Natural products

51

Connections

Building on:

- Stereochemistry ch16
- Conformational analysis ch18
- Enolate chemistry and synthesis ch24–ch30
- Pericyclic reactions ch35–ch36
- Rearrangement and fragmentation ch37–ch38
- Radicals ch39
- Chemistry of life ch49
- Mechanisms in biological chemistry ch50

Arriving at:

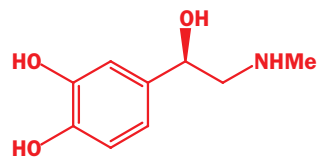
- Natural products are made by secondary metabolism
- Natural products come in enormous variety, but fall mainly into four types: alkaloids, polyketides, terpenes, and steroids
- Alkaloids are amines made from amino acids
- Pyrrolidine alkaloids from ornithine; benzylisoquinoline alkaloids from tyrosine
- Morphine alkaloids are made by radical cyclizations
- Fatty acids are built up from acetyl CoA and malonyl CoA subunits
- Polyketides are unreduced variants of fatty acids
- Terpenes are made from mevalonic acid
- Steroids are tetracyclic terpene derivatives
- Biomimetic synthesis: learning from Nature

Looking forward to:

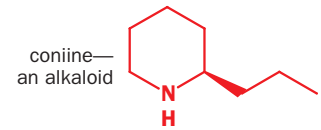
- Organic synthesis ch53

Introduction

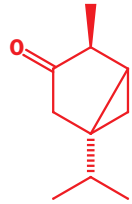
By **natural products**, we mean the molecules of nature. Of course, all life is made of molecules, and we will not be discussing in great detail the major biological molecules, such as proteins and nucleic acids, which we looked at in Chapters 49 and 50. In this chapter we shall talk much more about molecules such as adrenaline (epinephrine). Adrenaline is a human hormone. It is produced in moments of stress and increases our blood pressure and heart rate ready for 'fight or flight'. You've got to sit an exam tomorrow—surge of adrenaline. To an organic chemist adrenaline is intensely interesting because of its remarkable biological activity—but it is also a molecule whose chemical reactions can be studied, whose NMR spectrum can be analysed, which can be synthesized, and which can be imitated in the search for new medicines.



adrenaline



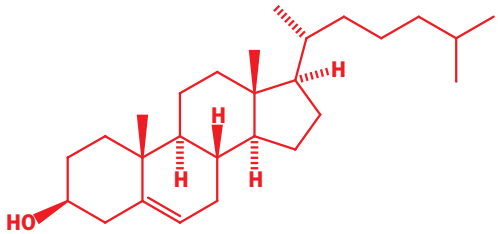
coniine—
an alkaloid



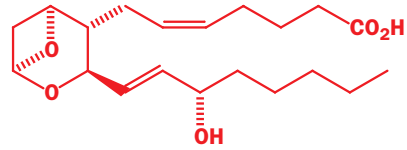
thujone—
a terpene

By the end of this chapter we hope you will be able to recognize some basic classes of natural products and know a bit about their chemistry. We will meet **alkaloids** such as coniine, the molecule in hemlock that killed Socrates, and **terpenes** such as thujone, which was probably the toxin in absinthe that killed the nineteenth-century artists in Paris.

Then there are the ambiguous natural products such as the **steroid** cholesterol, which may cause innumerable deaths through heart disease but which is a vital component of cell walls, and the **polyketide** thromboxane, one drop of which would instantaneously clot all the blood in your body but without which you would bleed to death if you cut yourself.

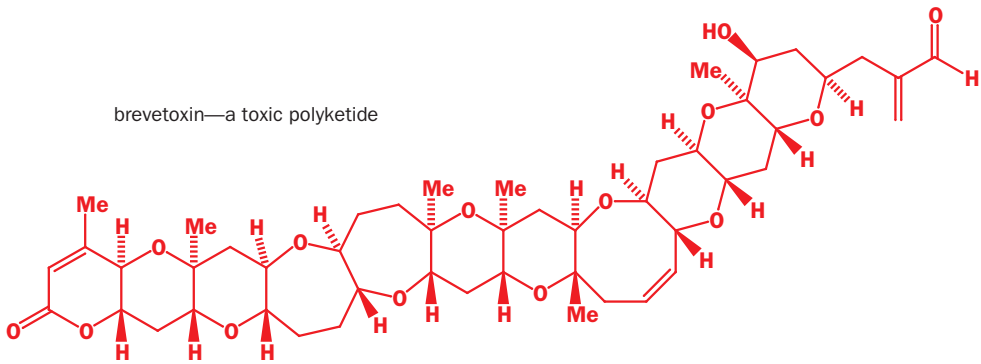


cholesterol—a steroid



thromboxane A_2—a polyketide

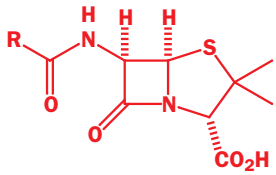
We will look at the structural variety within these four important classes and beyond, from perhaps the smallest natural product, nitric oxide, NO (which controls penile erections in men), to something approaching the largest—the polyketide brevetoxin, the algal product in 'red tides', which appear in coastal waters from time to time and kill fish and those who eat the fish.



brevetoxin—a toxic polyketide

Many natural products are the source of important life-saving drugs—consider the millions of lives saved by penicillin, a family of **amino acid** metabolites.

▶ Before moving on, just pause to admire brevetoxin, a wonderful and deadly molecule. Look at the alternating oxygen atoms on the top and bottom faces of alternate rings. Look at the rings themselves—six-, seven-, and eight-membered but each with one and no more than one oxygen atom. Trace the continuous carbon chain running from the lactone carbonyl group in the bottom left-hand corner to the aldehyde carbonyl in the top right. There is no break in this chain and, other than the methyl groups, no branch. With 22 stereogenic centres, this is a beautiful piece of molecular architecture. If you want to read more about brevetoxin, read the last chapter in Nicolaou and Sorensen's *Classics in total synthesis*, VCH, 1996.



penicillin
e.g. penicillin G; R = $PhCH_2$

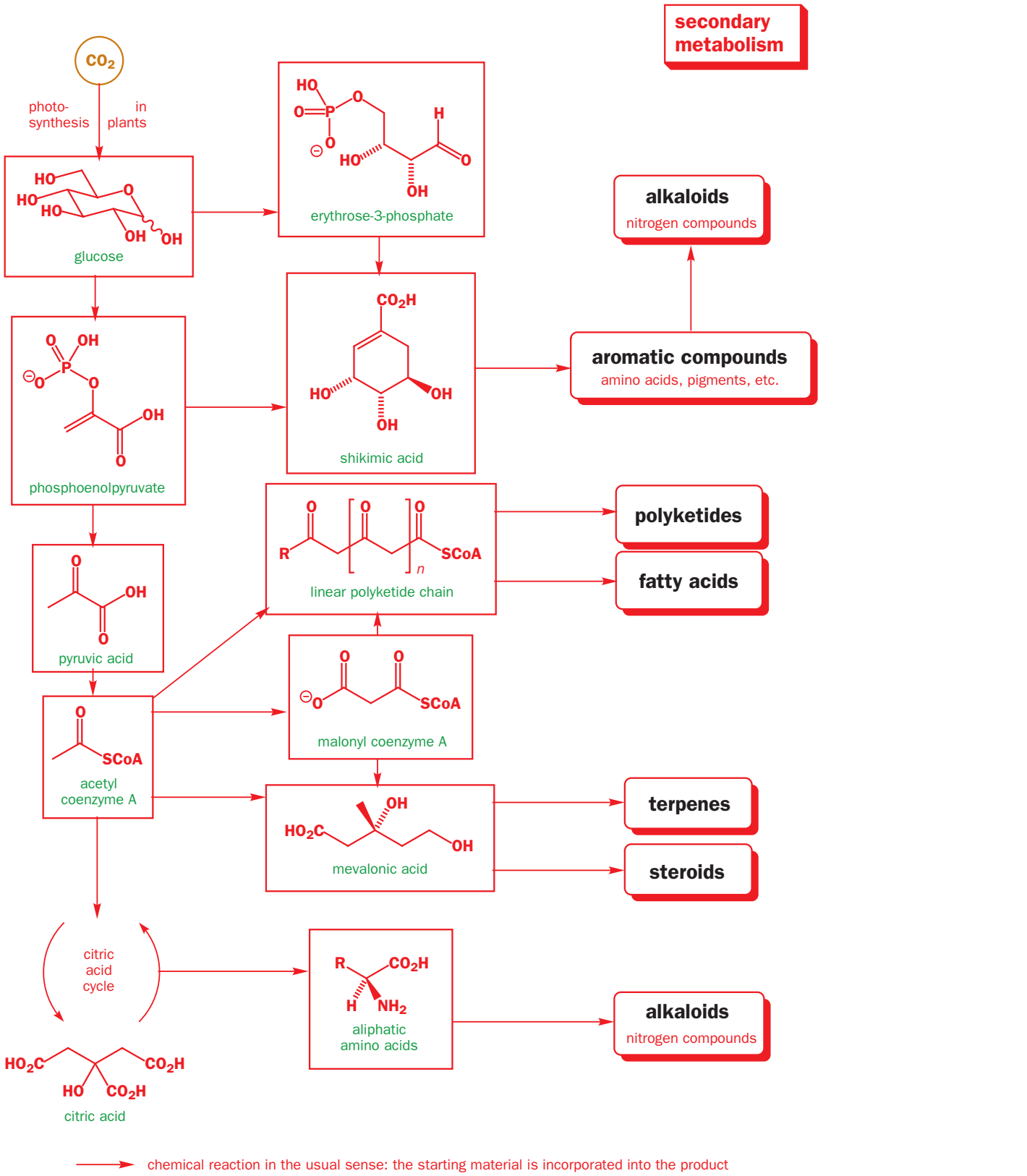
Natural products come from secondary metabolism

The chemical reactions common to all living things involve the **primary metabolism** of the 'big four' we met in Chapter 49—nucleic acids, proteins, carbohydrates, and lipids. Now we must look at chemical reactions that are more restricted. They occur perhaps in just one species, though more commonly in several. They are obviously, then, not essential for life, though they usually help survival. These are the products of **secondary metabolism**.

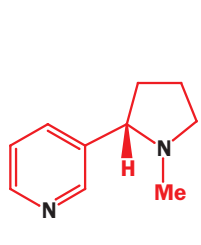
The exploration of the compounds produced by the secondary metabolism of plants, microorganisms, fungi, insects, mammals, and every other type of living thing has hardly begun. Even so, the variety and richness of the structures are overwhelming. Without some kind of classification the task of description would be hopeless. We are going to use a biosynthetic classification, grouping substances not by species but by methods of biological synthesis. Though every species is different, the basic chemical reactions are shared by all. The chart on p. 1415 relates closely to the chart of primary metabolism in the previous chapter.

Alkaloids are basic compounds from amino acid metabolism

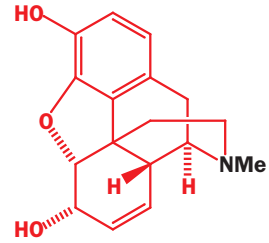
Alkaloids were known in ancient times because they are easy to extract from plants and some of them have powerful and deadly effects. Any plant contains millions of chemical compounds, but some plants, like the deadly nightshade, can be mashed up and extracted with aqueous acid to give a few compounds soluble in that medium, which precipitate on neutralization. These compounds were seen to be 'like alkali' and Meissner, the apothecary from Halle, in 1819 named them 'alkaloids'. Lucrezia Borgia already knew all about this and put the deadly nightshade extract atropine in her eyes (to make her look beautiful: atropine dilates the pupils) and in the drinks of her


secondary metabolism
CO_2
photo-synthesis
in plants
glucose
erythrose-3-phosphate
alkaloids
nitrogen compounds
shikimic acid
aromatic compounds
amino acids, pigments, etc.
phosphoenolpyruvate
linear polyketide chain
polyketides
fatty acids
pyruvic acid
malonyl coenzyme A
acetyl coenzyme A
terpenes
mevalonic acid
steroids
citric acid cycle
aliphatic amino acids
alkaloids
nitrogen compounds
citric acid
chemical reaction in the usual sense: the starting material is incorporated into the product

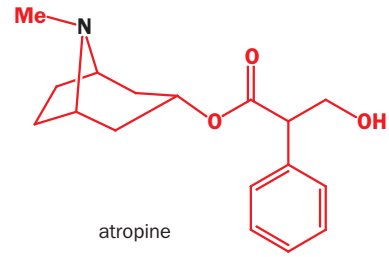
political adversaries to avoid any trouble in the future. Now, we would simply say that they are basic because they are amines. Here is a selection with the basic amino groups marked in black.



nicotine



morphine

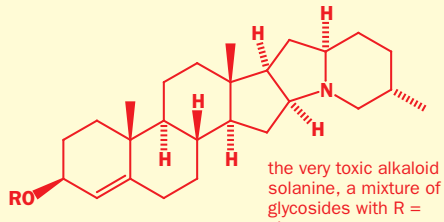


atropine

Natural products are often named by a combination of the name of the organism from which they are isolated and a chemical part name. These compounds are all am*ine*s so all their names end in '-ine'. They appear very diverse in structure but all are made in nature from amino acids, and we will look at three types.

Solanaceae alkaloids

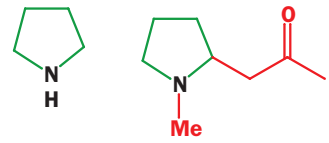
The Solanaceae family includes not only deadly nightshade (*Atropa belladonna*—hence atropine) plants but also potatoes and tomatoes. Parts of these plants also contain toxic alkaloids: for example, you should not eat green potatoes because they contain the toxic alkaloid solanine.



the very toxic alkaloid solanine, a mixture of glycosides with R = glucose, mannose, etc.

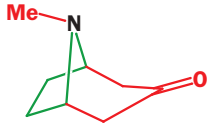
Atropine is a racemic compound but the (*S*)-enantiomer occurs in henbane (*Hyoscyamus niger*) and was given a different name, hyoscyamine, before the structures were known. In fact, hyoscyamine racemizes very easily just on heating in water or on treatment with weak base. This is probably what happens in the deadly nightshade plant.

Pyrrolidine alkaloids are made from the amino acid ornithine



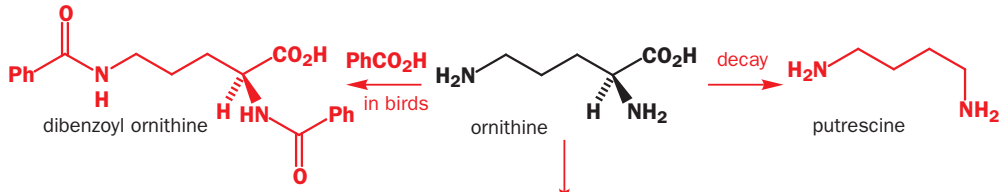
pyrrolidine

hygrine



tropinone

Pyrrolidine is the simple five-membered cyclic amine and pyrrolidine alkaloids contain this ring somewhere in their structure. Both nicotine and atropine contain a pyrrolidine ring as do hygrine and tropinone. All are made in nature from ornithine. Ornithine is an amino acid not usually found in proteins but most organisms use it, often in the excretion of toxic substances. If birds are fed benzoic acid ($PhCO_2H$) they excrete dibenzoyl ornithine. When dead animals decay, the decarboxylation of ornithine leads to putrescine which, as its name suggest, smells revolting. It is the 'smell of death'.



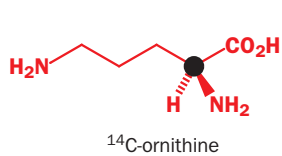
dibenzoyl ornithine

ornithine

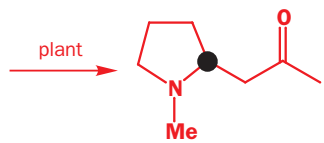
putrescine

pyrrolidine alkaloids

Biosynthetic pathways are usually worked out by isotopic labelling of potential precursors and we shall mark the label with a coloured blob. If ornithine is labelled with ^{14}C and fed to the plant, labelled hygrine is isolated.

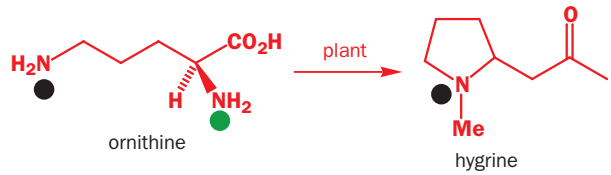


^{14}C-ornithine



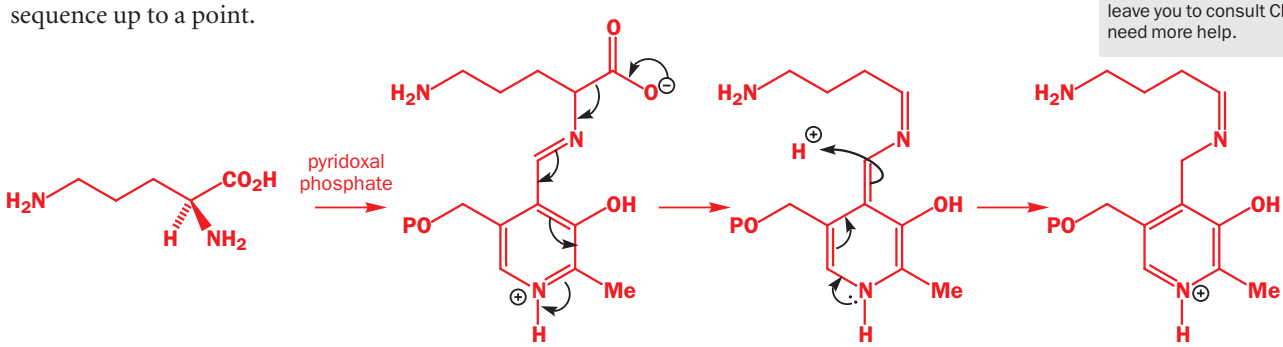
hygrine

If each amino group in ornithine is labelled in turn with ^{15}N, the α amino group is lost but the γ amino group is retained.

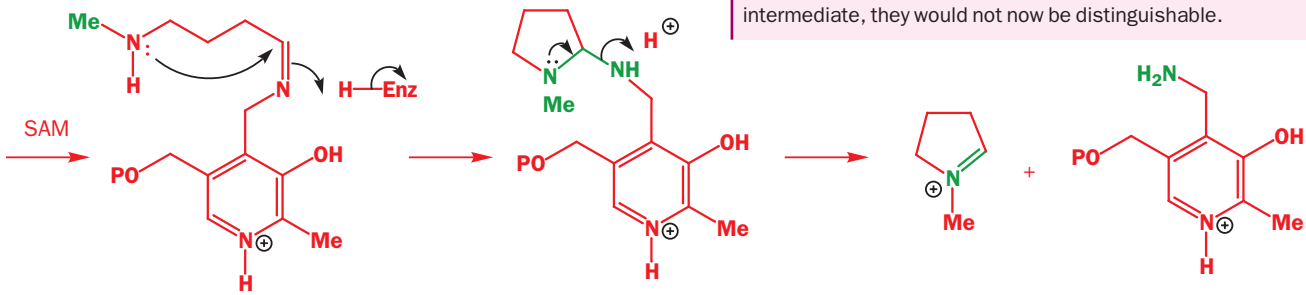


Further labelling experiments along these lines showed that the CO_2H group as well as the α amino group was lost from ornithine and that the rest of the molecule makes the pyrrolidine ring. The three-carbon side-chain in hygrine comes from acetate, or rather from acetyl CoA, and the *N*-methyl group comes from SAM. We can now work through the biosynthesis.

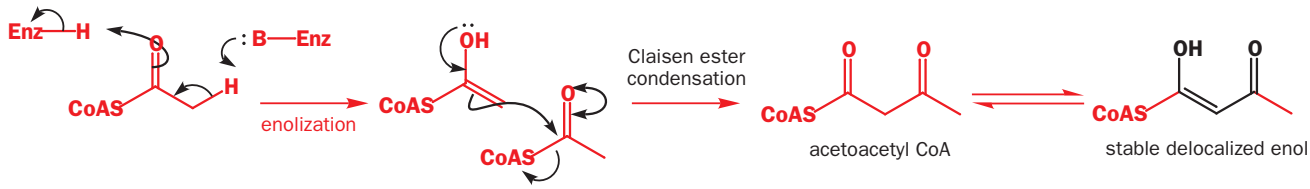
The first step is a pyridoxal-catalysed decarboxylation of ornithine, which follows the normal sequence up to a point.



Now the terminal amino group is methylated by SAM and the secondary amine cyclizes on to the pyridoxal imine to give an aminal. Decomposition of the aminal the other way round expels pyridoxamine and releases the salt of an electrophilic imine.

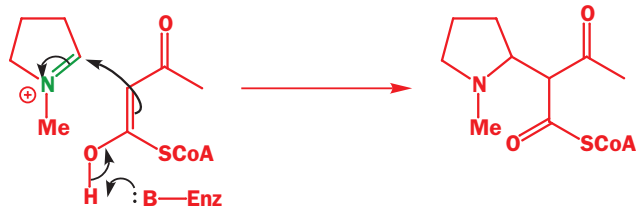


The rest of the biosynthesis does not need pyridoxal, but it does need two molecules of acetyl CoA. In Chapter 50 we noted that this thiol ester is a good electrophile and also enolizes easily. We need both reactivities now in a Claisen ester condensation of acetyl CoA.



The new keto-ester is very like the acetoacetates we used in Chapter 27 to make stable enolates and the CoA thiol ester will exist mainly as its enol, stabilized by conjugation.

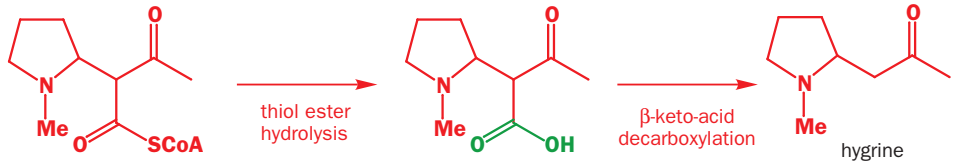
This enol reacts with the imine salt we have previously made and it will be easier to see this reaction if we redraw the enol in a different conformation. The imine salt does not have to wait around for acetoacetyl CoA to be made. The cell has a good stock of acetyl CoA and its condensation product.



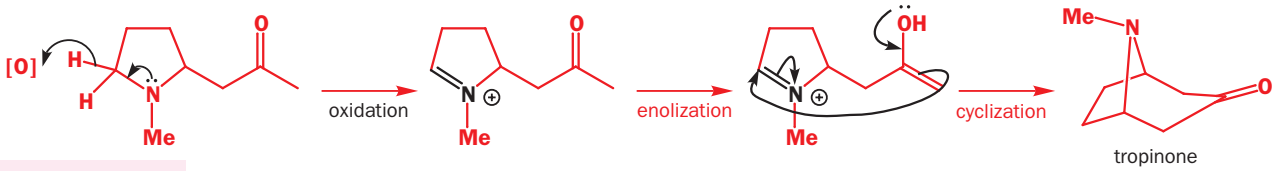
■ Both reagents SAM and acetyl CoA were discussed in Chapter 50. We will not be able to repeat at length the details of the chemistry of these and other common biochemical reagents already discussed there. In general, in this chapter we will give only the distinctive or interesting steps and leave you to consult Chapter 50 if you need more help.

▶ Notice that the methylation step means that the two carbon atoms that eventually become joined to nitrogen in the five-membered ring remain different throughout the sequence. If, say, putrescine had been an intermediate, they would not now be distinguishable.

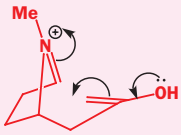
All that remains to form hygrine is the hydrolysis of the CoA thiol ester and decarboxylation of the keto-acid. This is standard chemistry, but you should ensure that you can draw the mechanisms for these steps.



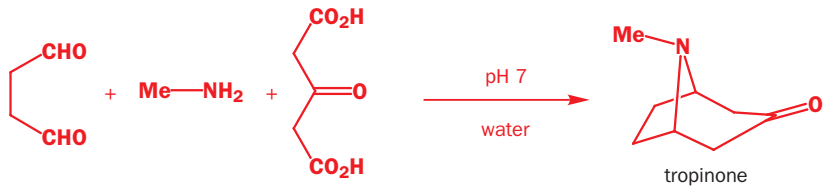
Tropinone is made from hygrine and it is clear what is needed. The methyl ketone must enolize and it must attack another imine salt resembling the first but on the other side of the ring. Such salts can be made chemically by oxidation with Hg(II) and biologically with an oxidizing enzyme and, say, NAD^+. The symbol [O] represents an undefined oxidizing agent, chemical or biological.



▶ The cyclization step looks dreadful when drawn on a flat molecule, but it looks much better in the conformation of tropinone shown below.

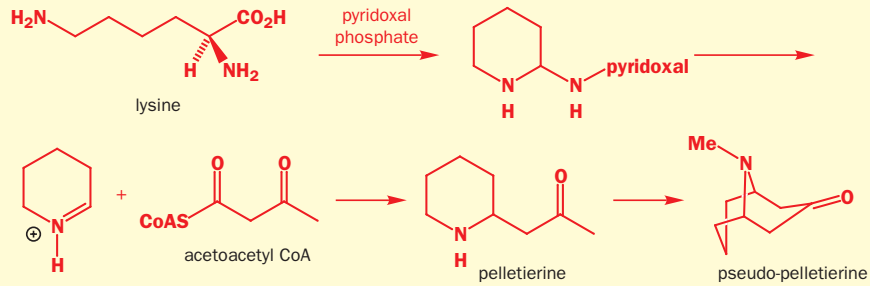


This complex route to tropinone was imitated as long ago as 1917 in one of the most celebrated reactions of all time, Robinson's tropinone synthesis. Robinson argued on purely chemical grounds that the sequence of imine salts and enols, which later (1970) turned out to be Nature's route, could be produced under 'natural' conditions (aqueous solution at pH 7) from a C_4 dialdehyde, $MeNH_2$ and acetone dicarboxylic acid. It worked and the intermediates must be very similar to those in the biosynthesis.



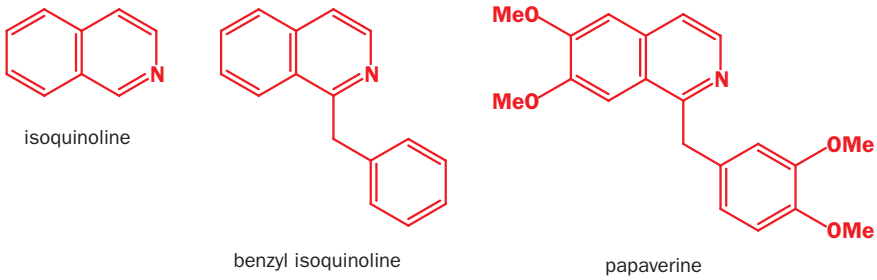
Other pyrrolidine alkaloids

There are many pyrrolidine alkaloids derived from ornithine and another large family of piperidine alkaloids derived from lysine by similar pathways involving decarboxylation and cyclization initiated by pyridoxal. We will not discuss these compounds in detail.

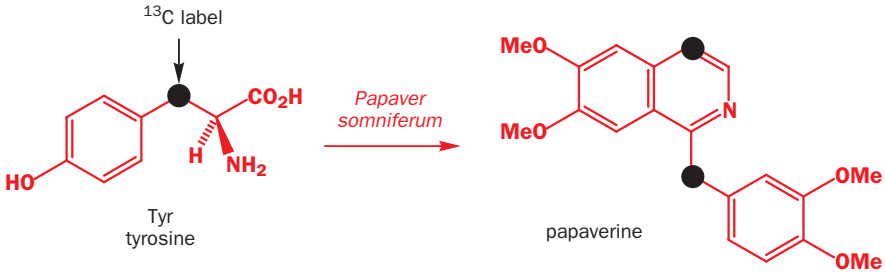


Benzyl isoquinoline alkaloids are made from tyrosine

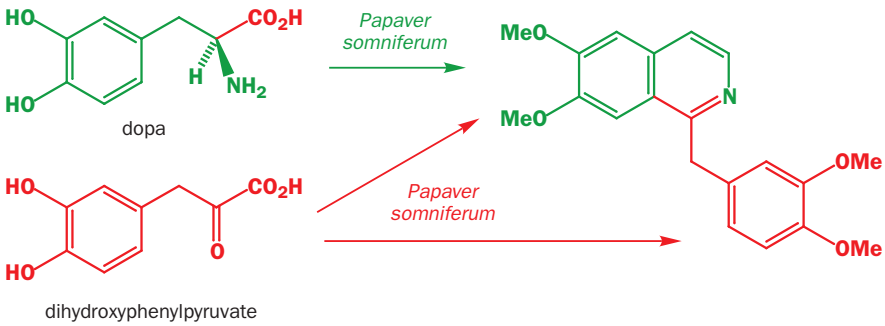
We switch to a completely different kind of alkaloid made from a different kind of amino acid. The **benzyl isoquinoline alkaloids** have a benzyl group attached to position 2 of an isoquinoline ring. Usually the alkaloids are oxygenated on the benzene ring and many are found in opium poppies (*Papaver somniferum*). For all these reasons papaverine is an ideal example.



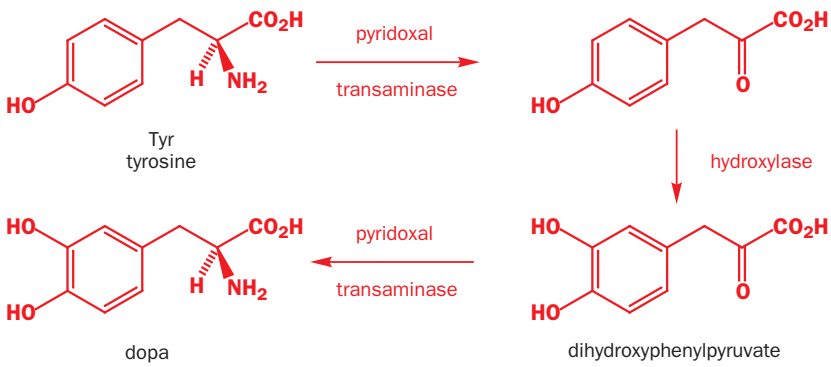
Labelling shows that these alkaloids come from two molecules of tyrosine. One must lose CO_2 and the other NH_3. We can easily see how to divide the molecule in half, but the details will have to wait a moment.



The question of when the extra OH groups are added was also solved by labelling and it was found that dihydroxyphenyl pyruvate was incorporated into both halves but the dihydroxyphenyl-alanine (an important metabolite usually called 'dopa') was incorporated only into the isoquinoline half.

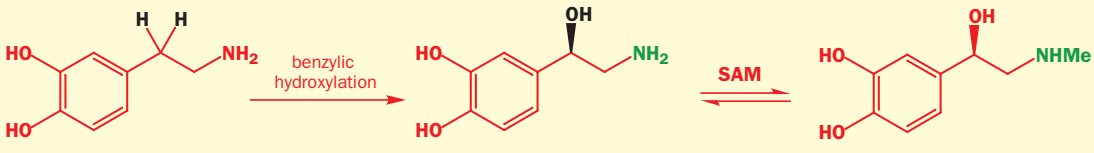


The amino acid and the keto-acid are, of course, related by a pyridoxal-mediated transaminase and the hydroxylation must occur right at the start. Both of these reactions are discussed in Chapter 50.



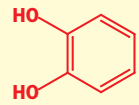
Catecholamines

Dopa and dopamine are important compounds because they are the precursors to adrenaline in humans. Decarboxylation of dopa gives dopamine, which an oxidase (Chapter 50) hydroxylates stereospecifically at the benzylic position to give noradrenaline (norepinephrine).



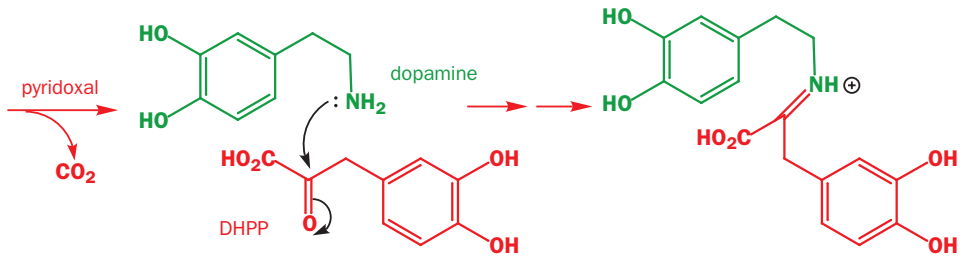
dopamine — noradrenaline (norepinephrine) — adrenaline (epinephrine)

The family of hormones that includes adrenaline and noradrenaline is often called the **catecholamines** (catechol is 1,2-dihydroxybenzene). The hormones are produced in the adrenal gland around the kidneys and regulate several important aspects of metabolism: they help to control the breakdown of stored sugars to release glucose and they have a direct effect on blood pressure, heart rate, and breathing. The relative proportion of noradrenaline and its *N*-methylated analogue, adrenaline, controls these things.



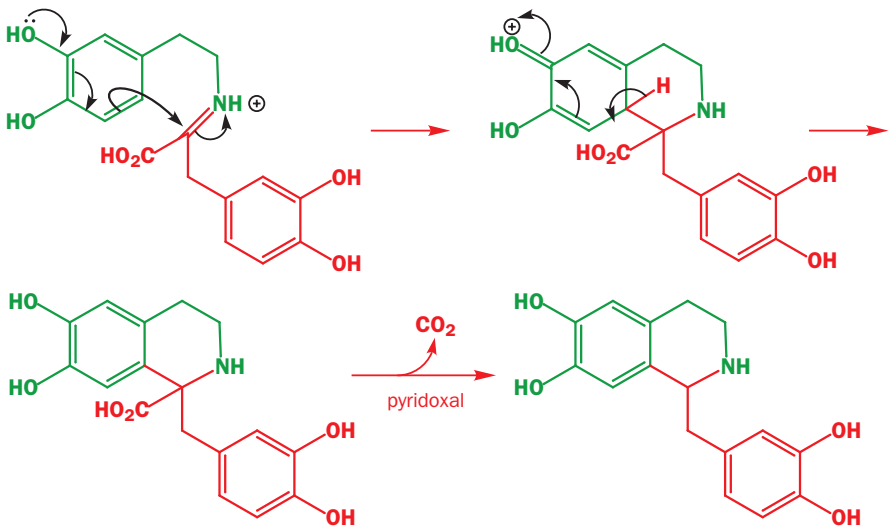
catechol

Pyridoxal-mediated decarboxylation of dopa gives dopamine and this reacts with the keto-acid to form an imine salt. This is an open-chain imine salt unlike the cyclic ones we saw in the pyrrolidine alkaloids, but it will prove to have similar reactivity.

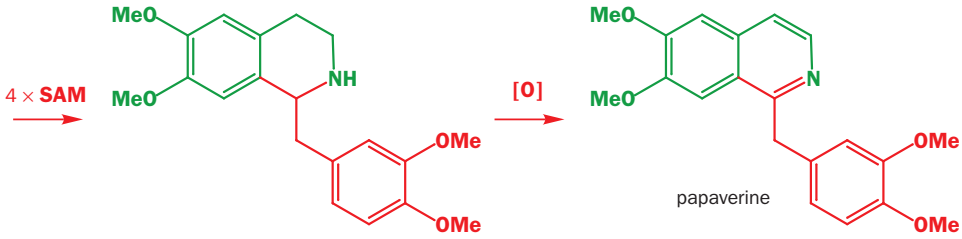


The imine salt is perfectly placed for an intramolecular electrophilic aromatic substitution by the electron-rich dihydroxyphenyl ring. This closes the isoquinoline ring in a Mannich-like process (Chapter 27) with the phenol replacing the enol in the pyrrolidine alkaloid biosynthesis.

■ Even in biological electrophilic aromatic substitutions, it is still important to remember to write in the hydrogen atom at the place of substitution (Chapter 22)!

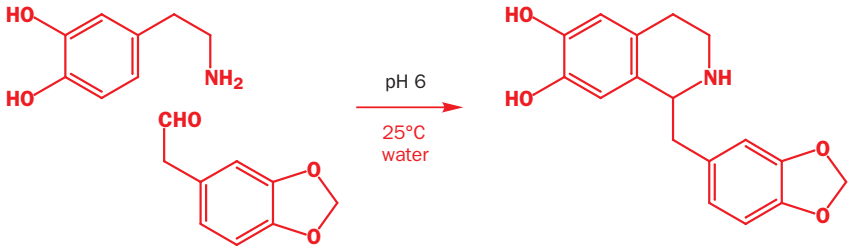


The cyclization product is still an amino acid and it can be decarboxylated by pyridoxal. Now we have something quite like papaverine but it lacks the methyl groups and the aromatic heterocyclic ring. Methylation needs SAM and is done in two stages for a reason we will discover soon. The final oxidation should again remind you of the closing stages of the tropinone route.



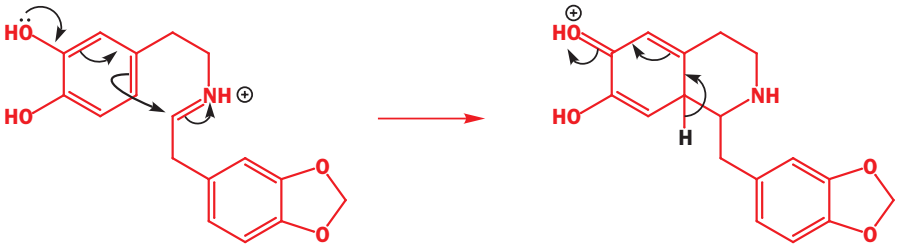
The reaction to make the isoquinoline ring can be carried out chemically under very mild conditions providing that we use an aldehyde as the carbonyl component. Then it works very well with rather similar compounds.

The reaction also works with an aryl pyruvic acid, but the decarboxylation is more difficult to organize without pyridoxal.



The mechanism is straightforward—the imine is formed and will be protonated at pH 6, ready for the C–C bond formation, which is both a Mannich reaction and an electrophilic aromatic substitution.

Notice that it was not necessary to protect the OH groups—the acetal on the lower ring is not for protection, and this group (methylenedioxy or dioxolan) is present in many benzyl isoquinoline alkaloids. It is formed in nature by oxidation of an MeO group *ortho* to an OH group on a benzene ring.

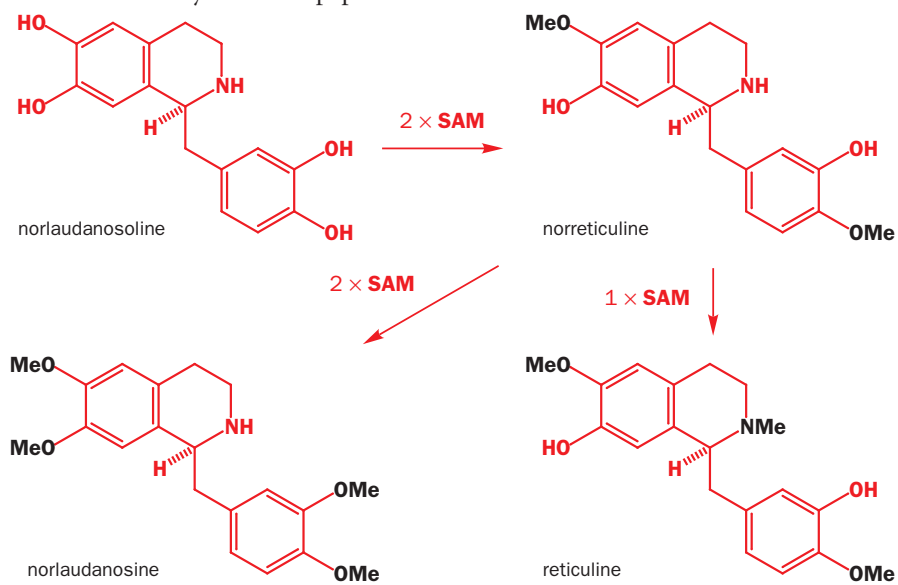


Complex benzyl isoquinoline alkaloids are formed by radical coupling

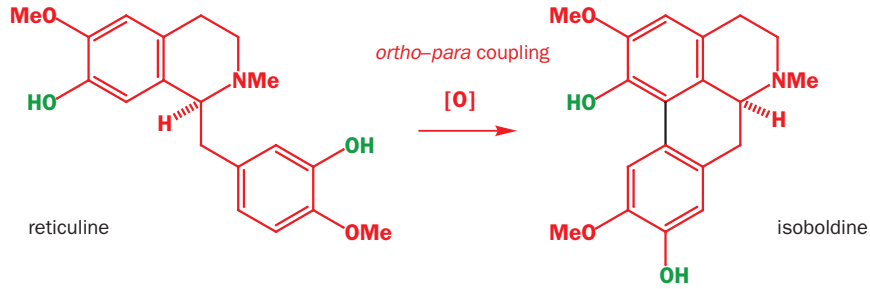
A more interesting series of alkaloids arises when benzyl isoquinoline alkaloids cyclize by radical reactions. Phenols easily form radicals when treated with oxidizing agents such as Fe(III), and benzyl isoquinoline alkaloids with free phenolic hydroxyl groups undergo radical reactions in an intramolecular fashion through a similar mechanism. Here are the details of some methylations of a class of alkaloids closely related to papaverine.

See Chapter 39.

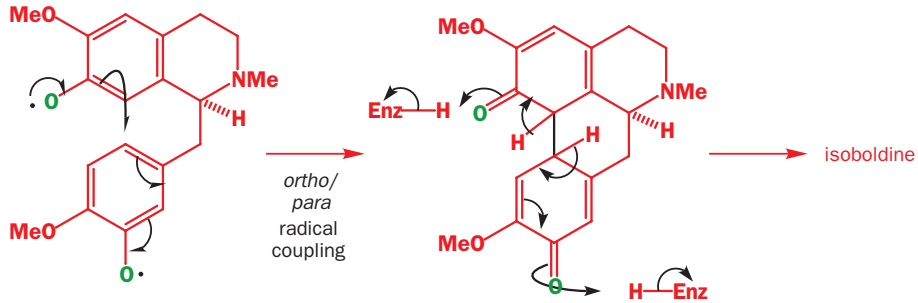
The names of the alkaloids should not, of course, be learned, but they are a convenient handle for quick reference. The prefix 'nor' means without a methyl group, in this case the *N*-Me group, as you can see with norreticuline and reticuline.



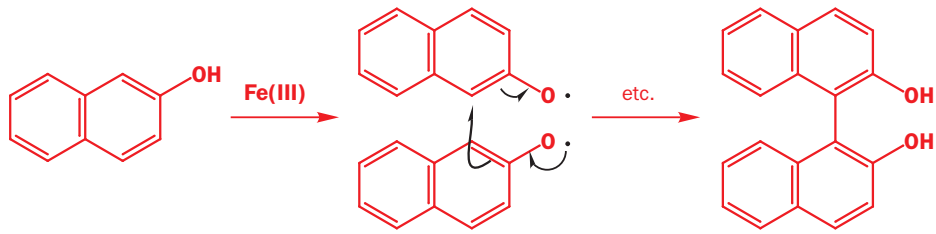
Methylating only one phenol on each ring of norreticuline leaves the other one free for radical coupling. Reticuline is oxidized in the plant to isoboldine by a radical cyclization with the formation of a new C–C bond.



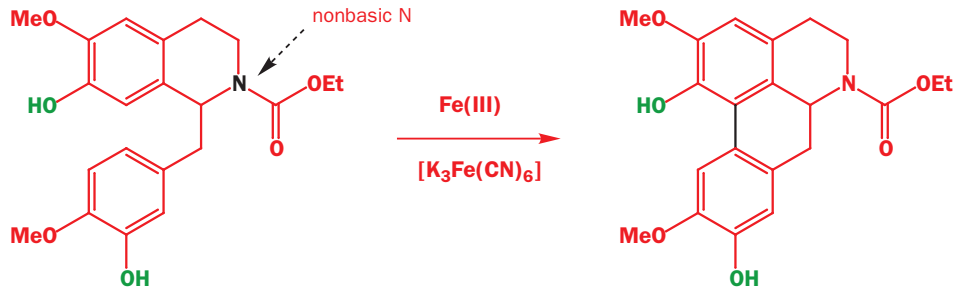
The new C–C bond is marked in black and the free phenolic OHs in green. Notice the relationship between them. The new bond is between a carbon atom *ortho* to one OH group and a carbon atom *para* to the other. We shall see in all these phenolic couplings that the *ortho* and *para* positions are the only activated ones (*ortho*/*ortho*, *ortho*/*para*, and *para*/*para* couplings are all possible). Oxidation occurs at the phenolic hydroxyl groups, and the resulting oxygen radicals couple.



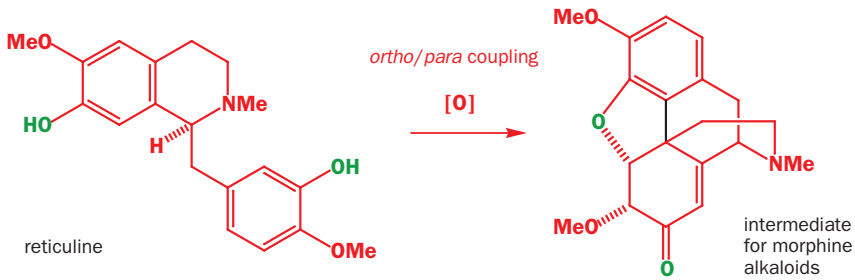
Phenol coupling occurs chemically under oxidation with Fe(III). The most famous example is the coupling of 2-naphthol to give binaphthol—an *ortho*/*ortho* coupling. The stereochemistry of binaphthyls like this was discussed in Chapter 45.



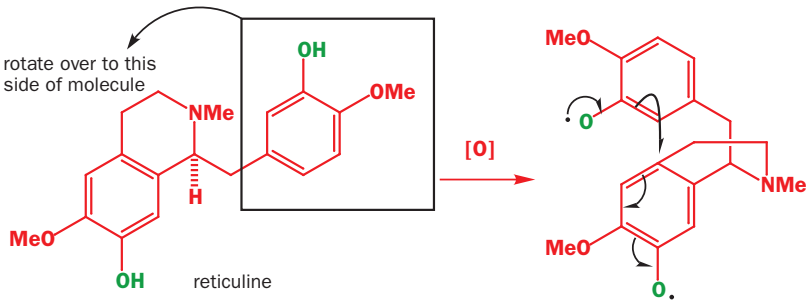
Similar phenol couplings have been attempted in the laboratory with compounds in the benzyl isoquinoline series but the nitrogen atom interferes if it is at all basic. When it has a carbonyl substituent the reactions do work reasonably well, but the yields are poor. Nature is still much better at this reaction than we are.

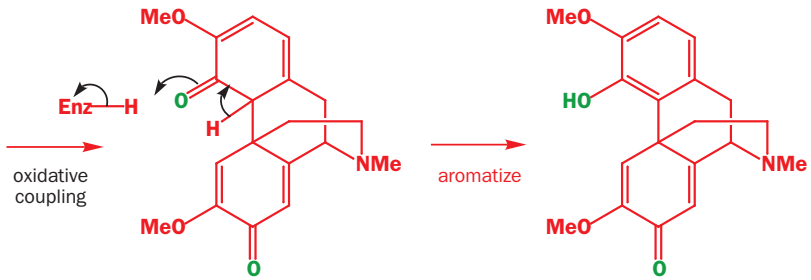


Reticuline is also the source of the morphine alkaloids by *ortho*/*para* radical coupling. The roles of the two rings are reversed this time and it is quite difficult to see at first how the structures are related.

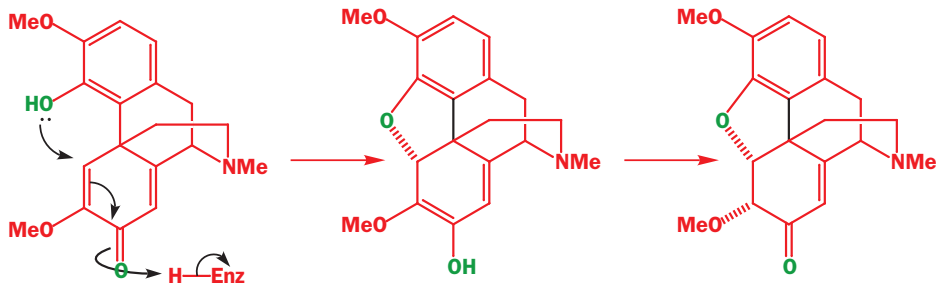


A great deal has happened in this reaction, but the new C–C bond (black) is *ortho* to the green oxygen atom in the top ring and *para* to the green oxygen atom in the bottom ring, so *ortho*/*para* coupling has occurred. To draw the reaction mechanism we need to draw reticuline in the right conformation.

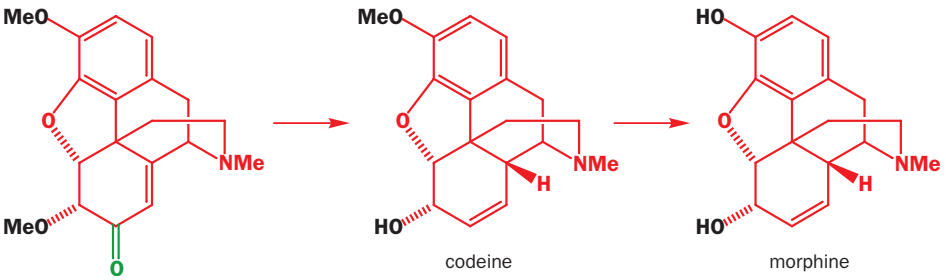




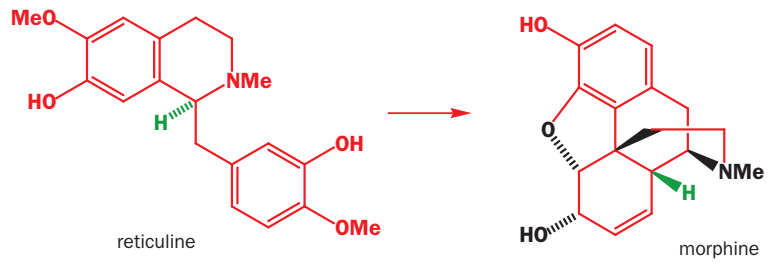
One of the two rings can re-aromatize but the other has a quaternary carbon atom so no proton can be lost from this site. Instead, the OH group in the top ring adds in conjugate fashion to the enone in the bottom ring.



This intermediate gives rise to the important alkaloids codeine and morphine, which differ only by a methyl group. Nature can remove methyl groups as well as add them.



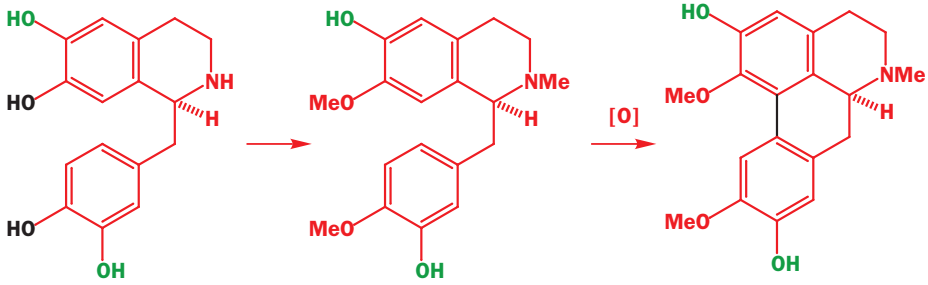
These alkaloids have plenty of stereochemistry. Indeed, if we compare the structures of reticuline and morphine, we can see that the one stereogenic centre in reticuline (marked in green) is still there in morphine (it hasn't been inverted—that part of the molecule has just been turned over) and that four new stereogenic centres marked in black have been added. These centres all result from the original twisting of reticuline to allow phenol coupling except for the one bearing an OH group, which comes from a stereoselective reduction.



reticuline morphine

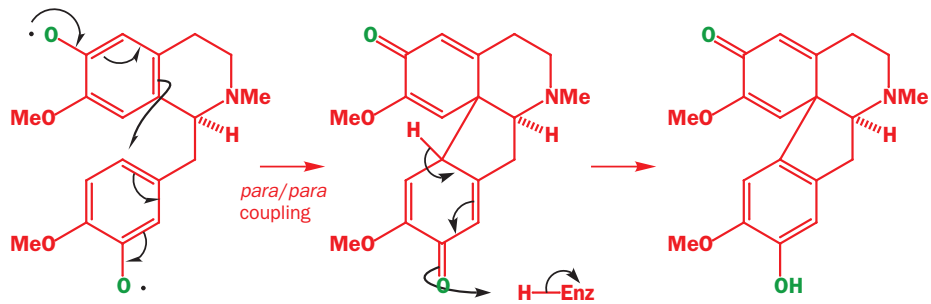
Boldine, an isomer of isoboldine, is formed by rearrangement

We mentioned isoboldine a while back, so there must be a boldine as well. This alkaloid is also formed from norlaudanosoline by a different methylation sequence and oxidative radical coupling. Looking at the structure of boldine you may see what appears to be a mistake on someone's part.

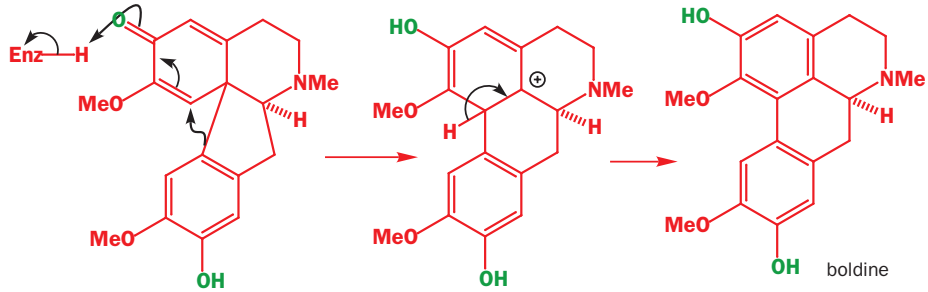


norlaudanosoline boldine

The coupling is correctly *para* in the bottom ring but is *meta* in the top ring. But there is no mistake (neither by the authors nor by Nature!)—this structure is correct and it has been made by *para*/*para* coupling.



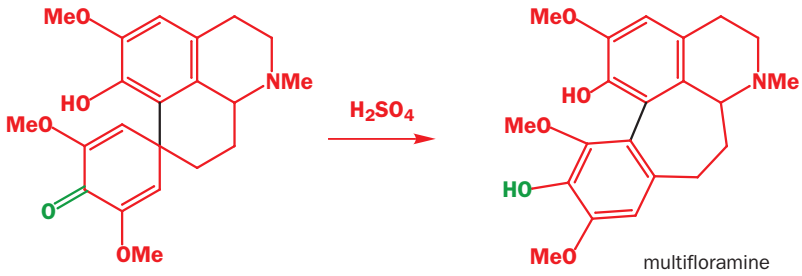
One of the rings has aromatized, but the other cannot—this should remind you of the morphine biosynthesis. However, there is no nucleophilic OH group here capable of conjugate addition to the enone so a rearrangement occurs instead. The new bond to the lower ring migrates across the top ring. You might even say that the lower ring does an intramolecular conjugate addition on the upper ring.



boldine

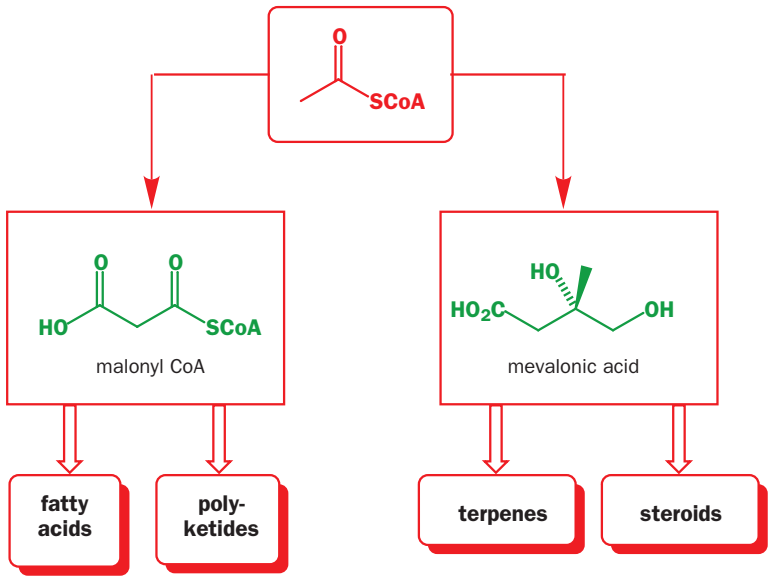
After the rearrangement there is a proton available to be lost and the cation can aromatize. The *para* relationship in the original coupling product has become a *meta* relationship by rearrangement. You should be able to recognize this rearrangement from Chapter 37: it is a dienone–phenol rearrangement.

In rearrangements like these with cationic intermediates, the group that can best support a positive charge usually prefers to migrate. The reasons for this are discussed in Chapter 37. Here is a purely chemical example of the same reaction, giving 82% yield in acidic solution. The bond that migrates is marked in black.



Fatty acids and other polyketides are made from acetyl CoA

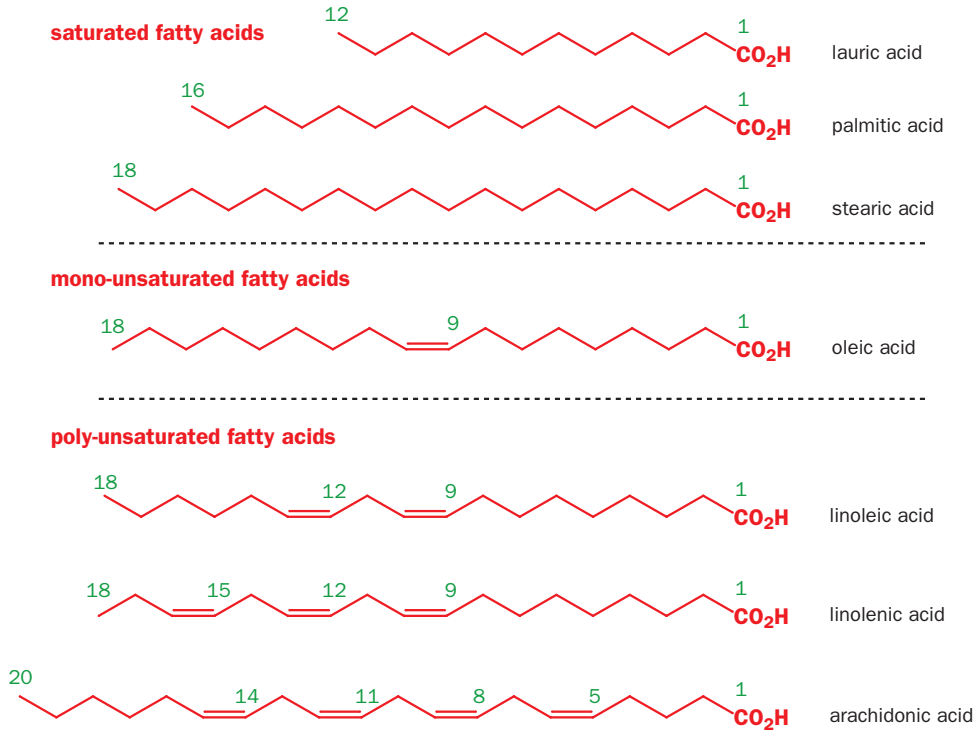
The sections that remain in this chapter show how Nature can take a very simple molecule—acetyl CoA—and build it up into an amazing variety of structures. There are two main pathways from acetyl CoA and each gives rise to two important series of natural products.



We shall discuss these four types of compounds in the order shown so that we start with the simplest, the fatty acids. You met these compounds in Chapter 49 as their glyceryl esters, but you now need to learn about the acids in more detail and outline their biosynthesis. Compare the structures of the typical fatty acids in the chart overleaf.

These are just a few of the fatty acids that exist, but all are present in our diet and you'll find many referred to on the labels of processed foods. You should notice a number of features.

- They have straight chains with no branching
- They have even numbers of carbon atoms
- They may be saturated with no double bonds in the chain, or
- They may have one or more C=C double bonds in the chain, in which case they are usually *cis* (*Z*) alkenes. If there is more than one C=C double bond, they are not conjugated (either with the CO_2H group or with each other)—there is normally one saturated carbon atom between them.

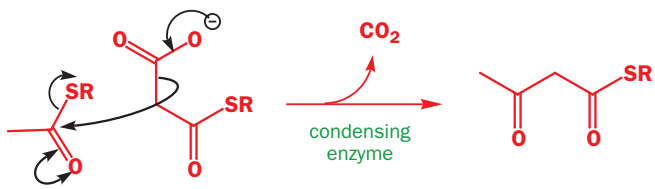


Palmitic acid (C_{16} saturated) is the most common fatty acid in living things. Oleic acid (C_{18} mono-unsaturated) is the major fatty acid in olive oil. Arachidonic acid (C_{20} tetra-unsaturated) is a rare fatty acid, which is the precursor of the very important prostaglandins, thromboxanes, and leukotrienes, of which more later.

The prevalence of fatty acids with even numbers of carbon atoms suggests a two-carbon building block, the most obvious being acetate. If labelled acetate is fed to plants, the fatty acids emerge with labels on alternate carbons like this.

The green blob might represent deuterium (as a CD_3 group) and the black blob ^{13}C. In fact, the reactions are more complex than this suggests as CO_2 is also needed as well as CoA and it turns out that only the first two-carbon unit is put in as acetyl CoA. The remainder are added as malonyl CoA. If labelled malonyl CoA is fed, the starter unit, as it is called, is not labelled.

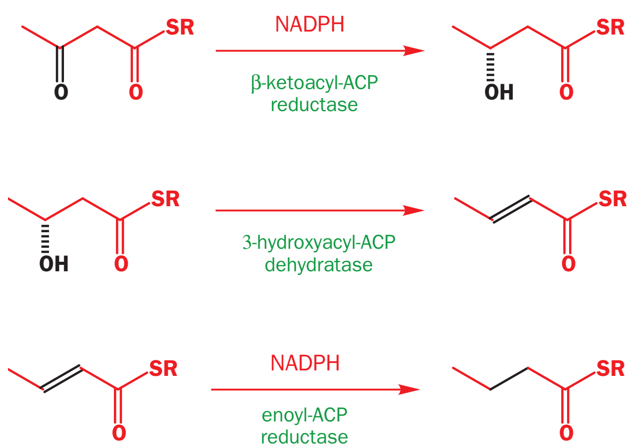
Malonyl CoA is made from acetyl CoA and CO_2 carried, as usual, on a molecule of biotin (Chapter 50). The first stage in the fatty acid biosynthesis proper is a condensation between acetyl CoA (the starter unit) and malonyl CoA with the loss of CO_2. This reaction could be drawn like this.



Notice that CO_2 is lost as the new C–C bond is formed. When chemists use malonates, we like to make the stable enol using both carbonyl groups, condense, and only afterwards release CO_2 (Chapter 26). Nature does this in making acetoacetyl CoA during alkaloid biosynthesis, but here she works differently.

The next step is reduction of the ketone group.

This NADPH reaction is typically stereo- and chemoselective, though the stereochemistry is rather wasted here as the next step is a dehydration, typical of what is now an aldol product, and occurring by an enzyme-catalysed E1cB mechanism.

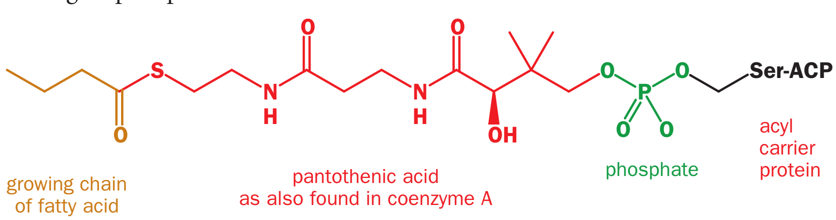


The elimination is known to be a *cis* removal of H and OH and the double bond is exclusively *trans* (*E*). Only later in the nonconjugated unsaturated fatty acids do we get *Z*-alkenes. Finally, in this cycle, the double bond is reduced using another molecule of NADPH to give the saturated side-chain.

Now the whole cycle can start again using this newly made C_4 fatty acid as the starter unit and building a C_6 fatty acid and so on. Each time the cycle turns, two carbon atoms are added to the acyl end of the growing chain.

Fatty acid synthesis uses a multienzyme complex

We have not told you the whole truth so far. Did you notice that 'SCoA' in the structures had been replaced by 'SR' and that a mysterious 'ACP' had crept into the enzyme names? That was because these reactions actually happen while the growing molecule is attached as a thiol ester to a long side-chain on an **acyl carrier protein** (**ACP**). The long side-chain is closely related to CoA and is attached through a phosphate to a serine residue of the ACP.



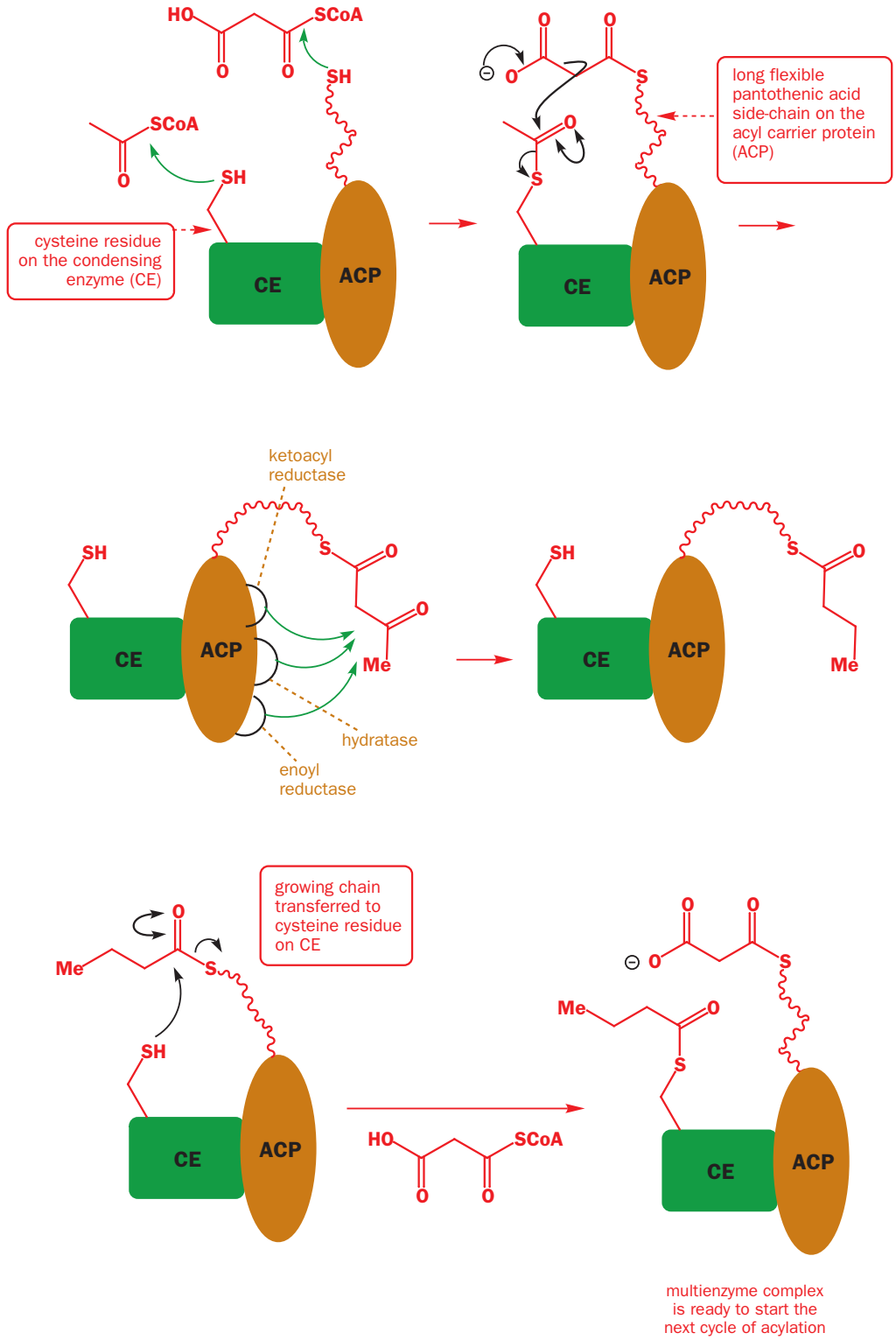
All of the enzymes needed for one cycle are clumped together to form two large proteins (ACP, the acyl carrier protein, and **CE**, the **condensing enzyme**) which associate in a stable dimer. The long side-chain passes the substrate from enzyme to enzyme so that synthesis can be continuous until the chain is finished and only then is the thiol ester hydrolysed. The chart on p. 1428 illustrates this.

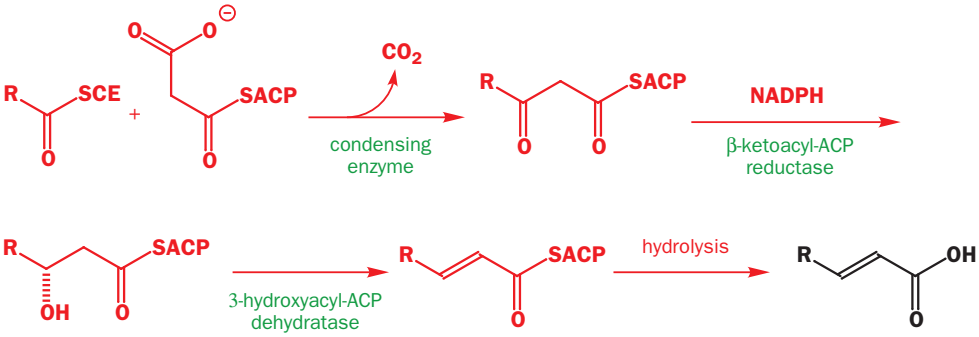
You saw a smaller multienzyme complex in Chapter 50 (p. 1395), but this one is much more complex. More are being discovered all the time—Nature invented the production line well before Henry Ford.

There are three ways of making unsaturated fatty acids

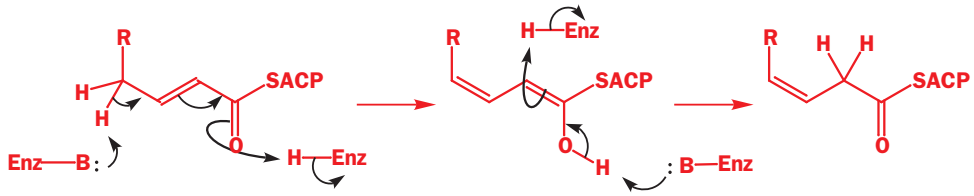
Conjugated unsaturated fatty acids are made simply by stopping the acylation cycle at that stage and hydrolysing the thiol ester linkage between the unsaturated acyl chain and ACP. They always have the *E* (*trans*) configuration and are the starting points for other biosynthetic pathways.

fatty acid biosynthesis: schematic diagram of the multienzyme dimer





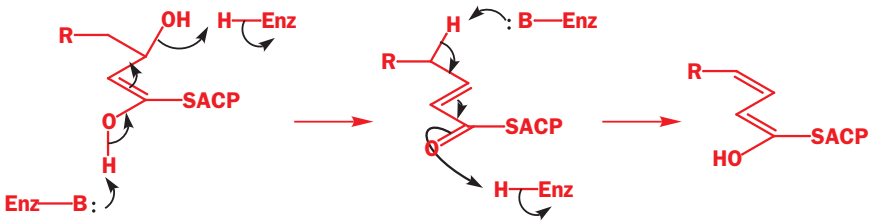
The second method makes *Z*-3,4-unsaturated acids by deconjugation from the *E*-2,3-unsaturated acids catalysed by an isomerase while the acyl chain is still attached to ACP. This is an anaerobic route as no oxidation is required (the double bond is already there—it just has to be moved) and is used by prokaryotes such as bacteria.



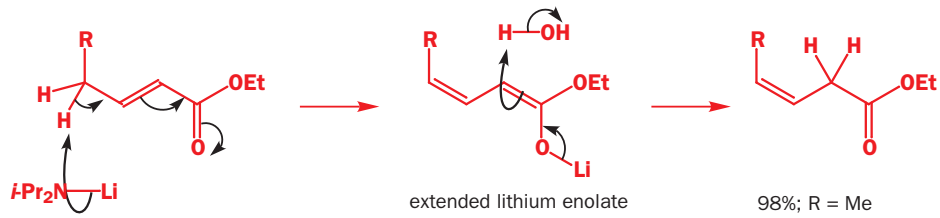
extended enol intermediate

Removal of a proton from C4 forms an extended enol, which can be protonated at C2 or C4. Protonation at C4 is thermodynamically favoured as it leads to the conjugated alkene. But protonation at C2 is kinetically favoured, and this leads to the nonconjugated alkene. The geometry of the new alkene depends on the conformation of the chain when the first (deprotonation) step occurs. It is thought that this is the best conformation for the previous reaction, the dehydration step, and that no rotation of the chain occurs before the isomerase gets to work.

For more on this, read a specialized book, such as Ian Fleming's, *Frontier orbitals and organic chemical reactions*, Wiley, Chichester, 1976. Similar regioselectivity is evident in the protonation of the Birch reduction products on p. 628.



You may think this a rather unlikely reaction, but the same thing can be done in the laboratory. If a simple unsaturated ester is converted into its lithium enolate and then reprotonated with water, the major product is the ester of the *Z*-3,4-enoic acid. Yields and steroselectivities are excellent.

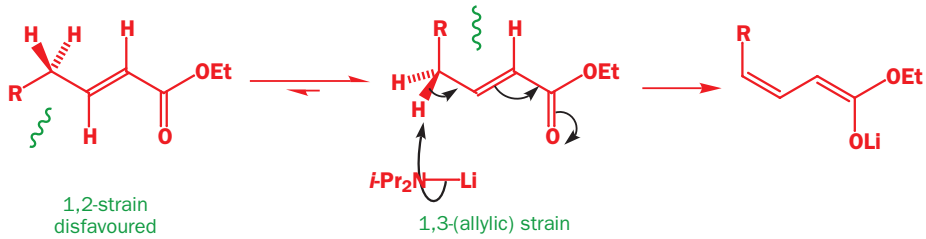


extended lithium enolate

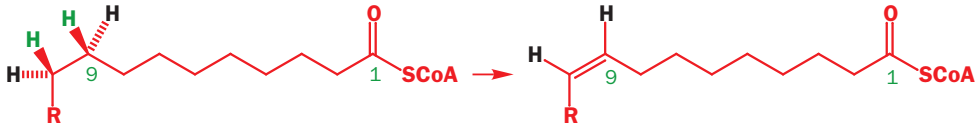
98%; R = Me

One explanation suggests that control is exercised by a favourable conformation in which 1,3-allylic strain is preferred to 1,2-strain. It looks as though Nature has again seized on a natural chemical preference and made it even better.

$A^{1,3}$ strain (1,3-allylic strain) was discussed in Chapter 34, p. 896.



The third method is a concerted stereospecific removal of two adjacent hydrogen atoms from the chain of a fatty acid after synthesis. This is an aerobic route as oxidation is required and is used by mammals such as ourselves. The stereochemistry of the reaction is known from labelling studies to be *cis* elimination.

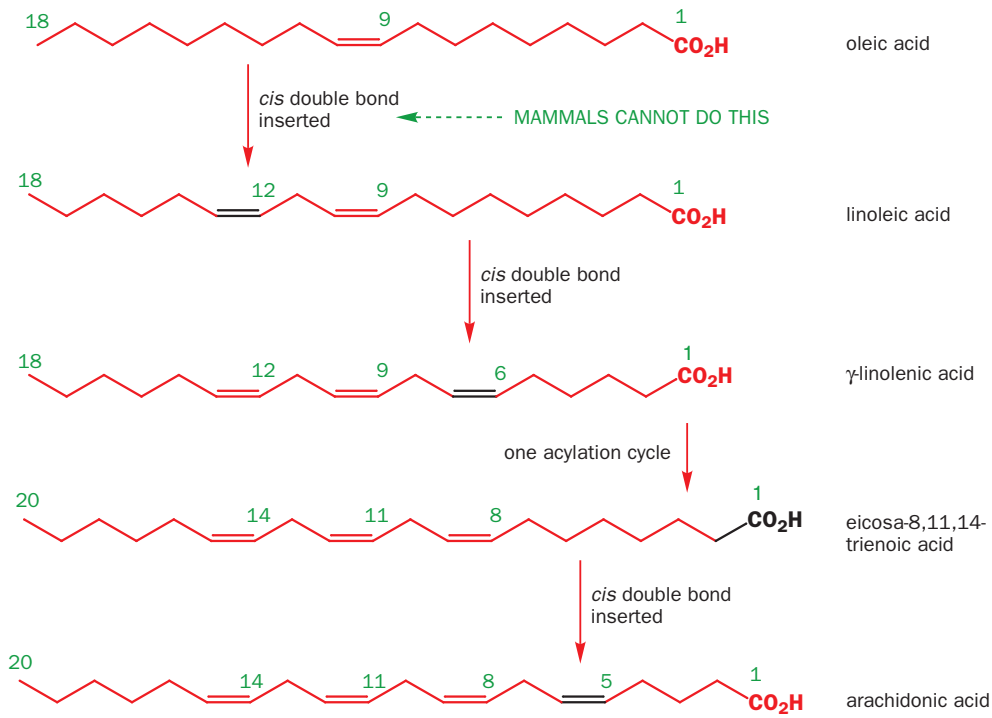


This oxidation involves a chain of reagents including molecular oxygen, Fe(III), FAD, and NAD^+. A hydroxylation followed by a dehydration or a sulfur-promoted dehydrogenation has been suggested for the removal of the hydrogen atoms. The chemical reaction corresponding to the biological reaction has not yet been discovered.

What is so important about unsaturated fatty acids?

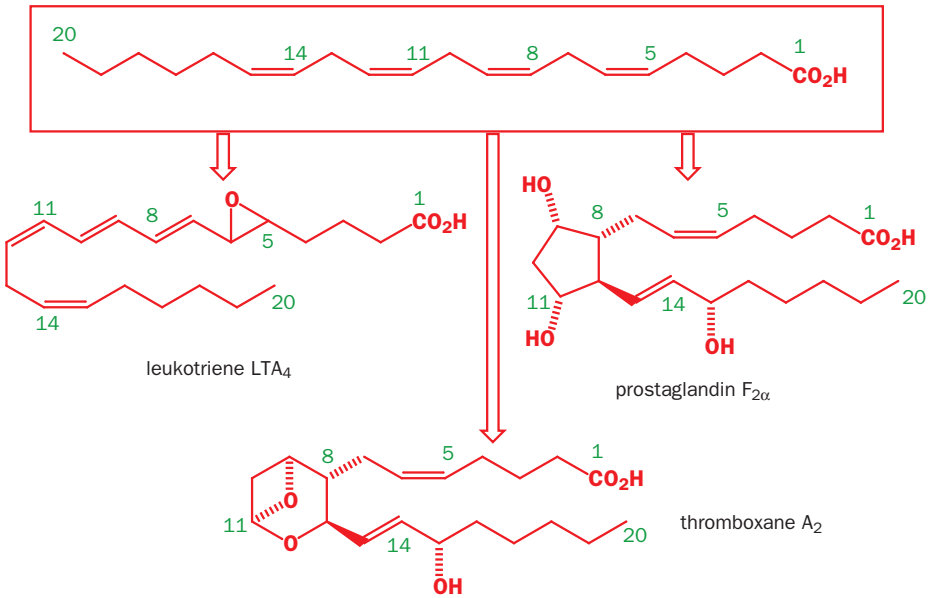
Mammals can insert a *cis*-alkene into the chain, providing that it is no further away from the carbonyl group than C9. We cannot synthesize linoleic or linolenic acids (see chart a few pages back) directly as they have alkenes at C12 and C15. These acids must be present in our diet. And why are we so keen to have them? They are needed for the synthesis of arachidonic acid, a C_{20} tetraenoic acid that is the precursor for some very interesting and important compounds. Here is the biosynthesis of arachidonic acid.

synthesis of unsaturated fatty acids

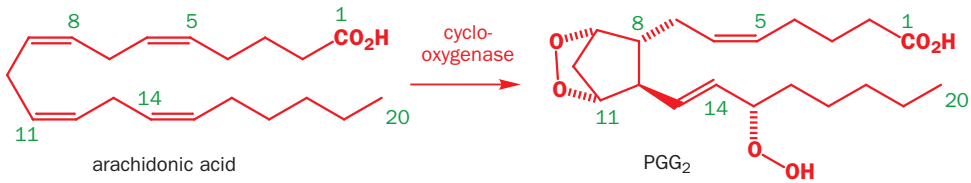


The final product of this chain of events—arachidonic acid—is one of the **eicosanoids**, so called because *eicosa* is Greek for 'twenty', and the systematic names for these compounds contain 'eicosanoic acid' in some form. The leukotrienes resemble arachidonic acid most closely, the prostaglandins have a closed chain forming a five-membered ring, and the thromboxanes resemble the prostaglandins but have a broken chain. All are C_{20} compounds with the sites of the alkenes (C5, C8, C11, and C14) marked by functionality or some other structural feature.

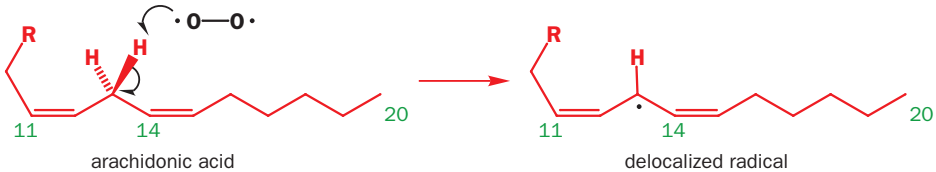
compounds synthesized from arachidonic acid



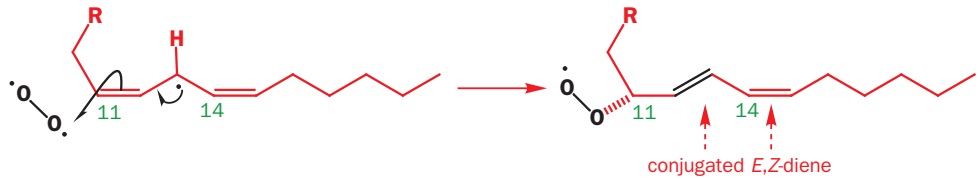
These compounds are all unstable and all are involved in transient events such as inflammation, blood clotting, fertilization, and immune responses. They are produced locally and decay quickly and are implicated in autoimmune diseases like asthma and arthritis. They are made by oxidation of arachidonic acid—you can see this best if you redraw the molecule in a different conformation.



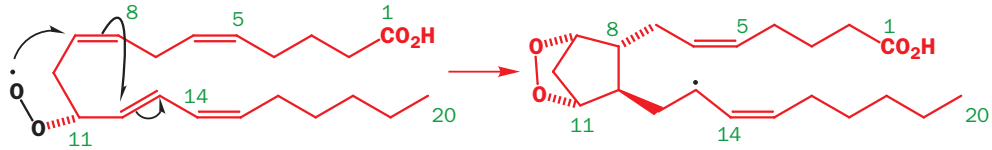
The first step is a radical abstraction of a hydrogen atom from an allylic position by oxygen (perhaps carried on an iron atom in a haem). The atom removed is between two alkenes so that the resulting radical is doubly allylic.



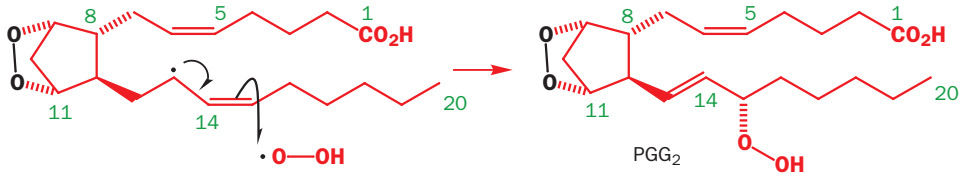
This allylic radical captures a molecule of oxygen at C11 to form a new oxyradical. The reaction occurs at one end of the delocalized radical so that the product is a conjugated diene and the new alkene is *trans* (*E*).



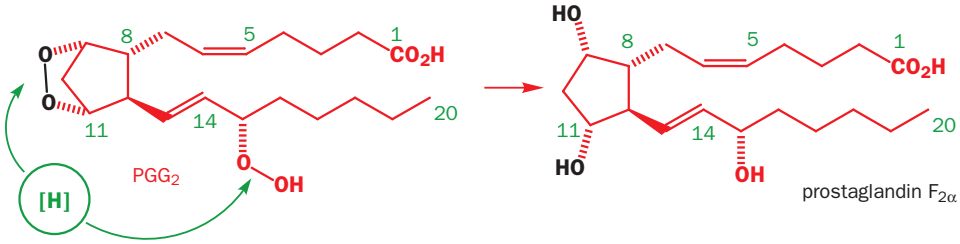
Now we need to resume the full structure of the intermediate because the oxyradical does an elaborate addition to the C8 alkene and then to the newly formed diene to form a new stable allylic radical.



Three new stereogenic centres are created in this cyclization, at C8, C9, and C12, and all are under full control both from the centre already present and from the way in which the molecule folds up under the guidance of the enzyme. Now the allylic radical reacts with oxygen to give the unstable hydroperoxide PGG_2.



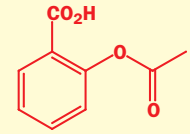
This unstable prostaglandin has been isolated from sheep but, as it has a half-life of only 5 minutes, this is no trivial matter. Both weak O–O bonds are now reduced enzymatically to give the first reasonably stable compound, $PGF_{2\alpha}$ (PG just means prostaglandin).



The best evidence for this pathway comes from labelled oxygen molecules. If a mixture of ^{16}O–^{16}O (ordinary oxygen) and ^{18}O–^{18}O is supplied to an organism making $PGF_{2\alpha}$, the product has either both black OHs as ^{16}O or both as ^{18}O but no molecules are formed with one ^{16}O and one ^{18}O. These isotopes are easily measured by mass spectrometry. Both black OHs then come from one and the same molecule of oxygen—not an obvious conclusion when you inspect the molecule of $PGF_{2\alpha}$, and thus good evidence for this pathway.

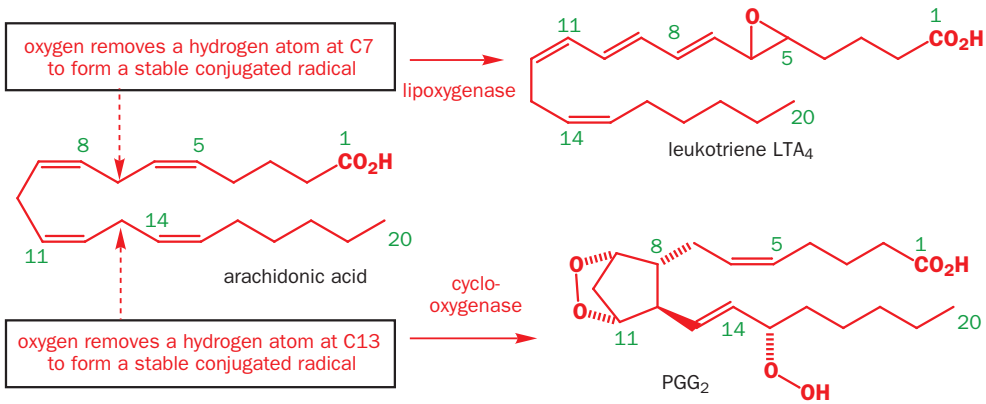
How aspirin works

The enzyme that catalyses these remarkable reactions, cyclooxygenase, is an important target for medicinal chemists. Inhibiting PG synthesis can bring about a reduction of inflammation and pain. In fact, this is how aspirin works. It was not, of course, *designed* to work that way and its mode of action was discovered decades after its use began. There is a price to pay for such a useful drug. PGs also control acid secretion in the stomach and aspirin inhibits their synthesis there too so stomach ulceration can result.

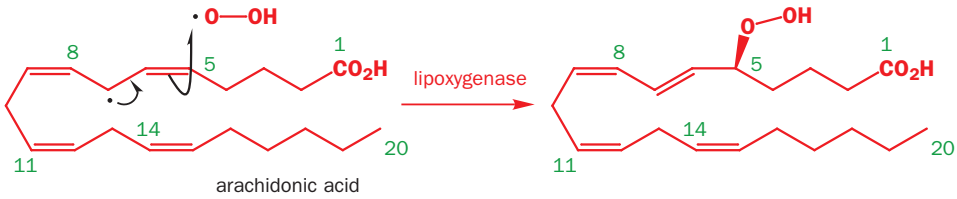


aspirin

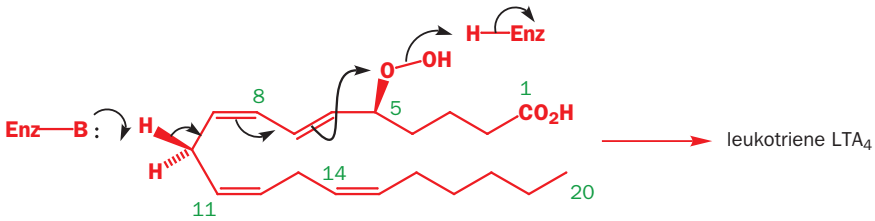
Each of the other families of eicosanoids—thromboxanes and leukotrienes—has interesting biosynthetic pathways too, but we will mention only one small detail. A completely different oxidation enzyme, lipoxygenase, initiates a separate pathway leading to the leukotrienes, but the first steps are very similar. They just occur elsewhere in the arachidonic acid molecule.



The initially formed radical is stabilized by two double bonds in the same way as that we have just seen and reacts with oxygen in the same way again to give a *trans*-alkene and a new hydroperoxide.



The next step is something quite new. No new C–C bond is formed: instead, the diene attacks the hydroperoxide to give an epoxide and a fully conjugated triene. The new double bond is *cis* this time, which is what we should expect from the conformation we have been using. This is LTA_4 and all the other leukotrienes are made from this compound.

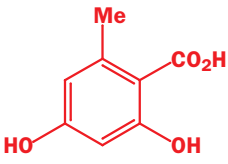


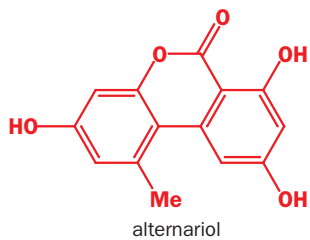
The relatively recent discovery of these unstable molecules of incredibly powerful biological activity means that we by no means know all about them yet. They are very important to our well-being and important medical advances are bound to follow from a better understanding.

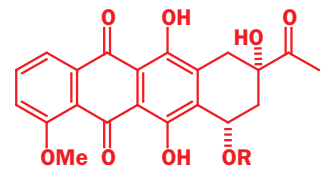
■ Prostaglandins and leukotrienes have appeared several times before in this book, and you can read about aspects of their laboratory synthesis on pp. 686, 1229, 1268, and 883.

Aromatic polyketides come in great variety

The **fatty acid pathway** or, as we should call it now, the **acyl polymalonate pathway**, also gives rise to an inexhaustible variety of aromatic and other compounds belonging to the family of the polyketides. You saw in Chapter 50 how the shikimic acid pathway makes aromatic compounds but the compounds below are from the polyketide route.

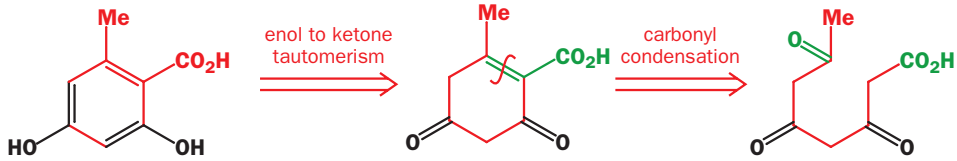

orsellinic acid


alternariol


daunomycin (R = sugar)

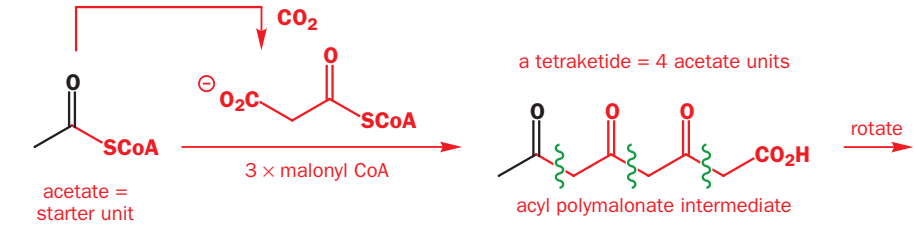
You might immediately be struck by the extent of oxygenation in these compounds. The shikimic acid route produced Ar–C_3 compounds with at most one OH group in the *para* position and others

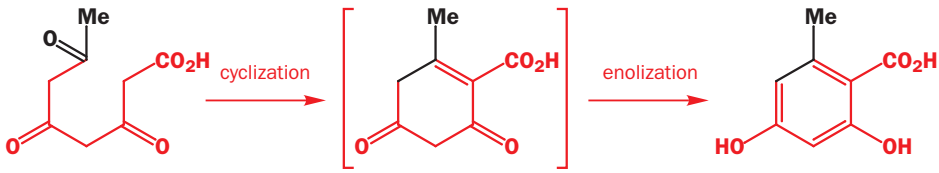
added *ortho* to that first OH group. Here we have multiple oxygenation with a predominant 1,3 pattern. If we try to arrange an acyl polymalonate product to make orsellinic acid, this is what we shall need.



orsellinic acid

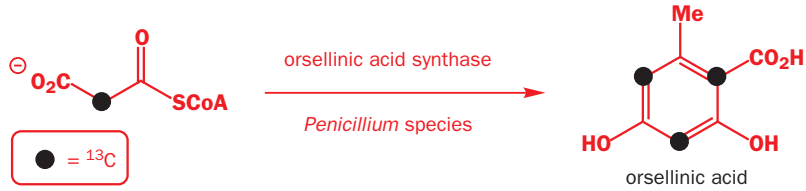
Merely by writing ketones instead of phenols and doing one disconnection corresponding to a simple carbonyl condensation, we have reached a possible starting material which is a typical acyl polymalonate product without any reductions. This is what polyketides are. The fatty acids are assembled with full reduction at each stage. Polyketides are assembled from the same process but without full reduction; indeed, as the name polyketide suggests, many are made without any reduction at all. This is the biosynthesis of orsellinic acid.





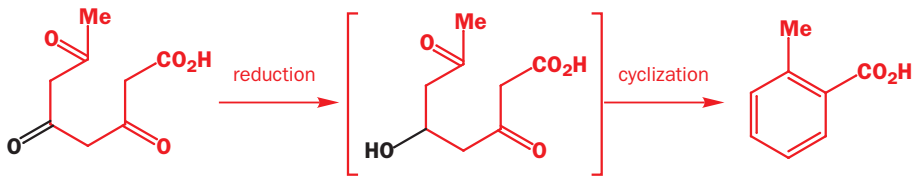
orsellinic acid

This route has been demonstrated by feeding ^{13}C-labelled malonyl CoA to a microorganism. The orsellinic acid produced has three ^{13}C atoms only, seen by an $M + 3$ peak in the mass spectrum. The location of the labels can be proved by NMR. The starter unit, acetate, is not labelled.



orsellinic acid

As the polyketide chain is built up, any of the reductions or eliminations from fatty acid biosynthesis can occur at any stage. The simple metabolite 6-methyl salicylic acid (6-MSA) is made in the microorganism *Penicillium patulum*, and it could come from the same intermediate as orsellinic acid with one reduction.

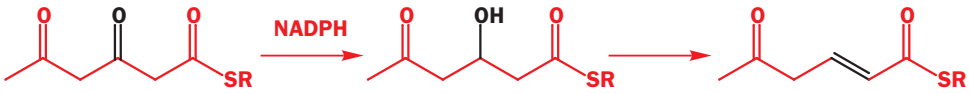


linear tetraketide

6-methyl salicylic acid

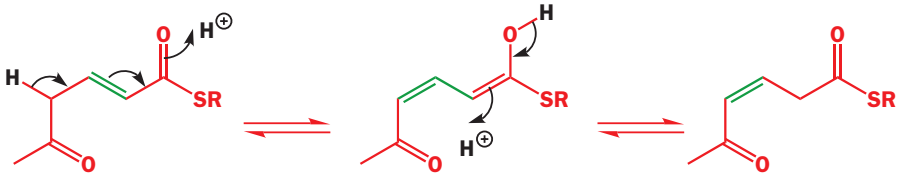
Reduction to the alcohol or to the unsaturated acid or ketone would give the right oxidation level and could occur as the chain is built, after it is completed, or after cyclization. In fact, reduction to

the conjugated unsaturated triketide occurs as the third acetate unit is added, just as the fatty acid route would lead us to expect.



linear triketide

This intermediate cannot cyclize as it has a *trans* double bond and the ends cannot reach each other. First, the double bond is moved out of conjugation with the COSR group, again as in the fatty acids, except that here the new *Z* double bond moves into conjugation with the remaining keto group.

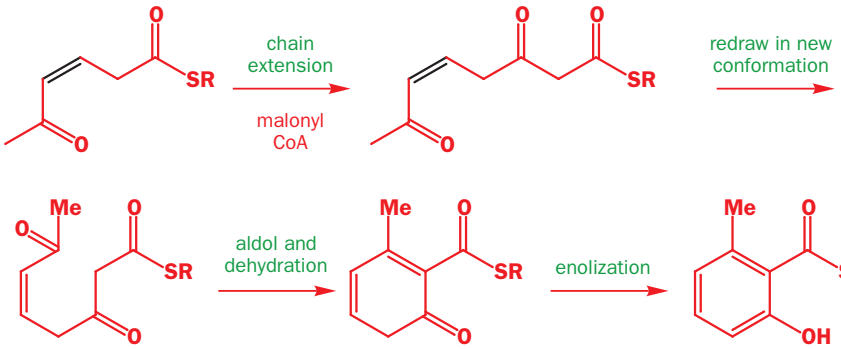


E-alkene conjugated with acid group

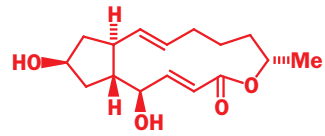
extended enol

Z-alkene conjugated with ketone group

Now the last chain extension occurs and the completed *Z*-tetraketide cyclizes to 6-methyl salicylic acid. Chemically, we would prefer not to carry the unstable *Z*-enone through several steps, but Nature controls these reactions very precisely.



This precise sequence was discovered only through very careful double labelling experiments and after the discovery of specific inhibitors for the enzyme. Since polyketides can be made from the acyl polymalonate pathway with or without reduction and elimination at any step, the number of possible structures is vast. With more reduction, no aromatic ring can be formed: macrolide antibiotics such as brefeldin A come from this route.



brefeldin A

If you examine this structure, you should be able to find a continuous carbon chain made from an acetate starter unit and seven malonyl CoA units with full or partial reduction occurring after many acylation steps.

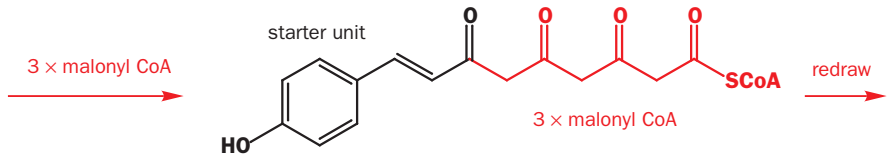
Other starter units

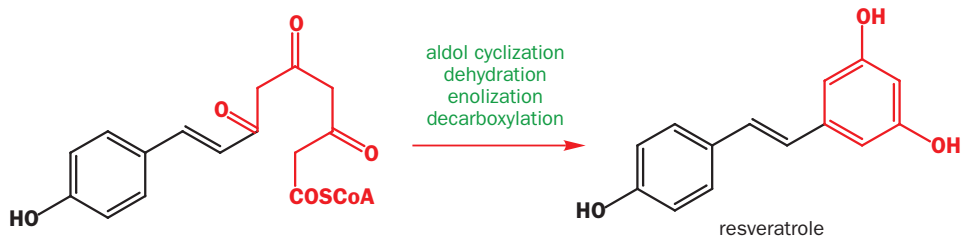
So far we have started the chain with acetate, but many other starter units are used. Some important groups of compounds use shikimic acid metabolites such as cinnamic acid (Chapter 50) as starter units. They include the widespread plant flavones and the anthocyanidin flower pigments.

4-hydroxy cinnamic acid from shikimic acid

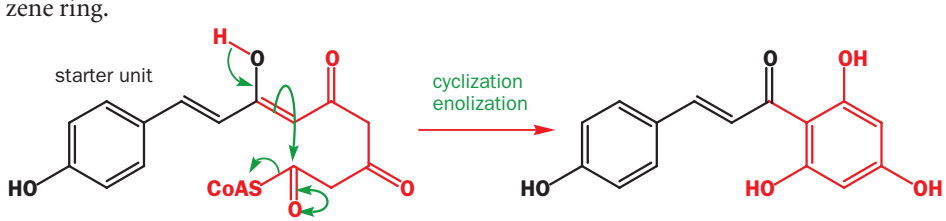
activated for polymalonate addition

The most common sequence uses three malonyl CoA acylations followed by cyclization to a new aromatic ring. The simplest type is exemplified by resveratrole, the compound in red wine that helps to prevent heart disease. Each step in this sequence is a simple reaction that you have met before.

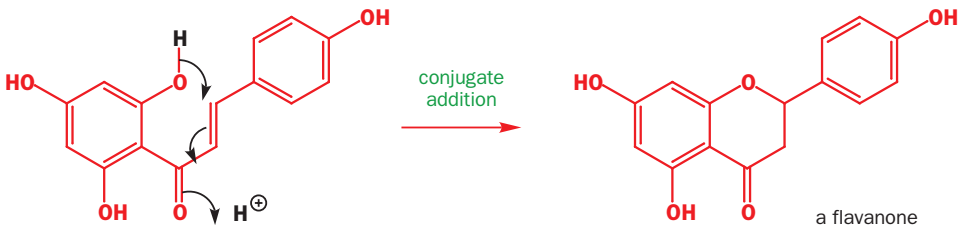




A different cyclization leads to the flavones and anthocyanidins. Reaction of the stable enol from a 1,3-diketone with the thiol ester as electrophile results in acylation at carbon in the manner of the Claisen ester condensation (Chapter 28) with loss of CoASH and the formation of a trihydroxybenzene ring.

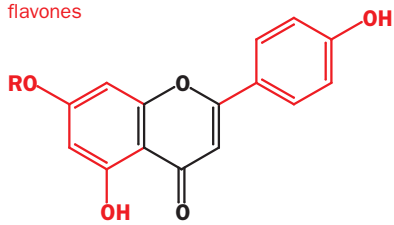


This cyclization is followed by a conjugate addition of an *ortho* phenolic OH group on to the enone system. The product is a flavanone structure, which is always drawn a different way up to the molecules we have just been discussing. Redrawing the last product shows the cyclization.



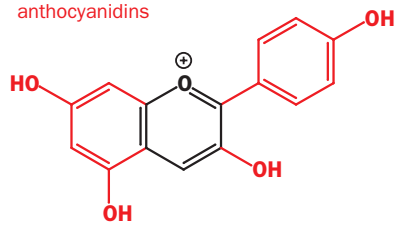
Aromatization of the central oxygen heterocycle by oxidation leads to the flavones, which are yellow or orange depending on their substituents. Dehydration leads to the red or blue anthocyanidins, pigments of flowers and fruit. This important group of molecules also includes plant growth hormones and defence compounds.

flavones



R = H; naringenin, R = glucose; naringin
—a bitter substance from grapefruit peel

anthocyanidins

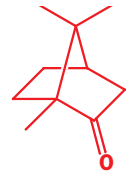


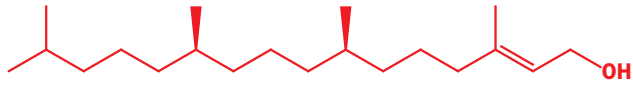
pelargonidin, pigment of raspberries,
geraniums, and red grape skins

Terpenes are volatile constituents of plant resins and essential oils

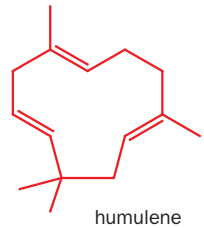
Terpenes were originally named after turpentine, the volatile oil from pine trees used in oil painting, whose major constituent is α-pinene. The term was rather vaguely used for all the volatile oily compounds, insoluble in water and usually with resiny smells from plants. The oils distilled from plants, which often contain perfumery or flavouring materials, are called **essential oils** and these too contain terpenes. Examples include camphor from the camphor tree, used to preserve clothes from moths, humulene from hops, which helps to give beer its flavour, and phytol, found in many plants.


α-pinene

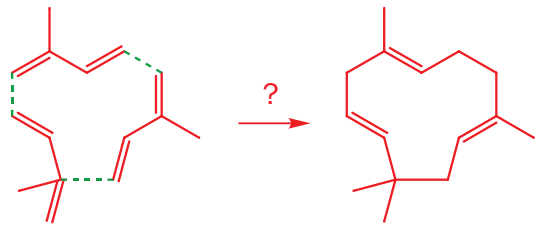

camphor



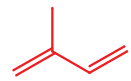
phytol


humulene

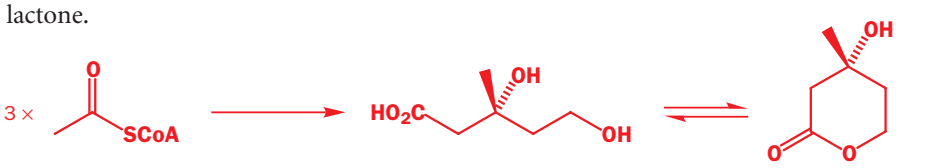
You will notice that they are all aliphatic compounds with a scattering of double bonds and rings, few functional groups, and an abundance of methyl groups. A better definition (that is, a biosynthetically based definition) arose when it was noticed that all these compounds have 5*n* carbon atoms. Pinene and camphor are C_{10} compounds, humulene is C_{15}, and phytol is C_{20}. It seemed obvious that terpenes were made from a C_5 precursor and the favourite candidate was isoprene (2-methylbuta-1,3-diene) as all these structures can be drawn by joining together 2-, 3-, or 4-isoprene skeletons end to end. Humulene illustrates this idea.



3 isoprene units humulene

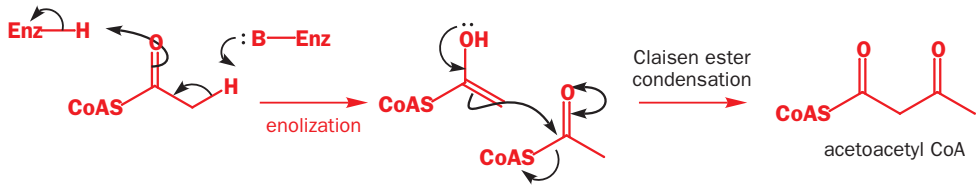

isoprene

In fact, this is not correct. Isoprene is not an intermediate, and the discovery of the true pathway started when acetate was, rather surprisingly, found to be the original precursor for all terpenes. The key intermediate is mevalonic acid, formed from three acetate units and usually isolated as its lactone.

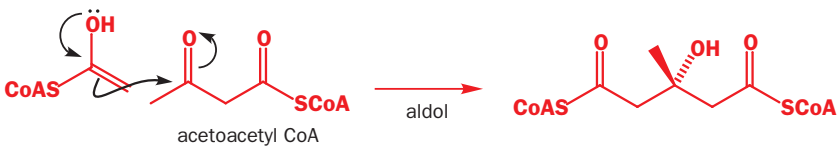


acetyl CoA mevalonic acid mevalonolactone

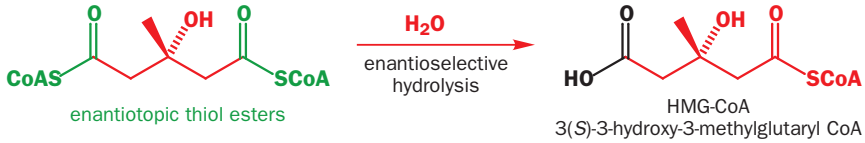
The first step is the Claisen ester condensation of two molecules of acetyl CoA, one acting as an enol and the other as an electrophilic acylating agent to give acetoacetyl CoA. We saw the same reaction in the biosynthesis of the pyrrolidine alkaloids earlier in this chapter.



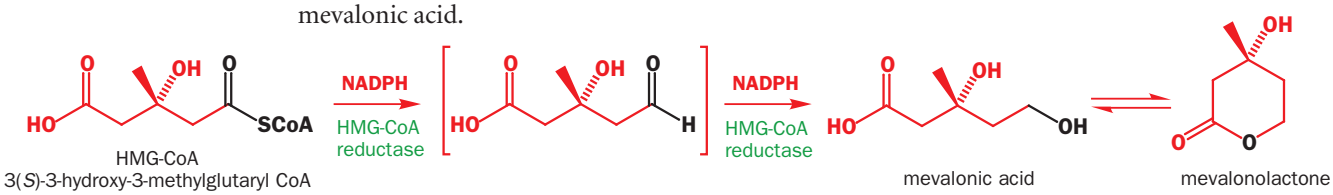
The third molecule of acetyl CoA also functions as a nucleophilic enol and attacks the keto group of acetoacetyl CoA. This is not a Claisen ester condensation—it is an aldol reaction between the enol of a thiol ester and an electrophilic ketone.



We have drawn the product with stereochemistry even though it is not chiral. This is because one of the two enantiotopic thiol esters is hydrolysed while this intermediate is still bound to the enzyme, so a single enantiomer of the half-acid/half-thiol ester results.

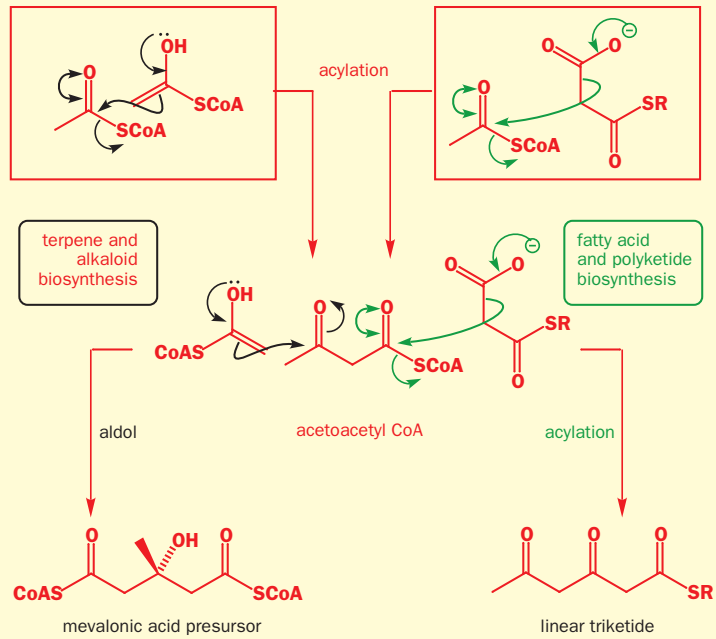


The remaining thiol ester is more electrophilic than the acid and can be reduced by the nucleophilic hydride from NADPH. Just as in $LiBH_4$ reductions of esters (Chapter 24), the reaction does not stop at the aldehyde level, and two molecules of NADPH are used to make the alcohol. This is mevalonic acid.



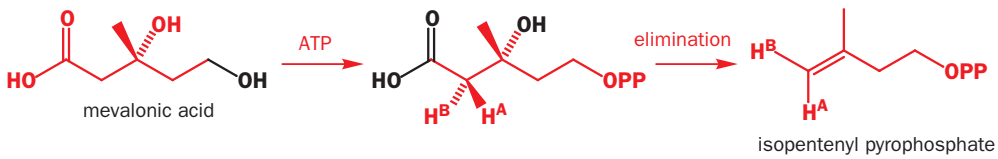
Different pathways; different reactivity

Acetyl CoA (as an enol) and malonyl CoA are both acylated by acetyl CoA as an electrophile, but the behaviour of the two nucleophiles is different when they react with acetoacetyl CoA. Malonyl CoA is acylated while acetyl CoA does the aldol reaction. This could be enzymatic control.

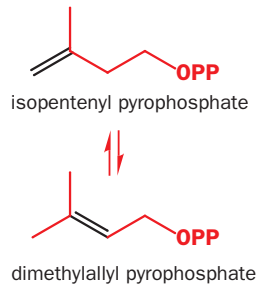


Mevalonic acid is indeed the true precursor of the terpenes but it is a C_6 compound and so it must lose a carbon atom to give the C_5 precursor. The spare carbon atom becomes CO_2 by an elimination reaction. First, the primary alcohol is pyrophosphorylated with ATP (Chapter 49); then the CO_2H group and the tertiary alcohol are lost in a concerted elimination. We know it is concerted because labelling the diastereotopic hydrogen atoms on the CH_2CO_2H group reveals that the elimination is stereospecific.

▶ 'PP' indicates the pyrophosphate group transferred from ATP.

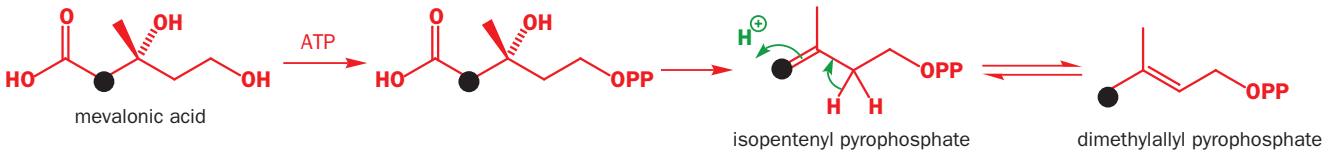


So is isopentenyl pyrophosphate the C_5 intermediate at last? Well, yes and no. There are actually two closely related C_5 intermediates, each of which has a specific and appropriate role in terpene biosynthesis. Isopentenyl pyrophosphate is in equilibrium with dimethylallyl pyrophosphate by a simple allylic proton transfer.

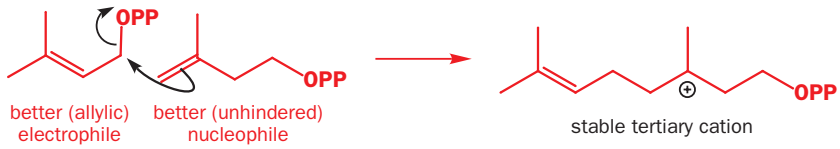


This is again a concerted reaction and again we know that by proton labelling. One of the two enantiotopic protons (H^S in the diagram) is lost from the bottom face of the allylic CH_2 group while the new proton is added to the top face of the alkene. This is an *anti* rearrangement overall.

The stereochemical details are interesting in establishing the mechanism but not important to remember. What is important is that the origin of the two methyl groups in dimethylallyl pyrophosphate is quite distinct and can easily be traced if you always draw the intermediates in the way we have drawn them. We will now switch to ^{13}C labelling to make the point.

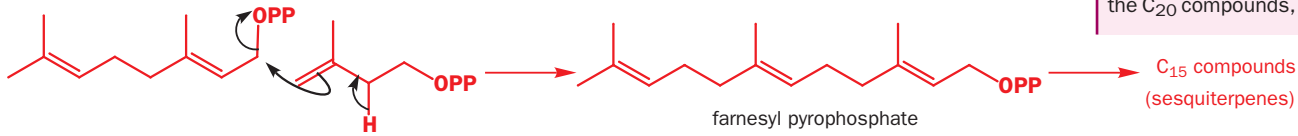


The two C_5 intermediates now react with each other. The dimethylallyl pyrophosphate is the better electrophile because it is allylic, and allylic compounds are good at both S_N1 and S_N2 reactions (Chapter 17). Isopentenyl pyrophosphate is the better nucleophile because it can react through an unhindered primary carbon atom to produce a tertiary cation. This is what we have in mind.



Though this idea reveals the thinking behind the reaction, in fact it does not go quite like this. The product is one particular positional and geometrical isomer of an alkene and the cation is not an intermediate. Indeed, the reaction is also stereospecific (discovered again by proton labelling, but we will not give the rather complex details) and this too suggests a concerted process.

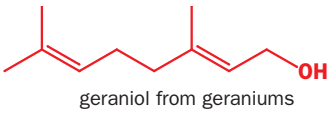
Geranyl pyrophosphate is the starting point for all the monoterpenes. It is still an allylic pyrophosphate and repeating the alkylation with another molecule of isopentenyl pyrophosphate gives farnesyl pyrophosphate, the starting point for the sesquiterpenes, and so on.



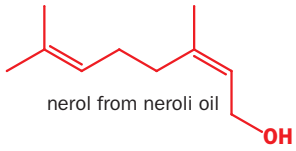
As soon as we start to make typical cyclic monoterpenes from geranyl pyrophosphate we run into a snag. We cannot cyclize geranyl pyrophosphate because it has a *trans* double bond! We *could* cyclize the *cis* compound (neryl pyrophosphate), and it used to be thought that this was formed from the *trans* compound as an intermediate.

▶ Though terpenes are made from C_5 units, they are classified in C_{10} units. The monoterpenes are the C_{10} compounds, the sesquiterpenes (*sesqui* is Latin for one-and-a-half) are the C_{15} compounds, the diterpenes are the C_{20} compounds, and so on.

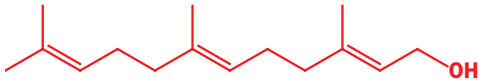
many of these names are derived from fragrant plant oils:



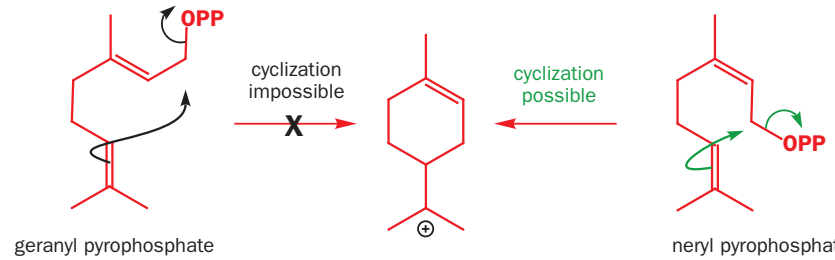
geraniol from geraniums



nerol from neroli oil

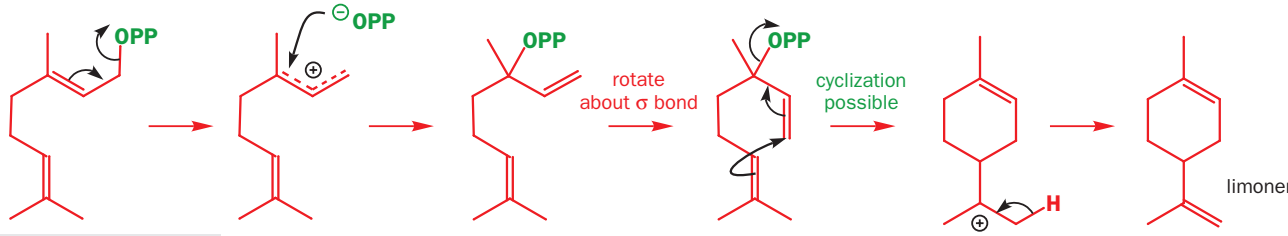


farnesol, also present in neroli oil



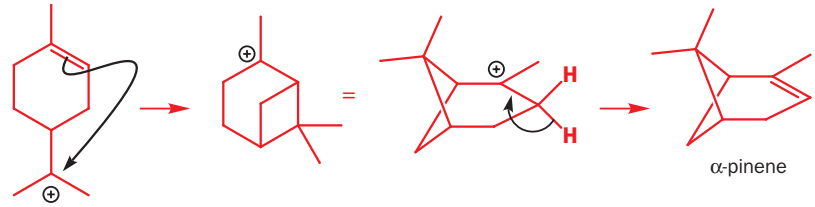
geranyl pyrophosphate — neryl pyrophosphate

It is now known that Nature gets round this problem without making neryl pyrophosphate. An allylic rearrangement occurs to move the pyrophosphate group to the tertiary centre. This is an unfavourable rearrangement thermodynamically and probably occurs via the allyl cation and catalysed by Mg(II). There is no longer any geometry about the alkene. The molecule can now rotate freely about a single bond and cyclization can occur. Even if only a small amount of the rearranged allylic pyrophosphate is present, that can rearrange and more can isomerize.

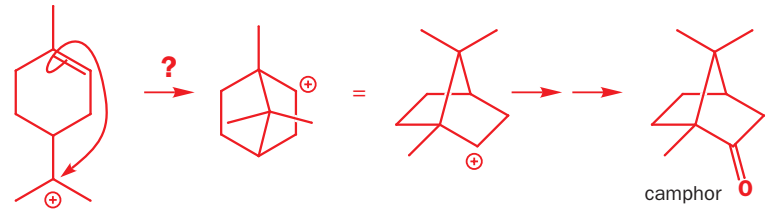


■ The product here is limonene—a terpene of the peel of citrus fruits. One enantiomer occurs in lemon peel—the other in orange peel. See Chapter 45.

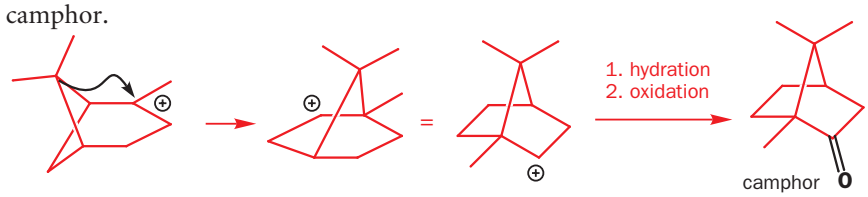
More interesting compounds come from the cyclization of the first formed cation. The remaining alkene can attack the cation to form what looks at first to be a very unstable compound but which is actually a tertiary carbocation with the pinene skeleton.



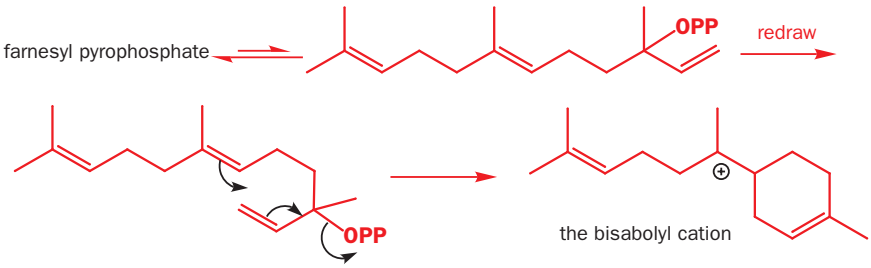
The camphor skeleton looks as though it might be formed by cyclization of the wrong end of the alkene on to the cation. This would certainly give the right skeleton but the intermediate secondary cation is rather unlikely.



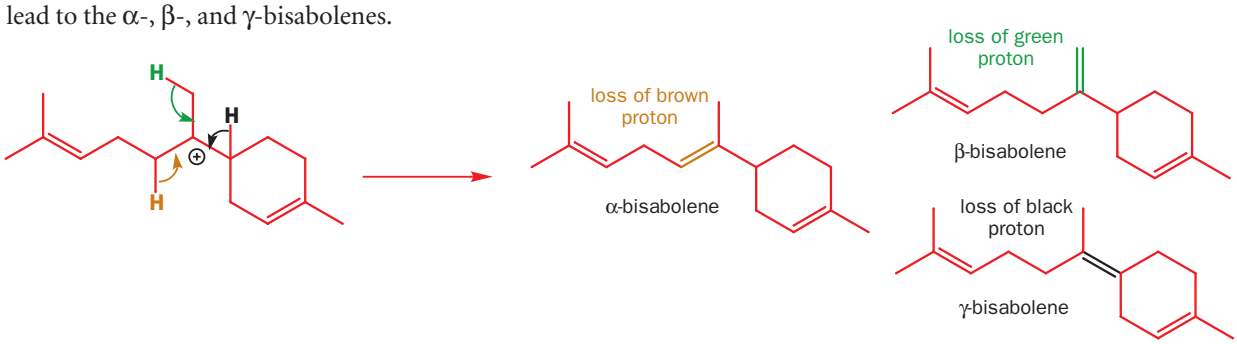
There is a better route. The more likely cation formed on the way to pinene could rearrange to the camphor cation. This is a known chemical reaction and is a simple 1,2-shift of the kind discussed in Chapter 37. However the new cation is formed, addition of water and oxidation would give camphor.



In the sesquiterpene series, similar cyclizations lead to an amazing variety of products. After the initial unfavourable allylic rearrangement of the pyrophosphate group, farnesyl pyrophosphate can give a six-membered ring cation known as the bisabolyl cation.



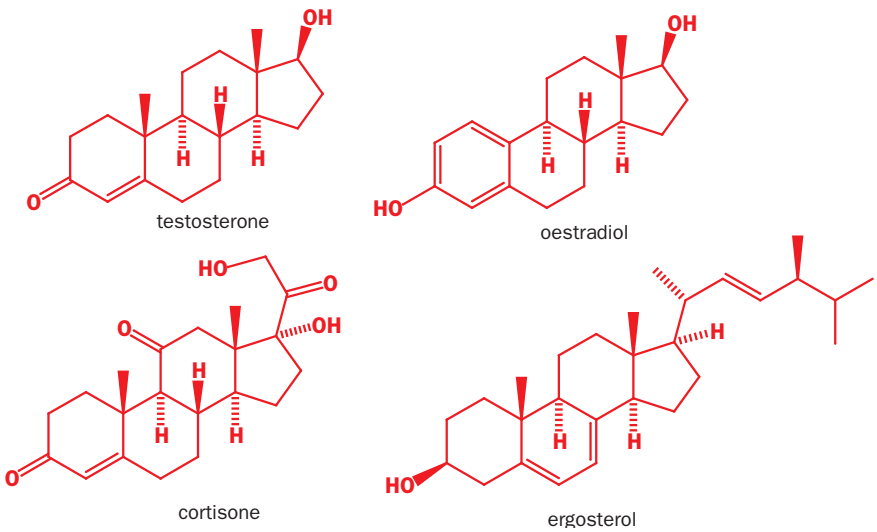
This cation does many things but it takes its name from the three fairly random proton losses that lead to the α-, β-, and γ-bisabolenes.



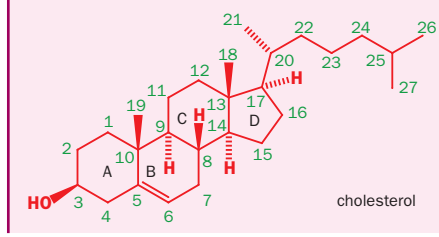
Many other reactions give even larger and more complex terpenes with a variety of functionalization but we will treat only one group in detail. These compounds are so important to us that they are given a different name.

Steroids are metabolites of terpene origin

Two types of human hormone are steroidal—the sex hormones such as oestradiol and testosterone and the adrenal hormones such as cortisone. Cholesterol is a steroid too, as is vitamin D, derived from ergosterol.

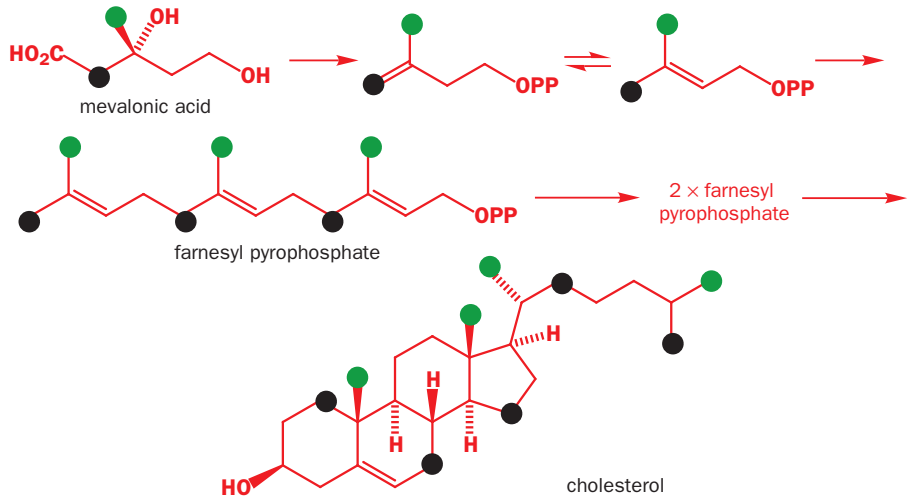


▶ For reference, here is the numbering of the steroid nucleus, not because we want you to learn it, but because it is often used without explanation in books and it is not obvious.

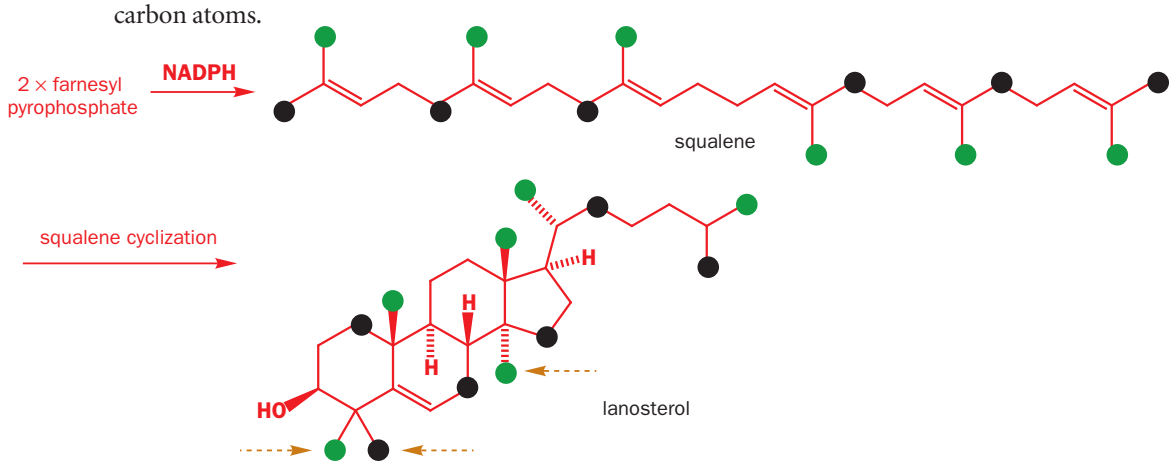


All share the skeleton of four fused rings, three six-membered and one five-membered and conventionally lettered A–D. Beyond the ring stereochemistry and some common oxygenation patterns they share little else. Some (such as the female sex hormones) have an aromatic A ring; some have side-chains on the five-membered ring.

At first glance, it is not at all clear that steroids are terpenoid in origin. The 5*n* numbers are absent—cholesterol is a C_{27} compound while the others variously have 20, 21, or 23 carbon atoms. Studies with labelled mevalonic acid showed that cholesterol is terpenoid, and that it is formed from two molecules of farnesyl pyrophosphate ($2 \times C_{15} = C_{30}$ so three carbon atoms must be lost). Labelling of one or other of the methyl groups (two experiments combined in one diagram!) showed that two of the green carbon atoms and one of the black carbon atoms were lost during the biosynthesis.



It is not obvious how the two farnesyl pyrophospate molecules could be combined to make the steroid skeleton, and the chemistry involved is extraordinary and very interesting. The first clues came from the discovery of the intermediates squalene and lanosterol. Squalene is obviously the farnesyl pyrophosphate dimer we have been looking for while lanosterol looks like cholesterol but still has all 30 carbon atoms.



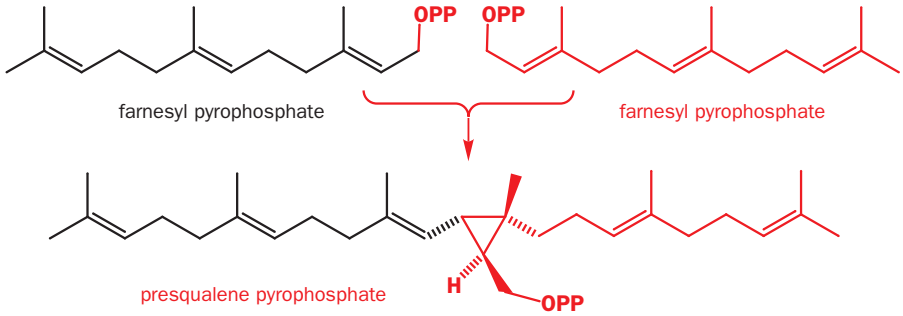
The three carbon atoms that are lost from lanosterol (C_{30}) in its conversion to cholesterol (C_{27}) are marked with brown arrows. Now at least we know which carbon atoms are lost. But many questions remain to be answered.

- How does farnesyl pyrophosphate dimerize so that two electrophilic carbon atoms (CH_2OPP) join together?
- Why does the formation of squalene require the reducing agent NADPH?
- How does squalene cyclize to lanosterol so that the very odd labelling pattern can be achieved?
- Where do the three lost carbon atoms go?
- How is the sterochemistry controlled?

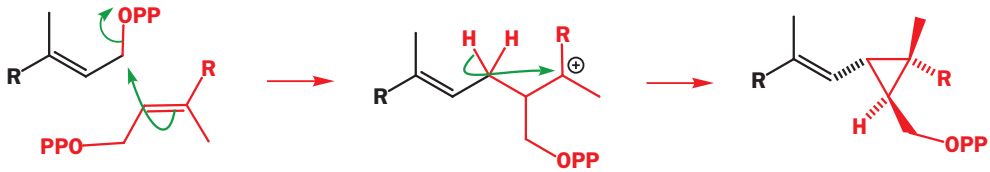
Before we tell you the answers, be warned: prepare for some surprises, and be ready to hold back outright disbelief!

The formation of squalene from farnesyl pyrophosphate

If the reducing agent NADPH is omitted from the cell preparation, squalene is not formed. Instead, another farnesyl pyrophosphate dimer accumulates—presqualene pyrophosphate—which has a three-membered ring and in which we can see that the two molecules of farnesyl pyrophosphate are joined in a slightly more rational way.



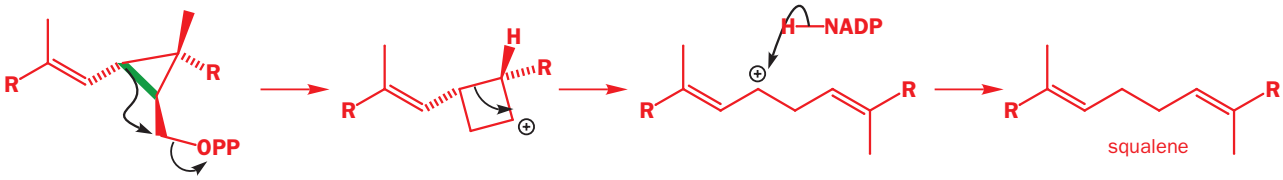
Maybe it's not so obvious that this is more rational! The first C–C bond formation is quite straightforward. The alkene in the red molecule attacks the allylic pyrophosphate in the black molecule in a simple S_N2 reaction. The product is a stable carbocation. Only one C–C bond remains to be formed to close the three-membered ring and this occurs by the loss of a proton from the black molecule.



■ We will abbreviate the long terpene side-chain to 'R' from now on.

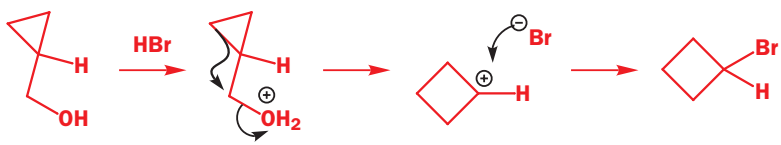
This is a very remarkable reaction. Such reactions do not occur chemically: this biological one occurs only because the molecule is held in the right shape by the enzyme and because the new ring is three-membered. Three-membered rings are very easily formed but also very easily opened—and that is what happens to this ring. In the presence of NADPH, a series of rearrangements gives a series of carbocations, the last of which is trapped by reduction.

The first step is the migration of one of the bonds (shown in green) of the three-membered ring to displace the pyrophosphate leaving group, expand the ring to four-membered, and release some strain. Now the cyclobutyl cation breaks down to give an open-chain allylic cation stabilized by one of the alkenes. This is the cation that is reduced by NADPH.



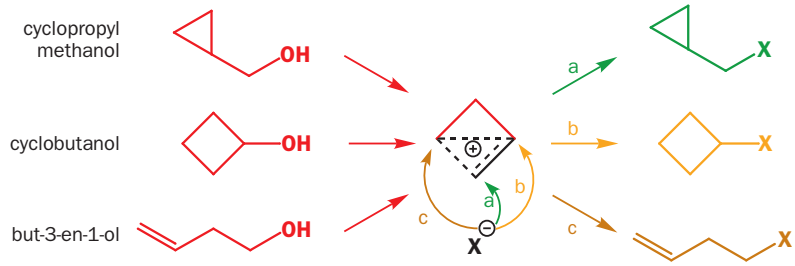
If you follow this sequence backwards, you will see that the originally formed 'rational' bond (shown in green) is the one that migrated and is retained in squalene, while the second bond is cleaved in the last step.

This may all seem far-fetched, but it happens in laboratory reactions too! Treatment of the simplest cyclopropyl alcohol with HBr gives cyclobutyl bromide by a similar rearrangement.



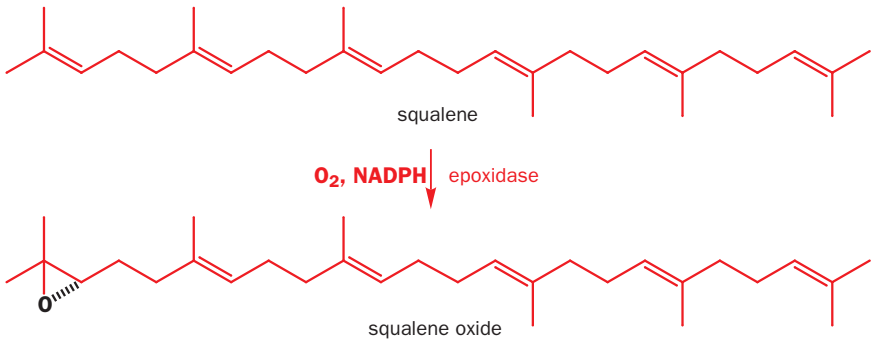
In fact, cyclopropylmethyl compounds, cyclobutyl compounds, and homoallyl compounds are all in equilibrium in acid solution and mixtures of products are often formed. The delocalized cation

shown has been suggested as an intermediate. Make sure that you can draw mechanisms for each starting material to give the intermediate cation and from the cation to each product.

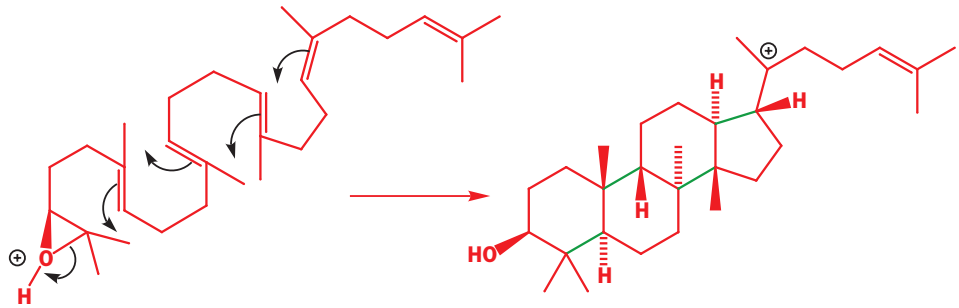


Squalene to lanosterol

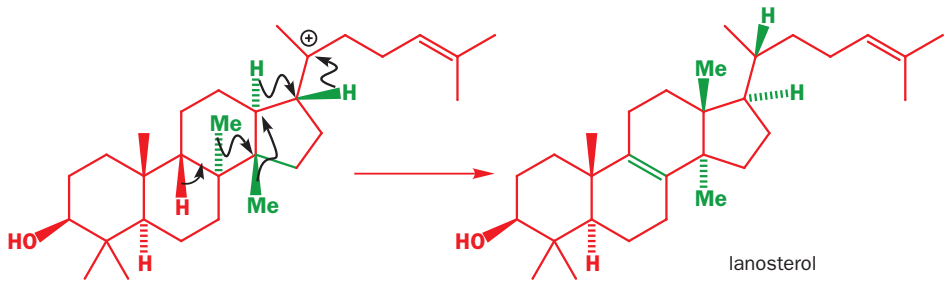
The next step is simple—the epoxidation of one of the terminal double bonds—but it leads to two of the most remarkable reactions in all of biological chemistry. Squalene is not chiral, but enzymatic epoxidation of one of the enantiotopic alkenes gives a single enantiomer of the epoxide with just one stereogenic centre.



We will start now to draw squalene in a coiled up way as the next step is the polycyclization of the epoxide. The basic reaction is best seen first in the flat, though we will draw the stereochemistry immediately. The first alkene cyclizes on to the epoxide and then each remaining alkene cyclizes on to the next to give a stable tertiary cation.

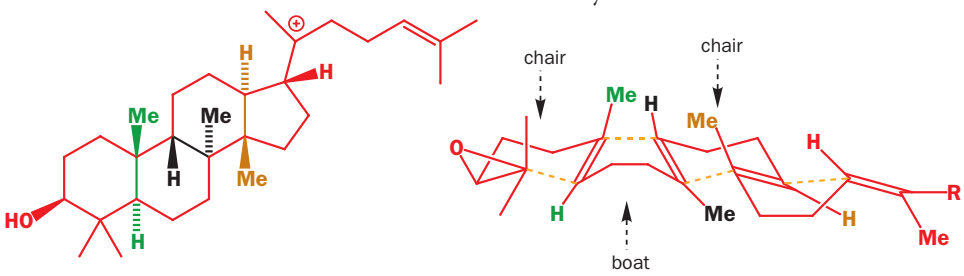


By analogy with what has gone before, you might now expect a tame hydration or reduction of this cation. Nothing of the sort! A rearrangement occurs in which *five* consecutive 1,2-shifts are followed by an elimination. Since this reaction organizes the backbone of the steroids, it is often called the **steroid backbone rearrangement**.



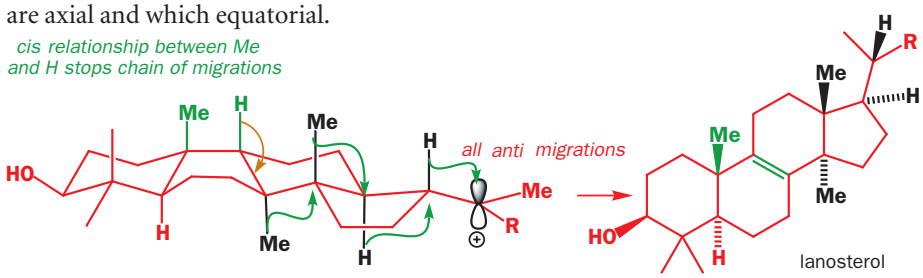
Finally, we have reached lanosterol. Now we will go back over these two steps and discuss them a bit more. Consider first the regiochemistry of the cyclization. The epoxide opens in the way we would expect to give positive charge at the more substituted carbon atom and then all the alkenes attack through their less substituted end (again as we would expect to give positive charge at the more substituted carbon atom)—all except one. The third alkene cyclizes the 'wrong' way—this is presumably a result of the way the molecule is folded.

We learn much more about the folding by examining the stereochemistry of the product cation. First, all of the stereochemistry of each alkene is faithfully reproduced in the product: the cyclization is stereospecific. This is emphasized in colour in the diagram. The green stereochemistry arises because the green Me and H were *trans* in the first alkene of squalene, the black Me and H *trans* in the second, and the brown *trans* in the third. But what about the relationship between the green methyl and the black H? Or between the black and brown methyls? These were determined by the folding and the key observation is that all the relationships are *trans* except that between the green Me and the black H. Now we can draw a conformation for the cyclization.



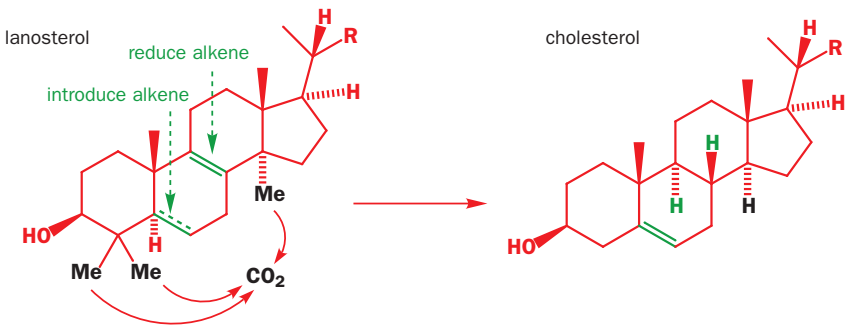
When the transition state for a ring closure forms a chair then a *trans* relationship results. This is the case for the black Me and brown Me. When a boat is formed a *cis* relationship results. This is the case for the green Me and black H. Squalene folds up in a chair–boat–chair conformation and that leads to the observed stereochemistry.

Next, we need to look at the stereochemistry of the rearrangement step. If we draw the product cation as nearly as possible in the conformation of folded squalene, we will see which substitutents are axial and which equatorial.



Each group that migrates (black) is axial and is anti-periplanar to the one before so that each migrating group does an S_N2 reaction on the migration terminus with inversion. The chain stops because of the *cis* relationship between the green Me and H in ring B and an elimination of the green H is all that can happen.

The remainder of the biosynthesis of cholesterol requires various redox reactions and is a bit of an anticlimax: the details are summarized in the scheme below.



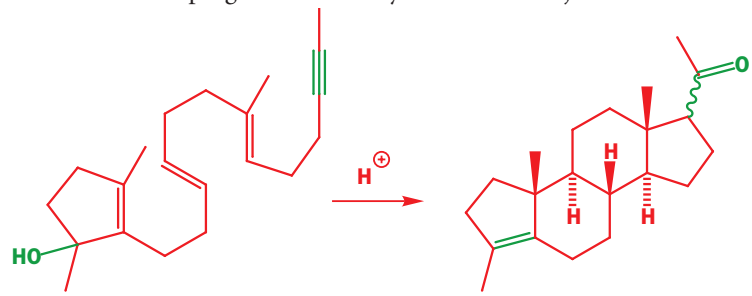
Biomimetic synthesis: learning from Nature

When new and academic-looking reactions are discovered in the laboratory, it often seems only a short time before they are found in nature as well. However, the development of polyolefin cyclization reactions in synthesis occurred by the reverse philosophy—it was inspiration from Nature that led W. S. Johnson to use the reactions in synthesis, including steroid synthesis. This is **biomimetic synthesis**, a strategy that is bound to work provided we can just master the practical details.

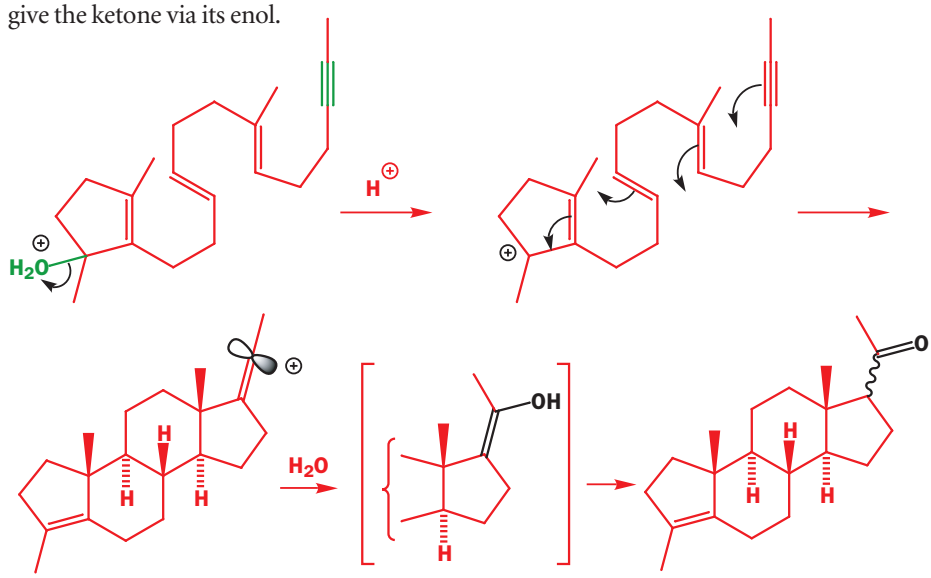
There are quite a lot of differences between the chemical and the biochemical versions so far—the chemical ones are less complex and less sophisticated but more versatile. The reactions are just cyclizations without the backbone rearrangements. The most important points of difference are:

- The cyclization is usually begun with a cation from treatment of a cyclic tertiary alcohol rather than an epoxide
- The cyclization sequence is terminated with an alkyne or an allyl silane rather then with simple alkene
- The substituents are placed in the correct positions in the starting material as no rearrangement follows cyclizations
- The cyclizations are all stereospecific as in nature but the rings coil up in an all-chair fashion rather than in a chair–boat–chair fashion as there is no enzyme to shape the molecule
- The product cation is quenched by addition of water rather than loss of a proton

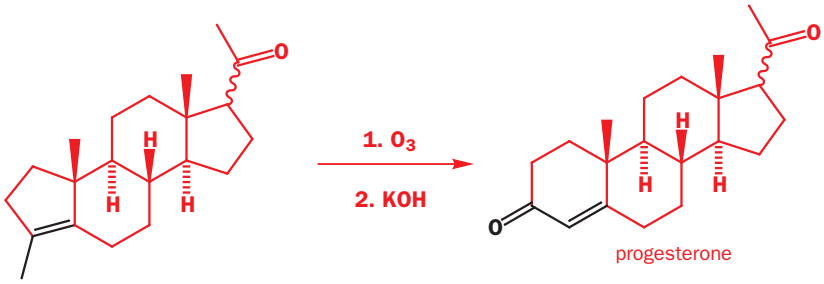
Here is one of Johnson's best examples which leads eventually to a biomimetic synthesis of the human hormone progesterone. The cyclization occurs just on treatment of the tertiary alcohol with acid.



The first step is the formation of a symmetrical allyl cation, which then initiates the cyclization. The next double bond is disubstituted so that it has no built-in regioselectivity but prefers to form a six-membered rather than a five-membered ring B. The next double bond is trisubstituted and directs the formation of a six-membered ring C. The alkyne, being linear, can reach only through its inner end and so a five-membered ring D is formed. The resulting linear vinyl cation picks up a molecule of water to give the ketone via its enol.

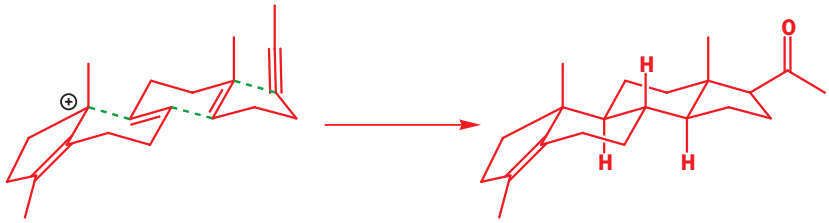


The five-membered ring A is there to ensure efficient initiation of the cyclization by the symmetrical allylic cation. It can easily be opened with ozone and the product cyclized to progesterone.



progesterone

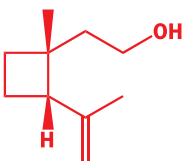
The conformation of the molecule in the moment of cyclization can be seen easily by working backwards from the product. The green dashed lines show new bonds that are being formed. All the six-membered rings in the transition state are chairs and all the ring junctions *trans*. This is an impressive result as there is no enzyme to help the molecule fold up in this way.



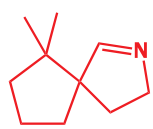
By studying the chemistry that Nature uses in living things we can learn new reactions as well as new ways in which to carry out known reactions. Many of the reactions in this chapter would be laughed at by worldly wise chemists if they appeared in a research proposal, but they have been evolved over millions of years to do precise jobs under mild conditions. Humans have been doing complex organic chemistry for only about a hundred years so that learning from Nature is one of the most important ways in which organic chemistry is advancing at the beginning of the twenty-first century.

Problems

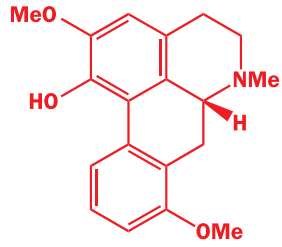
1. Assign each of these natural products to a general class (such as amino acid metabolite, terpene, polyketide) explaining what makes you choose that class. Then assign them to a more specific part of the general class (for example, tetraketide, sesquiterpene).



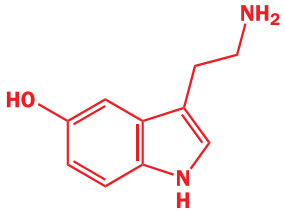
grandisol



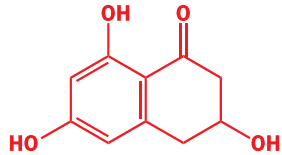
polyzonimine



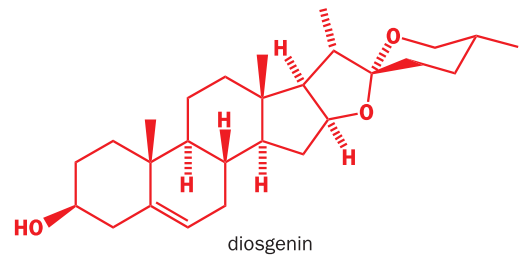
stephanine



serotonin

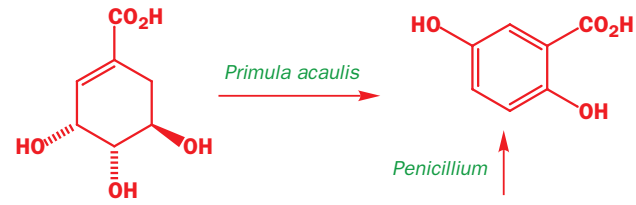


scytalone



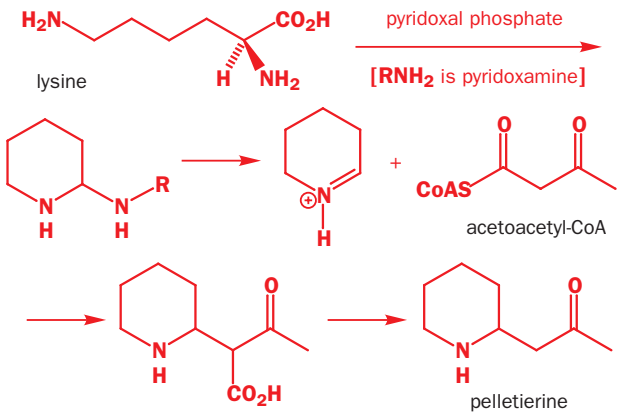
diosgenin

2. Some compounds can arise from different sources in different organisms. 2,5-Dihydroxybenzoic acid comes from shikimic acid (Chapter 50) in *Primula acaulis* but from acetate in *Penicillium* species. Outline details.

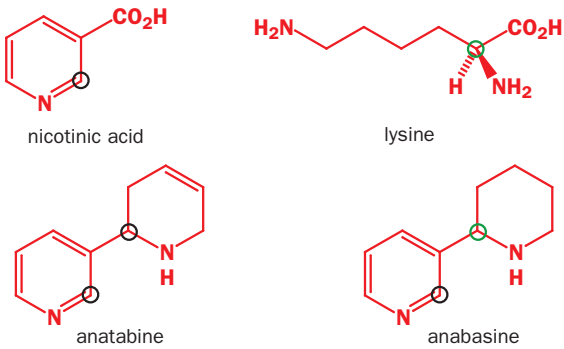


shikimic acid

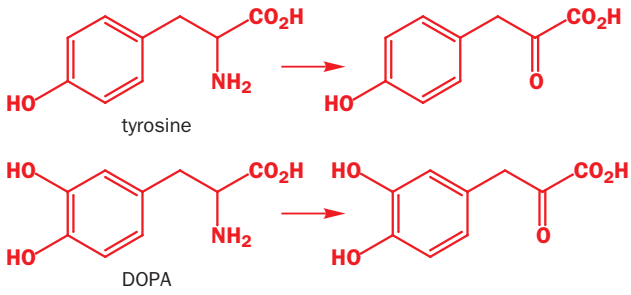
3. The piperidine alkaloid pelletierine was mentioned in the chapter but full details of its biosynthesis were not given. There follows an outline of the intermediates and reagents used. Fill in the details. Pyridoxal chemistry is discussed in Chapter 50.

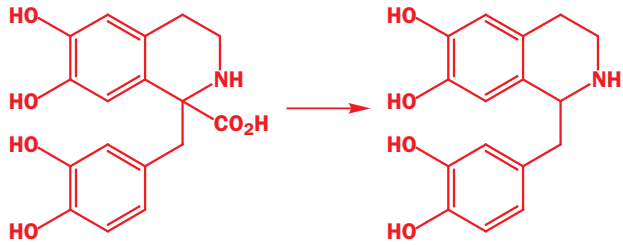


4. The rather similar alkaloids anabasine and anatabine come from different biosynthetic pathways. Labelling experiments outlined below show the origin of one carbon atom from lysine and others from nicotinic acid. Suggest detailed pathways. (*Hint.* Nicotinic acid and the intermediate you have been using in Problem 3 in the biosynthesis of the piperidine alkaloid are both electrophilic at position 2. You also need an intermediate derived from nicotinic acid which is nucleophilic at position 3. The biosynthesis involves reduction.)

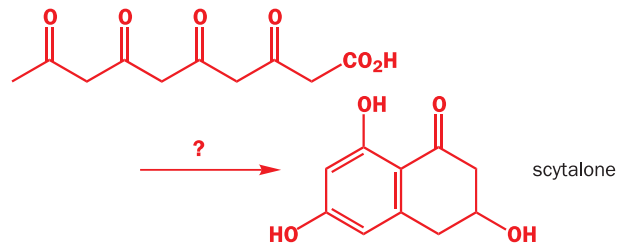


5. The three steps in the biosynthesis of papaverine set out below involve pyridoxal (or pyridoxamine). Write detailed mechanisms.

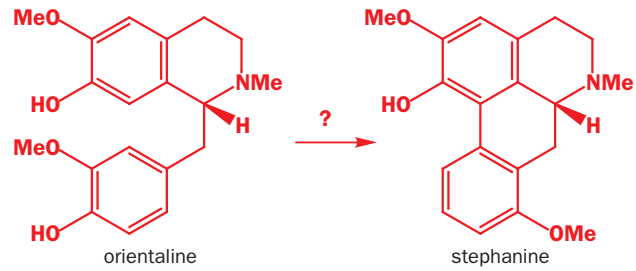




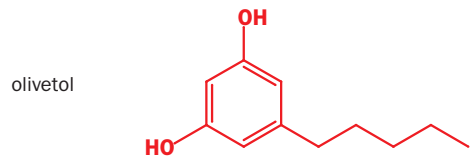
6. Concentrate now on the biosynthesis of scytalone in the first problem. You should have identified it as a pentaketide. Now consider how many different ways the pentaketide chain might be folded to give scytalone.



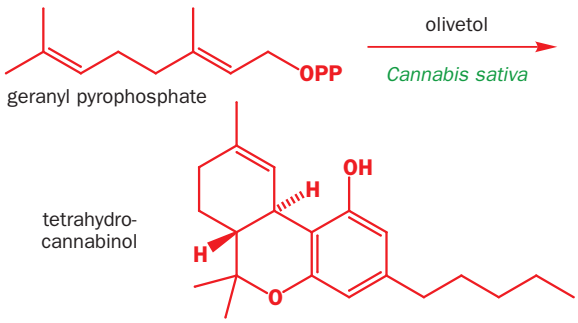
7. This question concerns the biosynthesis of stephanine, another compound mentioned in Problem 1. You should have deduced that it is a benzylisoquinoline alkaloid. Now suggest a biosynthesis from orientaline.



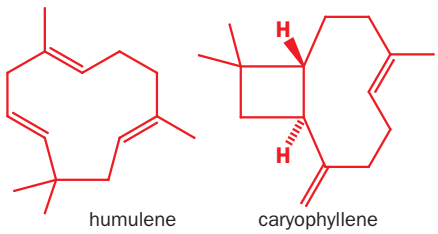
8. Suggest a biosynthesis of olivetol.



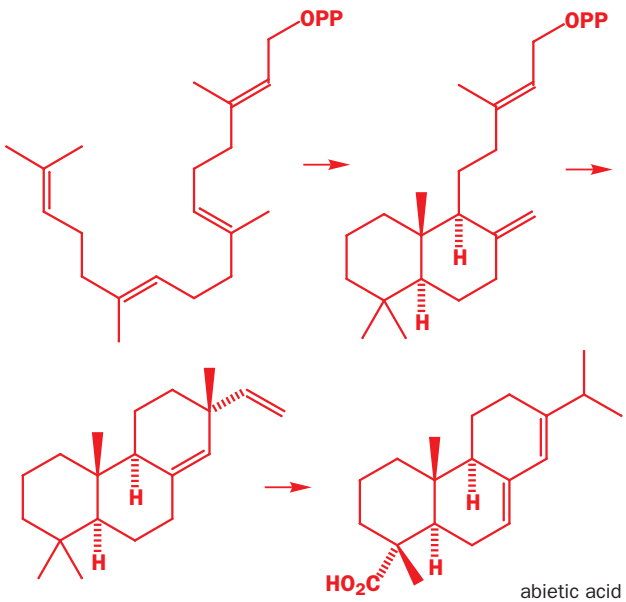
9. Tetrahydrocannabinol, the major psychoactive compound in marijuana, is derived in the *Cannabis* plant from olivetol and geranyl pyrophosphate. Details of the pathway are unknown. Make some suggestions and outline a labelling experiment to establish whether your suggestions are correct.



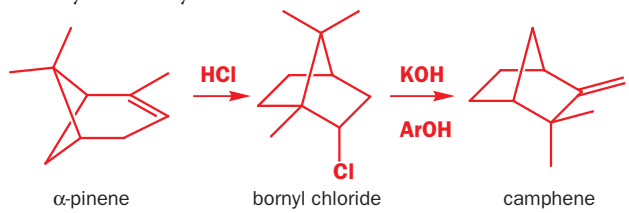
10. Both humulene, mentioned in the chapter, and caryophyllene are made in nature from farnesyl pyrophosphate in different plants. Suggest detailed pathways. How do the enzymes control which product is formed?



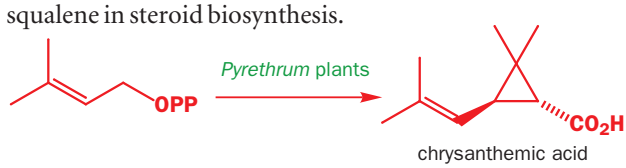
11. Abietic acid is formed in nature from mevalonate via the intermediates shown. Give some more details of the cyclization and rearrangement steps and compare this route with the biosynthesis of the steroids.



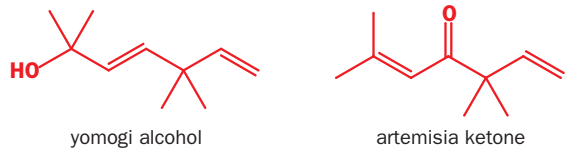
12. Borneol, camphene, and α-pinene are made in nature from geranyl pyrophosphate. The biosynthesis of α-pinene and the related camphor is described in the chapter. In the laboratory bornyl chloride and camphene can be made from α-pinene by the reactions described below. Give mechanisms for these reactions and say whether you consider them to be biomimetic.



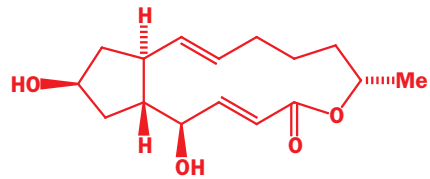
13. Suggest a biosynthetic route to the monoterpene chrysanthemic acid that uses a reaction similar to the formation of squalene in steroid biosynthesis.



How could the same route also lead to the natural products yomogi alcohol and artemisia ketone?



14. In the chapter we suggested that you could detect an acetate starter unit and seven malonate additional units in the skeleton of brefeldin. Give the mechanism of the addition of the first malonyl CoA unit to acetate. Draw out the structure of the complete acyl polymalonate chain and state clearly what must happen to each section of it (reduction, elimination, etc.) to get brefeldin A.



15. This chemical experiment aims to imitate the biosynthesis of terpenes. A mixture of products results. Draw a mechanism for the reaction. To what extent is it biomimetic, and what can the natural system do better?