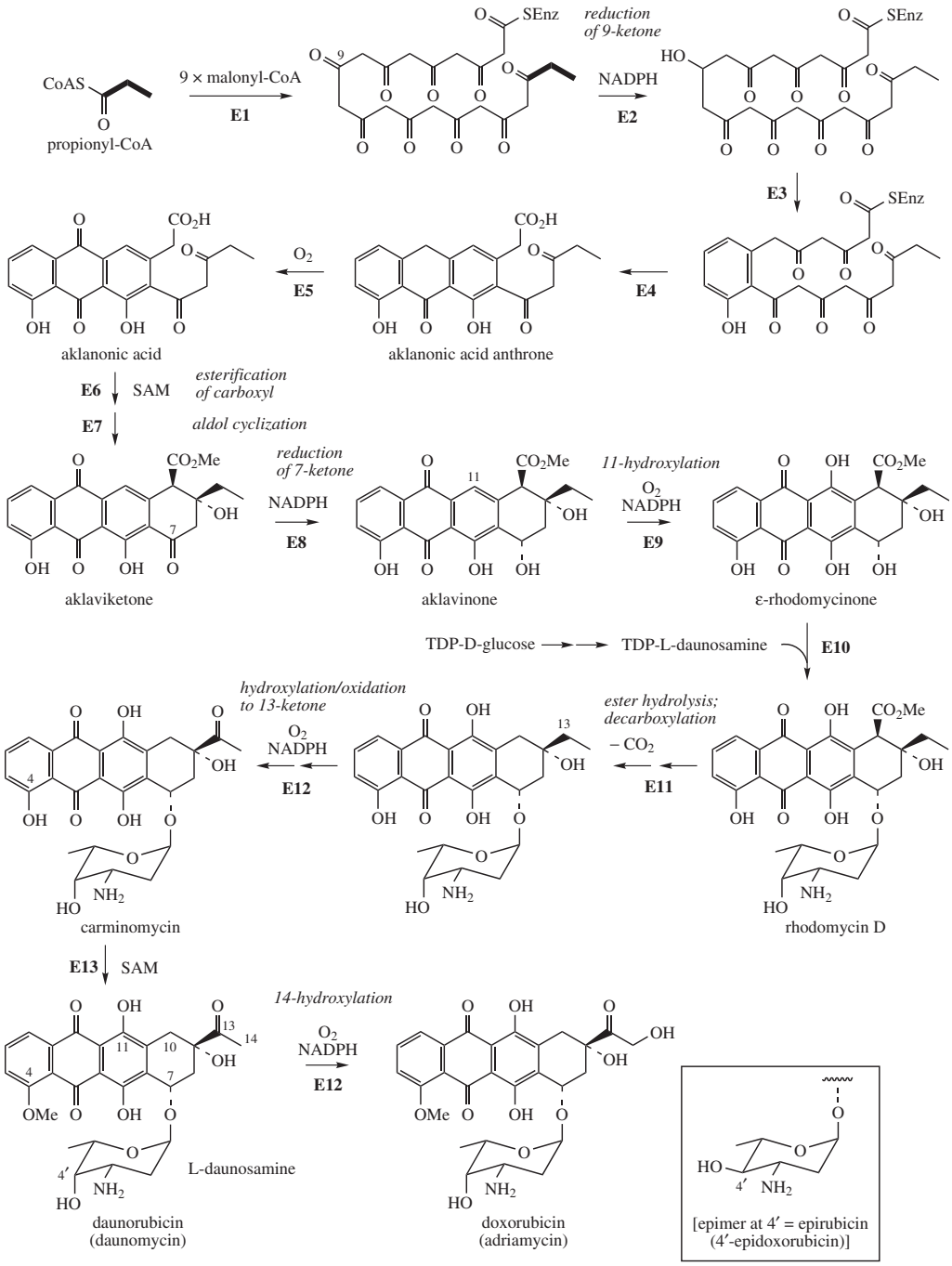
tigecycline

Figure 3.100

developing their adult teeth, and in pregnant women, since the tetracyclines become deposited in growing teeth and bone. In children, this would cause unsightly and permanent staining of teeth with the chelated yellow tetracycline.

Although the tetracycline antibiotics have a broad spectrum of activity spanning Gram-negative and Gram-positive bacteria, their value has decreased as bacterial resistance has developed in pathogens such as *Pneumococcus, Staphylococcus, Streptococcus*, and *E. coli*. These organisms appear to have evolved two main mechanisms of resistance: bacterial efflux and ribosome protection. In bacterial efflux, a membrane-embedded transport protein exports the tetracycline out of the cell before it can exert its effect. Ribosomal protection releases tetracyclines from the ribosome, the site of action (see below). Nevertheless, tetracyclines are the antibiotics of choice for infections caused by *Chlamydia, Mycoplasma, Brucella*, and *Rickettsia*, and are valuable in chronic bronchitis due to activity against *Haemophilus influenzae*. They are also used systemically to treat severe cases of acne, helping to reduce the frequency of lesions by their effect on skin flora. There is little significant difference in the antimicrobial properties of the various agents, except for **minocycline**, which has a broader spectrum of activity and, being active against *Neisseria meningitides*, is useful for prophylaxis of meningitis. The individual tetracyclines do have varying bioavailabilities, however, which may influence the choice of agent. **Tetracycline** and **oxytetracycline** are probably the most commonly prescribed agents, whilst chlortetracycline and methacycline have both been superseded. Tetracyclines are formulated for oral application or injection, as ear and eye drops, and for topical use on the skin. **Doxycycline** is the agent of choice for treating Lyme disease (caused by the spirochaete *Borellia burgdorferi*), and also finds use as a prophylactic against malaria in areas where there is widespread resistance to chloroquine and mefloquine (see page 382). A new generation of tetracyclines known as glycylcyclines has been developed to counter resistance and provide higher antibiotic activity. The first of these in general use is **tigecycline** (Figure 3.100), a glycylamido derivative of minocycline. It is synthesized via nitration of minocycline, itself produced from tetracycline. Tigecycline maintains antibiotic spectrum and potency, whilst overcoming resistance arising from both ribosomal protection and efflux mechanisms.

Their antimicrobial activity arises by inhibition of protein synthesis. This is achieved by interfering with the binding of aminoacyl-tRNA to acceptor sites on the ribosome by disrupting the codon–anticodon interaction (see page 422). Evidence points to a single strong binding site on the smaller 30S subunit of the ribosome. Although tetracyclines can also bind to mammalian ribosomes, there appears to be preferential penetration into bacterial cells, and there are few major side-effects from using these antibiotics.



E1: DpsA, DpsB, DpsC, DpsD, DpsG
E2: DpsE (9-ketoreductase)
E3: DpsF (cyclase/aromatase)
E4: DpsY (cyclase)
E5: DnrG (oxygenase)
E6: DnrC (*O*-methyltransferase)
E7: DnrD (cyclase)
E8: DnrH (reductase)
E9: DnrF (hydroxylase)
E10: DnrS (glycosyltransferase)
E11: DnrP (methylesterase)
E12: DoxA (oxygenase)
E13: DnrK (*O*-methyltransferase)

Figure 3.101

A number of **anthracycline antibiotics**, e.g. **doxorubicin** (Figure 3.101) from *Streptomyces peuceticus* and **daunorubicin** from *Streptomyces coeruleorubicus*, have structurally similar tetracyclic skeletons and would appear to be related to the tetracyclines. However, anthraquinone derivatives are intermediates in anthracycline biosynthesis, and the fourth ring is constructed later. The folding of the poly-β-keto chain is also rather different from that seen with tetracyclines; as a result, the end-of-chain carboxyl is ultimately lost through decarboxylation. This carboxyl is actually retained for a considerable portion of the pathway, and is even protected against decarboxylation by methylation to the ester, until no longer required. The starter group for the type II PKS is **propionyl-CoA** (Figure 3.101) and a specific 9-ketoreductase component acts upon the full-length polyketide chain, as seen with the tetracyclines. The initial enzyme-free product is the anthrone, which is oxidized to the anthraquinone **alklanonic acid**. The carboxylic acid is esterified, and then the fourth ring can be elaborated by a simple aldol reaction. Most of the modifications which subsequently occur during the biosynthetic pathway are easily predictable. A feature of note in molecules such as doxorubicin and daunorubicin is the amino sugar L-daunosamine which originates from TDPglucose (thymidine diphosphoglucose; compare UDPglucose, page 31) and is introduced in the latter stages of the sequence. Hydroxylation of daunorubicin to doxorubicin is the very last step. Doxorubicin and daunorubicin are used as antitumour drugs rather than antimicrobial agents [Box 3.21]. They act primarily at the DNA level and so also have cytotoxic properties. Doxorubicin in particular is a highly successful and widely used antitumour agent, employed in the treatment of leukaemias, lymphomas and a variety of solid tumours.

The production of modified aromatic polyketides by genetically engineered type II PKSs is not quite so 'obvious' as with the modular type I enzymes, but significant progress has been made in many systems. Each type II PKS contains a minimal set of three protein subunits, two β-ketoacyl synthase (KS) subunits and an ACP to which the growing chain is attached. Additional subunits, including KRs, cyclases (CYC), and aromatases

Box 3.21

Anthracycline Antibiotics

Doxorubicin (**adriamycin**; Figure 3.101) is produced by cultures of *S. peucetius* var *caesius* and is one of the most successful and widely used antitumour drugs. The organism is a variant of *S. peucetius*, a producer of daunorubicin (see below), in which mutagen treatment resulted in expression of a latent hydroxylase enzyme and thus synthesis of doxorubicin by 14-hydroxylation of daunorubicin. Doxorubicin has one of the largest spectrum of antitumour activity shown by antitumour drugs and is used to treat acute leukaemias, lymphomas, and a variety of solid tumours. It is administered by intravenous injection and is largely excreted in the bile. The planar anthracycline molecule intercalates between base pairs on the DNA helix. The sugar unit provides further binding strength and also plays a major role in sequence recognition for the binding. Intercalation is central to doxorubicin's primary mode of action, inhibition of the enzyme topoisomerase II, which is responsible for cleaving and resealing of double-stranded DNA during replication (see page 155). Common toxic effects include nausea and vomiting, bone marrow suppression, hair loss, and local tissue necrosis, with cardiotoxicity at higher dosage.

Daunorubicin (Figure 3.101) is produced by *S. coeruleorubidus* and *S. peucetius*. Though similar to doxorubicin in its biological and chemical properties, it is no longer used therapeutically to any extent. It has a much less favourable therapeutic index than doxorubicin, and the markedly different effectiveness as an antitumour drug is not fully understood, though differences in metabolic degradation may be responsible. **Epirubicin** (Figure 3.101), the semi-synthetic 4′-epimer of doxorubicin, is particularly effective in the treatment of breast cancer, producing lower side-effects than doxorubicin. The antileukaemics **aclarubicin** from *Streptomyces galilaeus*, a complex glycoside of aklavinone (Figure 3.101), and the semi-synthetic **idarubicin** are shown in Figure 3.102. These compounds are structurally related to doxorubicin but can show increased activity with less cardiotoxicity. The principal disadvantage of all of these agents is their severe cardiotoxicity, which arises through inhibition of cardiac Na^+, K^+-ATPase.

Mitoxantrone (**mitozantrone**) (Figure 3.102) is a synthetic analogue of the anthracyclinones in which the non-aromatic ring and the aminosugar have both been replaced with aminoalkyl side-chains. This agent has reduced toxicity compared with doxorubicin and is effective in the treatment of solid tumours and leukaemias. In addition, it is currently proving useful in multiple sclerosis treatment, where it can reduce the frequency of relapses. Despite considerable synthetic research, however, relatively few anthracycline analogues have proved superior to doxorubicin itself, though some newer ones, e.g. **sabarubicin** and **galarubicin** (Figure 3.102), are in clinical trials.

Box 3.21 (continued)

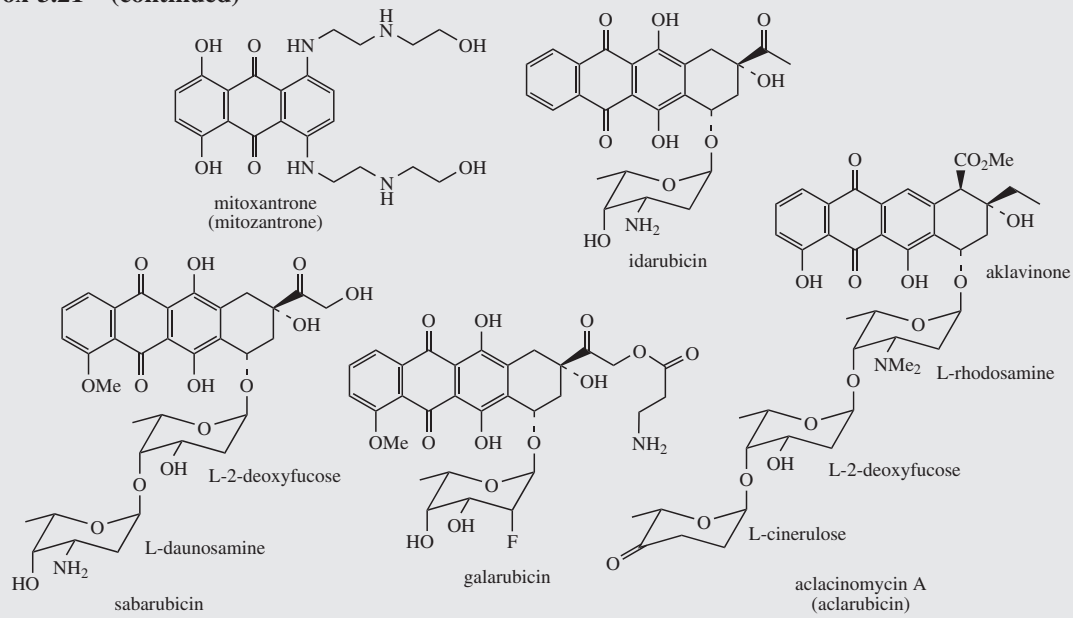


Figure 3.102

(ARO) are responsible for modification of the nascent chain to form the final cyclized structure. Novel polyketides have been generated by manipulating type II PKSs, exchanging KS, CYC and ARO subunits among different systems. However, because of the highly reactive nature of poly-β-keto chains, the cyclizations that occur with the modified gene product frequently vary from those in the original compound. Compared with type I PKSs, the formation of new products with predictable molecular structures has proven less controllable.

FURTHER READING

Fatty Acid Synthase

Jenni S, Leibundgut M, Maier T and Ban N (2006) Architecture of a fungal fatty acid synthase at 5 Å resolution. *Science* **311**, 1263–1267.

Kolter T (2007) The fatty acid factory of yeasts. *Angew Chem Int Ed* **46**, 6772–6775.

Lu H and Tonge PJ (2008) Inhibitors of FabI, an enzyme drug target in the bacterial fatty acid biosynthesis pathway. *Acc Chem Res* **41**, 11–20.

Maier T, Jenni S and Ban N (2006) Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science* **311**, 1258–1262.

Smith S (2006) Architectural options for a fatty acid synthase. *Science* **311**, 1251–1252.

Smith S and Tsai S-C (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep* **24**, 1041–1072.

Surolia A, Ramya TNC, Ramya V and Surolia N (2004) 'FAS't inhibition of malaria. *Biochem J* **383**, 401–412; corrigenda **384**, 655.

Tasdemir D (2006) Type II fatty acid biosynthesis, a new approach in antimalarial natural product discovery. *Phytochem Rev* **5**, 99–108.

Wright HT and Reynolds KA (2007) Antibacterial targets in fatty acid biosynthesis. *Curr Opin Microbiol* **10**, 447–453.

Fatty Acids and Fats

Behrouzian B and Buist PH (2003) Bioorganic chemistry of plant lipid desaturation. *Phytochem Rev* **2**, 103–111.

Brouwer IA, Geelen A and Katan MB (2006) *n*–3 Fatty acids, cardiac arrhythmia and fatal coronary heart disease. *Prog Lipid Res* **45**, 357–367.

Buist PH (2004) Fatty acid desaturases: selecting the dehydrogenation channel. *Nat Prod Rep* **21**, 249–262.

Buist PH (2007) Exotic biomodification of fatty acids. *Nat Prod Rep* **24**, 1110–1127.

Chatgilialoglu C and Ferreri C (2005) Trans lipids: the free radical path. *Acc Chem Res* **38**, 411–448.

Cunnane SC (2003) Problems with essential fatty acids: time for a new paradigm? *Prog Lipid Res* **42**, 544–568.

Domergue F, Abbadi A and Heinz E (2005) Relief for fish stocks: oceanic fatty acids in transgenic oilseeds. *Trends Plant Sci* **10**, 112–116.

Drexler H, Spiekermann P, Meyer A, Domergue F, Zank T, Sperling P, Abbadi A and Heinz E (2003) Metabolic engineering of fatty acids for breeding of new oilseed crops: strategies, problems and first results. *J Plant Physiol* **160**, 779–802.

Kaulmann U and Hertweck C (2002) Biosynthesis of polyunsaturated fatty acids by polyketide synthases. *Angew Chem Int Ed* **41**, 1866–1869.

Murphy DJ (2002) Biotechnology and the improvement of oil crops – genes, dreams and realities. *Phytochem Rev* **1**, 67–77.

Sayanova OV and Napier JA (2004) Molecules of interest. Eicosapentaenoic acid: biosynthetic routes and the potential for synthesis in transgenic plants. *Phytochemistry* **65**, 147–158.

Singh SP, Zhou X-R, Liu Q, Stymne S and Green AG (2005) Metabolic engineering of new fatty acids in plants. *Curr Opin Plant Biol* **8**, 197–203.

Tziomalos K, Athyros VG and Mikhailidis DP (2007) Fish oils and vascular disease prevention: an update. *Curr Med Chem* **14**, 2622–2628.

Valentine RC and Valentine DL (2004) *Omega*-3 fatty acids in cellular membranes: a unified concept. *Prog Lipid Res* **43**, 383–402.

Echinacea

Woelkart K and Bauer R (2008) The role of alkamides as an active principle of *Echinacea*. *Planta Med* **73**, 615–623.

Lipstatin

Eisenreich W, Kupfer E, Stohler P, Weber W and Bacher A (2003) Biosynthetic origin of a branched chain analogue of the lipase inhibitor, lipstatin. *J Med Chem* **46**, 4209–4212.

Schuhr CA, Eisenreich W, Goese M, Stohler P, Weber W, Kupfer E and Bacher A (2002) Biosynthetic precursors of the lipase inhibitor lipstatin. *J Org Chem* **67**, 2257–2262.

Prostaglandins, Thromboxanes, Leukotrienes

Blobaum AL and Marnett LJ (2007) Structural and functional basis of cyclooxygenase inhibition. *J Med Chem* **50**, 1425–1441.

Das S, Chandrasekhar S, Yadav JS and Grée R (2007) Recent developments in the synthesis of prostaglandins and analogues. *Chem Rev* **107**, 3286–3337.

De Leval X, Hanson J, David J-L, Masereel B, Pirotte B and Dogne J-M (2004) New developments on thromboxane and prostacyclin modulators. Part II: prostacyclin modulators. *Curr Med Chem* **11**, 1243–1252.

Dogne J-M, de Leval X, Hanson J, Frederich M, Lambermont B, Ghuysen A, Casini A, Masereel B, Ruan K-H, Pirotte B and Kolh P (2004) New developments on thromboxane and prostacyclin modulators. Part I: thromboxane modulators. *Curr Med Chem* **11**, 1223–1241.

Dogné J-M, Supuran CT and Pratico D (2005) Adverse cardiovascular effects of the coxibs. *J Med Chem* **48**, 2251–2257.

Fam SS and Morrow JD (2003) The isoprostanes: unique products of arachidonic acid oxidation – a review. *Curr Med Chem* **10**, 1723–1740.

Garg R, Kurup A, Mekapati SB and Hansch C (2003) Cyclooxygenase (COX) inhibitors: a comparative QSAR study. *Chem Rev* **103**, 703–731.

Jampilek J, Dolezal M, Opletalova V and Hartl J (2006) 5-Lipoxygenase, leukotrienes biosynthesis and potential antileukotrienic agents. *Curr Med Chem* **13**, 117–129.

McGinley CM and van der Donk WA (2003) Enzymatic hydrogen atom abstraction from polyunsaturated fatty acids. *Chem Commun* 2843–2846.

Montuschi P, Barnes P and Roberts LJ (2007) Insights into oxidative stress: the isoprostanes. *Curr Med Chem* **14**, 703–717.

Murakami M and Kudo I (2004) Recent advances in molecular biology and physiology of the prostaglandin E_2-biosynthetic pathway. *Prog Lipid Res* **43**, 3–35.

Roberts SM, Santoro MG and Sickle ES (2002) The emergence of the cyclopentenone prostaglandins as important, biologically active compounds. *J Chem Soc, Perkin Trans 1* 1735–1742.

Rouzer CA and Marnett LJ (2003) Mechanism of free radical oxygenation of polyunsaturated fatty acids by cyclooxygenases. *Chem Rev* **103**, 2239–2304.

Valmsen K, Järving I, Boeglin WE, Varvas K, Koljak R, Pehk T, Brash AR and Samel N (2001) The origin of 15*R*-prostaglandins in the Caribbean coral *Plexaura homomalla*: molecular cloning and expression of a novel cyclooxygenase. *Proc Natl Acad Sci USA* **98**, 7700–7705.

Van der Donk WA, Tsai A-L and Kulmacz RJ (2002) The cyclooxygenase reaction mechanism. *Biochemistry* **41**, 15451–15458.

Warner TD and Mitchell JA (2002) Cyclooxygenase-3 (COX-3): filling in the gaps toward a COX continuum? *Proc Natl Acad Sci USA* **99**, 13371–13373.

Polyketide Synthases: General

Austin MB and Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. *Nat Prod Rep* **20**, 79–110.

Cox RJ (2007) Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org Biomol Chem* **5**, 2010–2026.

Hill AM (2006) The biosynthesis, molecular genetics and enzymology of the polyketide-derived metabolites. *Nat Prod Rep* **23**, 256–320.

Liou GF and Khosla C (2003) Building-block selectivity of polyketide synthases. *Curr Opin Chem Biol* **7**, 279–284.

Moore BS and Hertweck C (2002) Biosynthesis and attachment of novel bacterial polyketide synthase starter units. *Nat Prod Rep* **19**, 70–99.

Shen B (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* **7**, 285–295.

Staunton J and Weissman KJ (2001) Polyketide biosynthesis: a millennium review. *Nat Prod Rep* **18**, 380–416.

Staunton J and Wilkinson B (2001) Combinatorial biosynthesis of polyketides and nonribosomal peptides. *Curr Opin Chem Biol* **5**, 159–164.

Walsh CT (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* **303**, 1805–1810.

Watanabe K and Oikawa H (2007) Robust platform for *de novo* production of heterologous polyketides and nonribosomal peptides in *Escherichia coli*. *Org Biomol Chem* **5**, 593–602.

Polyketide Synthases: Macrolides

Fischbach MA and Walsh CT (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem Rev* **106**, 3468–3496.

Katz L and Ashley GW (2005) Translation and protein synthesis: macrolides. *Chem Rev* **105**, 499–527.

Kopp F and Marahiel MA (2007) Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat Prod Rep* **24**, 735–749.

McDaniel R, Welch M and Hutchinson CR (2005) Genetic approaches to polyketide antibiotics 1. *Chem Rev* **105**, 543–558.

Menzella HG and Reeves CD (2007) Combinatorial biosynthesis for drug development. *Curr Opin Microbiol* **10**, 238–245.

Pohl NL (2002) Nonnatural substrates for polyketide synthases and their associated modifying enzymes. *Curr Opin Chem Biol* **6**, 773–778.

Macrolide Antibiotics

Pal S (2006) A journey across the sequential development of macrolides and ketolides related to erythromycin. *Tetrahedron* **62**, 3171–3200.

Rodriguez E and McDaniel R (2001) Combinatorial biosynthesis of antimicrobials and other natural products. *Curr Opin Microbiol* **4**, 526–534.

Wu Y-J and Su W-G (2001) Recent developments on ketolides and macrolides. *Curr Med Chem* **8**, 1727–1758.

Avermectins

Ikeda H, Nonomiya T and Omura S (2001) Organization of biosynthetic gene cluster for avermectin in *Streptomyces avermitilis*: analysis of enzymatic domains in four polyketide synthases. *J Ind Microbiol Biotechnol* **27**, 170–176.

Polyene Antifungals

Aparicio JF, Mendes MV, Anton N, Recio E and Martin JF (2004) Polyene macrolide antibiotic biosynthesis. *Curr Med Chem* **11**, 1645–1656.

Caffrey P, Lynch S, Flood E, Finnan S and Oliynyk M (2001) Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem Biol* **8**, 713–723; errata (2003) **10**, 93–94.

Zotchev SB (2003) Polyene macrolide antibiotics and their applications in human therapy. *Curr Med Chem* **10**, 211–223.

Ascomycin, Tacrolimus, Rapamycin

Gregory MA, Hong H, Lill RE, Gaisser S, Petkovic H, Low L, Sheehan LS, Carletti I, Ready SJ, Ward MJ, Kaja AL, Weston AJ, Challis IR, Leadlay PF, Martin CJ, Wilkinson B and Sheridan RM (2006) Rapamycin biosynthesis: elucidation of gene product function. *Org Biomol Chem* **4**, 3565–3568.

Lowden PAS, Wilkinson B, Böhm GA, Handa S, Floss HG, Leadlay PF and Staunton J (2001) Origin and true nature of the starter unit for the rapamycin polyketide synthase. *Angew Chem Int Ed* **40**, 777–779.

Mann J (2001) Natural products as immunosuppressive agents. *Nat Prod Rep* **18**, 417–430.

Wu K, Chung L, Revill WP, Katz L and Reeves CD (2000) The FK520 gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* **251**, 81–90.

Epothilones

Altmann K-H (2004) The merger of natural product synthesis and medicinal chemistry: on the chemistry and chemical biology of epothilones. *Org Biomol Chem* **2**, 2137–2152.

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.

Feyen F, Cachoux F, Gertsch J, Wartmann M and Altmann K-H (2008) Epothilones as lead structures for the synthesis-based discovery of new chemotypes for microtubule stabilization. *Acc Chem Res* **41**, 21–31.

Nicolaou KC, Ritzén A and Namoto K (2001) Recent developments in the chemistry, biology and medicine of the epothilones. *Chem Commun* 1523–1535.

Rivkin A, Chou T-C and Danishefsky SJ (2005) On the remarkable antitumor properties of fludelone: how we got there. *Angew Chem Int Ed* **44**, 2838–2850.

Rifamycins

Floss HG and Yu T-W (2005) Rifamycins – mode of action, resistance, and biosynthesis. *Chem Rev* **105**, 621–632.

Zearalenone

Gaffoor I and Trail F (2006) Characterization of two polyketide synthase genes involved in zearalenone biosynthesis in *Gibberella zeae*. *Appl Environ Microbiol* **72**, 1793–1799.

Winssinger N and Barluenga S (2007) Chemistry and biology of resorcylic acid lactones. *Chem Commun* 22–36.

Discodermolide

Burlingame MA, Mendoza E, Ashley GW and Myles DC (2006) Synthesis of discodermolide intermediates from engineered polyketides. *Tetrahedron Lett* **47**, 1209–1211.

Smith AB and Freeze BS (2008) (+)-Discodermolide: total synthesis, construction of novel analogues, and biological evaluation. *Tetrahedron* **64**, 261–298.

Mupirocin

El-Sayed AK, Hothersall J, Cooper SM, Stephens E, Simpson TJ and Thomas CM (2003) Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. *Chem Biol* **10**, 419–430.

Polyethers

Bhatt A, Stark CBW, Harvey BM, Gallimore AR, Demydchuk Y, Spencer JB, Staunton J and Leadlay PF (2005) Accumulation of an *E,E,E*-triene by the monensin-producing polyketide synthase when oxidative cyclization is blocked. *Angew Chem Int Ed* **44**, 7075–7078.

Dounay AB and Forsyth CJ (2002) Okadaic acid: the archetypal serine/threonine protein phosphatase inhibitor. *Curr Med Chem* **9**, 1939–1980.

Fernández JJ, Candenas ML, Souto ML, Trujillo MM and Norte M (2002) Okadaic acid, useful tool for studying cellular processes. *Curr Med Chem* **9**, 229–262.

Inoue M (2004) Convergent syntheses of polycyclic ethers. Illustrations of the utility of acetal-linked intermediates. *Org Biomol Chem* **2**, 1811–1817.

Nakata T (2005) Total synthesis of marine polycyclic ethers. *Chem Rev* **105**, 4314–4347.

Zaragozic Acids (Squalestatins)

Armstrong A and Blench TJ (2002) Recent synthetic studies on the zaragozic acids (squalestatins). *Tetrahedron* **58**, 9321–9349.

Cox RJ, Glod F, Hurley D, Lazarus CM, Nicholson TP, Rudd BAM, Simpson TJ, Wilkinson B and Zhang Y (2004) Rapid cloning and expression of a fungal polyketide synthase gene involved in squalestatin biosynthesis. *Chem Commun* 2260–2261.

Statins

Luo J-D and Chen AF (2003) Perspectives on the cardioprotective effects of statins. *Curr Med Chem* **10**, 1593–1601.

Müller M (2005) Chemoenzymatic synthesis of building blocks for statin side chains. *Angew Chem Int Ed* **44**, 362–365.

Oikawa H and Tokiwano T (2004) Enzymatic catalysis of the Diels–Alder reaction in the biosynthesis of natural products. *Nat Prod Rep* **21**, 321–352.

Rajanikant GK, Zemke D, Kassab M and Majid A (2007) The therapeutic potential of statins in neurological disorders. *Curr Med Chem* **14**, 103–112.

Roche VF (2005) Antihyperlipidemic statins: a self-contained, clinically relevant medicinal chemistry lesson. *Am J Pharm Educ* **69**, 546–560.

Sorensen JL and Vederas JC (2003) Monacolin N, a compound resulting from derailment of type I iterative polyketide synthase function en route to lovastatin. *Chem Commun* 1492–1493.

Stocking EM and Williams RM (2003) Chemistry and biology of biosynthetic Diels–Alder reactions. *Angew Chem Int Ed* **42**, 3078–3115.

Polyketide Synthases: Aromatics

Hertweck C, Luzhetskyy A, Rebets Y and Bechthold A (2007) Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* **24**, 162–190.

Rix U, Fischer C, Remsing LL and Rohr J (2002) Modification of post-PKS tailoring steps through combinatorial biosynthesis. *Nat Prod Rep* **19**, 542–580.

Thomas R (2001) A biosynthetic classification of fungal and Streptomycete fused-ring aromatic polyketides. *ChemBioChem* **2**, 612–627.

Watanabe K, Praseuth AP and Wang CCC (2007) A comprehensive and engaging overview of the type III family of polyketide synthases. *Curr Opin Chem Biol* **11**, 279–286.

Anthraquinones

Gill M (2001) The biosynthesis of pigments in Basidiomycetes. *Aust J Chem* **54**, 721–734.

Hypericum/St John's Wort

Adam P, Arigoni D, Bacher A and Eisenreich W (2002) Biosynthesis of hyperforin in *Hypericum perforatum*. *J Med Chem* **45**, 4786–4793.

Barnes J, Anderson LA and Phillipson JD (2001) St John's wort (*Hypericum perforatum* L.): a review of its chemistry, pharmacology and clinical properties. *J Pharm Pharmac* **53**, 583–600.

Beerhues L (2006) Molecules of interest. Hyperforin. *Phytochemistry* **67**, 2201–2207.

Singh YN (2005) Potential for interaction of kava and St. John's wort with drugs. *J Ethnopharmacol* **100**, 108–113.

Usnic Acid

Ingólfsdóttir K (2002) Molecules of interest. Usnic acid. *Phytochemistry* **61**, 729–736.

Cannabis

Amar MB (2006) Cannabinoids in medicine: a review of their therapeutic potential. *J Ethnopharmacol* **105**, 1–25.

Ashton JC, Wright JL, McPartland JM and Tyndall Joel DA (2008) Cannabinoid CB1 and CB2 receptor ligand specificity and the development of CB2-selective agonists. *Curr Med Chem* **15**, 1428–1443.

Drysdale AJ and Platt B (2003) Cannabinoids: mechanisms and therapeutic applications in the CNS. *Curr Med Chem* **10**, 2719–2732.

Mechoulam R, Sumariwalla PF, Feldmann M and Gallily R (2005) Cannabinoids in models of chronic inflammatory conditions. *Phytochem Rev* **4**, 11–18.

Merzouki A and Mesa JM (2002) Concerning kif, a *Cannabis sativa* L. preparation smoked in the Rif mountains of northern Morocco. *J Ethnopharmacol* **81**, 403–406.

Raharjo TJ and Verpoorte R (2004) Methods for the analysis of cannabinoids in biological materials: a review. *Phytochem Anal* **15**, 79–94.

Taura F, Sirikantaramas S, Shoyama Y, Yoshikai K, Shoyama Y and Morimoto S (2007) Cannabidiolic-acid synthase, the chemotype-determining enzyme in the fiber-type *Cannabis sativa*. *FEBS Lett* **581**, 2929–2934.

Endocannabinoids

Ahn K, McKinney MK and Cravatt BF (2008) Enzymatic pathways that regulate endocannabinoid signaling in the nervous system. *Chem Rev* **108**, 1687–1707.

Di Marzo V, Bisogno T and De Petrocellis L (2007) Endocannabinoids and related compounds: walking back and forth between plant natural products and animal physiology. *Chem Biol* **14**, 741–756.

Lambert DM and Fowler CJ (2005) The endocannabinoid system: drug targets, lead compounds, and potential therapeutic applications. *J Med Chem* **48**, 5059–5087.

Lambert DM, Vandevoorde S, Jonsson K-O and. Fowler CJ (2002) The palmitoylethanolamide family: a new class of anti-inflammatory agents? *Curr Med Chem* **9**, 663–674.

Ueda N, Okamoto Y and Tsuboi K (2005) Endocannabinoid-related enzymes as drug targets with special reference to *N*-acylphosphatidylethanolamine-hydrolyzing phospholipase D. *Curr Med Chem* **12**, 1413–1422.

Wendeler M and Kolter T (2003) Inhibitors of endocannabinoid degradation: potential therapeutics for neurological disorders. *Angew Chem Int Ed* **42**, 2938–2941.

Aflatoxins

Bhatnagar D, Ehrlich KC and Cleveland TE (2003) Molecular genetic analysis and regulation of aflatoxin biosynthesis. *Appl Microbiol Biotechnol* **61**, 83–93.

Watanabe CMH and Townsend CA (2002) Initial characterization of a type I fatty acid synthase and polyketide synthase multienzyme complex NorS in the biosynthesis of aflatoxin B_1. *Chem Biol* **9**, 981–988.

Yabe K and Nakajima H (2004) Enzyme reactions and genes in aflatoxin biosynthesis. *Appl Microbiol Biotechnol* **64**, 745–755.

Tetracyclines

Petković H, Cullum J, Hranueli D, Hunter IS, Perić-Concha N, Pigac J, Thamchaipenet A, Vujaklija D and Long PF (2006) Genetics of *Streptomyces rimosus*, the oxytetracycline producer. *Microbiol Mol Biol Rev* **70**, 704–728.

Zhang W, Wenjun S, Wang CCC and Tang Y (2007) Investigation of early tailoring reactions in the oxytetracycline biosynthetic pathway. *J Biol Chem* **282**, 25717–25725.

Anthracyclines

Lee TS, Khosla C and Tang Y (2005) Engineered biosynthesis of aklanonic acid analogues. *J Am Chem Soc* **127**, 12254–12262.