# 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:

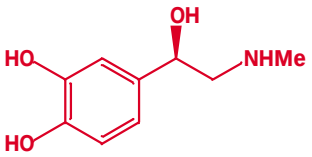
- Natural products are made by secondary metabolism
- Natural products come in enormous variety, but fall mainly into four types: alkaloids, polyketides, terpenes, and steroids
- Alkaloids are amines made from amino acids
- Pyrrolidine alkaloids from ornithine; benzylisoquinoline alkaloids from tyrosine
- 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:

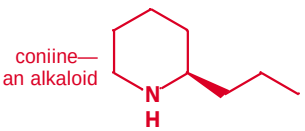
- Organic synthesis ch53

## Introduction

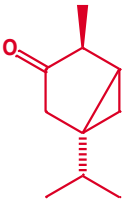
By natural products, we mean the molecules of nature. Of course, all life is made of molecules, and we will not be discussing in great detail the major biological molecules, such as proteins and nucleic acids, which we looked at in Chapters 49 and 50. In this chapter we shall talk much more about molecules 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



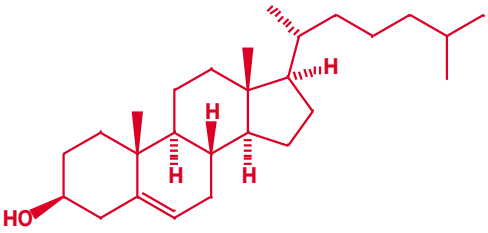
coniine—
an alkaloid



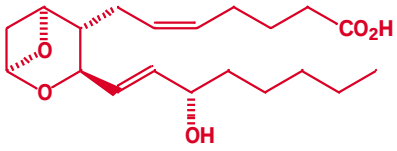
thujone—
a terpene

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.

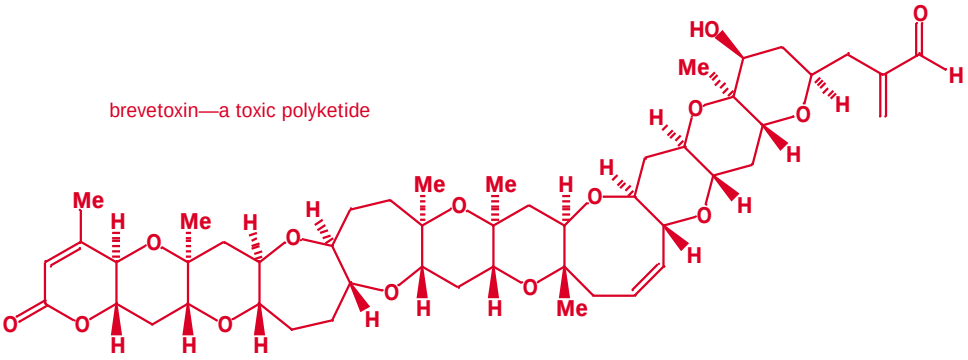



cholesterol—a steroid



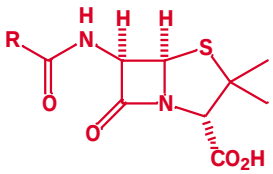

thromboxane $A_2$—a polyketide

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.




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.

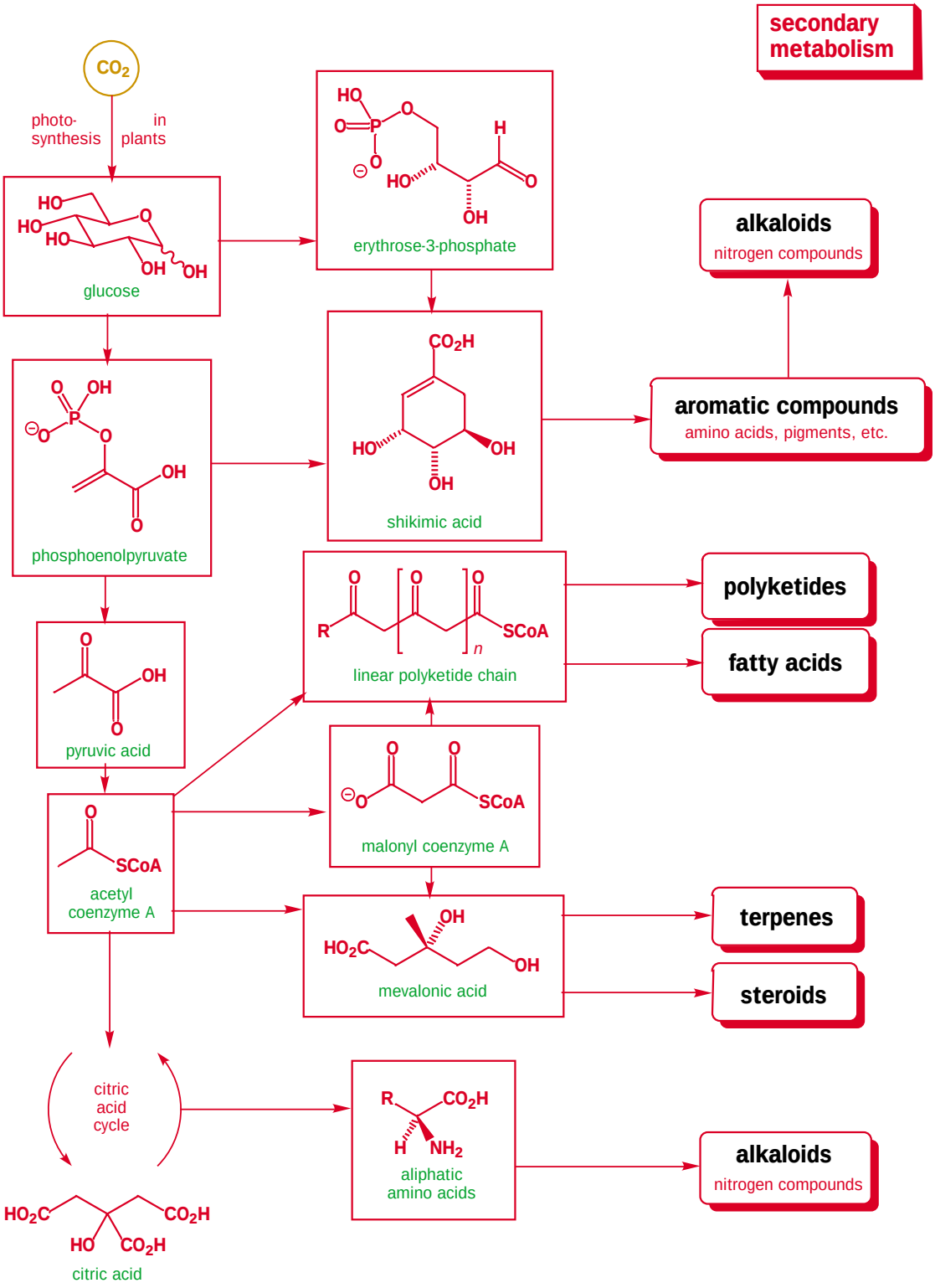
The exploration of the compounds produced by the secondary metabolism of plants, microorganisms, fungi, insects, mammals, and every other type of living thing has hardly begun. Even so, the variety and richness of the structures are overwhelming. Without some kind of classification the task of description 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. 000 relates closely to the chart of primary metabolism in the previous chapter.

## Alkaloids are basic compounds from amino acid metabolism

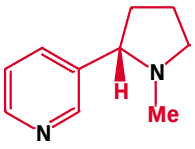
Alkaloids were known in ancient times because they are easy to extract from plants and some of them have powerful and deadly effects. Any plant contains millions of chemical compounds, but some plants, like the deadly nightshade, can be mashed up and extracted with aqueous acid to give a few compounds 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

▶

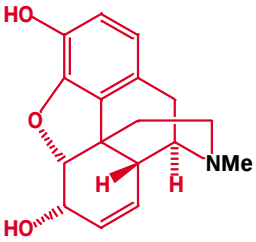
Before moving on, just pause to admire brevetoxin, a wonderful and deadly molecule. Look at the alternating oxygen atoms on the top and bottom faces of alternate rings. Look at the rings themselves—six-, seven-, and eight-membered but each with one and no more than one oxygen atom. Trace the continuous 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*


secondary metabolism
CO2
photo-synthesis
in plants
glucose
erythrose-3-phosphate
alkaloids
nitrogen compounds
aromatic compounds
amino acids, pigments, etc.
shikimic acid
phosphoenolpyruvate
polyketides
fatty acids
linear polyketide chain
pyruvic acid
malonyl coenzyme A
acetyl coenzyme A
terpenes
steroids
mevalonic acid
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

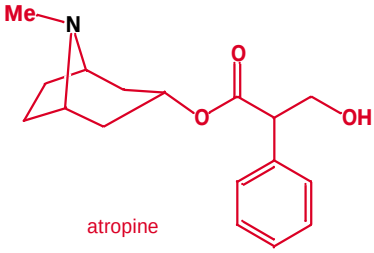
political adversaries to avoid any trouble in the future. Now, we would simply say that they are basic because they are amines. Here is a selection with the basic amino groups marked in black.



nicotine



morphine

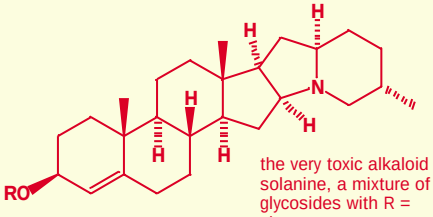


atropine

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 acid, and we will look at three types.

## Solanaceae alkaloids

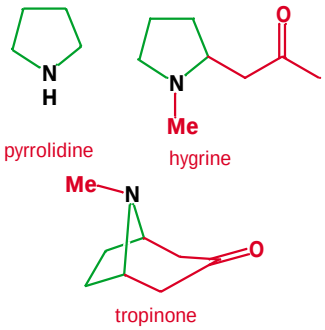
The Solanaceae family includes not only deadly nightshade (*Atropa belladonna*—hence atropine) plants but also potatoes and tomatoes. Parts of these plants also contain toxic alkaloids: for example, you should not eat green potatoes because they contain the toxic alkaloid solanine.



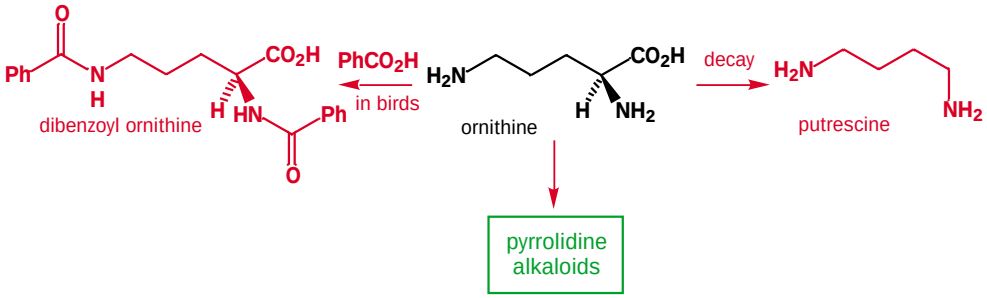
the very toxic alkaloid solanine, a mixture of glycosides with R = glucose, mannose, etc.

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.

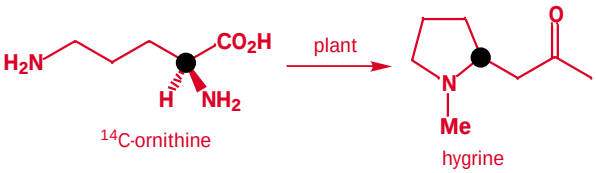
## Pyrrolidine alkaloids are made from the amino acid ornithine



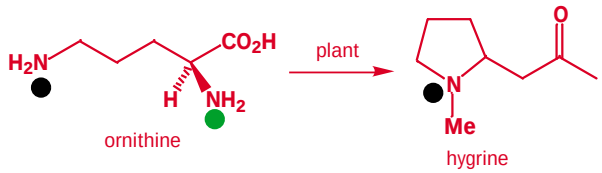

Pyrrolidine is the simple five-membered cyclic amine and pyrrolidine alkaloids contain this ring somewhere in their structure. Both nicotine and atropine contain a pyrrolidine ring as do hygrine and tropinone. All are made in nature from ornithine. Ornithine is an amino acid not usually found in proteins 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'.




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.



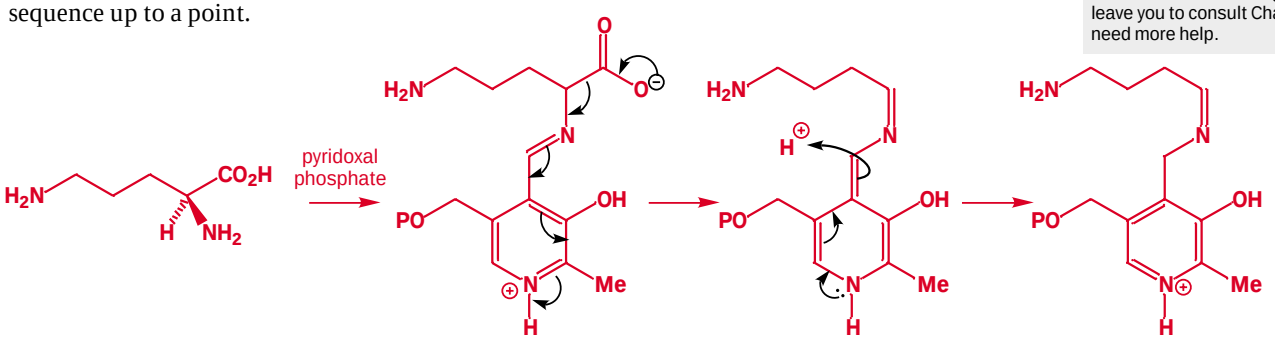
If each amino group in ornithine is labelled in turn with $^{15}N$, the α amino group is lost but the γ amino group is retained.




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.

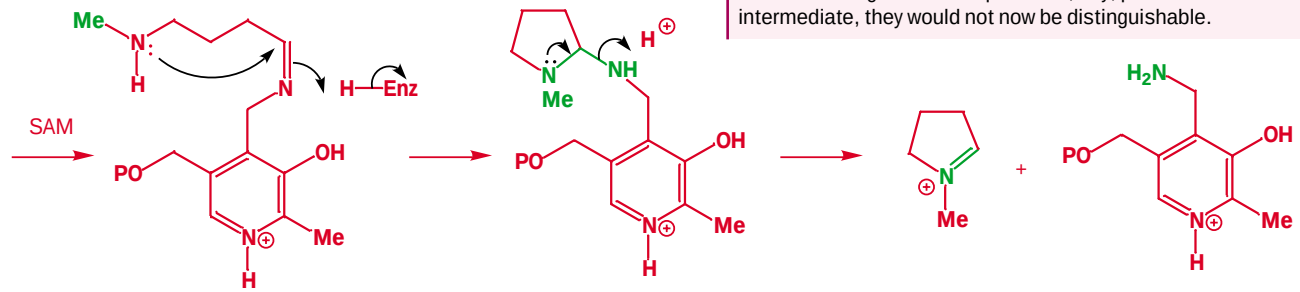
■ Both reagents SAM and acetyl CoA were discussed in Chapter 50. We will not be able to repeat at length the details of the chemistry of these and other common biochemical reagents already discussed there. In general, in this chapter we will give only the distinctive or interesting steps and leave you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

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

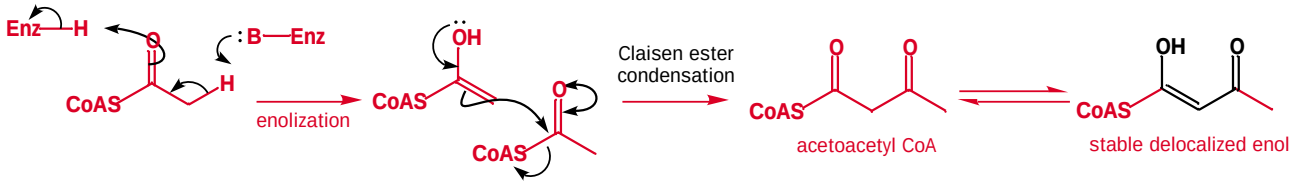



Now the terminal amino group is methylated by SAM and the secondary amine cyclizes on to the pyridoxal imine to give an aminal. Decomposition of the aminal the other way round expels pyridoxamine and releases the salt of an electrophilic imine.

▶ 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.

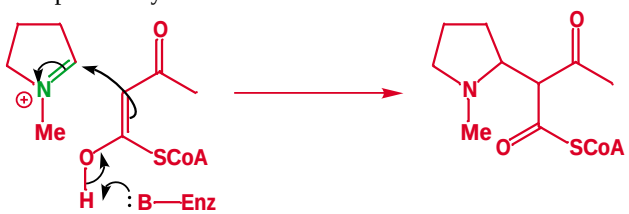



The rest of the biosynthesis does not need pyridoxal, but it does need two molecules of acetyl CoA. In Chapter 50 we noted that this thiol ester is a good electrophile and also enolizes easily. We need both reactivities now in a Claisen ester condensation of acetyl CoA.

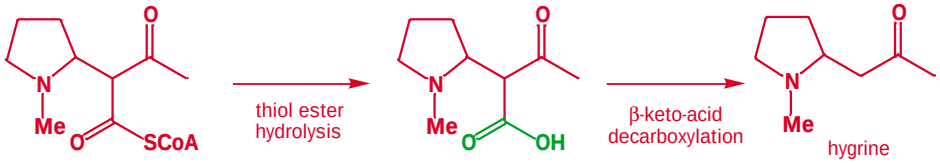



The new keto-ester is very like the acetoacetates we used in Chapter 27 to make stable enolates and the CoA thiol ester will exist mainly as its enol, stabilized by conjugation.

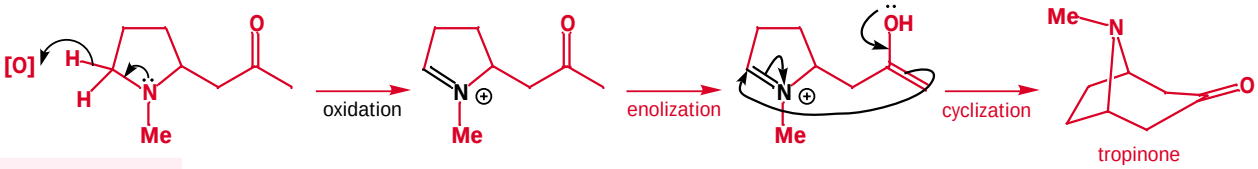
This enol reacts with the imine salt we have previously made and it will be easier to see this reaction if we redraw the enol in a different conformation. The imine salt does not have to wait around for acetoacetyl CoA to be made. The cell has a good stock of acetyl CoA and its condensation product.



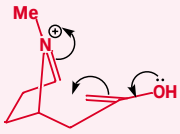
All that remains to form hygrine is the hydrolysis of the CoA thiol ester and decarboxylation of the keto-acid. This is standard chemistry, but you should ensure that you can draw the mechanisms for these steps.



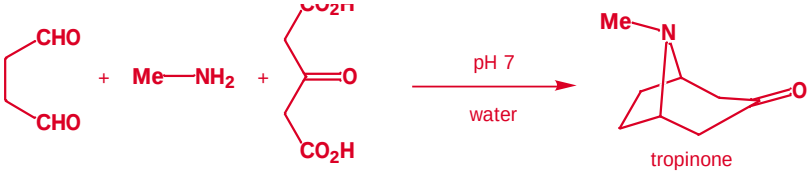

Tropinone is made from hygrine and it is clear what is needed. The methyl ketone must enolize and it must attack another imine salt resembling the first but on the other side of the ring. Such salts can be made chemically by oxidation with Hg(II) and biologically with an oxidizing enzyme and, say, $NAD^+$.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

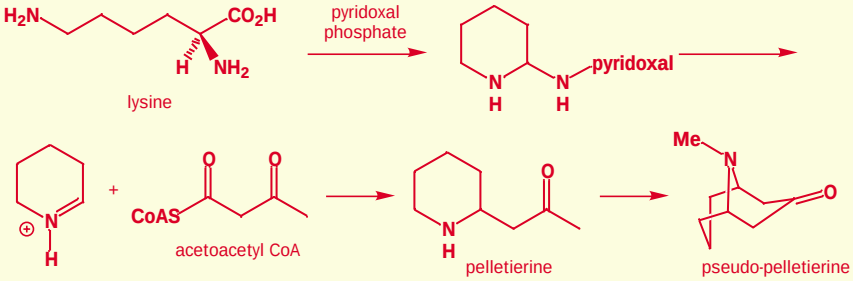


This complex route to tropinone was imitated as long ago as 1917 in one of the most celebrated reactions of all time, Robinson's tropinone synthesis. Robinson argued on purely chemical grounds that the sequence of imine salts and enols, which later (1970) turned out to be Nature's route, could be produced 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.



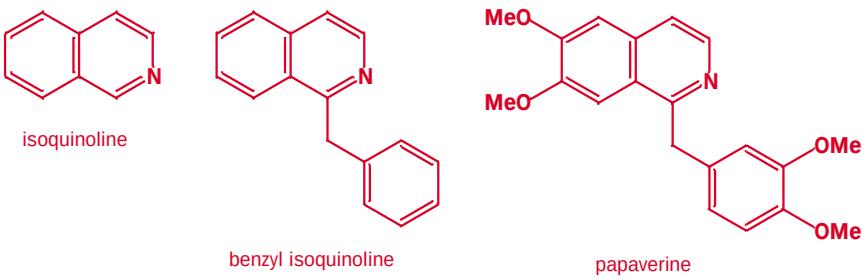

## Other pyrrolidine alkaloids

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.

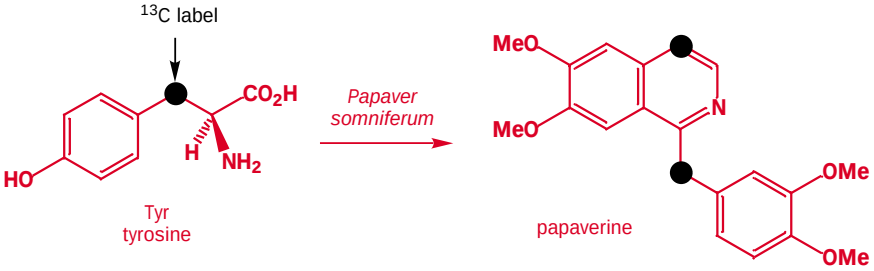



## Benzyl isoquinoline alkaloids are made from tyrosine

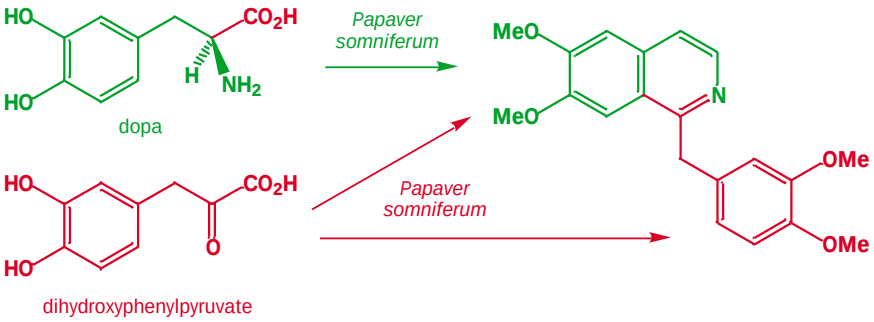
We switch to a completely different kind of alkaloid made from a different kind of amino acid. The benzyl isoquinoline alkaloids have a benzyl group attached to position 2 of an isoquinoline ring. Usually the alkaloids are oxygenated on the benzene ring and many are found in opium poppies (*Papaver somniferum*).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



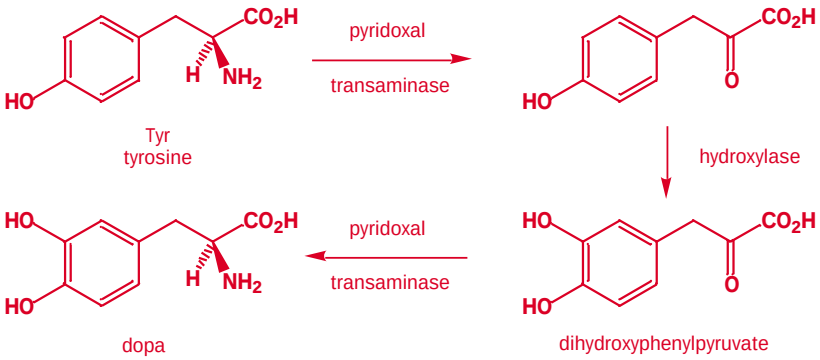

Labelling shows that these alkaloids come from two molecules of tyrosine. One must lose $CO_2$ and the other $NH_3$. We can easily see how to divide the molecule in half, but the details will have to wait a moment.




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.

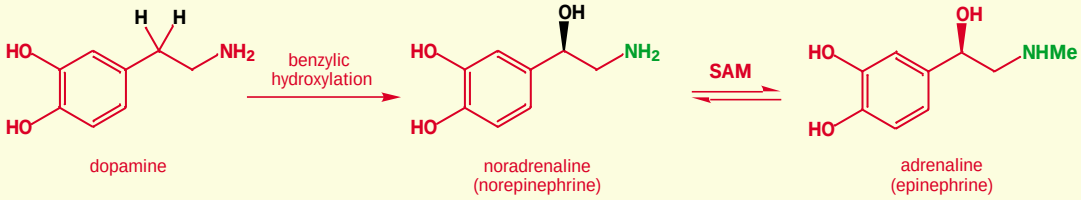



The amino acid and the keto-acid are, of course, related by a pyridoxal-mediated transaminase and the hydroxylation must occur right at the start. Both of these reactions are discussed in Chapter 50.

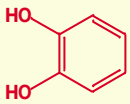


## Catecholamines

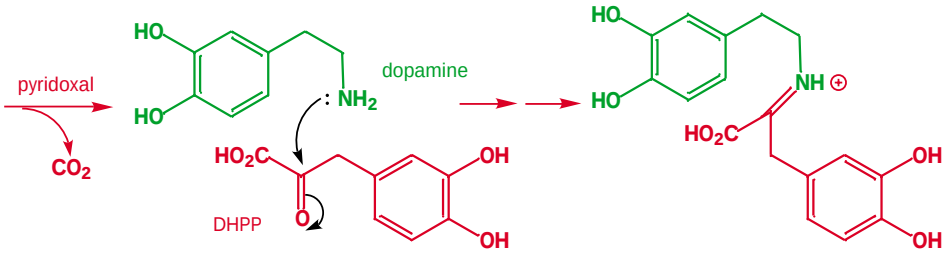
Dopa and dopamine are important compounds because they are the precursors to adrenaline in humans. Decarboxylation of dopa gives dopamine, which an oxidase (Chapter 50) hydroxylates stereospecifically at the benzylic position to give noradrenaline (norepinephrine).




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.

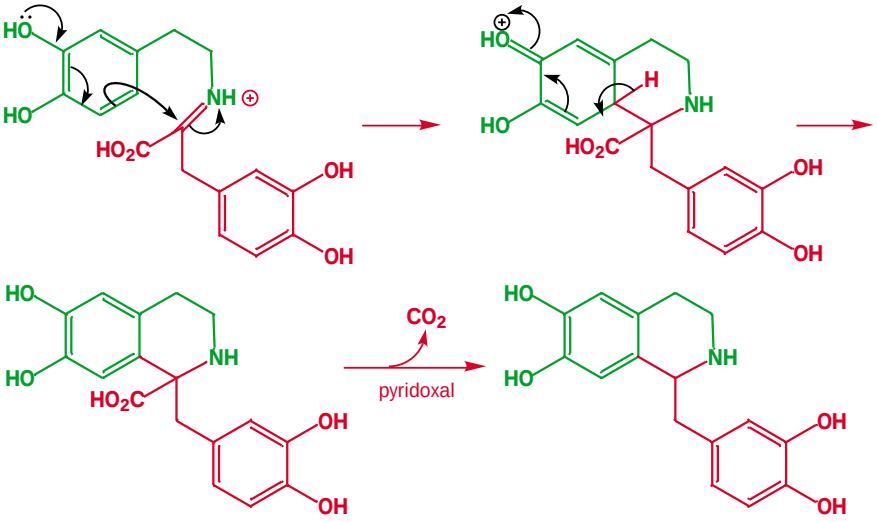



Pyridoxal-mediated decarboxylation of dopa gives dopamine and this reacts with the keto-acid to form an imine salt. This is an open-chain imine salt unlike the cyclic ones we saw in the pyrrolidine alkaloids, but it will prove to have similar reactivity.

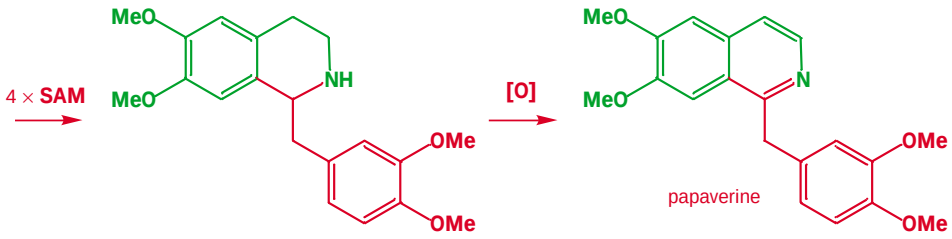



The imine salt is perfectly placed for an intramolecular electrophilic aromatic substitution by the electron-rich dihydroxyphenyl ring. This closes the isoquinoline ring in a Mannich-like process (Chapter 27) with the phenol replacing the enol in the pyrrolidine alkaloid biosynthesis.

■ 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)!

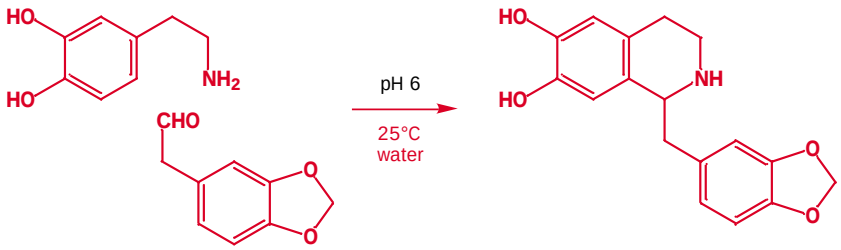



The cyclization product is still an amino acid and it can be decarboxylated by pyridoxal. Now we have something quite like papaverine but it lacks the methyl groups and the aromatic heterocyclic ring. Methylation needs SAM and is done in two stages for a reason we will discover soon. The final oxidation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



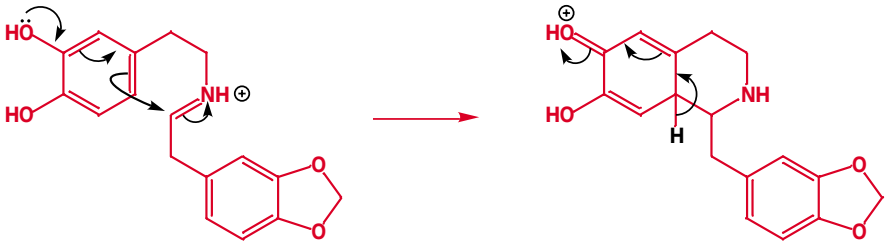

The reaction to make the isoquinoline ring can be carried out chemically under very mild conditions providing that we use an aldehyde as the carbonyl component. Then it works very well with rather similar compounds.

▶ 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.

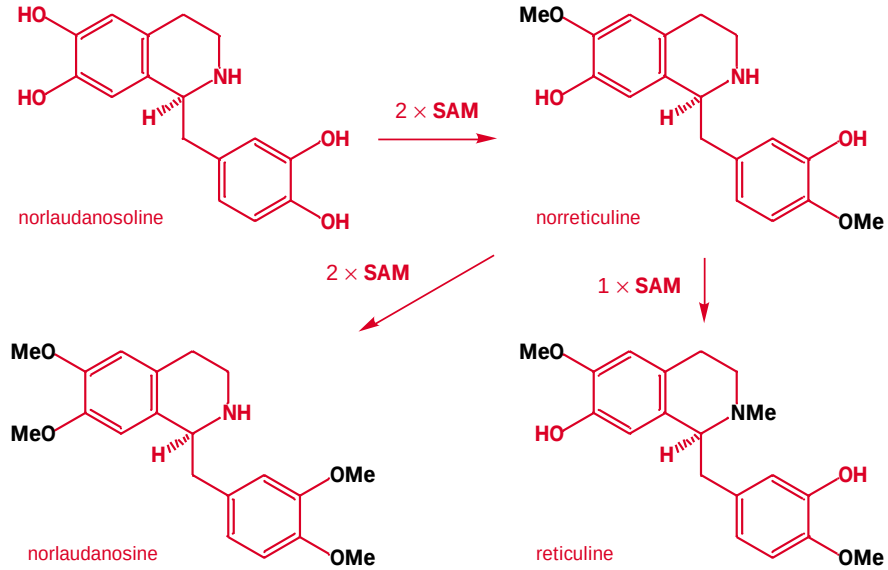



## Complex benzyl isoquinoline alkaloids are formed by radical coupling

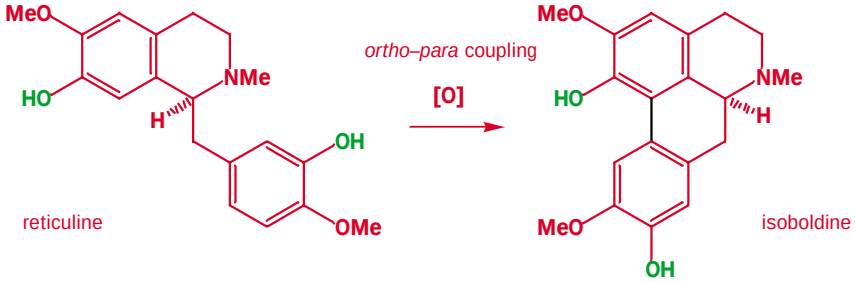
A more interesting series of alkaloids arises when benzyl isoquinoline alkaloids cyclize by radical reactions. Phenols easily form radicals when treated with oxidizing agents such as Fe(III), and benzyl isoquinoline alkaloids with free phenolic hydroxyl groups undergo radical reactions in an intramolecular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

■ See Chapter 39.

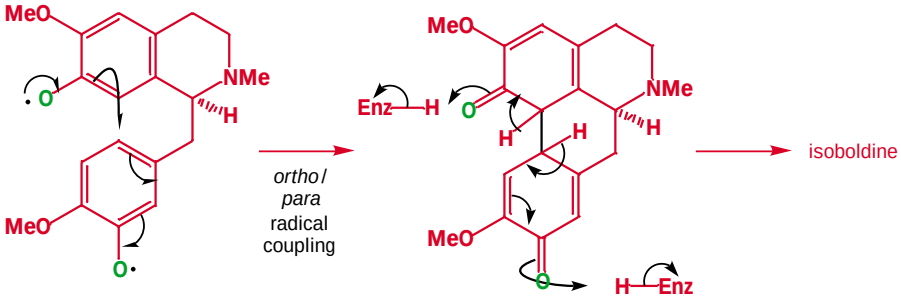
▶ The names of the alkaloids should not, of course, be learned, but they are a convenient handle for quick reference. The prefix 'nor' means without a methyl group, in this case the *N*-Me group, as you can see with norreticuline and reticuline (see p. 000).



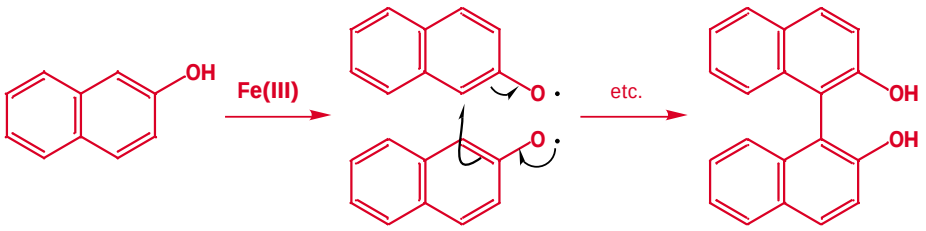
Methylating only one phenol on each ring of norreticuline leaves the other one free for radical coupling. Reticuline is oxidized in the plant to isoboldine by a radical cyclization with the formation of a new C–C bond.



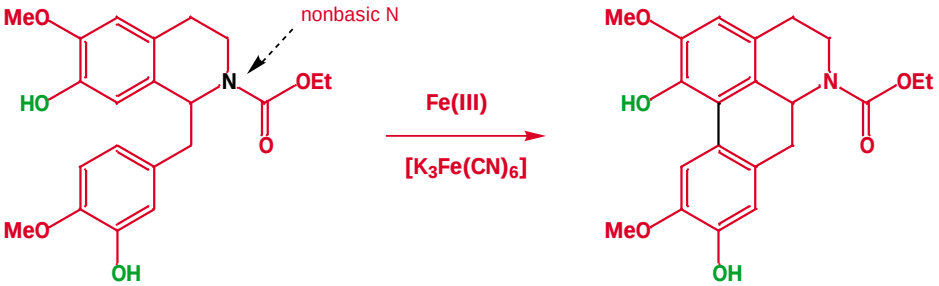

The new C–C bond is marked in black and the free phenolic OHs in green. Notice the relationship between them. The new bond is between a carbon atom *ortho* to one OH group and a carbon atom *para* to the other. We shall see in all these phenolic couplings that the *ortho* and *para* positions are the 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.



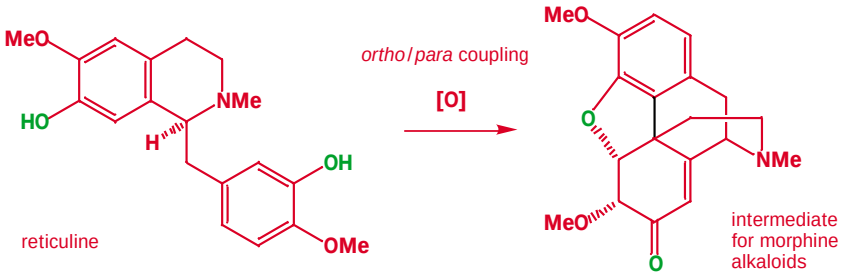

Phenol coupling occurs chemically under oxidation with Fe(III). The most famous example is the coupling of 2-naphthol to give binaphthol—an *ortho*/*ortho* coupling. The stereochemistry of binaphthyls like this was discussed in Chapter 45.



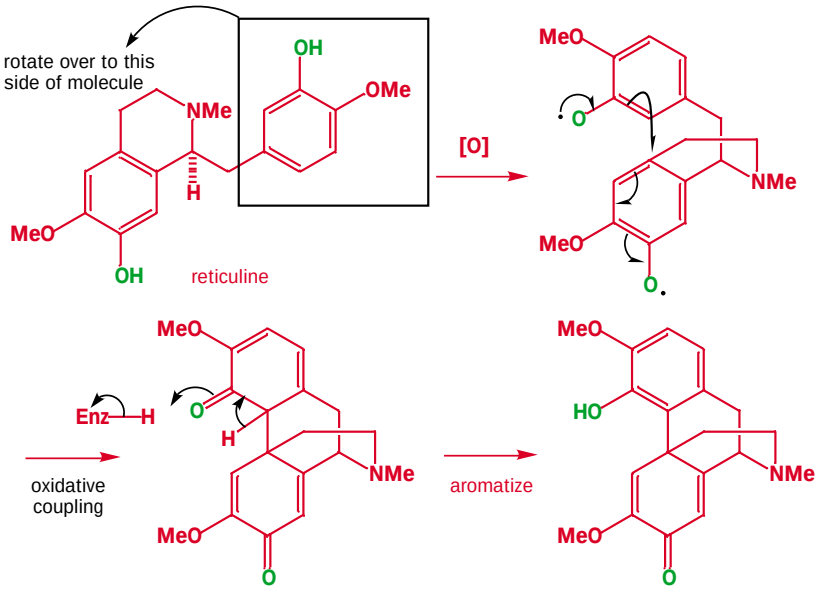

Similar phenol couplings have been attempted in the laboratory with compounds in the benzyl isoquinoline series but the nitrogen atom interferes if it is at all basic. When it has a carbonyl substituent the reactions do work reasonably well, but the yields are poor. Nature is still much better at this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



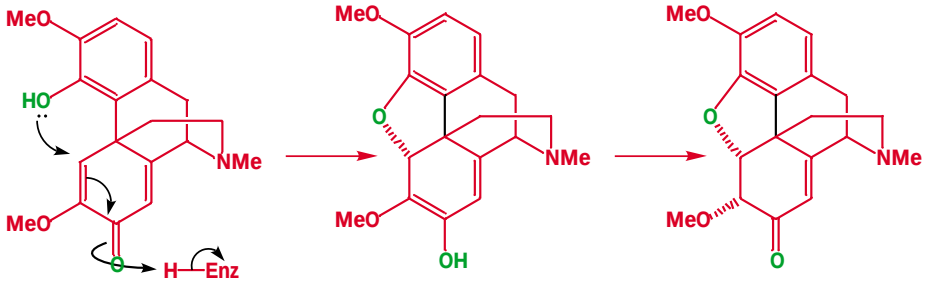

Reticuline is also the source of the morphine alkaloids by *ortho*/*para* radical coupling. The roles of the two rings are reversed this time and it is quite difficult to see at first how the structures are related.



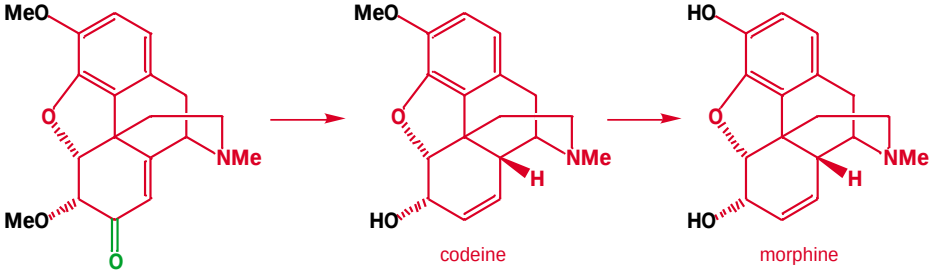

A great deal has happened in this reaction, but the new C–C bond (black) is *ortho* to the green oxygen atom in the top ring and *para* to the green oxygen atom in the bottom ring, so *ortho*/*para* coupling has occurred. To draw the reaction mechanism we need to draw reticuline in the right conformation.

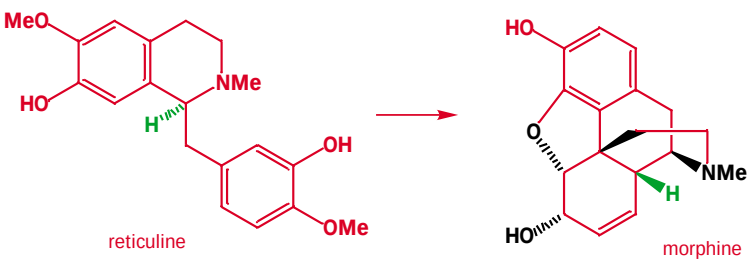



One of the two rings can re-aromatize but the other has a quaternary carbon atom so no proton can be lost from this site. Instead, the OH group in the top ring adds in conjugate fashion to the enone in the bottom ring.




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.

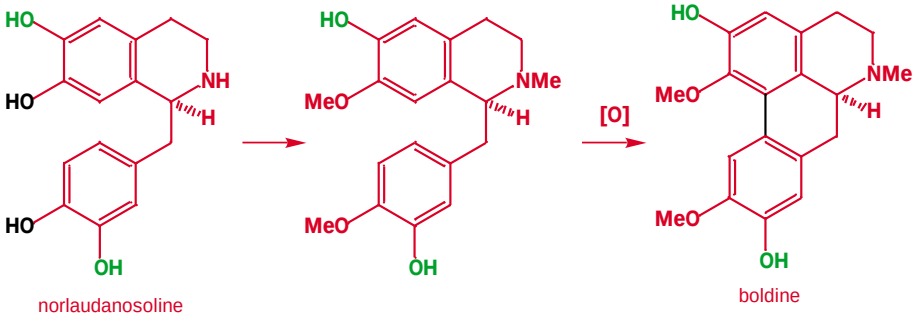




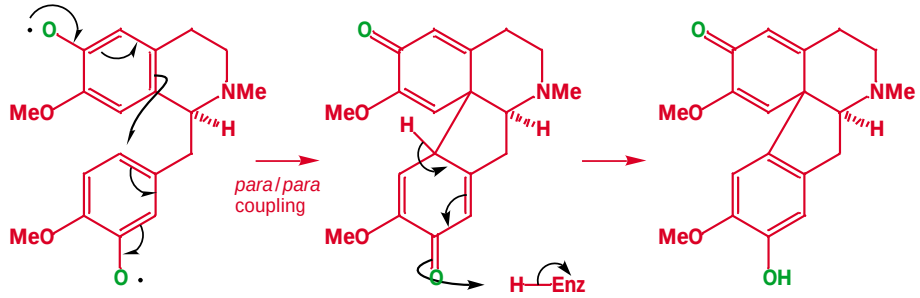

These alkaloids have plenty of stereochemistry. Indeed, if we compare the structures of reticuline and morphine, we can see that the one stereogenic centre in reticuline (marked in green) is still there in morphine (it hasn't been inverted—that part of the molecule has just been turned over) and that 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.

## Boldine, an isomer of isoboldine, is formed by rearrangement

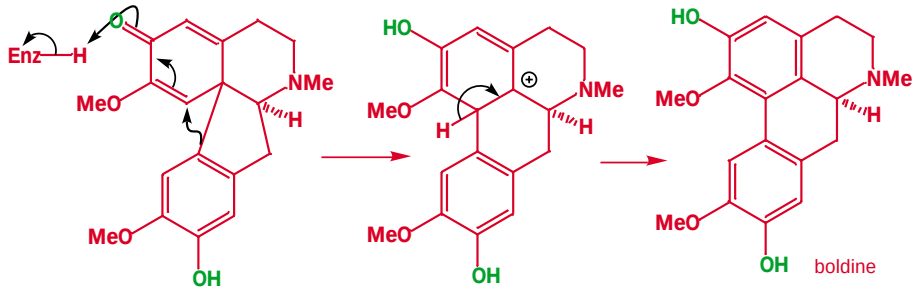
We mentioned isoboldine a while back, so there must be a boldine as well. This alkaloid is also formed from norlaudanosoline by a different methylation sequence and oxidative radical coupling. Looking at the structure of boldine you may see what appears to be a mistake on someone's part.




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.

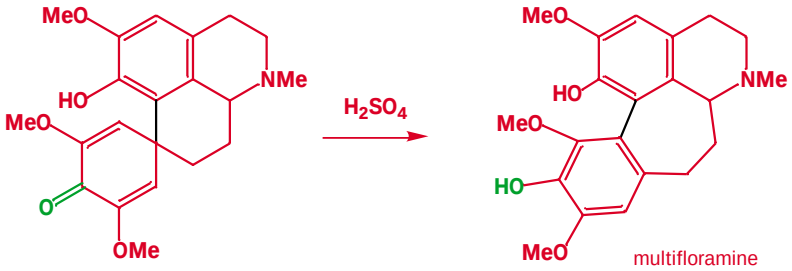



One of the rings has aromatized, but the other cannot—this should remind you of the morphine biosynthesis. However, there is no nucleophilic OH group here capable of conjugate addition to the enone so a rearrangement occurs instead. The new bond to the lower ring migrates across the top ring. You might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



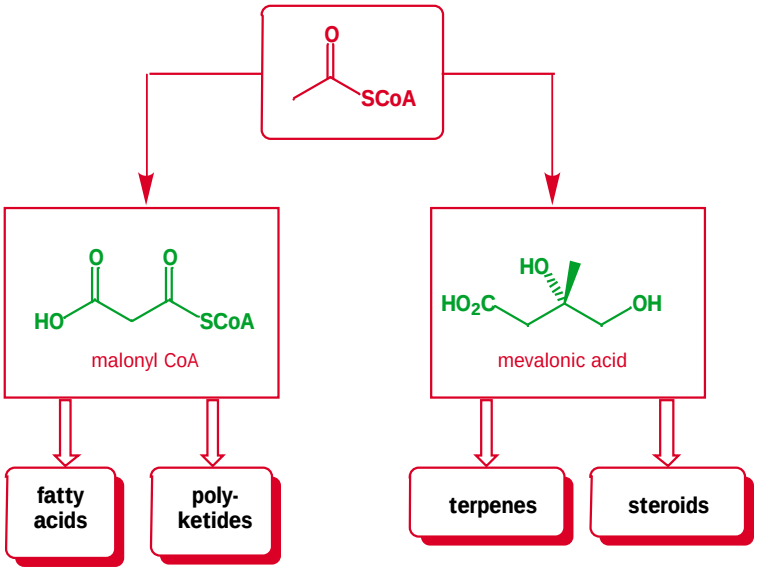
After the rearrangement there is a proton available to be lost and the cation can aromatize. The *para* relationship in the original coupling product has become a *meta* relationship by rearrangement. You should be able to recognize this rearrangement from Chapter 37: it is a dienone–phenol rearrangement.

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

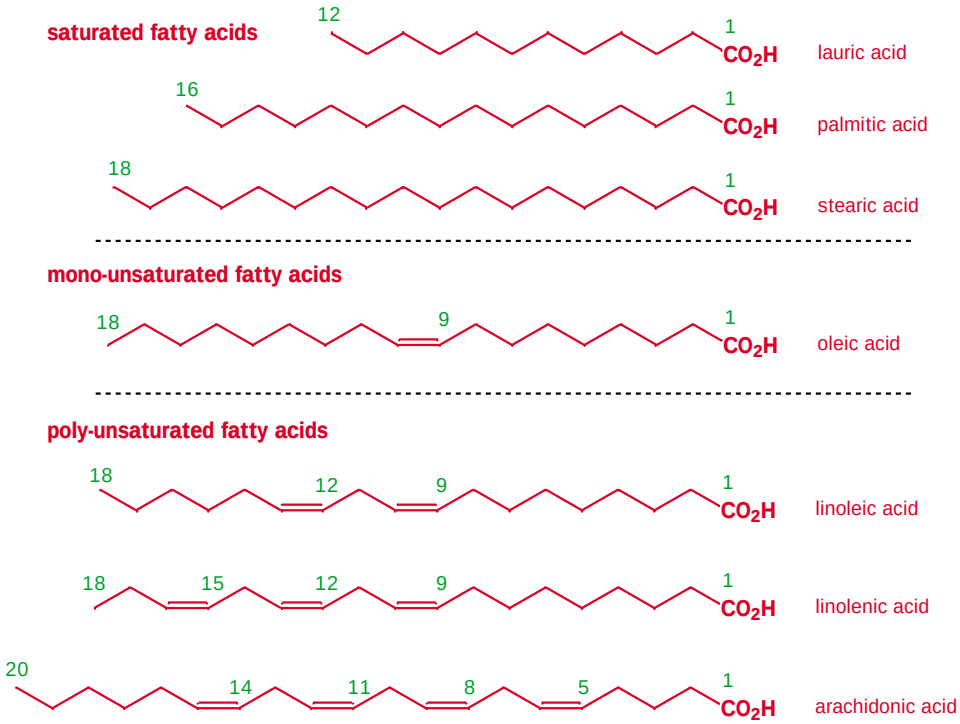
The sections that remain in this chapter show how Nature can take a very simple molecule—acetyl CoA—and build it up into an amazing variety of structures. There are two main pathways from acetyl CoA and each gives rise to two important series of natural products.




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.

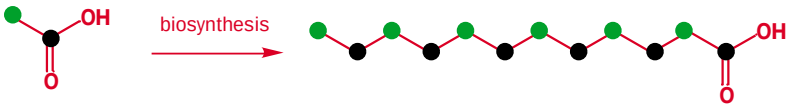
These are just a few of the fatty acids that exist, but all are present in our diet and you'll find many referred to on the labels of processed foods. You should notice a number of features.

- 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.



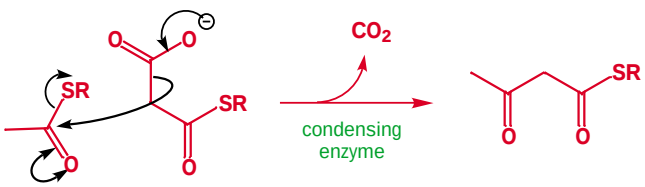

Palmitic acid ($C_{16}$ saturated) is the most common fatty acid in living things. Oleic acid ($C_{18}$ mono-unsaturated) is the major fatty acid in olive oil. Arachidonic acid ($C_{20}$ tetra-unsaturated) is a rare fatty acid, which is the precursor of the very important prostaglandins, thromboxanes,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

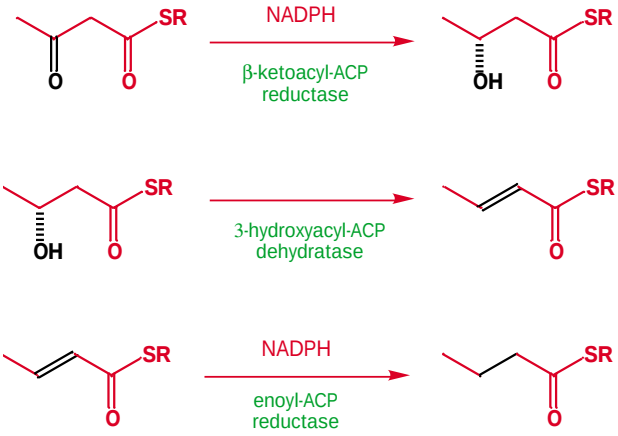


Notice that $CO_2$ is lost as the new C–C bond is formed. When chemists use malonates, we like to make the stable enol using both carbonyl groups, condense, and only afterwards release $CO_2$ (Chapter 26). Nature does this in making acetoacetyl CoA during alkaloid biosynthesis, but here she works differently.

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.

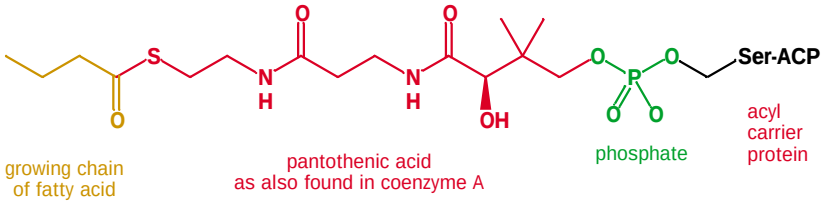




Now the whole cycle can start again using this newly made $C_4$ fatty acid as the starter unit and building a $C_6$ fatty acid and so on. Each time the cycle turns, two carbon atoms are added to the acyl end of the growing chain.

## 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.



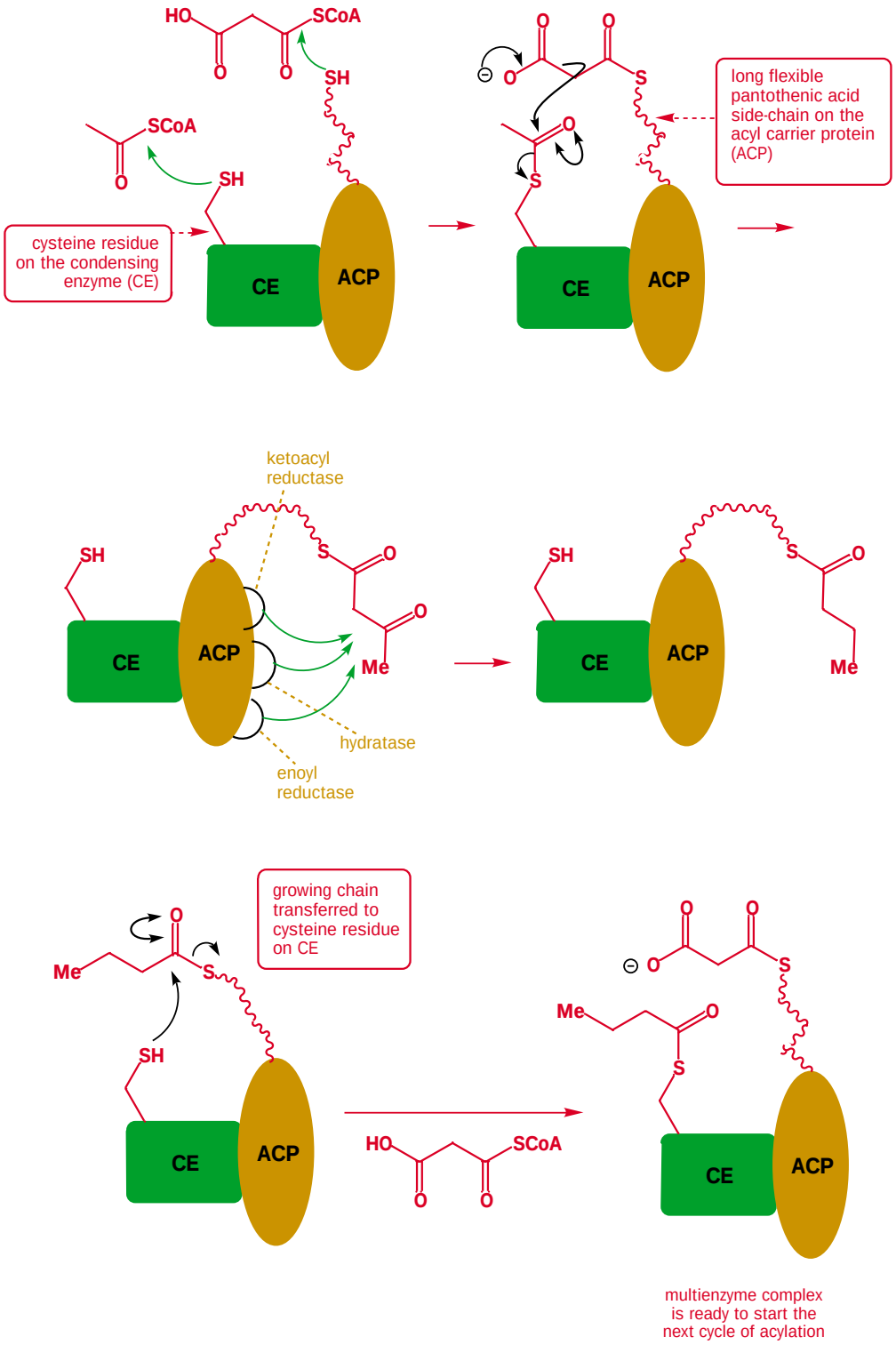

All of the enzymes needed for one cycle are clumped together to form two large proteins (ACP, the acyl carrier protein, and CE, the condensing enzyme) which associate in a stable dimer. The long side-chain passes the substrate from enzyme to enzyme so that synthesis can be continuous until the chain is finished and only then is the thiol ester hydrolysed. The chart on p. 000 illustrates this.

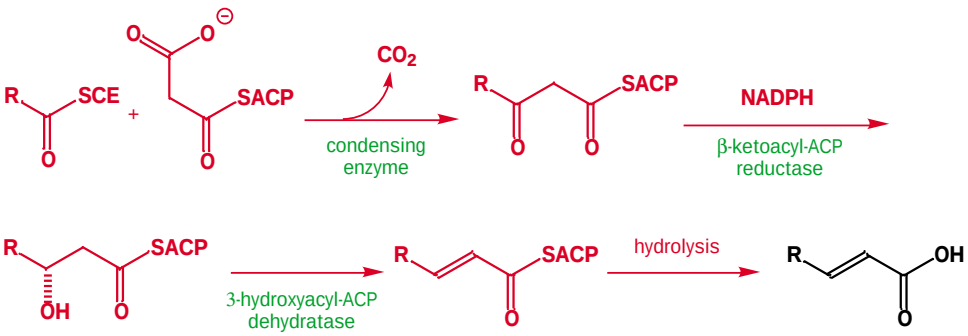
■ You saw a smaller multienzyme complex in Chapter 50 (p. 000), but this one is much more complex. More are being discovered all the time—Nature invented the production line well before Henry Ford.

## There are three ways of making unsaturated fatty acids

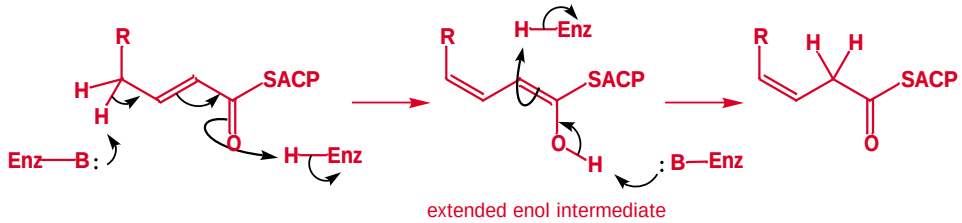
Conjugated unsaturated fatty acids are made simply by stopping the acylation cycle at that stage and hydrolysing the thiol ester linkage between the unsaturated acyl chain and ACP. They always have the *E* (*trans*) configuration and are the starting points for other biosynthetic pathways.

fatty acid biosynthesis: schematic diagram of the multienzyme dimer





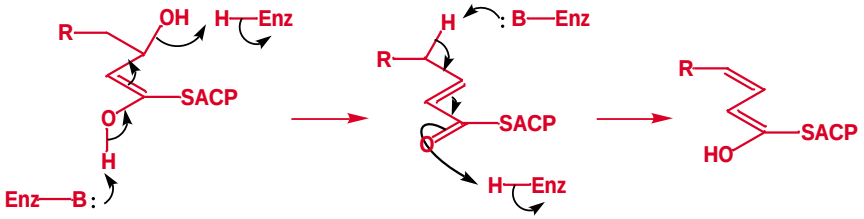

The second method makes *Z*-3,4-unsaturated acids by deconjugation from the *E*-2,3-unsaturated acids catalysed by an isomerase while the acyl chain is still attached to ACP. This is an anaerobic route as no oxidation is required (the double bond is already there—it just has to be moved) and is used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



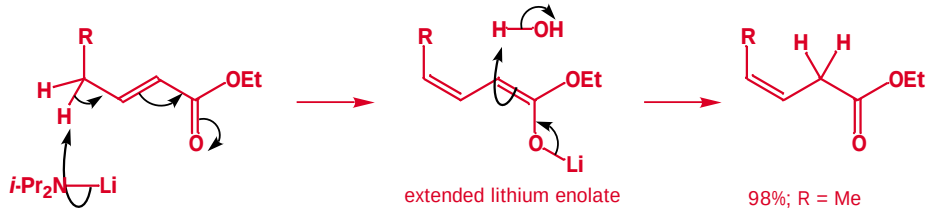

extended enol intermediate

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. 000.




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.

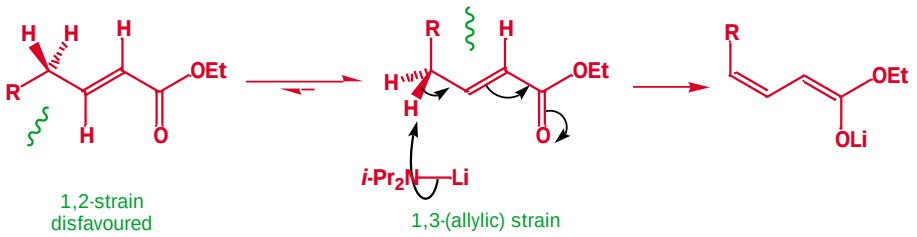



extended lithium enolate

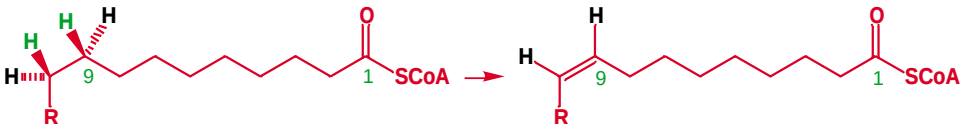
98%; R = Me

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. 000.



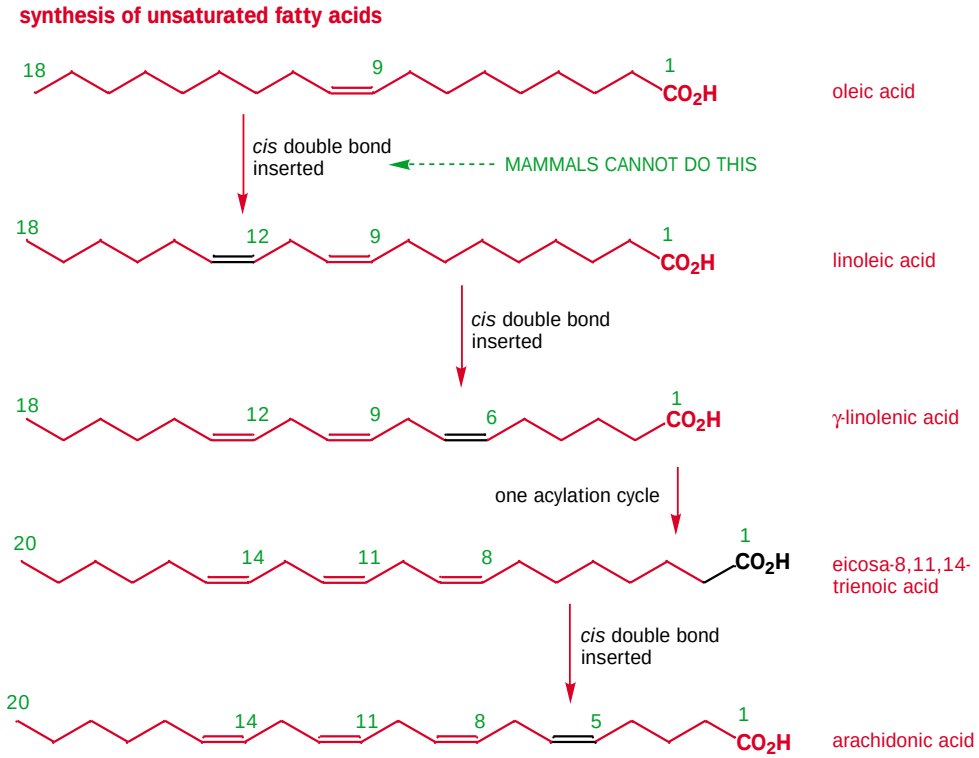

The third method is a concerted stereospecific removal of two adjacent hydrogen atoms from the chain of a fatty acid after synthesis. This is an aerobic route as oxidation is required and is used by mammals such as ourselves. The stereochemistry of the reaction is known from labelling studies to be *cis*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

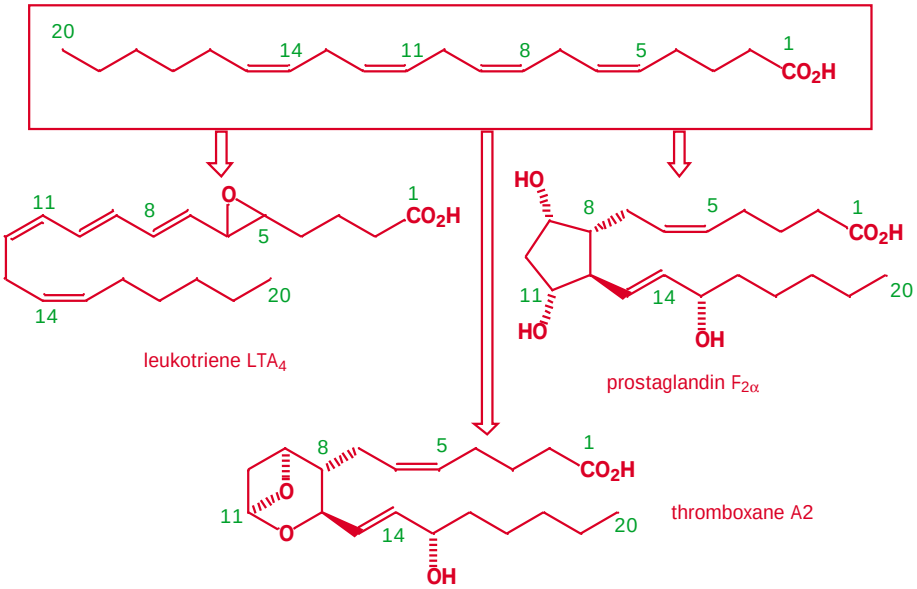
## What is so important about unsaturated fatty acids?

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.

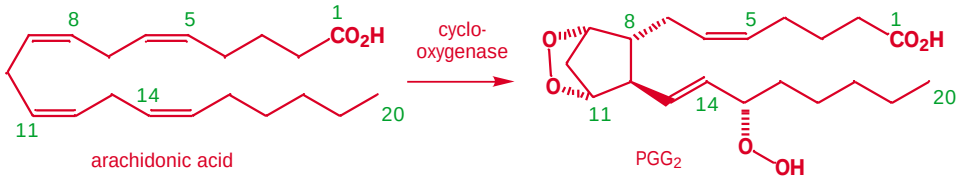


The final product of this chain of events—arachidonic acid—is one of the eicosanoids, so called because *eicosa* is Greek for 'twenty', and the systematic names for these compounds contain 'eicosanoic acid' in some form. The leukotrienes resemble arachidonic acid most closely, the prostaglandins have 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.

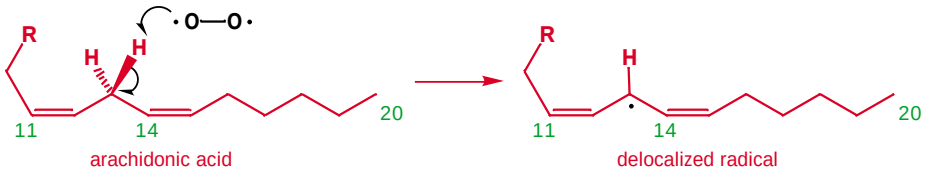
**compounds synthesized from arachidonic acid**



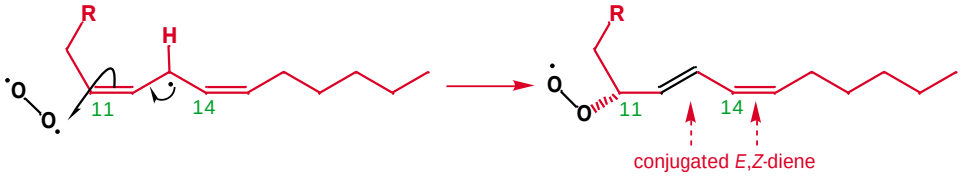

These compounds are all unstable and all are involved in transient events such as inflammation, blood clotting, fertilization, and immune responses. They are produced locally and decay quickly and are implicated in autoimmune diseases like asthma and arthritis. They are made by oxidation of arachidonic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



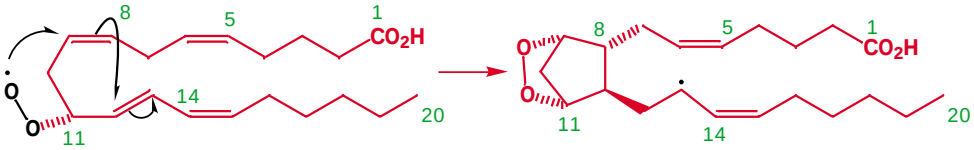

The first step is a radical abstraction of a hydrogen atom from an allylic position by oxygen (perhaps carried on an iron atom in a haem). The atom removed is between two alkenes so that the resulting radical is doubly allylic.



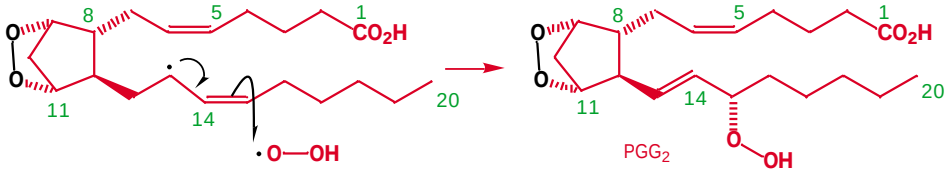

This allylic radical captures a molecule of oxygen at C11 to form a new oxyradical. The reaction occurs at one end of the delocalized radical so that the product is a conjugated diene and the new alkene is *trans* (*E*).



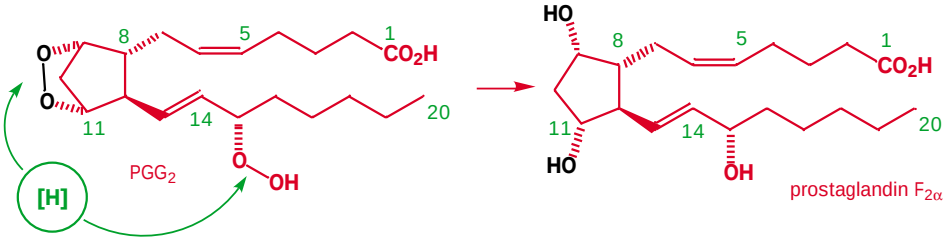

Now we need to resume the full structure of the intermediate because the oxyradical does an elaborate addition to the C8 alkene and then to the newly formed diene to form a new stable allylic radical.




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 $PGG_2$.



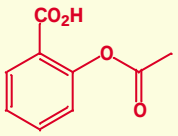

This unstable prostaglandin has been isolated from sheep but, as it has a half-life of only 5 minutes, this is no trivial matter. Both weak O–O bonds are now reduced enzymatically to give the first reasonably stable compound, $PGF_{2\alpha}$ (PG just means prostaglandin).




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.

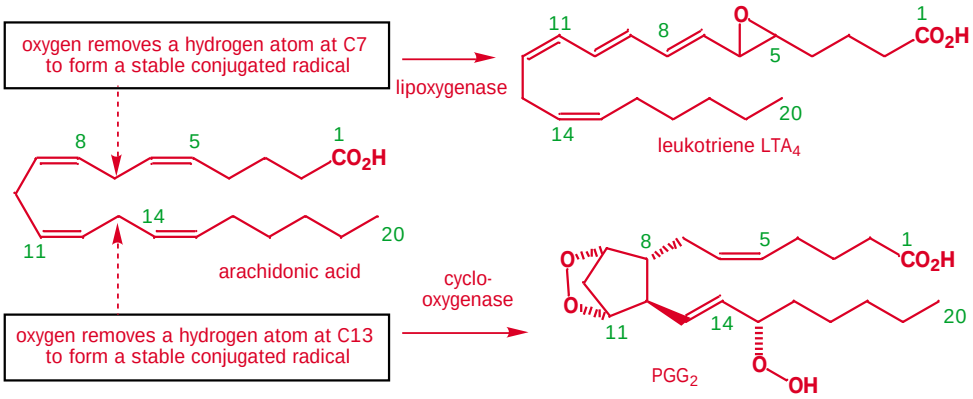
## How aspirin works

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.

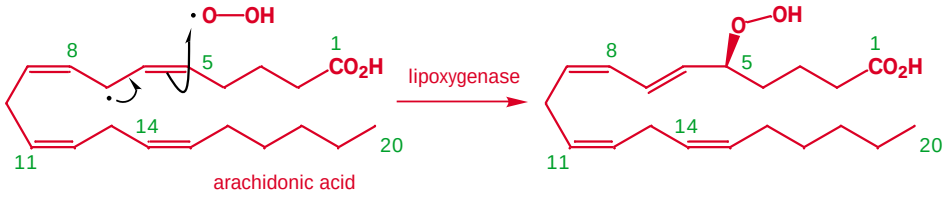


aspirin

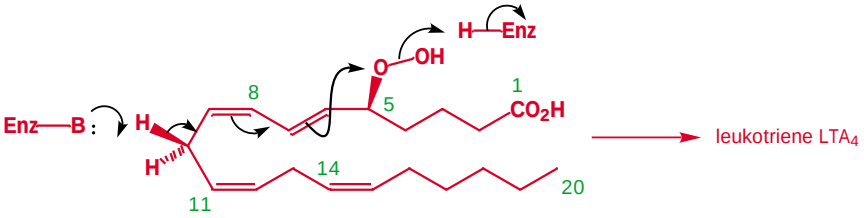
Each of the other families of eicosanoids—thromboxanes and leukotrienes—has interesting biosynthetic pathways too, but we will mention only one small detail. A completely different oxidation enzyme, lipoxygenase, initiates a separate pathway leading to the leukotrienes, but the first steps are very similar.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

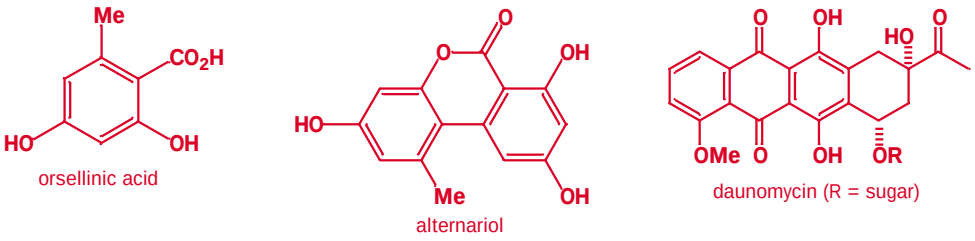



The relatively recent discovery of these unstable molecules of incredibly powerful biological activity means that we by no means know all about them yet. They are very important to our well-being and important medical advances are bound to follow from a better understanding.

■ Prostaglandins and leukotrienes have appeared several times before in this book, and you can read about aspects of their laboratory synthesis on pp. 000 and 000.

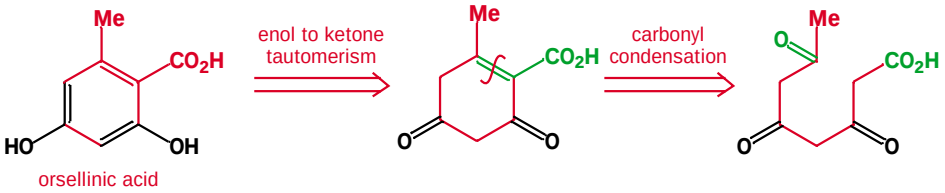
## Aromatic polyketides come in great variety

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.

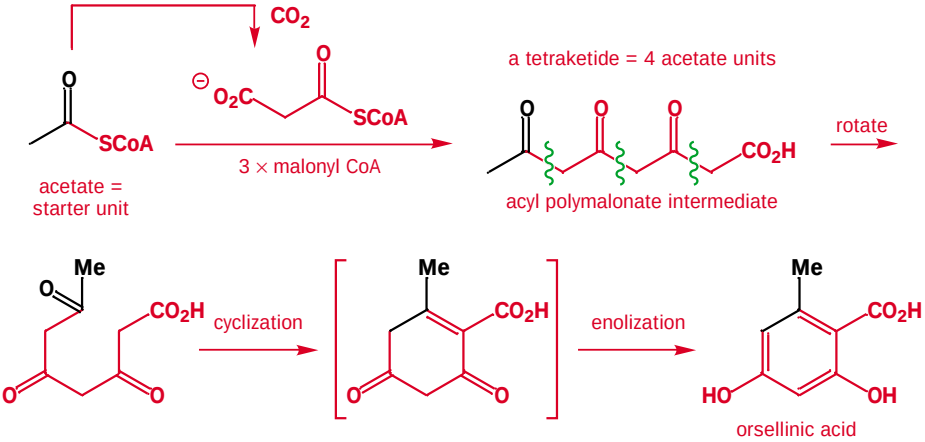



You might immediately be struck by the extent of oxygenation in these compounds. The shikimic acid route produced $Ar–C_3$ compounds with at most one OH group in the *para* position and others

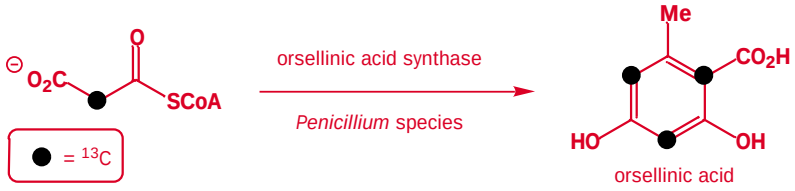
added *ortho* to that first OH group. Here we have multiple oxygenation with a predominant 1,3 pattern. If we try to arrange an acyl polymalonate product to make orsellinic acid, this is what we shall need.



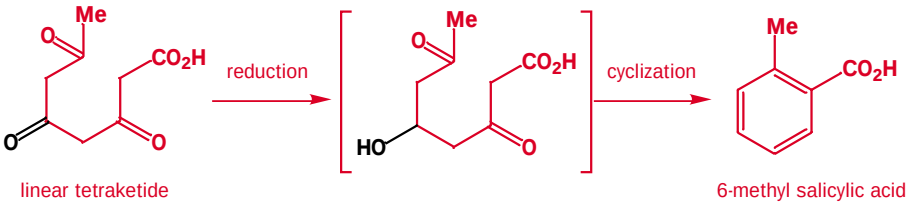

Merely by writing ketones instead of phenols and doing one disconnection corresponding to a simple carbonyl condensation, we have reached a possible starting material which is a typical acyl polymalonate product without any reductions. This is what polyketides are. The fatty acids are assembled with 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.




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.

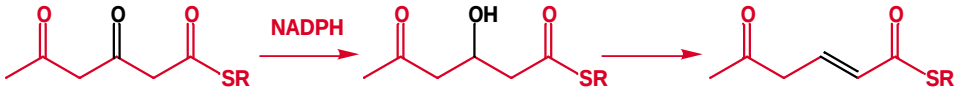



As the polyketide chain is built up, any of the reductions or eliminations from fatty acid biosynthesis can occur at any stage. The simple metabolite 6-methyl salicylic acid (6-MSA) is made in the microorganism *Penicillium patulum*, and it could come from the same intermediate as orsellinic acid with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



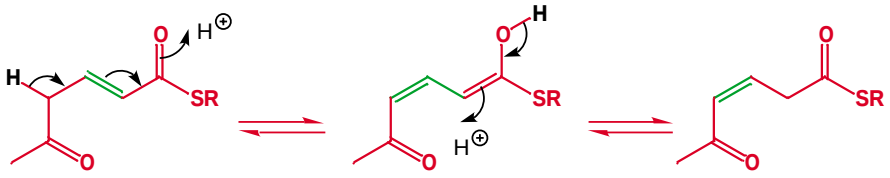

Reduction to the alcohol or to the unsaturated acid or ketone would give the right oxidation level and could occur as the chain is built, after it is completed, or after cyclization. In fact, reduction to

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.

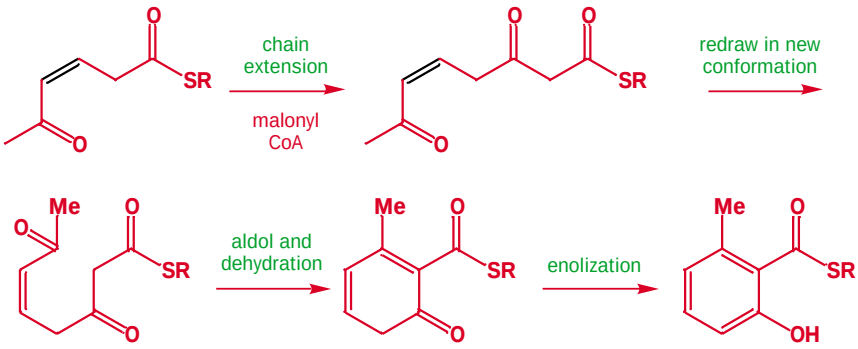



*E*-alkene conjugated with acid group

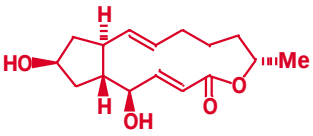
extended enol

*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.

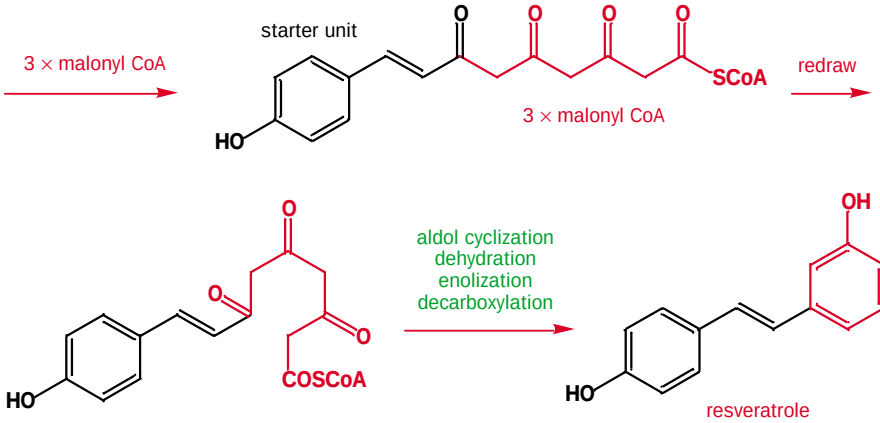
## Other starter units

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.

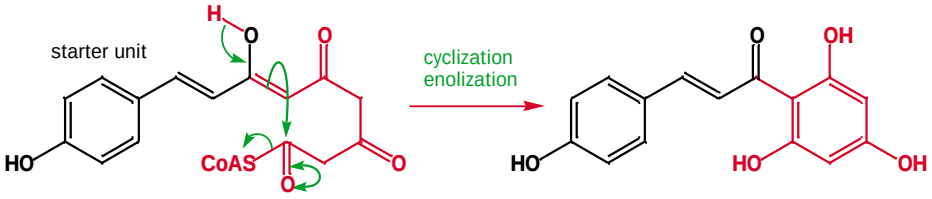

4-hydroxy cinnamic acid from shikimic acid

activated for polymalonate addition

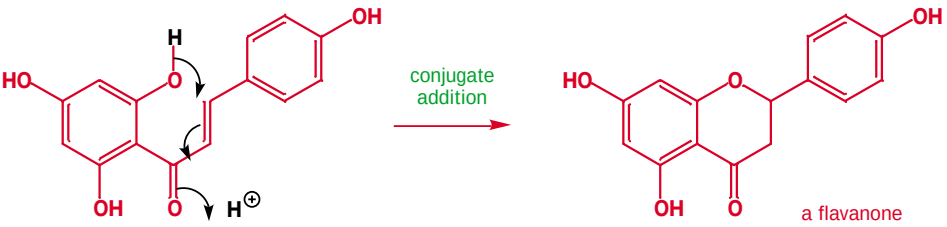
The most common sequence uses three malonyl CoA acylations followed by cyclization to a new aromatic ring. The simplest type is exemplified by resveratrole, the compound in red wine that helps to prevent heart disease. Each step in this sequence is a simple reaction that you have met before.




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.

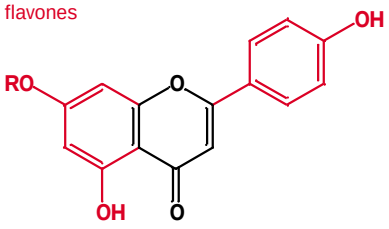



This cyclization is followed by a conjugate addition of an *ortho* phenolic OH group on to the enone system. The product is a flavanone structure, which is always drawn a different way up to the molecules we have just been discussing. Redrawing the last product shows the cyclization.



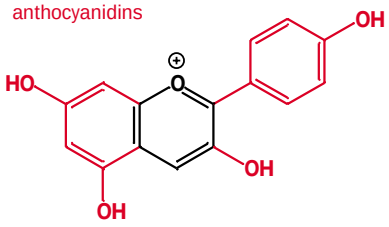

Aromatization of the central oxygen heterocycle by oxidation leads to the flavones, which are yellow or orange depending on their substituents. Dehydration leads to the red or blue anthocyanidins, pigments of flowers and fruit. This important group of molecules also includes plant growth hormones and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

flavones



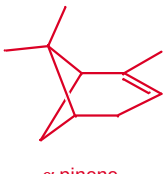
R = H; naringenin, R = glucose; naringin
—a bitter substance from grapefruit peel

anthocyanidins



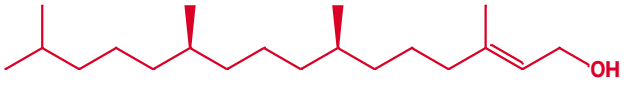
pelargonidin, pigment of raspberries,
geraniums, and red grape skins

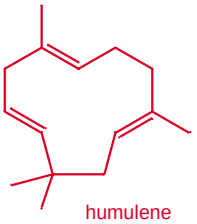
# Terpenes are volatile constituents of plant resins and essential oils



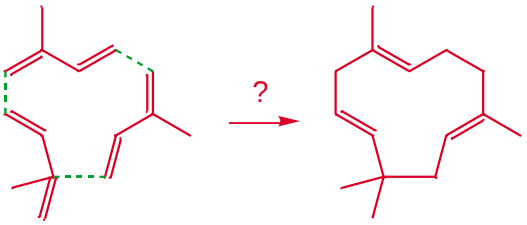


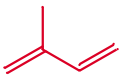
Terpenes were originally named after turpentine, the volatile oil from pine trees used in oil painting, whose major constituent is α-pinene. The term was rather vaguely used for all the volatile oily compounds, insoluble in water and usually with resiny smells from plants. The oils distilled from plants, 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.



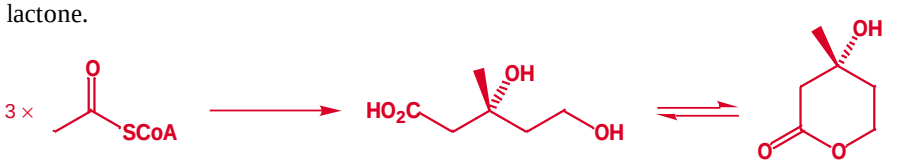




You will notice that they are all aliphatic compounds with a scattering of double bonds and rings, few functional groups, and an abundance of methyl groups. A better definition (that is, a biosynthetically based definition) arose when it was noticed that all these compounds have 5*n* carbon atoms. Pinene 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.

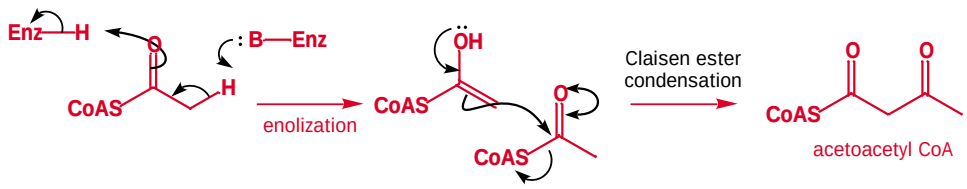



In fact, this is not correct. Isoprene is not an intermediate, and the discovery of the true pathway started when acetate was, rather surprisingly, found to be the original precursor for all terpenes. The key intermediate is mevalonic acid, formed from three acetate units and usually isolated as its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

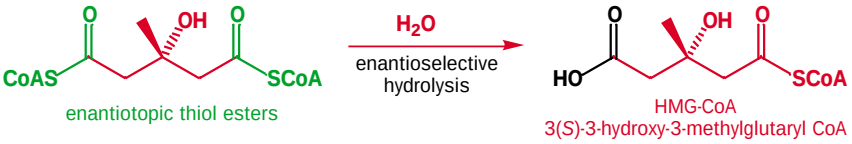



The first step is the Claisen ester condensation of two molecules of acetyl CoA, one acting as an enol and the other as an electrophilic acylating agent to give acetoacetyl CoA. We saw the same reaction in the biosynthesis of the pyrrolidine alkaloids earlier in this chapter.

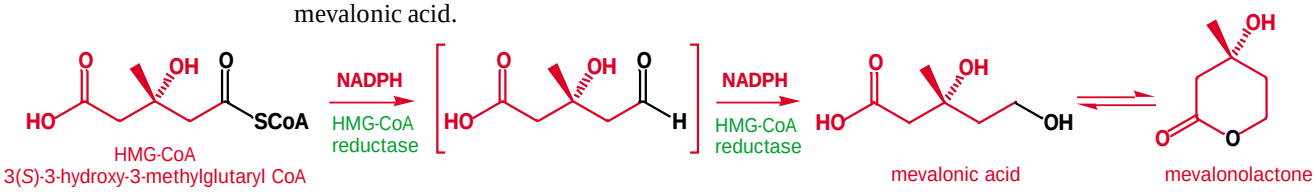



The third molecule of acetyl CoA also functions as a nucleophilic enol and attacks the keto group of acetoacetyl CoA. This is not a Claisen ester condensation—it is an aldol reaction between the enol of a thiol ester and an electrophilic ketone.

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.

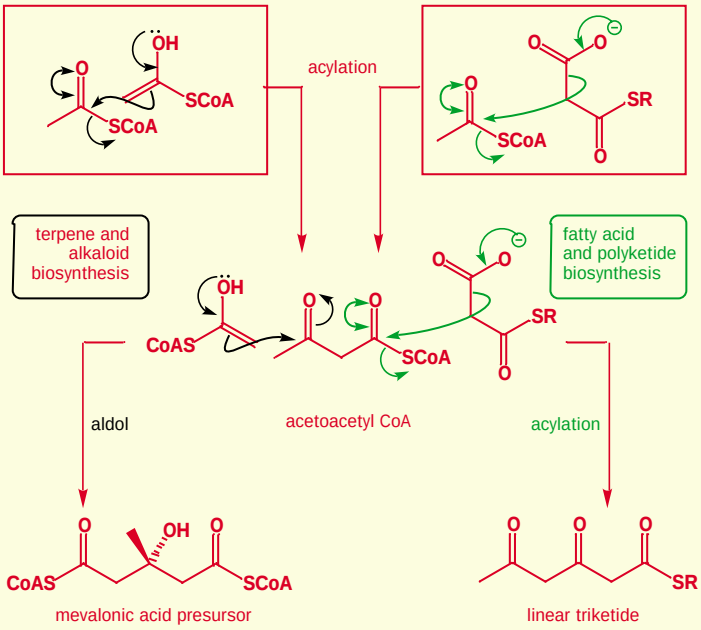



The remaining thiol ester is more electrophilic than the acid and can be reduced by the nucleophilic hydride from NADPH. Just as in $LiBH_4$ reductions of esters (Chapter 24), the reaction does not stop at the aldehyde level, and two molecules of NADPH are used to make the alcohol. This is mevalonic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



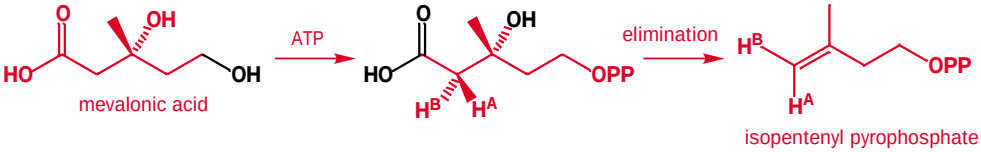

## Different pathways; different reactivity

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.

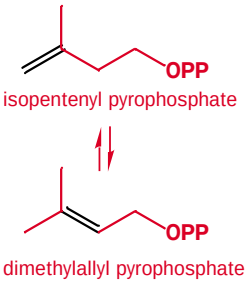



Mevalonic acid is indeed the true precursor of the terpenes but it is a $C_6$ compound and so it must lose a carbon atom to give the $C_5$ precursor. The spare carbon atom becomes $CO_2$ by an elimination reaction. First, the primary alcohol is pyrophosphorylated with ATP (Chapter 49); then the $CO_2H$ 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.

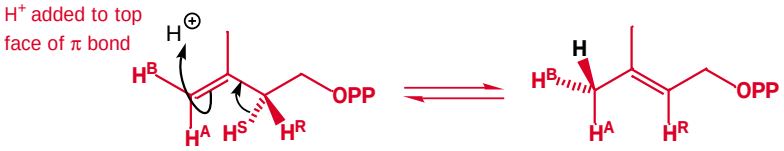
▶ 'PP' indicates the pyrophosphate group transferred from ATP.



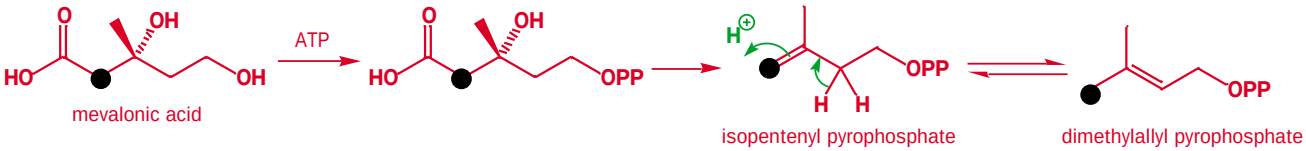
So is isopentenyl pyrophosphate the $C_5$ intermediate at last? Well, yes and no. There are actually two closely related $C_5$ intermediates, each of which has a specific and appropriate role in terpene biosynthesis. Isopentenyl pyrophosphate is in equilibrium with dimethylallyl pyrophosphate by a simple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



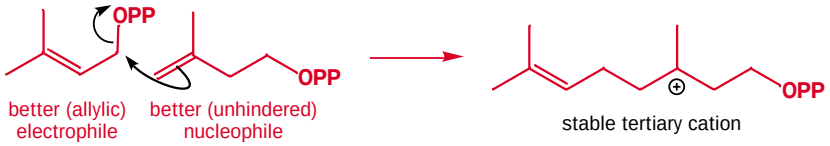

This is again a concerted reaction and again we know that by proton labelling. One of the two enantiotopic protons ($H^S$ in the diagram) is lost from the bottom face of the allylic $CH_2$ group while the new proton is added to the top face of the alkene. This is an *anti* rearrangement overall.



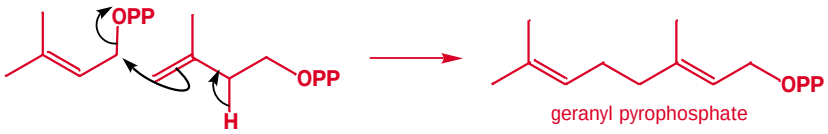

The stereochemical details are interesting in establishing the mechanism but not important to remember. What is important is that the origin of the two methyl groups in dimethylallyl pyrophosphate is quite distinct and can easily be traced if you always draw the intermediates in the way we have drawn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



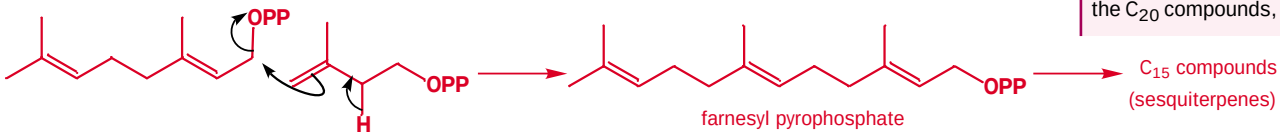

The two $C_5$ intermediates now react with each other. The dimethylallyl pyrophosphate is the better electrophile because it is allylic, and allylic compounds are good at both $S_N1$ and $S_N2$ reactions (Chapter 17). Isopentenyl pyrophosphate is the better nucleophile because it can react through an 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.



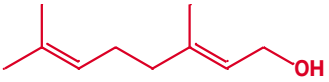

Geranyl pyrophosphate is the starting point for all the monoterpenes. It is still an allylic pyrophosphate and repeating the alkylation with another molecule of isopentenyl pyrophosphate gives farnesyl pyrophosphate, the starting point for the sesquiterpenes, and so on.



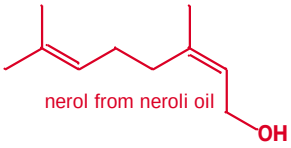

As soon as we start to make typical cyclic monoterpenes from geranyl pyrophosphate we run into a snag. We cannot cyclize geranyl pyrophosphate because it has a *trans* double bond! We *could* cyclize the *cis* compound (neryl pyrophosphate), and it used to be thought that this was formed from the *trans*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

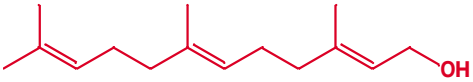
many of these names are derived from fragrant plant oils:



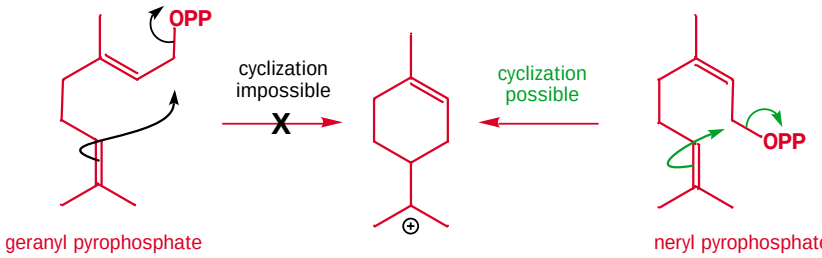

geraniol from geraniums




nerol from neroli oil

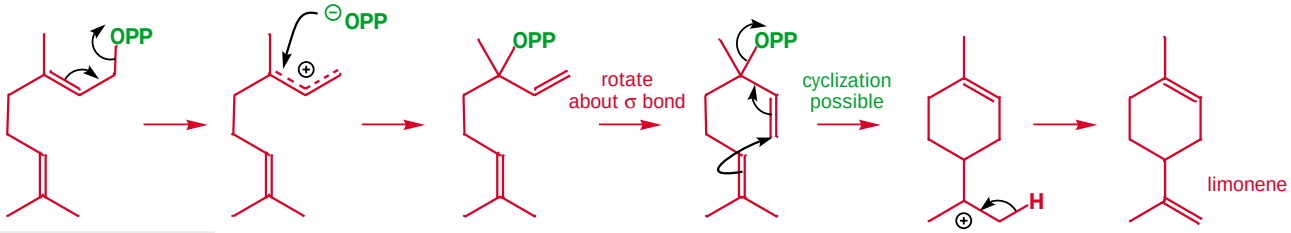



farnesol, also present in neroli oil



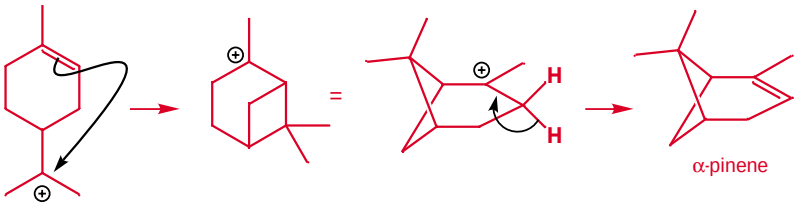

geranyl pyrophosphate — neryl pyrophosphate

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.

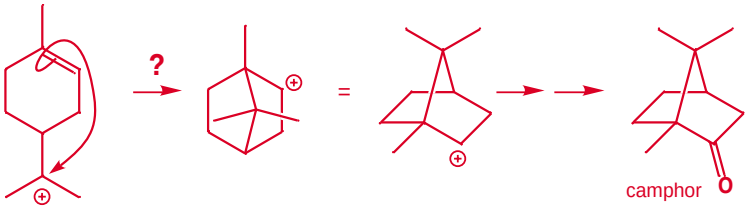



■ The product here is limonene—a terpene of the peel of citrus fruits. One enantiomer occurs in lemon peel—the other in orange peel. See Chapter 45.

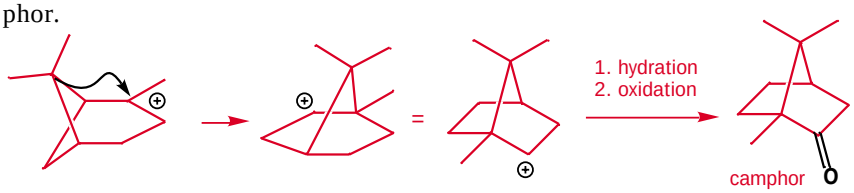
More interesting compounds come from the cyclization of the first formed cation. The remaining alkene can attack the cation to form what looks at first to be a very unstable compound but which is actually a tertiary carbocation with the pinene skeleton.



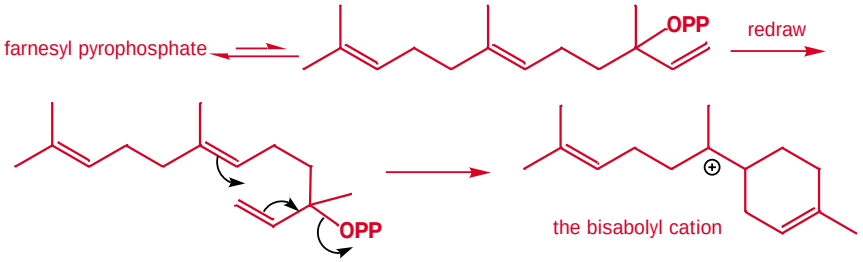

The camphor skeleton looks as though it might be formed by cyclization of the wrong end of the alkene on to the cation. This would certainly give the right skeleton but the intermediate secondary cation is rather unlikely.



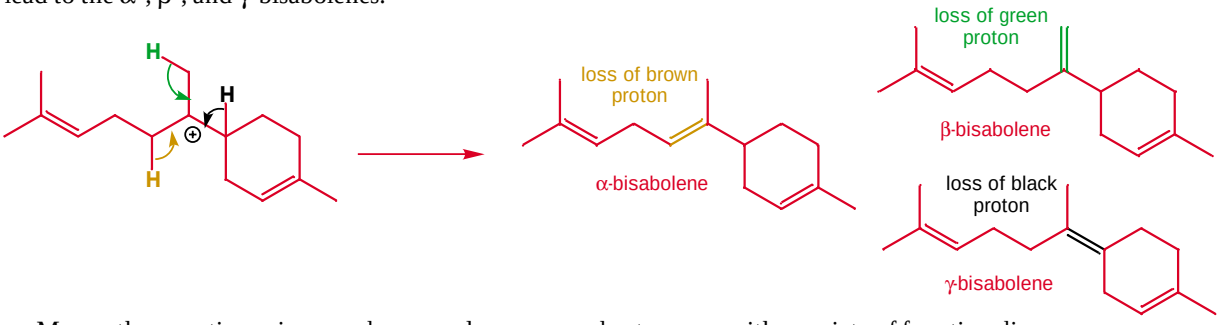

There is a better route. The more likely cation formed on the way to pinene could rearrange to the camphor cation. This is a known chemical reaction and is a simple 1,2-shift of the kind discussed in Chapter 37. However the new cation is formed, addition of water and oxidation would give camphor.



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.



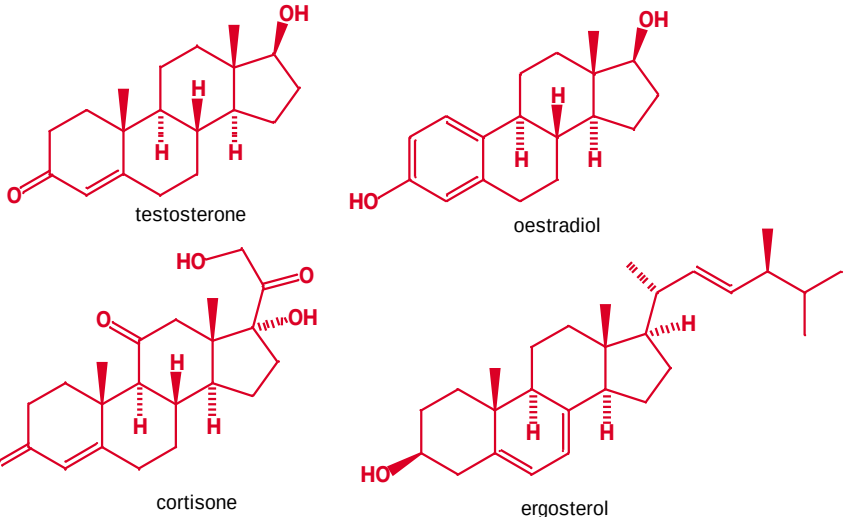

This cation does many things but it takes its name from the three fairly random proton losses that lead to the α-, β-, and γ-bisabolenes.



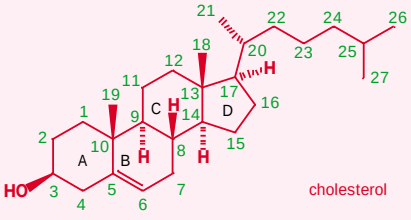

Many other reactions give even larger and more complex terpenes with a variety of functionalization but we will treat only one group in detail. These compounds are so important to us that they are given a different name.

## 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.

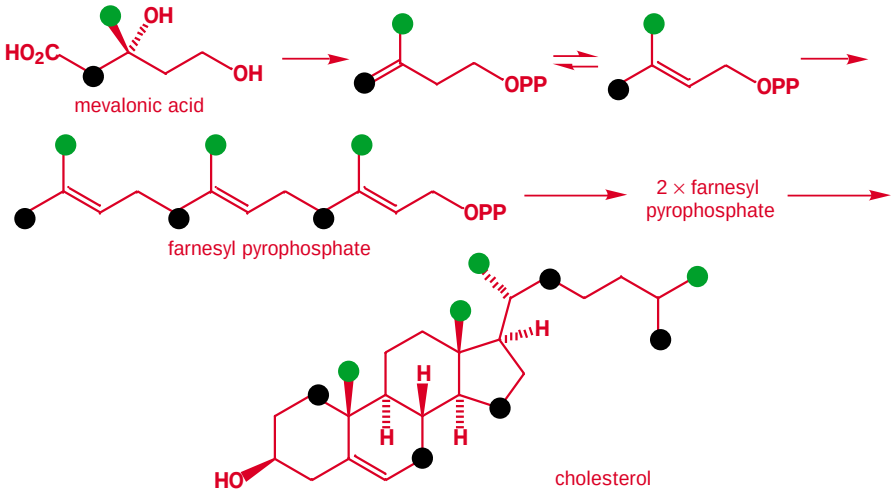



▶ For reference, here is the numbering of the steroid nucleus, not because we want you to learn it, but because it is often used without explanation in books and it is not obvious.

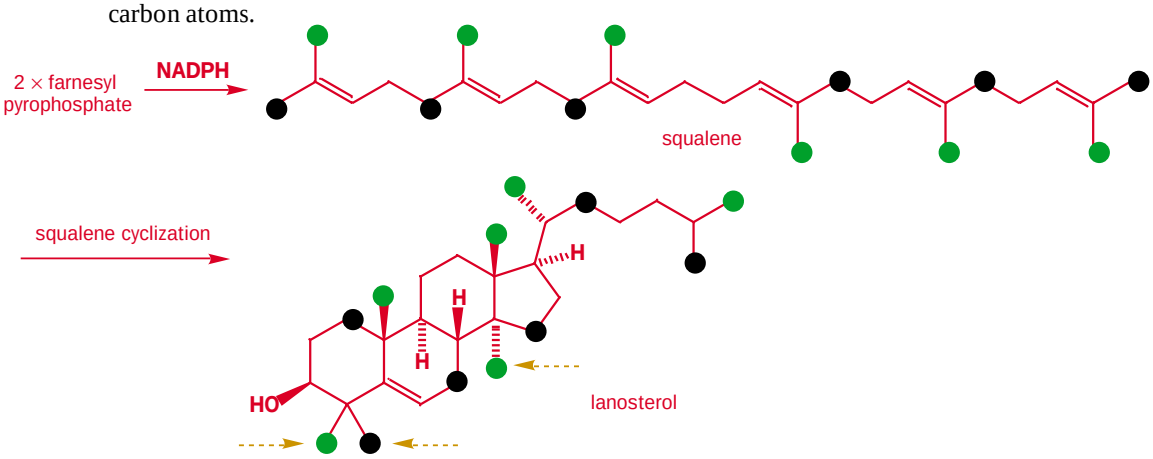



All share the skeleton of four fused rings, three six-membered and one five-membered and conventionally lettered A–D. Beyond the ring stereochemistry and some common oxygenation patterns they share little else. Some (such as the female sex hormones) have an aromatic A ring; some have side-chains on the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



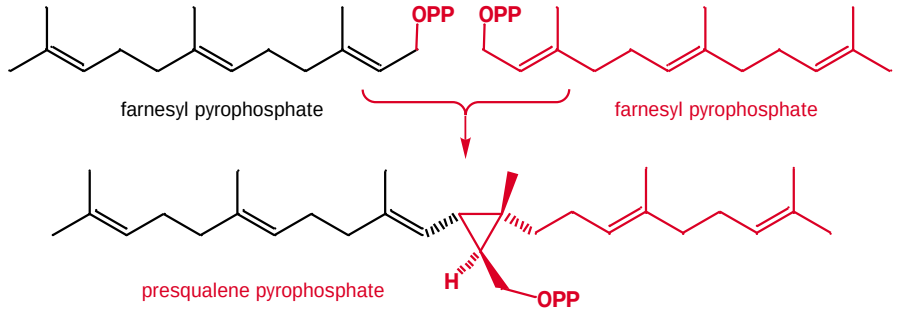

The three carbon atoms that are lost from lanosterol ($C_{30}$) in its conversion to cholesterol ($C_{27}$) are marked with brown arrows. Now at least we know which carbon atoms are lost. But many questions remain to be answered.

- 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?

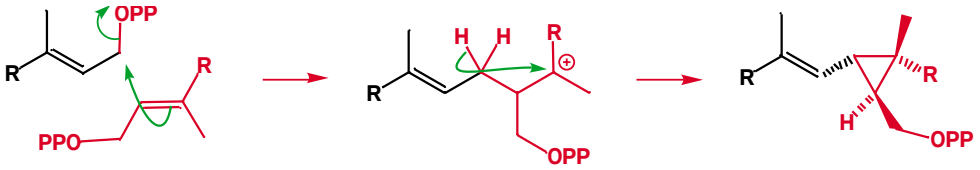
Before we tell you the answers, be warned: prepare for some surprises, and be ready to hold back outright disbelief!

## 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.



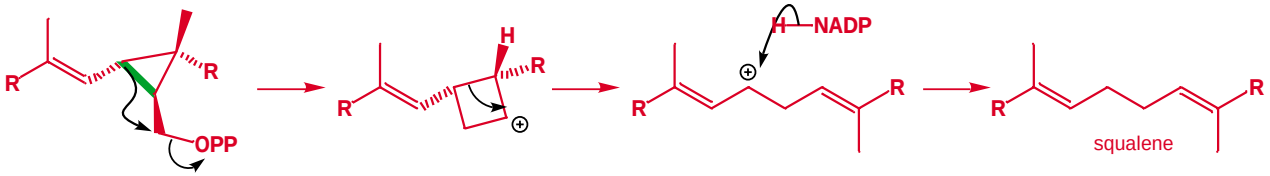

Maybe it's not so obvious that this is more rational! The first C–C bond formation is quite straightforward. The alkene in the red molecule attacks the allylic pyrophosphate in the black molecule in a simple $S_N2$ reaction. The product is a stable carbocation. Only one C–C bond remains to be formed 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.

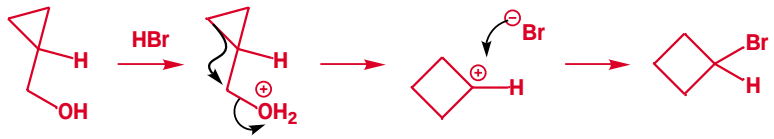
This is a very remarkable reaction. Such reactions do not occur chemically: this biological one occurs only because the molecule is held in the right shape by the enzyme and because the new ring is three-membered. Three-membered rings are very easily formed but also very easily opened—and that is what 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.



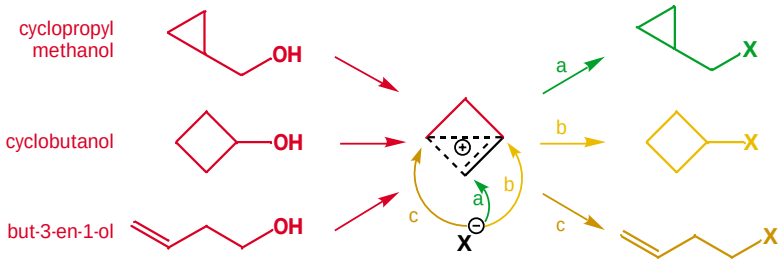

If you follow this sequence backwards, you will see that the originally formed 'rational' bond (shown in green) is the one that migrated and is retained in squalene, while the second bond is cleaved in the last step.

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.



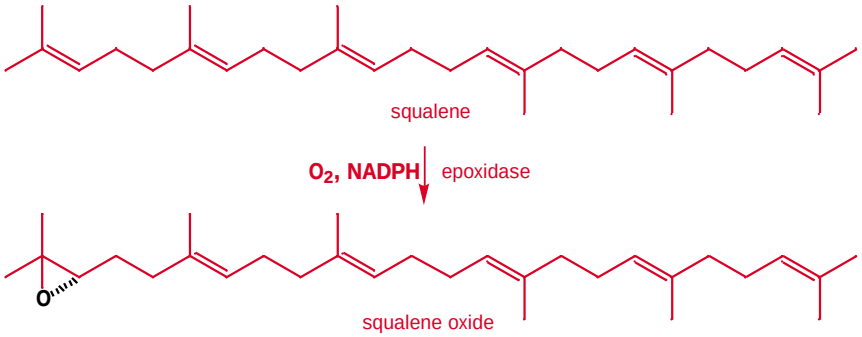

In fact, cyclopropylmethyl compounds, cyclobutyl compounds, and homoallyl compounds are all in equilibrium in acid solution and mixtures of products are often formed. The delocalized cation

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.

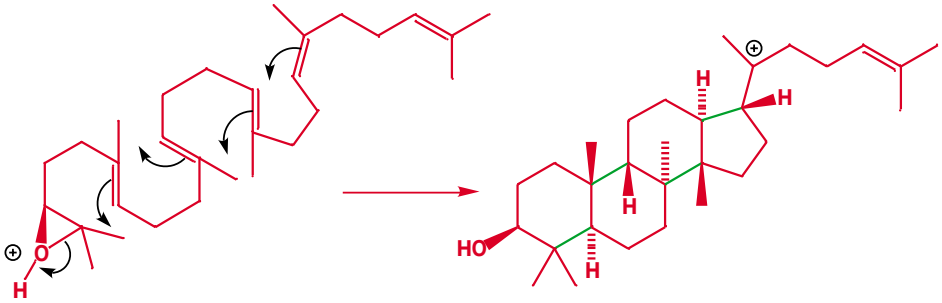



## Squalene to lanosterol

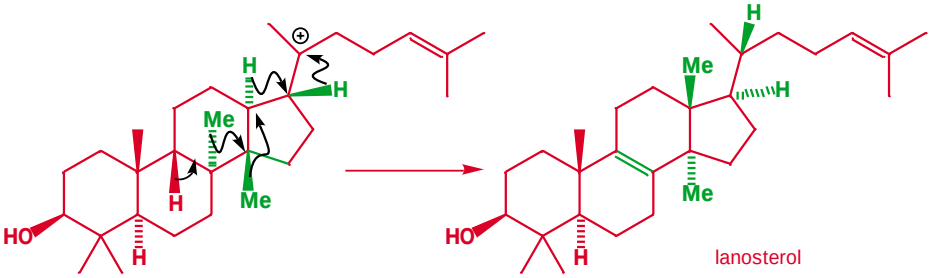
The next step is simple—the epoxidation of one of the terminal double bonds—but it leads to two of the most remarkable reactions in all of biological chemistry. Squalene is not chiral, but enzymatic epoxidation of one of the enantiotopic alkenes gives a single enantiomer of the epoxide with just one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

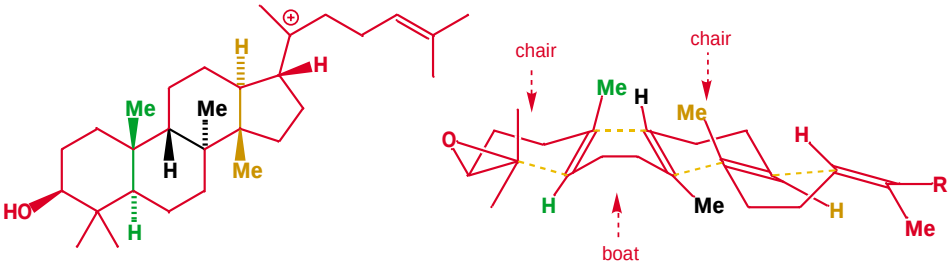



By analogy with what has gone before, you might now expect a tame hydration or reduction of this cation. Nothing of the sort! A rearrangement occurs in which *five* consecutive 1,2-shifts are followed by an elimination. Since this reaction organizes the backbone of the steroids, it is often called the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



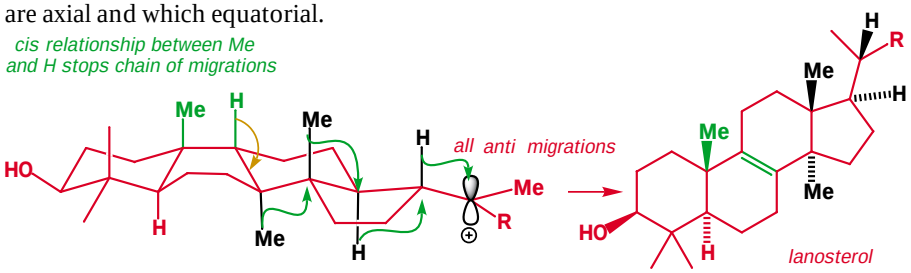
Finally, we have reached lanosterol. Now we will go back over these two steps and discuss them a bit more. Consider first the regiochemistry of the cyclization. The epoxide opens in the way we would expect to give positive charge at the more substituted carbon atom and then all the alkenes attack through 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.



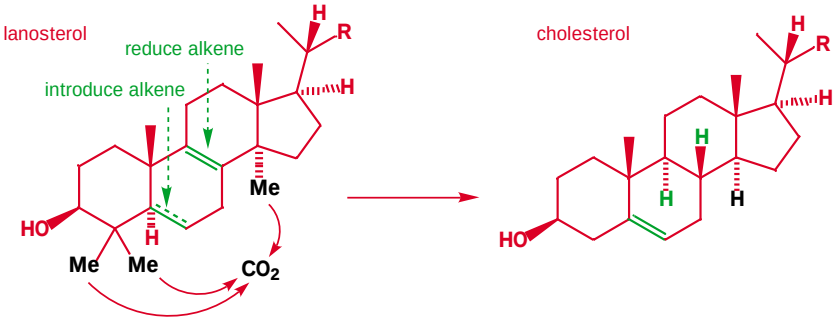

When the transition state for a ring closure forms a chair then a *trans* relationship results. This is the case for the black Me and brown Me. When a boat is formed a *cis* relationship results. This is the case for the green Me and black H. Squalene folds up in a chair–boat–chair conformation and that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



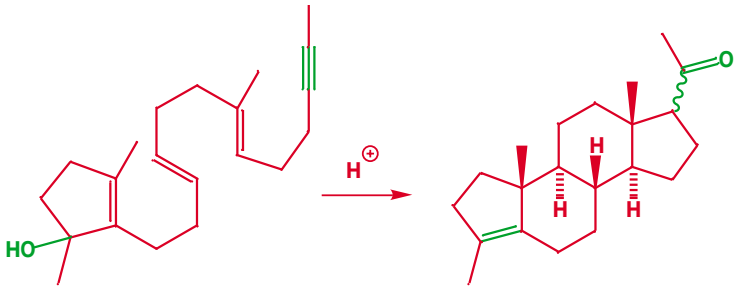
## Biomimetic synthesis: learning from Nature

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.

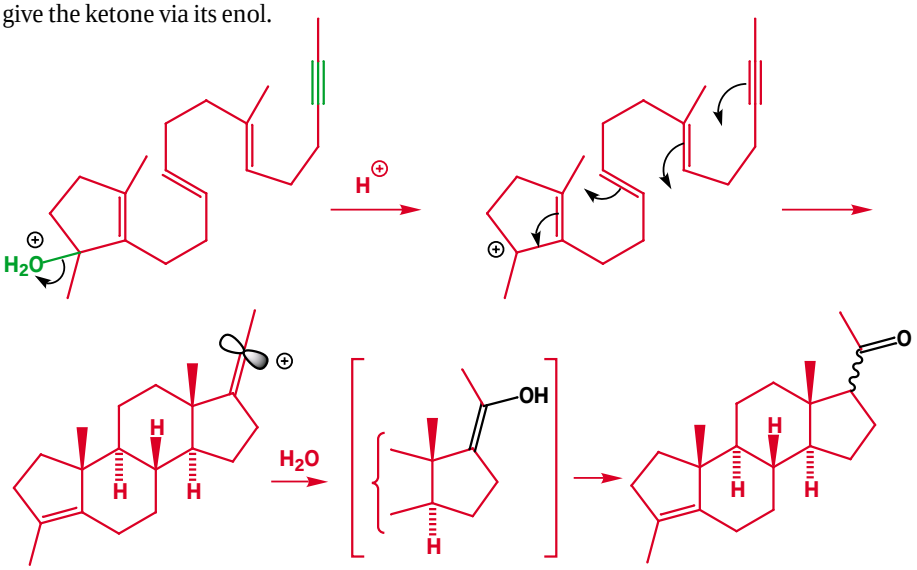
There are quite a lot of differences between the chemical and the biochemical versions so far—the chemical ones are less complex and less sophisticated but more versatile. The reactions are just cyclizations without the backbone rearrangements. The most important points of difference are:

- 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

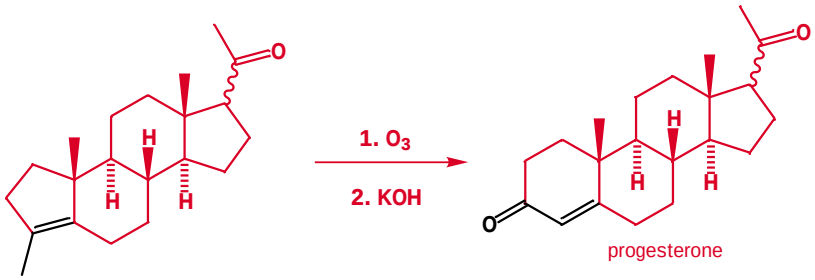
Here is one of Johnson's best examples which leads eventually to a biomimetic synthesis of the human hormone progesterone. The cyclization occurs just on treatment of the tertiary alcohol with acid.




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.

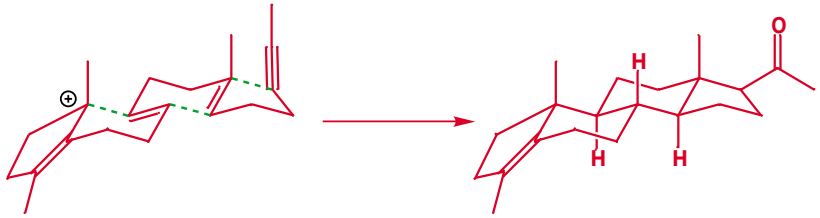


The five-membered ring A is there to ensure efficient initiation of the cyclization by the symmetrical allylic cation. It can easily be opened with ozone and the product cyclized to progesterone.




progesterone

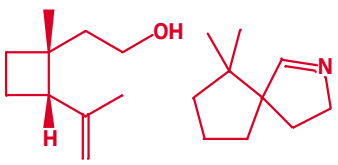
The conformation of the molecule in the moment of cyclization can be seen easily by working backwards from the product. The green dashed lines show new bonds that are being formed. All the six-membered rings in the transition state are chairs and all the ring junctions *trans*. This is an impressive 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

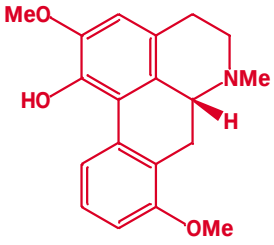
**1.** Assign each of these natural products to a general class (such as amino acid metabolite, terpene, polyketide) explaining what makes you choose that class. Then assign them to a more specific part of the general class (for example, tetraketide, sesquiterpene).



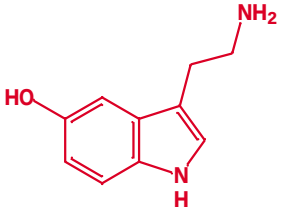

grandisol



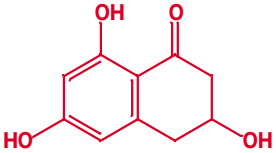

polyzonimine



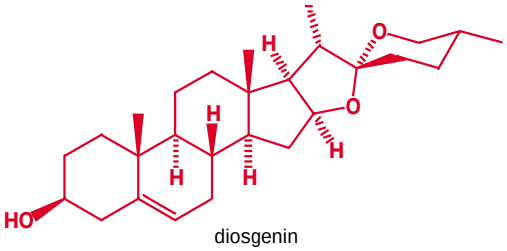

stephanine




serotonin

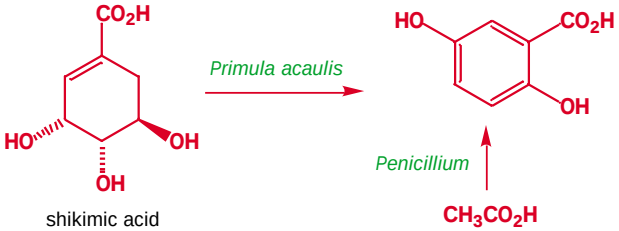



skytalone



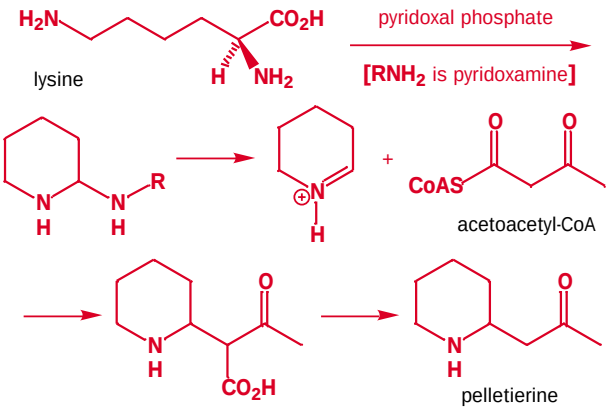

diosgenin

**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.

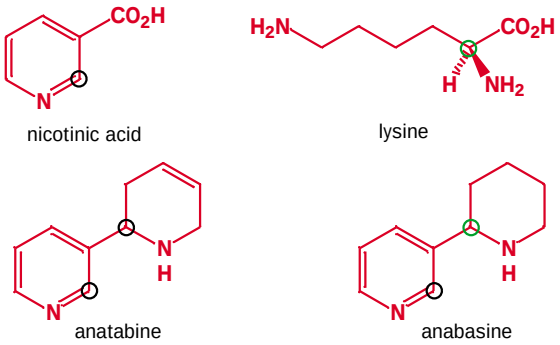



shikimic acid

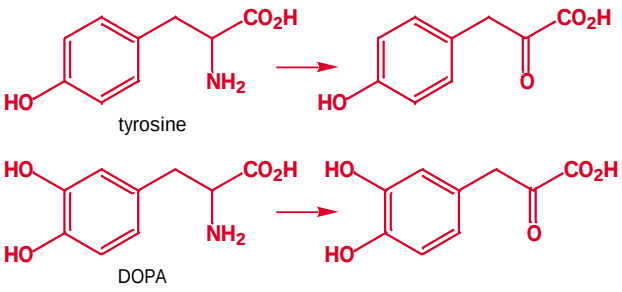
**3.** The piperidine alkaloid pelletierine was mentioned in the chapter but full details of its biosynthesis were not given. There follows an outline of the intermediates and reagents used. Fill in the details. Pyridoxal chemistry is discussed in Chapter 50.

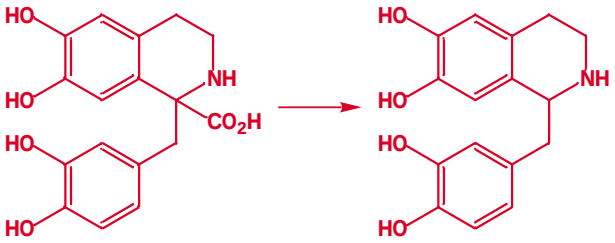



**4.** The rather similar alkaloids anabasine and anatabine come from different biosynthetic pathways. Labelling experiments outlined below show the origin of one carbon atom from lysine and others from nicotinic acid. Suggest detailed pathways. (*Hint*. Nicotinic acid and the intermediate you have been 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.)

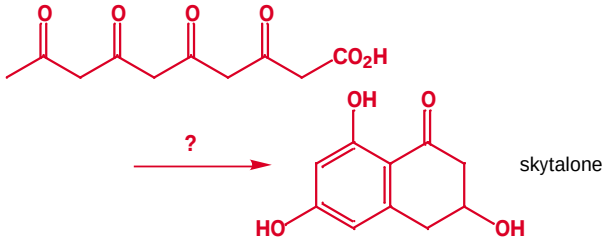



**5.** The three steps in the biosynthesis of papaverine set out below involve pyridoxal (or pyridoxamine). Write detailed mechanisms.

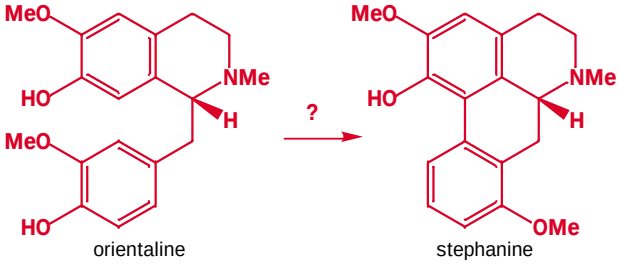


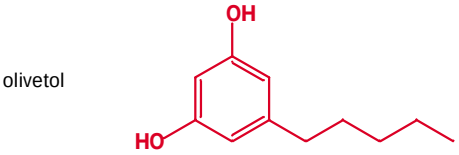
**6.** Concentrate now on the biosynthesis of skytalone in the first problem. You should have identified it as a pentaketide. Now consider how many different ways the pentaketide chain might be folded to give skytalone.



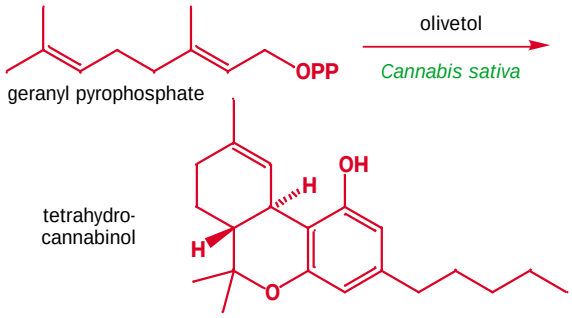

**7.** This question concerns the biosynthesis of stephanine, another compound mentioned in Problem 1. You should have deduced that it is a benzylisoquinoline alkaloid. Now suggest a biosynthesis from orientaline.



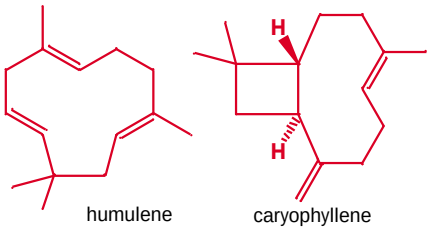

**8.** Suggest a biosynthesis of olivetol.



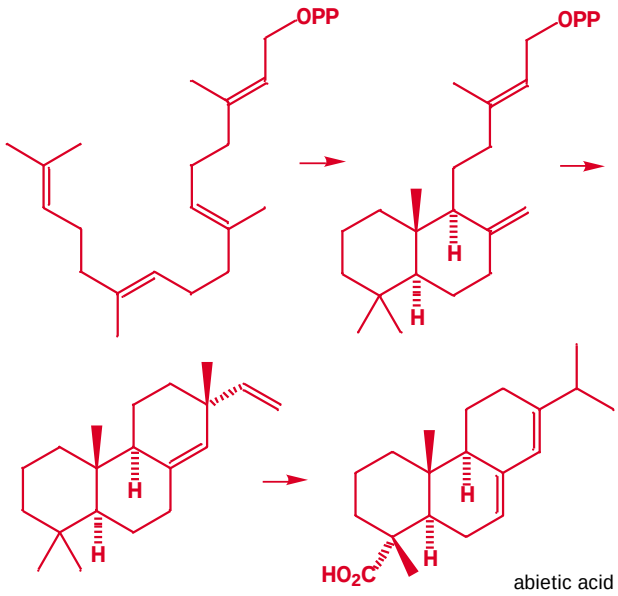
**9.** Tetrahydrocannabinol, the major psychoactive compound in marijuana, is derived in the *Cannabis* plant from olivetol and geranyl pyrophosphate. Details of the pathway are unknown. Make some suggestions and outline a labelling experiment to establish whether your suggestions are correct.



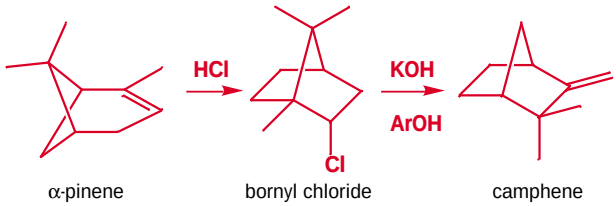

**10.** Both humulene, mentioned in the chapter, and caryophyllene are made in nature from farnesyl pyrophosphate in different plants. Suggest detailed pathways. How do the enzymes control which product is formed?



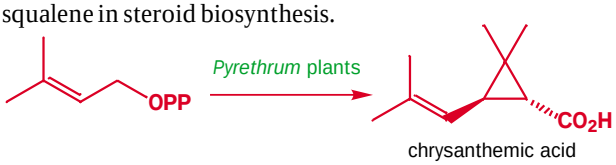

**11.** Abietic acid is formed in nature from mevalonate via the intermediates shown. Give some more details of the cyclization and rearrangement steps and compare this route with the biosynthesis of the steroids.



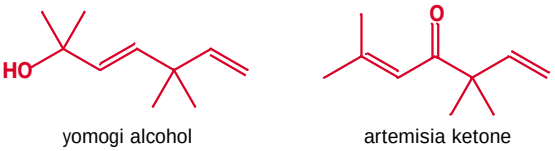

**12.** Borneol, camphene, and α-pinene are made in nature from geranyl pyrophosphate. The biosynthesis of α-pinene and the related camphor is described in the chapter. In the laboratory bornyl chloride and camphene can be made from α-pinene by the reactions described below. Give mechanisms for these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



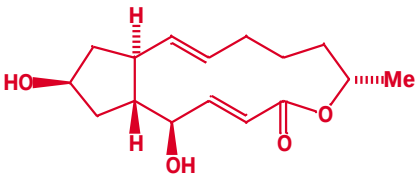

**13.** Suggest a biosynthetic route to the monoterpene chrysanthemic acid that uses a reaction similar to the formation of squalene in steroid biosynthesis.



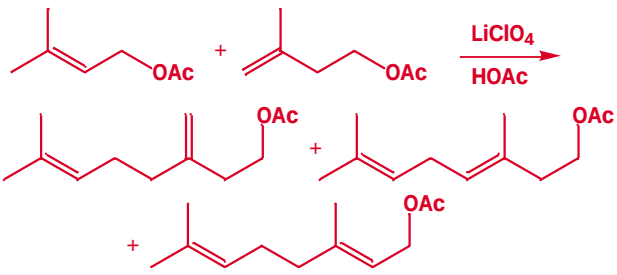

How could the same route also lead to the natural products yomogi alcohol and artemisia ketone?




**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?



# Polymerization

# 52

## Connections

### Building on:

- Carbonyl chemistry ch12 & ch14
- Substitution reactions ch17
- Radical reactions ch39
- Protecting groups and synthesis ch24–ch25
- The aldol reaction ch27
- Making double bonds ch31
- Cycloadditions ch35
- Heterocycles ch43–ch44
- Organometallics ch48
- The chemistry of life ch49
- Natural products ch51

### Arriving at:

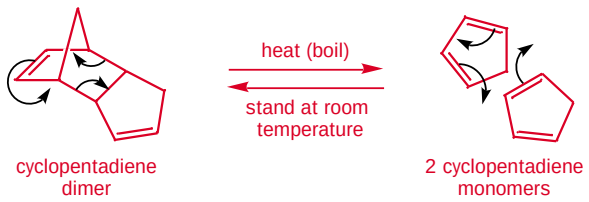
- Some molecules react together to form oligomers
- Some molecules spontaneously polymerize
- Polyamides, polyesters, and polycarbonates are formed by substitution reactions at carbonyl groups
- Polyurethane foams are formed by nucleophilic attack on isocyanates
- Epoxy adhesives work by polymerization via substitution reactions at saturated carbon
- The most important polymers are derived from alkene monomers
- Alkenes can be polymerized by radical, cationic, anionic, or organometallic methods
- Cross-linking or co-polymerization changes the physical properties of polymers
- Reactions on polymers are involved in paint drying, rubber strengthening, and the chemical synthesis of peptides

Most of the things you can see about you at this moment are made of organic polymers. Skin, clothes, paper, hair, wood, plastic, and paint are among them. Teeth, muscle, glue, cling film, starch, crab shells, and marmalade are all polymer-based too. In this chapter we will explore the world of polymers. We will ask questions like these:

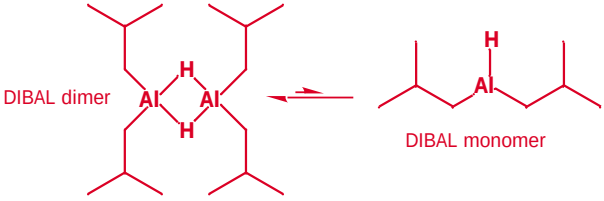
- What makes a molecule prefer to react with others of its kind to form a polymer?
- What mechanisms are available for polymerization reactions?
- How can polymerization reactions be controlled?
- How are the properties of polymers related to their molecular structure?

## Monomers, dimers, and oligomers

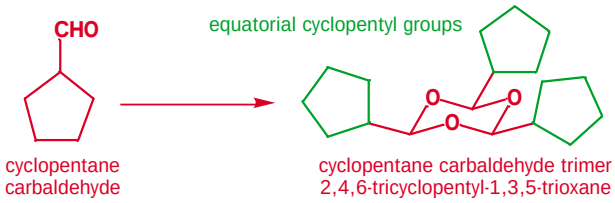
Cyclopentadiene featured in Chapter 35 as an important diene in the Diels–Alder reaction. If you try to buy 'cyclopentadiene' you will find that the catalogues list only 'dicyclopentadiene' or 'cyclopentadiene dimer'. The dimerization of cyclopentadiene is reversible: the monomer dimerizes by a Diels–Alder reaction at room temperature to give the dimer and the reaction is reversed on heating. So the dimer is a good source of the monomer.



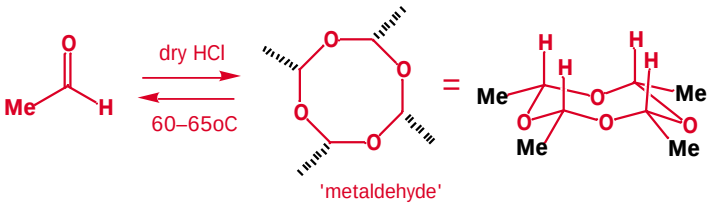

Other familiar cases of stable dimers are neutral boron and aluminium hydrides. DIBAL, for example, exists as two molecules linked by Al–H–Al bonds in a four-membered ring. Again, the dimer is a practical source of monomer for chemical reactions.

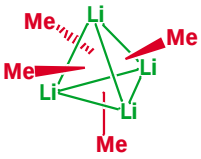


Simple aldehydes easily form trimers. When cyclopentanecarbaldehyde is prepared, it is a colourless liquid. On standing, particularly with traces of acid, it forms the crystalline trimer. The trimer is a stable six-membered heterocycle with all substituents equatorial



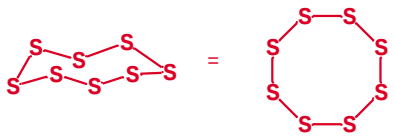

Acetaldehyde (ethanal) forms a liquid trimer called 'paraldehyde', which reverts to the monomer on distillation with catalytic acid. More interesting is 'metaldehyde', the common slug poison, which is an all-*cis* tetramer (2,4,6,8-tetramethyl-1,3,5,7-tetroxocane) formed from acetaldehyde with dry HCl at below 0 °C. Metaldehyde is a white crystalline solid that has all the methyl groups pseudoequatorial, and it reverts to acetaldehyde on heating.



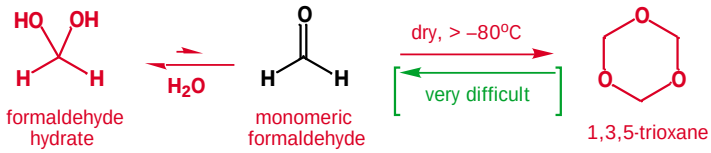




Another tetramer is methyl lithium. MeLi is a very reactive compound in the monomeric state, and it crystallizes as a tetramer: a tetrahedron of lithium atoms with a methyl group 'plugged in' to the centre of each face.

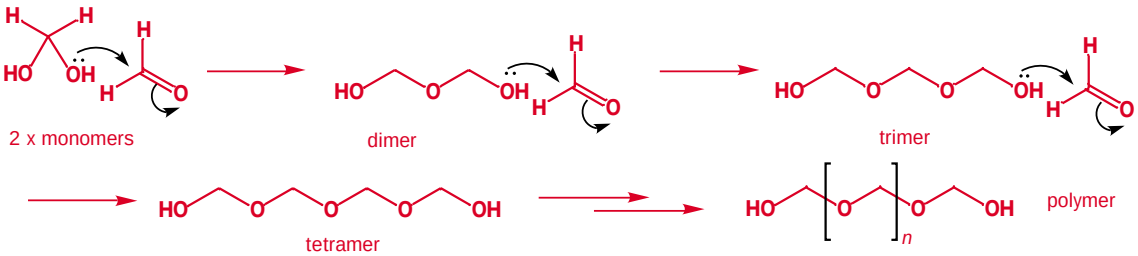
Whereas oxygen gas consists of diatomic molecules $O_2$, crystalline sulfur is $S_8$, a cyclic octomer. Such multiples are usually called **oligomers** (*oligo* = a few). The monomer in this case would be the sulfur atom. The shape of the $S_8$ ring is very similar to that of the eight-membered ring of metaldehyde.




If you buy formaldehyde (methanal), which is in fact a gas, b.p. –19 °C, you have four choices. You can buy a 37% aqueous solution 'formalin' which is mostly hydrate in equilibrium with a small amount of formaldehyde, or the crystalline trimer (1,3,5-trioxane), or a white solid called (misleadingly) 'paraformaldehyde', or another white solid called polyoxymethylene.

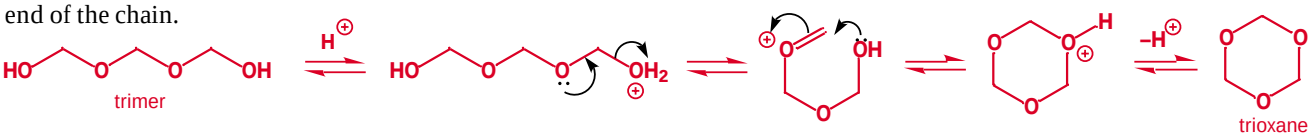



Trioxane is not a good source of formaldehyde as it is very stable but the two other solids are good sources. Both paraformaldehyde and, more obviously, polyoxymethylene are polymers. Each molecule of either polymer consists of a large number of formaldehyde molecules reacted together.




Paraformaldehyde is made by evaporation of aqueous formaldehyde to dryness and is a water-soluble polymer. Polyoxymethylene is made by heating formaldehyde with catalytic sulfuric acid and is *not* soluble in water. They are both linear polymers of formaldehyde, so how can they be so different? The answer is in the polymer chain length—the *n* in the diagram. Paraformaldehyde is water-soluble because it has short chain lengths, about $n = 8$ on average, and so it has many hydrophilic OH groups. Polyoxymethylene has much longer chain lengths, $n > 100$ on average, and so has very few OH groups per monomer of formaldehyde.

Trioxane is formed when the trimer cyclizes instead of continuing to polymerize. All the oligomers and polymers of formaldehyde have this potential as there is a hemiacetal group at each end of the chain.




**Summary of what we know so far**

Not much, you might think. Actually we have mentioned some important things about polymerization, which we will discuss further in the pages that follow.

- Polymerization tends to occur at low temperature
- Depolymerization tends to occur at high temperature
- Polymerization competes with cyclic oligomer formation
- Different polymers of the same monomer can have different chain lengths
- The chain length varies about a mean value in a given polymer
- The properties of polymers depend on chain length (among other things!)

Check back over these last few pages to make sure you see which pieces of evidence establish each of these points.

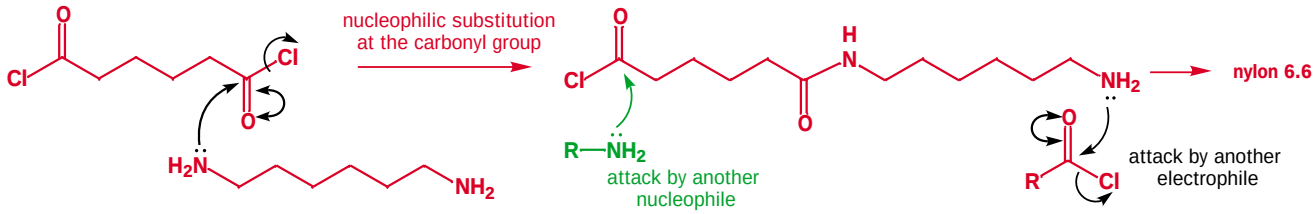
There is no exact limit to the terms oligomer and polymer. You have just seen us refer to paraformaldehyde—on average an octomer—as a polymer. The terms monomer, dimer, trimer, tetramer, etc. do have exact meanings. Oligomer usually means > 3 and < 25 but different authors will use the term in different ways.

## Polymerization by carbonyl substitution reactions

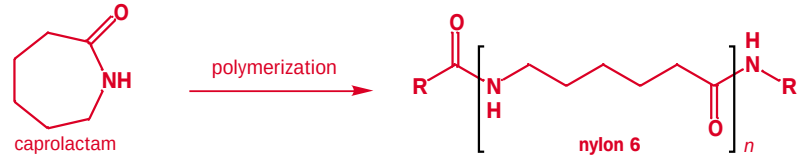
In general, carbonyl compounds do not polymerize by themselves. It is only the exceptional reactivity of formaldehyde as an electrophile that allows repeated nucleophilic addition of hemiacetal intermediates. A more common way to polymerize carbonyl compounds is to use two different functional groups that react together by carbonyl substitution to form a stable functional group such as an amide or an ester. Nylon is just such a polymer.

### Polyamides

You may have carried out the nylon rope trick in a practical class. The diacid chloride of adipic acid is dissolved in a layer of a heavy organic solvent such as $CCl_4$ and a layer of aqueous hexane-1,6-diamine is carefully placed on top. With a pair of tweezers you can pick up the film of polymer that forms at the interface and draw it out to form a fibre. The reaction is a simple amide formation.

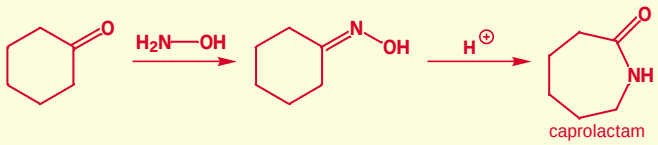



After the first amide is formed, one end of the new molecule is nucleophilic and the other electrophilic so that it can grow at both ends. The polymer is made up of alternating $–NH(CH_2)_6NH–$ and $–(CH_2)_4CO–$ units, each having six carbon atoms, and is called 'nylon 6.6'. Another and much simpler way to make nylon is to polymerize caprolactam. This monomer is a cyclic amide and the polymer does not have alternating units—instead, each unit is the same.



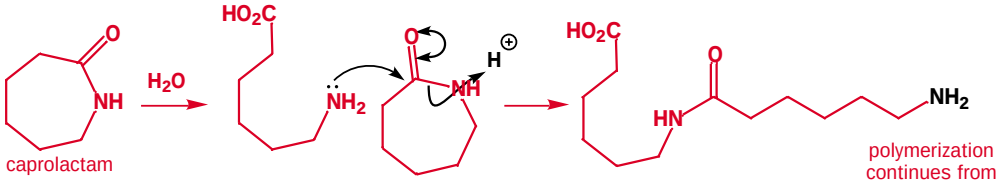
## Caprolactam

Caprolactam can be made by the Beckmann rearrangement of the oxime of cyclohexanone. (Check that you can draw the mechanisms, of both these reactions and look at Chapters 14 and 37 if you find you can't.) Cyclohexanone used to be made by the oxidation of cyclohexane with molecular oxygen until the explosion at Flixborough in Lincolnshire on 1 June 1974 that killed 28 people. Now cyclohexanone is made from phenol.



caprolactam

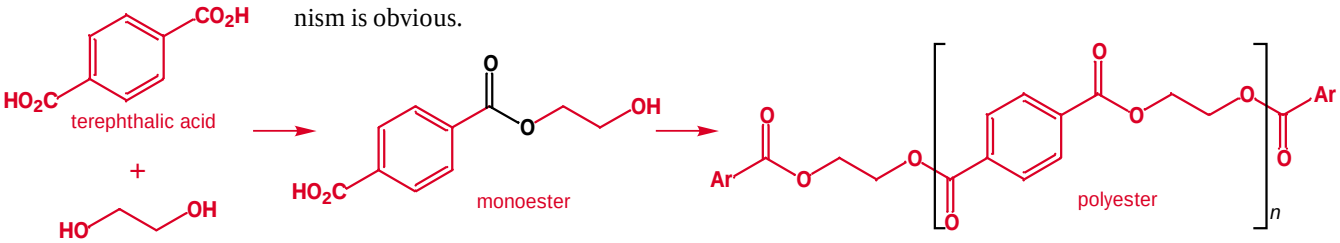
So how is this polymerization initiated? A small amount of water is added to hydrolyse some of the caprolactam to 6-aminohexanoic acid. The amino group can then attack another molecule of caprolactam and so on. The amount of water added influences the average chain length of the polymer.




These synthetic polyamides are made up of the same repeating unit but will inevitably have a range of molecular weights as the polymer length will vary. This is a different story from that of the natural polyamides—peptides and proteins—that you met in Chapter 49. Those polymers were made of twenty or so different monomers (the amino acids) combined in a precise order with a precise stereochemistry and all molecules of the same protein have the same length. Nonetheless, some of their uses are almost identical: both nylon and wool are polyamides, for example.

## Polyesters

Much the same act can be carried out with dicarboxylic acids and diols. The most famous example is the polymer of ethylene glycol (ethane-1,2-diol) and terephthalic acid, which can be made simply by melting the two components together so that water is lost in the esterification reaction. The mechanism is obvious.

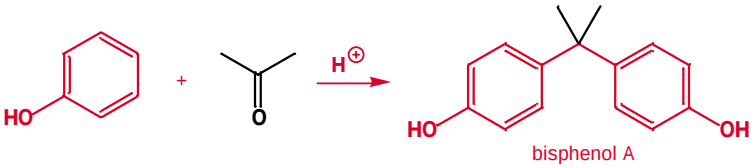



This linear polymer, like nylon, is well shaped for making long fibres and is now so important for making clothes that it is usually just called 'polyester' rather than by the older names such as 'Terylene'.

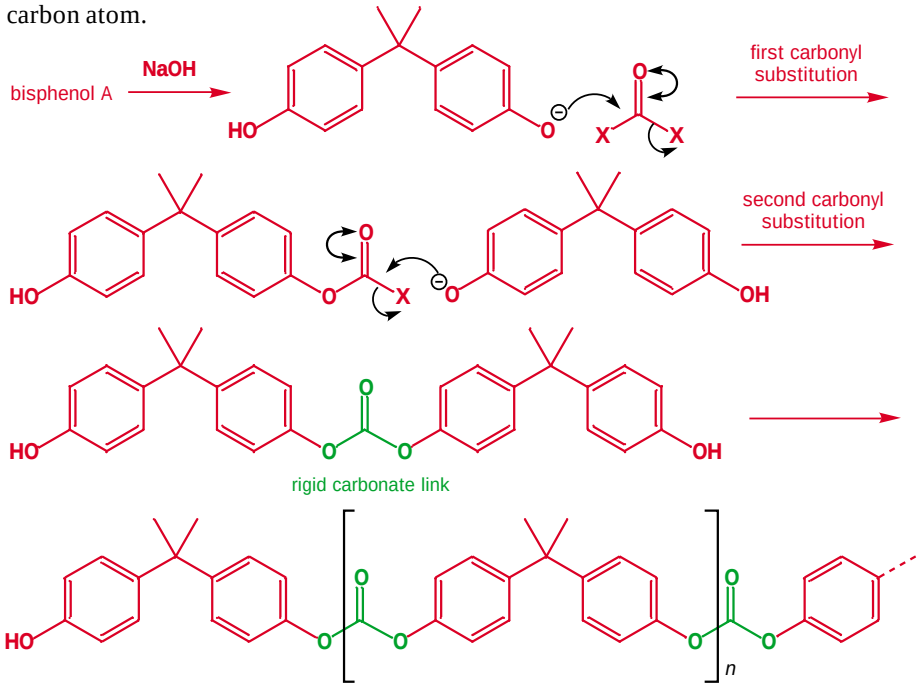
## Polycarbonates

These too are made by carbonyl substitution reactions, but this time the nucleophile is aromatic and the electrophile is an aliphatic derivative of carbonic acid such as phosgene ($COCl_2$) or a carbonate diester [$CO(OR)_2$]. The aromatic nucleophile is a diphenol but the two OH groups are on separate rings joined together by an electrophilic aromatic substitution. This compound is called bisphenol A and has many other applications.

■ Make sure that you can draw the mechanism for this reaction—two electrophilic aromatic substitutions are involved (Chapter 22). If you need a hint, look at the synthesis of Bakelite on the next page.



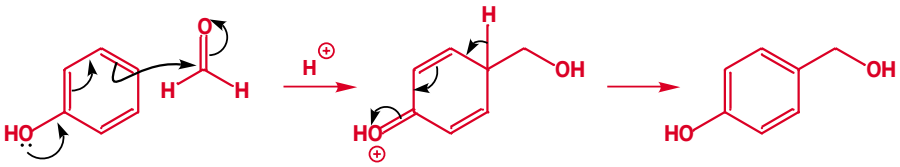
The diphenol reacts with the carbonic acid derivative, which is doubly electrophilic at the same carbon atom.



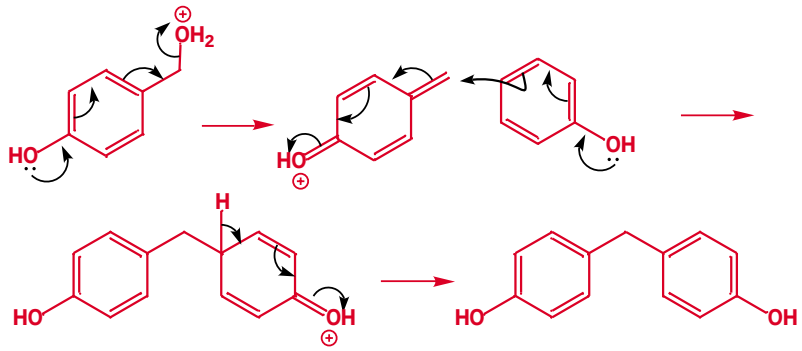

After two carbonyl substitutions the rigid carbonate ester group is formed. This polymer is neither as flexible nor as linear as the previous examples. The carbonate portion is conjugated to the benzene rings and held rigidly in the conformation shown by the anomeric effect (Chapter 42). The only flexibility is where the $CMe_2$ group links the two benzene rings. This is a polymer that combines transparency, lightness, and strength with just enough flexibility not to be brittle. Your safety glasses are probably made of polycarbonate.

## Polymerization by electrophilic aromatic substitution

The first synthetic polymers to be of any use were the 'phenol formaldehyde resins' of which the most famous, Bakelite, was discovered by Bäkeland at the turn of the century. He combined phenol and formaldehyde in acid solution and got a reaction that starts like the bisphenol A synthesis.



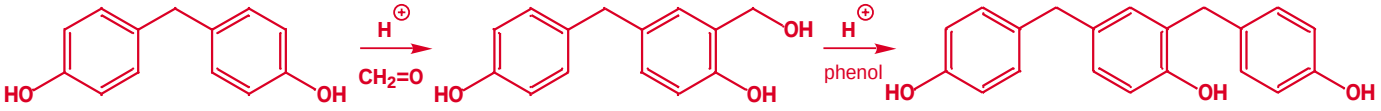
A second acid-catalysed electrophilic aromatic substitution now occurs to link a second phenol to the first. The rather stable benzylic cation makes a good intermediate.



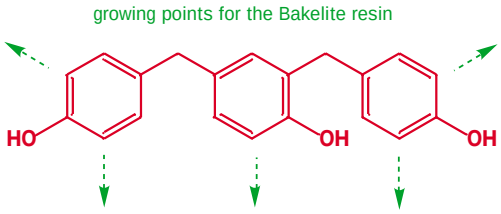
■

If you tried a moment ago, as we suggested, to write the mechanism for the formation of bisphenol A, this is what you should have done (but with acetone, of course, instead of formaldehyde).

Formaldehyde is reactive enough to continue and put another substituent ortho to the OH group in one of the rings. The mechanisms are the same as those we have just written.



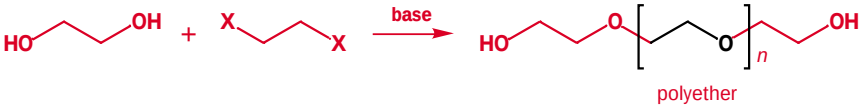

The carbon chains are *meta* related on the central ring so for the first time we have a branched polymer. Complexity can rapidly increase as more phenols linked through more formaldehydes can be joined on to this core structure at several points. Each benzene ring could, in theory, form three new C–C bonds.




These polymers have the useful property of being thermosetting—they are made from liquid mixtures that polymerize on heating to form a solid polymer, and can therefore be moulded easily.

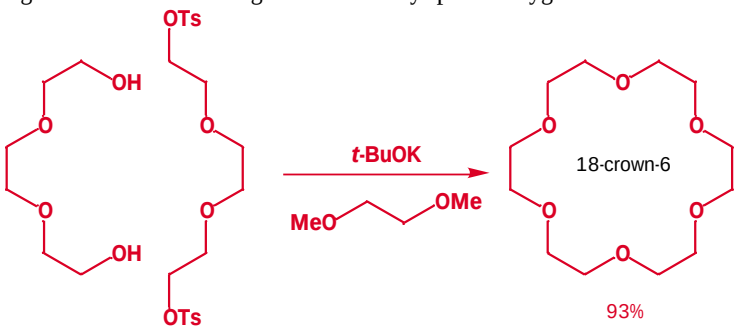
## Polymerization by the $S_N2$ reaction

In principle, co-polymerization of a 1,2-diol and a 1,2-dihalide might lead to a polyether.

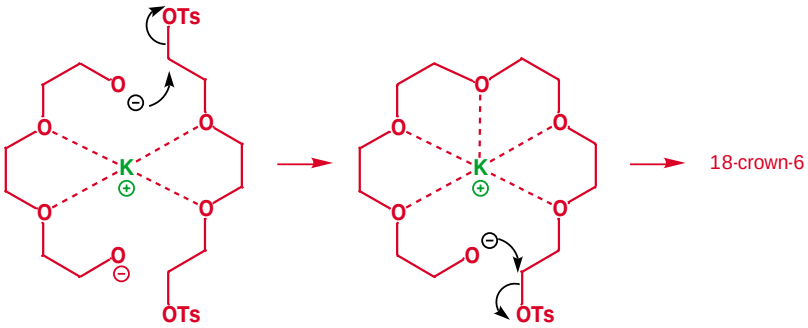



This route is not used because of the large amounts of base needed. One molecule of base is consumed for each new C–O bond made, and these reactions terminate quickly before long chains are made. It is more useful for making the cyclic oligomers called 'crown ethers'. 18-Crown-6 has an eighteen-membered ring with six evenly spaced oxygen atoms.

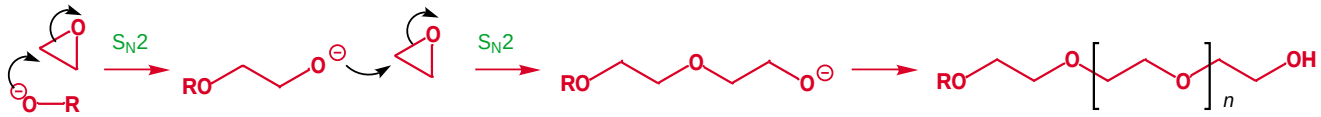
■ We discussed the use of crown ethers on p. 000.



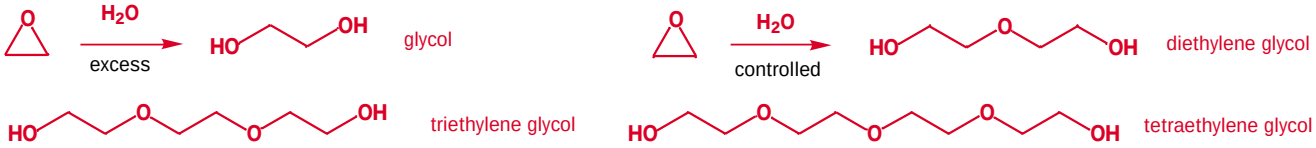

These crown ethers have cavities ideal for complex formation with metal ions. They can even carry metal ions into solution in organic solvents. This one, 18-crown-6, is the right size for potassium ions, and a solution of $KMnO_4$ and 18-crown-6 in benzene, so-called 'purple benzene', is a useful oxidizing agent. The high-yielding oligomerization is a template reaction with a potassium ion holding the two reagents together. If a base such as $Bu_4N^+OH^-$ (which cannot form complexes) is used with the same reagents, linear polymers result.



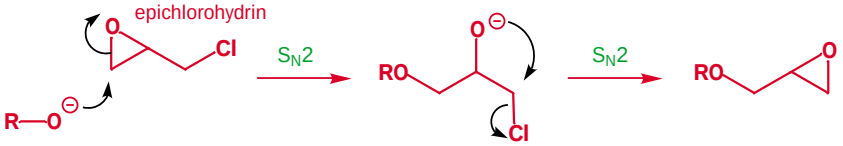
A more practical way to make linear polyethers is by polymerization of epoxides. Each time an epoxide is opened by a nucleophile, it releases a nucleophilic oxyanion that can attack another epoxide, and so on. The whole process can be initiated by just a catalytic amount of a nucleophile such as an alkoxide or an amine.



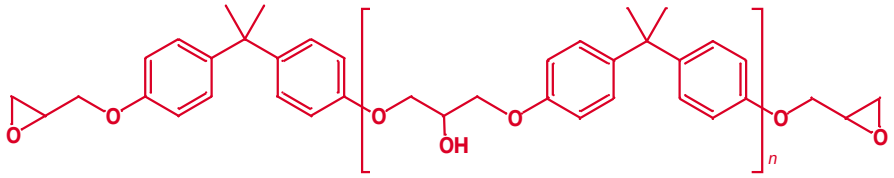

This reaction cannot be controlled—once it is initiated, it runs to completion. Treatment of ethylene oxide with controlled amounts of water does lead to the important coolant ethylene glycol (excess water) and the oligomers di-, tri-, and tetraethylene glycol. These are important solvents for polar compounds. Triethylene glycol is also the starting material for the synthesis of 18-crown-6 above.



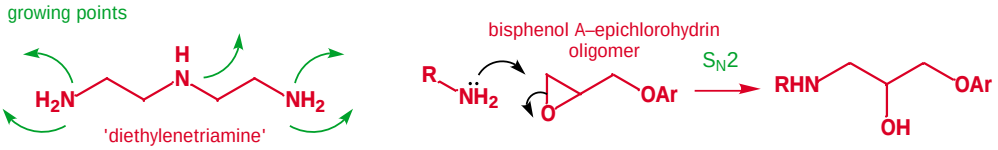

A subtle method of controlling the reaction so that it can be made to run at will is to use bisphenol A as the diol and epichlorohydrin as the epoxide. Epichlorohydrin reacts with nucleophiles at the epoxide end, but the released alkoxide ion immediately closes down at the other end to give a new epoxide.



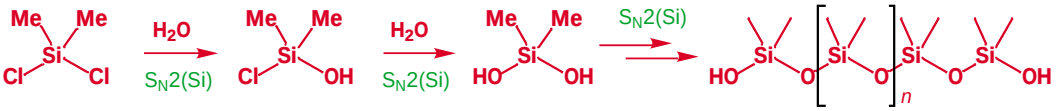

With bisphenol A in alkaline solution, this reaction happens twice and a bis adduct is formed. Further reaction with more bisphenol A creates oligomers with about 8–10 bisphenol A molecules and an epoxide at each end. This is a reasonably stable neutral compound with two terminal epoxides, just waiting for initiation for polymerization to start.




In the CIBA–Geigy glue Araldite, strong enough to glue aeroplane wings on to the fuselage, a solution of this oligomer is mixed with a solution of a polyfunctional amine such as diethylenetriamine. Since each $NH_2$ group can react twice and the NH group once with epoxides, the final polymer has a densely cross-linked structure and is very strong. The reaction is again a simple $S_N2$ process.



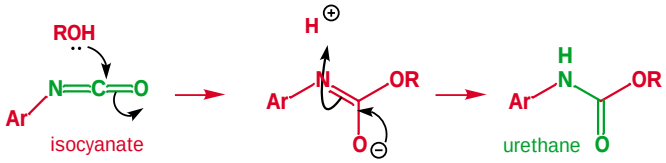

A totally different kind of polymer is a poly-silylether. Dimethylsilyl dichloride polymerizes easily on treatment with hydroxide. Silicon is more susceptible to the $S_N2$ reaction than is carbon and long chains grow quickly.



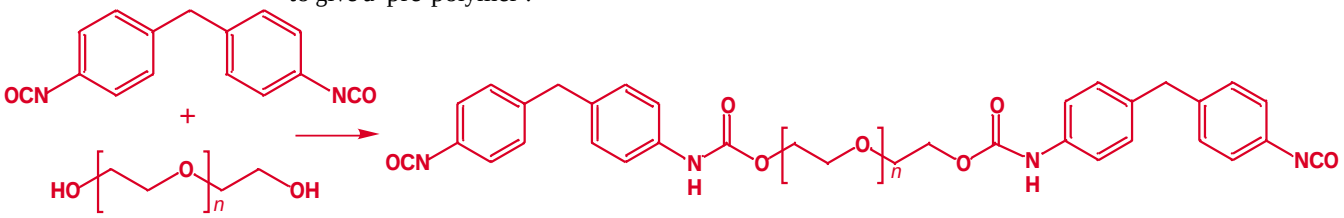

This linear poly(dimethylsiloxane) is an oil and is used in the lab in oil baths as it is more stable and less smelly than conventional paraffin baths at high temperatures.

## Polymerization by nucleophilic attack on isocyanates

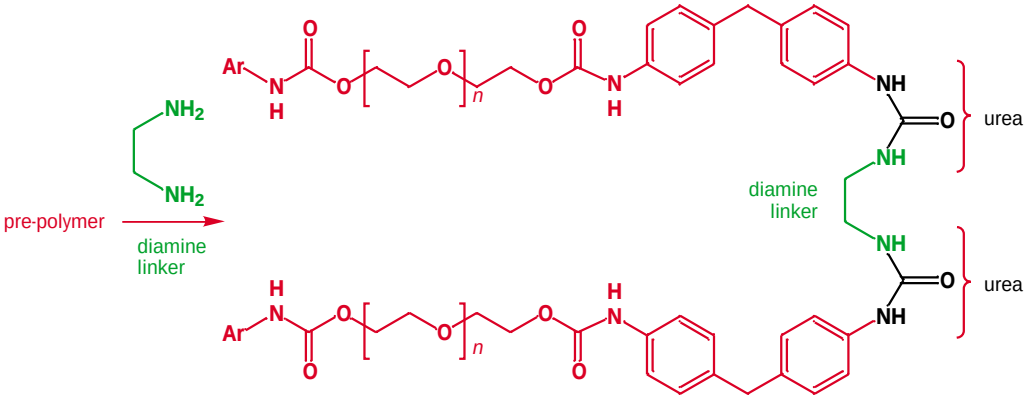
Isocyanates react with alcohol nucleophiles to give urethanes—hybrids between carbonates and ureas—half-esters and half-amides of carbonic acid. Nucleophilic attack occurs at the very reactive linear (sp) carbon in the centre of the isocyanate.




To make a polymer it is necessary to react aryl diisocyanates with diols. Some important polymers of the type, called elastanes, are made by using long-chain aliphatic diols from partly polymerized epoxides, rather like those discussed in the last section, and reacting them with diaryl diisocyanates to give a 'pre-polymer'.




The next stage is to initiate an exothermic linking of the residual terminal isocyanates with simple diamines. The reaction is again nucleophilic attack on the isocyanate, but the new functional group is now a urea rather than a urethane. Showing just one end of the growing polymer:




These polymers have short rigid portions (the aromatic rings and the ureas) joined by short flexible 'hinges' (the diamine linker and the $CH_2$ group between the aromatic ring) and long very flexible portions (the polyether) whose length can be adjusted. The polymer is easily stretched and regains its shape on relaxation—it is an elastomer.

Why should it matter that the second polymerization is exothermic? If the diamine linker is added as a solution in a volatile hydrocarbon such as heptane, the heat of the polymerization causes the heptane to boil and the polymer becomes a foam. What is more, the length of the polyether chain determines what kind of foam results. Shorter (~500 $–OCH_2CH_2O–$ units) chains give rigid foams but longer chains (>1000 $–OCH_2CH_2O–$ units) give soft foams. This is only a bare outline of one of the many skills polymer chemists now have in the design of materials. The results are all around us.

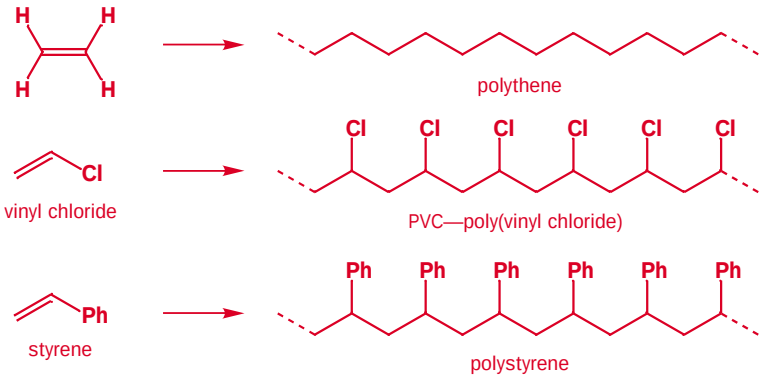
So far we have discussed polymerization that has been essentially of one kind—bifunctional molecules have combined in normal ionic reactions familiar from the rest of organic chemistry where a nucleophilic functional group attacks an electrophilic functional group. The new bonds have generally been C–O or C–N. We need now to look at the polymerization of alkenes. In these reactions, C–C bonds will be formed and many of the reactions may be new to you.

# Polymerization of alkenes

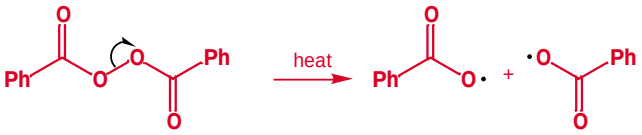
Formaldehyde polymerizes because the two resulting C–O σ bonds are very slightly more stable than its C=O π bond, but the balance is quite fine. Alkenes are different: two C–C σ bonds are always considerably more stable than an alkene, so thermodynamics is very much on the side of alkene polymerization. However, there is a kinetic problem. Formaldehyde polymerizes without our intervention, but alkenes do not. We will discuss four quite distinct mechanisms by which alkene polymerization can be initiated—two ionic, one organometallic, and one radical.

## Radical polymerization of alkenes: the most important polymerization of all

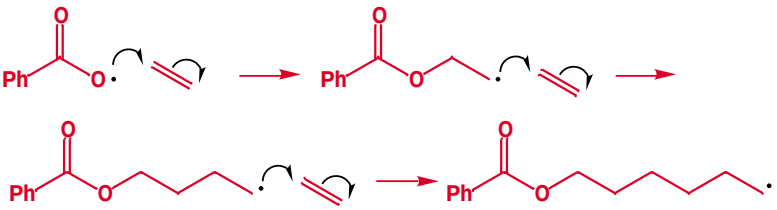
We will start with the radical mechanism simply because it is the most important. A bigger tonnage of polymers is made by this method than by any other, including the three most familiar ones—polythene (polyethylene), PVC (poly(vinyl chloride)), and polystyrene.



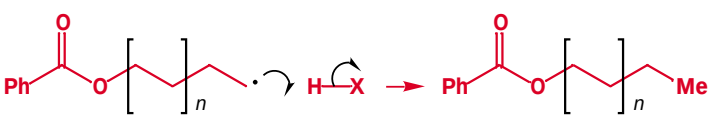

Polythene is difficult to make and was discovered only when chemists at ICI were attempting to react ethylene with other compounds under high pressure. Even with the correct reagents, radical initiators like AIBN or peroxides (Chapter 39), high pressures and temperatures are still needed. At 75 °C and 1700 atmospheres pressure ethylene polymerization, initiated by dibenzoyl peroxide, is a radical chain reaction. The peroxide is first cleaved homolytically to give two benzoate radicals.



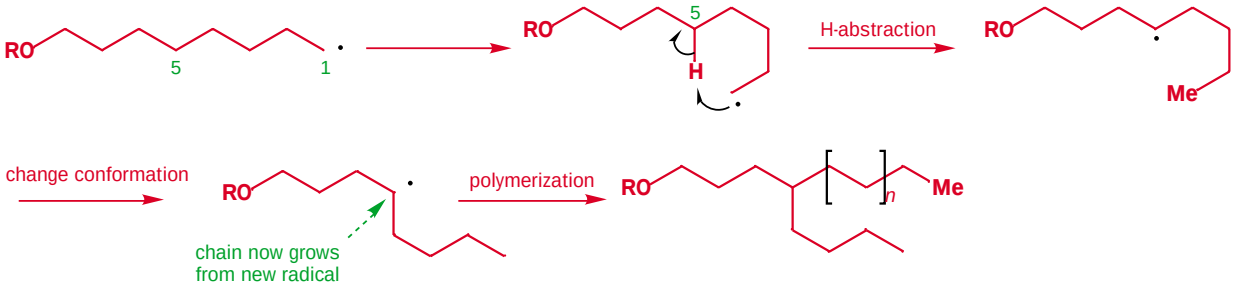

These oxyradicals add to the alkene to give an unstable primary carbon radical that adds to another molecule of alkene, and so on.



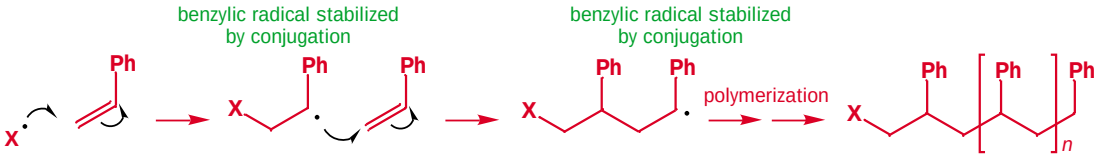

Eventually, the chain is terminated by combination with another radical (unlikely) or by hydrogen abstraction from another polymer molecule. This approach to polythene synthesis, using ethylene liquefied by pressure and small amounts (<0.005% by weight) of peroxide, produces relatively low molecular weight polymer as a white solid.

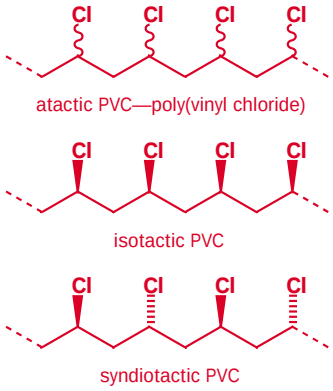


Radical polymerization can lead to branched polymers by intramolecular hydrogen atom transfer, a process sometimes called backbiting. Removal of H through a six-membered transition state moves the growing radical atom five atoms back down the chain, and leads to butyl side-chains. A more stable secondary radical is produced and chain growth then occurs from that point.



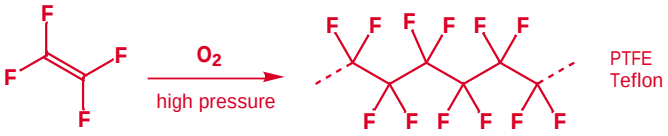

Radical polymerization of vinyl chloride and styrene is much easier than that of ethylene because the intermediate radicals are more stable. You saw in Chapter 39 that any substituent stabilizes a radical, but Cl and Ph are particularly good because of conjugation of the unpaired electron with a lone pair on chlorine or the π bonds in the benzene ring.



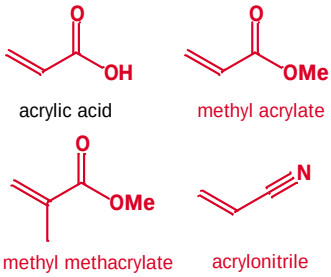

Neither PVC nor polystyrene is very crystalline and polystyrene often has poor mechanical strength. Both of these may be results of the stereorandom nature of the polymerization process. The substituents (Cl or Ph) are randomly to one side or other of the polymer chain and so the polymer is a mixture of many diastereoisomers as well as having a range of chain lengths. Such polymers are called atactic. In some polymerizations, it is possible to control stereochemistry, giving (instead of atactic polymers) isotactic (where all substituents are on the same side of the zig-zag chain) or syndiotactic (where they alternate) polymers.

A unique polymer is formed by the radical polymerization of tetrafluoroethylene and is called PTFE or Teflon. The outside of the polymer consists of a layer of fluorine atoms which repel all other molecules. It is used as the coating in nonstick pans and as a bearing that needs no lubrication. Two pieces of Teflon slide across one another almost without friction.

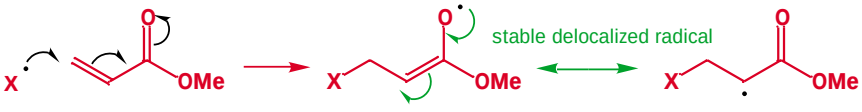



Something else is special about this polymerization—it is done in solution. Normally, no solvent is used because it would be difficult to separate from the polymer product. However, PTFE interacts with no other molecules. It precipitates from all known solvents and can be isolated easily by filtration.

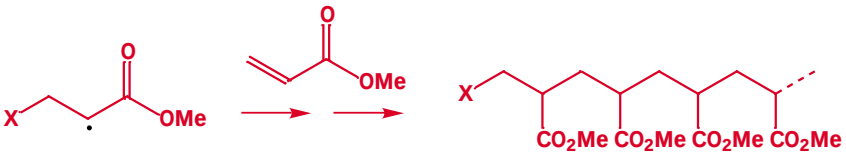
## Acrylics—easily made polymers of acrylate esters



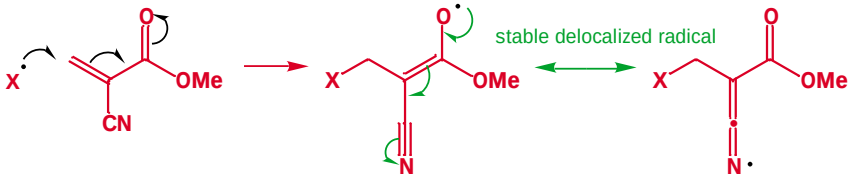

Alkenes conjugated with carbonyl groups, such as acrylates (derivatives of acrylic acid), are easily polymerized by a variety of mechanisms. Indeed, these compounds are often difficult to store because they polymerize spontaneously when traces of weak nucleophiles (even water) or radicals (even oxygen) are present. Radical polymerization occurs very easily because the intermediate carbon radical is stabilized by conjugation with the carbonyl group.



Polymerization follows the mechanism that we have seen several times already, and each radical has the same additional stabilization from the carbonyl group.



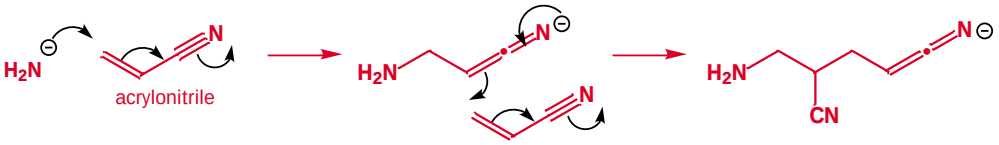

With two stabilizing groups on the carbon radical, polymerization becomes even easier. A famous example is 'SuperGlue', which is methyl 2-cyanoacrylate. The monomer in the tube polymerizes on to any surface (wood, metal, plastic, fingers, eyelids, lips, ...) catalysed by traces of moisture or air, and the bonds, once formed, are very difficult to break. The intermediate radical in this polymerization is stabilized by both CN and $CO_2Me$ groups.



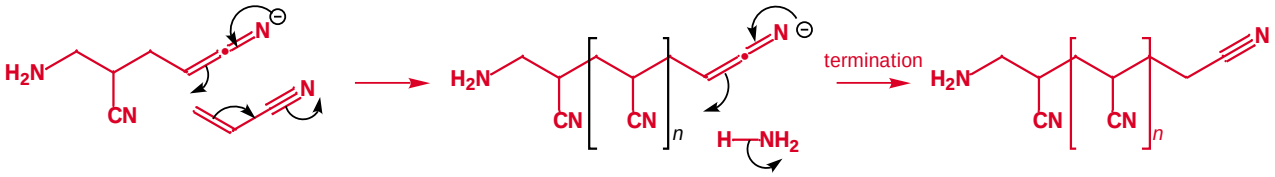

Though there are many other polymers made by radical pathways, we need now to look at the two main ionic routes—anionic and cationic polymerization.

## Anionic polymerization is multiple conjugate addition

We have seen in Chapter 23 how alkenes conjugated with electron-withdrawing groups undergo conjugate addition to give an enolate anion as an intermediate. This enolate anion is itself nucleophilic and could attack another molecule of the conjugate alkene. Acrylonitrile is polymerized in liquid ammonia at low temperature by this method. Small amounts of alkali metal are added to generate $NH_2^-$, initiating polymerization.

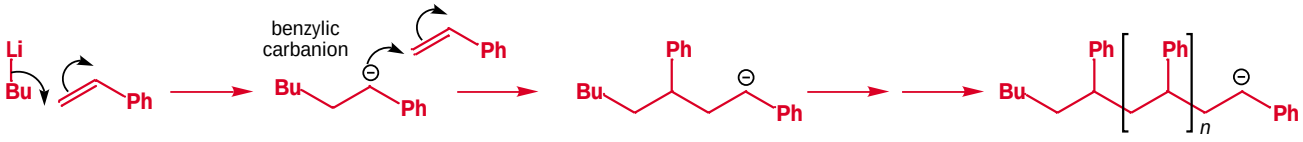



The chain grows by repetition of the last step: each new C–C bond-forming step produces a new anion stabilized by the nitrile group. Termination probably occurs most frequently by proton capture from the solvent. The result is poly(acrylonitrile).

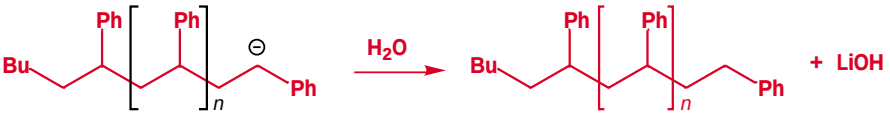



## 'Living polymers' by the anionic polymerization of styrene

Nucleophilic addition to styrene is possible only because the intermediate carbanion is stabilized by conjugation into the benzene ring. It needs a more reactive carbanion than the benzyl anion to initiate the polymerization, and an unstabilized nonconjugated organolithium compound like butyl lithium is the answer.

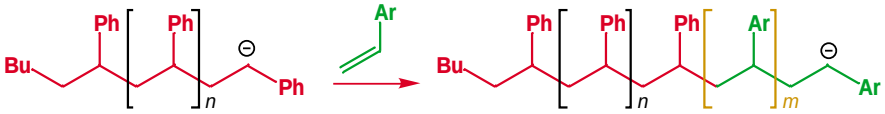


It is clear enough how the chain is propagated, but how is it terminated? You might expect protonation to bring things to a close, but there cannot be any acid (even a weak one) present—if there were, it would have already been destroyed by the butyl lithium. To terminate the polymerization, a weak acid must be added in a separate step—water will do.



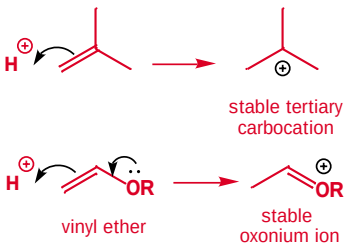

When this polystyrene sample is analysed, it is found to consist of a remarkably narrow range of chain lengths—almost all the chains are the same. Such polymers are known as monodisperse. This result must mean that all the BuLi molecules must add immediately to a styrene molecule and that chain growth then occurs at the same rate for each chain until the styrene is used up.

There is a useful expansion of this idea. Under the conditions of the polymerization (before the water is added), these almost identical chain lengths all end with a carbanion. If, instead of adding water, we add another monomer (say, 4-chlorostyrene) it too will add to the end of the chain and polymerize until it is used up, producing new chains again of about the same length. This will be the situation after the second polymerization.



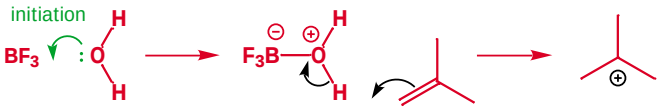

And still the polymer is active towards further polymerization. Indeed, these polymers are called 'living polymers' because they can go on growing when a new monomer is added. The final result, after as many monomers have been added as is required and the living polymer has been quenched, is a polymer with blocks of one monomer followed by blocks of another. These polymers are called block co-polymers for obvious reasons.

## Cationic polymerization requires stabilized carbocations

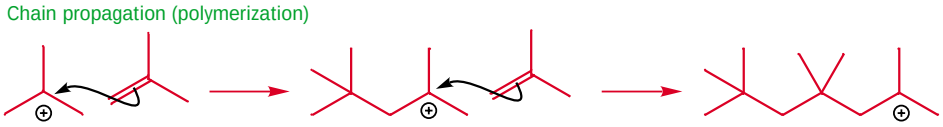



Cationic polymerization is used only for alkenes that can give a tertiary carbocation on protonation or for vinyl ethers that can give an oxonium ion. In other words, the cation intermediate must be quite stable. If it isn't, the chain is terminated too quickly by loss of a proton.

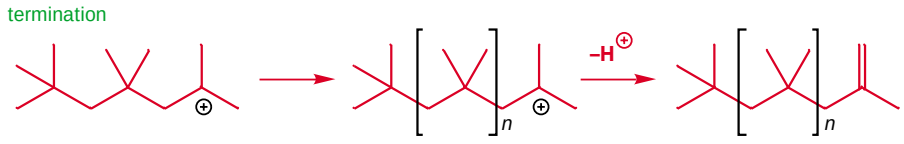
The initiator for isobutene (2-methylpropene) polymerization is usually a Lewis acid with a proton source. We shall illustrate isobutene polymerization with $BF_3$ as the Lewis acid and water as the proton source.



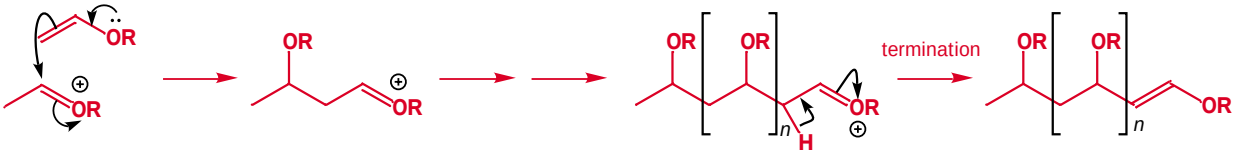

The tertiary carbocation can now act as an electrophile and attack the alkene to form another tertiary carbocation of similar stability and reactivity to the first. So the polymerization continues.




The termination will be the loss of a proton to form an alkene (an E1 reaction). Providing that the tertiary carbocation is reasonably stable, this will be a slower process than chain elongation, especially as there are no good bases around, and long polymer chains may result.



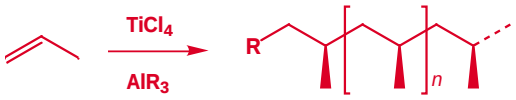
The polymerization of vinyl ethers follows much the same mechanism, using the oxonium ion as an intermediate instead of the tertiary carbocation. Termination might again be by loss of a proton or by picking up a nucleophile at the oxonium ion centre.




One of the best polymers for building strong rigid heat-resistant objects is polypropylene but this can be made by none of the methods we have examined so far. We need now to look at the polymerization of alkenes in the coordination sphere of a transition metal.

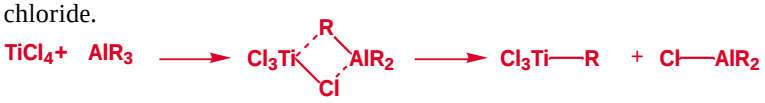
## Ziegler–Natta polymerization gives isotactic polypropylene

Propylene can be polymerized by a titanium/aluminium catalyst developed by Ziegler and Natta. The mere fact that polymerization is possible is remarkable, but this polymer also has stereoregularity and can be isotactic. The overall process is shown on the right.




■ The organometallic principles relating to this section can be found in Chapter 48.

The mixed metal compounds react to form a titanium σ complex that is the true catalyst for the polymerization. An alkyl group is transferred from aluminium to titanium in exchange for a chloride.

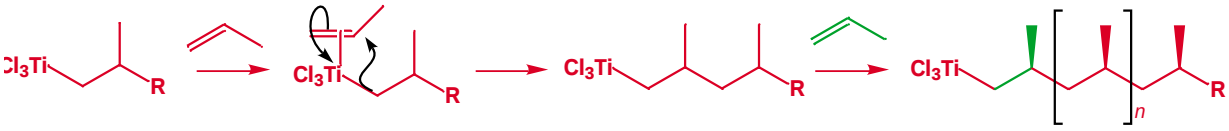



The alkyl-Ti σ complex can form a π complex with the first molecule of propene and then carry out a carbo-titanation of the π bond. This establishes the first C–C bond.

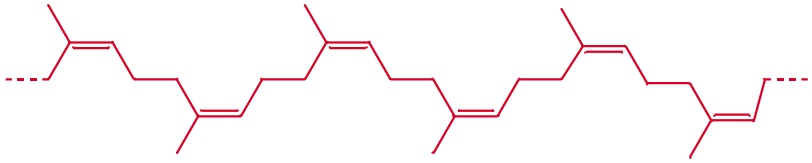

▶ This is a simplification as the catalyst is a solid and the active Ti atom almost certainly Ti(III) rather than Ti(IV) as we have shown here. The third Cl ligand is in fact shared with other Ti atoms in the crystal. Coordination of the active Ti(III) atom must be such that each σ complex is a 16-electron species while the π complexes are 18-electron species.

Insertion of the next propene by a repeat of the previous step now starts the polymerization. Each new C-C bond is formed on the coordination sphere of the Ti atom by transformation of a π complex into a σ complex. Repetition of this process leads to polymerization. We have shown the polymer with isotactic stereochemistry, and this control over the stereochemistry reflects the close proximity of the new propene molecule and the growing polymer.

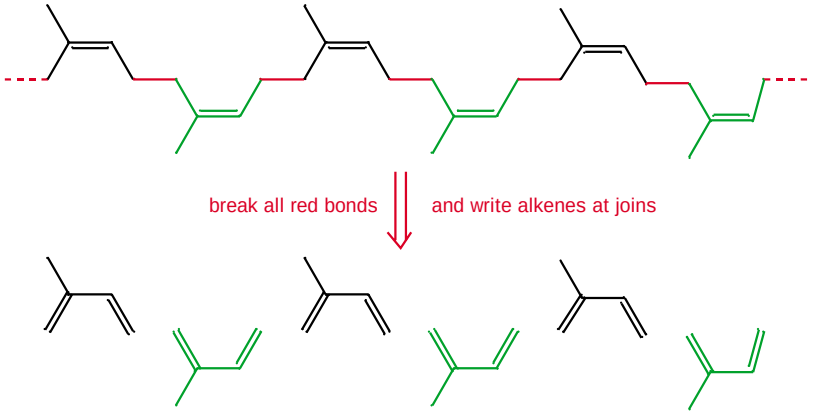
▶ In fact, the reaction can lead either to isotactic or syndiotactic polymer depending on the detailed structure of the catalyst.



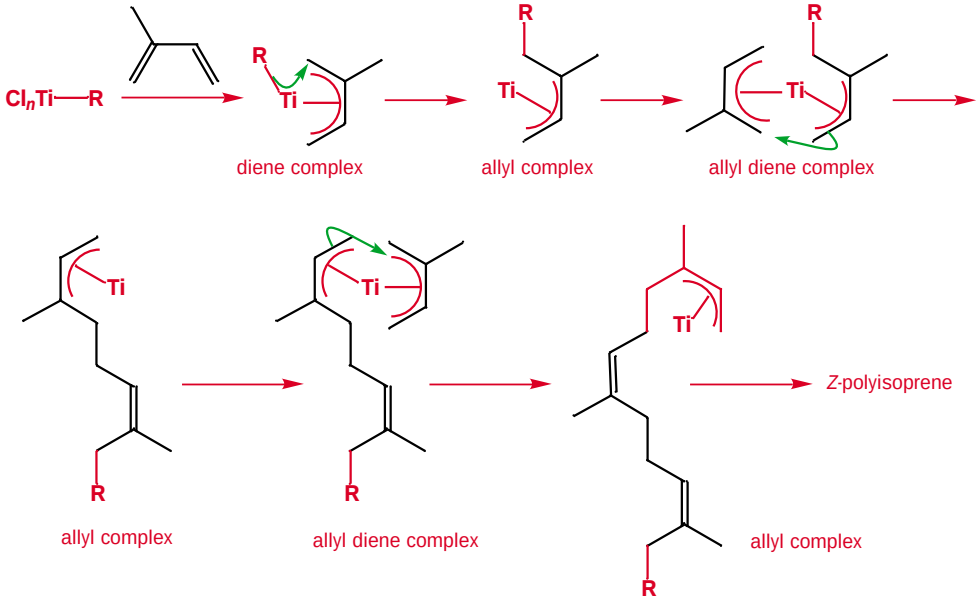

One important elastane polymer that can be made by polymerization in a Ziegler–Natta fashion is rubber. Natural rubber is a polymeric terpene (Chapter 51) made from mevalonic acid and has a branched structure with regular trisubstituted alkenes, which are all in the *Z*-configuration.



Looked at as a polymer, rubber is made up of $C_5$ units joined together by C–C bonds. We should naturally expect to make a hydrocarbon polymer from alkenes, so if we separate these $C_5$ units we find that they are dienes rather than simple alkenes. If you have read Chapter 51, they might be familiar to you as the isoprene units from which terpenes were originally supposed to be made.




The all-*cis* structure of natural rubber is vital to its elasticity. The all-*trans* compound is known and it is hard and brittle. Though dienes such as isoprene can easily be polymerized by cationic methods, the resulting 'rubber' is not all-*cis* and has poor elasticity and durability. However, polymerization of isoprene in the Ziegler–Natta way gives an all-*cis* (90–95% at least) polyisoprene very similar to natural rubber.




One possible explanation is that each isoprene unit adds to the titanium (and we will drop the pretence at this point that we have any idea which other ligands are on the Ti atom) to form an $\eta^4$ diene complex. This must necessarily have the s-*cis* conformation. Addition of R to one end of this complex gives an $\eta^3$ allyl complex still maintaining the *cis* configuration. The next diene then adds to form a new $\eta^4$ diene complex, couples to the allyl complex, and so on. As the chain grows, each diene is added as an $\eta^4$ complex and an all-*cis* polymer results.

## Co-polymerization

If two or more monomers polymerize to give a single polymer containing different subunits, the product is a co-polymer and the process is called co-polymerization. Protein synthesis is an example from nature: amino acids are polymerized stepwise to give proteins of precise sequence and precise length. We can do the same thing chemically providing that we do it in a stepwise fashion—we shall discuss this later. In most cases, chemical co-polymerization cannot be precisely ordered, but still gives useful results.

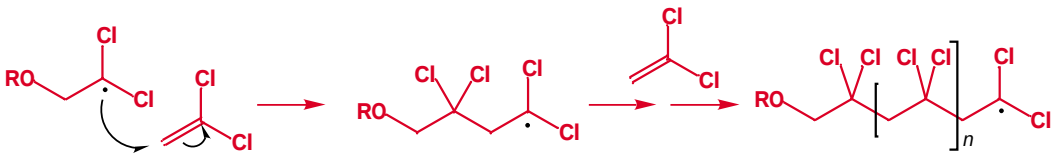
It may have surprised you, when you read the fine print on packaging, that some quite different materials are made out of the same polymer. PVC, for example, is widely used in clothing, 'vinyl' floor and seat coverings, pipework, taps, and lab stopcocks. Some of these applications require strength and rigidity; others flexibility. How is this possible with the same polymer? Some variation can be achieved by the addition of plasticizers—additives that are blended into the polymer mixture but are not chemically bonded to it. Another approach is to use a co-polymer with a smaller amount of a different (but often similar) monomer built randomly into the growing polymer chain. This is quite different from the alternating co-polymers that we saw under carbonyl substitution polymerization, such as nylon 6.6 or the block co-polymers we met a page or two back.

▶ A polymer is a chemical compound while a plastic is a mixture of a polymer and other substances (plasticizers, pigments, fibres, etc.), which allow it to be used in a certain way.

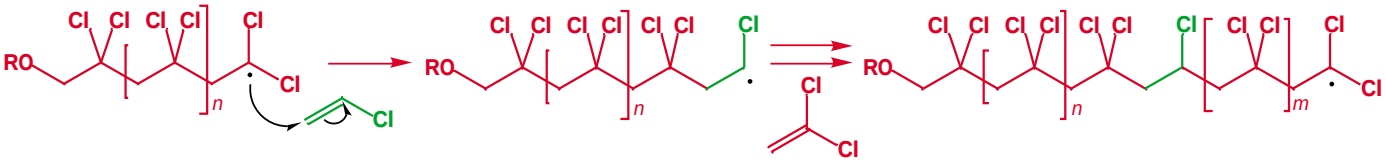
We will choose the example of elastane films for food wrapping—'ClingFilm'. These can be made from poly(vinylidene dichloride) (this is poly(1,1-dichloroethene)) into which a small amount of vinyl chloride is co-polymerized. The method is radical polymerization and the initiator usually a peroxide in aqueous suspension.



Polymerization continues adding vinyl chloride or vinylidene dichloride more or less at random. At first, several dichloroalkene molecules will add, simply because there are more of them.



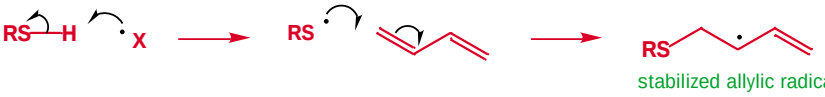
Every now and then a vinyl chloride adds in, followed again by a number of dichloroalkenes to give the co-polymer.



Eventually, polymerization will be terminated by the usual methods and the final co-polymer will have a random mixture of dichloroalkene (mostly) and monochloroalkene, roughly in proportion to their availabilities in the polymerization mixture. The precise properties of the resulting polymer will depend on the ratio of the two monomers.

## Synthetic rubbers can be made by co-polymerization of alkenes and dienes

Radical co-polymerization of styrene and butadiene produces a material that is very like natural rubber. The initiator is a one-electron oxidizing agent, and a thiol (RSH) is used to start the polymerization process. The mixture is about 3:1 butadiene:styrene so there are no long runs of one monomer in the product. We will use butadiene as the starter unit.




The first radical is an allylic radical, stabilized by conjugation with the remaining alkene in the old butadiene molecule. Addition could now occur to another butadiene or to styrene.

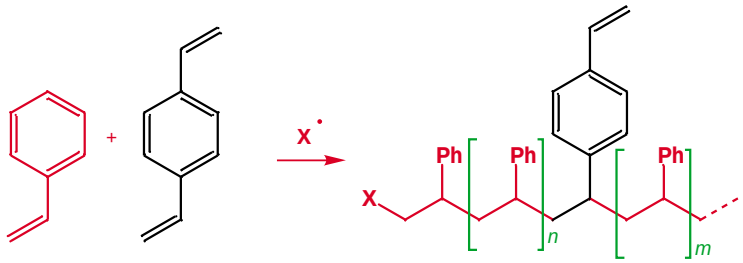

The product is the stabilized benzylic radical with the more stable *trans* double bond. Stabilization of radicals in allylic and benzylic groups is about the same, so the two monomers will react roughly in proportion to their concentration. The final product will be a random co-polymer of about 3:1 buta-

diene to styrene with mostly *E*-alkenes. It is an elastomer used for tyres and other applications where a tough and flexible 'rubber' is needed.

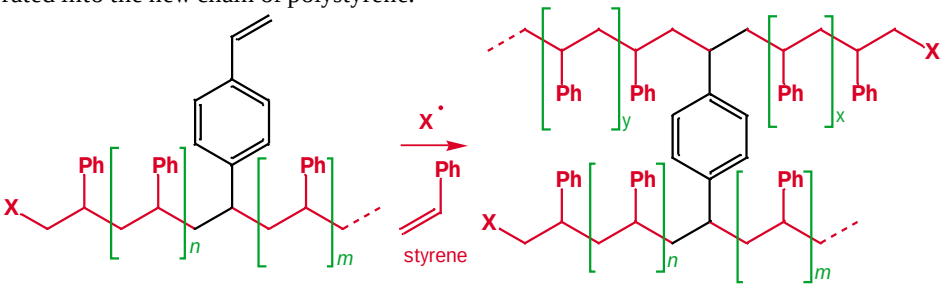
## Cross-linked polymers

Many linear polymers are too flexible to be of use in making everyday objects because they lack the strength, the rigidity, or the elasticity for the job. Linear polymers can be stiffened and strengthened by bonds between the chains. This process is known as cross-linking and we will look now at some ways in which this can be achieved.

All that is really needed is a co-polymer with a small amount of a compound similar to the main monomer but with at least one more functional group than is strictly necessary to form a linear polymer. For example, a small amount of 1,4-divinylbenzene co-polymerized with styrene leads to a linear polymer in which some of the phenyl rings carry a 4-vinyl group.

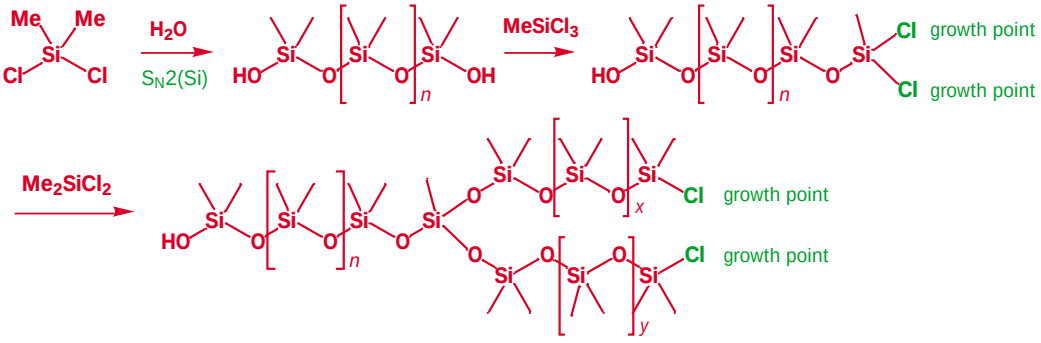



When another chain polymerizes nearby, the spare vinyl group in the first chain may be incorporated into the new chain of polystyrene.

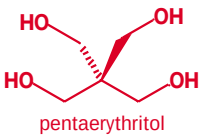



Not all of the spare vinyl groups will be caught up in a new chain of polymerizing styrene, but that need not matter if there are enough of them. It is simply a question of adding enough 1,4-divinyl benzene to get the required degree of cross-linking. These cross-linked styrenes are often made into small beads for polymer-supported reagents, as described below.

Divinyl benzene has two identical 'arms', which become growing points in polymerization. In the polymerization of $Me_2SiCl_2$ we had two growing points (the two chlorine atoms) on each monomer. To get cross-linking we need a third, provided by (a small amount of) $MeSiCl_3$.



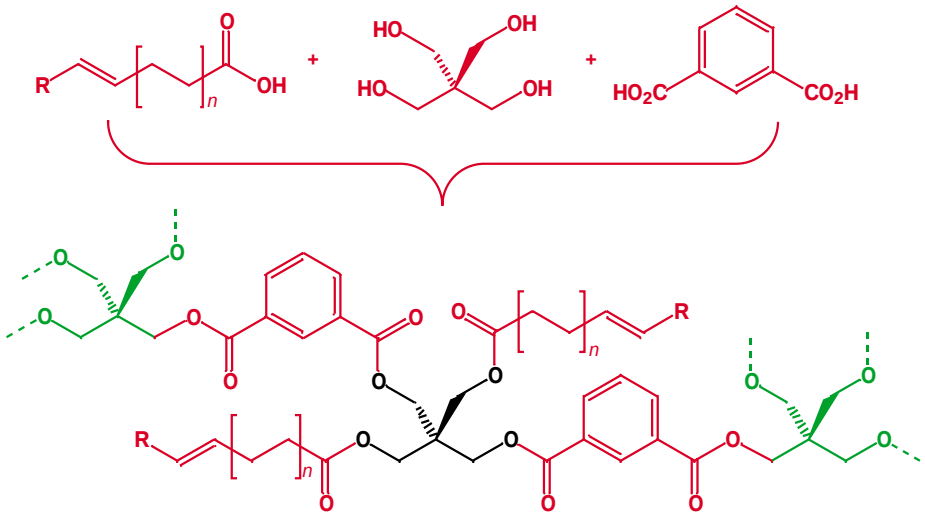
The four-armed cross-linking agent known as pentaerythritol is made from acetaldehyde and formaldehyde in aqueous base. The four arms are arranged in a tetrahedron around a quaternary carbon atom.



pentaerythritol

Co-polymerization of pentaerythritol and two other monomers—an unsaturated acid and benzene 1,3-dicarboxylic acid—gives a network of polymer chains branching out from the quaternary carbon atom at the centre of pentaerythritol. The reaction is simply ester formation by a carbonyl substitution reaction at high temperature (> 200 °C). Ester formation between acids and alcohols is an equilibrium reaction but at high temperatures water is lost as steam and the equilibrium is driven over to the right.

■ Pentaerythritol is made by a Cannizzaro reaction: see Chapter 27, p. 000.



The black pentaerythritol at the centre of the polymer is shown with two each of the ester side chains, though this need not be the case, of course. The green pentaerythritol molecules are the growing points of the network of polymer chains. It is obvious why the benzene dicarboxylic acid is helpful in linking growing points together, but what is the point of the long-chain unsaturated acid? These are naturally occurring acids as described in Chapter 51 and the alkenes are used for further cross-linking under oxidative conditions as described in the next section. Such polymers are called 'alkyd resins' and are used in paints. They form emulsions in water ('emulsion paints') and the ester groups do not hydrolyse under these conditions as water cannot penetrate the polymer network. As the paint 'dries' it is cross-linked by oxygen in the air.

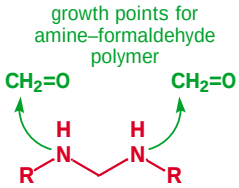
It is not necessary to have quite such a highly branched cross-linking agent to make a network of polymer chains. A triply branched compound is the basis for one of the strongest polymers known—one that we take for granted every time we use the kitchen. It is made by a very simple reaction.

## Melamine

You saw a carbonyl addition reaction forming a polymer right at the beginning of the chapter—the polymerization of formaldehyde. If an amine is added to formaldehyde, condensation to form imines and imine salts occurs readily. These intermediates are themselves electrophilic so we have the basis for ionic polymerization—electrophilic and nucleophilic molecules present in the same mixture. Reaction with a second molecule of amine gives an aminal, the nitrogen equivalent of an acetal.

■ This is also the first step of the Mannich reaction: Chapter 27, p. 000.

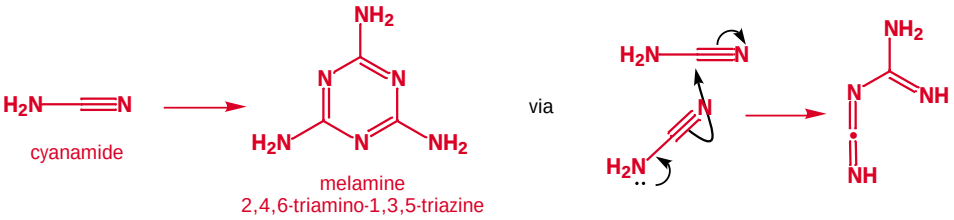

There are now two nucleophilic atoms in the molecule. Each can react with formaldehyde to form more C–N bonds and so on, making two growth points for the polymer.



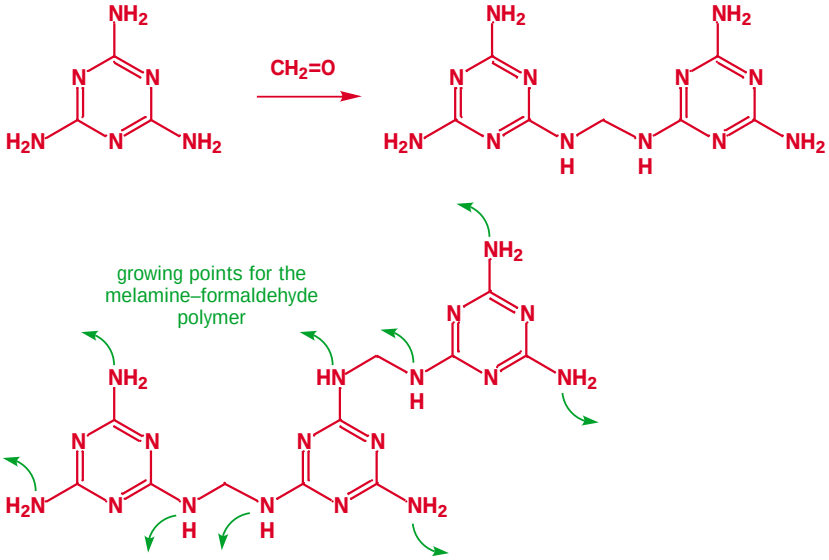

We do better if we have two or even three nucleophilic amino groups present in the same molecule. With three amino groups we will produce a branching polymer of great strength

and the most important of the triamines is melamine. This compound is itself produced by the trimerization of a simple compound, cyanamide $H_2N–CN$, and has given its name to a group of plastics.

■ You should be able to write the full mechanism for the formation of melamine: the first step is given.




When the triamine reacts with formaldehyde, branched polymerization can occur by the same mechanism as the one we drew above for simple amines. Further condensations with formaldehyde allow amines to be attached in many places, and each new amine itself adds many new growing points. An exceptionally strong polymer results.




These resins are used to make ‘unbreakable’ plastic plates and for the famous kitchen surface ‘Formica’. Partly polymerized melamine–formaldehyde mixtures are layered with other polymers such as cellulose (Chapter 49) and phenol–formaldehyde resins and the polymerization is completed under pressure with heat. The result is the familiar, tough, heat-resistant surface.

## Reactions of polymers

We have so far given the impression that all polymers are formed fully armed, as it were, from monomers already having the correct functionality. This is, indeed, often the case because it can be very difficult to persuade polymers to carry out any reactions—reagents cannot penetrate their interiors. Polyester fabrics can be washed without any of the ester linkages being hydrolysed in the washing machine because the water cannot penetrate the fibres. However, some useful reactions, including ester hydrolysis, can be carried out on complete polymers.

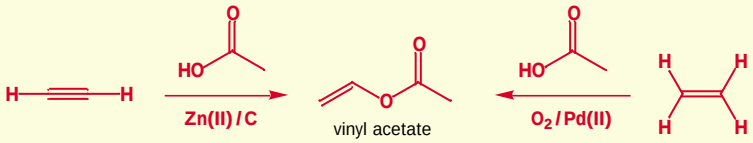

Poly(vinyl alcohol) is an important example. Inspection of the structure reveals that this is a typical alkene polymer but the monomer would have to be vinyl alcohol—the unstable enol of acetaldehyde. The way to make the polymer is to start with something else and only later

convert the polymer product into poly(vinyl alcohol). The most common method of doing this is to use radical polymerization of vinyl acetate, the enol ester of acetaldehyde, and hydrolyse the ester afterwards.

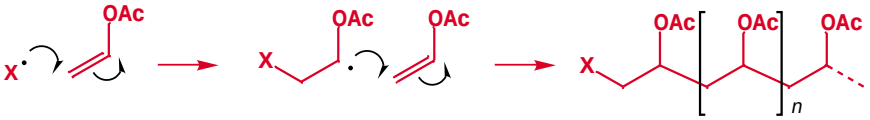
### Vinyl acetate

Vinyl acetate is manufactured on a large scale by two routes.

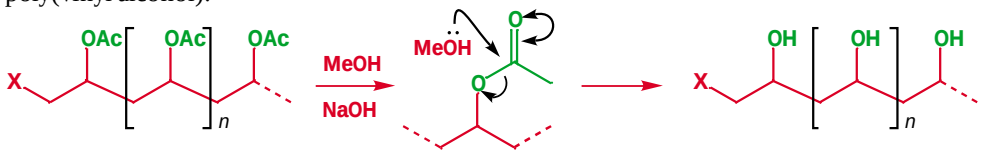
Satisfy yourself that you can at least see what is happening here—if you are stuck on the Pd(II)-catalysed reaction, refer to Chapter 48 and look at oxypalladation and the Wacker reaction for clues.




The polymerization of the enol acetate goes in the usual way.



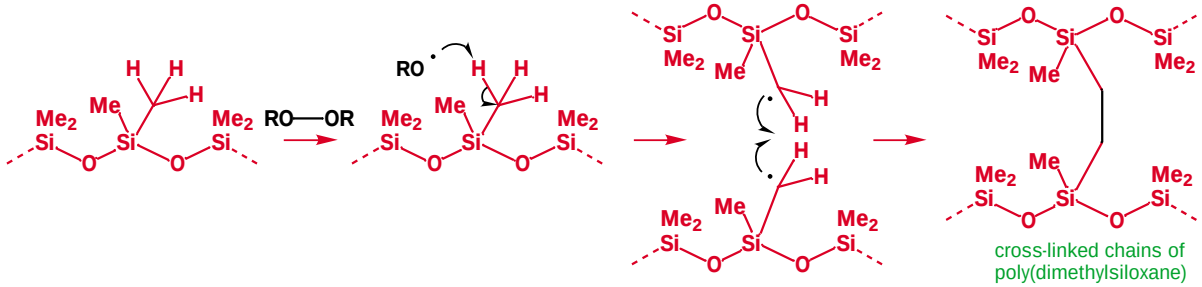

The complete polymer may now be attacked by reagents that cleave the ester groups. Water is a possibility, but methanol penetrates the polymer better and ester exchange in alkaline solution gives poly(vinyl alcohol).




Poly(vinyl alcohol) is soluble in water, unlike almost all other polymers, and that gives it many uses in glues and even as a solublizing agent in chemical reactions to make other polymers. Poly(vinyl acetate) is used in paints.

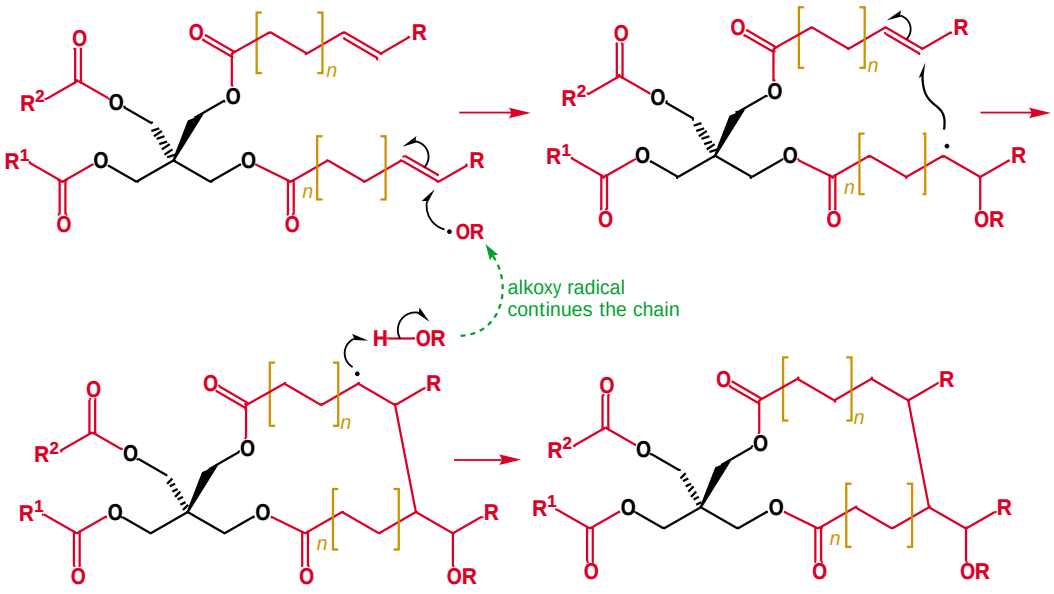
## Cross-linking of pre-formed polymers

We have already discussed cross-linking during polymerization but cross-linking is often carried out after the initial polymer is made. You saw earlier how poly(dimethylsiloxane) can be cross-linked by co-polymerization with $MeSiCl_3$. An alternative way of cross-linking the linear polymer uses radical reactions to convert silicone oil into silicone putty. Peroxides are used in this process.

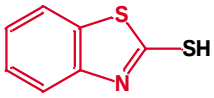



A similar sort of reaction occurs during the cross-linking of alkyd resins for paint manufacture. You may recall that the alkenes are incorporated in these resins for a reason not yet made clear. Now these alkene units come into their own. Oxygen is the reagent and it works by radical dimerization of the chains (see overleaf).

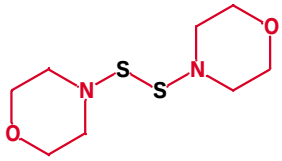
The most important of all of these types of reactions is the vulcanization of rubber. Originally, the raw rubber was just heated with sulfur ($S_8$) and cross-linking of the polyisoprene chains with short chains of sulfur atoms gave it extra strength without destroying the elasticity. Nowadays, a vulcanizing initiator, usually a thiol or a simple disulfide, is added as well. Some examples are



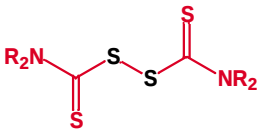

vulcanization initiators



benzothiazole-2-thiol

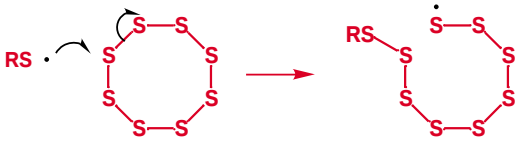


dithio-*bis*-morpholine

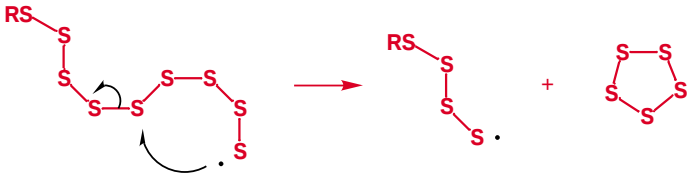


tetraalkyl thiuram disulfide

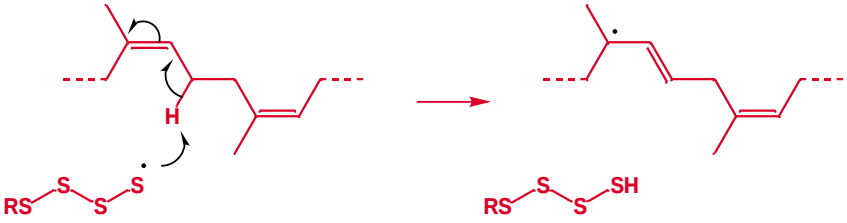
shown in the margin. The thiols give sulfur radicals with oxygen and the disulfides cleave easily as the S–S bond is weak (about 140 kJ mol$^{-1}$ in $S_8$). We will write all these as RS$^{\bullet}$. The initiators either attack the rubber directly or attack sulfur to open the $S_8$ ring.



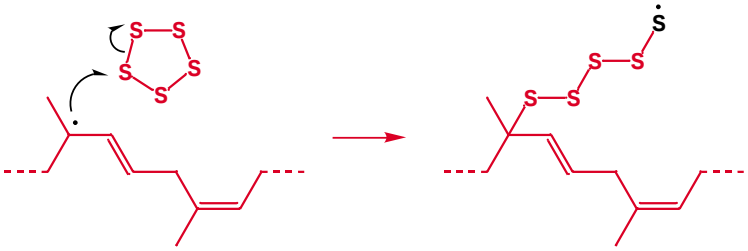
The newly released sulfur radical can bite back on to the sulfur chain and close a ring of 5–7 sulfur atoms, releasing a short chain of sulfur atoms attached to the initiator and terminating in a sulfur radical.



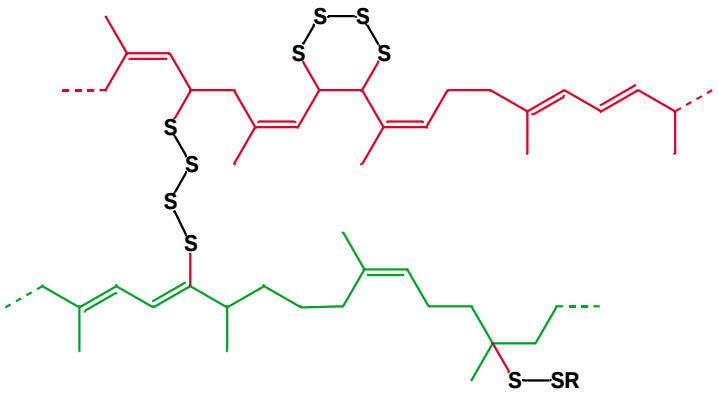
Now the attack on rubber can start. We know that vulcanized rubber has many *E*-alkenes, whereas unvulcanized rubber is all *Z*-alkenes. This suggests that the sulfur radicals do not add to the alkenes but rather abstract allylic hydrogen atoms. Writing only a small section of rubber, we have:



The new allylic radical can do many things, but it might, for example, capture one of the sulfur rings present ($S_5$ to $S_8$). We will use the $S_5$ ring we have just made.



This sulfur radical can attack another chain to give a cross-link or bite back to give a link within the same chain. Many different sulfur links are formed and the next diagram summarizes a part of the vulcanized rubber structure. There is some license here: in reality the links would not be as dense as this, and more than two chains would be involved. Notice the two chains joined by one cross-link, the internal cross-link in the black chain, the attachment of the initiator (RS) to the green chain, and the (*E*,*E*)-dienes in both chains.




We have not given compositions of complete plastics in general, but you might like to know the typical composition of a motor tyre. Notice that the ratio of sulfur to rubber is about 1:40—that gives an idea of how many cross-links there are. Notice also that the rubber contains a great deal of carbon to improve the wear of the rubber. The roles of the other materials are explained in the table.

**Typical composition of rubber motor tyre**

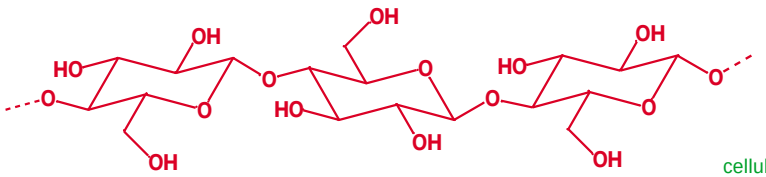
| Component | Parts by weight, % | Function |
|---|---|---|
| rubber | 61 | basic structure |
| carbon black | 27 | reinforcement |
| oils and waxes | 4.9 | processing aid |
| sulfur | 1.5 | vulcanizing agent |
| organic disulfide | 0.4 | accelerates vulcanization |
| zinc oxide (ZnO) | 3 | activates vulcanization |
| stearic acid | 0.6 | activates vulcanization |

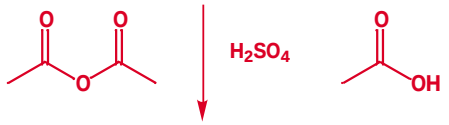
This makes only 98.4% in total and there are small amounts of other materials such as antioxidants to prolong the life of the rubber.

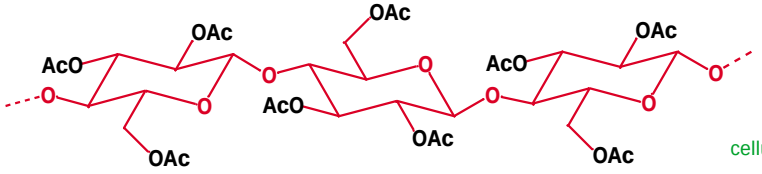
Though synthetic diene polymers have now replaced natural rubber in many applications, they too need to be cross-linked by vulcanization using essentially the same reactions, though the details vary from product to product and from company to company.

## Chemical reactions of cellulose

We met cellulose, the bulk polysaccharide of woody plants, in Chapter 49. It is a strong and flexible polymer but no use for making fabrics or films as it cannot be processed. One solution to this problem is to carry out chemical reactions that transform its properties. Acid-catalysed acetylation with acetic anhydride gives a triacetate with most of the free OH groups converted into esters.

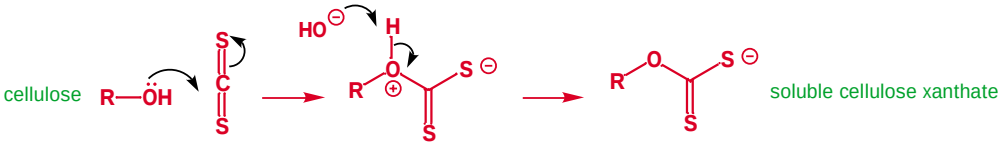


The starting material for this process is wood pulp, cloth, or paper waste and the acetic acid is added first to 'swell' the material and allow it to take up the reagents better. Organic solvents often do this to polymers. The anhydride now carries out the acid-catalysed acetylation and the cellulose triacetate, unlike the cellulose, dissolves in the reaction mixture. The new polymer is often known simply as 'acetate'.

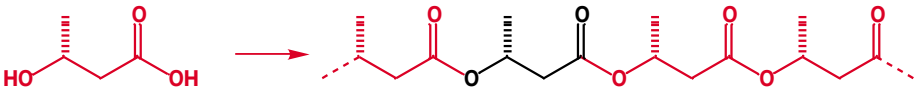
Another cellulose product is rayon. This is really cellulose itself, temporarily modified so that it can be dissolved and processed to give films or fibres. The starting material (from wood, cloth, or paper) is impregnated with concentrated NaOH solution. Addition of $CS_2$ allows some of the OH groups to react to give a 'xanthate' salt that is soluble in water.




Injection of the viscous solution of cellulose xanthate into an acidic ($H_2SO_4$) bath regenerates the cellulose by the reverse of this reaction, as a film or a fibre depending on the process. The result is known as 'cellophane' if it is a film or 'viscose rayon' if it is a fibre.

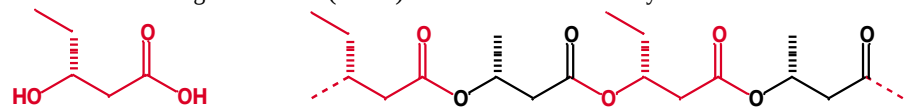
## Biodegradable polymers and plastics

It is necessary to take only a short walk in most cities to see that plastics are not very easily degraded biologically, and it is becoming more important to design plastics, for packaging at least, that have built-in susceptibility to bacteria or fungi. Natural polymers based on proteins and polysaccharides do have that advantage, and one approach is to use a near-natural polymer, poly(hydroxybutyrate) or P(3-HB). This compound is found in some microorganisms as massive (by microorganism standards!) whitish granules occupying substantial parts of the cell—up to 80% of its dry weight of the cell. It seems that it is used as a storage compound (like starch or fat in our case) for excess carbohydrates in the diet.



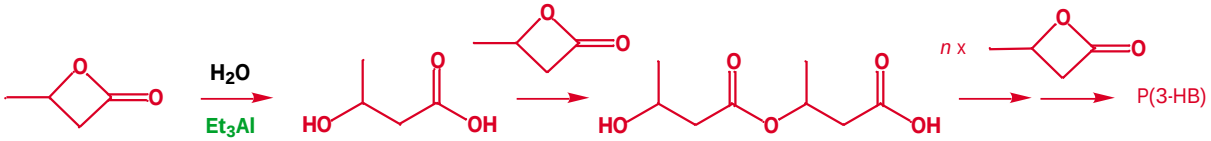
3-hydroxybutyric acid — polyhydroxybutyrate [P(3-HB)]

A co-polymer of P(3-HB) and poly(hydroxyvalerate) P(3-HV) is also found in microorganisms and performs the same function. This polyester forms the basis for a good strong but flexible plastic for containers such as toiletries, and is produced by ICI under the name 'BIOPOL'. Microorganisms must be able to degrade both P(3-HB) and BIOPOL since they themselves use them to store energy.



3-hydroxyvaleric acid (3-hydroxypentanoic acid) — BIOPOL P(3-HB/3-HV)

BIOPOL and the two simple polymers P(3-HB) and P(3-HV) are manufactured by fermentation. They can also be produced chemically by the polymerization of a four-membered lactone (β-butyrolactone). The polymerization is initiated by a water molecule that opens the first lactone ring. The reaction is catalysed by $Et_3Al$ and continues by repeated esterification of the released OH group.




β-butyrolactone

Biological degradation requires that fungi or microorganisms can attack the polymer with their enzymes. This happens efficiently with very few polymers (because these enzymes do not exist) and is, of course, the reason that they are used: people tolerate ugly plastic window frames because they don't rot.

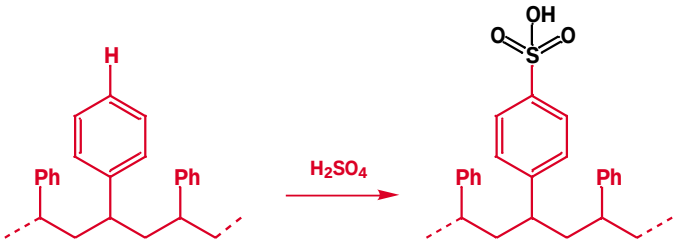
One way in which most polymers do decay is by the action of oxygen in the air and of light. You will be familiar with the way that some polymers go yellow after a time and some become brittle. Coloured plastics, in particular, absorb light and oxygen-induced radical reactions follow. The polymer becomes too cross-linked and loses flexibility. One ingenious application of this natural process helps to degrade the polythene rings that hold cans of beer in packs. These are often discarded and decay quite quickly because some carbon monoxide has been incorporated into the polyethylene to make it more sensitive to photolysis.

## Chemical reagents can be bonded to polymers

We have left this subject to the end of the chapter because it uses all of the principles we have established earlier on. It requires an understanding of radical polymerization, co-polymerization, cross-linking, functionalization of polymers after they have been made, and so on. This is a rapidly growing subject and we can only outline the basics.

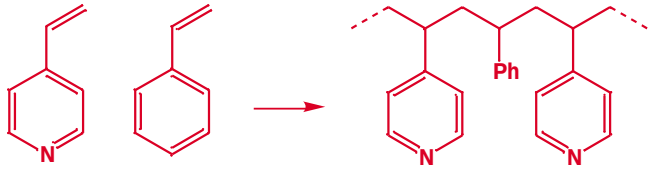
If you are already wondering why anyone would bother to attach reagents to polymers, just think of the problems you have had in the lab in separating the product you want from the other products of the reaction, often the spent reagent and inorganic by-products. If the reagent is attached to a polymer, the work-up becomes easier as the spent reagent will still be attached and can just be filtered off. Polymer-supported reagents can often be reused and their reactions can even be automated.

You may already be familiar with ion-exchange resins and we will start with them. They are commonly based on the co-polymer of styrene and 1,4-divinyl benzene we discussed earlier. The polymerization is carried out in an emulsion in water so that the organic molecules are in tiny droplets. The resulting polymer forms as more or less spherical beads of less than a millimetre in diameter. They can be put through a series of sieves to ensure even sizes if required. The surface of each bead bristles with benzene rings (attached to the polymer backbone) that can be sulfonated in the *para* position just like toluene.

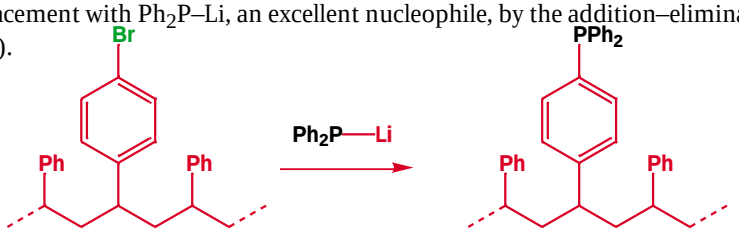



A good proportion of the rings become sulfonated, and the outside of each bead is now coated with strongly acidic sulfonic acid groups. The polymer is an acidic reagent that is not soluble in any normal solvent. It can be packed into a column or simply used as a heterogeneous reagent. In any case, whatever reaction we are doing, there is no difficulty in separating the organic product from the acid.

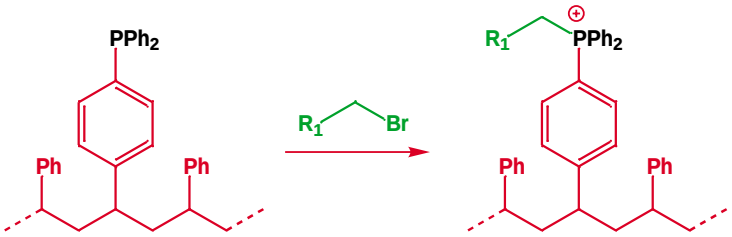
A useful *basic* polymer is made by co-polymerization of 4-vinyl pyridine and styrene.



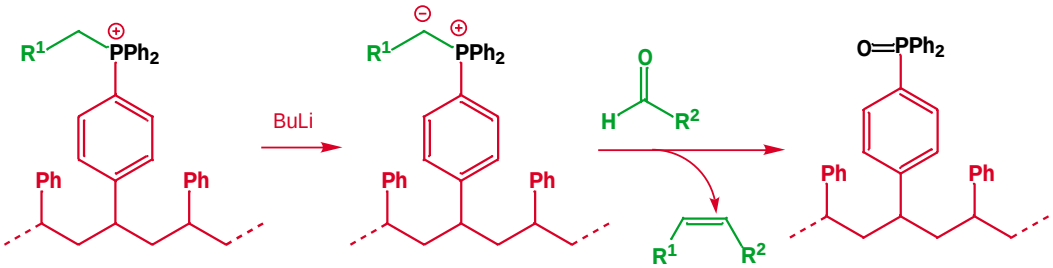

These polymers are reagents in themselves, but a new style of chemistry is being developed around the idea of attaching reagents to the polymer. Poly 4-bromostyrene (or a co-polymer with styrene itself) allows a number of different groups to be attached in the place of the bromine atom. One example is a polymer-bound Wittig reagent. The phosphine can be introduced by nucleophilic displacement with $Ph_2P$–Li, an excellent nucleophile, by the addition–elimination mechanism (Chapter 23).



Though we have shown only one bromine atom and hence only one $Ph_2P$ group on the polymer, almost all of the benzene rings in polystyrene can be functionalized if the bromopolymer is made by bromination of polystyrene in the presence of a Lewis acid. Now the phosphine can be alkylated with an alkyl halide of your choice to form a phosphonium salt, still on the polymer.

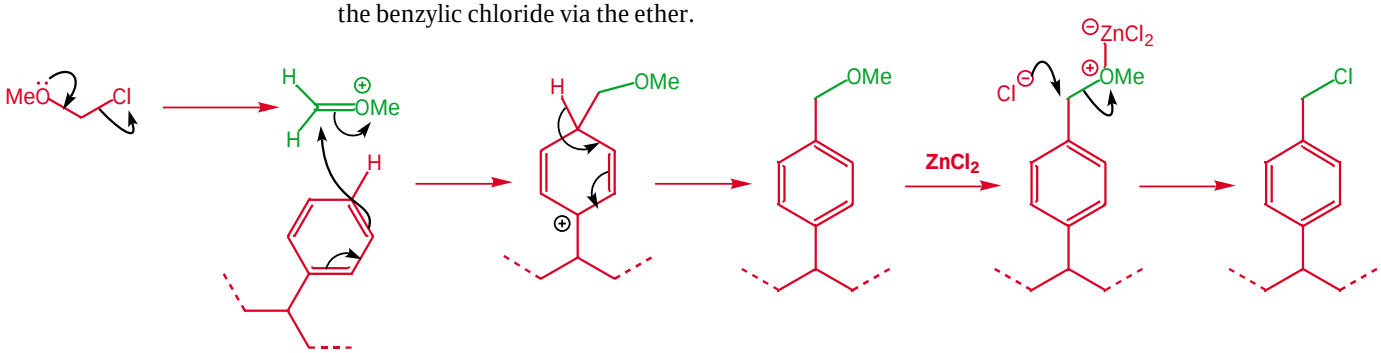



Treatment of the polymer with BuLi and then the aldehyde gives a Wittig reaction (Chapter 31) that releases the alkene product but leaves the phosphine oxide bound to the polymer.

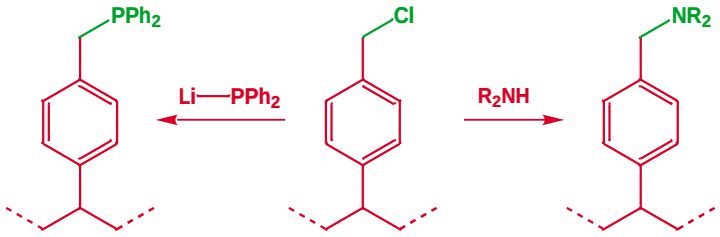



The phosphine oxide can be reduced back to the phosphine (for example, with $Cl_3SiH$) while still bound to the polymer and the polymer-bound reagent can be used again. Separation of $Ph_3P{=}O$ from alkene products after a Wittig reaction can be quite a nuisance so the ease of work-up alone makes this an attractive procedure.

It is not necessary to attach the functional group directly to the benzene ring. There are some advantages in separating the reaction from the polymer by a 'spacer', normally a chain of aliphatic carbon atoms. It may allow reagents to approach more easily and it may allow a higher 'loading' of functional groups per bead. Even a spacer of one $CH_2$ group makes $S_N2$ reactions not only possible but favourable at the benzylic position and the most important of these spacers is introduced by chloromethylation. Reaction of the cross-linked polystyrene with $MeOCH_2Cl$ and a Lewis acid gives the benzylic chloride via the ether.




The chloromethylated resin can now be combined with many different nucleophiles. Amines give basic ion-exchange resins while $Ph_2P$–Li gives a phosphine suitable for complexation to transition metals.



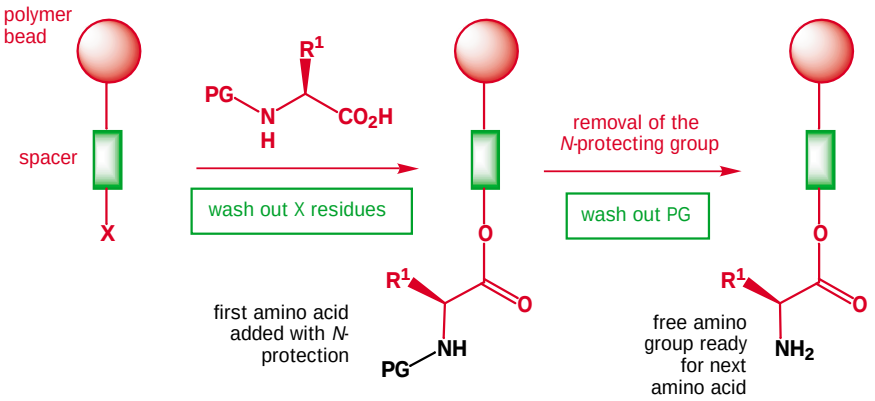
## Automated peptide synthesis uses polymer-bound reagents

Automated polymer-based synthesis comes into its own when a stepwise polymerization is required with precise control over the addition of particular monomers in a specific sequence. This is almost a definition of peptide synthesis. Nature attaches each amino acid to a different 'polymer' (transfer RNA) and uses a 'computer program' (the genetic code) to assemble the polymers in the right order so that the amino acids can be joined together while bound to another polymer (a ribosome). No protection of any functional groups is necessary in this process.

Chemical synthesis of peptides uses a similar approach but our more primitive chemistry has not yet escaped from the need for full protection of all functional groups not involved in the coupling step. The idea is that the first amino acid is attached to a polymer bead through its carboxyl group (and a spacer) and then each *N*-protected amino acid is added in turn. After each addition, the *N*-protection must be removed before the next amino acid is added. The growing peptide chain is attached to the polymer so that all waste products, removed protecting groups, excess reagents, and inorganic rubbish can be washed out after each operation.

■ This subject was introduced in Chapter 25 and we will not repeat here all the details of how protecting groups are added and removed. Please refer to that chapter if you need more explanations of the reactions. We will concentrate here on the role of the polymer.

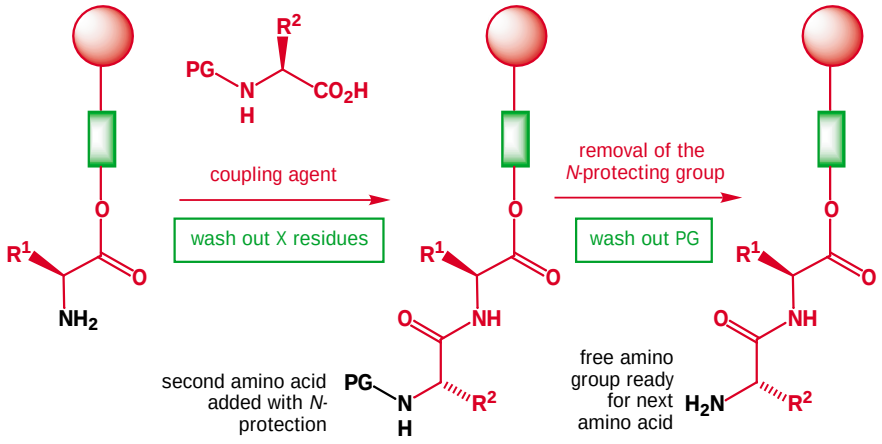
stage 1: attachment of the first (C-terminal) amino acid




Stage 1 involves two chemical reactions—linking the first amino acid to the polymer and removing the *N*-protecting group—and two washing operations. These four steps would take time if everything were in solution but, with the compounds attached to polystyrene beads, they can be carried out simply by packing the beads into a column chromatography-style and passing reagents and solvents through.

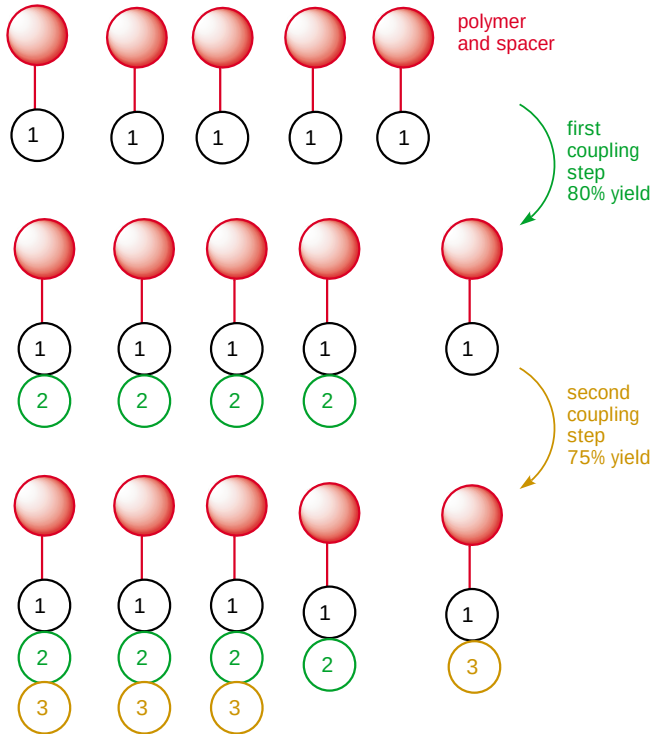
Stage 2 involves the addition of the second *N*-protected amino acid with a reagent to couple it to the free amino group of the amino acid already in place. Removal of the protecting group from the new amino acid is needed, followed by washes, as in stage 1.

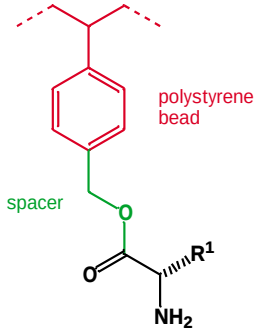
stage 2: formation of the first peptide bond



This process must now be repeated until all of the amino acids have been added. Finally, all the side-chain protecting groups must be removed and the bond joining the peptide chain to the polymer must be broken to give the free peptide. That is the process in outline, but we need now to look at some of the chemistry involved.

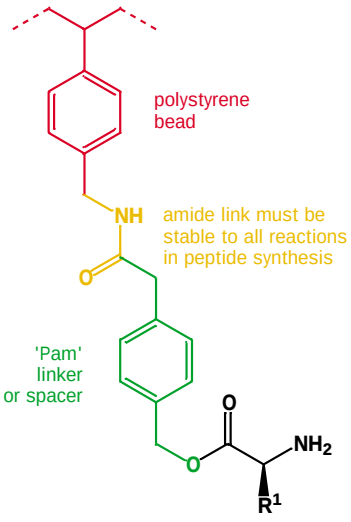
It is obviously important that all reactions are very efficient. Suppose that the coupling step joining the second amino acid on to the first goes in 80% yield. This may not seem bad for a chemical reaction, but it would mean that 20% of the chains consisted of only the first amino acid while 80% contained correctly both first and second. Now what happens when the third amino acid is added?




first amino acid

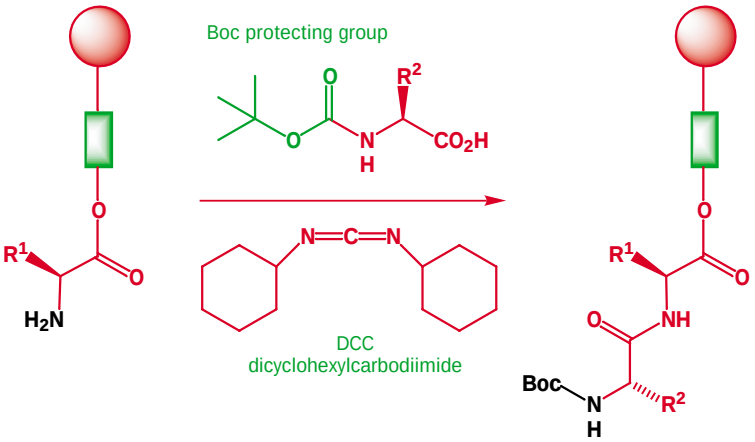
The diagram shows that four out of five growing chains will be right (1–2) after the first coupling step, but after the second (we have put this one at 75% yield for convenience) only three of the five are correct (1–2–3). One of the others has the sequence 1–2 and the other 1–3. This situation will rapidly deteriorate and the final peptide will be a mixture of thousands of different peptides. So, for a start, each reaction must occur in essentially 100% yield. This can be achieved with efficient reactions and an excess of reagents (which are not a problem in polymer-supported reactions as the excess is washed away).



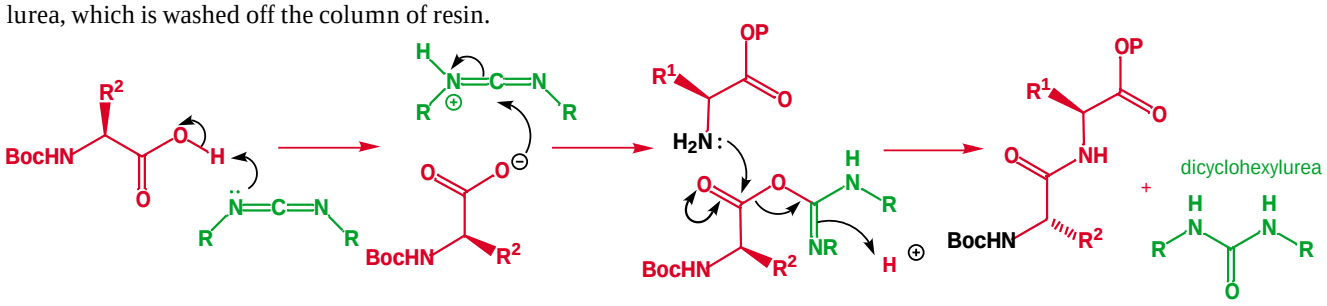

first amino acid

Now some detail—and we will discuss the Merrifield version of peptide synthesis. Spherical cross-linked polystyrene beads of about 50 μm in diameter are used and attached to various spacers of which the simplest is just a $CH_2$ group from the chloromethylated polystyrene we have just discussed. The caesium (Cs) salt of the amino acid is used to displace the chloride as it is a better nucleophile than the Na or K salts. A better alternative is 'Pam' (shown in the margin). It can be used as the nucleophile to displace the chloride first. The amino acid is then added after purification. No chloromethyl groups can remain on the polymer with this spacer.

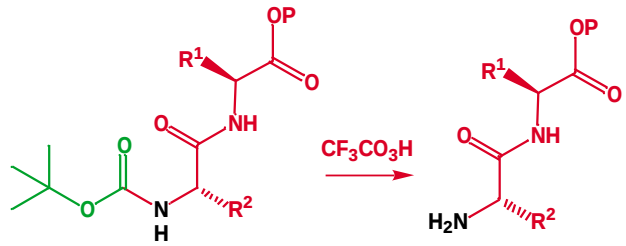
The next stage is to link the carboxyl group of the second amino acid on to the amino group of the first. The Boc group (Chapter 24) is usually used for amino group protection in the Merrifield method and DCC (dicyclohexylcarbodiimide) is used to activate the new amino acid. Here is a summary of this step, using symbols again for polymer and spacer.




The details of the reaction mechanism with DCC were given in Chapter 43, p. 000, and can be shown more easily if we mark the polymer and spacer as 'P' and the cyclohexyl groups as 'R'. The DCC is protonated by the free carboxylic acid and is then attacked by the carboxylate anion. The intermediate is rather like an anhydride with a C=NR group replacing one of the carbonyl groups. It is attacked by the amino group of the polymer-bound amino acid. The by-product is dicyclohexylurea, which is washed off the column of resin.




Now the Boc group must be removed with acid (such as $CF_3CO_2H$ in $CH_2Cl_2$) and washed off the column leaving the free $NH_2$ group of amino acid number two ready for the next step.



The mechanism of this reaction is discussed in Chapter 25.

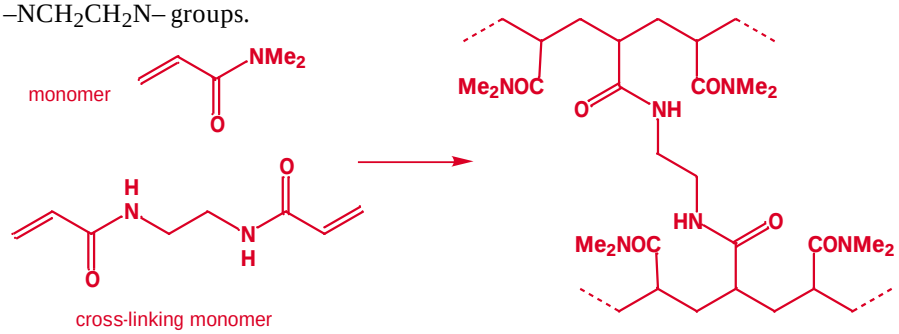
The synthesis continues with repetition of these two steps until the peptide chain is complete. The peptide is cleaved from the resin, usually with HF in pyridine or $CF_3SO_2OH$ in $CF_3CO_2H$ and given a final purification from small amounts of peptides of the wrong sequence by chromatography, usually HPLC.

This process is routinely automated in commercially available machines. Solutions of all of the protected amino acids required are stored in separate containers and a programmed sequence of coupling and deprotection leads rapidly to the complete peptide in days rather than the years needed for solution chemistry. The most dramatic illustration of this came with the publication of a heroic traditional synthesis of bovine pancreatic ribonuclease A (an enzyme with 124 amino acids) by Hirschmann, side-by-side with one by Merrifield using functionalized polystyrene as we have described. The traditional method required 22 co-workers, while the Merrifield method needed only one.

## Peptide synthesis on polyacrylamide gel

Another method of polymer-supported peptide synthesis has been developed by Sheppard. Most things are different in this approach, which is better adapted for polar solvents and automated

operation. The polymer is a polyacrylamide cross-linked with bis-acrylamides joined by $-NCH_2CH_2N-$ groups.

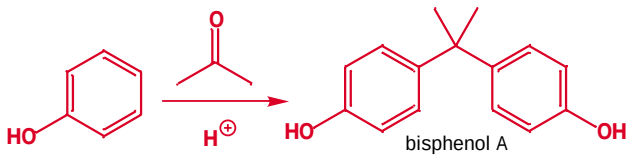



Polar solvents such as water or DMF penetrate the beads, making them swell much more than do the polystyrene resins. This exposes more reactive groups and increases the loading of peptide chains on each bead. The first amino acid is attached through its carboxyl group to an amino group on the polymer, added during or after polymerization by incorporating more 1,2-diaminoethane. The favoured amino protecting group is now Fmoc (see Chapter 24), which has the advantage that it can be removed under basic conditions (piperidine) which do not affect acid-labile side-chain protecting groups.

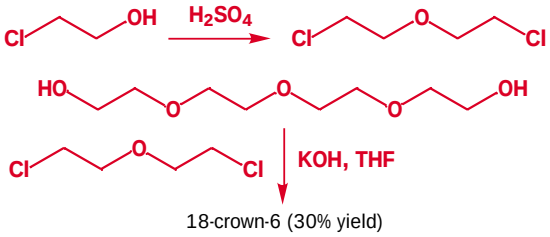
Methods like these have made polymer-supported synthesis so valuable a method that it is now being developed for many reactions old and new. A recent (1999) issue of the journal *Perkin Transactions 1* reported two syntheses of natural products in which every step was carried out using a polymer-supported reagent. Polymers are vital to us in everyday life in a multitude of ways and new polymers are being invented all the time. We have done no more than scratch the surface of this subject and you should turn to more specialized books if you want to go further.

## Problems

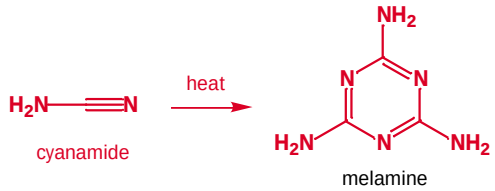
**1.** The monomer bisphenol A is made by the following reaction. Suggest a detailed mechanism.

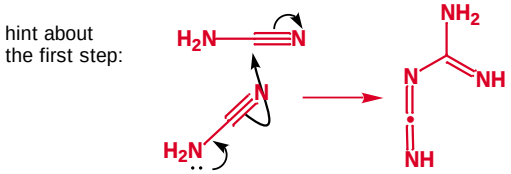



**2.** An alternative synthesis of 18-crown-6 to the one given in the chapter is outlined below. How would you describe the product in polymer terms? What is the monomer? How would you make 15-crown-5?

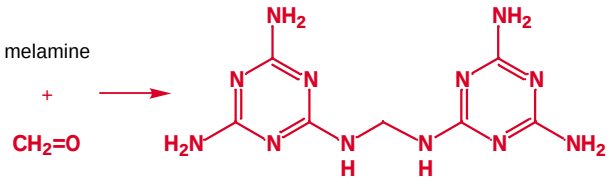



**3.** Melamine is formed by the trimerization of cyanamide and a hint was given in the chapter as to the mechanism of this process. Expand that hint into a full mechanism.

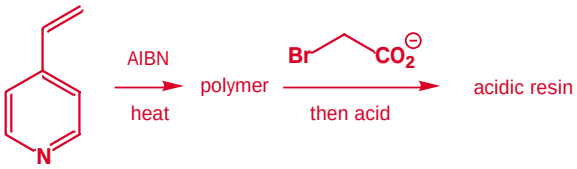



Melamine is polymerized with formaldehyde to make formica. Draw a mechanism for the first step in this process.

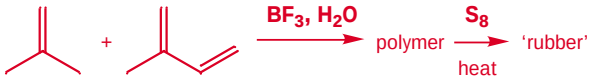



**4.** An acidic resin can be made by the polymerization of 4-vinylpyridine initiated by AIBN and heat followed by treatment of the polymer with bromoacetate. Explain what is happening and give a representative part structure of the acidic resin.



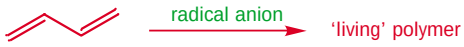

**5.** An artificial rubber may be made by cationic polymerization of isobutene using acid initiation with $BF_3$ and water. What is the mechanism of the polymerization, and what is the structure of the polymer?


This rubber is too weak to be used commercially and 5–10% isoprene is incorporated into the polymerizing mixture to give a different polymer that can be cross-linked by heating with sulfur (or other radical generators). Draw representative structures for sections of the new polymer and show how it can be cross-linked with sulfur.

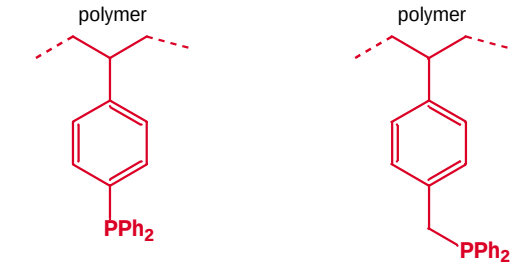



**6.** When sodium metal is dissolved in a solution of naphthalene in THF, a green solution of a radical anion is produced. What is its structure?


This green solution initiates the polymerization of butadiene to give a 'living polymer'. What is the structure of this polymer and why is it called 'living'?



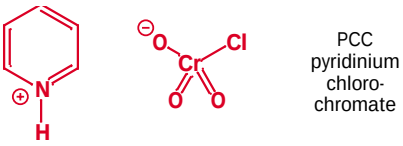

**7.** We introduced the idea of a spacer between a benzene ring (in a polystyrene resin) and a functional group in the chapter. If a polymer is being designed to do Wittig reactions, why would it be better to have a $Ph_2P$ group joined directly to the benzene ring than to have a $CH_2$ spacer between them?



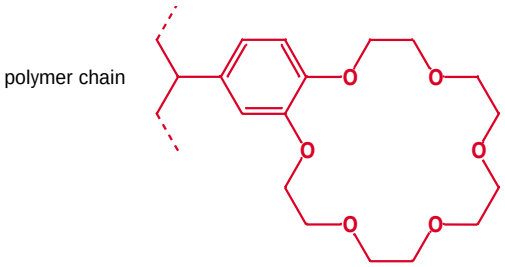

useful for Wittig reactions useless for Wittig reactions

If you need a hint, draw out the reagents that you would add to the polymer to do a Wittig reaction and work out what you would get in each case.

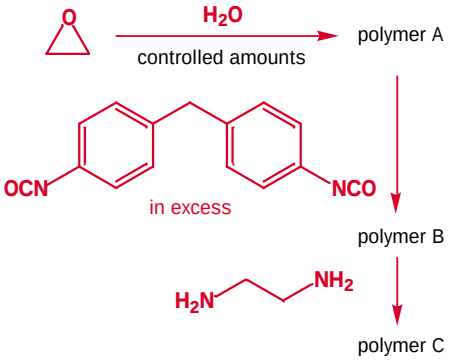
**8.** A useful reagent for the oxidation of alcohol is 'PCC' (pyridinium chlorochromate). Design a polymeric (or at least polymer-bound) reagent that should show similar reactivity. What would be the advantage of the polymer-bound reagent over normal PCC?




**9.** A polymer that might bind specifically to metal ions and be able to extract them from solution would be based on a crown ether. How would you make a polymer such as this?

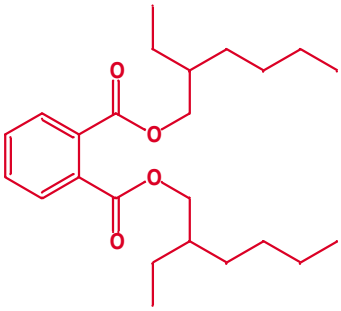



**10.** What is a 'block co-polymer'? What polymer would be produced by this sequence of reactions? What special physical properties would it have?




**11.** Why does polymerization occur only at relatively low temperatures often below 200 °C? What occurs at higher temperatures? Formaldehyde polymerizes only below about 100 °C but ethylene still polymerizes up to about 500 °C. Why the difference?

**12.** Poly(vinyl chloride) (PVC) is used for rigid structures like window frames and gutters with only small amounts of additives such as pigments. If PVC is used for flexible things like plastic bags, about 20–30% of dialkyl phthalates such as the compound below are incorporated during polymerization. Why is this?



# Organic chemistry today

# 53

## Connections

**Building on:**

- The rest of the book ch1–ch52

**Arriving at:**

- How organic chemistry produced an AIDS treatment in collaboration with biologists

**Looking forward to:**

- Life as a chemist

## Modern science is based on interaction between disciplines

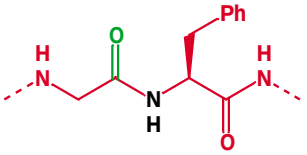
Organic chemistry has transformed the materials of everyday life, as we have seen in Chapter 52, but this is merely a glimpse of the future of organic materials where light-emitting polymers, polymers that conduct electricity, self-reproducing organic compounds, molecules that work (nano-engineering), and even molecules that think may transform our world in ways not yet imagined. These developments are the result of cooperation between organic chemists and physicists, engineers, material scientists, computer experts, and many others.

The most dramatic developments at the beginning of the twenty-first century are new methods in medicine from collaborations between organic chemists and biologists. (The biochemical background is sketched out in Chapters 49–51.) The media's favourite 'a cure for cancer' is already not just '*a* cure' but hundreds of successful cures for the hundreds of diseases collectively called 'cancer'. A newspaper headline in 1999 revealed that there was *some* chance of survival for all known types of childhood cancer. We are going to discuss just one equally dramatic medical development, the treatment of AIDS. Like the treatment of cancer, this is a story that is only just starting, but enough is known to make it a gripping story full of hope.

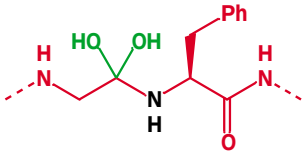
When AIDS (Acquired Immune Deficiency Syndrome) first came into the news in the 1980s it was a horror story of mysterious deaths from normally harmless diseases after the patient's immune system had been weakened and eventually destroyed. The cause was identified by biologists as a new virus: HIV (Human Immunodeficiency Virus) and antiviral drugs, notably AZT (Chapter 49), were used with some success. These drugs imitate natural nucleosides (AZT imitates deoxythymidine) and inhibit the virus from copying its RNA into DNA inside human cells by inhibiting the enzyme 'reverse transcriptase'.

These drugs also inhibit our own enzymes and are very toxic. Biologists then discovered an alternative point of attack. An enzyme unique to the virus cuts up long proteins into small pieces essential for the formation of new HIV particles. If this enzyme could be inhibited, no new viruses would be formed, and the inhibitor should not damage human chemistry. Several companies invented HIV protease inhibitors, which looked more like small pieces of proteins with the weak link of the amide bond replaced by a more stable C–C bond.

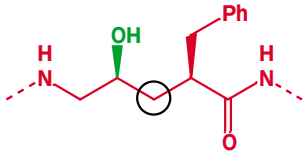
Real peptides are usually poor drugs because we have our own peptidases which quickly cut up ingested proteins into their constituent amino acids by hydrolysis of the amide link. Drugs that imitate peptides may avoid this ignominious fate by replacing the amide bond with another bond less susceptible to hydrolysis. This part structure of one HIV protease inhibitor makes the point.



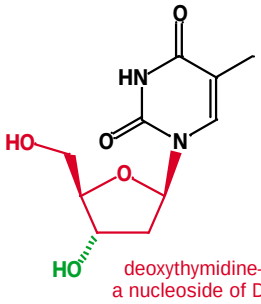

section of protein



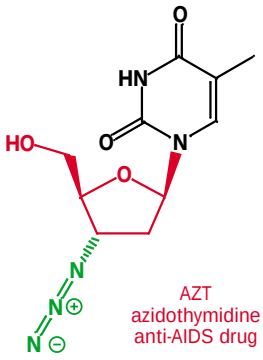

intermediate in amide hydrolysis




section of inhibitor




deoxythymidine—
a nucleoside of DNA

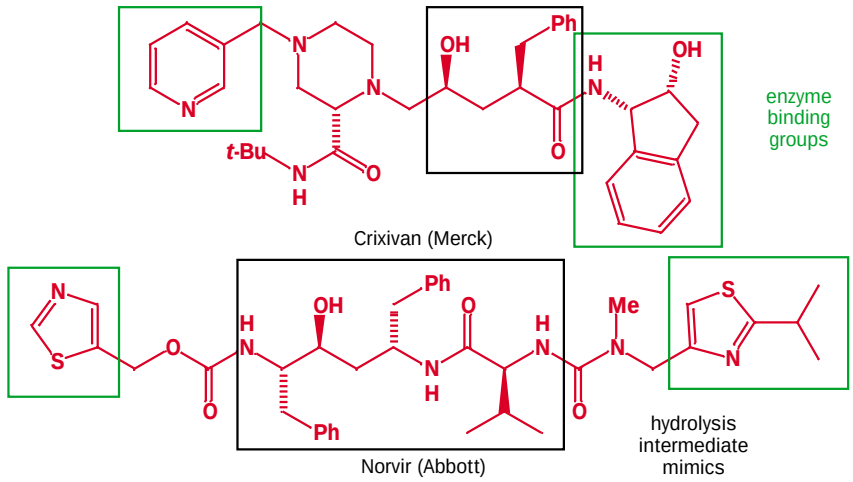


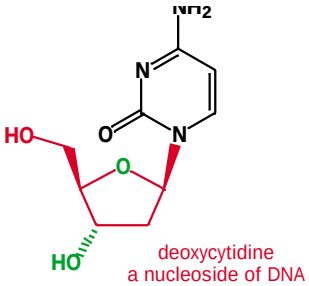

AZT
azidothymidine
anti-AIDS drug

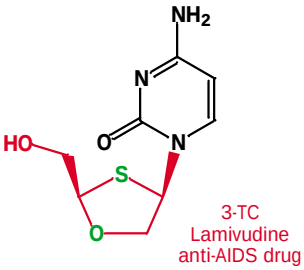
On the left is a section of normal protein with glycine and phenylalanine residues (Chapter 49). In the middle is the intermediate formed when a molecule of water attacks the amide carbonyl group. On the right is a piece of the HIV protease inhibitor. The amide nitrogen atom has been replaced by a $CH_2$ group (ringed in black) so that no 'hydrolysis' of the C–C bond can occur. The inhibitor may bind but it cannot react.

Enzymes ideally bind their substrates strongly and the product of the reaction much more weakly. If they are to accelerate the reaction they need to lower the energy of the transition state (Chapters 13 and 41) and they can do this by binding the transition state of the reaction strongest of all. We cannot literally synthesize a transition state analogue because transition states are by definition unstable, but intermediate analogues can be synthesized. The inhibitor above has one OH group instead of the two in the genuine intermediate but this turns out to be the vital one. This knowledge was acquired from an X-ray crystal structure showing how the enzyme binds the substrate. The inhibitor binds well to the enzyme but cannot react so it blocks the active site.

These compounds are a good deal more sophisticated than this simple analysis suggests. For example, HIV protease is a dimeric enzyme and experience with this class of protease suggested correctly that more or less symmetrically placed heterocyclic rings (Chapters 42–44) would greatly improve binding. Here are two of the inhibitors with the active site binding portion framed in black and the heterocyclic binding portions framed in green.



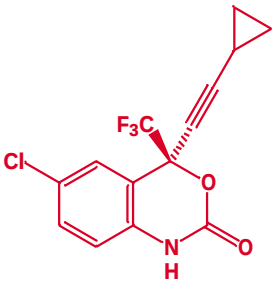




These developments looked so promising that Merck even set up a completely new research station at West Point, Pennsylvania, dedicated to this work. The biochemist in charge, Dr Irving Segal, was one of the victims of the Lockerbie bombing in 1988 but his work lives on as Crixivan (indinavir) is now one of the cocktail of three drugs (AZT and 3TC, shown with the nucleoside it imitates, are the others) that has revolutionized the treatment of HIV. Before this treatment most HIV victims were dead within 2 years. Now no one knows how long they will survive as the combination of the three drugs reduces the amount of virus below detectable levels.

Crixivan was not the first compound that Merck discovered. Many others fell by the wayside because they were not active enough, were too toxic, didn't last long enough in the body, or for other reasons. Crixivan was developed from cooperation between biochemists, virologists, X-ray crystallographers, and molecular modellers as well as organic chemists. When the choice of Crixivan from the various drug candidates had been made and the chemists were trying to make enough of it for trials and use, theirs was an exceptionally urgent task. They knew that a kilo of compound was needed to keep each patient alive and well for a year. Merck built a dedicated plant for the manufacture of Crixivan at Elkton, Virginia, in 1995. Within 1 year, production was running at full blast and there are thousands of people alive today as a result.

The AIDS crisis led to cooperation between the pharmaceutical companies unparalleled since the development of penicillin during the Second World War. Fifteen companies set up an AIDS drug development collaboration programme and government agencies and universities have all joined in.

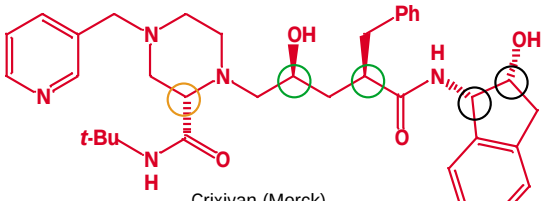
The battle is not yet won, of course, but the HIV protease inhibitors are being followed by a new generation of nonnucleoside reverse transcriptase inhibitors, which promise to be less toxic to humans. An example is the DuPont–Merck compound DMP-266, made as a single enantiomer and now under clinical trials. This compound, though it contains a most unusual cyclopropane and alkyne combination, is nevertheless a much simpler compound than Crixivan. We shall devote most of this final chapter to the synthesis of the established and chemically more interesting drug Crixivan.



DMP-266

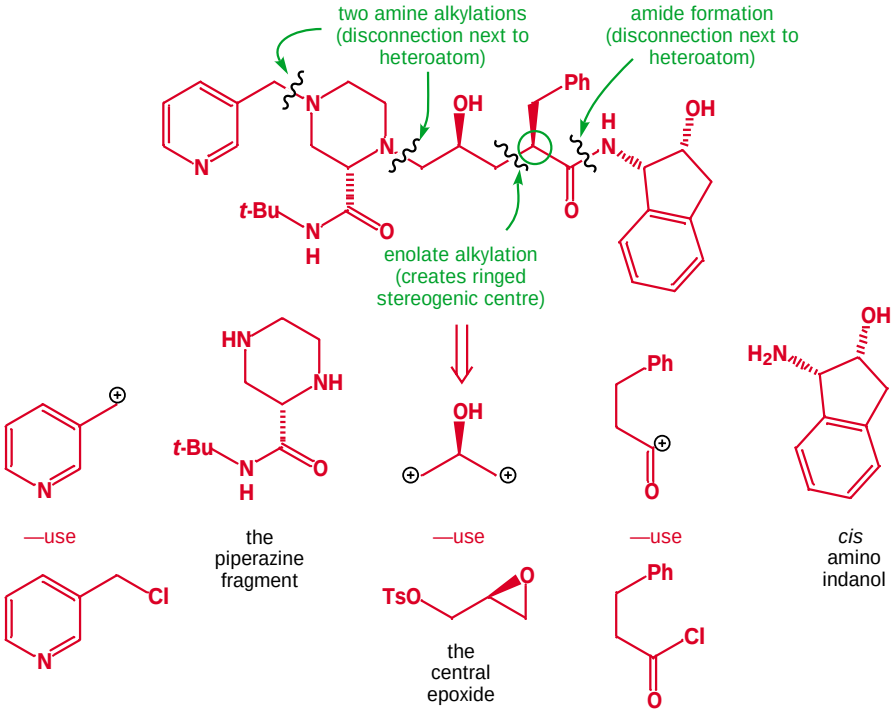
## The synthesis of Crixivan

Crixivan is a formidable synthetic target. It is probably the most complex compound ever made in quantity by organic synthesis and very large amounts must be made because one kilo is needed per patient per year. The complexity largely arises from the stereochemistry. There are five stereogenic centres, marked with coloured circles on this diagram, and their disposition means that three separate pieces of asymmetric synthesis must be devised. There are, of course, also many functional groups and four different rings.



Crixivan (Merck)
stereogenic centres marked with circles

The two black centres are 1,2-related and we have already discussed them in part at the end of Chapter 41. The green centres are 1,3-related and we saw in Chapter 45 that this type of control is possible though difficult. The orange centre is 1,4-related to the nearer green centre and must be considered separately.

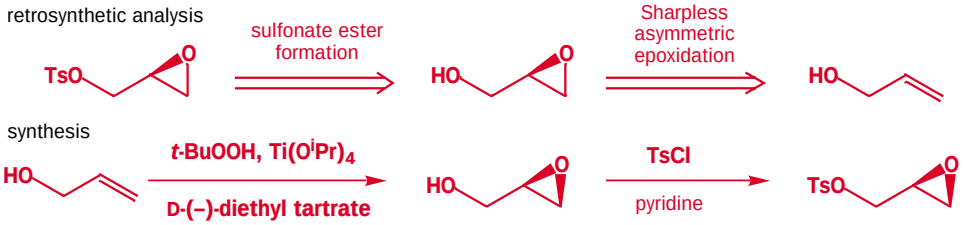



The challenge with Crixivan, as with any drug, is to make it efficiently—high yields; few steps. It has five stereogenic centres, so the chemists developing the synthesis needed to address the issue of diastereoselectivity. And it is a single enantiomer, so an asymmetric synthesis was required. We can start by looking at some likely disconnections, summarized in the scheme above. They are all disconnections of the sorts you met in Chapter 30, and they all correspond to reliable reactions.

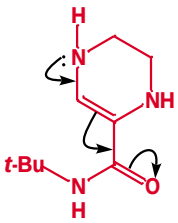
These disconnections split the molecule into five manageable chunks (synthons), three of which contain stereogenic centres and will have to be made as single enantiomers. The final stereogenic

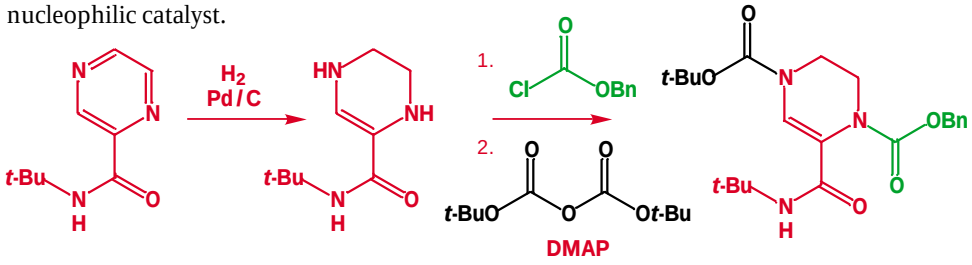
centre (ringed in the disconnection diagram) would have to be made in the enolate alkylation step, so this step will have to be done diastereoselectively.

Let's take these three chiral synthons in turn. First, the simplest one: the central epoxide. The reagent we need here will carry a leaving group, such as a tosylate, and it can easily be made from the epoxy-alcohol. This gives a very good way of making this compound as a single enantiomer—a Sharpless asymmetric epoxidation of allyl alcohol.



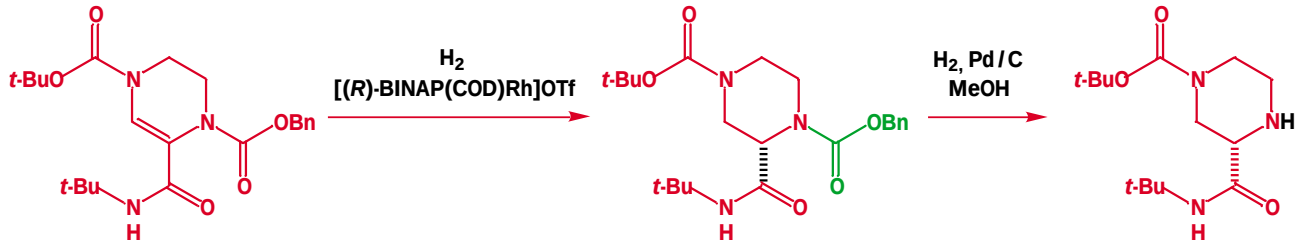

Next, the piperazine fragment. This has two nucleophilic nitrogen atoms and they will both need protecting with different protecting groups to allow them to be revealed one at a time. It will also need to be made as a single enantiomer. In an early route to Crixivan, this was done by resolution, but enantioselective hydrogenation provides a better alternative. Starting from a pyridine derivative, a normal hydrogenation over palladium on charcoal could be stopped at the tetrahydropyrazine stage. The two nitrogens in this compound are quite different because one is conjugated with the amide while one is not (the curly arrows in the margin show this). The more nucleophilic nitrogen—the one *not* conjugated with the amide—was protected with benzyl chloroformate to give the Cbz derivative. Now the less reactive nitrogen can be protected with a Boc group, using DMAP as a nucleophilic catalyst.





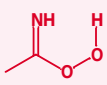

You met asymmetric hydrogenation using BINAP–metal complexes in Chapter 45 as a method for the synthesis of amino acids. The substrate and catalyst are slightly different here, but the principle is the same: the chiral ligand, BINAP, directs addition of hydrogen across the double bond with almost perfect enantioselectivity and in very high yield. In Chapter 45 we described this as addition to one enantiotopic face of the alkene. A further hydrogenation step allowed selective removal of the Cbz group, preparing one of the two nitrogen atoms for alkylation.

■ COD = cyclooctadiene.



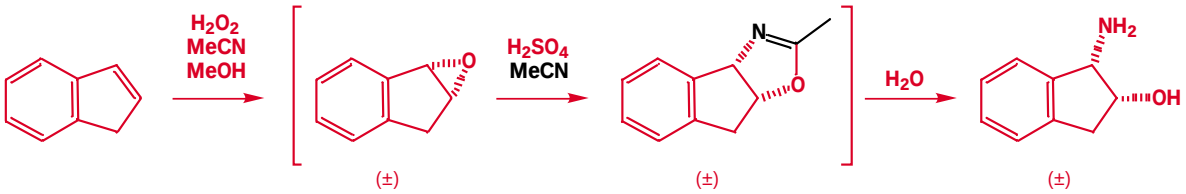

96% yield; 99% ee

▶ $H_2O_2$ and MeCN react to give a 'peroxyimidic acid'—the C=N analogue of a peroxy-acid—as the true epoxidizing agent.



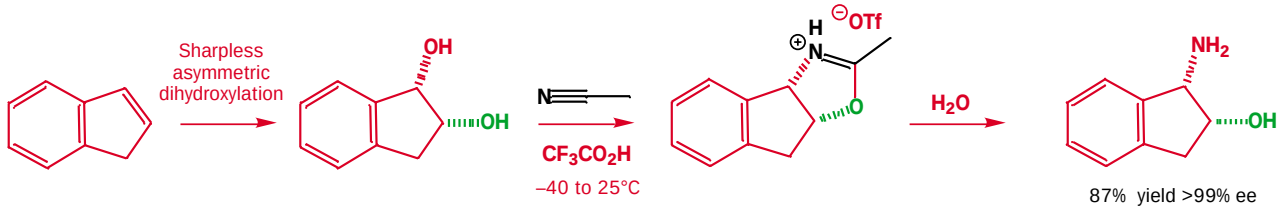
peroxyimidic acid

The remaining chiral fragment is a compound whose synthesis was discussed in Chapter 41, and you should turn to p. 000 for more details of the mechanisms in the reaction sequence. It can be made on a reasonably large scale (600 kg) in one reaction vessel, starting from indene. First, the double bond is epoxidized, not with a peroxy-acid but with the cheaper hydrogen peroxide in an acetonitrile–methanol mixture. Acid-catalysed opening of the epoxide leads to a cation, which takes part in a reversible Ritter reaction with the acetonitrile solvent, leading to a single diastereoisomer of a heterocyclic intermediate which is hydrolysed to the amino-alcohol.

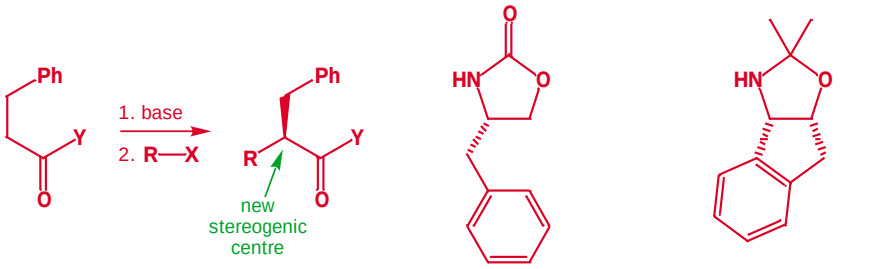



The product is, of course, racemic but, as it is an amine, resolution with an acid should be straightforward. Crystallization of its tartrate salt, for example, leads to the required single enantiomer in 99.9% ee. With such cheap starting materials, resolution is just about acceptable, even though it wastes half the material. It would be better to oxidize the indene enantioselectively, and retain the enantiomeric purity through the sequence: it is indeed possible to carry out a very selective Sharpless asymmetric dihydroxylation (Chapter 45) of indene, and the diol serves as an equally good starting material for the Ritter reaction. The stereogenic centre carrying the green hydroxyl group remains firmly in place throughout the route, and controls the absolute configuration of the final product.

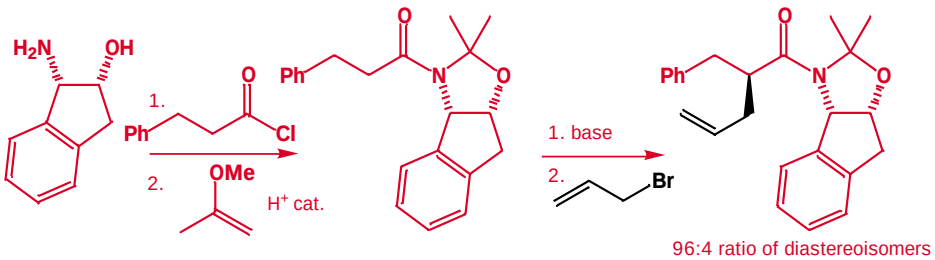
■ The Ritter reaction was described in Chapter 16, p. 000. The reason for the formation of the *cis* diastereoisomer in this example is discussed in Chapter 41, p. 000.

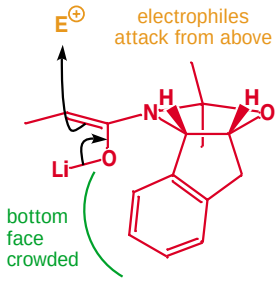



Both resolution and Sharpless asymmetric dihydroxylation were successful in the synthesis of Crixivan but the best method is one we shall keep till later. Only one stereogenic centre remains, and its stereoselective formation turns out to be the most remarkable reaction of the whole synthesis. The centre is the one created in the planned enolate alkylation step.

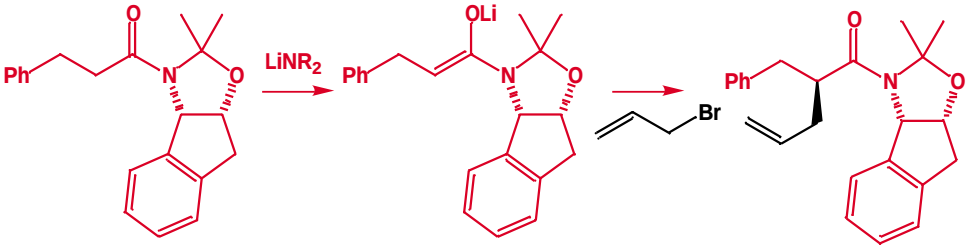



The obvious way to make this centre is to make Y a chiral auxiliary; the required acyl chloride could be used to acylate the auxiliary, which would direct a diastereoselective alkylation, before being removed and replaced with the amino-alcohol portion. But the amino-alcohol itself, certainly once protected, has a remarkable similarity to Evans' oxazolidinone auxiliaries (Chapter 45), and it turns out that this amino-alcohol will function very successfully as a chiral auxiliary, which does not need to be removed, avoiding waste and saving steps! The amino-alcohol was acylated with the acyl chloride, and the amide was protected as the nitrogen analogue of an acetonide by treating with 2-methoxypropene (the methyl enol ether of acetone) and an acid catalyst. The enolate of this amide reacts highly diastereoselectively with alkylating agents, including, for example, allyl bromide.



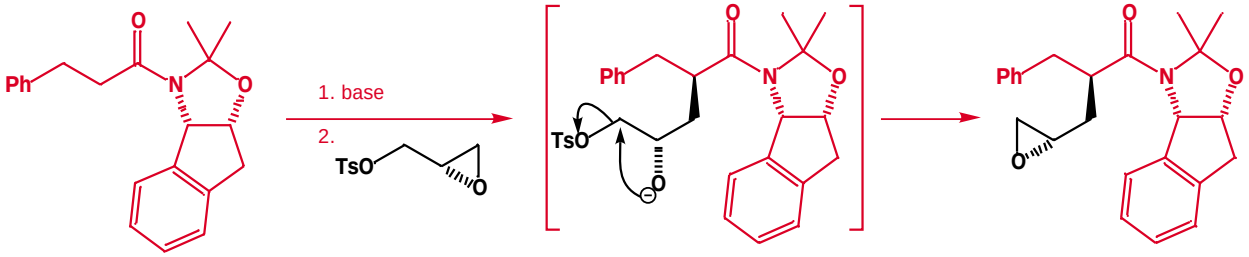



The reason for the stereoselectivity is not altogether clear, but we would expect the bulky nitrogen substituents to favour formation of the *cis* enolate. With the amino-alcohol portion arranged as shown, the top face is more open to attack by electrophiles.

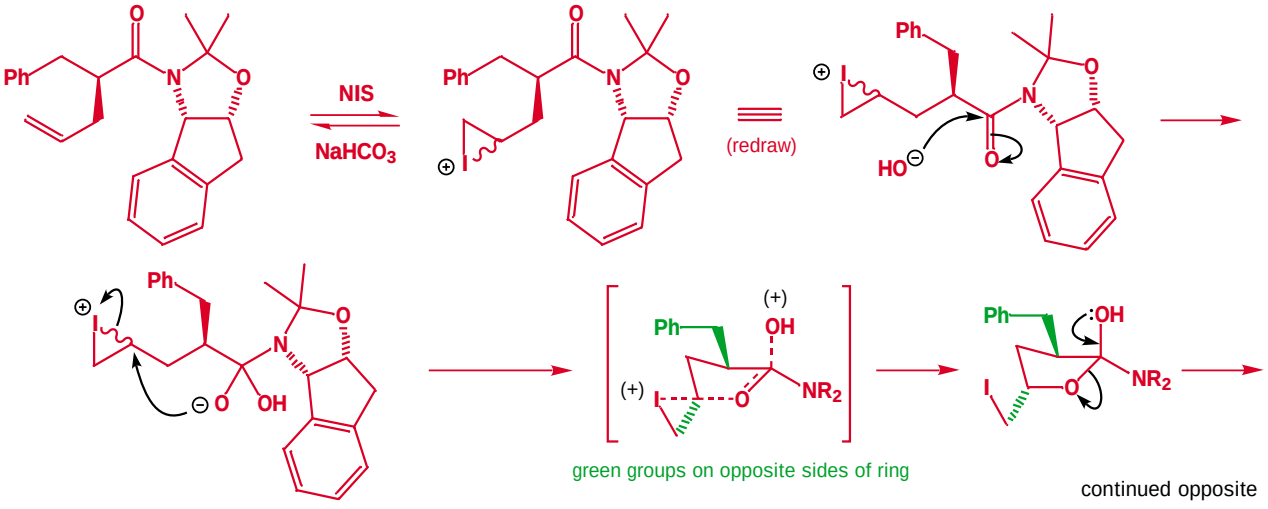



■ How do we know that this happens, and that the reaction does not go simply via direct displacement of tosylate?

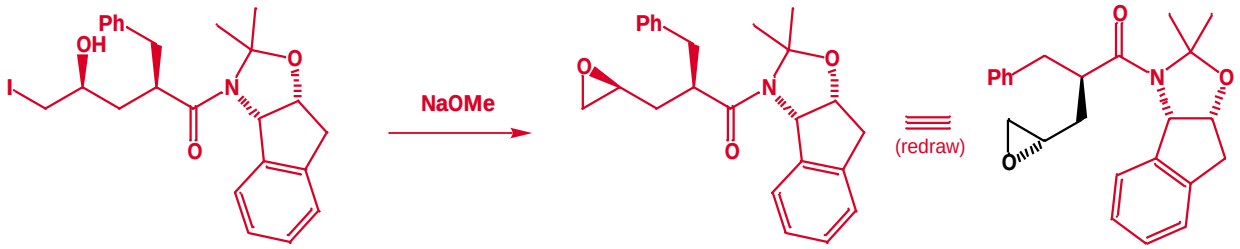
The enolate also reacted diastereoselectively with the epoxy-tosylate prepared earlier. The epoxide, being more electrophilic than the tosylate, is opened first, giving an alkoxide, which closes again to give a new epoxide.



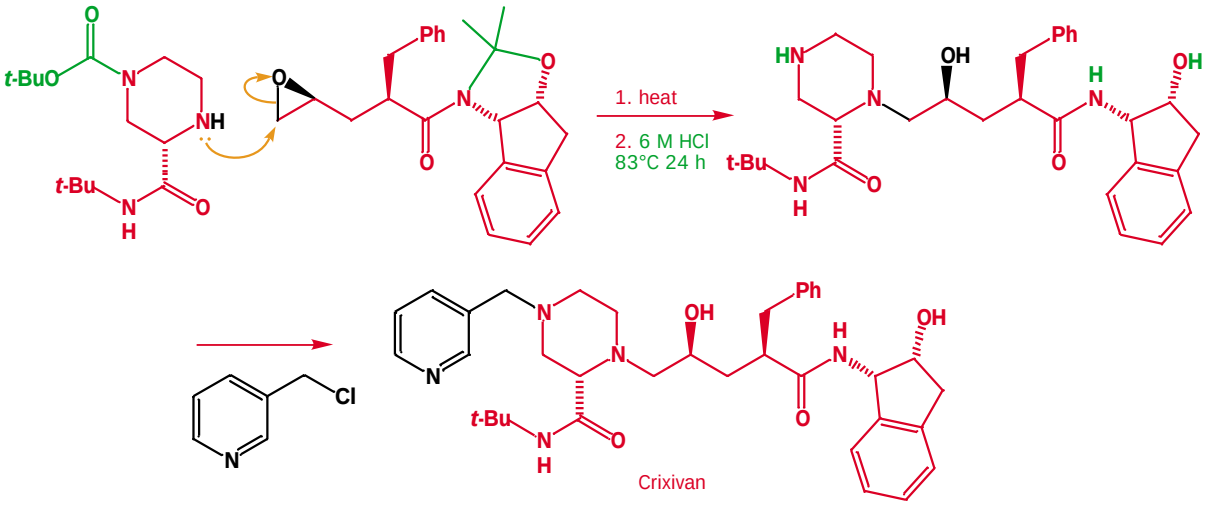

The absolute configuration at the stereogenic centre in the epoxide was, of course, already fixed (by the earlier *enantioselective* Sharpless epoxidation). However, it also turned out to be possible to make this compound by a different route involving a *diastereoselective* reaction of the alkylation product from allyl bromide, again directed by the amino-alcohol-derived auxiliary. The reagents make the reaction look like an iodolactonization—and, indeed, there are many similarities with the diastereoselective iodolactonizations of Chapter 33. NIS (*N*-iodosuccinimide, the iodine analogue of NBS) provides an '$I^+$' source, reacting reversible and non-stereoselectively with the alkene. Of the two diastereoisomeric iodonium ions, one may cyclize rapidly by intramolecular attack of the amide carbonyl group. Cyclization of the other diastereoisomer is prevented by steric hindrance between the parts of the molecule coloured green. Opening of the five-membered ring gives a single diastereoisomer of the iodoalcohol, which was closed to the epoxide by treatment with base.




continued opposite




Three of the five fragments have now been assembled, and only the two amine alkylations remain. The first alkylation makes use of the epoxide to introduce the required 1,2-amino-alcohol functionality. The protected enantiomerically pure piperazine reacted with the epoxide, and the product was treated with acid to deprotect both the second piperazine nitrogen and the 'acetonide' group left over from the earlier chiral auxiliary step. The newly liberated secondary amine was alkylated with the reactive electrophile 3-chloromethyl pyridine, and the final product was crystallized as its sulfate salt.

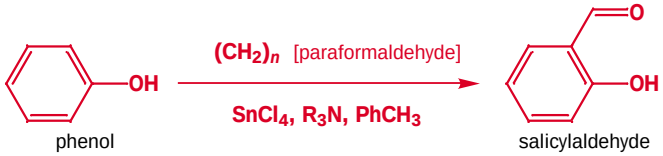



## The future of organic chemistry

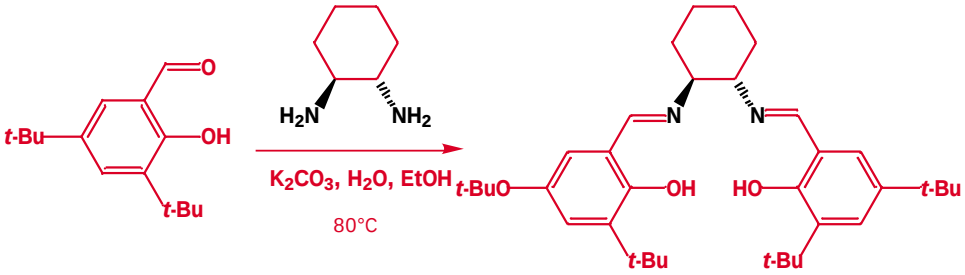
Not all organic chemists can be involved in such exciting projects as the launching of a new anti-AIDS drug. But the chemistry used in this project was invented by chemists in other institutions who had no idea that it would eventually be used to make Crixivan. The Sharpless asymmetric epoxidation, the catalytic asymmetric reduction, the stereoselective enolate alkylation, and the various methods tried out for the enantiomerically pure amino indanol (resolution, enzymatic kinetic resolution) were developed by organic chemists in research laboratories. Some of these famous chemists like Sharpless invented new methods, some made new compounds, some studied new types of molecules, but all built on the work of other chemists.

In 1980 Giovanni Casiraghi, a rather less famous chemist from the University of Parma, published a paper in the *Journal of the Chemical Society* about selective reactions between phenols and formaldehyde. He and his colleagues made the modest discovery that controlled reactions to give salicylaldehydes could be achieved in toluene with $SnCl_4$ as catalyst. The reaction is regioselective for the *ortho* isomer and the paper described the rather precise conditions needed to get a good yield.

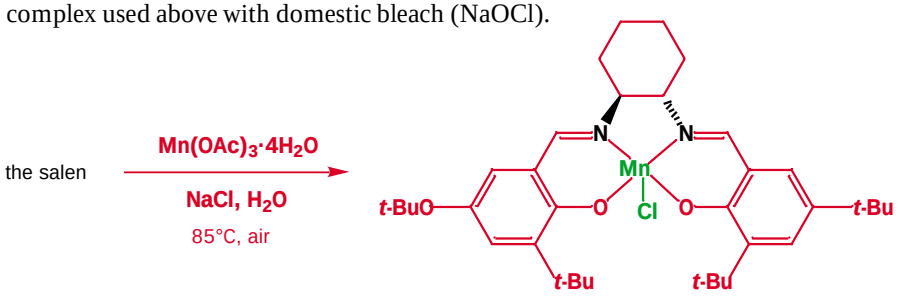
■ In Chapter 52 you met Bakelite, the first synthetic polymer, which results from unselective reactions between these two compounds.



The reaction was also successful for substituted salicylaldehydes. When Jacobsen came to develop his asymmetric epoxidation, which, unlike the Sharpless asymmetric epoxidation, works for simple alkenes and not just for allylic alcohols, he chose 'salens' as his catalysts, partly because they could be made so easily from salicylaldehydes. For example:

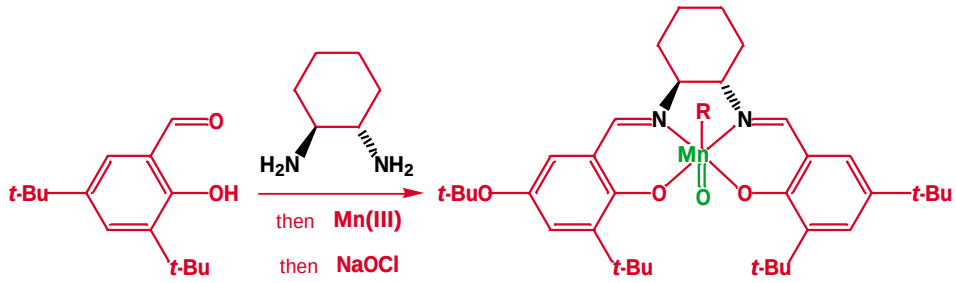



This 'salen' is the ligand for manganese in the asymmetric epoxidation. The stable brown Mn(III) complex can be made from it with $Mn(OAc)_3$ in excellent yield and this can be oxidized to the active complex used above with domestic bleach (NaOCl).

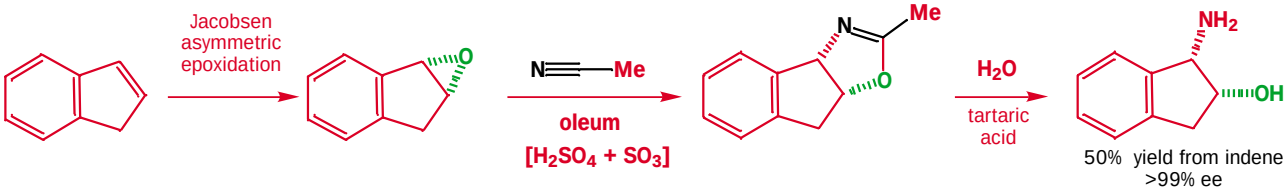



stable Mn(III) complex

Jacobsen epoxidation turned out to be the best large-scale method for preparing the *cis*-amino-indanol for the synthesis of Crixivan. This process is very much the cornerstone of the whole synthesis. During the development of the first laboratory route into a route usable on a very large scale, many methods were tried and the final choice fell on this relatively new type of asymmetric epoxidation. The Sharpless asymmetric epoxidation works only for allylic alcohols (Chapter 45) and so is no good here. The Sharpless asymmetric dihydroxylation works less well on *cis*-alkenes than on *trans*-alkenes. The Jacobsen epoxidation works best on *cis*-alkenes. The catalyst is the Mn(III) complex easily made from a chiral diamine and an aromatic salicylaldehyde (a 2-hydroxybenzaldehyde).



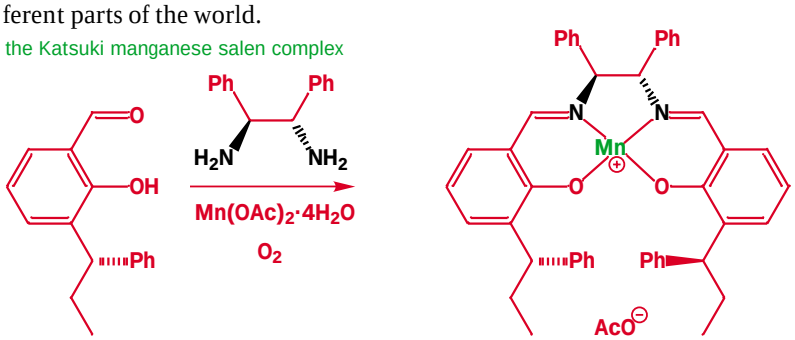

The chirality comes from the diamine and the oxidation from ordinary domestic bleach (NaOCl), which continually recreates the Mn=O bond as it is used in the epoxidation. Only 0.7% catalyst is needed to keep the cycle going efficiently. The epoxide is as good as the diol in the Ritter reaction and the whole process gives a 50% yield of enantiomerically pure *cis*-amino-indanol on a very large scale.




50% yield from indene
>99% ee

In the same year (1990) that Jacobsen reported his asymmetric epoxidation, a group led by Tsutomu Katsuki at the University of Kyushu in Japan reported a closely related asymmetric epoxidation. The chiral catalyst is also a salen and the metal manganese. The oxidant is iodosobenzene (PhI=O) but this method works best for *E*-alkenes. It is no coincidence that Katsuki and Jacobsen both worked for Sharpless. It is not unusual for similar discoveries to be made independently in different parts of the world.

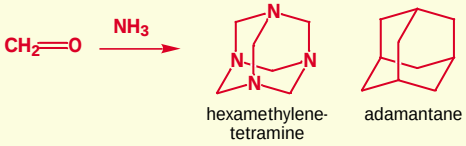
the Katsuki manganese salen complex




It did not enter Casiraghi's wildest dreams that his work might some day be useful in a matter of life and death. Nor did his four co-workers nor Jacobsen's more numerous co-workers see clearly the future applications of their work. By its very nature it is impossible to predict the outcome or the applications of research. But be quite sure of one thing. Good research and exciting discoveries come from a thorough understanding of the fundamentals of organic chemistry and require chemists to work as a team . The Italian work is a model of careful experimentation and a thorough study of reaction conditions together with sensible explanations of their discoveries using the same curly arrows we have been using. The Harvard team probably had a clearer idea that they were into something significant and worked with equal care and precision. Jacobsen's name is famous but both teams at Parma and Harvard Universities were needed to make the work available to Merck.

## Hexamethylenetetramine

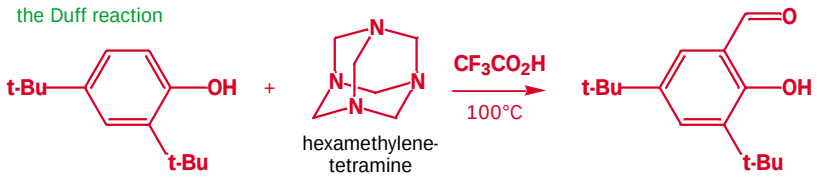
Hexamethylenetetramine is a co-polymer (oligomer really such as those we met in Chapter 52) of formaldehyde and ammonia containing six formaldehyde and four ammonia molecules. It has a beautifully symmetrical cage structure belonging to the adamantane series.




Hexamethylenetetramine is a crystalline compound used as a convenient source of formaldehyde for, among other things, polymerization reactions. It has a tetrahedral symmetry, as does adamantane, which might be regarded as the basic structural unit (not the same as the monomer!) of diamonde. Diamond is of course a polymer of carbon atoms.

When Jacobsen's epoxidation was fully described in 1998–99, the Casiraghi method was abandoned in favour of an even older method discovered in the 1930s by Duff. The remarkable Duff reaction uses hexamethylenetetramine, the oligomer of formaldehyde and ammonia, to provide the extra carbon atom. The otherwise unknown Duff worked at Birmingham Technical College. Later in 1972, a William E. Smith, working in the GEC chemical laboratories at Schenectady, New York, found how to make the Duff reaction more general and better yielding by using $CF_3CO_2H$ as catalyst. Even so, this method gives a lower yield than the Casiraghi method but it uses no dangerous reagents (particularly no stoichiometric tin) and is more suitable for large-scale work. When Duff was inventing his reaction or Smith was modifying the conditions, asymmetric synthesis was not even a gleam in anyone's eyes. It is impossible even for the inventor to predict whether a discovery is important or not.

the Duff reaction



The Sharpless asymmetric dihydroxylation works best for *trans* disubstituted alkenes, while the Jacobsen epoxidation works best for *cis* disubstituted alkenes. Even in this small area, there is a need for better and more general methods. Organic chemistry has a long way to go.

If you continue your studies in organic chemistry beyond the scope of this book, you will want to read of modern work in more specialized areas. Your university library should have a selection of books on topics such as: orbitals and chemical reactions; NMR spectroscopy; enzyme mechanisms; organometallic chemistry; biosynthesis; asymmetric synthesis; combinatorial chemistry; and molecular modelling. This book should equip you with enough fundamental organic chemistry to explore these topics with understanding and enjoyment and, perhaps, to discover what you want to do for the rest of your life. All of the chemists mentioned in this chapter and throughout the book began their careers as students of chemistry at universities somewhere in the world. You have the good fortune to study chemistry at a time when more is understood about the subject than ever before, when information is easier to retrieve than ever before, and when organic chemistry is more interrelated with other disciplines than ever before. Duff, Smith, and Casiraghi felt themselves part of an international community of organic chemists in industry and universities but never has that community been so well founded as it is nowadays. Travel to laboratories in other countries is commonplace for students of organic chemistry now and even at home you can travel on the internet to other countries and see what is going on in chemistry there. You might try the web pages of our institutions for a start: Cambridge is http://www.ch.cam.ac.uk/; Liverpool is http://www.liv.ac.uk/Chemistry/; and Manchester is http://www.ch.man.ac.uk/. There is a general index to chemistry all over the world on http://www.ch.cam.ac.uk/ChemSitesIndex.html.

■

If you want to read more about these discoveries we suggest: 'Practical asymmetric synthesis', I. W. Davies and P. J. Reider, *Chemistry and Industry (London)*, 1996, 412–15. The reference for the Parma work is: G. Casiraghi, G. Casnati, G. Puglia, and G. Terenghi, *J. Chem. Soc., Perkin Trans. 1*, 1980, 1862–65. These journals will be in your department or university library.