Chapter 5

BIOCHEMISTRY OF TERPENOIDS: MONOTERPENES, SESQUITERPENES AND DITERPENES*

Mohamed Ashour\(^1\), Michael Wink\(^1\) and Jonathan Gershenzon\(^2\)

\(^1\)Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Heidelberg, Germany
\(^2\)Max-Planck-Institute of Chemical Ecology, Jena, Germany

Abstract: Terpenoids represent the largest class of secondary metabolites and usually do not contain nitrogen or sulfur in their structures. Many terpenoids serve as defence compounds against microbes and herbivores and/or are signal molecules to attract pollinating insects, fruit-dispersing animals or predators which can destroy insect herbivores. As a consequence, many terpenoids have pronounced pharmacological activities and are therefore interesting for medicine and biotechnology. The first part of the biosynthesis is the generation of a C5 unit, such as isopentenyl diphosphate (IPP) or dimethylallyl diphosphate (DMAPP). Two independent pathways have been discovered that can produce the C5 unit: the mevalonate and the methylerythritol phosphate (MEP) pathway. Depending on the number of C5 units, we distinguish homoterpenes C5, monoterpenes including iridoids (C10), sesquiterpenes (C15), diterpenes (C20), sesterterpenes (C25), triterpenes (including steroids) (C30), tetraterpenes (C40) and polyterpenes (>C40). The biosynthesis (including enzymes, genes and their regulation) of mevalonate and the methylerythritol phosphate pathway and the consecutive pathways leading to mono-, sesqui- and diterpenes are discussed in this chapter in detail.

Keywords: biosynthesis; genes; monoterpenes; sesquiterpenes; diterpenes; mevalonate pathway; methylerythritol phosphate pathway

* This chapter is an update from an earlier version from J. Gershenzon and W. Kreis printed in the first edition in 1999.
5.1 Introduction

The largest class of plant secondary metabolites is undoubtedly that of the terpenoids or isoprenoids. Over 36,000 individual members of this class have been reported (Buckingham, 2007) and new structures are currently being added at the rate of about 1000 every year. Compilations of newly described terpenoids appear periodically in *Natural Product Reports* (e.g. Grayson, 2000; Hanson, 2005; Connolly and Hill, 2008; Fraga, 2008). Terpenoids are not only numerous, but also extremely variable in structure, exhibiting hundreds of different carbon skeletons and a large assortment of functional groups. In spite of such diversity, all terpenoids are unified by a common mode of biosynthesis: the fusion of C₅ units with an isopentenoid structure.

Since the origins of organic chemistry, terpenoids have been a source of fascination for many practitioners of this discipline. However, the basic structural unity of terpenoids has only been appreciated since the end of the past century, when pioneers, such as the German Otto Wallach, discovered that some members of this class could be pyrolyzed to give isoprene, a C₅ diene with an isopentenoid skeleton (Fig. 5.1). These studies gave rise to the so-called isoprene rule, which states that all terpenoids are derived from the ordered, head-to-tail joining of isoprene units. More recent workers have refined the original concept, recognizing that non-head-to-tail condensations of isoprene units also occur in a few secondary metabolites (pyrethrins) which seem to be limited almost exclusively to members of family Asteraceae. Substantial structural rearrangements or loss of carbons during biosynthesis have been observed (Dewick, 2002). Nevertheless, the original isoprene rule was a very useful concept in determining the structures of many unknown substances and assessing their biogenetic origin. In this context, terpenoids have frequently been referred to as isoprenoids, and the terms isoprenoids, terpenoids and terpenes are now used interchangeably.

The classification of terpenoids is based on the number of isoprenoid units present in their structure. The largest categories are those made up of compounds with two isoprenoid units (monoterpenes), three isoprenoid units (sesquiterpenes), four isoprenoid units (diterpenes), five isoprenoid units (sesterterpenes), six isoprenoid units (triterpenes) and eight isoprenoid units (tetraterpenes) (Table 5.1). Although the biosynthesis is based on a unit of five-carbon atoms terpenoid nomenclature is based on a unit of ten carbon atoms since the C₁₀ terpenoids were once thought to be the smallest naturally occurring representatives of this class. Designation of the C₁₀ terpenoids as mono-(‘one’)-terpenes made it necessary to name the subsequently described C₅ terpenes as hemi-(‘half’)-terpenes, the C₁₅ terpenes as sesqui-(‘one-and-a-half’)-terpenes and so on. In this section, the biosynthesis and functional significance of the lower (C₅–C₂₀) terpenes are surveyed, with emphasis on the major advances in the past five years. Triterpenes (C₃₀), cardiac glycosides and steroid saponins are treated in Chapter 6. Relevant monographs of outstanding coverage and quality (Cane, 1998; Leeper
Figure 5.1 Examples of terpenoids that are of commercial importance or whose functional role in plants has recently been investigated. Isoprene may stabilize membranes at high temperatures. Camphor, artemisinin and paclitaxel (taxol) are valuable pharmaceuticals. The other three compounds appear to be involved in plant defence: pulegone is toxic to herbivores; polygodial is a herbivore feeding deterrent; and (3E)-4,8-dimethyl-1,3,7-nonatriene, a C₁₁ homoterpene, functions to attract herbivore enemies to herbivore-damaged plants.

and Vederas 2000) and several excellent individual reviews (Chappell, 1995; McCaskill and Croteau, 1997) have covered many aspects of this subject.

5.2 Function

The enormous structural diversity of the terpenoids is almost matched by their functional variability. Terpenoids have well-established roles in almost all basic plant processes, including growth, development, reproduction and
Table 5.1 The classification of terpenoids is based on the number of C$_5$ isoprenoid units in their structures

<table>
<thead>
<tr>
<th>Isoprene units $n$</th>
<th>Carbon atoms $n$</th>
<th>Name</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>Hemiterpenes</td>
<td>Isoprene</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>Monoterpenes</td>
<td>Pulegone</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>Sesquiterpenes</td>
<td>Polygodial</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>Diterpenes</td>
<td>Paclitaxel</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>Sesterterpenes</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>Triterpenes</td>
<td>$\beta$-Amyrin</td>
</tr>
<tr>
<td>8</td>
<td>40</td>
<td>Tetraterpenes</td>
<td>$\beta$-Carotene</td>
</tr>
<tr>
<td>9—30 000</td>
<td>&gt; 40</td>
<td>Polyterpenes</td>
<td>Rubber</td>
</tr>
</tbody>
</table>

defence (Wink and van Wyk, 2008). Among the best-known lower (C$_5$–C$_{20}$) terpenes are the gibberellins, a large group of diterpene plant hormones involved in the control of seed germination, stem elongation and flower induction (Thomas et al., 2005). Another terpenoid hormone, the C$_{15}$ compound, abscisic acid (ABA), is not properly considered a lower terpenoid, since it is formed from the oxidative cleavage of a C$_{40}$ carotenoid precursor (Schwartz et al., 1997).

Several important groups of plant compounds, including cytokinins, chlorophylls and the quinone-based electron carriers (the plastoquinones and ubiquinones), have terpenoid side chains attached to a non-terpenoid nucleus. These side chains facilitate anchoring to or movement within membranes. In the past decade, proteins have also been found to have terpenoid side chains attached. In fact, all eukaryotic cells appear to contain proteins that have been post-translationally modified by the attachment of C$_{15}$ and C$_{20}$ terpenoid side chains via a thioether linkage.

Prenylation substantially increases protein hydrophobicity and serves to target proteins to membranes or direct protein–protein interactions (Zhang and Casey, 1996). In plants, prenylated proteins may be involved in the control of the cell cycle (Qian et al., 1996; Crowell, 2000), nutrient allocation (Zhou et al., 1997) and ABA signal transduction (Clark et al., 2001).

The most abundant hydrocarbon emitted by plants is the hemiterpene (C$_5$) isoprene, 2-methyl-1,3-butadiene (Fig. 5.1). Emitted from many taxa, especially woody species, isoprene has a major impact on the redox balance of the atmosphere, affecting levels of ozone, carbon monoxide and methane (Lerdau et al., 1997). The release of isoprene from plants is strongly influenced by light and temperature, with the greatest release rates typically occurring under conditions of high light and high temperature (Lichtenthaler, 2007). Although the direct function of isoprene in plants themselves has been a mystery for many years, there are now indications that it may serve to prevent cellular damage at high temperatures, perhaps by reacting with free radicals to stabilize membrane components (Sasaki et al., 2007). Instead of isoprene,
some plant species emit large amounts of monoterpene (C\(_{10}\)) hydrocarbons, which may function in a similar fashion (Loreto et al., 1998).

Most of the thousands of terpenoids produced by plants have no discernible role in growth and development and are, therefore, often classified as ‘secondary’ metabolites. Although comparatively few of these substances have been investigated in depth, they are thought to serve primarily in ecological roles, providing defence against herbivores or pathogens (Wittstock and Gershenson, 2002; Wink, 2007) and acting as attractants for animals that disperse pollen or seeds or as inhibitors of germination and growth of neighbouring plants (Harborne and Tomas-Barberan, 1991; Langenheim, 1994; Wink, 2010). One of the best-known examples of a lower terpene involved in plant defence is polygodial, a drimane-type sesquiterpene dialdehyde found in Polygonum hydropiper (Fig. 5.1). Among the most potent deterrents to insect feeding known, polygodial has been shown to inhibit the feeding of a diverse assortment of herbivorous insects (Moreno-Osorio et al., 2008). The deterrent effect appears to be a direct result of the action of polygodial on taste receptors. In lepidopteran larvae, polygodial and other drimane dialdehydes block the stimulatory effects of glucose and sucrose on chemosensory receptor cells found on the mouthparts (Frazier, 1986; Jansen and de Groot, 2004). The aldehyde groups can covalently bond with the free amino group of proteins under physiological conditions; such modification can change the three-dimensional structure of proteins and thus alter their bioactivity (Wink, 2008).

Although a few lower terpenes have been studied in as much detail as polygodial, many other members of this group serve as toxins, feeding deterrents or oviposition deterrents to herbivores, and so are also thought to function in plant defence. As toxins or deterrents, these substances possess many diverse modes of action on herbivores. For example, the monoterpene ketone, pulegone (Fig. 5.1), is a liver toxin in mammals (Chen et al., 2001); the pyrethrins, monoterpene esters, function as insect nerve poisons by interacting with the voltage-gated sodium channel proteins found in insect nerve cell membranes, leading to paralysis and eventual death (Davies et al., 2007); and the diterpene, atractyloside, inhibits ADP/ATP translocation in the mitochondria (Stewart and Steenkamp, 2000).

In the past few years, a new role for lower terpenes in plant defence has emerged. Certain plant species respond to herbivore attack by emitting volatile terpenes that attract the enemies of herbivores. For example, lima bean (Phaseolus lunatus) plants damaged by the spider mite, Tetranychus urticae, emit a mixture of monoterpenes, C\(_{11}\) and C\(_{16}\) homoterpenes (Fig. 5.1) and methyl salicylate, which attracts a carnivorous mite, Phytoseiulus persimilis, that preys on spider mites (Dicke et al., 1990; Dicke, 1994). When maize or cotton is fed upon by lepidopteran larvae, a blend of monoterpenes, sesquiterpenes, homoterpenes and other compounds is released, which attracts parasitic wasps that oviposit on the larvae (Turlings et al., 1990, 1995). The majority of these volatiles are emitted only by arthropod-damaged plants
and not by unattacked or artificially damaged plants. The terpenoids released are largely synthesized de novo following an initial herbivore attack (Pare and Tumlinson, 1997) and are released systemically throughout the plant (Dicke et al., 1993; Rose et al., 1996). The use of volatile terpenoids to attract the enemies of herbivores may be a valuable complement to the more direct modes of antitherbivore defence.

The functions of the lower terpenes are not limited to the natural world. Many play important roles in human society, such as the myriad of monoterpane and sesquiterpene flavour and fragrance agents that are added to foods, beverages, perfumes, soaps, toothpaste, tobacco and other products (Berger, 2007). Some lower terpenes find use in industry as raw materials in the manufacture of adhesives, coatings, emulsifiers and speciality chemicals, whilst others, such as limonene and the pyrethrins, are of increasing commercial importance as insecticides because of their low toxicity to mammals and lack of persistence in the environment. The pharmaceutical importance of plant lower terpenes has steadily increased in the past decade. In addition to the well-known roles of camphor (Fig. 5.1) and cineole in preparations to relieve the pain of burns, strains and other inflammations, the past few years have seen the acceptance of artemisinin, a sesquiterpene endoperoxide derived from the traditional Chinese medicinal plant, *Artemisia annua* (Fig. 5.1), as a valuable antimalarial compound (Balint, 2001), and the development of paclitaxel (Fig. 5.1), a highly functionalized diterpene from yew (*Taxus* spp.), as a new drug for the treatment of ovarian and breast cancer (Kingston and Newman, 2007). Recently, and after the thorough research carried by Thomas Efferth and co-workers, it was proven that artemisinin and its semisynthetic artemether derivatives have not only antimalarial activity, but also antiviral and cytotoxic activities against different cancer cells (Efferth et al., 2007, 2008; Youns et al., 2009). These sesquiterpene lactone drugs with the highly active endoperoxide bridge can bind covalently to DNA and protein leading to permanent inactivation of many molecular targets (Wink, 2008). Paclitaxel, also known as taxol, enhances the polymerization of tubulin, a protein component of the microtubules of the mitotic spindle, resulting in stabilized, non-functional tubules and blocking the cell cycle. The potential of other lower terpenes in the therapy and prevention of cancer is currently under active investigation (Gould, 1995). Several sesquiterpene lactones have exocyclic methylene groups that are highly reactive. They can easily make covalent bonds with SH-groups of proteins or glutathione and thus alter their bioactivity (Wink, 2008). These interactions can explain the activity of several sesquiterpene lactones against inflammation and as antiinfectants.

### 5.3 Biosynthesis

The biosynthetic pathway to terpenoids (Fig. 5.2) is conveniently treated as comprising four stages, the first of which involves the formation of IPP,
Biochemistry of Plant Secondary Metabolism

Figure 5.2 Overview of terpenoid biosynthesis in plants, showing the basic stages of this process and major groups of end products. CoA, coenzyme A; GAP, glyceraldehyde-3-phosphate.

The biological C5 isoprene unit. Plants synthesize IPP and its allylic isomer, DMAPP, by one of two routes: the well-known mevalonic acid pathway, or the newly discovered methylerythritol phosphate (MEP) pathway. In the second stage, the basic C5 units condense to generate three larger prenyl diphosphates, geranyl diphosphate (GPP, C10), farnesyl diphosphate (FPP, C15) and geranylgeranyl diphosphate (GGPP, C20). In the third stage, the C10–C20 diphosphates undergo a wide range of cyclizations and rearrangements to produce the parent carbon skeletons of each terpene class. GPP is converted to the monoterpenes, FPP is converted to the sesquiterpenes and GGPP is converted to the diterpenes. FPP and GGPP can also dimerize in a head-to-head fashion to form the precursors of the C30 and the C40 terpenoids,
respectively. The fourth and final stage encompasses a variety of oxidations, reductions isomerizations, conjugations and other transformations by which the parent skeletons of each terpene class are converted to thousands of distinct terpene metabolites. This section discusses the latest findings concerning each of the four stages of terpenoid biosynthesis in plants. The portions of the third and fourth stages that are not involved in the formation of the lower (C\textsubscript{5}–C\textsubscript{20}) terpenes are dealt with in Section 5.2.

### 5.3.1 Formation of the basic C\textsubscript{5} unit: the mevalonate pathway

The classic route for the formation of the C\textsubscript{5} building blocks of terpenoid biosynthesis in plants is via the reactions of the mevalonate pathway, first demonstrated in yeast and mammals. This well-characterized sequence (Fig. 5.3) involves the stepwise condensation of three molecules of acetyl coenzyme A (AcCoA) to form the branched C\textsubscript{6} compound, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). Following the reduction of HMG-CoA to mevalonic acid, two successive phosphorylations and a decarboxylationelimination yield the C\textsubscript{5} compound, IPP.

Among the most recent developments in mevalonate pathway research is the successful cloning of the plant genes encoding all the enzymes (acetyl-CoA acetyltransferase genes, HMG-CoA synthase genes and HMG-CoA reductase) that catalyse the initial steps of the mevalonic acid (MVA) pathway in rubber (Sando et al., 2008). Along with the work done earlier on acetoacetyl-CoA thiolase (Vollack and Bach, 1996) and HMG-CoA synthase (Montamant et al., 1995), it was obvious that the two sequences of acetyl-CoA acetyltransferase, HMG-CoA synthase are separated and distinct from each other, in contrast to an earlier report suggesting that in plants, in contrast to animals and microorganisms, both reactions are catalysed by a single protein (Weber and Bach, 1994). Each sequence is highly homologous to that of corresponding genes in the mevalonate pathways of mammals and microbes.

The third step of the mevalonic acid pathway is the conversion of HMG-CoA to mevalonic acid, a two-step, nicotinamide adenine diphosphate (reduced form) (NADPH)-requiring reduction catalysed by HMG-CoA reductase (HMGR) (Fig. 5.3). Researchers have lavished considerable attention on HMGR, since it catalyses a critical, rate-determining step in the biosynthesis of sterols in animals, and has been assumed to play a role of similar importance in the formation of plant terpenoids. Plant HMGR is a membrane-bound enzyme, a feature that has greatly hindered efforts to purify and characterize it. However, now that HMGR genes from more than ten species have been cloned and analysed (Table 5.2), our knowledge of this important catalyst has increased substantially. All plant genes isolated so far encode polypeptides of 60–65 kDa each, with three distinct regions: a very divergent NH\textsubscript{2}-terminal domain, a more conserved membrane-binding region with two membrane-spanning sequences and a highly conserved COOH-terminal domain containing the catalytic site.
Experiments with cloned genes have contributed to the resolution of a long-standing controversy concerning the subcellular location of HMGR in plants. Over the past 25 years, it has been claimed that HMGR is present in the endoplasmic reticulum (ER), the plastids and the mitochondria (Bach et al., 1991). However, HMGR gene products from both Arabidopsis thaliana (Enjuto et al., 1994; Campos and Boronat, 1995) and tomato (Denbow et al., 1996) have recently been demonstrated to be co-translationally inserted into ER-derived microsomal membranes in vitro. Since the insertion is mediated by the two transmembrane regions (Enjuto et al., 1994; Denbow et al., 1996; Re et al., 1997) whose sequences are conserved among all plant HMGR genes.
so far isolated, it seems probable that all known plant HMGRs are targeted to the ER (Campos and Boronat, 1995). Nevertheless, claims regarding the plastidial localization of HMGR have continued to appear (Nakagawara et al., 1993; Bestwick et al., 1995; Kim et al., 1996). While an as yet uncharacterized HMGR may be present in the plastids, reports of plastidial localization are more likely to be due to contamination of plastid fractions with microsomes (Gray, 1987). Marker enzymes or electron microscopy have seldom been used to verify the purity of subcellular fractions in such studies.

Evidence for the regulatory role of HMGR in the formation of plant terpenoids comes from numerous studies that have demonstrated a close correlation between changes in HMGR activity and alterations in the rate of terpenoid biosynthesis. For example, Heide and co-workers (Gaisser and Heide, 1996; Lange et al., 1998a) have been studying the formation of shikonin,
## Biochemistry of Plant Secondary Metabolism

### Table 5.2  Isolated genes encoding several major classes of enzymes in terpene biosynthesis

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A) HMG-CoA reductase</td>
<td>Arabidopsis thaliana</td>
<td>Cælles \textit{et al.} (1989)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D’Auria and Gershenzon (2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enjuto \textit{et al.} (1994)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learned and Fink (1989)</td>
</tr>
<tr>
<td></td>
<td>Camptotheca acuminata</td>
<td>Burnett \textit{et al.} (1993)</td>
</tr>
<tr>
<td></td>
<td>Catharanthus roseus</td>
<td>Maldonado-Mendoza \textit{et al.} (1992)</td>
</tr>
<tr>
<td></td>
<td>Corylus avellana</td>
<td>Wång \textit{et al.} (2007)</td>
</tr>
<tr>
<td></td>
<td>Euphorbia Pekinensis</td>
<td>Cao \textit{et al.} (2009)</td>
</tr>
<tr>
<td></td>
<td>Ginkgo biloba</td>
<td>Shen \textit{et al.} (2006)</td>
</tr>
<tr>
<td></td>
<td>Hevea brasiliensis</td>
<td>Sando \textit{et al.} (2008)</td>
</tr>
<tr>
<td></td>
<td>Lycopersicon esculentum</td>
<td>Narita and Gruissem (1989)</td>
</tr>
<tr>
<td></td>
<td>Nicotiana sylvestris</td>
<td>Park \textit{et al.} (1992)</td>
</tr>
<tr>
<td></td>
<td>Oryza sativa</td>
<td>Genschik \textit{et al.} (1992)</td>
</tr>
<tr>
<td></td>
<td>Raphanus sativus</td>
<td>Monfar \textit{et al.} (1990)</td>
</tr>
<tr>
<td></td>
<td>Triticum aestivum</td>
<td>Vollack \textit{et al.} (1994)</td>
</tr>
<tr>
<td>FPP synthase</td>
<td>Artemisia annua</td>
<td>Matsushita \textit{et al.} (1996)</td>
</tr>
<tr>
<td></td>
<td>Capsicum annuum</td>
<td>Huguenev \textit{et al.} (1996)</td>
</tr>
<tr>
<td></td>
<td>Centella asiatica</td>
<td>Kim \textit{et al.} (2005)</td>
</tr>
<tr>
<td></td>
<td>Hevea brasiliensis</td>
<td>Adiwiila and Kush (1996)</td>
</tr>
<tr>
<td></td>
<td>Oryza sativa</td>
<td>Sanmiya \textit{et al.} (1997)</td>
</tr>
<tr>
<td></td>
<td>Parthenium argentatum</td>
<td>Pan \textit{et al.} (1996)</td>
</tr>
<tr>
<td></td>
<td>Picea abies</td>
<td>Schmidt and Gershenzon (2007)</td>
</tr>
<tr>
<td></td>
<td>Zea mays</td>
<td>Li and Larkins (1996)</td>
</tr>
<tr>
<td>GGPP synthase</td>
<td>Arabidopsis thaliana</td>
<td>Scolnick and Bartley (1996)</td>
</tr>
<tr>
<td></td>
<td>Brassica campestris</td>
<td>Scolnick and Bartley (1994)</td>
</tr>
<tr>
<td></td>
<td>Capsicum annuum</td>
<td>Lim \textit{et al.} (1996)</td>
</tr>
<tr>
<td></td>
<td>Catharanthus roseus</td>
<td>Badillo \textit{et al.} (1995)</td>
</tr>
<tr>
<td></td>
<td>Ginkgo biloba</td>
<td>Kuntz \textit{et al.} (1992)</td>
</tr>
<tr>
<td></td>
<td>Solanum tuberosum</td>
<td>Schmidt and Gershenzon (2007)</td>
</tr>
</tbody>
</table>
Table 5.2 (Continued)

<table>
<thead>
<tr>
<th>Enzyme Description</th>
<th>Species</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>C) Terpene synthases</td>
<td>Abies grandis</td>
<td>Bohlmann et al. (1997)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bohlmann et al. (1998a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Steele et al. (1998a)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vogel et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Arabidopsis thaliana</td>
<td>Corey et al. (1993)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sun and Kamiya (1994)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yamaguchi et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>Artemisia annua</td>
<td>Bertea et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Cichorium intybus</td>
<td>Lu et al. (2002)</td>
</tr>
<tr>
<td></td>
<td>Citrus junos</td>
<td>Mercke et al. (2000)</td>
</tr>
<tr>
<td></td>
<td>Clarkia brewerii</td>
<td>Bouwmeester et al. (2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maruyama et al. (2001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dudareva et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Cucurbita maxima</td>
<td>Yamaguchi et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Gossypium arboreum</td>
<td>Chen et al. (1995)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chen et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Hyoscyamus muticus</td>
<td>Back and Chappell (1995)</td>
</tr>
<tr>
<td></td>
<td>Luffa cylindrical</td>
<td>Hayashi et al. (2001)</td>
</tr>
<tr>
<td></td>
<td>Lycopersicon esculentum</td>
<td>Colby et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>Mentha x piperita</td>
<td>Crock et al. (1997)</td>
</tr>
<tr>
<td></td>
<td>Mentha spicata</td>
<td>Colby et al. (1993)</td>
</tr>
<tr>
<td></td>
<td>Nicotiana tabacum</td>
<td>Facchini and Chappell (1992)</td>
</tr>
<tr>
<td></td>
<td>Penilla frutescens</td>
<td>Yuba et al. (1996)</td>
</tr>
<tr>
<td></td>
<td>Picea sitchensis</td>
<td>Mckay et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Pinus taeda</td>
<td>Phillips et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Ricinus communis</td>
<td>Mau and West (1994)</td>
</tr>
<tr>
<td></td>
<td>Salvia officinalis</td>
<td>Wise et al. (1998)</td>
</tr>
<tr>
<td></td>
<td>Taxus brevifolia</td>
<td>Wildung and Croteau (1996)</td>
</tr>
<tr>
<td></td>
<td>Vitis vinifera</td>
<td>Lucker et al. (2004)</td>
</tr>
<tr>
<td></td>
<td>Zea mays</td>
<td>Bensen et al. (1995)</td>
</tr>
<tr>
<td></td>
<td>Zingiber zerumbet</td>
<td>Yu et al. (2008)</td>
</tr>
</tbody>
</table>

HMG-CoA, 3-hydroxy-3 methylglutaryl coenzyme A; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate.

a naphtoquinone pigment constructed from a benzenoid ring and a molecule of GPP. In Lithospermum erythrorhizon cultures, they showed that increases in the level of HMGR enzyme activity under various light and inhibitor treatments were associated with greater accumulation of shikonin and its derivatives. Other recent examples include correlations between the level of HMGR and the formation of: sesquiterpenes in lettuce (Bestwick et al., 1995), sesquiterpenes in cotton (Joost et al., 1995), triterpenes in Tabernaemontana divaricata (Fulton et al., 1994) and rubber in guayule (Ji et al., 1993).

To obtain a more rigorous proof of the regulatory role of plant HMGR, researchers have used constitutive promoters to overexpress HMGR in various
species. For example, tobacco transformed with a constitutively expressed HMGR construct showed a three–eightfold increase in HMGR enzyme activity and a three–tenfold increase in total sterols (Chappell et al., 1995; Schaller et al., 1995). However, there was no change in the level of other terpenoid end products, including sesquiterpenes, phytol (the C20 side chain of chlorophyll) and carotenoids. Curiously, the sterol composition of these HMGR-overexpressing plants differed from that of untransformed tobacco in having a much higher proportion of biosynthetic intermediates, such as cycloartenol (often conjugated as esters), rather than end products, such as sitosterol or stigmasterol. A mutant tobacco cell line resistant to a sterol inhibitor showed a very similar phenotype (Gondet et al., 1992, 1994). Taken together, these results make a strong case for HMGR being a rate-determining step, at least in the formation of sterols, although later enzymes in the pathway also have a significant influence on the rate of sterol biosynthesis. However, this conclusion may not be applicable to all plant species, since the overexpression of HMGR in A. thaliana had no effect on the accumulation of sterols and other terpenoids (Re et al., 1997).

If HMGR activity limits the rate of terpenoid formation, it is important to understand the mechanism of this control. In mammals, HMGR activity is subject to feedback inhibition by sterols that regulates the rates of transcription and translation, and post-translational controls involving allosteric effects and reversible phosphorylation (Panda and Devi, 2004). HMGR activity in plants appears to be modulated in similar ways, although we are only just beginning to understand the mechanisms of control. The close correlation of HMGR activity with the abundance of HMGR mRNA in L. erythrorhizon (Lange et al., 1998a), tomato (Yang et al., 1991) and other species (Stermer et al., 1994) is good evidence for transcriptional control. At the post-translational level, HMGR from Brassica oleracea was shown to be inactivated by reversible phosphorylation, mediated by a specific kinase (MacKintosh et al., 1992; Dale et al., 1995). Since plants produce a much wider assortment of terpenoid end products than mammals do, they might be expected to regulate HMGR in unique ways not found in mammals. While only a single HMGR gene is known from each of the mammal species studied so far, all plants examined possess a small gene family with as many as nine members (Bhattacharyya et al., 1995; Joost et al., 1995). Detailed studies in tomato and potato reveal that different HMGR genes may be expressed in different organs or under different environmental conditions (Choi et al., 1994; Enjuto et al., 1995; Daraselia et al., 1996), raising the possibility that a differential expression of HMGR genes could serve as a major mechanism for the control of HMGR activity.

Mevalonic acid, the product of HMGR, is converted to IPP by the sequential action of three enzymes: mevalonate kinase, phosphomevalonate kinase and diphosphomevalonate decarboxylase (Fig. 5.3). These three catalysts have not previously been considered to be important control points in plant terpenoid biosynthesis, and little new information has appeared to alter this view. The
activities of all the three enzymes were shown to be higher than that of HMGR (Bianchini et al., 1996), similar to each other (Sandmann and Albrecht, 1994) and unrelated to fluctuations in the rate of terpenoid formation (Ji et al., 1993; Bianchini et al., 1996). A cDNA encoding mevalonate kinase was recently isolated from A. thaliana by genetic complementation in yeast (Riou et al., 1994). The lack of a transit peptide and the presence of only a single gene, as deduced from Southern blotting, make it appear that plant mevalonate kinase, like HMGR, is a cytosolic enzyme.

5.3.2 Formation of the basic C₅ unit: the methylerythritol phosphate pathway

The most exciting advance in the field of plant terpenoid biosynthesis is the discovery of a second route for making the basic C₅ building block of terpenes, completely distinct from the mevalonate pathway (Lichtenthaler, 2000). This new route, which starts from glyceraldehyde phosphate and pyruvate (Fig. 5.4), has also been detected in bacteria and other microorganisms. With the advantage of hindsight, one can list many observations made during the past 30 years that, taken together, should have persuaded researchers of the existence of a non-mevalonate pathway to terpenoids in higher plants. For example, it was demonstrated numerous times that mevalonate itself is a very poor precursor for many classes of terpenoids (Croteau and Loomis, 1972; Charlwood and Banthorpe, 1978). However, there was no reasonable alternative to the mevalonate pathway prior to the pioneering investigations of terpenoid biosynthesis in eubacteria, carried out by Michel Rohmer, Hermann Sahm and co-workers. These investigators discovered that the incorporation of ¹³C-labelled precursors, such as glucose, acetate and pyruvate, into bacterial terpenoids (hopanoids and ubiquinones) was not consistent with the operation of the mevalonate pathway (Flesch and Rohmer, 1988; Rohmer, 2008). In addition, when intermediates of the mevalonate pathway were fed to species such as Escherichia coli, they were not incorporated (Horbach et al., 1993). Analysis of the ¹³C incorporation patterns from labelled glucose and acetate allowed the deduction that a C₃-unit from glycolysis and a C₂-unit from pyruvate combined in some manner to form the basic C₅ isopentenoid unit (Rohmer, 1999). Subsequent experiments with E. coli mutants, blocked in specific steps of triose phosphate metabolism, pointed to glyceraldehyde phosphate and pyruvate as the actual precursors of this new pathway (Rohmer et al., 1996).

The existence of a similar non-mevalonate route to terpenoids in plants was first reported in 1994. When Duilio Arigoni and co-workers fed different ¹³C-labelled forms of glucose to Ginkgo biloba embryos, the ¹³C-nuclear magnetic resonance (NMR) spectra of the resulting diterpenes were not what would have been expected from the normal operation of the mevalonate pathway (Cartayrade et al., 1994), but showed an incorporation pattern identical to that seen with the E. coli terpenoids. Subsequent studies employing
Biochemistry of Plant Secondary Metabolism

Figure 5.4 Outline of the newly discovered glyceraldehyde phosphate/pyruvate pathway for the formation of C₅ isoprenoid units. None of the intermediates after 2-C-methyl-d-erythritol 4-phosphate is known. P indicates a phosphate moiety. TPP, thiamine pyrophosphate; NADP, nicotinamide adenine dinucleotide phosphate.

A similar methodology have demonstrated that an assortment of terpenoids from angiosperms, gymnosperms and bryophytes, including monoterpenes (Eisenreich et al., 1997; Adam et al., 1998), diterpenes (Knoss et al., 1997; Jennewein and Croteau, 2001), carotenoids (Lichtenthaler et al., 1997) and the side chains of chlorophyll (phytol) and quinones (Lichtenthaler et al., 1997; Adam et al., 1998) are formed in a non-mevalonate fashion, while the labelling of sesquiterpenes and sterols was consistent with their origin from the mevalonate pathway (Schwarz, 1994; Lichtenthaler et al., 1997; Adam et al., 1998).
Several research groups are now actively involved in elucidating the sequence of the new pathway. In 1996, Rohmer and co-workers refined their concept of the first step, proposing that hydroxyethylthiamine diphosphate, a C$_2$–unit derived from pyruvate, condenses with glyceraldehyde-3-phosphate to form 1-deoxy-D-xylulose 5-phosphate (Rohmer et al., 1996) (Fig. 5.4). This hypothesis was based on the pattern of labelling in terpenoids formed from $^{13}$C-pyruvate, $^{13}$C-glycerol and various $^{13}$C-glucoses, and the natural occurrence of 1-deoxy-D-xylulose, a precursor of the enzyme cofactors thiamine (vitamin B$_1$) diphosphate and pyridoxal (vitamin B$_6$) 5′-phosphate. Additional support comes from the high rate of 1-deoxy-D-xylulose incorporation into terpenoids measured in *E. coli* (Broers, 1994) and several plant species (Zeidler et al., 1997; Sagner et al., 1998b; Eisenreich et al., 2001). A more rigorous proof of the nature of the first step of the non-mevalonate pathway has become available in the past years, with the isolation of cDNAs for enzymes that catalyse the conversion of glyceraldehyde phosphate and pyruvate to 1-deoxy-D-xylulose 5-phosphate from *E. coli* (Sprenger et al., 1997; Lois et al., 1998), *Capsicum annuum* (Bouvier et al., 1998) and *Mentha x piperita* (Lange et al., 1998b). The encoded enzymes are novel transketolases that are distinct from other members of this enzyme family, such as the well-characterized transketolases of the pentose phosphate pathway.

After 1-deoxy-D-xylulose 5-phosphate, subsequent reactions of the new pathway must transform the linear five-carbon backbone of this sugar phosphate to a branched, isopentenoid carbon skeleton. Just recently, 1-deoxy-D-xylulose 5-phosphate has been shown to be converted to 2-C-methyl-D-erythritol 4-phosphate in *E. coli* (Duvold et al., 1997; Wanke et al., 2001; Kuzuyama, 2002) (Fig. 5.4), and the same reaction was demonstrated to occur in several species of plants (Sagner et al., 1998a). This intramolecular rearrangement involves the cleavage of the C3–C4 bond of the deoxyxylulose backbone and the establishment of a new bond between C2 and C4. Similar skeletal rearrangements are involved in both riboflavin and valine biosynthesis. While nothing is yet known of any additional intermediates in the pathway, several dehydration steps, reductions and at least one phosphorylation seem to be required to transform 2-C-methyl-D-erythritol 4-phosphate to IPP. Given the high level of interest in this work and the participation of several excellent research groups, it would be surprising if the remaining steps of this novel pathway were not rapidly elucidated.

The non-mevalonate route to terpenoids appears to be localized in the plastids. In plant cells, terpenoids are manufactured both in the plastids and the cytosol (Gray, 1987; Kleinig, 1989). As a general rule, the plastids produce monoterpines, diterpenes, phytol, carotenoids and the side chains of plastoquinone and α-tocopherol, while the cytosol/ER compartment produces sesquiterpenes, sterols and dolichols. In the studies discussed above, nearly all of the terpenoids labelled by deoxyxylulose (Sagner et al., 1998b; Eisenreich et al., 2001) and 2-C-methyl erythritol feeding (Duvold et al., 1997) or
showing $^{13}$C-patterns indicative of a non-mevalonate origin (Cartayrade et al., 1994; Eisenreich et al., 1997, 2001; Lichtenthaler et al., 1997; Adam et al., 1998) are thought to be plastid derived. Consistent with this generalization is the fact that the genes of the non-mevalonate pathway that have been isolated so far all encode plastid-targeting sequences (Bouvier et al., 1998; Lange et al., 1998b). In contrast, the mevalonate pathway appears to reside solely in the cytosol/ER compartment based on the sequence analysis and expression of genes encoding pathway enzymes, including acetoacetyl-CoA thiolase (Vollack and Bach, 1996), HMG-CoA synthase (Montamant et al., 1995), HMGR (discussed in Section 5.3.1) and mevalonate kinase (Riou et al., 1994). A third subcellular compartment, the mitochondrion, also participates in terpenoid biosynthesis, making the prenyl side chain of ubiquinone, an electron transport system component found in this organelle, using IPP derived from the cytosol/ER pathway (Hemmerlin et al., 2004).

It was once difficult to reconcile the terpenoid-manufacturing capabilities of the plastids with the usual absence of detectable HMGR activity in these organelles. Models proposed that the basic reactions of terpenoid biosynthesis are confined to the cytosol, with the preformed C$_5$-units being transferred to other subcellular compartments (Gray, 1987; Luetke-Brinkhaus and Kleinig, 1987). However, current knowledge suggests a more accurate generalization: the plastids biosynthesize terpenoids primarily via the methylerythritol phosphate pathway, while in the cytosol/ER terpenoid formation occurs largely via the mevalonate pathway. Reviewing the older literature with this paradigm in mind, it is not surprising that mevalonate was found to be so poorly incorporated into many plastid-formed terpenoids (Croteau and Loomis, 1972; Charlwood and Banthorpe, 1978; Keene and Wagner, 1985; Lunn, 2007), that levels of HMGR activity were often noted to be poorly correlated with the formation of plastidial terpenoids (Chappell et al., 1989; Narita and Gruissem, 1989) and that the HMGR inhibitor, mevinolin, was shown to have a negligible effect on the production of plastidial terpenoids (Bach and Lichtenthaler, 1983; Bach et al., 1999).

The existence of a non-mevalonate route to terpenoids also helps account for other puzzling observations, such as the complete failure of green algae to incorporate mevalonate into terpenoids (Lichtenthaler, 2000). Feeding experiments with $^{13}$C-labelled glucose and acetate have now shown that all terpenoids in *Scenedesmus obliquus* (Rohmer, 1999) and other green algae (Disch et al., 1998) are formed by the glyceraldehyde–pyruvate pathway. Among other photosynthetic microorganisms surveyed, the red alga, *Cyanidium caldarium*, and the chrysophyte, *Ochromonas danica* (Disch et al., 1998), use both pathways, *Euglena gracilis* (Disch et al., 1998) and the eubacterium, *Chloroflexus aurantiacus* (Rieder et al., 1998), use only the mevalonate pathway, while the cyanobacterium, *Synechocystis PCC 6714* (Disch et al., 1998), employs only the glyceraldehyde–pyruvate pathway, like the plastids of higher plants. These results are in accord with the endosymbiotic origin of higher plant plastids from a cyanobacterium-like symbiont.
A strict division between the mevalonate and non-mevalonate pathways may not always exist for a given end product. The biosynthesis of certain terpenoids appears to involve the participation of both routes (Schwarz, 1994; Nabeta et al., 1995; Adam and Zapp, 1998; Piel et al., 1998). For example, the first two C₅ units of the sesquiterpenes of chamomile (Matricaria recutita) are formed via the methylerythritol phosphate pathway, while the third unit is derived from both the mevalonate pathway and the glyceraldehyde–pyruvate pathway (Adam and Zapp, 1998). Joint participation of the two pathways may be a result of the transport of prenyl diphosphate intermediates between the different sites of terpenoid biosynthesis (Heintze et al., 1990; Soler et al., 1993; McCaskill and Croteau, 1997), or the actual presence of both pathways in the same compartment. While the preponderance of evidence argues for the localization of the mevalonate pathway in the cytosol and the glyceraldehyde–pyruvate pathway in the plastids, as discussed above, there are some indications that the mevalonate pathway may also be found in the plastids, at least in certain species (Kim et al., 1996) at certain developmental stages (Heintze et al., 1990, 1994).

The occurrence of both terpenoid pathways at the same subcellular site, or the exchange of prenyl diphosphates between sites, may also help explain other curious phenomena noted in previous biosynthetic studies, such as the unequal labelling of different C₅ units. Administration of mevalonate has frequently been shown to result in the IPP-derived portion of the molecule being much more heavily labelled than the portion derived from DMAPP (Croteau and Loomis, 1972; Charlwood and Banthorpe, 1978). Such asymmetry has been attributed to the existence of a large pool of DMAPP that dilutes any DMAPP formed from an exogenous, labelled precursor. However, asymmetric labelling could also be a consequence of having separate pathways to each of the two basic C₅ units. The actual C₅ product of the alternative pathway is not known, and might be DMAPP rather than IPP. If DMAPP arising from the non-mevalonate pathway condensed with mevalonate-derived IPP (produced in situ or transported from another compartment), this could result in the unequal labelling of C₅ units. More research is needed not only to identify the remaining intermediates in the methylerythritol phosphate pathway, but also to determine in which species, tissues and compartments it operates, as well as to understand its regulation.

5.3.3 Assembly of C₅ units into C₁₀, C₁₅ and C₂₀ prenyl diphosphates

The second stage of terpene biosynthesis involves the fusion of the basic C₅ building blocks to yield larger metabolic intermediates (Fig. 5.2). IPP and its more reactive allylic isomer, DMAPP, condense in a head-to-tail orientation to form C₁₀, C₁₅ and C₂₀ prenyl diphosphates (Fig. 5.5). The requisite DMAPP is derived directly from IPP by the action of IPP isomerase, which is also capable of catalysing the reverse reaction. In the past years, genes encoding this
Figure 5.5 The formation of $C_{10}$, $C_{15}$ and $C_{20}$ prenyl diphosphates from the fusion of $C_5$ isoprenoid units. PP indicates a diphosphate moiety.

e enzyme have been isolated from *A. thaliana* (Okada *et al.*, 2008) and *Clarkia breweri* (Blanc and Pichersky, 1995). The sequences reported exhibit high homology to the IPP isomerase gene sequences of other organisms, except at their N-termini, which seem to encode transit peptides for plastid localization. *Arabidopsis thaliana* possesses an IPP isomerase gene family consisting of at least two members (Phillips *et al.*, 2008), a finding consistent with the detection of multiple forms of this enzyme in cell cultures of several plant species (Ramosvaldivia *et al.*, 1998). In *Cinchona robusta*, e.g., the two isoforms of IPP isomerase had different kinetic parameters, different preferences for divalent metal ion cofactors and different patterns of occurrence; one form was present only after induction by a fungal elicitor (Ramosvaldivia *et al.*, 1997c). Although there is no strong evidence that IPP isomerase has any control of
flux through the terpenoid pathway (Ramosvaldivia et al., 1997b), the activity of this enzyme in maize increases significantly after stimulation of carotenoid biosynthesis by light (Sandmann, 2001), and activity in cell cultures of several species increases after induction of phytoalexin formation by treatment with fungal elicitors (Hanley et al., 1992; Fulton et al., 1994; Ramosvaldivia et al., 1997a).

The substrate (IPP) and the product (DMAPP) of IPP isomerase are both involved in the fundamental reactions by which C5 isopentenoid units are joined together. Enzymes known as prenyltransferases add varying numbers of IPP units to a DMAPP primer in sequential chain elongation steps. The initial head-to-tail (1′-4) condensation of IPP and DMAPP yields the C10 allylic diphosphate, GPP. Further 1′-4 condensations of IPP with the enlarging allylic diphosphate chain give the C15 allylic diphosphate, farnesyl diphosphate (FPP) and the C20 allylic diphosphate, GGPP. In plants, FPP and GGPP are produced by well-characterized, product-specific enzymes that catalyse two- or three-step elongation sequences starting with IPP and DMAPP (Fig. 5.5). For example, GGPP synthases convert DMAPP and IPP directly to GGPP (Spurgeon et al., 1984; Dogbo and Camara, 1987; Laskaris et al., 2000). The reaction proceeds through the intermediacy of GPP and FPP, but under normal conditions GGPP is the first product to leave the active site. In contrast to FPP and GGPP synthases, much less attention has been devoted to GPP synthases. In fact, the very existence of this class of prenyltransferases in plants was once doubted, in the belief that amounts of GPP sufficient to sustain monoterpene biosynthesis were released during the formation of the larger allylic diphosphates (Tello et al., 2008). However, prenyltransferases that synthesize GPP exclusively have now been discovered in several plant species that produce monoterpenes or natural products incorporating a monoterpene unit (Croteau and Purkett, 1989; Clastre et al., 1993; Tello et al., 2008).

In the past six years, cDNAs encoding FPP and GGPP synthases have been isolated from a diverse assortment of plant species (Table 5.2). The amino acid sequences deduced have a high degree of similarity to the FPP and GGPP synthases of other organisms (Chen et al., 1994), which means that the recent determination of the crystal structure of an avian FPP synthase has a considerable value for the study of plant prenyltransferases as well. The structure of FPP synthase from avian liver consists of a novel arrangement of ten parallel α-helices positioned around a large central cavity (Tarshis et al., 1994). Two aspartate-rich sequences (DDxxD) that are highly conserved among other prenyltransferases (Chen et al., 1994) and essential for catalysis (Joly and Edwards, 1993; Song and Poulter, 1994; Harris and Poulter, 2000) are found on opposite sides of the cavity, with their aspartate carboxyl side chains pointing towards the cavity centre. These aspartate residues had previously been suggested to bind the diphosphate moieties of the substrates via Mg2+ bridges (Harris and Poulter, 2000). Structural analysis of a samarium-containing heavy atom derivative of an avian FPP synthase (samarium commonly adheres to Mg2+-binding sites in enzymes) showed
samarium atoms bound to each of the two aspartate-rich regions, supporting the role of the aspartate residues in binding Mg$^{2+}$ (Tarshis et al., 1994). Work has now begun to identify other amino acid residues involved in the reaction mechanism. Prenyltransferases are one of the few groups of enzymes in which carbon–carbon bond formation results from electrophilic attack of a carbocationic species on a pre-existing double bond (Poulter and Rilling, 1981). The initial carbocation is formed by the ionization of the allylic substrate through hydrolysis of the diphosphate ester. Subsequently, addition to the double bond of IPP forms a new carbocation, which is then stabilized by proton elimination.

A long-standing goal in the study of prenyltransferases is to understand how these catalysts control the length of the growing chain during the reaction sequence. The availability of cloned prenyltransferase sequences and a three-dimensional structure for this enzyme class has provided new tools to approach this problem. Random and site-directed mutagenesis of bacterial FPP and GGPP synthases has demonstrated that several amino acid residues near the conserved aspartate-rich domains were most critical in determining chain length (Tarshis et al., 1996; Wang and Ohnuma, 2000). For example, when an avian FPP synthase was altered so that two phenylalanine residues, located just on the N-terminal side of the first aspartate-rich domain, were changed to serine and alanine, the mutant enzyme produced products up to C$_{35}$–C$_{40}$ (Tarshis et al., 1996). Structural analysis carried out in parallel with the mutagenesis revealed that the mutant FPP synthase had a larger binding pocket for allylic diphosphate substrates than native FPP synthase. Other amino acid residues involved in the substrate and product specificity of prenyltransferases are being actively sought.

The prenyltransferases that catalyse the syntheses of GPP, FPP and GGPP may be important regulatory enzymes in plant terpenoid biosynthesis since they are situated at the primary branch points of the pathway, directing flux among the various major classes of terpenoids. The level of prenyltransferase activity is, in fact, closely correlated with the rate of terpenoid formation in many experimental systems (Dudley et al., 1986; Hanley et al., 1992; Hugueney et al., 1996) consistent with the regulatory importance of these catalysts. The localization of specific prenyltransferases in particular types of tissue or subcellular compartments may control the flux and direction of terpenoid synthesis at these sites. For example, the GPP synthase in *Salvia officinalis* is restricted to the secretory cells of the glandular trichomes, which are the sole site of monoterpane biosynthesis in this species (Croteau and Purkett, 1989).

### 5.3.4 Formation of parent carbon skeletons

The prenyl diphosphates, GPP, FPP and GGPP, are the central intermediates of terpenoid biosynthesis. Under the catalysis of monoterpane, sesquiterpene and diterpene synthases, respectively, these substances are transformed into
the primary representatives of each terpene skeletal type. Recent progress in the area of terpene synthases has been remarkable. In the past years, many novel activities have been described (Guo et al., 1994; Dekraker et al., 1998; Lu et al., 2002; Dudareva et al., 2004), over 30 terpene synthase cDNAs have been isolated from plants (Table 5.2) and the first crystal structures of terpene synthases have been obtained (Starks et al., 1997; Christianson, 2006; Abe, 2007). These achievements have permitted new insights into the evolutionary origin and genetic regulation of terpene synthases and have provided unprecedented opportunities for exploring the reaction mechanisms of these catalysts.

A sequence comparison of the isolated terpene synthase cDNAs suggests that all appear to be derived from a single ancestral stock (Bohlmann et al., 1998b). Overall, the amino acid sequences deduced share a high degree of similarity, and the positions of many residues thought to be involved in catalysis are conserved. When genomic sequences are compared (Facchini and Chappell, 1992; Mau and West, 1994; Back and Chappell, 1995), a common pattern of intron–exon organization is evident. Within the terpene synthases, phylogenetic reconstruction divides the known sequences into six subfamilies, each of which has a minimum of 40% identity among its members (Bohlmann et al., 1998b). The pattern of sequence relationships is influenced by the taxonomic affinities of plant species, as well as by the chemical similarities among enzyme products and the reaction mechanism employed. For example, the limonene synthases of Abies grandis, a gymnosperm, are more closely related to other gymnosperm monoterpenes and sesquiterpenes synthases than they are to the limonene synthases from angiosperms.

Terpene synthases, also known as terpene cyclases because most of their products are cyclic, utilize a carbocationic reaction mechanism very similar to that employed by the prenyltransferases. Numerous experiments with inhibitors, substrate analogues and chemical model systems (Croteau, 1987; Cane, 1990, 1998) have revealed that the reaction usually begins with the divalent metal ion-assisted cleavage of the diphosphate moiety (Fig. 5.6). The resulting allylic carbocation may then cyclize by addition of the resonance-stabilized cationic centre to one of the other carbon–carbon double bonds in the substrate. The cyclization is followed by a series of rearrangements that may include hydride shifts, alkyl shifts, deprotonation, reprotonation and additional cyclizations, all mediated through enzyme-bound carbocationic intermediates. The reaction cascade terminates by deprotonation of the cation to an olefin or capture by a nucleophile, such as water. Since the native substrates of terpene synthases are all configured with trans (E) double bonds, they are unable to cyclize directly to many of the carbon skeletons found in nature. In such cases, the cyclization process is preceded by isomerization of the initial carbocation to an intermediate capable of cyclization.

The recently published crystal structure of tobacco epi-aristolochene synthase (a sesquiterpene synthase) has provided the first look at the three-dimensional configuration of a plant terpene synthase (Starks et al., 1997).
Figure 5.6  Proposed mechanism for the cyclization of geranyl diphosphate to sabinene and sabinene hydrate under catalysis by monoterpene synthases: the reaction begins with the hydrolysis of the diphosphate moiety to generate a resonance-stabilized carbocation (1); the carbocation then isomerizes to an intermediate capable of cyclization by return of the diphosphate (2); and rotation around a single bond (3); after a second diphosphate hydrolysis (4); the resulting carbocation undergoes a cyclization (5); a hydride shift (6); and a second cyclization (7); before the reaction terminates by deprotonation (8); or capture of the cation by water (9). Cyclizations, hydride shifts and a variety of other rearrangements of carbocationic intermediates are a characteristic of the mechanisms of terpene synthases. No known terpene synthase actually produces both sabinene and sabinene hydrate; these are shown to indicate the possibilities for reaction termination. PP indicates a diphosphate moiety.

The structure provides a physical basis for some of the proposed mechanistic features and reveals several elements responsible for controlling the course of reaction. The arrangement of the protein backbone, consisting of eight antiparallel α-helices that form a large cavity, is very similar to that reported for two other terpene synthases, a fungal sesquiterpene synthase (Lesburg et al., 1997) and a bacterial triterpene synthase (Wendt et al., 1997). It is also strongly reminiscent of the structure of avian liver FPP synthase (discussed in Section 5.3.3) despite only a low level of sequence similarity, reflecting the parallels in the reaction mechanism between terpene synthases and prenyltransferases. Among the notable features of the epi-aristolochene synthase structure is the presence of an aspartate-rich cluster, DDxxD, in the active site (just like those found in prenyltransferases) that serves to bind the diphosphate moiety of
the substrate via a Mg\textsuperscript{2+} bridge. Prenyltransferases, which simultaneously bind two different diphosphate-containing substrates, have two such clusters, while epi-aristolochene synthase and other terpene synthases, which bind only one diphosphate-containing substrate, have only one. The active site of epi-aristolochene synthase also contains a variety of aromatic amino acid residues that may serve to stabilize the enzyme-bound carbocationic intermediates by π–cation interactions (Wise and Croteau, 1998). Other amino acid residues were identified that direct the released diphosphate moiety away from the active site, that complex two additional Mg\textsuperscript{2+} ions and that participate in protonation and deprotonation.

Terpene synthases employ two other modes of generating the initial carbocationic intermediate in addition to hydrolysis of the diphosphate ester. The reaction may be initiated by protonation of an epoxide, as in the cyclization of oxidosqualene to sterols and triterpenes (Abe, 2007), or by protonation of the carbon–carbon double bond at the opposite end of the molecule from the diphosphate moiety. Mechanisms initiated by double-bond protonation are a characteristic of the formation of many diterpenes, such as copalyl diphosphate (West, 1981) (Fig. 5.7). Isolated cDNA sequences encoding copalyl diphosphate synthase have some homology to the sequences of terpene synthases in which the reaction is initiated by diphosphate hydrolysis, but lack the characteristic DDxxD motif, possessing instead an alternate aspartate-rich motif, DxDDTA, at a very different position in the sequence (Sun and Kamiya, 1994; Bensen et al., 1995; Ait-Ali et al., 1997). A second category of diterpene synthases has more in common with the majority of terpene synthases discussed above, catalysing diphosphate hydrolysis-initiated cyclizations while possessing typical DDxxD motifs (Yamaguchi et al., 1996, 1998). Notable members of this group include the ent-kaurene synthases involved in gibberellin biosynthesis, which use copalyl diphosphate as a substrate rather than a product. There is also a third type of diterpene synthase that seems to combine the properties of the other two classes. For example, abietadiene synthase from A. grandis catalyses two sequential cyclization steps: first cyclizing GGPP to copalyl diphosphate via a double-bond protonation-initiated cyclization and then converting copalyl diphosphate to the olefin, abietadiene, via a diphosphate hydrolysis-initiated process (Keeling and Bohlmann, 2006). Appropriately, the A. grandis abietadiene synthase cDNA has regions of sequence homologous to both other types of diterpene synthases and contains both DDxxD and DxDDTA elements (Vogel et al., 1996).

Not all terpene synthases catalyse complex reactions. Isoprene synthase converts DMAPP to the hemiterpene (C\textsubscript{5}), isoprene (Fig. 5.1), a comparatively simple process involving the ionization of the diphosphate group, followed by double-bond migration and proton elimination (Silver and Fall, 1991). Present in chloroplasts in both stromal and thylakoid-bound forms, isoprene synthase is a homodimer that differs from other terpene synthases in many properties, such as subunit architecture, optimum pH and kinetic parameters.
Biochemistry of Plant Secondary Metabolism

Figure 5.7  Proposed mechanism for the cyclization of geranylgeranyl diphosphate (GGPP) to the diterpene copalyl diphosphate, an example of terpene synthase-catalysed cyclization initiated by double-bond protonation, rather than by hydrolysis of the diphosphate ester. PP indicates a diphosphate moiety.

(Silver and Fall, 1995; Wildermuth and Fall, 1998). Its key role in the formation of isoprene, an abundant plant volatile with a major influence on atmospheric chemistry, has made it a popular target for cloning efforts.

An unusual feature of terpene synthases is the ability of a single enzyme to catalyse the formation of more than one product species. First suggested by the copurification of separate activities and differential inactivation studies, and later demonstrated by isotopically sensitive branching experiments (Wagschal et al., 1991; Rajaonarivony et al., 1992), this property has been unequivocally proved by cDNA cloning. Heterologous expression of many cloned terpene synthases, such as 1,8-cineole synthase from S. officinalis, leads to a mixture of products (Wise et al., 1998). In a spectacular, recently published example, two sesquiterpene synthases from A. grandis, δ-selinene synthase
and γ-humulene synthase, were shown to synthesize 34 and 52 different sesquiterpenes, respectively (Steele et al., 1998a). The tendency of terpene synthases to form multiple products is probably a consequence of their reaction mechanisms, which involve highly reactive carbocationic intermediates that may have more than one chemical fate. Interestingly, exon-swapping experiments on epi-aristolochene synthase converted this single product sesquiterpene synthase to one making multiple products (Back and Chappell, 1996). Further correlations between elements of protein structure and features of the reaction mechanism using three-dimensional structures will increase our understanding of how terpene synthases are able to make multiple products.

Terpene synthases are likely to serve as important agents of flux control in terpene biosynthesis because they operate at metabolic branch points where pathways diverge to different terpene types. However, there is still insufficient information available to assess the regulatory significance of these catalysts. Direct relationships between terpene synthase activity and changes in the rate of terpene formation have been noted on several occasions (Dudley et al., 1986; Gijzen et al., 1991; Zook et al., 1992; Bohlmann and Croteau, 1999; Sharkey and Yeh, 2001), but terpene synthase activity is not always well correlated with the accumulation of end products of the pathway (Keller et al., 1998a; Jennewein and Croteau, 2001; Pichersky et al., 2006). In evaluating the regulatory importance of terpene synthases, it is necessary to consider not only the level of activity, but also its subcellular location. As we have noted above, monoterpenes and diterpenes are generally formed in the plastids, while sesquiterpene and triterpene biosynthesis is restricted to the cytosol (Mettal et al., 1988; Kleinig, 1989; Turner et al., 1999). Based on subcellular fractionation studies and the presence or absence of plastid transit peptides, the distribution of most terpene synthases follows this pattern. Most monoterpene and diterpene synthases are localized in the plastids (Mau and West, 1994; Aach et al., 1995, 1997; Vogel et al., 1996; Wise et al., 1998; Yamaguchi et al., 1998), while sesquiterpene and triterpene synthases are cytosolic (Belingheri et al., 1988; Kleinig, 1989; Bohlmann et al., 1998a; Steele et al., 1998a). Terpene synthase activity itself seems to be regulated by the level of the corresponding mRNA (Facchini and Chappell, 1992; Chen et al., 1995; Dudareva et al., 1996; Keller et al., 1998a; Steele et al., 1998b). Reports of multi-gene families (Facchini and Chappell, 1992; Colby et al., 1993; Back and Chappell, 1995) may imply complex developmental and tissue-specific patterns of regulation or may just indicate the existence of different synthases with closely related sequences.

In addition to terpene synthases, the construction of terpenoid carbon skeletons in plants also involves a number of prenyltransferases distinct from those that make the C_{10}, C_{15} and C_{20} diphosphates. One class of prenyltransferases catalyses 1′-4 condensations of IPP with an FPP or GGPP starter unit to make long-chain polyterpenes, such as rubber, a linear hydrocarbon with cis (Z) double bonds and as many as 30 000 isoprene units. The
cis-polyisoprenyltransferase participating in rubber biosynthesis has been characterized in several species of plants (Mooibroek and Cornish, 2000; Cornish, 2001; Takahashi and Koyama, 2006), but efforts to purify this protein or clone the corresponding gene have not yet been successful. Another class of prenyltransferases mediates condensations between allylic diphosphates and non-isoprenoid substrates, in which dimethylallyl, geranyl, farnesyl or geranylgeranyl moieties are transferred to a nucleophilic acceptor. These are key reactions in the formation of many different prenylated compounds, including prenylated proteins, prenylated flavonoids, furanocoumarins, cytokinins, ubiquinone, plastoquinone and the tocopherols. Several of the enzymes responsible have been well studied and are similar in gross properties to other prenyltransferases (Laflamme et al., 1993; Cutler et al., 1996; Qian et al., 1996; Fellermeier and Zenk, 1998; Muhlenweg et al., 1998; Yamamoto et al., 2000).

### 5.3.5 Secondary transformations

The cyclic terpenes formed initially are subject to an assortment of further enzymatic modifications, including oxidations, reductions, isomerizations and conjugations, to produce the wide array of terpenoid end products found in plants. Unfortunately, few of these conversions have been well studied, and there is little evidence from most of the biosynthetic routes proposed, except in the case of the gibberellin (Yamaguchi, 2008) pathway. Many of the secondary transformations belong to a series of well-known reaction types that are not restricted to terpenoid biosynthesis. For example, the hydroxylation of terpenes by cytochrome P450-dependent oxygenases has been the subject of much investigation (Mihaliak et al., 1993) (Fig. 5.8a). This large family of membrane-bound enzymes catalyses the position-specific hydroxylation of many terpenoids, using molecular oxygen and NADPH (Hallahan et al., 1992; Hoshino et al., 1995; Kato et al., 1995; Winkler and Helentjaris, 1995; Helliwell et al., 1998; Ro and Bohlmann, 2006). The first cDNA encoding a cytochrome P450-dependent terpene hydroxylase has recently been isolated (Lupien et al., 1995).

A second group of oxidative enzymes, the 2-oxoglutarate-dependent dioxygenases, are soluble, nonhaeme iron-containing catalysts (Prescott and John, 1996) that participate in several reactions in terpene biosynthesis (Lange et al., 1994; Phillips et al., 1995; Xu et al., 1995; Yamaguchi, 2008) (Fig. 5.8b). Several other types of secondary transformation that have been characterized include the oxidation of acyclic monoterpene alcohols to their corresponding aldehydes during iridoid biosynthesis in *Nepeta racemosa* (Hallahan et al., 1995), the reduction of the geranylgeranyl moiety of chlorophylls, tocopherols and phyloquinone in *A. thaliana* (Keller et al., 1998b) and the glucosylation of diterpene alcohols by glucosyltransferases in *Stevia rebaudiana* (Shibata et al., 1995).
Figure 5.8 Examples of oxidative secondary transformations in terpenoid biosynthesis. (a) Hydroxylation of epi-aristolochene at the 3-position by a cytochrome P450-dependent terpene hydroxylase in Capsicum annuum (Hoshino et al., 1995). (b) Conversion of \( \text{GA}_1 \) to \( \text{GA}_9 \) by a 2-oxoglutarate-dependent dioxygenase involved in gibberellin biosynthesis. A single enzyme catalyses three successive oxidations leading to the loss of a methyl group and lactone formation (Lange et al., 1994; Phillips et al., 1995; Xu et al., 1995). (c) Oxidation of cis, cis-nepetalactol to cis, cis-nepetalactone by a nicotinamide adenine dinucleotide (NAD\(^+\))-dependent soluble oxidoreductase in Nepeta racemosa (Hallahan et al., 1998).

5.4 Conclusions

Research on the formation and function of plant terpenoids has flourished in the past decades. The greatest achievement has been the discovery of a new, non-mevalonate route for the synthesis of the \( C_5 \) building blocks of terpenoids. While many of the intermediates of the new glyceraldehyde phosphate/pyruvate pathway are still unidentified and most of the enzymes are completely unknown, such details should be rapidly elucidated, setting the stage for studies on the distribution of the new pathway in plants and its relationship to the ‘classical’ mevalonate pathway. At present, the glyceraldehyde phosphate/pyruvate route appears to be found in the plastids of all higher plant species and is the likely source of substrate for the plastid-associated terpenoids, including monoterpenes, diterpenes, phytol, plastoquinones and carotenoids. In contrast, the mevalonate pathway appears to be restricted to the cytosol/ER based on the finding that all known pathway genes are targeted to this compartment. The mevalonate route may be the chief source of substrate for cytosolic (sesquiterpenes, triterpenes, dolichols) and mitochondrial (ubiquinone) terpenoids. Further
research is urgently needed to confirm these generalizations concerning the subcellular compartmentation of terpenoid biosynthesis. The extent to which the two pathways interact must also be clarified and the existence of a cryptic mevalonate pathway in the plastids, at least in certain taxa or specific developmental stages, must be investigated. With the basic features of the new, non-mevalonate pathway coming into focus, it is also time to re-evaluate the regulation of terpenoid formation in general, especially the role of HMGR, to determine which steps are the main modulators of flux.

As in most other branches of plant science, the application of molecular biology to terpenoid biosynthesis has led to enormous progress. The cloning and heterologous expression of biosynthetic enzymes have permitted new inferences about the evolution of these catalysts and have opened the door to site-specific mutagenesis and X-ray structure determination, which in turn have revealed much new information on enzyme structure and mechanism. For prenyltransferases and terpene synthases, two major groups of terpenoid-synthesizing enzymes that catalyse complex reactions involving carbocationic intermediates, we will soon achieve a detailed understanding of not only how the enzyme directs the outcome of the reaction, but also how redesign of the protein can give altered product distributions.

As terpenoids constitute the largest class of plant secondary compounds, it is fitting that terpenoid metabolites play a wide assortment of roles in nearly all basic plant processes. Recent research has added to this list, suggesting new functions for terpenoids, such as isoprene (stabilizing membranes at high temperatures), prenylated proteins (control of the cell cycle, allocation of nutrients) and certain mono- and sesquiterpenes (attraction of the enemies of herbivores). Nevertheless, the roles of most terpenoids are completely unstudied. Many compounds are thought to be involved in protecting plants from herbivores and pathogens, but supporting data are often fragmentary and unconvincing. In the coming years, the use of molecular techniques to make precise alterations to the levels of individual compounds should facilitate more rigorous investigation of the functional significance of terpenoids and give us a greater appreciation of their roles in plants.

References


Biochemistry of Plant Secondary Metabolism


Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes


Biochemistry of Plant Secondary Metabolism


Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes


Biochemistry of Plant Secondary Metabolism


Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes


Biochemistry of Plant Secondary Metabolism


Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes


300  Biochemistry of Plant Secondary Metabolism


Biochemistry of Terpenoids: Monoterpenes, Sesquiterpenes and Diterpenes


