2

SECONDARY METABOLISM: THE BUILDING BLOCKS AND CONSTRUCTION MECHANISMS

Distinctions between primary and secondary are defined, and the basic building blocks used in the biosynthesis of secondary natural products are introduced. The chemistry underlying how these building blocks are assembled in nature is described, subdivided according to chemical mechanism, including alkylation reactions, Wagner–Meerwein rearrangements, aldol and Claisen reactions, Schiff base formation and Mannich reactions, transaminations, decarboxylations, oxidation and reduction reactions, phenolic oxidative coupling, and glycosylations.

PRIMARY AND SECONDARY METABOLISM

All organisms need to transform and interconvert a vast number of organic compounds to enable them to live, grow, and reproduce. They need to provide themselves with energy in the form of ATP, and a supply of building blocks to construct their own tissues. An integrated network of enzyme-mediated and carefully regulated chemical reactions is used for this purpose, collectively referred to as **intermediary metabolism**, and the pathways involved are termed **metabolic pathways**. Some of the crucially important molecules of life are carbohydrates, proteins, fats, and nucleic acids. Apart from fats, these are polymeric materials. Carbohydrates are composed of sugar units, whilst proteins are made up from amino acids, and nucleic acids are based on nucleotides. Organisms vary widely in their capacity to synthesize and transform chemicals. For instance, plants are very efficient at synthesizing organic compounds via photosynthesis from inorganic materials found in the environment, whilst other organisms such as animals and microorganisms rely on obtaining their raw materials in their diet, e.g. by consuming plants. Thus, many of the metabolic pathways are concerned with degrading materials taken in as food, whilst others are then required to synthesize specialized molecules from the basic compounds so obtained.

Despite the extremely varied characteristics of living organisms, the pathways for generally modifying and synthesizing carbohydrates, proteins, fats, and nucleic acids are found to be essentially the same in all organisms, apart from minor variations. These processes demonstrate the fundamental unity of all living matter, and are collectively described as **primary metabolism**, with the compounds involved in the pathways being termed **primary metabolites**. Thus degradation of carbohydrates and sugars generally proceeds via the well characterized pathways known as glycolysis and the Krebs/citric acid/tricarboxylic acid cycle, which release energy from the organic compounds by oxidative reactions. Oxidation of fatty acids from fats by the sequence called β-oxidation also provides energy. Aerobic organisms are able to optimize these processes by adding on a further process, oxidative phosphorylation. This improves the efficiency of oxidation by incorporating a more general process applicable to the oxidation of a wide variety of substrates rather than having to provide specific processes for each individual substrate. Proteins taken in via the diet provide amino acids, but the proportions of each will almost certainly vary from the organism's requirements. Metabolic pathways are thus available to

interconvert amino acids, or degrade those not required and thus provide a further source of energy. Most organisms can synthesize only a proportion of the amino acids they actually require for protein synthesis. Those structures not synthesized, so-called essential amino acids, must be obtained from external sources.

In contrast to these primary metabolic pathways, which synthesize, degrade, and generally interconvert compounds commonly encountered in all organisms, there also exists an area of metabolism concerned with compounds which have a much more limited distribution in nature. Such compounds, called **secondary metabolites**, are found in only specific organisms, or groups of organisms, and are an expression of the individuality of species. Secondary metabolites are not necessarily produced under all conditions, and in the vast majority of cases the function of these compounds and their benefit to the organism is not yet known. Some are undoubtedly produced for easily appreciated reasons, e.g. as toxic materials providing defence against predators, as volatile attractants towards the same or other species, or as colouring agents to attract or warn other species, but it is logical to assume that all do play some vital role for the well-being of the producer. It is this area of **secondary metabolism** that provides most of the pharmacologically active natural products. It is thus fairly obvious that the human diet could be both unpalatable and remarkably dangerous if all plants, animals, and fungi produced the same range of compounds.

The above generalizations distinguishing primary and secondary metabolites unfortunately leave a 'grey area' at the boundary, so that some groups of natural products could be assigned to either division. Fatty acids and sugars provide good examples, in that most are best described as primary metabolites, whilst some representatives are extremely rare and found only in a handful of species. Likewise, steroid biosynthesis produces a range of widely distributed fundamental structures, yet some steroids, many of them with pronounced pharmacological activity, are restricted to certain organisms. Hopefully, the blurring of the boundaries will not cause confusion; the subdivision into primary metabolism ($\equiv$ biochemistry) or secondary metabolism ($\equiv$ natural products chemistry)

is merely a convenience and there is considerable overlap.

THE BUILDING BLOCKS

The building blocks for secondary metabolites are derived from primary metabolism as indicated in Figure 2.1. This scheme outlines how metabolites from the fundamental processes of photosynthesis, glycolysis, and the Krebs cycle are tapped off from energy-generating processes to provide biosynthetic intermediates. The number of building blocks needed is surprisingly few, and as with any child's construction set a vast array of objects can be built up from a limited number of basic building blocks. By far the most important building blocks employed in the biosynthesis of secondary metabolites are derived from the intermediates acetyl coenzyme A (acetyl-CoA), shikimic acid, mevalonic acid, and 1-deoxyxylulose 5-phosphate. These are utilized respectively in the **acetate**, **shikimate**, **mevalonate**, and **deoxyxylulose phosphate** pathways, which form the basis of succeeding chapters. **Acetyl-CoA** is formed by oxidative decarboxylation of the glycolytic pathway product pyruvic acid. It is also produced by the β-oxidation of fatty acids, effectively reversing the process by which fatty acids are themselves synthesized from acetyl-CoA. Important secondary metabolites formed from the acetate pathway include phenols, prostaglandins, and macrolide antibiotics, together with various fatty acids and derivatives at the primary/secondary metabolism interface. **Shikimic acid** is produced from a combination of phosphoenolpyruvate, a glycolytic pathway intermediate, and erythrose 4-phosphate from the pentose phosphate pathway. The reactions of the pentose phosphate cycle may be employed for the degradation of glucose, but they also feature in the synthesis of sugars by photosynthesis. The shikimate pathway leads to a variety of phenols, cinnamic acid derivatives, lignans, and alkaloids. **Mevalonic acid** is itself formed from three molecules of acetyl-CoA, but the mevalonate pathway channels acetate into a different series of compounds than does the acetate pathway. **Deoxyxylulose phosphate** arises from a combination of two glycolytic pathway intermediates, namely pyruvic acid and glyceraldehyde 3-phosphate. The mevalonate and deoxyxylulose phosphate pathways are together
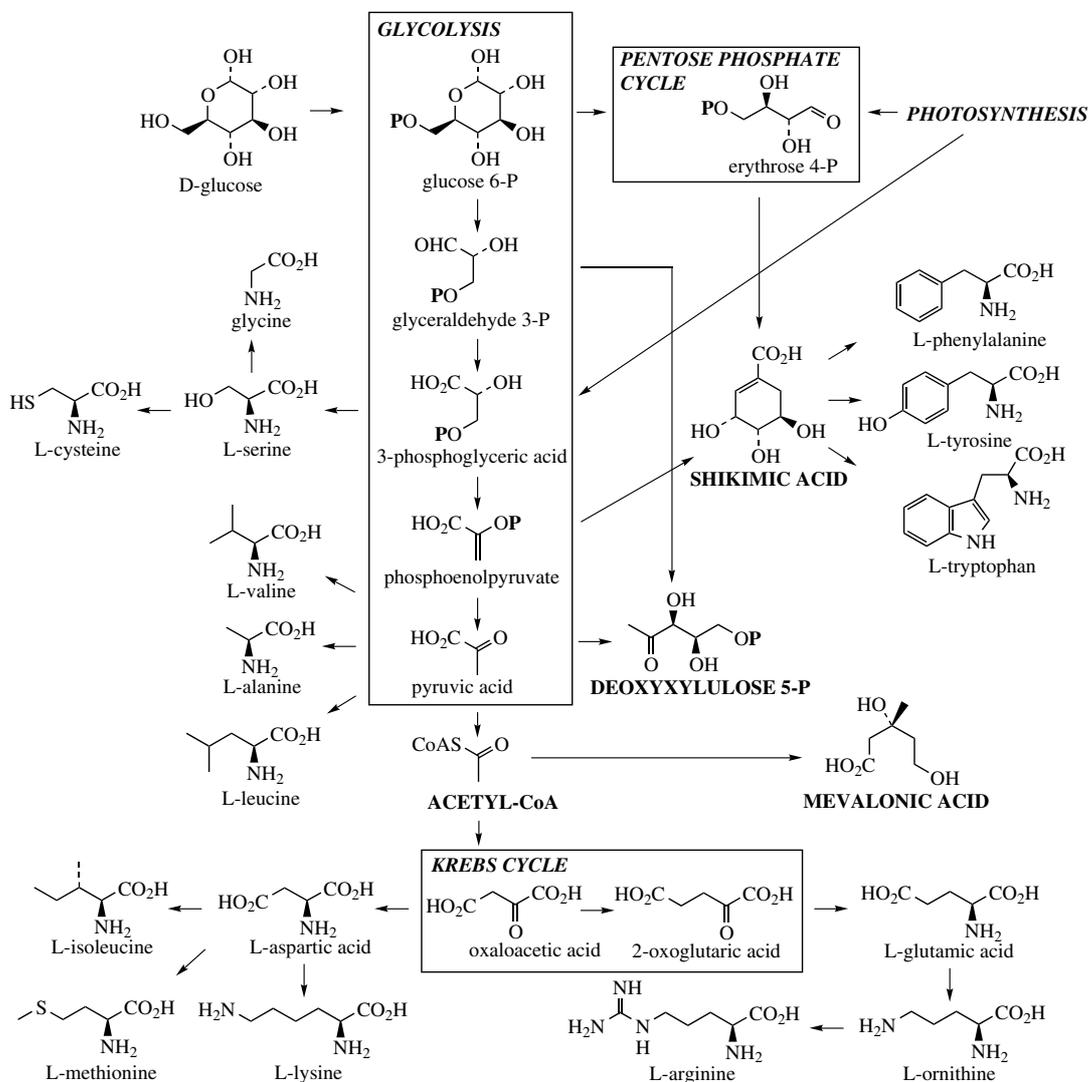
Figure 2.1

responsible for the biosynthesis of a vast array of terpenoid and steroid metabolites.

In addition to acetyl-CoA, shikimic acid, mevalonic acid, and deoxyxylulose phosphate, other building blocks based on amino acids are frequently employed in natural product synthesis. Peptides, proteins, alkaloids, and many antibiotics are derived from amino acids, and the origins of the most important amino acid components of these are briefly indicated in Figure 2.1. Intermediates from the glycolytic pathway and the Krebs cycle are used in constructing many of them, but the aromatic amino acids **phenylalanine**, **tyrosine**,

and **tryptophan** are themselves products from the shikimate pathway. **Ornithine**, a non-protein amino acid, along with its homologue **lysine**, are important alkaloid precursors having their origins in Krebs cycle intermediates.

Of special significance is the appreciation that secondary metabolites can be synthesized by combining several building blocks of the same type, or by using a mixture of different building blocks. This expands structural diversity, and consequently makes subdivisions based entirely on biosynthetic pathways rather more difficult. A typical natural product might be produced by combining elements

from the acetate, shikimate, and deoxyxylulose phosphate pathways. Many secondary metabolites also contain one or more sugar units in their structure, either simple primary metabolites such as glucose or ribose, or alternatively substantially modified and unusual sugars. To appreciate how a natural product is elaborated, it is of value to be able to dissect its structure into the basic building blocks from which it is made up, and to propose how these are mechanistically joined together. With a little experience and practice, this becomes a relatively simple process, and it allows the molecule to be rationalized, thus exposing logical relationships between apparently quite different structures. In this way, similarities become much more meaningful than differences, and an understanding of biosynthetic pathways allows rational connecting links to be established. This forms the basic approach in this book.

Relatively few building blocks are routinely employed, and the following list, though not comprehensive, includes those most frequently encountered in producing the carbon and nitrogen skeleton of a natural product. As we shall see, oxygen atoms can be introduced and removed by a variety of processes, and so are not considered in the initial analysis, except as a pointer to an acetate (see page 62) or shikimate (see page 123) origin. The structural features of these building blocks are shown in Figure 2.2.

- **C_1:** The simplest of the building blocks is composed of a single carbon atom, usually in the form of a methyl group, and most frequently it is attached to oxygen or nitrogen, but occasionally to carbon. It is derived from the *S*-methyl of L-**methionine**. The methylenedioxy group (OCH_2O) is also an example of a C_1 unit.

- **C_2:** A two-carbon unit may be supplied by **acetyl-CoA**. This could be a simple acetyl group, as in an ester, but more frequently it forms part of a long alkyl chain (as in a fatty acid) or may be part of an aromatic system (e.g. phenols). Of particular relevance is that in the latter examples, acetyl-CoA is first converted into the more reactive **malonyl-CoA** before its incorporation.

- **C_5:** The branched-chain C_5 'isoprene' unit is a feature of compounds formed from **mevalonate** or **deoxyxylulose phosphate**. Mevalonate itself is the product from three acetyl-CoA molecules, but only five of mevalonate's six carbons are used, the carboxyl group being lost. The alternative precursor deoxyxylulose phosphate, a straight-chain sugar derivative, undergoes a skeletal rearrangement to form the branched-chain isoprene unit.

- **C_6C_3:** This refers to a phenylpropyl unit and is obtained from the carbon skeleton of either L-**phenylalanine** or L-**tyrosine**, two of the shikimate-derived aromatic amino acids. This, of course, requires loss of the amino group. The C_3 side-chain may be saturated or unsaturated, and may be oxygenated. Sometimes the side-chain is cleaved, removing one or two carbons. Thus, C_6C_2 and C_6C_1 units represent modified shortened forms of the C_6C_3 system.

- **C_6C_2N:** Again, this building block is formed from either L-**phenylalanine** or L-**tyrosine**, L-tyrosine being by far the more common. In the elaboration of this unit, the carboxyl carbon of the amino acid is removed.

- **indole.C_2N:** The third of the aromatic amino acids is L-**tryptophan**. This indole-containing system can undergo decarboxylation in a similar way to L-phenylalanine and L-tyrosine so providing the remainder of the skeleton as an indole.C_2N unit.

- **C_4N:** The C_4N unit is usually found as a heterocyclic pyrrolidine system and is produced from the non-protein amino acid L-**ornithine**. In marked contrast to the C_6C_2N and indole.C_2N units described above, ornithine supplies not its α-amino nitrogen, but the δ-amino nitrogen. The carboxylic acid function and the α-amino nitrogen are both lost.

- **C_5N:** This is produced in exactly the same way as the C_4N unit, but using L-**lysine** as precursor. The ε-amino nitrogen is retained, and the unit tends to be found as a piperidine ring system.

These eight building blocks will form the basis of many of the natural product structures discussed in the following chapters. Simple examples of how compounds can be visualized as a combination of building blocks are shown in Figure 2.3. At this stage, it is inappropriate to justify why a particular combination of units is used, but this aspect should become clear as the pathways are described.
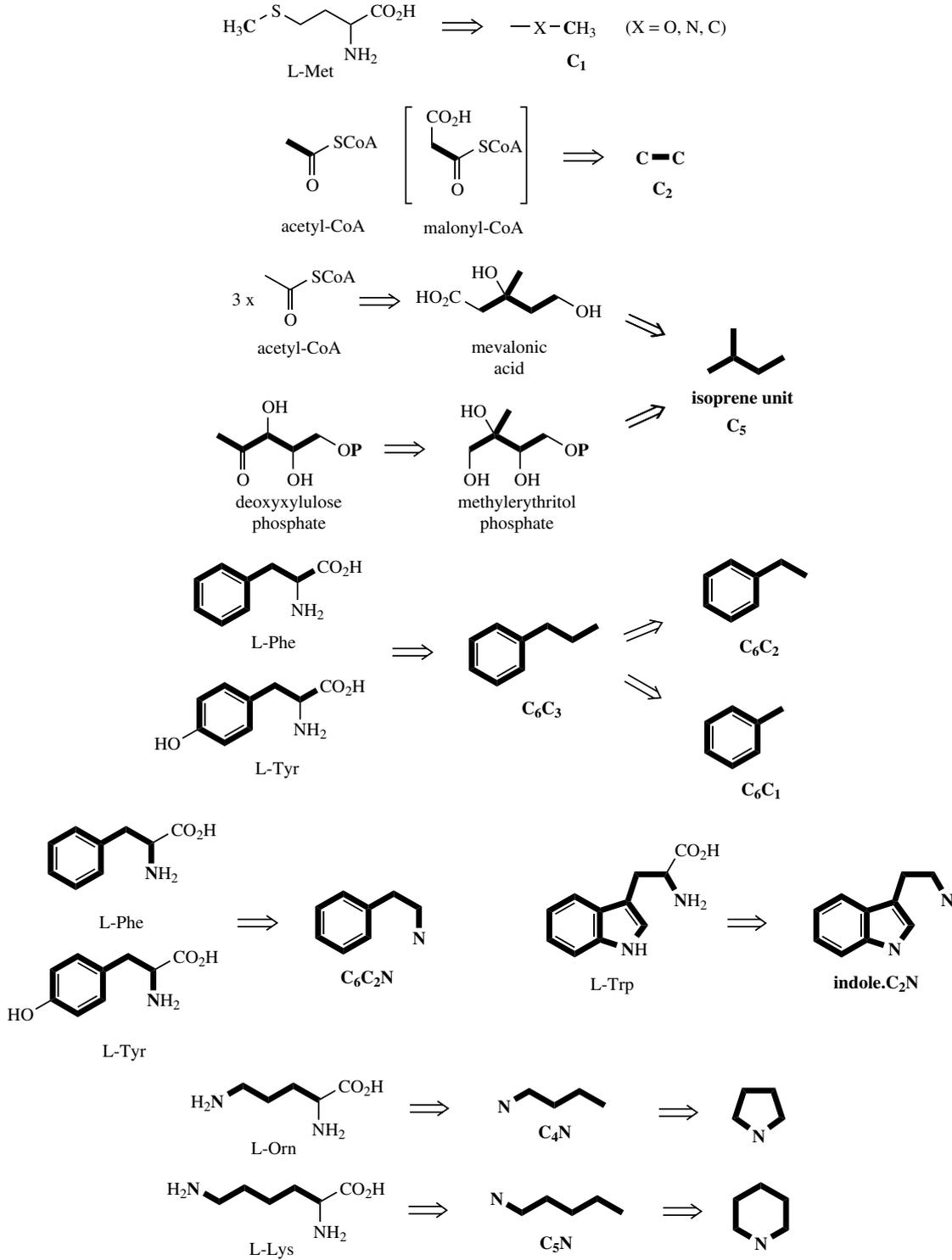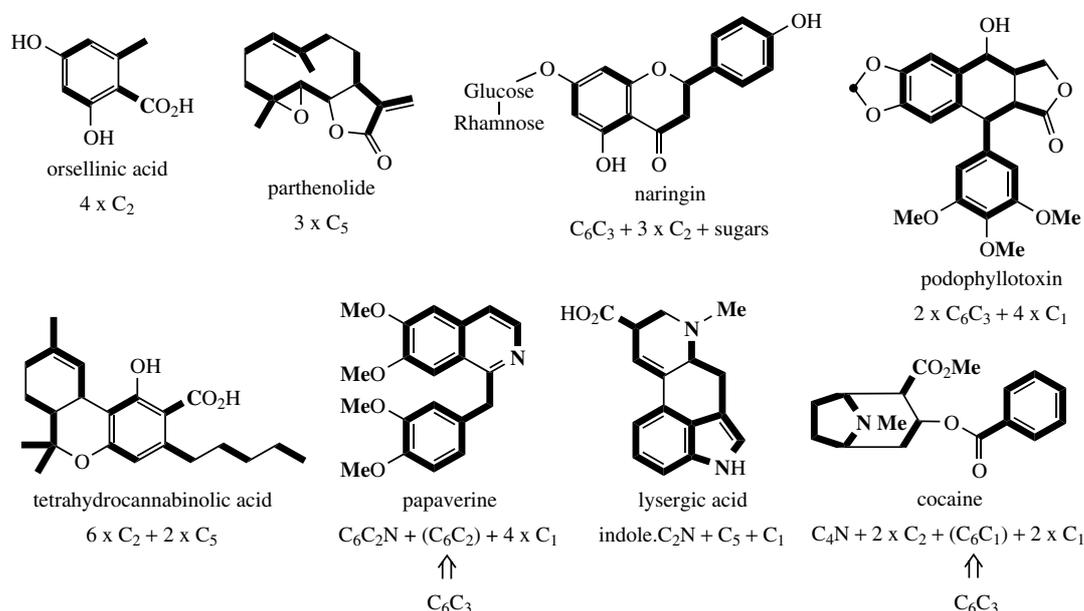
The building blocks

Figure 2.2

orsellinic acid

$4 \times C_2$

parthenolide

$3 \times C_5$

naringin

$C_6C_3 + 3 \times C_2 + sugars$

podophyllotoxin

$2 \times C_6C_3 + 4 \times C_1$

tetrahydrocannabinolic acid

$6 \times C_2 + 2 \times C_5$

papaverine

$C_6C_2N + (C_6C_2) + 4 \times C_1$

$\Uparrow$

C_6C_3

lysergic acid

$indole.C_2N + C_5 + C_1$

cocaine

$C_4N + 2 \times C_2 + (C_6C_1) + 2 \times C_1$

$\Uparrow$

C_6C_3

Figure 2.3

A word of warning is also necessary. Some natural products have been produced by processes in which a fundamental rearrangement of the carbon skeleton has occurred. This is especially common with structures derived from isoprene units, and it obviously disguises some of the original building blocks from immediate recognition. The same is true if one or more carbon atoms are removed by oxidation reactions.

THE CONSTRUCTION MECHANISMS

Natural product molecules are biosynthesized by a sequence of reactions which, with very few exceptions, are catalysed by enzymes. Enzymes are protein molecules which facilitate chemical modification of substrates by virtue of their specific binding properties conferred by the particular combination of functional groups in the constituent amino acids. In many cases, a suitable cofactor, e.g. NAD^+, PLP, HSCoA (see below), as well as the substrate, may also be bound to participate in the transformation. Although enzymes catalyse some fairly elaborate and sometimes unexpected changes, it is generally possible to account for the reactions using sound chemical principles and mechanisms. As we explore the pathways to a wide variety of natural products, the reactions will generally be discussed in terms of chemical analogies. Enzymes have the power to effect these transformations more efficiently and more rapidly than the chemical analogy, and also under very much milder conditions. Where relevant, they also carry out reactions in a stereospecific manner. Some of the important reactions frequently encountered are now described.

Alkylation Reactions: Nucleophilic Substitution

The C_1 methyl building unit is supplied from L-methionine and is introduced by a nucleophilic substitution reaction. In nature, the leaving group is enhanced by converting L-methionine into **S-adenosylmethionine** (**SAM**) [Figure 2.4(a)]. This gives a positively charged sulphur and facilitates the nucleophilic substitution (S_N2) type mechanism [Figure 2.4(b)]. Thus, O-methyl and N-methyl linkages may be obtained using hydroxyl and amino functions as nucleophiles. The generation of C-methyl linkages requires the participation of nucleophilic carbon. Positions *ortho* or *para* to a phenol group, or positions adjacent to one or more carbonyl groups, are thus candidates for C-methylation [Figure 2.4(c)].
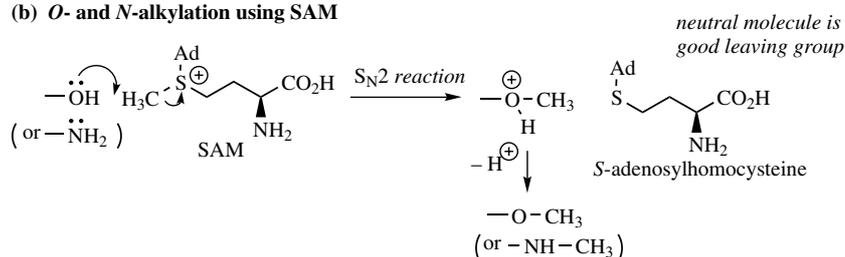
A C_5 isoprene unit in the form of **dimethylallyl diphosphate** (**DMAPP**) may also act as an

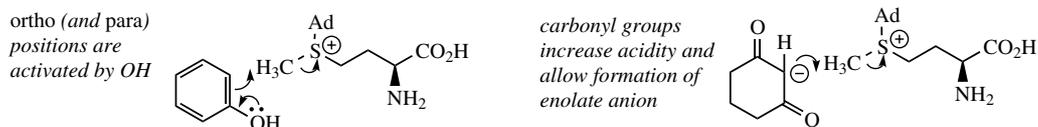**Alkylation reactions: nucleophilic substitution**

(a) formation of SAM
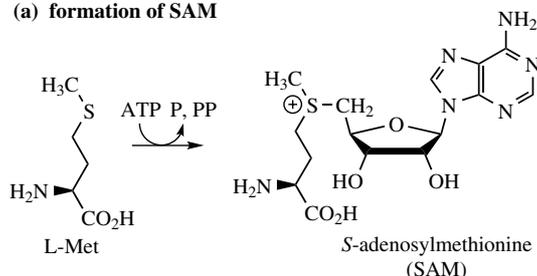
(b) O- and N-alkylation using SAM

(c) C-alkylation using SAM

(d) O-alkylation using DMAPP

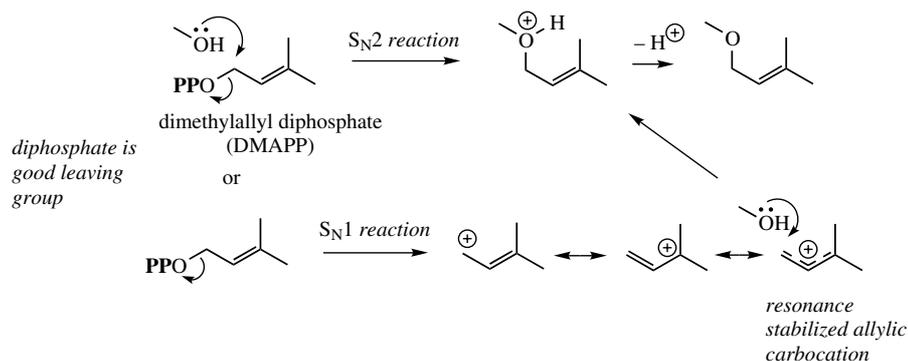**Figure 2.4**

alkylating agent, and a similar S_N2 nucleophilic displacement can be proposed, the diphosphate making a good leaving group [Figure 2.4(d)]. In some cases, there is evidence that DMAPP may ionize first to the resonance-stabilized allylic carbocation and thus an S_N1 process operates instead. C-Alkylation at activated positions using DMAPP is analogous to the C-methylation process above.

Alkylation Reactions: Electrophilic Addition

As indicated above, the C_5 isoprene unit in the form of **dimethylallyl diphosphate** (**DMAPP**) can be used to alkylate a nucleophile. In the elaboration of terpenoids and steroids, two or more C_5 units are joined together, and the reactions

are rationalized in terms of carbocation chemistry, including electrophilic addition of carbocations on to alkenes. DMAPP may ionize to generate a resonance-stabilized allylic carbocation as shown in Figure 2.4(d), and this can then react with an alkene [e.g. **isopentenyl diphosphate** (**IPP**)] as depicted in Figure 2.5(a). The resultant carbocation may then lose a proton to give the uncharged product geranyl diphosphate (GPP). Where the alkene and carbocation functions reside in the same molecule, this type

of mechanism can be responsible for cyclization reactions [Figure 2.5(a)].

The initial carbocation may be generated by a number of mechanisms, important examples being loss of a leaving group, especially diphosphate (i.e. S_N1 type ionization), protonation of an alkene, and protonation/ring opening of epoxides [Figure 2.5(b)]. *S*-**Adenosylmethionine** may also alkylate alkenes by an electrophilic addition mechanism, adding a C_1 unit, and generating an intermediate carbocation.
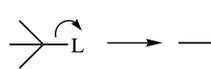
Alkylation reactions: electrophilic addition

(a) inter- and intra-molecular additions

electrophilic addition of cation on to alkene

isopentenyl
diphosphate
(IPP)

geranyl diphosphate
(GPP)
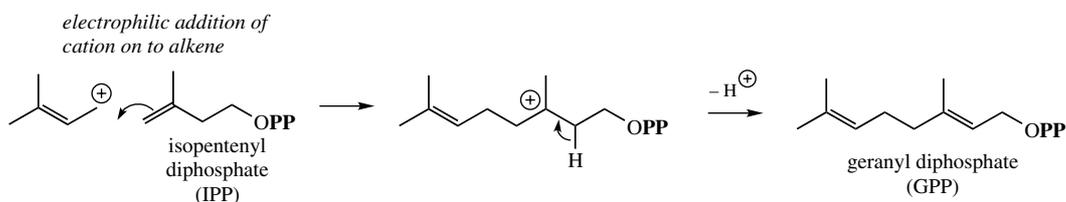
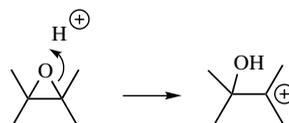*intramolecular
addition: cyclization*
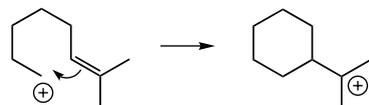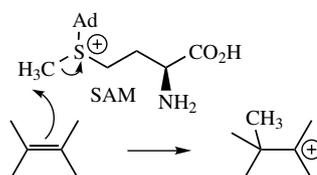
(b) generation of carbocation

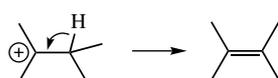*loss of leaving group*   *protonation of alkene*   *protonation and ring opening
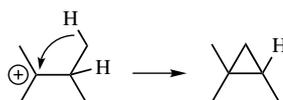of epoxide*

methylation of alkene via SAM

(c) discharge of carbocation

loss of proton   *cyclization / loss of proton*   *quenching with nucleophile (water)*

Figure 2.5

The final carbocation may be discharged by loss of a proton (giving an alkene or sometimes a cyclopropane ring) or by quenching with a suitable nucleophile, especially water [Figure 2.5(c)].

Wagner–Meerwein Rearrangements

A wide range of structures encountered in natural terpenoid and steroid derivatives can only be rationalized as originating from C_5 isoprene units if some fundamental rearrangement process has occurred during biosynthesis. These rearrangements have, in many cases, been confirmed experimentally, and are almost always consistent with the participation of carbocation intermediates. Rearrangements in chemical reactions involving carbocation intermediates, e.g. S_N1 and E1 reactions, are not uncommon, and typically consist of 1,2-shifts of hydride, methyl, or alkyl groups. Occasionally, 1,3- or longer shifts are encountered. These shifts, termed **Wagner–Meerwein rearrangements**, are readily rationalized in terms of generating a more stable carbocation, or relaxing ring strain (Figure 2.6). Thus, tertiary carbocations are favoured over secondary carbocations,

and the usual objective in these rearrangements is to achieve tertiary status at the positive centre. However, a tertiary to secondary transition might be favoured if the rearrangement allows a significant release of ring strain. These general concepts are occasionally ignored by nature, but it must be remembered that the reactions are enzyme-mediated and carbocations may not exist as discrete species in the transformations. An interesting feature of some biosynthetic pathways, e.g. that leading to steroids, is a remarkable series of concerted 1,2-migrations rationalized via carbocation chemistry, but entirely a consequence of the enzyme's participation (Figure 2.6).

Aldol and Claisen Reactions

The **aldol** and **Claisen** reactions both achieve carbon–carbon bond formation and in typical base-catalysed chemical reactions depend on the generation of a resonance-stabilized enolate anion from a suitable carbonyl system (Figure 2.7). Whether an aldol-type or Claisen-type product is formed depends on the nature of X and its potential as a leaving group. Thus, chemically, two molecules

Wagner–Meerwein rearrangements
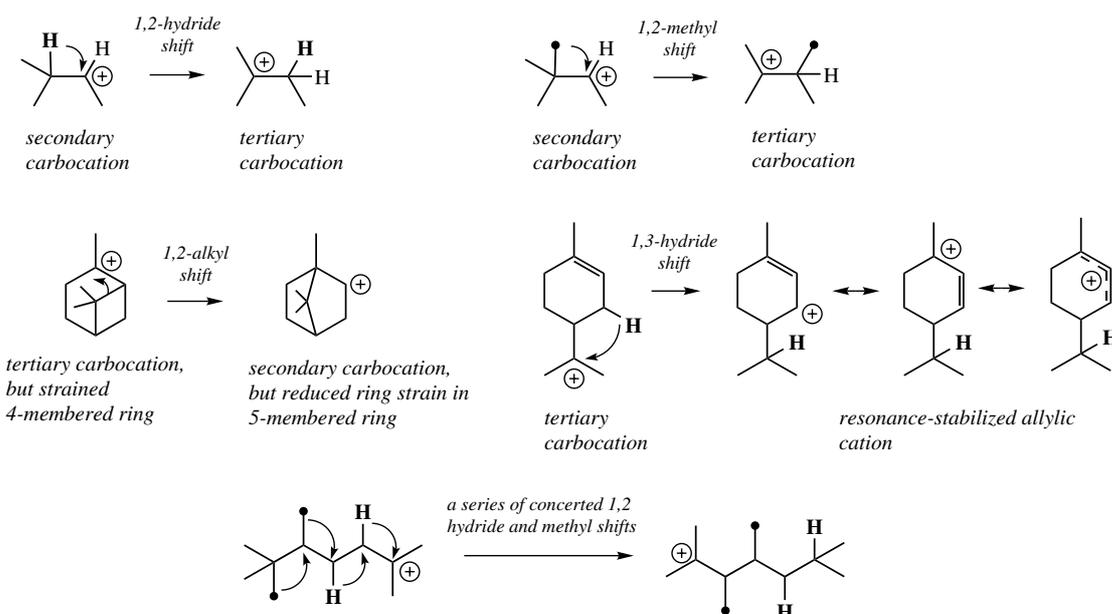

Figure 2.6

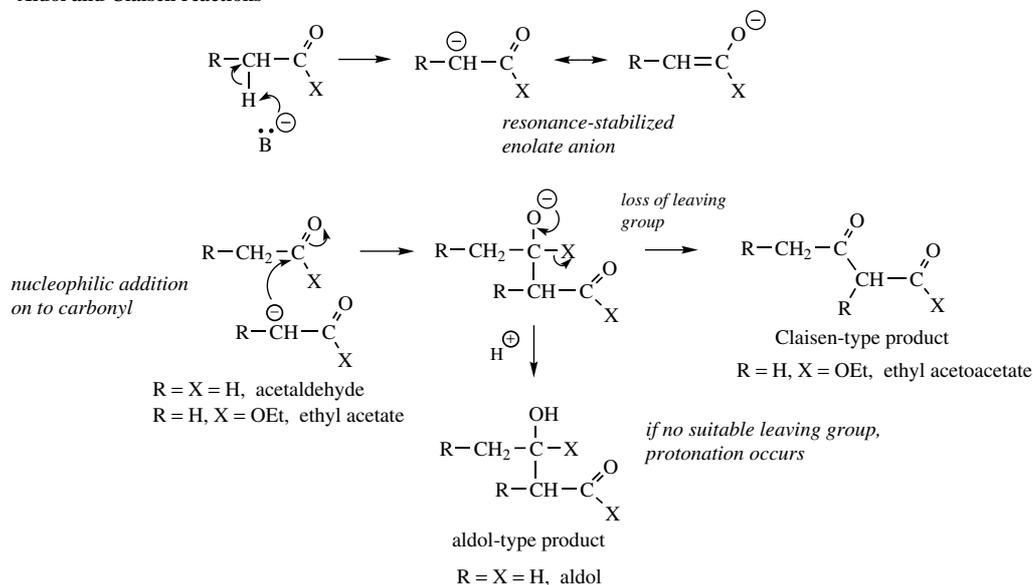**Aldol and Claisen reactions**
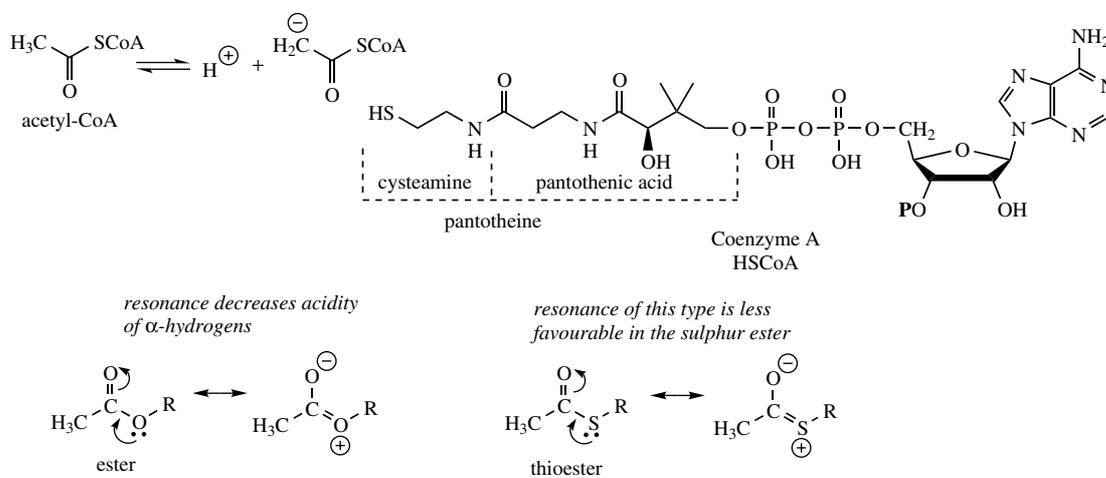

Figure 2.7

Figure 2.8

of acetaldehyde yield aldol, whilst two molecules of ethyl acetate can give ethyl acetoacetate. These processes are vitally important in biochemistry for the elaboration of both secondary and primary metabolites, but the enzyme catalysis obviates the need for strong bases, and probably means the enolate anion has little more than transitory existence. Nevertheless, the reactions do appear to parallel enolate anion chemistry, and are frequently responsible for joining together of C_2 acetate groups.

In most cases, the biological reactions involve coenzyme A esters, e.g. **acetyl-CoA** (Figure 2.8). This is a thioester of acetic acid, and it has significant advantages over oxygen esters, e.g. ethyl acetate, in that the α-methylene hydrogens are now more acidic, comparable in fact to those in the equivalent ketone, thus increasing the likelihood of generating the enolate anion. This is explained in terms of electron delocalization in the ester function (Figure 2.8). This type of delocalization is

more prominent in the oxygen ester than in the sulphur ester, due to oxygen's smaller size and thus closer proximity of the lone pair for overlap with carbon's orbitals. Furthermore, the thioester has a much more favourable leaving group than the oxygen ester, and the combined effect is to increase the reactivity for both the aldol and Claisen-type reactions.

Claisen reactions involving acetyl-CoA are made even more favourable by first converting acetyl-CoA into **malonyl-CoA** by a carboxylation reaction with CO_2 using ATP and the coenzyme biotin (Figure 2.9). ATP and CO_2 (as bicarbonate, HCO_3^-) form the mixed anhydride, which carboxylates the coenzyme in a biotin–enzyme complex. Fixation of carbon dioxide by biotin–enzyme complexes is not unique to acetyl-CoA, and another important example occurs in the generation of oxaloacetate from pyruvate in the synthesis of glucose from non-carbohydrate sources

(gluconeogenesis). The conversion of acetyl-CoA into malonyl-CoA means the α-hydrogens are now flanked by two carbonyl groups, and have increased acidity. Thus, a more favourable nucleophile is provided for the Claisen reaction. No acylated malonic acid derivatives are produced, and the carboxyl group introduced into malonyl-CoA is simultaneously lost by a decarboxylation reaction during the Claisen condensation (Figure 2.9). An alternative rationalization is that decarboxylation of the malonyl ester is used to generate the acetyl enolate anion without any requirement for a strong base. Thus, the product formed from acetyl-CoA as electrophile and malonyl-CoA as nucleophile is acetoacetyl-CoA, which is actually the same as in the condensation of two molecules of acetyl-CoA. Accordingly, the role of the carboxylation step is clear cut: the carboxyl activates the α-carbon to facilitate the Claisen condensation, and it is immediately removed on completion of this task.
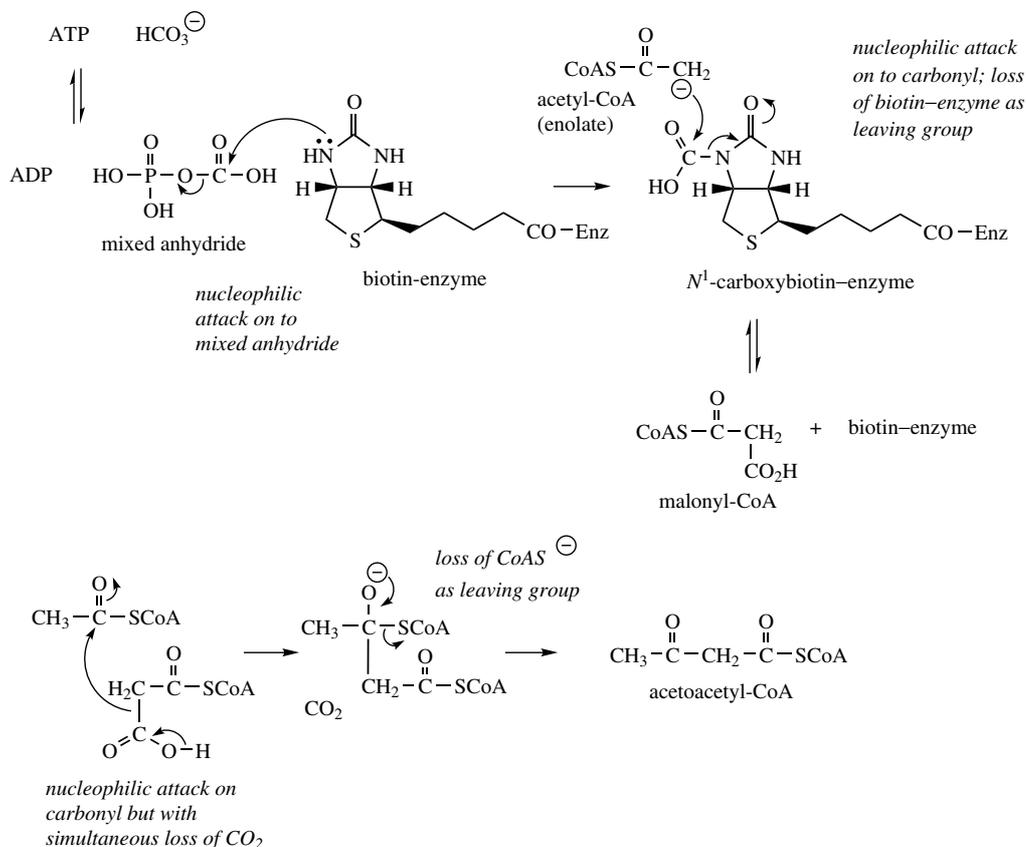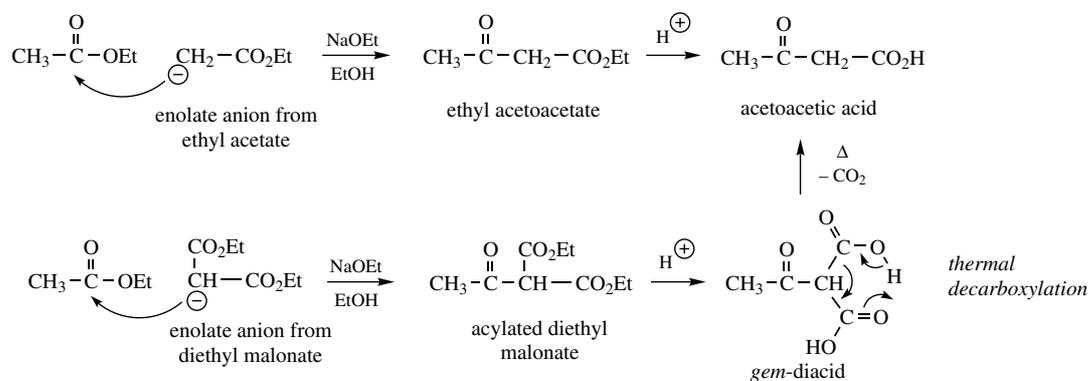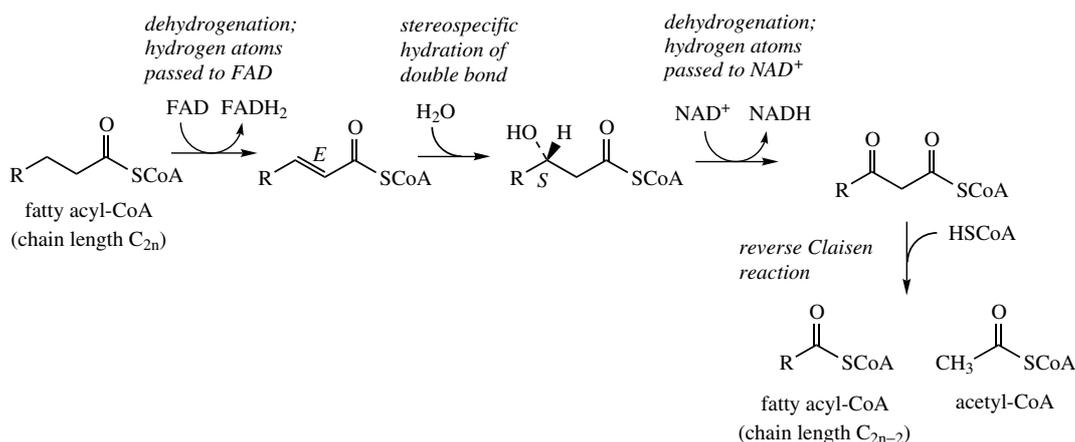

Figure 2.9

Figure 2.10

β-**Oxidation of fatty acids**

Figure 2.11

By analogy, the chemical Claisen condensation using the enolate anion from diethyl malonate in Figure 2.10 proceeds much more favourably than that using the enolate from ethyl acetate. The same acetoacetic acid product can be formed in the malonate condensation by hydrolysis of the acylated malonate intermediate and decarboxylation of the *gem*-diacid.

Both the **reverse aldol** and **reverse Claisen** reactions may be encountered in the modification of natural product molecules. Such reactions remove fragments from the basic skeleton already generated, but may extend the diversity of structures. The reverse Claisen reaction is a prominent feature of the **β-oxidation** sequence for the catabolic degradation of fatty acids (Figure 2.11),

in which a C_2 unit as acetyl-CoA is cleaved off from a fatty acid chain, leaving it two carbons shorter in length.

Schiff Base Formation and the Mannich Reaction

Formation of $C-N$ bonds is frequently achieved by condensation reactions between amines and aldehydes or ketones. A typical nucleophilic addition is followed by elimination of water to give an imine or **Schiff base** [Figure 2.12(a)]. Of almost equal importance is the reversal of this process, i.e. the hydrolysis of imines to amines and aldehydes/ketones [Figure 2.12(b)]. The imine so produced, or more likely its protonated form the

(a) Schiff base formation
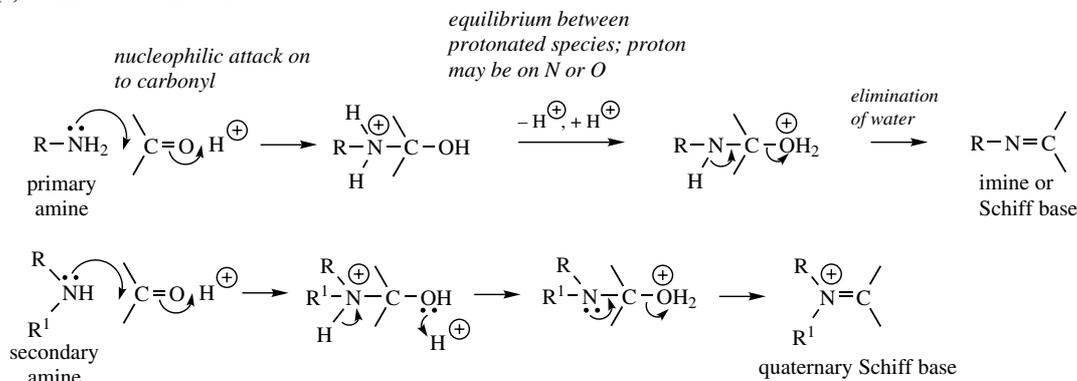
(b) Schiff base hydrolysis
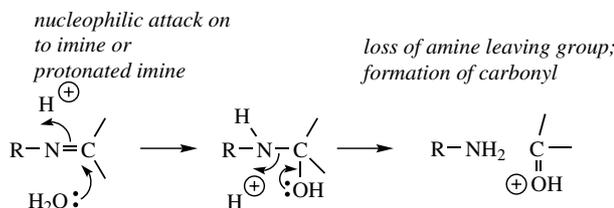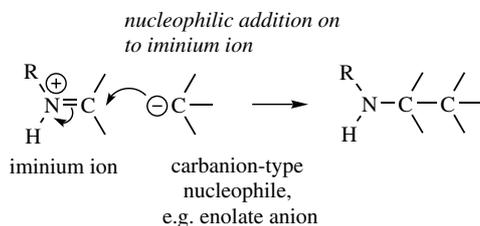

(c) Mannich reaction

Figure 2.12

iminium ion, can then act as an electrophile in a **Mannich reaction** [Figure 2.12(c)]. The nucleophile might be provided by an enolate anion, or in many examples by a suitably activated centre in an aromatic ring system. The Mannich reaction is encountered throughout alkaloid biosynthesis, and in its most general form involves combination of an amine (primary or secondary), an aldehyde or ketone, and a nucleophilic carbon. Secondary amines will react with the carbonyl compound to give an iminium ion (quaternary Schiff base) directly, and the additional protonation step is thus not necessary.

It should be appreciated that the Mannich-like addition reaction in Figure 2.12(c) is little different from nucleophilic addition to a carbonyl group. Indeed, the imine/iminium ion is merely acting as the nitrogen analogue of a carbonyl/protonated carbonyl. To take this analogy further, protons on carbon adjacent to an imine group will be acidic, as are those α to a carbonyl group, and the isomerization to the enamine shown in Figure 2.13 is analogous to keto–enol tautomerism. Just as two carbonyl compounds can react via an aldol reaction, so can two imine systems, and this is indicated in Figure 2.13. Often aldehyde/ketone substrates in enzymic reactions become covalently bonded to the enzyme through imine linkages; in so doing they lose none of the carbonyl activation as a consequence of the new form of bonding.
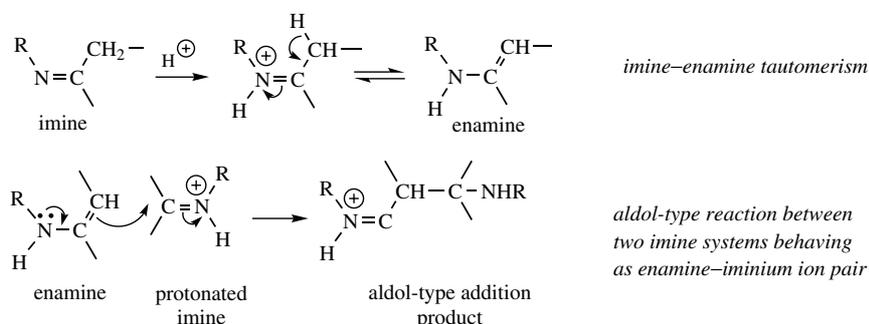
imine–enamine tautomerism

aldol-type reaction between two imine systems behaving as enamine–iminium ion pair

Figure 2.13

Transamination

Transamination is the exchange of the amino group from an amino acid to a keto acid, and provides the most common process for the introduction of nitrogen into amino acids, and for the removal of nitrogen from them. The couple **glutamic acid/2-oxoglutaric acid** are the usual donor/acceptor molecules for the amino group. Reductive amination of the Krebs cycle intermediate 2-oxoglutaric acid to glutamic acid (Figure 2.14) is responsible for the initial incorporation of nitrogen, a reaction which involves imine formation and subsequent reduction. Transamination then allows the amino group to be transferred from glutamic acid to a suitable keto acid, or in the reverse mode from an amino acid to 2-oxoglutaric acid. This reaction is dependent on the coenzyme **pyridoxal phosphate** (**PLP**) and features a Schiff base/imine intermediate (aldimine) with the aldehyde group of PLP (Figure 2.14). The α-hydrogen of the original amino acid is now made considerably more acidic and is removed, leading to the ketimine by a reprotonation process which also restores the aromaticity in the pyridine ring. The keto acid is then liberated by hydrolysis of the Schiff base function, which generates pyridoxamine phosphate. The remainder of the sequence is now a reversal of this process, and transfers the amine function from pyridoxamine phosphate to another keto acid.

Decarboxylation Reactions

Many pathways to natural products involve steps which remove portions of the carbon skeleton. Although two or more carbon atoms may be cleaved off via the reverse aldol or reverse Claisen reactions mentioned above, by far the most common degradative modification is loss of one carbon atom by a **decarboxylation** reaction. Decarboxylation is a particular feature of the biosynthetic utilization of amino acids, and it has already been indicated that several of the basic building blocks, e.g. C_6C_2N, indole.C_2N, are derived from an amino acid via loss of the carboxyl group. This decarboxylation of amino acids is also a **pyridoxal phosphate**-dependent reaction (compare transamination) and is represented as in Figure 2.15(a). This similarly depends on Schiff base formation and shares features of the transamination sequence of Figure 2.14. Decarboxylation is facilitated in the same way as loss of the α-hydrogen was facilitated for the transamination sequence. After protonation of the original α-carbon, the amine is released from the coenzyme by hydrolysis of the Schiff base function.

β-Keto acids are thermally labile and rapidly decarboxylated *in vitro* via a cyclic mechanism which proceeds through the enol form of the final ketone [Figure 2.15(b)]. Similar reactions are found in nature, though whether cyclic processes are necessary is not clear. *ortho*-Phenolic acids also decarboxylate readily *in vitro* and *in vivo*, and it is again possible to invoke a cyclic β-keto acid tautomer of the substrate. The corresponding decarboxylation of *para*-phenolic acids cannot have a cyclic transition state, but the carbonyl group in the proposed keto tautomer activates the system for decarboxylation. The acetate pathway frequently yields structures containing phenol and carboxylic acid functions, and decarboxylation reactions may thus feature as further modifications. Although the carboxyl group may originate by hydrolysis of the
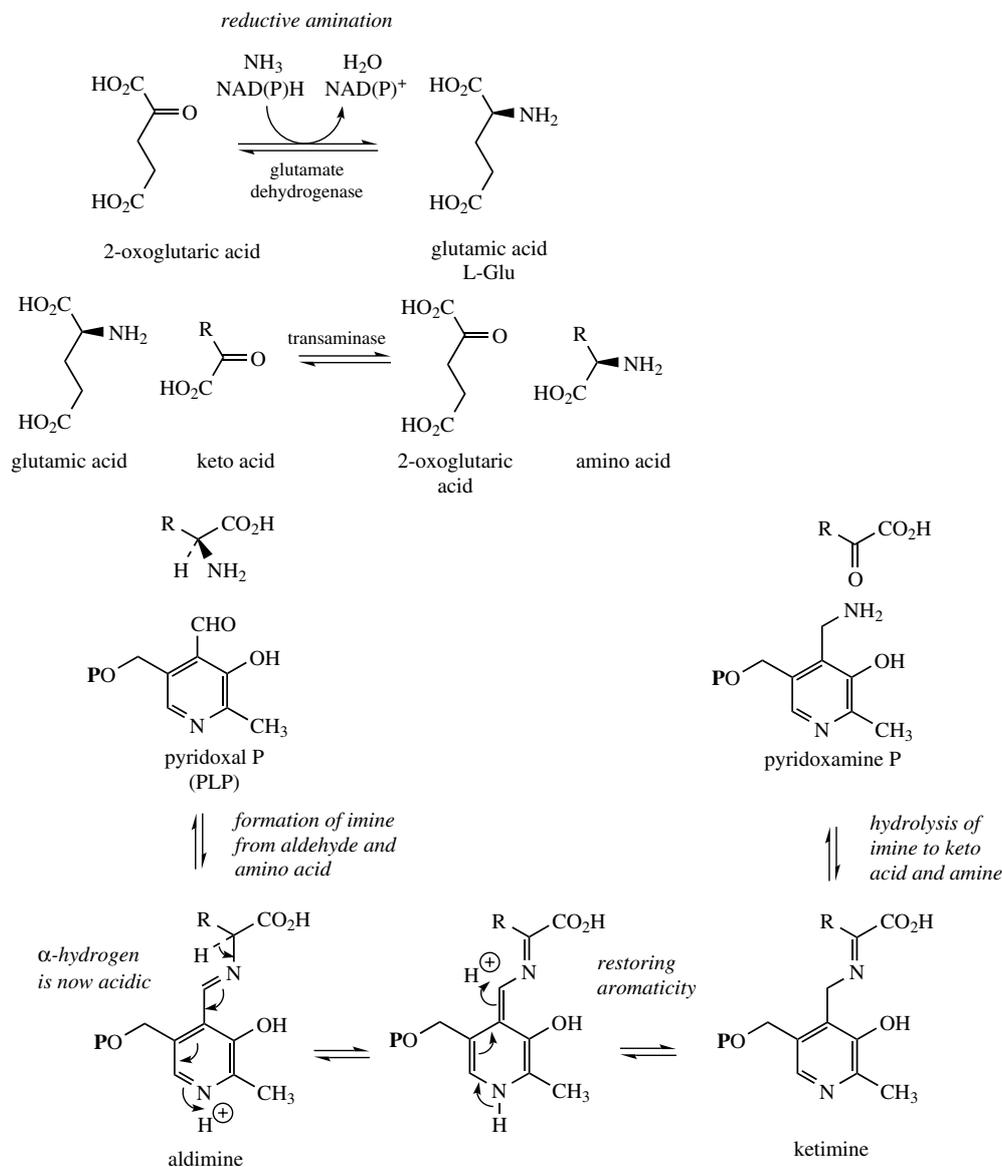
Transamination

Figure 2.14

thioester portion of the acetyl-CoA precursor, there are also occasions when a methyl group can be sequentially oxidized to a carboxyl, which then subsequently suffers decarboxylation.

Decarboxylation of **α-keto acids** is a feature of primary metabolism, e.g. pyruvic acid → acetaldehyde in glycolysis, and pyruvic acid → acetyl-CoA, an example of overall oxidative decarboxylation prior to entry of acetyl-CoA

into the Krebs cycle. Both types of reaction depend upon **thiamine diphosphate** (**TPP**). TPP is a coenzyme containing a thiazole ring, which has an acidic hydrogen and is thus capable of yielding the carbanion. This acts as a nucleophile towards carbonyl groups. Decarboxylation of pyruvic acid to acetaldehyde is depicted as in Figure 2.15(c), which process also regenerates the carbanion. In the oxidation step of oxidative
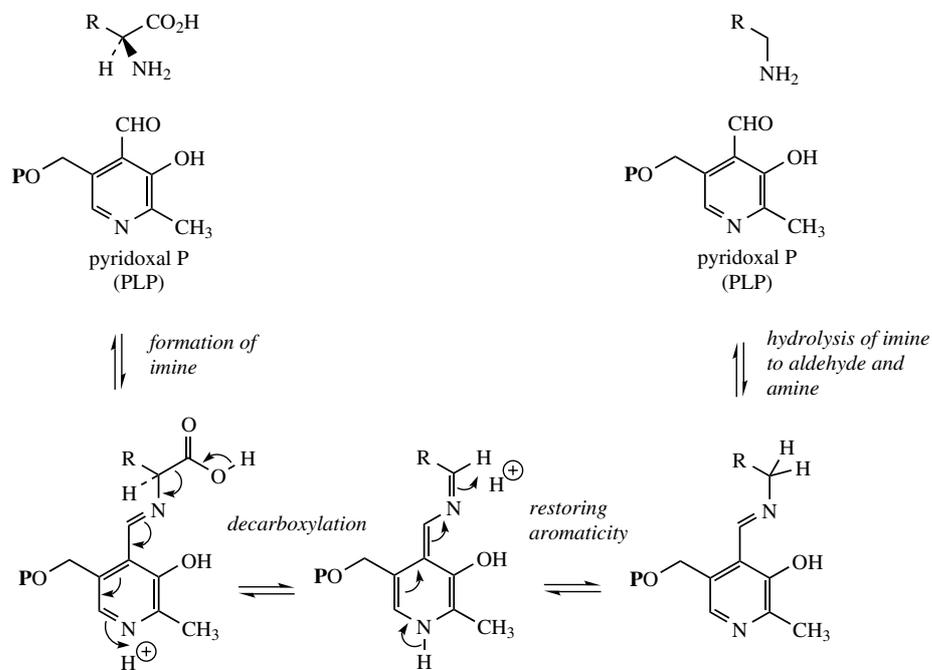
Decarboxylation reactions

(a) amino acids

pyridoxal P
(PLP)

pyridoxal P
(PLP)

formation of imine

hydrolysis of imine to aldehyde and amine

decarboxylation

restoring aromaticity

(b) β-keto acids

6-membered H-bonded system

keto-enol tautomerism

β-keto acid

intermediate enol

phenolic acid
(i.e. enol tautomer)

6-membered H-bonded system

keto tautomer
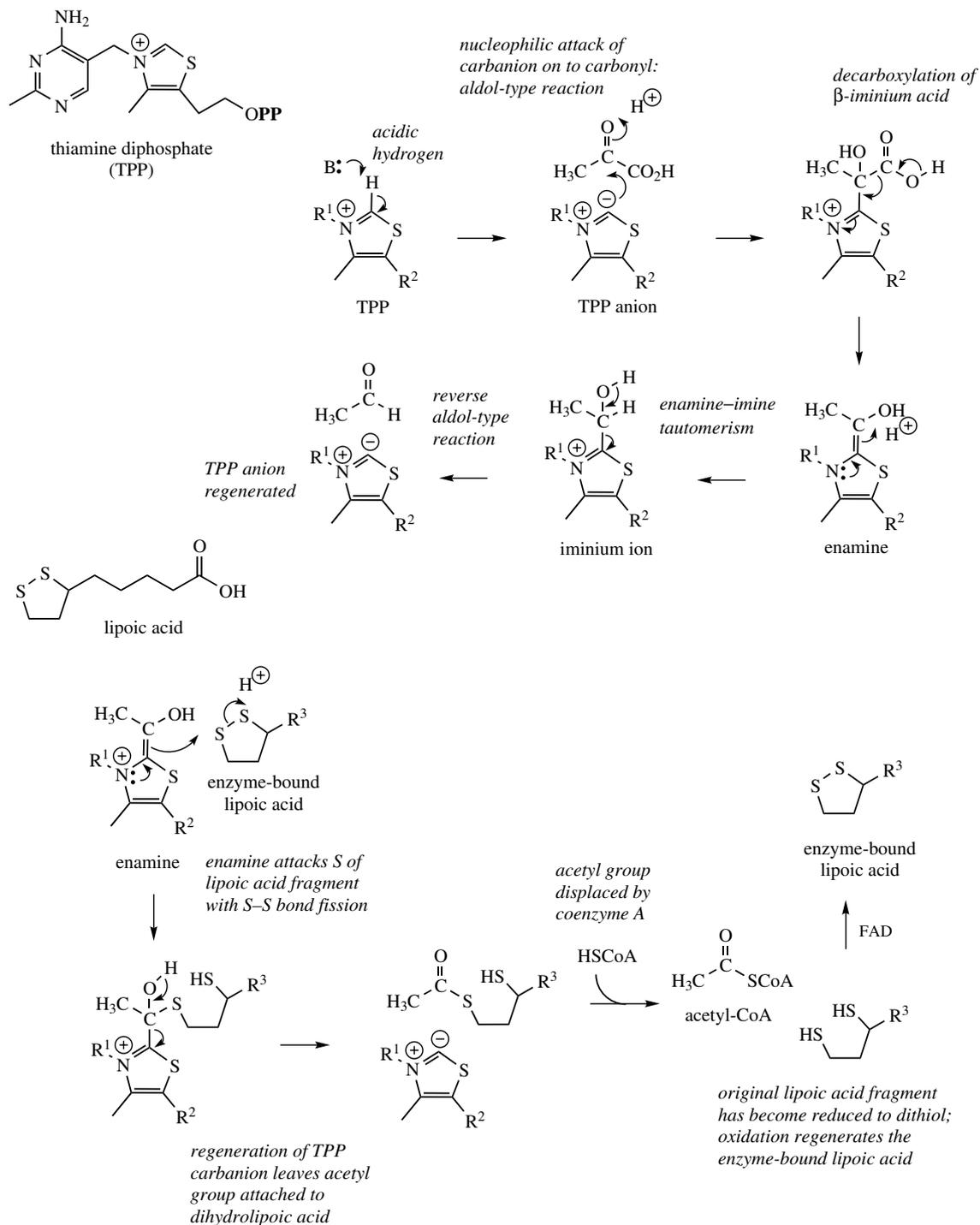≡ β-keto acid

keto tautomer
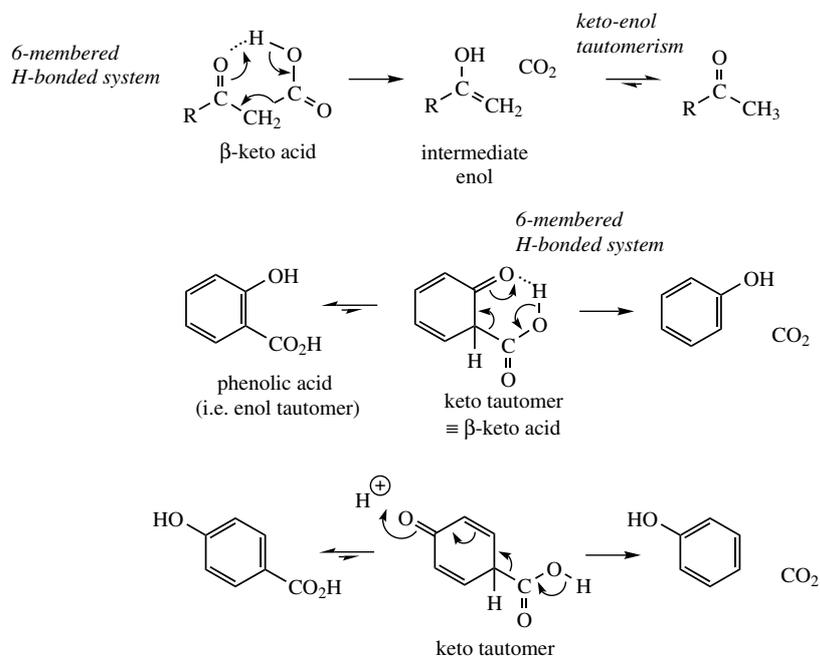
Figure 2.15

(c) α-keto acids

Figure 2.15 (*continued*)

decarboxylation, the enzyme-bound disulphide-containing coenzyme **lipoic acid** is also involved. The intermediate enamine in Figure 2.15(c), instead of accepting a proton, is used to attack a sulphur in the lipoic acid moiety with subsequent S−S bond fission, thereby effectively reducing the lipoic acid fragment. This allows regeneration of the TPP carbanion, and the acetyl group is bound to the dihydrolipoic acid. This acetyl group is then released as acetyl-CoA by displacement with the thiol coenzyme A. The bound dihydrolipoic acid fragment is then reoxidized to restore its function. An exactly equivalent reaction is encountered in the Krebs cycle in the conversion of 2-oxoglutaric acid into succinyl-CoA.

Oxidation and Reduction Reactions

Changes to the oxidation state of a molecule are frequently carried out as a secondary metabolite is synthesized or modified. The processes are not always completely understood, but the following general features are recognized. The processes may be classified according to the type of enzyme involved and their mechanism of action.
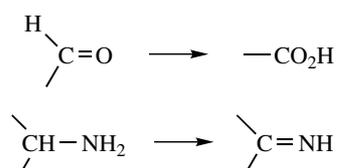
Dehydrogenases

Dehydrogenases remove two hydrogen atoms from the substrate, passing them to a suitable coenzyme acceptor. The coenzyme system involved can generally be related to the functional group being oxidized in the substrate. Thus if the oxidation process is

$$\overset{\diagdown}{\underset{\diagup}{C}}H - OH \longrightarrow \overset{\diagdown}{\underset{\diagup}{C}} = O$$
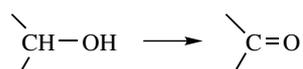
then a pyridine nucleotide, **nicotinamide adenine dinucleotide** (**NAD**$^+$) or **nicotinamide adenine dinucleotide phosphate** (**NADP**$^+$), tends to be utilized as hydrogen acceptor. One hydrogen from the substrate (that bonded to carbon) is transferred as hydride to the coenzyme, and the other, as a proton, is passed to the medium (Figure 2.16). NAD(P)$^+$ may also be used in the oxidations

$$\overset{H}{\underset{\diagup}{\diagdown}} C = O \longrightarrow -CO_2H$$

$$\overset{\diagdown}{\underset{\diagup}{C}}H - NH_2 \longrightarrow \overset{\diagdown}{\underset{\diagup}{C}} = NH$$

The reverse reaction, i.e. reduction, is also indicated in Figure 2.16, and may be compared with the chemical reduction process using complex metal hydrides, e.g. LiAlH$_4$ or NaBH$_4$, namely nucleophilic addition of hydride and subsequent protonation. The reduced forms NADH and NADPH are conveniently regarded as hydride-donating
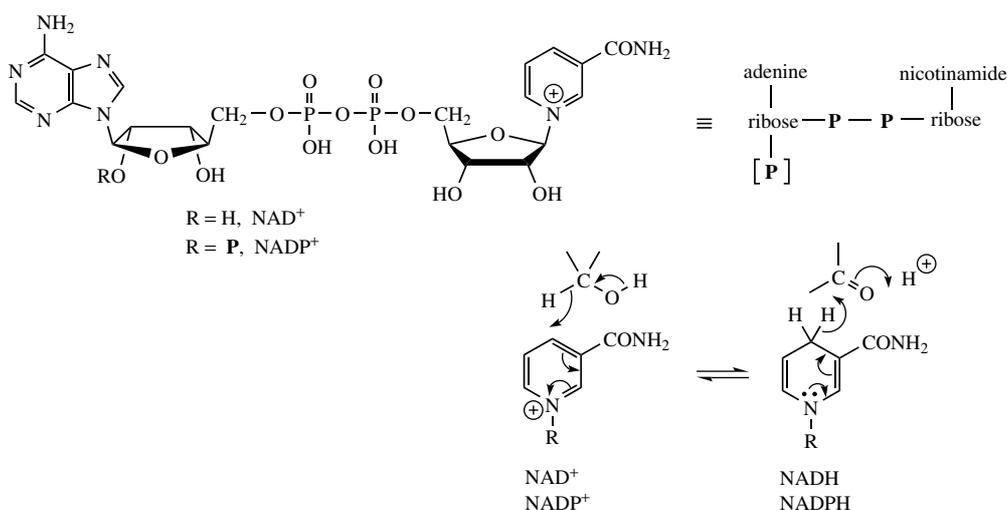
Dehydrogenases: NAD$^+$ and NADP$^+$

Figure 2.16
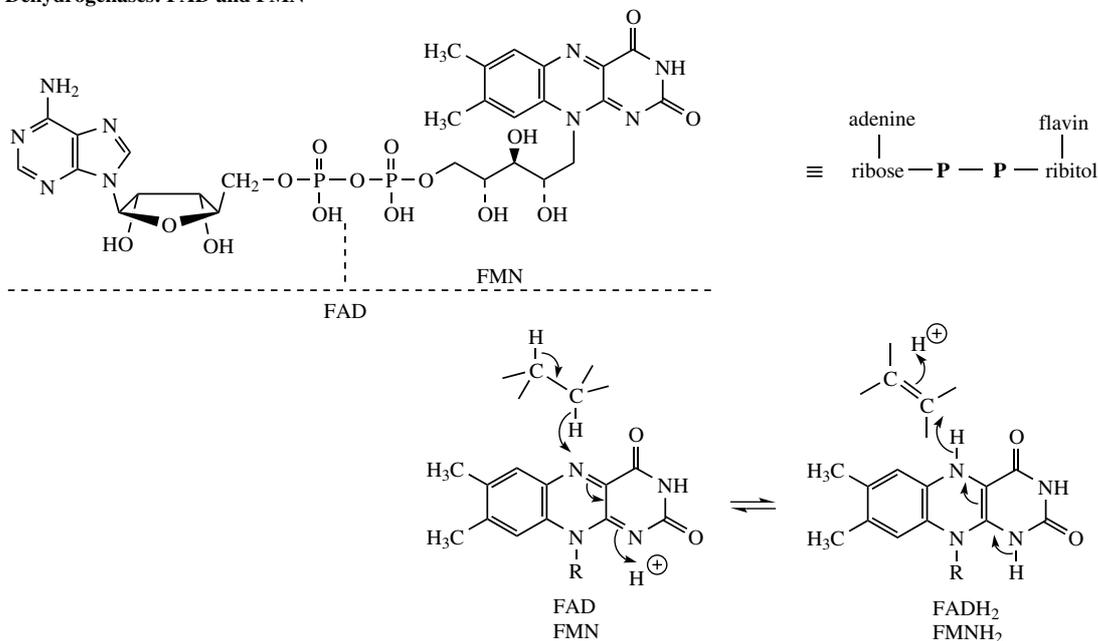
Dehydrogenases: FAD and FMN

Figure 2.17

reducing agents. In practice, NADPH is generally employed in reductive processes, whilst NAD^+ is used in oxidations.

Should the oxidative process be the conversion

$$-CH_2-CH_2- \longrightarrow -CH=CH-$$

the coenzyme used as acceptor is usually a flavin nucleotide, **flavin adenine dinucleotide** (**FAD**) or **flavin mononucleotide** (**FMN**). These entities are bound to the enzyme in the form of a flavoprotein, and take up two hydrogen atoms, represented in Figure 2.17 as being derived by addition of hydride from the substrate and a proton from the medium. Alternative mechanisms have also been proposed, however. Reductive sequences involving flavoproteins may be represented as the reverse reaction in Figure 2.17. NADPH may also be employed as a coenzyme in the reduction of a carbon–carbon double bond.

These oxidation reactions employing pyridine nucleotides and flavoproteins are especially important in primary metabolism in liberating energy from fuel molecules in the form of ATP. The reduced coenzymes formed in the process are normally reoxidized via the electron transport chain

of oxidative phosphorylation, so that the hydrogen atoms eventually pass to oxygen giving water.

Oxidases

Oxidases also remove hydrogen from a substrate, but pass these atoms to molecular oxygen or to hydrogen peroxide, in both cases forming water. Oxidases using hydrogen peroxide are termed **peroxidases**. Mechanisms of action vary and need not be considered here. Important transformations in secondary metabolism include the oxidation of *ortho-* and *para*-quinols to quinones (Figure 2.18), and the peroxidase-induced phenolic oxidative coupling processes (see page 28).

Mono-oxygenases

Oxygenases catalyse the direct addition of oxygen from molecular oxygen to the substrate. They are subdivided into mono- and di-oxygenases according to whether just one or both of the oxygen atoms are introduced into the substrate. With **mono-oxygenases**, the second oxygen atom from O_2 is reduced to water by an appropriate hydrogen

Oxidases

ortho-quinol   *ortho*-quinone

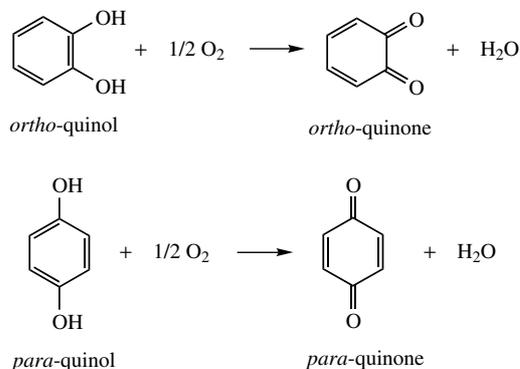*para*-quinol   *para*-quinone

Figure 2.18

donor, e.g. NADH, NADPH, or **ascorbic acid** (vitamin C). In this respect they may also be considered to behave as oxidases, and the term 'mixed-function oxidase' is also used for these enzymes. Especially important examples of these enzymes are the **cytochrome P-450-dependent mono-oxygenases**. These are frequently involved in biological hydroxylations, either in biosynthesis, or in the mammalian detoxification and metabolism of foreign compounds such as drugs, and such enzymes are thus termed '**hydroxylases**'. **Cytochrome P-450** is named after its intense absorption band at 450 nm when exposed to CO, which is a powerful inhibitor of these enzymes. It contains an iron−porphyrin complex (haem), which is bound to the enzyme, and a redox change involving the Fe atom allows binding and the

cleavage of an oxygen atom. Many such systems have been identified, capable of hydroxylating aliphatic or aromatic systems, as well as producing epoxides from alkenes (Figure 2.19). In most cases, NADPH features as hydrogen donor.

Aromatic hydroxylation catalysed by mono-oxygenases (including cytochrome P-450 systems) probably involves arene oxide (epoxide) intermediates (Figure 2.20). An interesting consequence of this mechanism is that when the epoxide opens up, the hydrogen atom originally attached to the position which becomes hydroxylated can migrate to the adjacent carbon on the ring. A high proportion of these hydrogen atoms is subsequently retained in the product, even though enolization allows some loss of this hydrogen. This migration is known as the **NIH shift**, having been originally observed at the National Institute of Health, Bethesda, MD, USA.
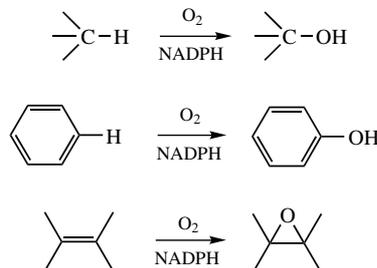
Mono-oxygenases

Figure 2.19

NIH shift

arene oxide

hydride migration NIH shift

enolization

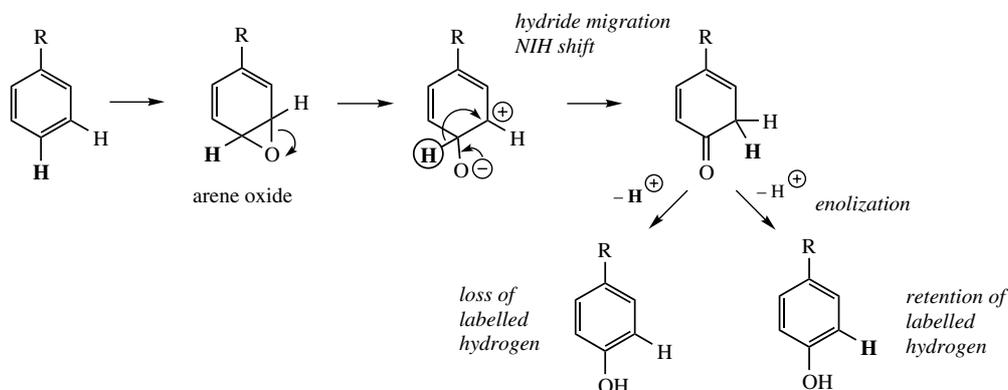*loss of labelled hydrogen*

retention of labelled hydrogen

Figure 2.20

The oxidative cyclization of an *ortho*-hydroxy-methoxy-substituted aromatic system giving a **methylenedioxy group** is also known to involve a cytochrome P-450-dependent mono-oxygenase. This enzyme hydroxylates the methyl to yield a formaldehyde hemiacetal intermediate, which can cyclize to the methylenedioxy bridge (the acetal of formaldehyde) by an ionic mechanism (Figure 2.21).

Dioxygenases

Dioxygenases introduce both atoms from molecular oxygen into the substrate, and are frequently involved in the cleavage of bonds, including aromatic rings. Cyclic peroxides (dioxetanes) are likely to be intermediates (Figure 2.22). Oxidative cleavage of aromatic rings typically employs catechol (1,2-dihydroxy) or quinol (1,4-dihydroxy) substrates, and in the case of catechols, cleavage may be between or adjacent to the two hydroxyls, giving products containing aldehyde and/or carboxylic acid functionalities (Figure 2.22).

Some dioxygenases utilize two acceptor substrates and incorporate one oxygen atom into each. Thus, **2-oxoglutarate-dependent dioxygenases** hydroxylate one substrate, whilst also transforming 2-oxoglutarate into succinate with the release of CO_2 (Figure 2.23). 2-Oxoglutarate-dependent dioxygenases also require as cofactors
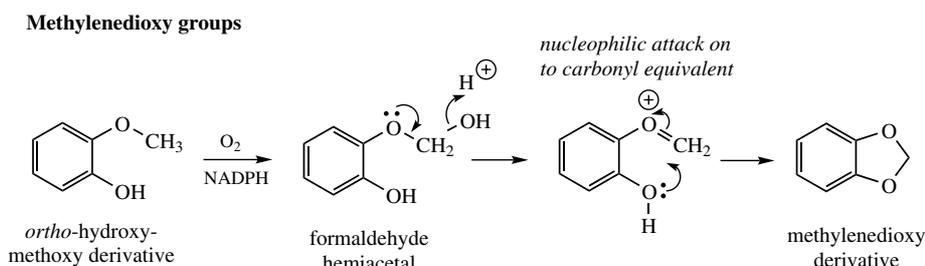
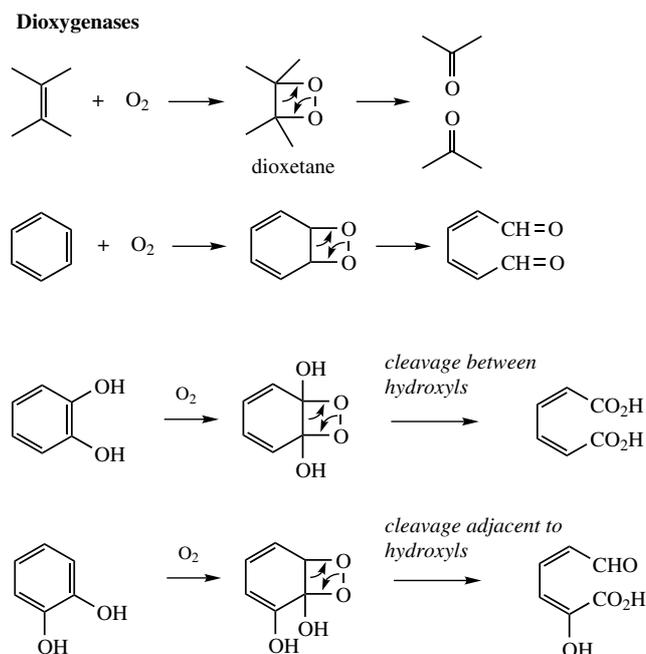**Methylenedioxy groups**

Figure 2.21
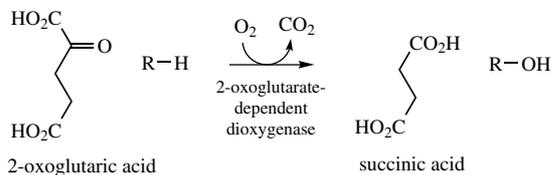
Dioxygenases

Figure 2.22

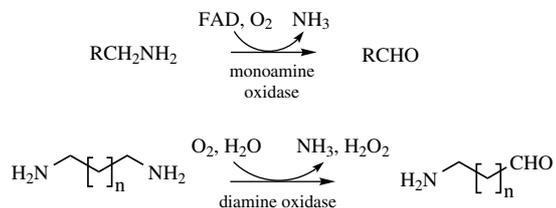**2-Oxoglutarate-dependent dioxygenases**

Figure 2.23

Amine oxidases

Figure 2.24

Fe^{2+} to generate an enzyme-bound iron−oxygen complex, and **ascorbic acid** (vitamin C) to subsequently reduce this complex.

Amine Oxidases

In addition to the oxidizing enzymes outlined above, those which transform an amine into an aldehyde, the **amine oxidases**, are frequently involved in metabolic pathways. These include **monoamine oxidases** and **diamine oxidases**. Monoamine oxidases utilize a flavin nucleotide, typically FAD, and molecular oxygen, and involve initial dehydrogenation to an imine, followed by hydrolysis to the aldehyde and ammonia (Figure 2.24). Diamine oxidases require a diamine substrate, and oxidize at one amino group using molecular oxygen to give the corresponding aldehyde. Hydrogen peroxide and ammonia are the other products formed. The aminoaldehyde so formed then has the potential to be transformed into a cyclic imine via Schiff base formation.

Baeyer−Villiger Oxidations

The chemical oxidation of ketones by peracids, the **Baeyer−Villiger oxidation**, yields an ester, and the process is known to involve migration of an alkyl group from the ketone (Figure 2.25). For comparable ketone → ester conversions known to occur in biochemistry, cytochrome-P-450- or FAD-dependent enzymes requiring NADPH and O_2 appear to be involved. This leads to formation of a peroxy−enzyme complex and a mechanism similar to that for the chemical Baeyer−Villiger oxidation may thus operate. The oxygen atom introduced thus originates from O_2.

Phenolic Oxidative Coupling

Many natural products are produced by the coupling of two or more phenolic systems, in a process readily rationalized by means of free radical reactions. The reactions can be brought
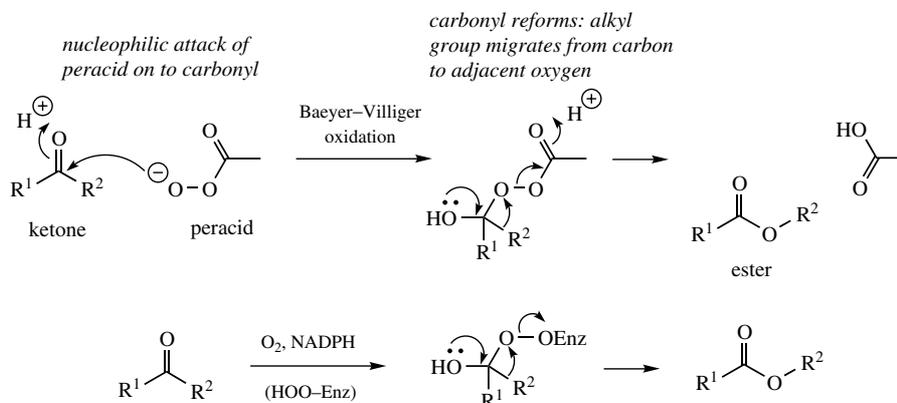
Baeyer−Villiger oxidations

Figure 2.25
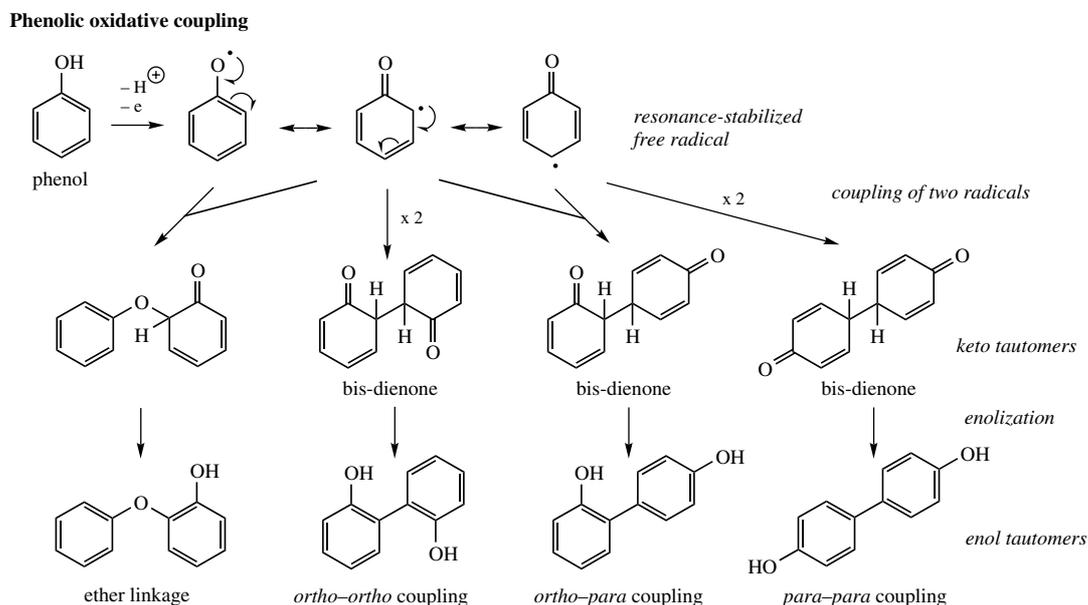
Phenolic oxidative coupling

Figure 2.26

about by oxidase enzymes, including peroxidase and laccase systems, known to be radical generators. Other enzymes catalysing **phenolic oxidative coupling** have been characterized as cytochrome P-450-dependent proteins, requiring NADPH and O_2 cofactors, but no oxygen is incorporated into the substrate. A one-electron oxidation of a phenol gives the free radical, and the unpaired electron can then be delocalized via resonance forms in which the free electron is dispersed to positions *ortho* and *para* to the original oxygen function (Figure 2.26). Coupling of two of these mesomeric structures gives a range of dimeric systems as exemplified in Figure 2.26. The final products indicated are then derived by enolization, which restores aromaticity to the rings. Thus, carbon–carbon bonds involving positions *ortho* or *para* to the original phenols, or ether linkages, may be formed. The reactive dienone systems formed as intermediates may in some cases be attacked by other nucleophilic groupings, extending the range of structures ultimately derived from this basic reaction sequence.

Glycosylation Reactions

The widespread occurrence of **glycosides** and **polysaccharides** requires processes for attaching sugar units to a suitable atom of an aglycone to give a glycoside, or to another sugar giving a polysaccharide. Linkages tend to be through oxygen, although they are not restricted to oxygen, since S-, N-, and C-glycosides are well known. The agent for glycosylation is a **uridine diphosphosugar**, e.g. **UDPglucose**. This is synthesized from glucose 1-phosphate and UTP, and then the glucosylation process can be envisaged as a simple S_N2 nucleophilic displacement reaction [Figure 2.27(a)]. Since UDPglucose has its leaving group in the α-configuration, the product has the β-configuration, as is most commonly found in natural glucosides. Note, however, that many important carbohydrates, e.g. sucrose and starch, possess α-linkages, and these appear to originate via double S_N2 processes (see page 470). Other UDPsugars, e.g. UDPgalactose or UDPxylose, are utilized in the synthesis of glycosides containing different sugar units.

The hydrolysis of glycosides is achieved by specific hydrolytic enzymes, e.g. β-glucosidase for β-glucosides and β-galactosidase for β-galactosides. These enzymes mimic the readily achieved acid-catalysed processes [Figure 2.27(b)]. Under acidic conditions, the α- and β-anomeric hemiacetal forms can also equilibrate via the open chain
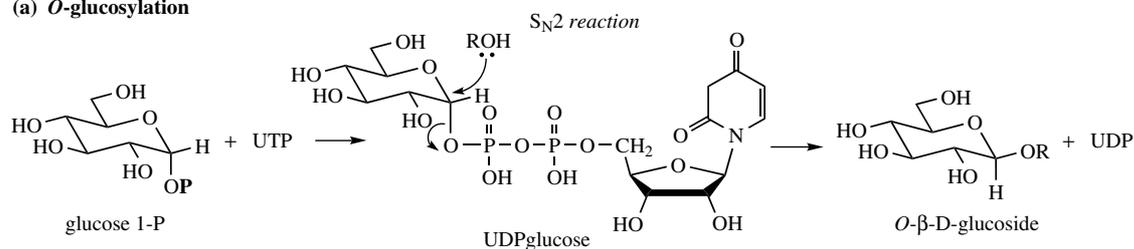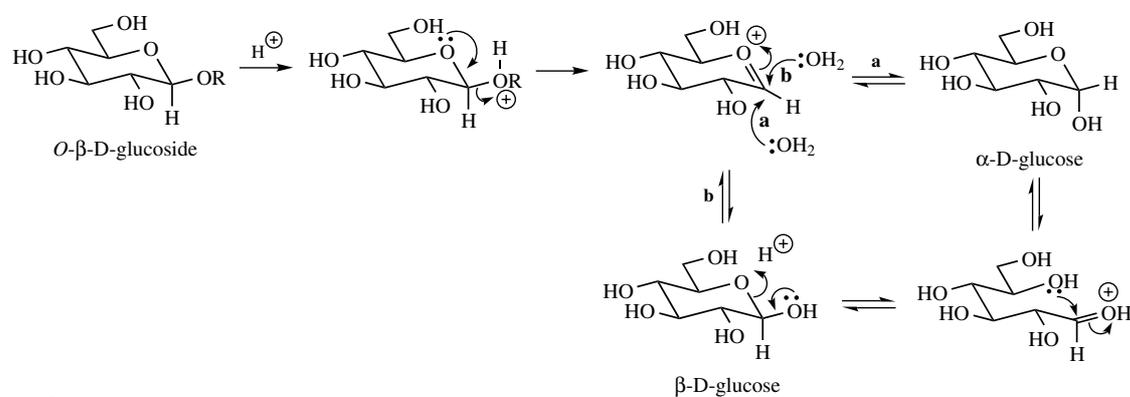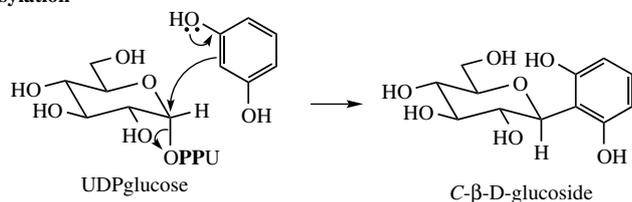
Glycosylation reactions

(a) *O*-glucosylation

glucose 1-P

UDPglucose

O-β-D-glucoside

S_N2 *reaction*

(b) hydrolysis of *O*-glucosides

O-β-D-glucoside

α-D-glucose

β-D-glucose

(c) *C*-glucosylation

UDPglucose

C-β-D-glucoside

Figure 2.27

sugar. Of particular importance is that although *O*-, *N*-, and *S*-glycosides may be hydrolysed by acid, *C*-glycosides are stable to acid. *C*-Glycosides are produced in a similar manner to the *C*-alkylation process described above, where a suitable nucleophilic carbon is available, e.g. aromatic systems activated by phenol groups [Figure 2.27(c)]. The resultant *C*-glycoside thus contains a new carbon–carbon linkage, and cleavage would require oxidation, not hydrolysis.

SOME VITAMINS ASSOCIATED WITH THE CONSTRUCTION MECHANISMS
Vitamin B₁

Vitamin B₁ (thiamine) (Figure 2.28) is a water-soluble vitamin with a pyrimidinylmethylthia-zolium structure. It is widely available in the diet, with cereals, beans, nuts, eggs, yeast, and vegetables providing sources. Wheat germ and yeast have very high levels. Dietary deficiency leads to beriberi, characterized by neurological disorders, loss of appetite, fatigue,
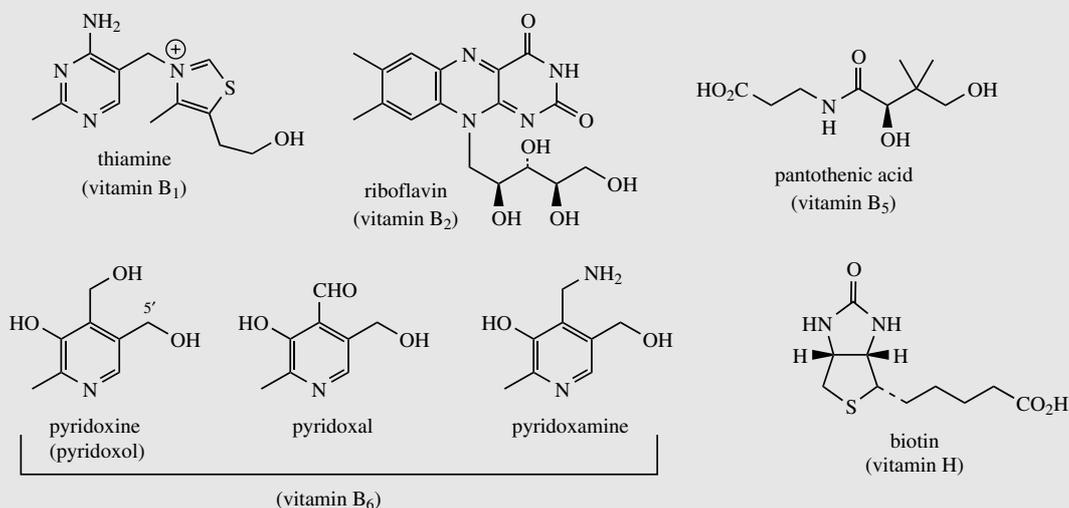
(*Continues*)

(*Continued*)

Figure 2.28

and muscular weakness. **Thiamine** is produced synthetically, and foods such as cereals are often enriched. The vitamin is stable in acid solution, but decomposes above pH 5, and is also partially decomposed during normal cooking. As thiamine diphosphate, vitamin B_1 is a coenzyme for pyruvate dehydrogenase which catalyses the oxidative decarboxylation of pyruvate to acetyl-CoA (see page 21), and also for transketolase which transfers a two-carbon fragment between carbohydrates in the pentose phosphate pathway (see page 446). Accordingly, this is a very important component in carbohydrate metabolism.

Vitamin B_2

Vitamin B_2 (riboflavin) (Figure 2.28) is a water-soluble vitamin having an isoalloxazine ring linked to D-ribitol. It is widely available in foods, including liver, kidney, dairy products, eggs, meat, and fresh vegetables. Yeast is a particularly rich source. It is stable in acid solution, not decomposed during cooking, but is sensitive to light. Riboflavin may be produced synthetically, or by fermentation using the yeastlike fungi *Eremothecium ashbyii* and *Ashbya gossypii*. Dietary deficiency is uncommon, but manifests itself by skin problems and eye disturbances. Riboflavin is a component of FMN (flavin mononucleotide) and FAD (flavin adenine dinucleotide), coenzymes which play a major role in oxidation–reduction reactions (see page 25). Many key enzymes containing riboflavin (flavoproteins) are involved in metabolic pathways. Since riboflavin contains ribitol and not ribose in its structure, FAD and FMN are not strictly nucleotides, though this nomenclature is commonly accepted and used.

Vitamin B_5

Vitamin B_5 (pantothenic acid) (Figure 2.28) is a very widely distributed water-soluble vitamin, though yeast, liver, and cereals provide rich sources. Even though animals must obtain the vitamin through the diet, pantothenic acid deficiency is rare, since most foods provide

(*Continues*)

(*Continued*)

adequate quantities. Its importance in metabolism is as part of the structure of coenzyme A (see page 16), the carrier molecule essential for carbohydrate, fat, and protein metabolism. Pantothenic acid is specifically implicated in enzymes responsible for the biosynthesis of fatty acids (see page 36), polyketides (page 62) and some peptides (page 421).

Vitamin B$_6$

Vitamin B$_6$ covers the three pyridine derivatives pyridoxine (pyridoxol), pyridoxal, and pyridoxamine, and also their 5′-phosphates (Figure 2.28). These are water-soluble vitamins, pyridoxine predominating in plant materials, whilst pyridoxal and pyridoxamine are the main forms in animal tissues. Meat, salmon, nuts, potatoes, bananas, and cereals are good sources. A high proportion of the vitamin activity can be lost during cooking, but a normal diet provides an adequate supply. Vitamin B$_6$ deficiency is usually the result of malabsorption, or may be induced by some drug treatments where the drug may act as an antagonist or increase its renal excretion as a side-effect. Symptoms of deficiency are similar to those of niacin (vitamin B$_3$) and riboflavin (vitamin B$_2$) deficiencies, and include eye, mouth, and nose lesions, and neurological changes. Synthetic **pyridoxine** is used for supplementation. Pyridoxal 5′-phosphate is a coenzyme for a large number of enzymes, particularly those involved in amino acid metabolism, e.g. in transamination (see page 20) and decarboxylation (see page 20). The production of the neurotransmitter γ-aminobutyric acid (GABA) from glutamic acid is an important pyridoxal-dependent reaction.

Vitamin B$_{12}$

Vitamin B$_{12}$ (cobalamins) (Figure 2.29) are extremely complex structures based on a corrin ring, which, although similar to the porphyrin ring found in haem, chlorophyll, and cytochromes,

R = CN, cyanocobalamin (vitamin B$_{12}$)
R = OH, hydroxocobalamin (vitamin B$_{12a}$)
R = H$_2$O, aquocobalamin (vitamin B$_{12b}$)
R = NO$_2$, nitritocobalamin (vitamin B$_{12c}$)
R = Me, methylcobalamin (methyl vitamin B$_{12}$)

5′-deoxyadenosylcobalamin (coenzyme B$_{12}$)

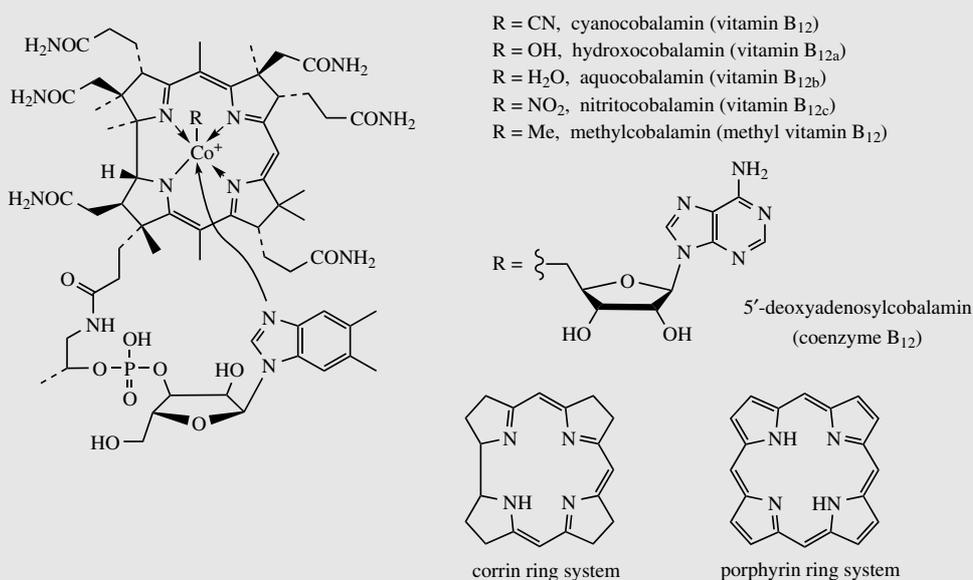corrin ring system

porphyrin ring system

Figure 2.29

(*Continues*)

(Continued)

has two of the pyrrole rings directly bonded. The central metal atom is cobalt; haem and cytochromes have iron, whilst chlorophyll has magnesium. Four of the six coordinations are provided by the corrin ring nitrogens, and a fifth by a dimethylbenzimidazole moiety. The sixth is variable, being cyano in **cyanocobalamin** (vitamin B_{12}), hydroxyl in **hydroxocobalamin** (vitamin B_{12a}), or other anions may feature. Cyanocobalamin is actually an artefact formed as a result of the use of cyanide in the purification procedures. The physiologically active coenzyme form of the vitamin is 5′-deoxyadenosylcobalamin (coenzyme B_{12}). Vitamin B_{12} appears to be entirely of microbial origin, with intestinal flora contributing towards human dietary needs. The vitamin is then stored in the liver, and animal liver extract has been a traditional source. Commercial supplies are currently obtained by semi-synthesis from the total cobalamin extract of *Streptomyces griseus*, *Propionibacterium* species, or other bacterial cultures. This material can be converted into cyanocobalamin or hydroxocobalamin. The cobalamins are stable when protected against light. Foods with a high vitamin B_{12} content include liver, kidney, meat, and seafood. Vegetables are a poor dietary source, and strict vegetarians may therefore risk deficiencies. Insufficient vitamin B_{12} leads to pernicious anaemia, a disease that results in nervous disturbances and low production of red blood cells, though this is mostly due to lack of the gastric glycoprotein (intrinsic factor) which complexes with the vitamin to facilitate its absorption. Traditionally, daily consumption of raw liver was used to counteract the problem. Cyanocobalamin, or preferably hydroxocobalamin which has a longer lifetime in the body, may be administered orally or by injection to counteract deficiencies. Both agents are converted into coenzyme B_{12} in the body. Coenzyme B_{12} is a cofactor for a number of metabolic rearrangements, such as the conversion of methylmalonyl-CoA into succinyl-CoA in the oxidation of fatty acids with an odd number of carbon atoms, and for methylations, such as in the biosynthesis of methionine.

Vitamin H

Vitamin H (biotin) (Figure 2.28) is a water-soluble vitamin found in eggs, liver, kidney, yeast, cereals, and milk, and is also produced by intestinal microflora so that dietary deficiency is rare. Deficiency can be triggered by a diet rich in raw egg white, in which a protein, avidin, binds biotin so tightly so that it is effectively unavailable for metabolic use. This affinity disappears by cooking and hence denaturing the avidin. Biotin deficiency leads to dermatitis and hair loss. The vitamin functions as a carboxyl carrier, binding CO_2 via a carbamate link, then donating this in carboxylase reactions, e.g. carboxylation of acetyl-CoA to malonyl-CoA (see page 17), of propionyl-CoA to methylmalonyl-CoA (see page 92), and of pyruvate to oxaloacetate during gluconeogenesis.

FURTHER READING

Natural Products, Biosynthesis

Mann J (1994) *Chemical Aspects of Biosynthesis*. Oxford Chemistry Primers, Oxford.

Mann J, Davidson RS, Hobbs JB, Banthorpe DV and Harborne JB (1994) *Natural Products: Their Chemistry and Biological Significance*. Longman, Harlow.

Torssell KBG (1997) *Natural Product Chemistry. A Mechanistic, Biosynthetic and Ecological Approach*. Apotekarsocieteten, Stockholm.

Vitamins

Battersby AR (2000) Tetrapyrroles: The pigments of life. *Nat Prod Rep* **17**, 507–526.

Burdick D (1998) Vitamins [pyridoxine (B_6)]. *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol **25**. Wiley, New York, pp 116–132.

Burdick D (1998) Vitamins [thiamine (B_1)]. *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol **25**. Wiley, New York, pp 152–171.

Kingston R (1999) Supplementary benefits? *Chem Brit* **35** (7), 29–32.

Outten RA (1998) Vitamins (biotin). *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol **25**. Wiley, New York, pp 48–64.

Rawalpally TR (1998) Vitamins (pantothenic acid). *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol **25**. Wiley, New York, pp 99–116.

Scott JW (1998) Vitamins (vitamin B_{12}). *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol **25**. Wiley, New York, pp 192–217.

Yoneda F (1998) Vitamins [riboflavin (B_2)]. *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th edn, Vol **25**. Wiley, New York, pp 132–152.